

Identification and Validation of a Novel Theranostic Target in Triple Negative Breast Cancer with Transcriptomics and Protein Analyses

Hyunjong Lee^{1,2}, Giro Kim^{1,2}, Mina Kim¹, Jung Lim Kim¹, Kyung-Ho Jung^{1,2}, Kyung-Han Lee^{1,2}

¹Department of Nuclear Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; ²Department of Health Sciences and Technology, SAIHST, Sungkyunkwan University, Seoul, 06355, Republic of Korea

Correspondence: Hyunjong Lee, Department of Nuclear Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul, 06351, Republic of Korea, Tel +82 2 3410 3424, Fax +82 2 3410 2639, Email nmhjlee@gmail.com

Background: Triple negative breast cancer (TNBC) poses a significant clinical challenge in imaging and therapy due to absence of conventional targets such as estrogen, progesterone, and HER2 receptors. The aim of this study is to identify potential targets for TNBC through comprehensive transcriptomic analyses and validation in TNBC cell lines and tissues.

Methods: Single-cell RNA sequencing (scRNA-seq) data obtained from 8 TNBC patients and 11 normal patients were analyzed to identify differentially expressed surface proteins using adjusted p-values based on Bonferroni correction. To verify the validity of potential targets identified through scRNA-seq, expression of selected proteins was evaluated in bulk RNA-seq data from 162 TNBC and 113 normal breast tissues in the TCGA cohort. Finally, expression of selected proteins was examined in representative human TNBC cell lines and xenograft tumors in immunodeficient mice.

Results: Nine proteins were revealed to be expressed more than twice as much in TNBC cells compared to normal cells: LY6E, LY6D, LAMP1, EMP2, TTYH1, CD74, BST2, HLA-DRA, and HLA-DRB1 ($P < 10^{-10}$ for all). Among those, LY6E, LY6D, BST2, and TTYH1 were selected as potential target proteins with significantly higher expression in TNBC tissues ($P < 0.0001$ for LY6E and LY6D; $P = 0.007$ for BST2; $P = 0.002$ for TTYH1). Validation experiments revealed that LY6E demonstrated high expression on the membrane of TNBC cell lines, while exhibiting low expression in normal breast epithelial cells. Consistently, Western blot analysis of tumors from an in vivo xenograft model derived from a TNBC cell line confirmed elevated LY6E expression.

Conclusion: In conclusion, this study suggests LY6E as a potential target for the selective imaging and therapy of TNBC, identified by transcriptomics analysis and cell line experiments. It lays the groundwork to augment clinical impact of future preclinical studies for the development of diagnostic and treatment approaches for TNBC.

Keywords: triple negative breast cancer, diagnosis, therapy, target protein, transcriptomics

Introduction

Breast cancer accounts for the largest proportion of newly diagnosed cancers in women and ranks second in mortality following lung cancer.¹ Triple negative breast cancer (TNBC) is a subtype of breast cancer, characterized by its inability to express the conventional estrogen, progesterone, and HER2 receptors. In the United States, TNBC constituted 12% of diagnosed breast cancer.² Owing to the absence of specific receptors and the aggressiveness of TNBC, its 5-year survival rate was observed to be 8% to 16% lower compared to that of hormone receptor-positive forms of the disease.² Despite the availability of a molecular imaging modality such as F-18 fluorodeoxyglucose positron emission tomography/computed tomography (FDG PET/CT),³ it falls short in assessing the expression of specific target proteins in TNBC. Furthermore, while potential targets like STAT3 and MELK have been suggested,^{4,5} their application in actual clinical fields remains limited. Therefore, the development of novel diagnostic and therapeutic targets for TNBC is imperative.

Traditional methods for proposing targets for cancers including TNBC have predominantly been based on known oncogenic mechanisms or characteristics identified through immunohistochemistry (IHC). In contrast to these

conventional approaches, with the emergence of single-cell RNA sequencing (scRNA-seq), a notable change is occurring in cancer research. This advanced technique is providing more detailed insights into the proteomic landscape. For breast cancer specifically, scRNA-seq analysis has explored the heterogeneity of tumor cells as well as the diverse immune profiles within the microenvironment, including subclasses of cancer-associated fibroblasts.^{6,7} Although previous studies have utilized scRNA-seq analysis to identify factors affecting aggressiveness and treatment response in TNBC,^{8,9} research specifically targeting diagnostic or therapeutic targets remains scarce. Furthermore, although LY6E has been studied in the context of immune modulation—including roles in T-cell signaling and associations with pathways such as TGF- β or PD-L1—its potential as a surface-accessible theranostic target has not been systematically explored.^{10,11} The discovery of potential target proteins using scRNA-seq analysis could provide crucial insights, paving the way for the development of innovative diagnostic and treatment methods.

Meanwhile, cell lines derived from specific cancer types, including TNBC, are pivotal in the preclinical testing of novel targeting agents. To boost the clinical applicability of tests, it is essential to evaluate whether target proteins identified through scRNA-seq in actual patient cells or tissues are similarly expressed in experimental cell lines. This approach serves as a vital link that not only validates the results of big data analysis but also facilitates their application in preclinical trials.

This study hypothesized that surface proteins enriched in malignant TNBC epithelial cells, relative to normal breast epithelium, would serve as tractable candidates for targeted imaging and therapy. To test this hypothesis, we analyze scRNA-seq data from 8 TNBC and 11 normal patients and identify surface proteins that are characteristically expressed in TNBC. Subsequently, we assess the level of expression of these surface proteins within The Cancer Genome Atlas (TCGA) dataset to select potential target proteins. Finally, we evaluate the expression of these proteins in representative TNBC cell lines and further validate in an in vivo xenograft tumor model. This comprehensive approach aims to identify novel targets for TNBC diagnosis and therapy, bridging the gap between molecular findings and clinical applications.

Methods

Preprocessing scRNA-Seq Data

Downstream processed R data objects from scRNA-seq data of TNBC patients were downloaded from Figshare at <https://doi.org/10.6084/m9.figshare.17058077>.¹² The objects included a total of 421,761 cells from 55 patients.¹³ Among objects, two data were selected for the purpose of identifying potential target proteins by comparing the transcriptomic profiles of TNBC cells with those of normal breast epithelial cells: “SeuratObject_TNBCtum.rds” and “SeuratObject_NormEpiSub.rds”. These objects contained a total of 83,252 cells after quality filtering from 8 TNBC and 11 normal patients. In this study, “TNBC cells” were defined as all malignant epithelial cells contained within the object “SeuratObject_TNBCtum.rds,” which represents the tumor epithelial compartment in the original annotation.

In the original study that generated these publicly available downstream processed objects, all scRNA-seq data underwent rigorous preprocessing, including filtering of low-quality cells (generally ≥ 500 detected genes with sample-specific adjustments), exclusion of cells with high mitochondrial read fractions (typically $\leq 20\%$), and removal of putative doublets based on upper bounds of library size and detected gene counts. Samples were integrated using Seurat’s anchor-based integration workflow, and malignant epithelial cells were identified through copy number variation analysis (inferCNV). These preprocessing steps and associated thresholds are described in detail in the original publication and its accompanying code repository.¹²

After merging two objects, data were then scaled by log-normalization after the read counts were divided by the total number of transcripts and multiplied by 10,000. Highly variable 2,000 genes were selected using the *FindVariableFeatures* function of the Seurat package (version 4.1.0). Dimension reduction was performed by uniform manifold approximation and projection (UMAP). The *FindMarkers* function of the Seurat package was used to identify differentially expressed genes (DEGs) between TNBC and normal epithelial cells from genes expressed in at least 10% of either population, ensuring significance in the downstream analysis.

Evaluating Expression of Alleged Imaging Targets

Alleged molecular targets and matched imaging tracers were selected as a previous study: glucose transporters (GLUT)/2- ^{18}F -fluoro-2-deoxy-d-glucose (^{18}F FDG), monocarboxylate transporters (MCT)/ ^{11}C -acetate, folate receptors (FOLR)/ ^{18}F -labeled folic acid derivatives, L-type amino acid transporter 1/ ^{11}C -methionine, translocator protein (TSPO)/ ^{11}C -PBR28, mannose receptor 1 (MRC1)/2-deoxy-2- ^{18}F -fluoro-D-mannose, somatostatin receptor subtype-2 (SSTR2)/ ^{68}Ga -DOTA-TATE, fibroblast-activated protein (FAP)/ ^{68}Ga -fibroblast-associated protein inhibitor, alpha-v-beta-3 integrin (ITGAV)/arginylglycylaspartic acid (RGD), CD8+ T-cells/ ^{89}Zr -radiolabeled human CD8-specific minibody, and granzyme B (GZMB)/ ^{68}Ga -NOTA-GZP.¹⁴ To visualize the differential expression of these molecular targets between TNBC and normal epithelial cells, volcano plots were generated using the *EnhancedVolcano* function of the *EnhancedVolcano* package. To visualize expression of expression levels, UMAP plots and violin plots were generated.

Differential expression between TNBC and normal epithelial cells was assessed using Seurat's default two-sided Wilcoxon rank-sum test, and P-values were adjusted using Bonferroni correction. Because cell-level testing does not explicitly model patient-level effects and may be susceptible to pseudoreplication, these analyses were used for exploratory candidate prioritization and subsequently cross-validated in TCGA bulk RNA-seq data to ensure consistency at the patient level. A p-value < 0.05 was considered significant.

Exploration of Target Proteins in an Open Access Database

It was hypothesized that proteins exhibiting high expression in TNBC cells and present on the cell surface would be ideal candidates for targeted imaging and therapy in TNBC. Thus, the in silico *human surfaceome*, a public resource of 2,886 surface proteins that are potential targets for ligands or antibodies, was utilized (version 1.0).¹⁵ Among the DEGs identified, those included in the *Surfaceome* database were further scrutinized. Specifically, genes exhibiting a log2 fold change of 1 or greater and an adjusted p-value of less than 10^{-10} were selected as potential candidates for imaging and therapeutic targets in TNBC. Visualization of these genes and their expression was achieved through the generation of volcano plots, UMAP plots, and violin plots. Wilcoxon rank-sum tests were performed to compare expression of genes as aforementioned subsection.

Validation of Target Proteins in Bulk RNA-Seq Data

To verify the validity of potential targets identified through scRNA-seq, bulk RNA-seq data were employed from TCGA project.¹⁶ It is crucial as significant imaging and therapy implications require substantial protein expression not only at the cellular level but also at the tissue level. The breast cancer database (TCGA-BRCA) was downloaded from TCGA projects by the *GDCquery* function of the TCGAbiolinks package. Clinical information including histopathological findings was downloaded from the Cancer Genomic Data Server using the *CGDS* function of the *cgdsr* package. Among 554 samples with known histopathological findings, data from 113 normal breast tissues and 162 TNBC tissues were selected, with TNBC cases defined based on PAM50 basal-like subtype annotation. DEG analysis was conducted using the *DESeq* function of the *DESeq2* package. Batch covariates were not included because TCGA-BRCA RNA-seq data were produced using a consistent sequencing and library preparation pipeline, and substantial batch variability is not anticipated. To ascertain the differential expression of potential target genes identified in scRNA-seq between normal and TNBC tissues, volcano plots were generated. Genes with a log2 fold change greater than 1 and an adjusted p-value less than 10^{-10} were selected for further analysis. To support downstream evaluation of diagnostic separability and effect size robustness, 95% confidence intervals and area-under-the-curve (AUC) metrics were subsequently computed for these genes using bulk RNA-seq data from the TCGA cohort. As aforementioned subsections, UMAP plots and violin plots were generated and Wilcoxon rank-sum tests were performed. These transcriptomic analyses using publicly available de-identified TCGA data were reviewed by the Samsung Medical Center Institutional Review Board and determined to be exempt from approval.

Cell Culture

Human Mammary Epithelial Cells (HMEC) were selected as a control group. MDA-MB-231, MDA-MB-468, HCC-1143, and HCC-1937 were selected as representative TNBC cell lines based on transcriptomics results from human cancer cell lines.^{17,18} HMEC, MDA-MB-231, and MDA-MB-468 cells were purchased from American Type Culture Collection (Manassas, VA, USA). HCC-1143 and HCC-1937 cells were obtained from Korean Cell Line Bank (Seoul, Korea). Breast cancer cell lines were grown in RPMI 1640 medium with 25 mM hydroxyethyl piperazine ethane sulfonic acid (HEPES; Cat. #22400089, Gibco, NY, USA) supplemented with 10% fetal bovine serum and 1% antibiotics. HMEC cells were cultured in Mammary Epithelial Cell Growth Medium BulletKit (Cat. #CC-3150, Lonza, Basel, Switzerland) with 1% antibiotics. All cells were incubated at the condition of 37°C and 5% CO₂. When the cells were 80% to 90% confluent, they were sub-cultured to a fresh media.

Western Blot

To prepare whole cell and tissue lysate, cultured cell lines and homogenized tumor tissues were washed cold phosphate buffered saline (PBS) and lysed by lysis buffer containing a cold protein extraction solution (PRO-PREP; Intron, Korea) and a protease inhibitor cocktail (Sigma-Aldrich, MO, USA). Cells were sonicated three times, centrifuged for 10 minutes at 14,000 rpm and 4°C.

The protein content of the supernatants was determined using a Bradford Protein Assay. About 1~20 micrograms of total protein were separated on a 12% SDS-PAGE and then transferred onto polyvinylidene fluoride (PVDF; Cat. #IPVH00010, Merck, Germany) membranes. The membranes were blocked with Tris-buffered saline and tween-20 (TBST) containing 5% skim milk powder for 1 hour, then incubated with primary antibodies against BST-2 (mouse IgG, 1:200; Cat. #sc-390719; Santa Cruz Biotechnology, Dallas, TX, USA), LY6D (mouse IgG, 1:200; Cat. #sc-373838; Santa Cruz Biotechnology, Dallas, TX, USA), LY6E (rabbit IgG, 1:1,000; Cat. #ab300399; Abcam, Cambridge, UK) and β -actin (mouse IgG, 1:10,000; Cat. #sc-4778; Santa Cruz Biotechnology, Dallas, TX, USA) overnight at 4°C. The membranes were washed by TBST three times and incubated with an anti-mouse immunoglobulin G-horseradish peroxidase (IgG-HRP; anti-mouse IgG, 1:5,000 to 1:20,000; Cat. #7076; Cell Signaling Technology, MA, USA) and an anti-rabbit IgG-HRP (anti-mouse IgG, 1:5,000; Cat. #7074; Cell Signaling Technology, MA, USA) for 2 hours at room temperature. After washing three times with TBST, an enhanced chemiluminescence detection system (Cat. #YA354271; Thermo Fisher Scientific, MA, USA) was used to visualize immunoreactive proteins, and band intensities were quantified by Quantity One software (Bio-Rad Laboratories, CA, USA). For Western blot analysis, three independent lysates were prepared for each cell line, and tumor lysates obtained from all five xenograft models were used. The Western blot experiment was performed once.

Flow Cytometry

MDA-MB-468 cells (1.5×10^6 cells) were washed with PBS twice and resuspended in PBS containing 5% FBS, 0.1% BSA and 0.1% sodium azide. The cells were incubated for 30 min with anti-LY6E IgG (Cat. #A324132; Antibodies.com, UK) and reacted with anti-human secondary antibody conjugated Alexa Fluor 594 (Cat. #A-11014; Invitrogen, MA, USA) for 30 min at 4°C in the dark. The cells were analyzed FACS Lyric cell sorting (BD Biosciences). A sequential gating strategy was applied, including exclusion of debris, singlet gating, and final gating on viable cells. Unstained samples and isotype-matched IgG controls were included to define background fluorescence and assess non-specific binding. The Flow Jo software was used for data acquisition and analysis.

Confocal Immunofluorescence

In brief, the cells were grown on 8-well chamber slides and fixed with 4% paraformaldehyde. The samples were incubated with anti-LY6E IgG (Cat. #HL1933; GeneTex, CA, USA) overnight at 4°C. The samples were incubated with anti-rabbit FITC (Cat. #sc-2359; Santa Cruz Biotechnology) for 10 min. And then the samples were stained with DAPI solution and mounted with coverslip. The samples were examined using an LSM780 confocal microscope (Zeiss).

Tumor-Bearing Mouse Model

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of Samsung Medical Center (Approval No. 20250306002) and conducted in an AAALAC-accredited facility in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. This study was carried out in compliance with the ARRIVE guidelines. Tumor models were prepared in female Balb/c nude mice ($n = 5$, Orient Bio, Korea) by subcutaneous injection of 1.4×10^7 MDA-MB-468 cells mixed with Matrigel (Cat No.# CLS354234, Corning, MA, USA) at 1:1 ratio into the right shoulder of 4–5 weeks old. Experiments were performed when the tumor diameter reached 0.8 cm at approximately 20 days after cell inoculation. Anesthesia was induced and maintained with isoflurane (Hana Pharm, Korea), and euthanasia was performed by CO₂ inhalation in accordance with the American Veterinary Medical Association (AVMA) Guidelines for the Euthanasia of Animals. Tumors harvested from all five xenograft-bearing mice were used for Western blot analysis, and the assay was performed once per tumor lysate without additional technical replicates.

Results

Expression of Alleged Imaging Targets

UMAP analysis was conducted on merged scRNA-seq data from normal breast epithelial cells and TNBC cells. The UMAP plot revealed a clear distinction between normal and TNBC cells ([Figure 1](#)). DEG analysis was performed, with detailed results presented in [Supplementary Table 1](#). A comparison of expression for each protein showed variations between nine alleged imaging targets, as depicted in a volcano plot ([Figure 2](#)). While significant differences were observed for SLC2A1, FOLR1, TSPO, ITGAV, and SLC7A5, log₂ fold changes of these five proteins did not exceed an absolute value of 1, indicating their limited potential for selective imaging or treatment in TNBC. Furthermore, ITGAV and SLC7A5 were observed to have slightly higher expression levels in normal cells compared to TNBC cells. The expression of these five proteins were visualized in a UMAP plot and further represented in a violin plot, confirming that none were sufficiently expressed higher in TNBC ([Supplementary Figure 1](#)). The remaining four alleged imaging targets

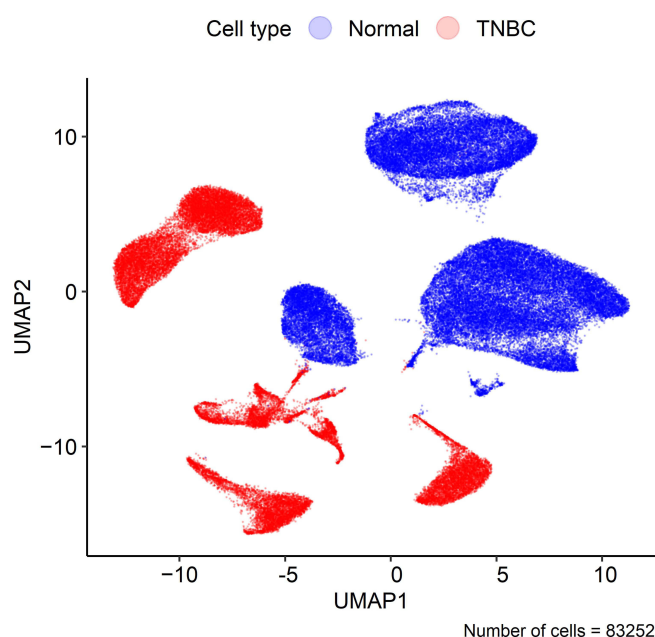


Figure 1 UMAP analysis results in scRNA-seq data. UMAP analysis was performed on combined scRNA-seq data encompassing both normal breast epithelial cells and TNBC cells. The UMAP plot illustrates distinct clustering patterns between normal and TNBC cells, highlighting their differential expression profiles.

Abbreviations: UMAP, uniform manifold approximation and projection; scRNA-seq, single-cell RNA-sequencing; TNBC, triple negative breast cancer.

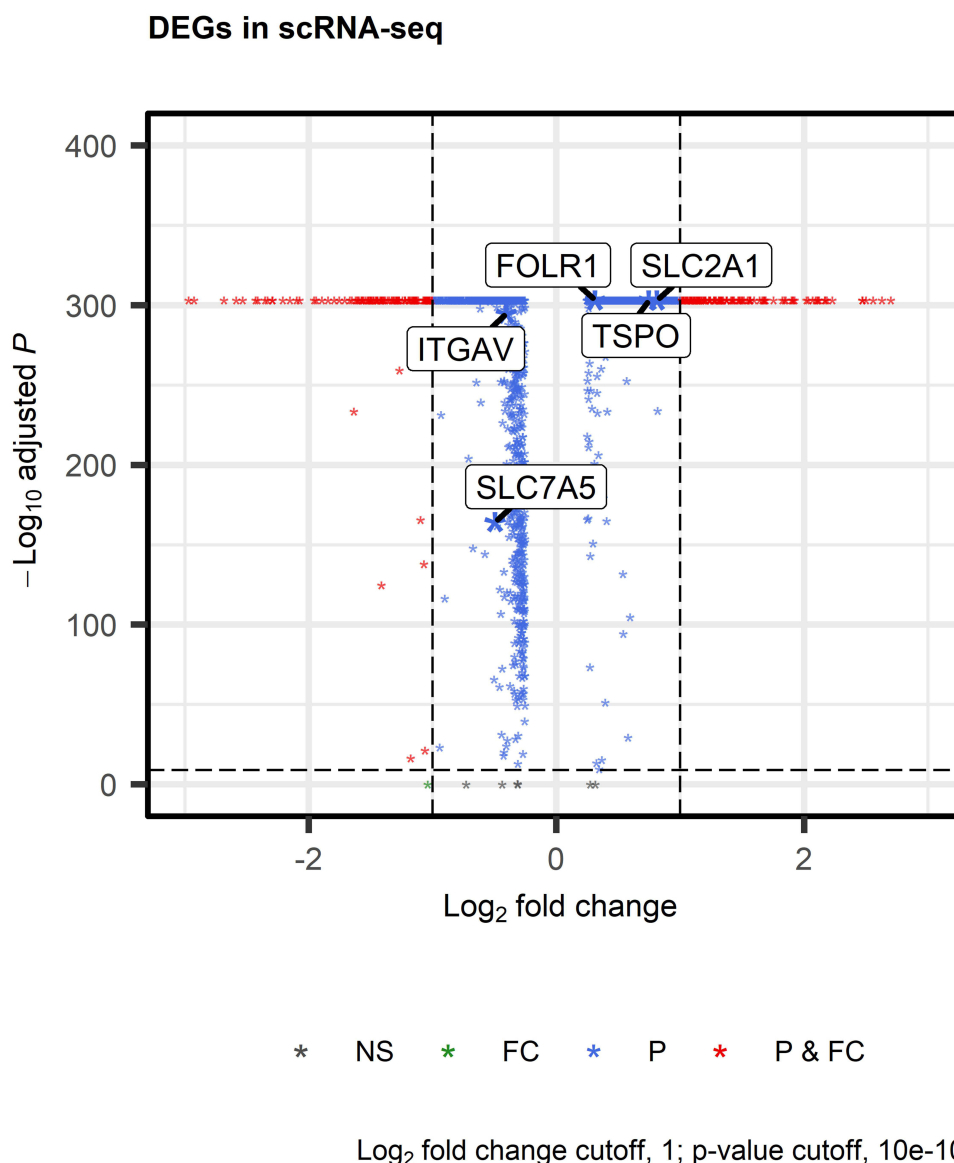


Figure 2 DEG analysis results in scRNA-seq data. A volcano plot demonstrates significant differences in expression of SLC2A1, FOLR1, TSPO, ITGAV, and SLC7A5 ($P < 10^{-10}$ for all). However, the log₂ fold changes of these proteins did not exceed an absolute value of 1. The remaining four targets with expression in less than 10% of cells in each group were excluded in the plot. Other genes with larger fold changes were not annotated because they were not part of the predefined set of previously suggested imaging targets.

Abbreviations: DEG, differentially expressed gene; scRNA-seq, single-cell RNA-sequencing.

were expressed in less than 10% of cells in each group, leading to their automatic exclusion during the execution of the *FindMarkers* function.

Exploration of Target Proteins

For the exploration of target proteins, the DEG results were filtered to select only surface proteins, and their expression differences were depicted in a volcano plot (Figure 3). Nine proteins were revealed to be expressed more than twice as much in TNBC cells compared to normal cells: LY6E, LY6D, LAMP1, EMP2, TTYH1, CD74, BST2, HLA-DRA, and HLA-DRB1. The expression levels of these proteins were also visualized on a UMAP plot and represented in a violin plot, further confirming their higher expression in TNBC cells (Figure 4).

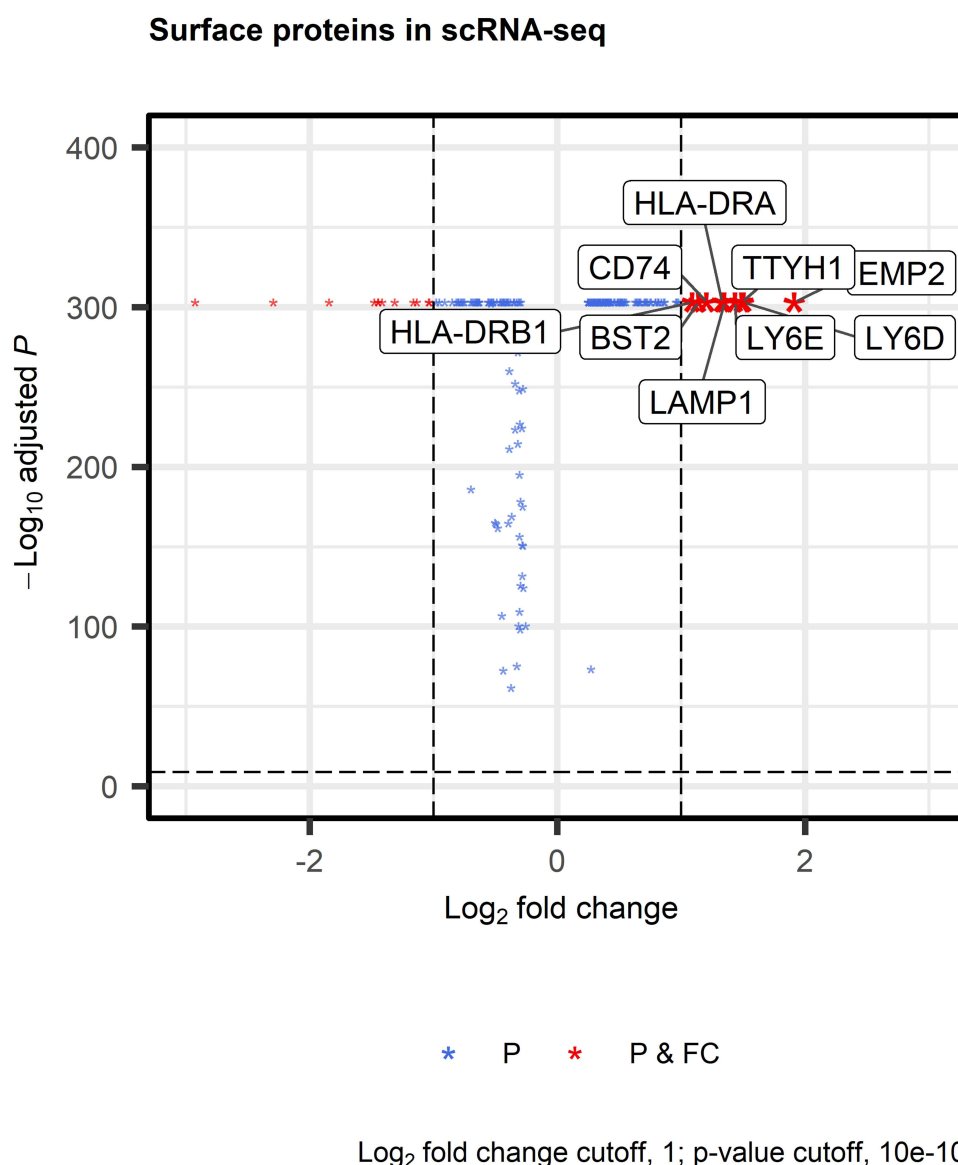


Figure 3 DEG analysis results for surface proteins in scRNA-seq data. A volcano plot demonstrates nine proteins with significantly higher expression levels in TNBC cells compared to normal cells, surpassing a twofold increase: LY6E, LY6D, LAMP1, EMP2, TTYH1, CD74, BST2, HLA-DRA, and HLA-DRB1 ($P < 10^{-10}$ for all).

Abbreviations: DEG, differentially expressed gene; scRNA-seq, single-cell RNA-sequencing; TNBC, triple negative breast cancer.

Validation of Target Proteins in Bulk RNA-Seq Data

UMAP analysis was conducted on bulk RNA-seq data, and a clear distinction was identified between TNBC and normal breast tissues as observed in scRNA-seq data ([Supplementary Figure 2](#)). Focusing on surface proteins from the DEG analysis results, their expression differences were illustrated in a volcano plot ([Figure 5](#)). Among the nine candidate proteins previously identified in scRNA-seq data, LY6E, LY6D, BST2, and TTYH1 exhibited a log₂ fold change greater than 1 with statistically significant p-values. For these genes, log₂ fold change estimates, corresponding 95% confidence intervals, and AUC values reflecting TNBC–normal tissue separability are summarized in [Table 1](#). The elevated expression of these proteins in TNBC tissues was reconfirmed in UMAP and violin plots ([Figure 5](#)). Consequently, LY6E, LY6D, BST2, and TTYH1 were selected as potential target proteins for validation in TNBC cell lines. Additionally, high expression—defined as $\log_2(\text{expression} + 1) > 1$ —was evaluated on a per-patient basis, and the prevalence of high-expressing cells for each candidate gene is reported in [Supplementary Table 2](#).

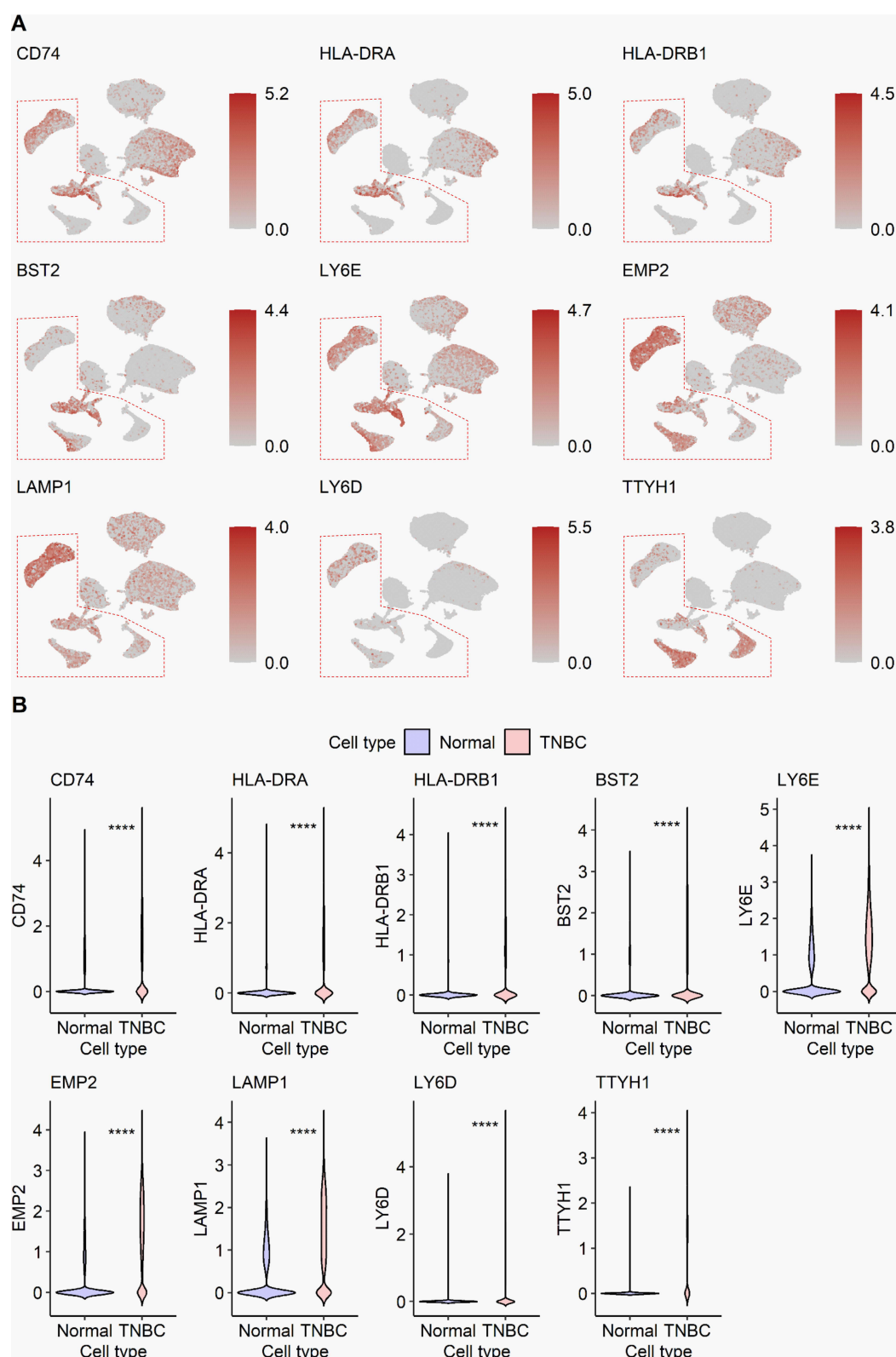


Figure 4 UMAP and violin plots for candidate proteins identified from scRNA-seq data. **(A)** The UMAP plot visually confirms significantly higher expression of the candidate proteins in TNBC. Clusters corresponding to TNBC cells are outlined with red dashed boundaries to aid visual distinction from normal epithelial cells. **(B)** Violin plots indicate significant differences for all proteins in the Wilcoxon rank-sum test ($P < 0.0001$ for all). **** denotes $P < 0.0001$.

Abbreviations: UMAP, uniform manifold approximation and projection; scRNA-seq, single-cell RNA-sequencing; TNBC, triple negative breast cancer.

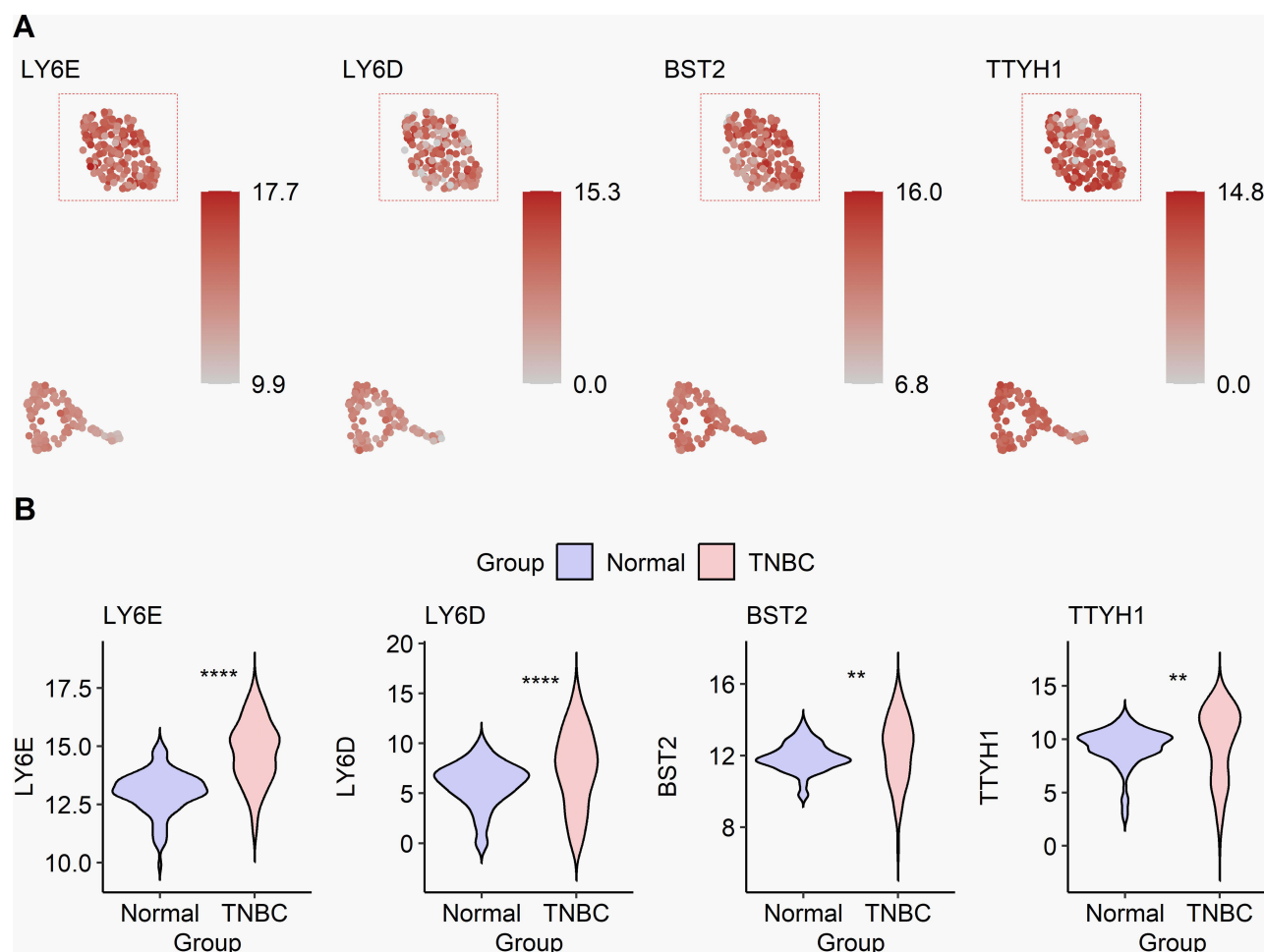


Figure 5 UMAP and violin plots for candidate proteins in bulk RNA-seq data. **(A)** The UMAP plot visually confirms significantly higher expression of the candidate proteins in TNBC. Samples corresponding to TNBC tissues are highlighted with red dashed boundaries to distinguish them from normal breast tissues. **(B)** Violin plots indicate significant differences for all proteins in the Wilcoxon rank-sum test ($P < 0.0001$ for LY6E and LY6D; $P = 0.007$ for BST2; $P = 0.002$ for TTYH1). ** and **** denote $0.001 \leq P < 0.01$ and $P < 0.0001$, respectively. **Abbreviations:** UMAP, uniform manifold approximation and projection; RNA-seq, RNA-sequencing; TNBC, triple negative breast cancer.

Validation of Target Proteins in TNBC Cell Lines

To validate the expression of potential target proteins in TNBC cell lines, Western blot analysis was conducted in both normal breast cells and TNBC cell lines. Three out of the four potential target proteins, LY6E, LY6D, and BST2, were successfully validated using monoclonal antibodies, while the validation of TTYH1 was impeded due to the

Table 1 Bulk RNA-Seq Validation Metrics for Candidate Surface Proteins Identified from Single-Cell Transcriptomic Analysis, Including Effect Size Estimates and Tissue-Level Separability

Genes	Log2 Fold Change (95% CI)	Adjusted P-value	AUC (95% CI)
LY6E	2.088 (1.828–2.348)	6.833×10^{-55}	0.141 (0.098–0.185)
LY6D	4.115 (3.508–4.721)	1.400×10^{-39}	0.355 (0.290–0.420)
BST2	1.905 (1.447–2.362)	8.883×10^{-16}	0.405 (0.337–0.473)
TTYH1	1.094 (0.897–1.399)	4.3458×10^{-12}	0.392 (0.323–0.460)

Abbreviation: CI, confidence interval.

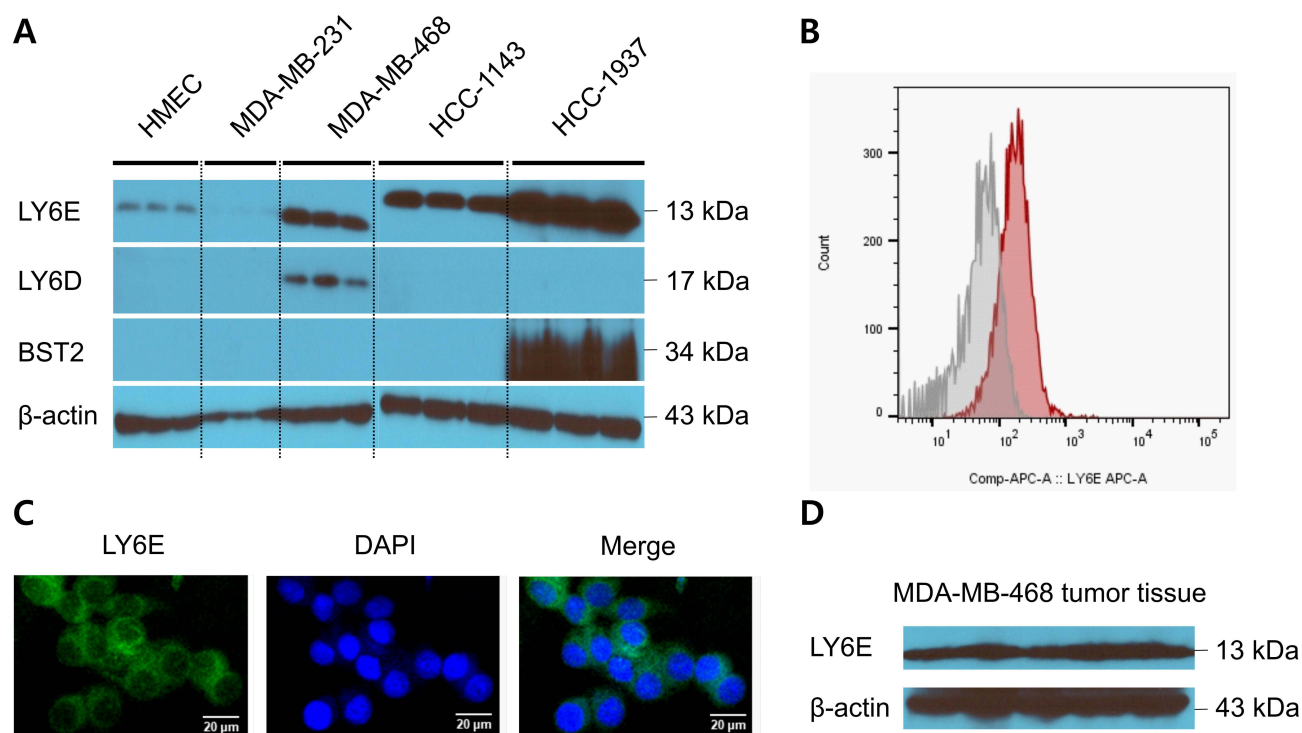


Figure 6 Validation of LY6E expression in TNBC cell lines and xenograft tumors. **(A)** Western blot analysis of whole-cell lysates from HMEC and TNBC cell lines. LY6E was detected in MDA-MB-468, HCC-1143, and HCC-1937, and showed weak expression in HMEC. **(B)** Flow cytometry analysis of MDA-MB-468 cells confirmed membrane localization of LY6E (red) compared with isotype control (gray). **(C)** Immunofluorescence confocal microscopy of MDA-MB-468 cells showing LY6E staining (green) localized to the cell membrane. Nuclei were counterstained with DAPI (blue). **(D)** Western blot analysis of tumors derived from MDA-MB-468 xenografts demonstrating consistent LY6E expression.

unavailability of a commercially accessible monoclonal antibody; a monoclonal antibody is essential for future diagnostic and therapeutic applications.

In Western blot of proteins from whole cell lysate, MDA-MB-468, HCC-1143, and HCC-1937 expressed LY6E (Figure 6A). LY6D and BST2 were detected in MDA-MB-468 and HCC-1937, respectively. In other cell lines including HMEC, there was no expression of LY6D and BST2, whereas expression was not observed in the other cell lines. The densitometry results for all proteins, normalized to β -actin, are summarized in Table 2. Among the tested TNBC cell lines, MDA-MB-468 was selected for following experiments because it is a widely used and well-characterized TNBC model that reliably supports membrane-targeted assays and xenograft establishment. To further assess membrane localization, flow cytometry and immunofluorescence confocal microscopy were performed in MDA-MB-468 cells, confirming membrane expression of LY6E (Figure 6B and C). In addition, Western blot analysis of tumors derived from an MDA-MB-453 xenograft model demonstrated consistent LY6E expression (Figure 6D).

Table 2 Densitometry Analysis of Candidate Proteins Normalized to β -Actin (Mean $\pm$ SD)

Proteins	HMEC	MDA-MB-231	MDA-MB-468	HCC-1143	HCC-1937
LY6E	0.844 $\pm$ 0.044	0.208 $\pm$ 0.041	3.800 $\pm$ 1.098	4.647 $\pm$ 0.469	8.843 $\pm$ 1.313
LY6D	0	0	1.895 $\pm$ 0.645	0	0
BST2	0	0	0	0	5.623 $\pm$ 0.141

Discussion

In this study, analyses of scRNA-seq and bulk RNA-seq were conducted to identify potential imaging and therapeutic targets for TNBC. Exploring beyond the expression levels of previously alleged targets, additional new candidate proteins were investigated using the *Surfaceome* database. It led to the identification of LY6E, LY6D, BST2, and TTYH1 as surface proteins with high expression in TNBC cells and tissues. Subsequent validation using Western blot, flow cytometry, and immunofluorescence confocal microscopy confirmed that LY6E was expressed in TNBC cell lines and localized to the cell membrane. LY6E expression was further verified in xenograft tumor tissues, underscoring its potential as a robust target for imaging and therapeutic applications.

Numerous studies have explored the transcriptomic landscape of TNBC through scRNA-seq analysis, revealing tumor heterogeneity or microenvironment which are vital for understanding tumor biology and treatment response.^{6,8,9} Furthermore, there have been previous works to evaluate aggressiveness and prognosticate outcomes by integrating bulk RNA-seq with scRNA-seq analyses.^{19,20} Despite extensive research in transcriptomics, there has been a significant lack of studies aimed at identifying proteins for potential use as imaging and therapeutic targets. By combining transcriptomics with experimental validation, this study addresses that translational gap.

This study also investigated whether proteins identified as candidates through RNA-seq analysis are highly expressed in TNBC cell lines, a crucial approach for preclinical research aimed at targeted imaging and therapy. Preclinical studies typically utilize specific cell lines to test imaging and therapeutic strategies. While existing methods focus on targeting proteins specifically expressed in cell lines, there may be a lack in verifying such expression in actual patient cells or tissues. The methodology employed in this study is of great significance because it enhances the impact of preclinical studies by providing evidence of target protein expression in the cells or tissues of patients.

The initial part of the study aimed to evaluate the expression levels of previously identified metabolic or immunologic target proteins. However, findings from scRNA-seq indicated that these proteins did not exhibit higher levels of expression in TNBC cells than normal breast cells. For targeted therapy to be effective, it is crucial that target proteins are highly expressed in tumor cells, not just in the tissue overall. Therefore, it was concluded that additional validation using bulk RNA sequencing was unnecessary. The relatively low expression of these alleged targets in TNBC cells underscores the urgent need for the discovery of new targets.

The comprehensive analysis integrating scRNA-seq, bulk RNA-seq, and the *Surfaceome* database identified LY6E, LY6D, BST2, and TTYH1 as potential target proteins. However, validation experiments demonstrated that only LY6E showed consistent expression in TNBC cell lines, with additional confirmation of its membrane localization through flow cytometry and immunofluorescence confocal microscopy. Furthermore, in vivo xenograft tumor analysis confirmed LY6E expression, supporting its robustness as a clinically relevant target. In contrast, BST2 and LY6D were not sufficiently expressed in TNBC cell lines, and TTYH1 could not be validated due to the lack of an available monoclonal antibody. These results collectively indicate that LY6E represents the most promising candidate for preclinical imaging and therapeutic studies. Notably, LY6E was not expressed in MDA-MB-231. This cell line is widely recognized as a claudin-low/mesenchymal TNBC subtype, a transcriptionally distinct state within the broader basal-like TNBC spectrum.²¹ Claudin-low tumors exhibit strong EMT signatures, low epithelial marker expression, and altered cell-surface antigen profiles.²² Thus, the absence of LY6E in MDA-MB-231 likely reflects subtype-dependent heterogeneity rather than inconsistency with scRNA-seq and TCGA findings of this study, which primarily represent classical basal-like TNBC.

LY6E is biologically relevant to breast cancer progression and immune modulation. Prior studies have associated LY6E with activation of TGF- β -related signaling pathways and with modulation of PD-L1 expression, implicating LY6E in immune evasion mechanisms and tumor aggressiveness.^{10,11} Such links suggest two potential advantages of LY6E as a theranostic target: first, its surface localization permits direct targeting for imaging or antibody-based delivery; second, its association with immune regulatory pathways raises the possibility that LY6E-targeted imaging could provide information about tumor immune status and predict responsiveness to immunotherapy.

To further establish the functional and clinical utility of LY6E, several follow-up experiments are recommended and are currently being planned or undertaken. First, preclinical evaluation of LY6E-targeted immuno-PET tracers in

xenograft or patient-derived xenograft models should be prioritized, including imaging, biodistribution, and dosimetry studies. Notably, a previous report suggested that LY6E-targeting antibodies undergo internalization, indicating potential suitability for antibody–drug conjugates.²³ Therefore, determining whether candidate immuno-PET probes or therapeutic antibodies similarly internalize—and whether this property influences imaging performance or therapeutic efficacy—is necessary prior to translational development. Second, quantitative IHC across a larger TNBC patient cohort is needed to determine target prevalence. Finally, therapeutic proof-of-concept studies using antibody-based or radioimmunotherapy constructs should evaluate antitumor efficacy and off-target toxicity in relevant models. As demonstrated in our analysis, LY6E expression is not completely absent in normal epithelial cells, suggesting the possibility of toxicity. Therefore, future validation should include safety profiling in preclinical therapeutic studies and quantitative assessment of immuno-PET performance, with criteria such as achieving a tumor-to-blood uptake ratio ≥ 3 considered a minimum requirement for clinical feasibility. Collectively, these efforts will clarify whether LY6E can serve both as an imaging biomarker predictive of immune response and as a safe, effective therapeutic target.

This study is subject to several limitations, highlighting necessity of further investigation. Firstly, the unavailability of a monoclonal antibody prevented the verification of TTYH1. While existing literature is limited in establishing a direct link between TTYH1 and TNBC, it is anticipated that future studies with adequate verification reveal its potential relevance. Secondly, scRNA-seq analysis has inherent limitations. While a powerful tool for unraveling the protein or cell landscape, it cannot unequivocally guarantee a proportional relationship between abundance of mRNA and protein expression. Therefore, validation through IHC results on TNBC tissue from patients becomes crucial to overcome this limitation. Thirdly, although the identified proteins show promise as suitable targets on the surface of TNBC cells, practical utility in preclinical or clinical trials remains uncertain. Verification is essential to confirm the feasibility of imaging agents or drug delivery targeting the identified proteins. Notably, AUC values from bulk RNA-seq did not demonstrate strong diagnostic discrimination. This outcome is expected, as bulk transcriptomic measurements average mixed cellular populations—including immune and stromal compartments—thereby diluting malignant epithelial-specific expression signals observed at the single-cell level. Furthermore, low-level physiological expression in normal tissue may reduce binary classifier performance while still being compatible with the requirements of molecular imaging or targeted therapy, where membrane accessibility and relative tumor enrichment are more relevant than absolute diagnostic separability. To address these limitations, an immuno-PET and radioimmunotherapy study targeting the proteins is being devised. Additionally, plans are underway to recruit actual TNBC patients to explore the potential applicability of BST2 or LY6E through tissue IHC. These steps are essential for a more comprehensive understanding and practical implementation of the identified target proteins in TNBC.

Conclusions

In conclusion, the integrative transcriptomic and experimental pipeline described here identified LY6E as a promising surface target for TNBC. The confirmation of LY6E expression and membrane localization in cell lines and xenograft tumors supports its prioritization for future preclinical studies, offering insights into the development of innovative diagnostic and treatment approaches for TNBC.

Abbreviations

TNBC, Triple negative breast cancer; FDG PET/CT, F-18 fluorodeoxyglucose positron emission tomography/computed tomography; IHC, immunohistochemistry; scRNA-seq, single-cell RNA sequencing; TCGA, The Cancer Genome Atlas; UMAP, uniform manifold approximation and projection; DEG, differentially expressed gene; GLUT, glucose transporter; MCT, monocarboxylate transporter; FOLR, folate receptor; TSPO, translocator protein; MRC1, mannose receptor 1; SSTR2, somatostatin receptor subtype-2; FAP, fibroblast-activated protein; ITGAV, alpha-v-beta-3 integrin; RGD, arginylglycylaspartic acid; GZMB, granzyme B; HMEC, Human Mammary Epithelial Cells; PBS, phosphate buffered saline; EDTA, ethylene-diamine-tetraacetic acid; PVDF, polyvinylidene fluoride; TBST, Tris-buffered saline and tween-20; PMSF, phenylmethylsulfonyl fluoride.

Data Sharing Statement

The processed scRNA-seq data of TNBC patients analyzed in this study is available in Figshare at <https://doi.org/10.6084/m9.figshare.17058077>. The bulk RNA-seq data of TNBC patients analyzed in this study is available in TCGA. Relevant analysis code and computational details are available from the corresponding author upon reasonable request.

Ethics Approval and Consent to Participate

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of Samsung Medical Center (Approval No. 20250306002). Transcriptomic analyses using publicly available de-identified TCGA data were reviewed by the Samsung Medical Center Institutional Review Board and determined to be exempt from approval, as obtaining informed consent was impracticable, there was no reasonable basis to assume refusal of participation, and the study posed minimal risk to research participants.

Funding

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2023-00239231).

Disclosure

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *Ca Cancer J Clin*. 2023;73(1):17–48. doi:10.3322/caac.21763
2. Howard FM, Olopade OI. Epidemiology of triple-negative breast cancer: a review. *Cancer J*. 2021;27(1):8–16. doi:10.1097/PPO.0000000000000500
3. Ulaner GA, Castillo R, Goldman DA, et al. 18 F-FDG-PET/CT for systemic staging of newly diagnosed triple-negative breast cancer. *Eur J Nucl Med Mol Imaging*. 2016;43:1937–1944. doi:10.1007/s00259-016-3402-9
4. Qin -J-J, Yan L, Zhang J, Zhang W-D. STAT3 as a potential therapeutic target in triple negative breast cancer: a systematic review. *J Exp Clin Cancer Res*. 2019;38(1):195. doi:10.1186/s13046-019-1206-z
5. Pitner MK, Taliaferro JM, Dalby KN, Bartholomeusz C. MELK: a potential novel therapeutic target for TNBC and other aggressive malignancies. *Expert Opin Ther Targets*. 2017;21(9):849–859. doi:10.1080/14728222.2017.1363183
6. Chung W, Eum HH, Lee H-O, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun*. 2017;8(1):15081. doi:10.1038/ncomms15081
7. Bartoschek M, Oskolkov N, Bocci M, et al. Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat Commun*. 2018;9(1):5150. doi:10.1038/s41467-018-07582-3
8. Karaayvaz M, Cristea S, Gillespie SM, et al. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat Commun*. 2018;9(1):3588. doi:10.1038/s41467-018-06052-0
9. Zhang Y, Chen H, Mo H, et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell*. 2021;39(12):1578–1593.e1578. doi:10.1016/j.ccell.2021.09.010
10. AlHossiny M, Luo L, Frazier WR, et al. Ly6E/K signaling to TGFβ promotes breast cancer progression, immune escape, and drug resistance. *Cancer Res*. 2016;76(11):3376–3386. doi:10.1158/0008-5472.CAN-15-2654
11. Hailin L, Yiting C, Yue W, et al. Ly6E on tumor cells impairs anti-tumor T-cell responses: a novel mechanism of tumor-induced immune exclusion. *Cancer Immunol Immunother*. 2024;74(1):4. doi:10.1007/s00262-024-03851-x
12. Chen Y, Pal B, Lindeman GJ, Visvader JE, Smyth GK. R code and downstream analysis objects for the scRNA-seq atlas of normal and tumorigenic human breast tissue. *Sci Data*. 2022;9(1):96. doi:10.1038/s41597-022-01236-2
13. Pal B, Chen Y, Vaillant F, et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J*. 2021;40(11):e107333. doi:10.15252/embj.2020107333
14. Lee H, Park J, Im H-J, Na KJ, Choi H. Discovery of potential imaging and therapeutic targets for severe inflammation in COVID-19 patients. *Sci Rep*. 2021;11(1):14151. doi:10.1038/s41598-021-93743-2
15. Bausch-Fluck D, Goldmann U, Müller S, et al. The in silico human surfaceome. *Proc Natl Acad Sci USA*. 2018;115(46):E10988–E10997. doi:10.1073/pnas.1808790115
16. Weinstein JN, Collisson EA, Mills GB, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet*. 2013;45(10):1113–1120. doi:10.1038/ng.2764
17. Jin H, Zhang C, Zwahlen M, et al. Systematic transcriptional analysis of human cell lines for gene expression landscape and tumor representation. *Nat Commun*. 2023;14(1):5417. doi:10.1038/s41467-023-41132-w
18. Human Protein Atlas. Available from: <https://www.proteinatlas.org>. Accessed January 28, 2026.
19. Bao X, Shi R, Zhao T, et al. Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels tumour heterogeneity plus M2-like tumour-associated macrophage infiltration and aggressiveness in TNBC. *Cancer Immunol Immunother*. 2021;70(1):189–202. doi:10.1007/s00262-020-02669-7

20. Wang X, Chen H. Prognosis prediction through an integrated analysis of single-cell and bulk RNA-sequencing data in triple-negative breast cancer. *Front Genet.* **2022**;13:928175. doi:10.3389/fgene.2022.928175
21. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* **2011**;121(7):2750–2767. doi:10.1172/JCI45014
22. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* **2010**;12(5):R68. doi:10.1186/bcr2635
23. Asundi J, Crocker L, Tremayne J, et al. An antibody–drug conjugate directed against lymphocyte antigen 6 complex, locus E (LY6E) provides robust tumor killing in a wide range of solid tumor malignancies. *Clin Cancer Res.* **2015**;21(14):3252–3262. doi:10.1158/1078-0432.CCR-15-0156

Breast Cancer: Targets and Therapy

Publish your work in this journal

Breast Cancer - Targets and Therapy is an international, peer-reviewed open access journal focusing on breast cancer research, identification of therapeutic targets and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/breast-cancer—targets-and-therapy-journal>

Dovepress
Taylor & Francis Group