Optical genome mapping and multi-modal characterization of breast tumor and organoid samples reveal structural variation and gene fusion signatures

Rahul Kumar[1,2], Osama Shiraz Shah[1,3], Brent T. Schlegel[1,3], Jagmohan Hooda[1], Daniel D. Brown[4,1], Hung-Ching Chang[5], Alexander Chih-Chieh Chang[6,1], Jennifer Xavier[6,1], Polly E Mattila[4,1], Liane Fairfull[1], George C. Tseng[5], Rohit Bhargava[7,1], Steffi Oesterreich[1,2*], Adrian V. Lee[1,2,4*]

[1]Women's Cancer Research Center at UPMC Hillman Cancer Center and Magee-Women's Research Institute, Pittsburgh, PA, USA, [2]Department of Pharmacology & Chemical Biology, University of Pittsburgh, Pittsburgh, PA, USA, [3]Integrative System Biology, School of Medicine, University of Pittsburgh, PA, USA
[4]Institute for Precision Medicine, PA, USA, [5]Department of Biostatistics, University of Pittsburgh, PA, USA, [6]Department of Medicine, UPMC, Pittsburgh, PA, USA, [7]Department of Pathology, Magee Women's Hospital, Pittsburgh, PA, USA
*Co-corresponding Author

Introduction

- Structural variations (SVs) including deletions, duplications, inversions, and translocations are critical drivers of genomic instability and oncogenesis in breast cancer.
- While short-read sequencing has enabled large-scale cataloging of somatic alterations, but it often lacks the resolution to accurately detect complex or large SVs, especially in repetitive or rearranged genomic regions.
- Optical genome mapping (OGM) provides an orthogonal, high-resolution platform to capture genome-wide SVs with exceptional sensitivity and precision.
- Unlike sequencing-based methods, OGM can resolve long-range genomic architecture, enabling improved characterization of chromosomal rearrangements and fusion events.
- In this study, we integrated OGM with bulk RNA-seq across breast tumor and patient-derived organoid (PDO) samples to uncover the landscape of SVs, genomic instability and putative functionally relevant gene fusions.

Hypothesis

We hypothesize that integrating Optical Genome Mapping with bulk RNA-seq enables the identification of putative functional gene fusions, defined as genomic rearrangements where at least one fusion partner gene shows significant expression change. By applying this multi-modal approach across independent cohorts of breast tumors and patient-derived organoids (PDOs), we aim to uncover recurrent structural variants and transcriptionally active gene fusions that are conserved across sample types, reflective of underlying genomic instability, and informative of breast cancer subtype-specific biology.

Approach

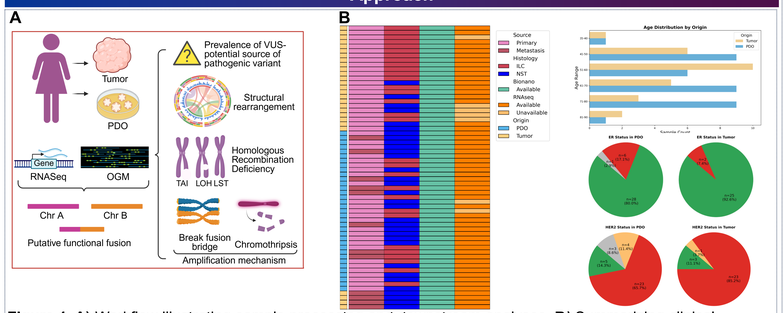


Figure 1. A) Workflow illustrating sample processing and downstream analyses. **B)** Summarizing clinical metadata and data availability across patient-matched tumor and PDO samples.

Results

Figure 2. Structural variant (SV) landscape across patient-derived organoids (PDOs) and tumor samples

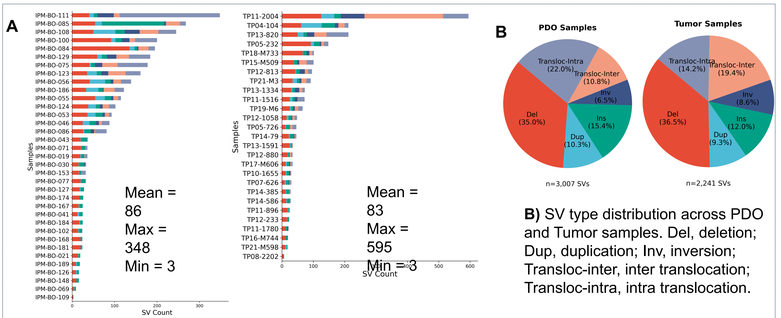


Figure 2. A) SV frequency & types across PDO (on left) and tumor (on right) samples.

B) SV type distribution across PDO and Tumor samples. Del, deletion; Dup, duplication; Inv, inversion; Transloc-inter, inter translocation; Transloc-intra, intra translocation.

Results- Continued

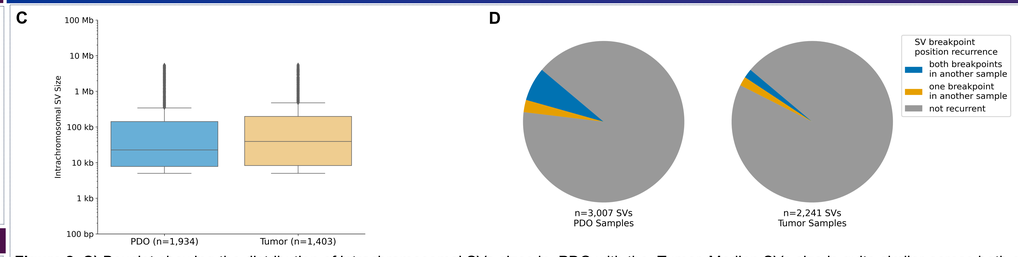


Figure 2. C) Boxplot showing the distribution of intrachromosomal SVs sizes in PDO with the Tumor. Median SVs size is quite similar across both groups, with a broad range spanning from a few kilobases to over 10 megabases **D)** Pie charts illustrate the recurrence of SV breakpoint position across all identified SVs in PDO samples (left) and tumor samples (right) .

Figure 3. Comprehensive profiling of top altered breast cancer genes and functional annotations

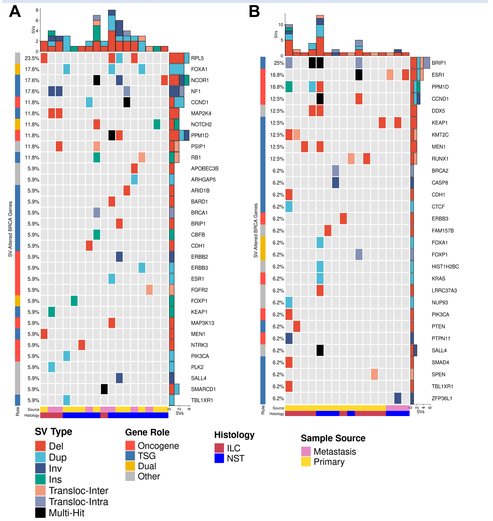


Figure 3. Heatmap showing the distribution and frequency of SVs affecting breast cancer genes curated from the OncoVar database across PDO **(A)** and Tumor **(B)** samples. Each column represents an individual sample; each row corresponds to a gene have SVs. The number of SVs per sample is indicated in the bar plots above each panel. Bar plots to the right of each panel shows number of SVs in each gene. Each panel further annotated by gene role, SV type, sample source and histology.

Figure 4. Assessing structural instability metrics

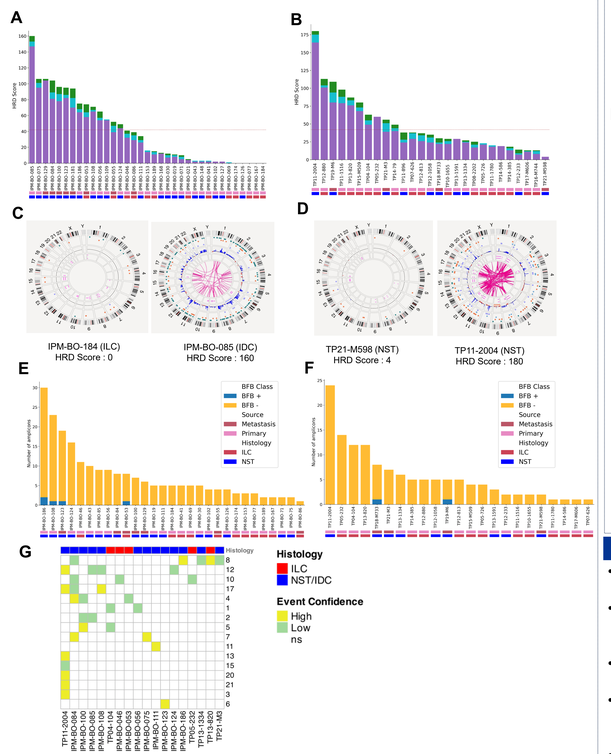


Figure 4. (A–B) Comparative analysis of homologous recombination deficiency (HRD) components—Large-Scale State Transitions (LST, purple), Loss of Heterozygosity (LOH, blue), and Telomeric Allelic Imbalance (TAI, green)—across (A) PDOs and (B) tumor samples, stratified by histological subtype. Samples with HRD scores >42 are classified as HRD-high. **(C–D)** Representative circos plots depicting structural variants (SVs) complexity in (C) PDOs and (D) tumor samples, selected based on the extremes of HRD scores. **(E–F)** Distribution of breakage-fusion-bridge (BFB) amplicons identified in (E) PDOs (N=216) and (F) tumor samples (N=131), illustrating the prevalence and variation of BFB-driven genomic instability across sample types. **(G)** Chromothripsis events detected across PDO and tumor samples. Each row represents a chromosome, and each column corresponds to a sample. Colored squares indicate chromothripsis events, with green denoting high-confidence and yellow representing low-confidence events.

Figure 5. Fusion events in PDO and tumor samples.

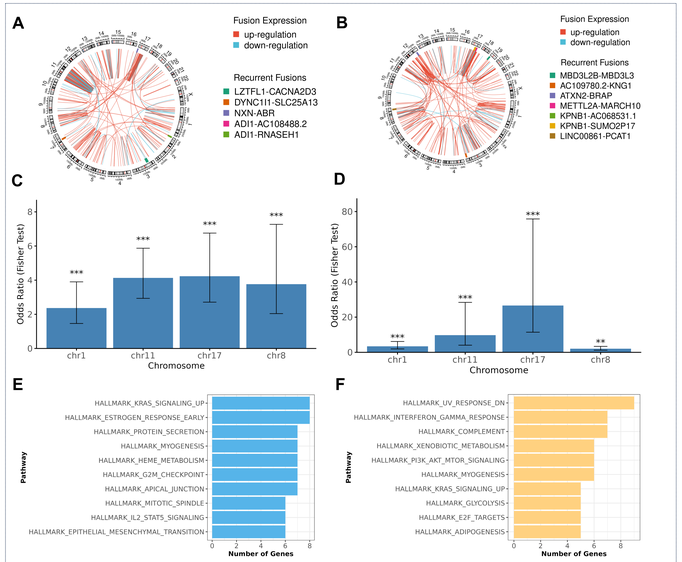


Figure 5. (A-B) Circos plot visualizes all the putative functional fusion events across the genome in (A) PDO and (B) tumor samples. Each arc represents a fusion event, color-coded by expression status of the fusion transcript: red for up-regulation and blue for down-regulation. Recurrently observed fusions are labeled on the right of plot, highlighting their potential relevance in tumor biology or PDO modeling. **(C–D)** Chromosomal enrichment of expression-relevant gene fusions among chromosome having high fusion burden. Bar plot summarizes the odds ratio for the enrichment of expression-relevant gene fusions compared to all fusions within each high fusion–burden chromosome (chr1, chr8, chr11, chr17) in (C) PDOs and (D) Tumors samples. **(E–F)** Bar plots display the top 10 pathways containing the highest number of fused genes involved in putative functional fusions in (E) PDO and (F) Tumor samples. Fusion-associated genes were mapped to Hallmark pathways from the MSigDB database. The x-axis indicates the number of fused genes mapped per pathway, and the y-axis shows the pathway names.

Summary and future direction

- This study uncovered the SV landscape in breast tumors and PDOs using OGM, revealing recurrent alterations in key cancer genes such as *ESR1*, *PPM1D*, *CCND1*, and *ERBB3*.
- We quantified genomic instability across samples by assessing HRD scores, BFB cycles and Chromothripsis demonstrating heterogeneity in structural complexity between tumors and organoids.
- Integration with RNA-seq enabled multi-omics characterization of over 670 putative functional fusions across tumor and PDO supported by differential gene expression.
- Notably, a significant fraction of these fusions involved long non-coding RNAs (lncRNAs)—a class of molecules increasingly implicated in cancer regulation but still poorly understood in the context of structural variation.
- Future work will explore the regulatory roles, subtype specificity, and therapeutic implications of lncRNA-containing fusions, potentially uncovering new targets and biomarkers in breast cancer.

Acknowledgements:
This work was supported by the Shear Family Foundation, The Dynami Foundation, NCI R01CA252378 and P30 Cancer Center Support Grant (P30CA047904). This research was also supported in part by the University of Pittsburgh Center for Research Computing and Data, RRID:SCR_022735, through the resources provided. Specifically, this work used the HTC cluster, which is supported by NIH award number S10OD028483.