Development of a Novel Platelet RNA Based Biomarkers for Breast Cancer Detection:

Jong-Ho Cheun[1], Changhoon Lee[2], Dong Seung Shin[2], Sangik Park[3], TaeJin Ah[4,5], Wonshik Han[2], Hong-Kyu Kim[2,*]

[1]Department of Surgery, SMG-SNU Boramae Medical Center, [2]Department of Surgery, Seoul National University Hospital, [3]Department of integrated science, Handong University, [4]Department of Life Science, Handong University, [5]Foretell My Health.

Introduction

- Current breast cancer screening relies mainly on mammography and breast ultrasonography, which remain the standard diagnostic tools in clinical practice. However, unlike imaging-based approaches, **there is still no reliable blood-based test for breast cancer detection**, highlighting an unmet need for minimally invasive biomarkers that can complement existing screening strategies.
- Emerging evidence suggests that **tumor-educated platelets (TEPs)** undergo dynamic RNA splicing changes in response to cancer-derived signals, **offering a minimally invasive avenue for cancer detection.** Platelet RNA, therefore, **represents a promising but underexplored liquid biopsy source for early cancer detection.**
- In this study, we **leveraged platelet RNA sequencing (RNA-seq) data to identify breast cancer-associated exon-exon splice junctions** whose expression is significantly altered in cancer patients compared to healthy controls.

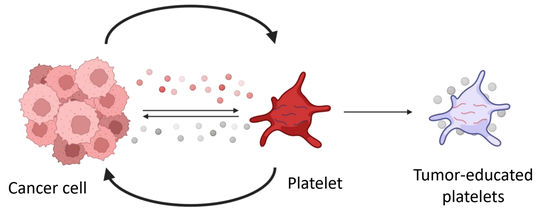


Fig. 1. The process by which cancer cells alter platelets

Methods and Materials

- We collected platelet **RNA-seq data from public datasets** and **prospectively collected clinical samples** from breast cancer (BC) patients and healthy controls.
- Breast cancer–associated splice junctions were identified using a stepwise feature selection process guided by the **Akaike Information Criterion (AIC).**
- **A logistic regression classifier** was then trained using the selected junction features and evaluated on both public test data and independent clinical samples using AUC as the primary metric.

Results

- Using the stepwise feature selection process guided by the AIC, we identified **11 splice junction candidates** from publicly available platelet RNA-seq datasets (control 274; BC 93) and prospectively collected clinical samples (control 87; BC 7). **(Fig.2)**
- Logistic regression models were trained on public training data and evaluated on both public and clinical test cohorts.
- The logistic regression classifier demonstrated **strong predictive accuracy, achieving an AUC of 0.873 on public test data and 0.949 on newly generated clinical samples**, highlighting robust generalizability across datasets

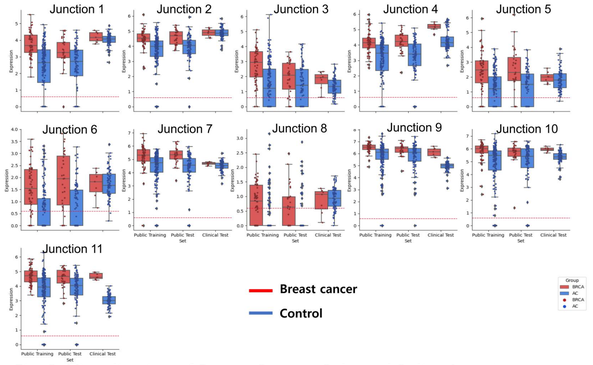


Fig. 2. Visualization of Breast Cancer–Selected Splice Junctions

- Dimensionality reduction using **t-SNE revealed a clear separation between breast cancer and control samples** in both public and clinical cohorts, indicating distinct splice-junction expression patterns associated with cancer. **(Fig.3 upper)**
- **Heatmap visualization further confirmed consistent differential expression** across the selected junctions, with coherent clustering of cancer versus non-cancer individuals in all datasets **(Fig.3 lower)**

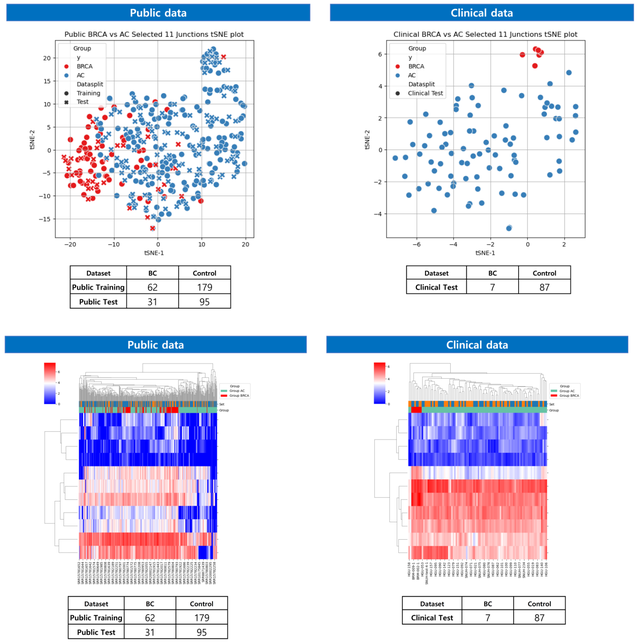


Fig. 3. Visualization of Breast Cancer–Selected Splice Junctions

- Gene Ontology analysis of the genes containing these junctions showed significant enrichment in pathways related to **endothelial cell development, establishment of endothelial barrier, and regulation of protein localization to the cell surface** (FDR < 0.05).
- These pathways suggest that **platelet splicing alterations reflect tumor-driven vascular and immune modulation**

chrom_junction	AUC	cancer_mean	AC_mean	mean_diff	cancer_median	AC_median	median_diff
Junction 1	0.811047	3.768551	2.309065	1.459486	3.676888	2.652484	1.024404
Junction 2	0.768787	4.489968	3.66543	0.824538	4.543535	3.981698	0.561836
Junction 3	0.759641	2.70614	1.365398	1.340742	2.953272	1.217705	1.735567
Junction 4	0.769688	4.138978	2.949778	1.1892	4.136051	3.460487	0.675565
Junction 5	0.755902	2.41865	1.173058	1.245592	2.458342	1.215289	1.243053
Junction 6	0.775275	1.576218	0.608072	0.968146	1.54388	0	1.54388
Junction 7	0.753469	5.204932	4.372174	0.832758	5.311977	4.738499	0.573478
Junction 8	0.750766	0.784371	0.201805	0.582566	0.839776	0	0.839776
Junction 9	0.778158	6.487154	5.74266	0.744494	6.525255	6.110434	0.414821
Junction 10	0.752838	5.817335	4.695349	1.121986	6.008291	5.477899	0.530391
Junction 11	0.772301	4.681863	3.558685	1.123178	4.714594	3.9246	0.789994

Table 1. AUC and Group-wise Mean Expression of the 11 Breast Cancer–Selected Splice Junctions

Conclusions

- Platelet-derived splice junction profiles **provide a robust and minimally invasive approach for distinguishing breast cancer** from healthy controls.
- The logistic regression classifier built on these junction features shows **strong predictive performance** across public and clinical datasets, supporting their **potential utility as blood-based biomarkers for breast cancer detection**

Contact

Jong-Ho, Cheun, MD

Division of Breast Surgery, Department of Surgery, Seoul National University, Boramae Medical Center, Seoul, Korea

Email: chun89aaa@naver.com / Phone: +82-10-8754-3054

This presentation is the intellectual property of the author/presenter. Contact at <chun89aaa@naver.com> for permission to reprint and/or distribute