



Research Article

Single-cell and multi-omics integration reveals cholesterol biosynthesis as a synergistic target with HER2 in aggressive breast cancer

Tzu-Yang Tseng^{a,} , Chiao-Hui Hsieh^{a,} , Jie-Yu Liu^a, Hsuan-Cheng Huang^{b,*,} ,
Hsueh-Fen Juan^{a,c,d,e,**}

^a Department of Life Science, National Taiwan University, Taipei, Taiwan

^b Institute of Biomedical Informatics, National Yang Ming Chiao Tung University, Taipei, Taiwan

^c Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

^d Center for Computational and Systems Biology, National Taiwan University, Taipei, Taiwan

^e Center for Advanced Computing and Imaging in Biomedicine, National Taiwan University, Taipei, Taiwan

ARTICLE INFO

Keywords:

Single-cell RNA sequencing
Combinatorial therapeutic targets
Cholesterol biosynthesis
Breast cancer

ABSTRACT

Breast cancer stands as one of the most prevalent malignancies affecting women. Alterations in molecular pathways in cancer cells represent key regulatory disruptions that drive malignancy, influencing cancer cell survival, proliferation, and potentially modulating therapeutic responsiveness. Therefore, decoding the intricate molecular mechanisms and identifying novel therapeutic targets through systematic computational approaches are essential steps toward advancing effective breast cancer treatments. In this study, we developed an integrative computational framework that combines single-cell RNA sequencing (scRNA-seq) and multi-omics analyses to delineate the functional characteristics of malignant cell subsets in breast cancer patients. Our analyses revealed a significant correlation between cholesterol biosynthesis and HER2 expression in malignant breast cancer cells, supported by proteomics data, gene expression profiles, drug treatment scores, and cell-surface HER2 intensity measurements. Given previous evidence linking cholesterol biosynthesis to HER2 membrane dynamics, we proposed a combinatorial strategy targeting both pathways. Experimental validation through clonogenic and viability assays demonstrated that simultaneous inhibition of cholesterol biosynthesis (via statins) and HER2 (via Neratinib) synergistically reduced malignant breast cancer cells, even in HER2-negative contexts. Through systematic analysis of scRNA-seq and multi-omics data, our study computationally identified and experimentally validated cholesterol biosynthesis and HER2 as novel combinatorial therapeutic targets in breast cancer. This data-driven approach highlights the potential of leveraging multiple molecular profiling techniques to uncover previously unexplored treatment strategies.

1. Introduction

Breast cancer is the most prevalent malignancy among women and is generally curable in its early, non-metastatic stages; however, treatment becomes challenging in advanced stages with distant organ metastases

[1]. Clinically, breast cancer patients are categorized into several subtypes based on the expression of hormone receptors (estrogen receptor and progesterone receptor), human epidermal growth factor receptor 2 (HER2), and other breast cancer-related genes [2]. These molecular subtypes include Luminal A (LumA), Luminal B (LumB), HER2-enriched

Abbreviations: **HER2**, Human epidermal growth factor receptor 2; **LumA**, Luminal A; **LumB**, Luminal B; **scRNA-seq**, Single-cell RNA-sequencing; **CCL**, Cancer Cell Line Encyclopedia; **EGA**, European Genome-phenome Archive; **MAD**, Median absolute deviation; **HVGs**, Highly variable genes; **PCA**, Principal component analysis; **PCs**, Principal components; **UMAP**, Unified manifold approximation and projection; **DEGs**, Differentially expressed genes; **AUC**, Area under the curve; **KEGG**, Kyoto Encyclopedia of Genes and Genomes; **TCGA**, The Cancer Genome Atlas; **ATCC**, American Type Culture Collection; **DMEM**, Dulbecco's Modified Eagle Medium; **FBS**, Fetal bovine serum; **OD**, Optical density; **ZIP**, Zero interaction potency; **HMG-CoA**, 3-hydroxy-3-methylglutaryl-CoA; **HMGCR**, 3-hydroxy-3-methylglutaryl-CoA reductase; **SQLE**, Squalene epoxidase; **CAV1**, Caveolin-1; **FDA**, Food and Drug Administration.

* Correspondence to: Institute of Biomedical Informatics, National Yang Ming Chiao Tung University, No.155, Sec.2, Linong Street, Taipei 112, Taiwan.

** Correspondence to: Department of Life Science, Graduate Institute of Biomedical Electronics and Bioinformatics, Center for Computational and Systems Biology, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan.

E-mail addresses: hsuancheng@nycu.edu.tw (H.-C. Huang), yukijuan@ntu.edu.tw (H.-F. Juan).

<https://doi.org/10.1016/j.csbj.2025.04.030>

Received 26 December 2024; Received in revised form 17 April 2025; Accepted 23 April 2025

Available online 24 April 2025

2001-0370/© 2025 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and triple-negative/basal-like [3], and this classification facilitates treatment decisions and prognosis prediction for breast cancer patients [4]. Current breast cancer therapies include a combination of surgery, chemotherapy, radiation therapy, hormonal therapy, and targeted therapy, tailored according to these subtypes [5,6]. While these treatments have improved outcomes for many patients, the emergence of drug resistance and adverse side effects underscores the need for novel therapeutic options that offer improved efficacy and reduced toxicity.

Recently, single-cell RNA-sequencing (scRNA-seq) has emerged as a powerful technology allowing gene expression profiling at single-cell resolution, providing unprecedented insights into the heterogeneity and complexity of tumor cell populations and underlying molecular mechanisms [7,8]. Inspecting the gene expression profiles of individual cells enables researchers to pinpoint differentially expressed genes and pathways within malignant tumor cells, revealing the extent of tumor heterogeneity. Such molecular characterization is invaluable for drug discovery, as it facilitates the identification of potential therapeutic agents capable of reversing the aberrant expression patterns, thereby paving the way for more effective cancer treatments [9]. Furthermore, combination therapies have been recognized as a promising approach to cancer treatment, offering improved outcomes, reduced toxicity, and delayed onset of resistance compared to single-agent therapies [10]. Consequently, leveraging scRNA-seq data in conjunction with combination therapies holds significant promise for advancing cancer treatment strategies and improving patient prognosis and outcomes.

Here, we aim to identify novel combinatorial treatment targets by specifically investigating differentially expressed pathways within malignant breast cancer cells. We analyze single-cell datasets from breast cancer patients and utilize well-known models, including PAM50 and Gene70, to classify cells according to subtype and malignancy, respectively [11]. We observe that cholesterol biosynthesis-related pathways are highly expressed in malignant cells. Elevated cholesterol biosynthesis in breast cancer has been identified as a key characteristic associated with increased tumor growth, invasion, stem cell propagation, and poor prognosis [12,13]. Additionally, we analyze the proteomics data from the Cancer Cell Line Encyclopedia (CCLE) to glean deeper insights into the relationship between cholesterol biosynthesis and ERBB family in breast cancer [14]. We discover a novel connection between cholesterol biosynthesis and HER2 in malignant breast cancer. HER2 is a well-studied protein in breast cancer that, when over-expressed, drives tumor aggression and increases sensitivity to targeted therapies [15]. Finally, we collect gene expression profiles, drug treatment scores, and cell-surface HER2 intensity data from breast cancer cell lines to reveal the impact of cholesterol biosynthesis on HER2 signaling, highlighting these interconnected pathways as potential novel combinatorial treatment targets.

2. Materials and methods

2.1. Data collection

Single-cell RNA sequencing read count data from 31 untreated breast cancer patients were retrieved from <http://biokey.lambrechtslab.org> and are also accessible from the European Genome-phenome Archive (EGA) under accession number EGAS00001004809 (Supplementary Fig. 1A) [16]. For external validation, we obtained single-cell RNA-sequencing read count data, integrated across 26 breast cancer patients, from Gene Expression Omnibus (GEO) under accession number GSE176078 (Supplementary Fig. 1B) [17]. The transcriptomics data of 65 breast cancer cell lines, proteomics data (reverse phase protein assay) of 59 breast cancer cell lines, and drug score of Neratinib, an EGFR and ERBB2 dual inhibitor, were downloaded from the CCLE and PRISM databases via the DepMap public data portal (<https://depmap.org/portal/>) [14,18].

2.2. scRNA-seq preprocessing and clustering

The single-cell gene expression matrices from 31 untreated patients were subjected to quality control using the R package *Scater* (v1.22.0) [19]. Following the pre-processing pipeline, we first calculated the quality control matrix, specifically subsetting mitochondria-related genes. Cells deviating more than three times the median absolute deviation (MAD) from the median were considered as outliers and removed. Furthermore, we removed low-quality cells with a mitochondrial gene count ratio greater than 10 % or gene counts per cell lower than 100. Genes expressed in fewer than 100 cells were also filtered out. Subsequently, we utilized the R package *Seurat* (v4.0.3) for normalization, cell clustering implementation, and further analysis [20]. We first performed SCTransform with 2000 selected highly variable genes (HVGs) and set the percentage of mitochondrial genes as the variable to be regressed. To integrate single cells from different patients, we conducted data integration using the R package *Harmony* (v0.1.0) after principal component analysis (PCA) [21]. The 30 most informative principal components (PCs) identified through Harmony integration were used for two-dimensional visualization using Unified Manifold Approximation And Projection (UMAP) [22]. Cell clustering was performed by constructing a shared nearest neighbor graph, and clusters were assigned to each cell using the Leiden algorithm implemented in Seurat [23].

2.3. Cell type annotation and carcinoma cell identification

To identify the breast carcinoma cells, we followed the annotation provided in the original publication and checked it against well-known cell type markers in breast cancer, such as carcinoma cells (*EPCAM*, *KRT18* and *KRT19*), B cells (*CD19*, *IGHG1* and *IGKC*), T cells (*CD3D*, *CD3E* and *CD3G*), myeloid cells (*CD33*, *CD68* and *ITGAX*), fibroblasts (*COL1A2*, *DCN* and *FAP*), endothelial cells (*CD34*, *CLDN5* and *PECAM1*), plasmacytoid dendritic cells (*CLEC4C*, *IRF7* and *IRF8*) and mast cells (*TPSAB1*, *TPSB2* and *MS4A2*). The carcinoma cells were subsetting, normalized, re-integrated, re-clustered, and visualized on UMAP based on the 20 most informative PCs, as described in the previously mentioned pipeline.

2.4. Copy number variation analysis

To validate the accuracy of our cancer cell clustering and cell type annotation, we applied inferCNV, a computational tool designed to infer large-scale chromosomal copy number variations (CNV) from single-cell RNA sequencing data [24]. Using myeloid cells as the reference population, we estimated relative CNV signals across the genome in each cell, allowing us to distinguish malignant cancer cells from non-malignant populations.

2.5. Breast cancer subtype classification using the PAM50 model

We categorized 24 carcinoma clusters into 5 distinct breast cancer subtypes, including basal-like, HER2-enriched, LumA, LumB, and normal-like, using the PAM50 model [25]. We calculated the average expression level of each carcinoma cluster and utilized *molecular.subtyping* function in the R package *genefu* (v2.26.0) for breast cancer subtype classification [26]. In *molecular.subtyping*, we first calculated the correlation coefficient of each sample for each breast cancer subtype:

$$r_s = \frac{\sum (T_g - T)(C_{s,g} - C_s)}{\sqrt{\sum (T_g - T)^2 \sum (C_{s,g} - C_s)^2}} \quad (1)$$

where r_s is the correlation coefficient for subtype s , T_g is the expression of gene g in sample T , T is the mean expression across all genes in sample T , $C_{s,g}$ is the centroid value for subtype s and gene g , and C_s is the mean centroid value for subtype s .

Next, we performed subtype assignment based on the correlation coefficient of each sample:

$$\text{Predicted subtype} = \text{argmax}(r_s) \quad (2)$$

where $s \in \{\text{Basal, LumA, LumB, HER2, Normal}\}$.

We then mapped the cluster-based classification results back to the carcinoma Seurat object and used the R package *ComplexHeatmap* (v2.10.0) to draw a heatmap visualizing the z-score transformed expression levels of PAM50 genes [27].

2.6. SCSubtype analysis

To ensure accurate molecular subtyping of breast cancer cells in our single-cell dataset, we utilized SCSubtype, a method designed for subtyping scRNA-seq data [17]. Following the default parameters, we applied SCSubtype to our datasets and scaled the SCSubtype scores based on established gene expression signatures. The highest scoring subtype was then assigned to each cell, ensuring accurate classification of breast cancer subtypes at the single-cell level.

2.7. Cell-cell communication analysis

To reveal subtype-specific differences in cell-cell communication, we performed analyses using the R package *CellChat* (v2.1.2) [28]. We calculated the interactions and quantified the strength between each cell type and breast cancer subtype using the Hill function:

$$\text{Interaction strength} = \frac{LR^n}{K_h^n + LR^n} \quad (3)$$

where LR is the cross-product of ligand and receptor expression, K_h is the half-maximal activation constant, and n is the Hill coefficient.

In the CellChat analysis, patterns refer to the global communication trends observed between different cell populations and are categorized into incoming and outgoing patterns. To determine the number of outgoing and incoming patterns, we calculated the Cophenetic and Silhouette values:

$$\text{Cophenetic correlation coefficient} = \frac{\text{Cov}(D, D_{\text{cophenetic}})}{\sigma(D)\sigma(D_{\text{cophenetic}})} \quad (4)$$

where D is the distance matrix converted from the consensus matrix $D = 1 - C$ and $D_{\text{cophenetic}}$ is the cophenetic distances of hierarchical clustering performed on D .

$$\text{Silhouette width} = \frac{1}{n} \sum_{i=1}^n \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (5)$$

where $a(i)$ is the average distance and $b(i)$ is the minimum distance between i and all points in that cluster.

2.8. Characterization of carcinoma malignancy

To characterize the malignancy of carcinoma cell clusters, we utilized the 70-gene signature from the Gene70 model, also known as MammaPrint, from geneu [26]. Specifically, we focused on genes associated with poor prognosis ($\text{average.good.prognosis.profile} < 0$) and calculated the module score using the *AddModuleScore* function in Seurat to quantify the malignancy of each single-cell cluster [11,29]. We utilized the module score to measure the relative expression enrichment of the poor prognosis gene signature in individual tumor cell clusters. This score does not directly correspond to the MammaPrint risk classification, which integrates weighted gene expression patterns into a prognostic model for patient stratification. Instead, our module score serves as a quantitative measure of pathway activation, enabling the identification of malignant clusters based on the expression of poor prognosis genes.

$$\text{Module score} = \frac{1}{|M|} \sum_{j \in M} X_{ij} - \frac{1}{|C|} \sum_{j \in C} X_{ij} \quad (6)$$

where M is a set of genes in the module, C is the control features randomly selected from each bin, and X_{ij} is the expression level of gene j in cell i .

2.9. Differential expression analysis

To identify the differentially expressed genes (DEGs) that are highly expressed in malignant clusters, we performed the Wilcoxon rank-sum test using the *FindMarkers* function in Seurat. This function selects DEGs between two clusters or groups of cells based on their p-value ($p < 0.05$) and fold change ($FC > 1.5$). In our study, we conducted the differential expression analysis on malignant clusters, specifically clusters 8 and 19, compared to the normal-like cluster.

2.10. Functional enrichment analysis

To identify enriched pathways at the single-cell level, we employed the R package *AUCCell* (v1.16.0) for pathway enrichment analysis [30]. *AUCCell* quantifies the enrichment level of a gene set by calculating the area under the curve (AUC) based on a ranked gene list. Gene sets were obtained from the Molecular Signature Database using the *getGeneSets* function in the R package *escape* (v1.4.0), which includes 186 curated gene sets derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) and 1615 gene sets from the Reactome pathway database [31–33]. Genes were ranked based on the normalized count matrix, and AUC scores for each gene set in each cell were calculated using the *AUCCell_calcAUC* function with default parameters. These scores were then visualized using violin and bubble plots to illustrate the fold enrichment and statistical significance between the two groups.

2.11. Survival analysis

Performing survival analysis to assess the impact of specific genes on patient prognostic outcomes, we utilized GEPIA, a tool that examines gene expression profiles, and overall survival data of 1071 breast cancer patients from The Cancer Genome Atlas (TCGA) database [34]. Furthermore, to combine the expression level of cholesterol biosynthesis-related genes, including *HMGCR*, *MVD*, *FDFT1*, and *SQLE*, we utilized the Kaplan-Meier plotter and compared the difference between breast cancer subtypes [35].

2.12. Proteomic analysis of breast cancer cell lines

We performed PCA on the proteomics data from 59 breast cancer cell lines using the R package *PCAtools* (v2.6.0). The results were first visualized on scatter plots depicting PC1 and PC2, where cell lines are colored according to the expression level of estrogen and HER2. To explore the similarities between cell lines and malignant clusters, we extracted the basal (ER-, HER2-) and luminal breast cancer cell lines (ER+, HER2-) and calculated the Pearson correlation to elucidate the relationship between cholesterol biosynthesis-related proteins (HMDH, KIME, MVD1, GGPPS, FDFT and ERG1) and ERBB family proteins (EGFR, HER2, ERBB3 and ERBB4) in the two groups of breast cancer cell lines, respectively.

2.13. Gene set variation analysis on CCLE transcriptomics data

To reveal the activation level of cholesterol biosynthesis level of each HER2-negative breast cancer cell line, we first collected the gene expression profiles of 26 HER2-negative breast cancer cell lines from the CCLE database. Then, we used the Reactome and KEGG databases to conduct gene set variation analysis (GSVA) using the R package *GSVA*

(v1.42.0) on the cholesterol biosynthesis pathway and steroid pathway, respectively [36].

2.14. Relationship between cholesterol biosynthesis and cell-surface HER2

We collected the PRISM score for Neratinib, a dual inhibitor of EGFR and ERBB2, via the DepMap portal and the cell-surface HER2 intensity data on breast cancer cell lines from a published study [37]. We performed regression analysis using linear model to examine the relationship between cholesterol biosynthesis and cell-surface HER2 expression.

2.15. Cell culture

Our study primarily focused on Basal-like and Luminal A subtypes due to their distinct reliance on cholesterol biosynthesis. To investigate this, we utilized basal-like cell line MDA-MB-231 (HTB-26) and luminal A cell line T-47D (HTB-133), both of which were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA). All cell lines tested negative for mycoplasma and were maintained in Dulbecco's Modified Eagle Medium (DMEM; Thermo Scientific, Waltham, MA, USA) supplemented with 10 % fetal bovine serum (FBS; Thermo Scientific, Waltham, MA, USA), at 37°C in a humidified incubator with 5 % CO₂.

2.16. Clonogenic assay

Cells (3000 cells per well for the MDA-MB-231 cell line and 6000 cells per well for the T-47D cell line) were seeded in 6-well plates and treated with two drug combination: either NB-598 (HY-16343C; MCE, Princeton, NJ, USA), an inhibitor of SQLE, and Neratinib (HY-32721; MCE, Princeton, NJ, USA) or Atorvastatin (HY-17379; MCE, Princeton, NJ, USA), an inhibitor of HMGCR, and Neratinib. The MDA-MB-231 and T-47D cell lines underwent clonogenic assay for 9 and 12 days, respectively. The culture media (with or without drugs) were refreshed every 3 days. Colonies were fixed with 100 % methanol and stained with 1 % crystal violet (Sigma-Aldrich, Burlington, MA, USA) in 30 % MeOH for 10 minutes. Each well was then washed three times with ddH₂O to remove excess crystal violet. A digital scanner was used to record the colony formations. Colonies were extracted with 10 % acetic acid, and the optical density (OD) values were measured by a Model 680 Microplate Reader (BIO-RAD, Hercules, CA, USA) at 590 nm. Relative growth rate was calculated as the OD value of each treatment relative to the control (DMSO). Growth inhibition was calculated as one minus the relative growth. Technical replicates were performed for each of the two independent biological replicates.

2.17. Analysis of drug synergy

Drug synergy refers to the phenomenon where the combined effect of two or more drugs is greater than the sum of their individual effects. To verify the combinatorial effect of two drugs, we calculated the synergy scores using the ZIP function in the R package *synergyfinder* (v3.2.10), following the zero interaction potency (ZIP) as described below [38]:

$$ZIP_{synergy} = \frac{1}{n} \sum_{i=1}^n Pred_i - \left(1 - \prod_{i=1}^n \left(1 - \frac{E_i}{100} \right) \right) \times 100 \quad (7)$$

where $Pred_i$ is the predicted response for drug i at a given combination of doses, E_i is the predicted response of drug i at a given dose, determined by fitting a dose-response curve for each single drug, and n is the number of drugs in the combination.

2.18. Statistical analysis

We employed the two-tailed Wilcoxon rank-sum test to assess the

significance of differences between clusters in terms of malignant scores, gene expressions, and pathways. The Wilcoxon rank-sum test was conducted using the *FindMarkers* function in Seurat. Additionally, survival outcomes and the prognostic value of marker genes were determined using the Kaplan-Meier method and the corresponding log-rank test. Regression analysis was performed using the *geom_smooth* function with a linear model in the R package ggplot2 (v3.3.5).

3. Results

3.1. scRNA-seq profiling of breast cancer patient samples

Single-cell sequencing offers a powerful approach to profile molecular characteristics at single-cell resolution, providing valuable insights into cellular identity and functionality. In this study, we focused on breast cancer, a prevalent malignancy affecting millions of individuals globally. To uncover the mechanisms underlying malignant breast cancer cells and inform prognosis-guided therapeutic approaches, we analyzed the scRNA-seq profiles of tumor tissues from 31 treatment-naïve breast cancer patients [16]. After rigorous quality control and normalization, we clustered the cells into 27 distinct clusters, visualized through two-dimensional UMAP. The top three differentially expressed genes in each cluster were identified, further revealing key molecular signatures associated with each cell population (Fig. 1A, Supplementary Table 1). We manually selected well-known cell type markers, including carcinoma markers (*EPCAM*, *KRT18*, and *KRT19*), and visualized their expression levels on UMAP (Fig. 1B, C). Cells were then categorized into 8 cell types (Fig. 1D). Moreover, we revealed distinct compositions of each cell type among patients, underscoring differences that merited deeper exploration through single-cell analysis (Fig. 1E). To distinguish cancer cells from normal cell populations, we utilized inferCNV to infer large-scale DNA copy number variations at the single-cell level. Our analysis revealed that cancer cells exhibited higher levels of CNVs compared to non-malignant cell types, especially in regions on chromosomes 8 and 15, confirming the accuracy of our clustering and cell type annotation results (Supplementary Fig. 2).

3.2. Identification of malignant breast cancer clusters

Focusing specifically on the malignant breast carcinomas, we sub-setted, re-clustered, and mapped the cancer cells onto UMAP (Fig. 2A). Each cancer cell cluster consisted of cells from multiple patients, demonstrating that the clustering was driven by shared transcriptional features rather than patient-specific effects, thereby capturing common molecular characteristics of malignant breast cancer (Supplementary Fig. 3A). Additionally, we identified the top three differentially expressed genes in each cluster, further highlighting key molecular signatures associated with distinct cancer cell populations (Supplementary Table 2). We then utilized the PAM50 model for breast cancer subtype prediction. The PAM50 method efficiently classifies breast cancer subtypes, including basal, HER2-enriched, LumA, LumB, and normal-like, based on gene expression profiles, and is widely used in research and clinical practice. We conducted cluster-based breast cancer subtyping, and the expression levels of 50 signatures were visualized in a heatmap, showing that cluster 8 (basal-like) exhibited elevated expression of numerous PAM50 genes. Subtype prediction results were mapped and visualized on UMAP (Fig. 2B, C). To validate the accuracy of our breast cancer subtyping, we additionally performed SCS subtype analysis, a method specifically designed for scRNA-seq data. The results were consistent with our initial PAM50 classification, confirming that cluster 8 remained predominantly composed of basal-like cells, while cluster 19 was primarily luminal A cells (Supplementary Fig. 3B). Furthermore, we performed CellChat analysis to investigate cell-cell communication. We discovered that LumA and basal breast cancer subtypes exhibited significantly stronger intercellular interactions compared to other subtypes, as illustrated by circle plots showing the interaction numbers and

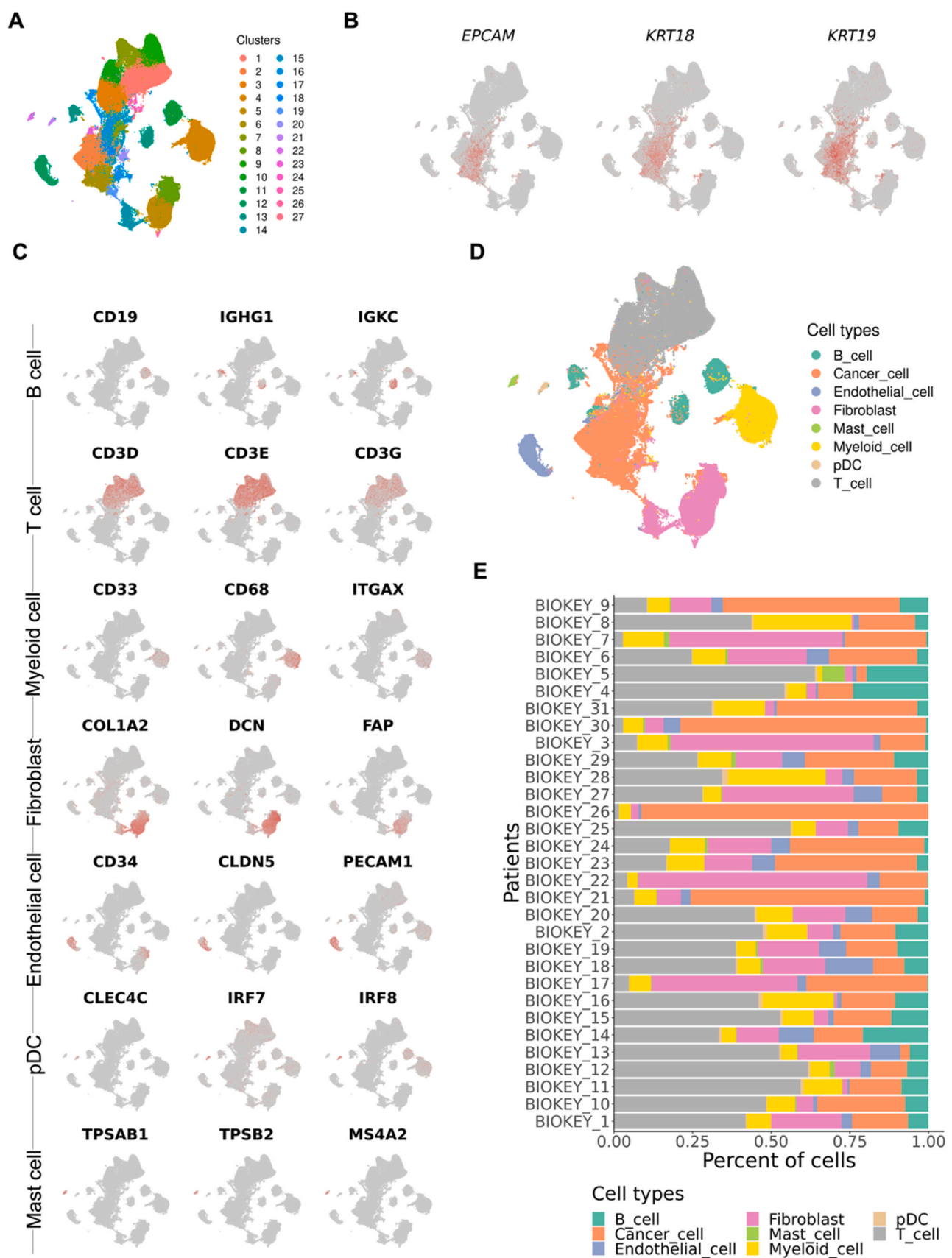


Fig. 1. scRNA-seq profiling of breast cancer cells from 31 untreated patients. (A) UMAP visualization shows 27 distinct clusters from the single-cell data of 31 breast cancer patients. (B) Cells were colored by the expression level of three well-known breast cancer cell markers, EPCAM, KRT18, and KRT19. (C) Well-known cell type markers, including B cells, T cells, and fibroblasts, were manually selected and mapped onto UMAP. (D) UMAP unveils the annotation results of each cell type in the single-cell dataset. (E) Bar plot shows the cell type composition of each patient.

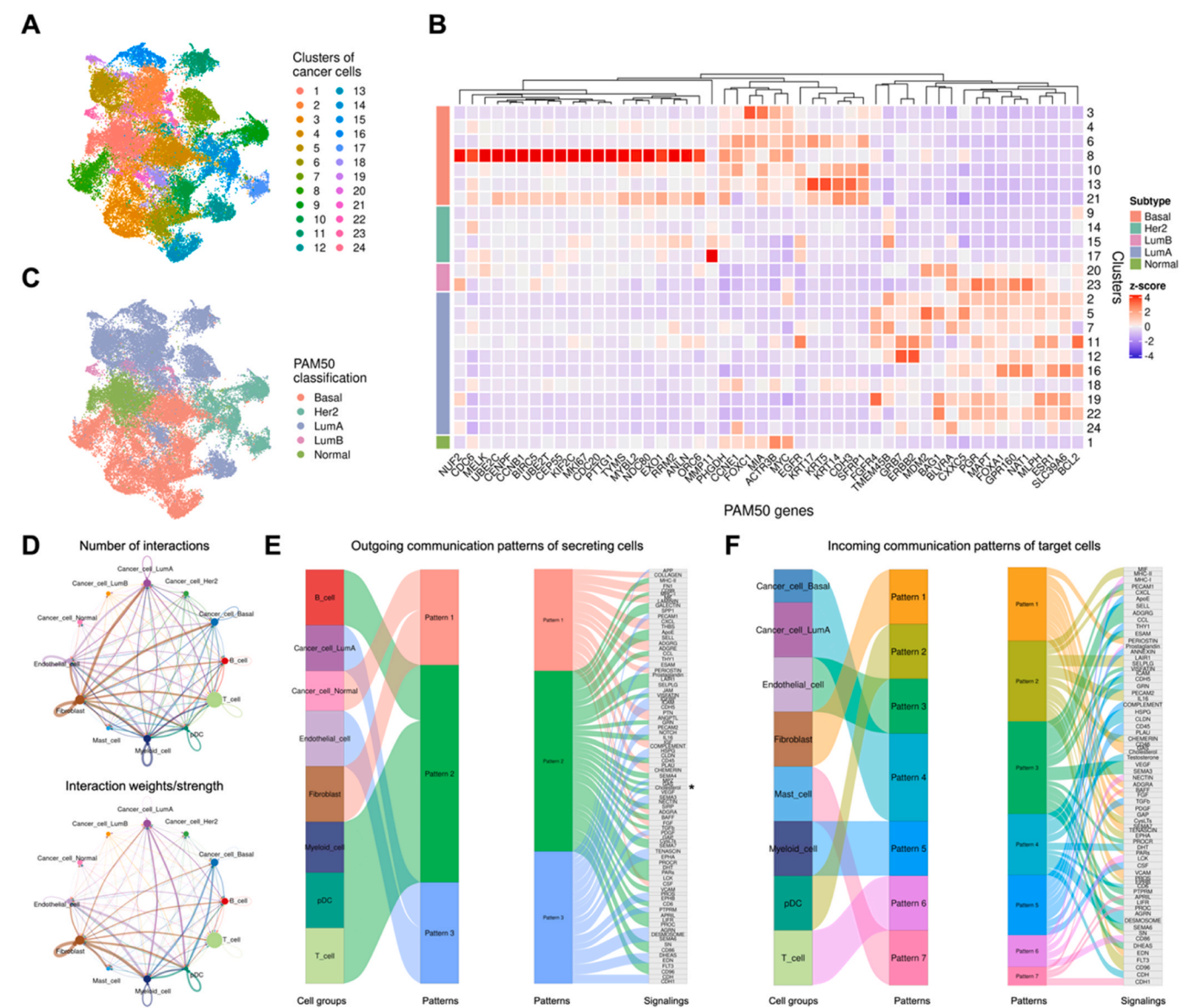


Fig. 2. Annotation of breast cancer subtype and identification of malignant breast cancer clusters. (A) Cancer cells were re-clustered through the Leiden algorithm and projected onto UMAP. (B) We applied the PAM50 model to classify single-cell clusters into 5 breast cancer subtypes: basal, HER2, LumA, LumB, and normal. (C) Cancer cells were recolored based on the results of subtype annotation and projected on UMAP. (D) The circle plot reveals the number and strength of interactions between each cell type and breast cancer subtype. (E) The river plot shows the outgoing communication patterns of secreting cells (number of patterns = 3). (F) The river plot displays the incoming communication patterns of target cells (number of patterns = 7).

strengths (Fig. 2D, Supplementary Fig. 4A). A two-dimensional plot of cell types and subtypes based on outgoing and incoming strength also corroborated these pronounced interactions (Supplementary Fig. 4B). Cophenetic and silhouette values were calculated to determine outgoing and incoming communication patterns (Supplementary Fig. 4C). We observed that LumA and basal malignant subtypes displayed a unique incoming communication pattern of target cells (Fig. 2E, F). Notably, we did not detect any communication in HER2-positive breast cancer cells. Together, these findings provided a comprehensive characterization of intercellular communication networks across breast cancer subtypes, highlighting distinctive communication patterns of each subtype.

3.3. Characterization of malignant breast cancer clusters

To further assess the malignant status of breast cancer clusters, we used the 70-gene signature test Gene70, also known as MammaPrint, which comprises 70 genes predictive of clinical prognosis in breast

cancer, to quantify malignancy and predict clinical outcome. The expression levels of each gene associated with poor prognosis were quantified and treated as modules. The calculated module scores showed significant differences between cluster 8 (basal), cluster 19 (LumA), and cluster 1 (normal-like), confirming two malignant clusters: cluster 8 and 19 (Fig. 3A, Supplementary Fig. 5A-C). We also examined the gene expression levels of clinically assessed markers, including *ESR1*, *PGR*, *ERBBQ*, *CDH1*, and *MKI67* (Supplementary Fig. 5D). Cluster 8 exhibited high *MKI67* expression, indicative of a highly proliferative phenotype. To further support its malignant nature beyond proliferation, we also evaluated other cell cycle-related genes such as *PCNA* and *MCM6*, which were likewise highly expressed in cluster 8 (Supplementary Fig. 5E). While high *MKI67* levels could suggest a proliferative cluster, the consistent expression of multiple proliferation-associated genes, combined with poor-prognosis markers, indicated that cluster 8 represented a malignant population rather than a general cell cycle-related cluster. In contrast, cluster 19 showed elevated levels of

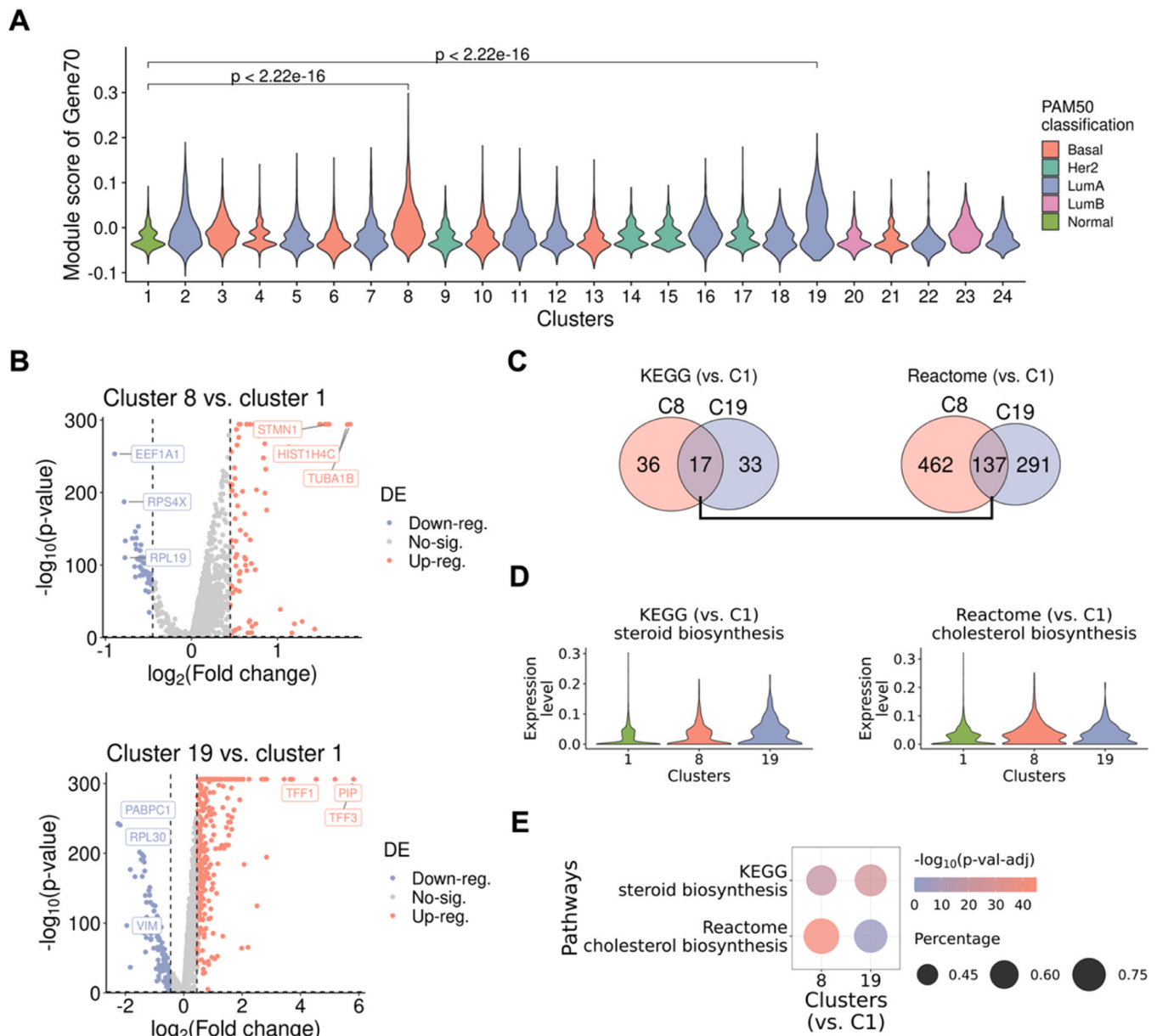


Fig. 3. Cholesterol-related pathways are highly expressed in malignant single-cell clusters. (A) We calculated the module score based on the poor prognosis genes from Gene70 and identified two malignant clusters, clusters 8 and 19. (B) Volcano plots reveal the DEGs of clusters 8 and 19 compared with the only normal-like cluster, cluster 1. (C) Venn diagrams show several activated pathways in both clusters 8 and 19 from KEGG and Reactome databases, where cholesterol-related pathways were found to be highly enriched in these malignant single-cell clusters. (D) Violin plots indicate that the expression levels of KEGG steroid biosynthesis and Reactome cholesterol biosynthesis are higher in malignant cells. (E) We conducted hypothesis testing on two cholesterol-related pathways of clusters 8 and 19 compared to cluster 1, and demonstrated significant activation of these pathways in malignant cells.

luminal-related markers, reinforcing its malignant nature. Given that both malignant clusters (clusters 8 and 19) identified in our analysis predominantly fall within HER2-negative subtypes, our study focused on targeting these populations rather than HER2-positive cancers. We further conducted differential expression analysis to identify the top three DEGs between the two malignant clusters (clusters 8 and 19) and the normal-like cluster (cluster 1) (Fig. 3A and Supplementary Tables 3–6). *STMN1* and *TUBA1B*, both playing key roles in regulating microtubule dynamics and associated with aggressive phenotypes in breast cancer, were highly expressed in cluster 8 [39,40]. On the other hand, *TFF1* and *TFF3*, two members of the trefoil factor family known to promote metastasis in breast cancer, were highly expressed in cluster 19 [41,42]. We then performed pathway enrichment analysis using KEGG and Reactome databases. The number of significantly activated

pathways shared between clusters 8 and 19 was visualized using a Venn diagram (Fig. 3B). The analysis revealed 17 and 137 significantly activated pathways in clusters 8 and 19, respectively (Supplementary Tables 7 and 8). Through manual comparison, we identified the cholesterol biosynthesis-related pathways, including steroid biosynthesis and cholesterol biosynthesis, were the only correlated pathways between KEGG and Reactome databases that were significantly activated in both clusters 8 and 19 (Fig. 3C). Both steroid and cholesterol biosynthesis commence with the conversion of acetyl-CoA and HMG-CoA to cholesterol through a series of enzymatic reactions known as the mevalonate pathway [43]. Violin and bubble plots (Fig. 3D, E) demonstrated the significant differences and differential expression levels between cluster 8 (basal), cluster 19 (LumA), and cluster 1 (normal-like). Additionally, our results indicated that Basal and Luminal A tumor cells exhibit higher

cholesterol biosynthesis levels compared to other subtypes (Supplementary Fig. 6).

3.4. Cross-dataset validation of cholesterol biosynthesis in malignant breast cancer

To validate our findings regarding cholesterol biosynthesis activation in malignant breast cancer, we analyzed an independent single-cell RNA-seq dataset from 26 breast cancer patients. This external dataset underwent similar preprocessing, including quality control, normalization, and dimensionality reduction, followed by clustering and cell type annotation based on known lineage markers (Supplementary Fig. 7). We further re-clustered the cancer cells and assigned molecular subtypes to malignant clusters, using the PAM50 classification model and SCS subtype (Supplementary Fig. 8A–D). We then applied the Gene70 poor prognostic signatures and identified malignant basal-like cluster 9 (Supplementary Fig. 8E). To assess the activation of cholesterol biosynthesis in malignant cancer cells, we conducted pathway enrichment analysis using KEGG and Reactome databases. Consistent with our initial dataset, we observed significant upregulation of cholesterol biosynthesis-related pathways in malignant breast cancer cells (Supplementary Fig. 8F, G).

3.5. Targeting cholesterol biosynthesis in breast cancer

We focused on the shared mevalonate pathway between KEGG steroid biosynthesis and Reactome cholesterol biosynthesis activated in malignant clusters (Fig. 4A). *De novo* cholesterol biosynthesis serves as a critical step for cancer development through a cascade of enzymatic reactions such as 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase (*HMGCR*) and squalene epoxidase (*SQLE*), both acting as rate-limiting enzymes that control the pathway's reaction rate [44]. To target cholesterol biosynthesis, we identified four significantly

differentially expressed genes, including *HMGCR*, *MVD*, *FDFT1* and *SQLE*, in cholesterol biosynthesis between malignant clusters and the normal-like cluster (Fig. 4B–E, Supplementary Fig. 9). We also conducted survival analysis on breast cancer patients from TCGA, demonstrating that patients with high expression levels of these genes have significantly worse clinical outcomes (Fig. 4F–I). By analyzing the combined expression levels of these four key cholesterol biosynthesis genes, we found that elevated expression of these genes is associated with poor prognosis, particularly in HER2-negative breast cancer patients (Supplementary Fig. 10). Collectively, these findings suggest that cholesterol biosynthesis serves as an important therapeutic target in breast cancer treatment.

3.6. Discovery of combinatorial treatment targets

Combinatorial therapies in cancer treatment have shown superior efficacy and fewer adverse effects than single-drug administration, prompting significant interest for researchers and scientists [45]. Our comprehensive single-cell analysis revealed compelling insights into cholesterol signaling in malignant breast cancer, while CellChat analysis demonstrated activated outgoing communication patterns, with cholesterol signaling prominently featured in luminal A breast cancer (Fig. 2E). Further investigation using Gene70 scoring focused on malignant clusters 8 and 19 (basal and luminal A subtypes) also indicated the significant activation of cholesterol biosynthesis pathways. Given the strong activation of cholesterol signaling in HER2-negative breast cancer subtypes, we posited an underlying relationship between cholesterol metabolism and HER2 receptor dynamics, which presents a unique opportunity to investigate how cholesterol pathway modulation might interact with or potentially influence HER2 receptor availability and signaling, thus opening new avenues for combinatorial treatment strategies in these challenging breast cancer subtypes. Previous studies have implicated cholesterol and caveolin-1 (CAV1) in the regulation of

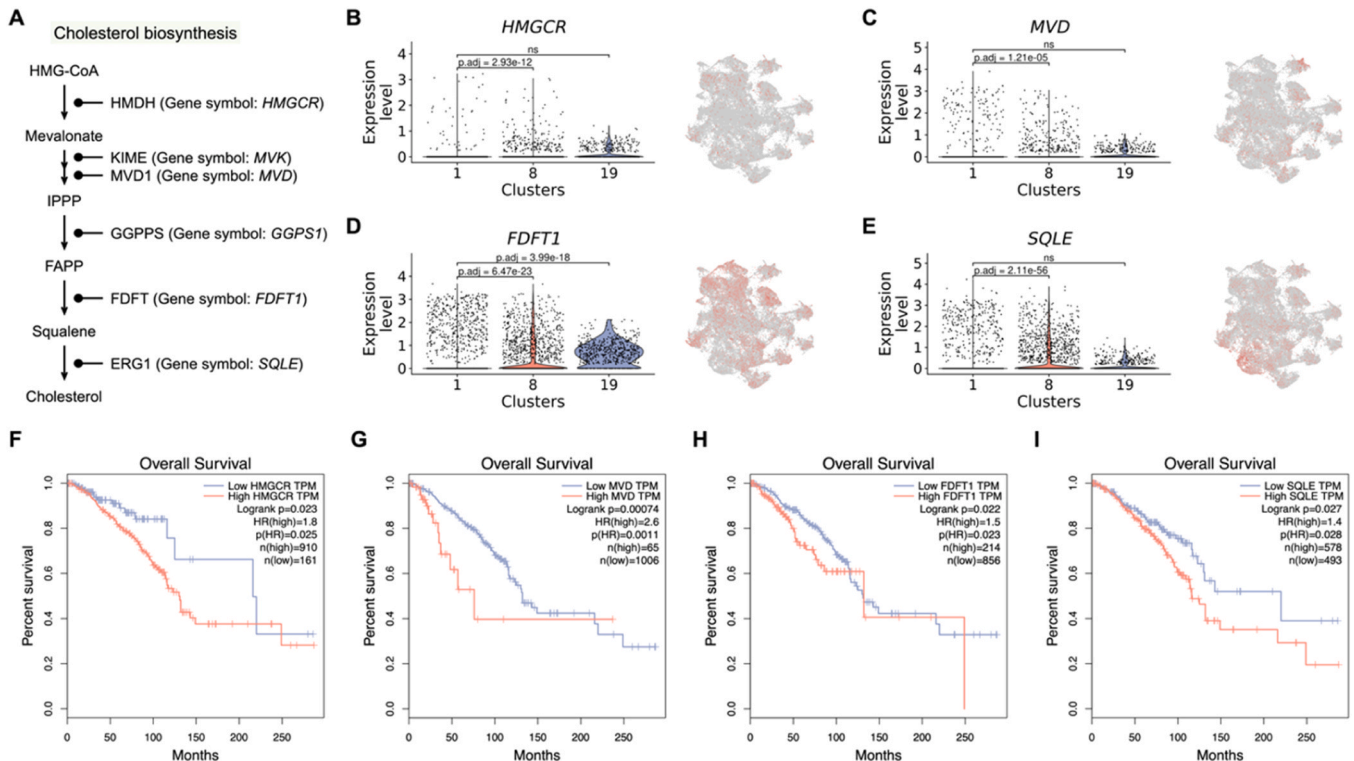


Fig. 4. Characteristics of cholesterol-related pathways. (A) Schematic illustration of the cholesterol biosynthesis pathway. (B–E) Violin plots and gene expression level colored on UMAP reveal that *HMGCR*, *MVD*, *FDFT1* and *SQLE* are highly expressed in malignant cell clusters 8 and 19 compared to the normal-like cluster 1. (F–I) Overall survival analysis of *HMGCR*, *MVD*, *FDFT1* and *SQLE* showed that patients with higher expression of these genes had worse prognosis and lower survival rate.

HER2 cell membrane dynamics, particularly in the context of HER2 endocytosis [46–48]. Meanwhile, statins have been reported to inhibit cholesterol biosynthesis and enhance the availability of HER2 on the cell surface, thereby improving the efficacy of trastuzumab therapy, which specifically targets HER2 in gastric cancer [49]. Our motivation for exploring this connection stemmed from the observation that cholesterol biosynthesis was highly activated in malignant breast cancer cells, and this metabolic alteration may influence HER2 activity and signaling dynamics in ways not strictly limited to HER2-positive cancers.

Motivated by these insights, we proposed a combinatorial treatment strategy that combines the inhibition of cholesterol biosynthesis with targeting cell-surface HER2 in breast cancer (Fig. 5A). Our findings suggested that inhibiting cholesterol biosynthesis may enhance HER2 expression levels, potentially creating an opportunity to eliminate malignant breast cancer cells more effectively. This combinatorial approach offers a promising strategy to improve treatment outcomes for breast cancer patients by targeting both biological and therapeutic mechanisms. To validate the potential of this set of combination therapy targets, we analyzed proteomics data of 59 breast cancer cell lines in the CCLE database. First, PCA was performed and cell lines were visualized based on the top two PCs (Fig. 5B). We focused on basal and luminal breast cancer subtypes to mimic the malignant clusters 8 and 19, respectively. The HER2-positive subtype was not the main focus of this analysis, as our study primarily centered on luminal A and basal-like subtypes due to their differential reliance on cholesterol biosynthesis, which is central to our previous investigation. In luminal breast cancer cell lines, we calculated the Spearman correlation between proteins involved in the cholesterol biosynthesis pathway and the ERBB family. The analysis revealed a negative correlation between cholesterol biosynthesis proteins and HER2 expression in luminal breast cancer cell lines, suggesting that inhibiting cholesterol biosynthesis increased HER2 availability (Fig. 5C). Notably, 3-hydroxy-3-methylglutaryl-CoA reductase (HMDH, gene symbol: *HMGCR*) exhibited the strongest negative correlation with HER2 in luminal breast cancer cell lines (Fig. 5D). A similar negative correlation was observed in basal breast cancer cell lines, with the top two proteins showing the lowest correlation

highlighted in scatter plots but not in HER2-positive breast cancer cell lines (Fig. 5E–G, Supplementary Fig. 11). These findings suggest that our proposed combinatorial treatment strategy, which targets cholesterol biosynthesis and HER2, may be effective across multiple breast cancer subtypes, underscoring its broader applicability in clinical settings.

3.7. Validation of cholesterol biosynthesis-HER2 combinatorial targets

We employed a comprehensive multi-omics approach to validate the proposed combinatorial treatment strategy targeting cholesterol biosynthesis and HER2 in breast cancer. Initially, we analyzed the transcriptomics data of HER2-negative breast cancer cell lines from the CCLE database, performing gene set variation analysis on cholesterol biosynthesis pathways using Reactome and KEGG databases. We visualized the results using a heatmap and ranked the cell lines based on their GSVA scores (Fig. 6A). To further explore the potential of this combinatorial approach, we investigated HER2 cell surface intensity from previous study and found a negative correlation between HER2 membrane levels and cholesterol biosynthesis ranking (Fig. 6B) [37]. We selected Neratinib as an HER2 inhibitor since it has been widely used in combinatorial treatment, e.g., the combination of neratinib with capecitabine in NALA trial, with fulvestrant in MutHER trial, and with trastuzumab [50–52]. Additionally, we leveraged the PRISM drug screening database to investigate the efficacy of Neratinib, a HER2 tyrosine kinase inhibitor, on HER2-negative breast cancer cell lines. By comparing the cholesterol biosynthesis pathway ranking with Neratinib scores, we also discovered a negative correlation in HER2-negative cell lines, suggesting that Neratinib's efficacy is intimately linked to the strength of cholesterol biosynthesis in HER2-negative breast cancer cell lines (Fig. 6C). Integrating transcriptomics, proteomics, HER2 membrane level and drug screening database, we provided a robust validation of cholesterol biosynthesis and HER2 as promising combinatorial targets in breast cancer. Lastly, we conducted a clonogenic assay to evaluate the efficiency of combining inhibitors of cholesterol biosynthesis (Atorvastatin and NB-598) with a well-known dual inhibitor of HER2 and EGFR (Neratinib) on two HER2-negative breast cancer cell

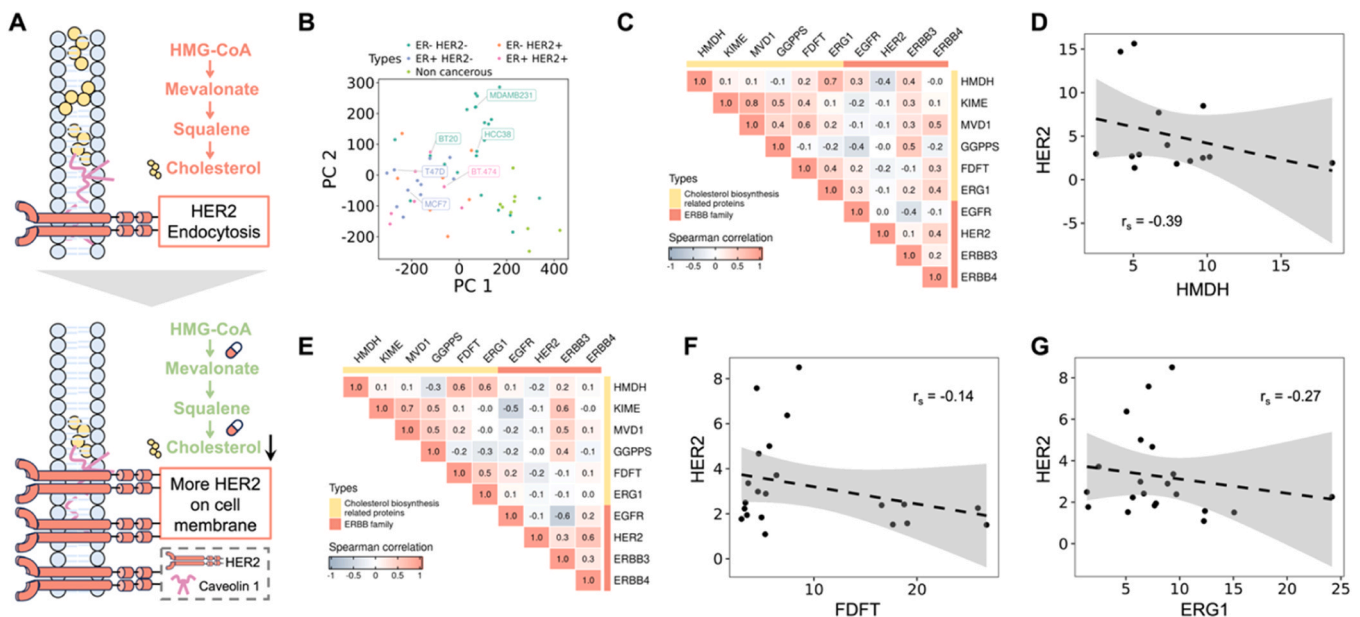


Fig. 5. Cholesterol biosynthesis pathway negatively correlates with HER2 in breast cancer cell lines. (A) Combinatorial treatment strategy of targeting cholesterol-related pathways and HER2. (B) Principal component analysis was performed on the proteomics profiles from breast cancer cell lines ($n = 59$). (C) Heatmap reveals the Spearman correlation between proteins involved in the cholesterol biosynthesis pathway and the ERBB family in luminal breast cancer cell lines. (D) Negative correlations were observed between HMDH and HER2 in luminal breast cancer cell lines. (E) Heatmap reveals the Spearman correlation between proteins involved in the cholesterol biosynthesis pathway and the ERBB family in basal breast cancer cell lines. (F-G) Negative correlations were observed between FDFT, ERG1, and HER2 in basal breast cancer cell lines.

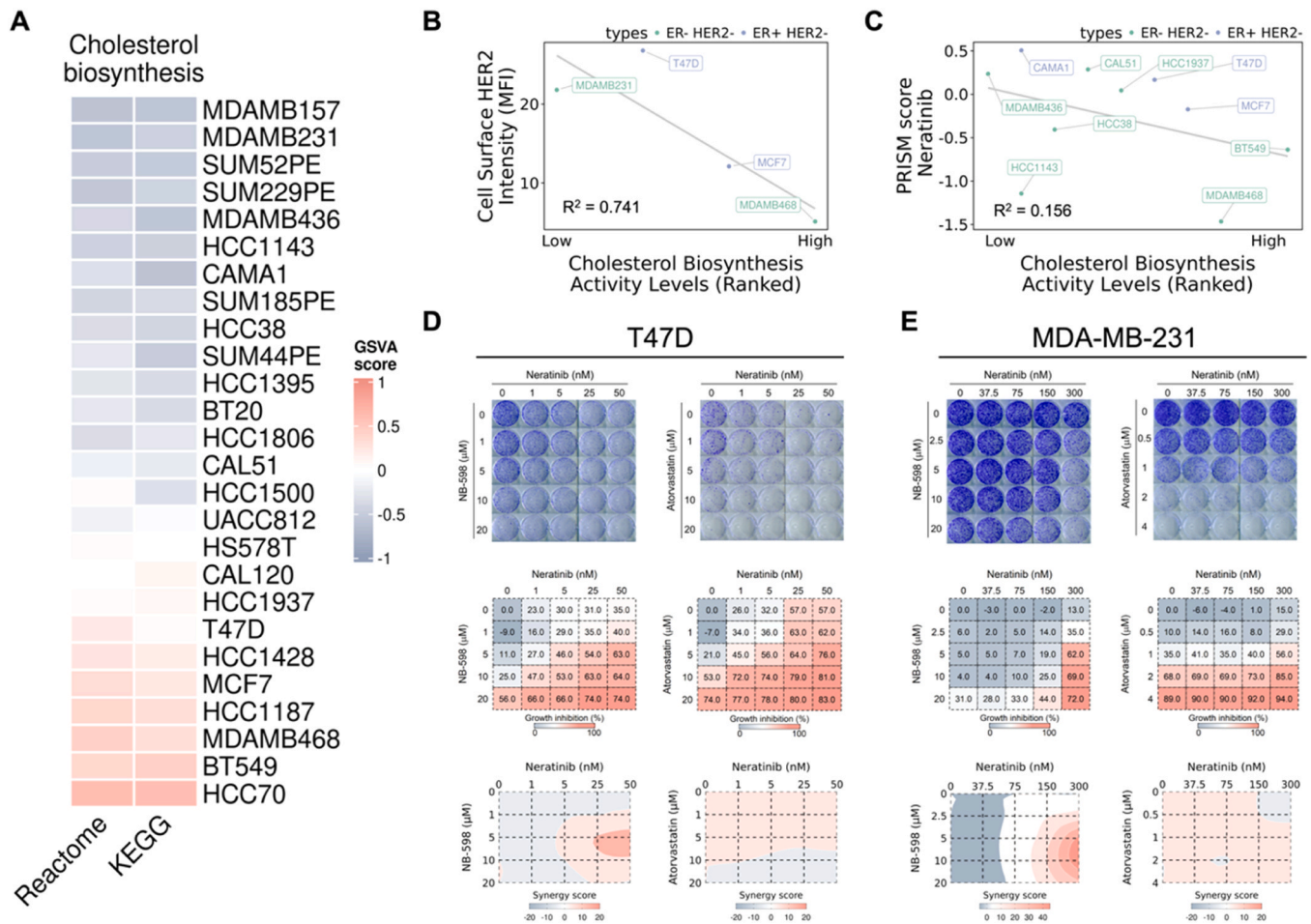


Fig. 6. Computational and experimental validation of cholesterol-HER2 combinatorial targeting. (A) Heatmap shows the ranking of cholesterol biosynthesis strength in HER2-negative breast cancer cell lines. (B) The scatter plot unveils the correlation between the ranked activity levels of cholesterol biosynthesis and HER2 cell surface intensity. (C) The scatter plot reveals the correlation between the ranked activity levels of cholesterol biosynthesis and PRISM score of HER2 inhibitor Neratinib. (D-E) Colony formation assays were conducted on T-47D (D) and MDA-MB-231 (E) cells treated with a combination of Neratinib and SQLE inhibitor (NB-598) or HMGCR inhibitor (Atorvastatin) for 9–12 days. The emergence of synergistic effects prominently inhibited colony formation, achieving more effective cell suppression. The drug synergistic effects were evaluated by the ZIP model (bottom).

lines, T-47D and MDA-MB-231 (Fig. 6D, E). In the clonogenic assay, we noted augmented growth inhibition with combinatorial treatment across both breast cancer cell lines and drug pairs. Additionally, we employed the ZIP model to capture drug-interaction relationships, comparing potency changes in dose-response curves between individual drugs and combinations to determine the drug synergy [38]. The results of synergy score calculation revealed that Atorvastatin (inhibitor of HMGCR) or NB-598 (inhibitor of SQLE) both synergized with Neratinib (an inhibitor of HER2 family) in the two breast cancer cell lines, suggesting that the combinatorial strategy had synergistic anticancer effects to inhibit cancer colony formation.

4. Discussion

HER2 has emerged as a crucial therapeutic target in breast cancer treatment, profoundly transforming clinical management strategies. HER2 + and HER2- breast cancers exhibit distinct clinical characteristics and require different treatment approaches [53]. Historically, HER2 + breast cancer has been considered more aggressive with a poorer prognosis, characterized by rapid tumor growth and higher metastasis risk; however, the advent of targeted therapies like HER2-directed monoclonal antibodies and tyrosine kinases has dramatically improved outcomes [54]. Here, we elucidated the significance of cholesterol biosynthesis in malignant breast cancer cells, cluster

8 (basal) and cluster 19 (LumA), through comprehensive analyses of patient scRNA-seq data. Our findings suggest that simultaneous targeting of cholesterol biosynthesis and HER2 could further optimize treatment strategies, potentially improving prognosis and managing HER2-negative breast cancer.

In cancer, alterations in biosynthesis and metabolism play a crucial role in tumor malignancy, influencing various aspects of cancer progression, including cell proliferation, invasion, and metastasis [55,56]. Through pathway enrichment analysis, numerous pathways were also identified as highly activated in the two malignant clusters (Supplementary Tables 7 and 8). In addition to cholesterol biosynthesis pathways, we discovered that folate biosynthesis, cysteine and methionine metabolism, and glutathione metabolism were also highly activated in these malignant clusters. All these mechanisms have been reported to promote cancer cell progression and survival, aligning with our classification of malignant clusters [56–59]. Moreover, cysteine has been linked to glutathione, a critical component of the glutathione metabolism that plays a central role in cellular redox homeostasis, making it a target for drug development in various cancers, including colon, liver, lung, and breast cancer [60–62].

Cholesterol biosynthesis is gaining increasing attention in cancer research due to its importance in cellular homeostasis and its potential as a targetable pathway. Fluctuations in cellular and blood cholesterol levels have been associated with tumor carcinogenesis, and their

accumulation can enhance tumor aggressiveness and promote angiogenesis [63]. Furthermore, accelerated cholesterol production can affect cell division and membrane synthesis, which are essential for tumor progression [12]. In cholesterol biosynthesis, HMGCR and SQLE serve as rate-limiting enzymes that regulate the pathway's reactions [44]. Prior studies have shown that dysregulation of HMGCR contributes to tumorigenesis, while dysregulation of SQLE has been significantly correlated with progression in several types of cancer, such as prostate cancer and hepatocellular carcinoma [64,65]. Additionally, a previous retrospective cohort study showed that post-diagnosis statin use is associated with improved cancer-specific survival, particularly in hormone receptor-positive/HER2-negative breast cancer patients, suggesting a potential therapeutic benefit that warrants further investigation [66]. Therefore, targeting HMGCR and SQLE presents promising strategies in cancer treatment, offering potent combination drug candidates to inhibit cholesterol accumulation.

In this study, we identified cholesterol biosynthesis pathways that were highly activated in malignant clusters compared with a normal-like cluster via single-cell analysis. We also comprehensively validated the potential of targeting cholesterol biosynthesis and HER2 as combinatorial targets in breast cancer through an integrative multi-omics strategy combining transcriptomics, proteomics, membrane protein intensity, and pharmacological screening. Statin-induced HER2 upregulation offers a potential strategy to enhance HER2-targeted antibody-drug conjugates (ADCs), such as T-Dxd, by improving drug delivery in tumors with low or heterogeneous HER2 expression [67]. This approach aligns with emerging strategies to optimize ADC performance through biomarker modulation, supporting further investigation into combinatorial treatments with cholesterol pathway inhibitors. In this study, we demonstrated that Atorvastatin or NB-598 (inhibitors of cholesterol biosynthesis) combined with Neratinib had synergistic effects on two breast cancer cell lines, representing the two malignant clusters from scRNA-seq data. Additionally, we utilized the ZIP model to rigorously assess drug-interaction relationships, enabling continuous prediction of synergistic effects by comparing potency changes in dose-response curves between individual drugs and their combinations. Beyond quantifying the synergistic interactions between cholesterol biosynthesis inhibitors and Neratinib, this approach offers a robust computational framework for predicting potential combinatorial therapeutic strategies in cancer treatment.

5. Conclusions

In this study, we leveraged scRNA-seq data to identify malignant breast cancer subsets and proposed a combinatorial therapy strategy targeting both cholesterol biosynthesis and HER2. Through rigorous multi-omics analysis on breast cancer cell lines, including transcriptomics, proteomics, membrane protein intensity, and drug screening data, we uncovered a significant relationship between cholesterol biosynthesis and HER2 expression in HER2-negative cell lines. Our computational approach revealed a novel therapeutic strategy that extends beyond conventional targeting methods. Also, experimental validation through clonogenic assays revealed that simultaneous inhibition of cholesterol biosynthesis and HER2 can effectively eliminate malignant breast cancer cells, even in HER2-negative contexts. Overall, the identification of combinatorial therapy from single-cell analysis, targeting both cholesterol biosynthesis and HER2, demonstrated promising synergistic effects in breast cancer treatment.

CRedit authorship contribution statement

Tzu-Yang Tseng: Software, Formal analysis, Investigation, Visualization, Writing-Original Draft. **Chiao-Hui Hsieh:** Validation, Visualization, Writing-Original Draft. **Jie-Yu Liu:** Validation, Visualization, Writing-Original Draft. **Hsuan-Cheng Huang:** Conceptualization, Methodology, Writing-Review and Editing, Supervision, Project

administration, Funding acquisition. **Hsueh-Fen Juan:** Conceptualization, Methodology, Writing-Review and Editing, Supervision, Project administration, Funding acquisition.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Author statement

All authors have read and approved the submission of this manuscript to *Computational and Structural Biotechnology Journal*.

Funding

This work was supported by the National Science and Technology Council (109-2221-E-002-161-MY3, 112-2320-B-002-022, 109-2221-E-010-011-MY3, 112-2221-E-A49-061-MY3, 113-2320-B-002-025-MY3, 113-2221-E-002-149-MY3), the Higher Education Sprout Project (NTU-CC-113L890302, NTU-113L8503, NTU-CC-113L894502, NTU-CC-114L892702) and Center for Advanced Computing and Imaging in Biomedicine (NTU-113L900701, NTU-114L900701) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Jason Lee at Jason's Language inspirations for proof-reading the manuscript.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2025.04.030.

Data availability

The single-cell RNA-seq dataset analyzed during the current study is available from the European Genome-phenome Archive (EGA) under accession number EGAS00001004809, and the bulk sequencing datasets of breast cancer cell lines, including transcriptomics and proteomics, can be downloaded from DepMap (<https://depmap.org/portal/>). Additionally, all the codes used in this study are available on our GitHub page (https://github.com/yangtseng/BRCA_chole_HER2).

References

- [1] Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, et al. Breast cancer. *Nat Rev Dis Prim* 2019;5(1):66. <https://doi.org/10.1038/s41572-019-0111-2>.
- [2] Yersal O, Barutca S. Biological subtypes of breast cancer: prognostic and therapeutic implications. *World J Clin Oncol* 2014;5(3):412–24. <https://doi.org/10.5306/wjco.v5.i3.412>.
- [3] Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001;98(19):10869–74. <https://doi.org/10.1073/pnas.191367098>.

- [4] Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, et al. Metastatic behavior of breast cancer subtypes. *J Clin Oncol* 2010;28(20):3271–7. <https://doi.org/10.1200/jco.2009.25.9820>.
- [5] Waks AG, Winer EP. Breast cancer treatment: a review. *JAMA* 2019;321(3):288–300. <https://doi.org/10.1001/jama.2018.19323>.
- [6] Burguin A, Diorio C, Durocher F. Breast cancer treatments: updates and new challenges. *J Pers Med* 2021;11(8). <https://doi.org/10.3390/jpm11080808>.
- [7] Kinker GS, Greenwald AC, Tal R, Orlova Z, Cuoco MS, et al. Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat Genet* 2020;52(11):1208–18. <https://doi.org/10.1038/s41588-020-00726-6>.
- [8] Chung W, Eum HH, Lee HO, Lee KM, Lee HB, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* 2017;8:15081. <https://doi.org/10.1038/ncomms15081>.
- [9] Van de Sande B, Lee JS, Mutasa-Gottgens E, Naughton B, Bacon W, et al. Applications of single-cell RNA sequencing in drug discovery and development. *Nat Rev Drug Discov* 2023;22(6):496–520. <https://doi.org/10.1038/s41573-023-00688-4>.
- [10] Fisusi FA, Akala EO. Drug combinations in breast cancer therapy. *Pharm Nanotechnol* 2019;7(1):3–23. <https://doi.org/10.2174/2211738507666190122111224>.
- [11] Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med* 2016;375(8):717–29. <https://doi.org/10.1056/NEJMoa1602253>.
- [12] Llaverias G, Danilo C, Mercier I, Daumer K, Capozza F, et al. Role of cholesterol in the development and progression of breast cancer. *Am J Pathol* 2011;178(1):402–12. <https://doi.org/10.1016/j.ajpath.2010.11.005>.
- [13] Ehmsen S, Pedersen MH, Wang G, Terp MG, Arslanagic A, et al. Increased cholesterol biosynthesis is a key characteristic of breast cancer stem cells influencing patient outcome. *Cell Rep* 2019;27(13):3927–3938.e6. <https://doi.org/10.1016/j.celrep.2019.05.104>.
- [14] Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* 2019;569(7757):503–8. <https://doi.org/10.1038/s41586-019-1186-3>.
- [15] Iqbal N, Iqbal N. Human epidermal growth factor receptor 2 (HER2) in cancers: overexpression and therapeutic implications. *Mol Biol Int* 2014;2014:852748. <https://doi.org/10.1155/2014/852748>.
- [16] Bassez A, Vos H, Van Dyck L, Floris G, Arijis I, et al. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat Med* 2021;27(5):820–32. <https://doi.org/10.1038/s41591-021-01323-8>.
- [17] Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet* 2021;53(9):1334–47. <https://doi.org/10.1038/s41588-021-00911-1>.
- [18] Corsello SM, Nagari RT, Spangler RD, Rossen J, Kocak M, et al. Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. *Nat Cancer* 2020;1(2):235–48. <https://doi.org/10.1038/s43018-019-0018-6>.
- [19] McCarthy DJ, Campbell KR, Lun AT, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 2017;33(8):1179–86. <https://doi.org/10.1093/bioinformatics/btw777>.
- [20] Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, et al. Integrated analysis of multimodal single-cell data. *Cell* 2021;184(13):3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.
- [21] Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16(12):1289–96. <https://doi.org/10.1038/s41592-019-0619-0>.
- [22] Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* 2018. <https://doi.org/10.1038/nbt.4314>.
- [23] Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 2019;9(1):5233. <https://doi.org/10.1038/s41598-019-41695-z>.
- [24] Tickle T, Tirosch I, Georgescu C, Brown M, Haas B. inferCNV of the Trinity CTAT Project. 2019 Available from: (<https://github.com/broadinstitute/infercnv>).
- [25] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;27(8):1160–7. <https://doi.org/10.1200/jco.2008.18.1370>.
- [26] Gendoo DM, Ratanasirigulchai N, Schröder MS, Paré L, Parker JS, et al. Genefu: an R/bioconductor package for computation of gene expression-based signatures in breast cancer. *Bioinformatics* 2016;32(7):1097–9. <https://doi.org/10.1093/bioinformatics/btv693>.
- [27] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32(18):2847–9. <https://doi.org/10.1093/bioinformatics/btw313>.
- [28] Jin S, Plikus MV, Nie Q. CellChat for systematic analysis of cell–cell communication from single-cell transcriptomics. *Nat Protoc* 2024. <https://doi.org/10.1038/s41596-024-01045-4>.
- [29] van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415(6871):530–6. <https://doi.org/10.1038/415530a>.
- [30] Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017;14(11):1083–6. <https://doi.org/10.1038/nmeth.4463>.
- [31] Borcherdig N, Vishwakarma A, Voigt AP, Bellizzi A, Kaplan J, et al. Mapping the immune environment in clear cell renal carcinoma by single-cell genomics. *Commun Biol* 2021;4(1):122. <https://doi.org/10.1038/s42003-020-01625-6>.
- [32] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28(1):27–30. <https://doi.org/10.1093/nar/28.1.27>.
- [33] Jassal B, Matthews L, Viteri G, Gong C, Lorente P, et al. The reactome pathway knowledgebase. *Nucleic Acids Res* 2020;48(D1):D498–d503. <https://doi.org/10.1093/nar/gkz1031>.
- [34] Tang Z, Li C, Kang B, Gao G, Li C, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45(W1):W98–w102. <https://doi.org/10.1093/nar/gkx247>.
- [35] Györfy B. Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput Struct Biotechnol J* 2021;19:4101–9. <https://doi.org/10.1016/j.csbj.2021.07.014>.
- [36] Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform* 2013;14(1):7. <https://doi.org/10.1186/1471-2105-14-7>.
- [37] Stanley A, Ashrafi GH, Seddon AM, Modjtahedi H. Synergistic effects of various Her inhibitors in combination with IGF-1R, C-MET and Src targeting agents in breast cancer cell lines. *Sci Rep* 2017;7(1):3964. <https://doi.org/10.1038/s41598-017-04301-8>.
- [38] Yadav B, Wennerberg K, Aittokallio T, Tang J. Searching for drug synergy in complex dose-response landscapes using an interaction potency model. *Comput Struct Biotechnol J* 2015;13:504–13. <https://doi.org/10.1016/j.csbj.2015.09.001>.
- [39] Wang Y, Li Y, Jing Y, Yang Y, Wang H, et al. Tubulin alpha-1b chain was identified as a prognosis and immune biomarker in pan-cancer combining with experimental validation in breast cancer. *Sci Rep* 2024;14(1):8201. <https://doi.org/10.1038/s41598-024-58982-z>.
- [40] Obayashi S, Horiguchi J, Higuchi T, Katayama A, Handa T, et al. Stathmin1 expression is associated with aggressive phenotypes and cancer stem cell marker expression in breast cancer patients. *Int J Oncol* 2017;51(3):781–90. <https://doi.org/10.3892/ijo.2017.4085>.
- [41] Pandey V, Wu ZS, Zhang M, Li R, Zhang J, et al. Trefoil factor 3 promotes metastatic seeding and predicts poor survival outcome of patients with mammary carcinoma. *Breast Cancer Res* 2014;16(5):429. <https://doi.org/10.1186/s13058-014-0429-3>.
- [42] Ahmed ARH, Griffiths AB, Tilby MT, Westley BR, May FEB. TFF3 is a normal breast epithelial protein and is associated with differentiated phenotype in early breast cancer but predisposes to invasion and metastasis in advanced disease. *Am J Pathol* 2012;180(3):904–16. <https://doi.org/10.1016/j.ajpath.2011.11.022>.
- [43] Rone MB, Fan J, Papadopoulos V. Cholesterol transport in steroid biosynthesis: role of protein-protein interactions and implications in disease states. *Biochim Biophys Acta* 2009;1791(7):646–58. <https://doi.org/10.1016/j.bbalip.2009.03.001>.
- [44] Zou Y, Zhang H, Bi F, Tang Q, Xu H. Targeting the key cholesterol biosynthesis enzyme squalene monooxygenase for cancer therapy. *Front Oncol* 2022;12:938502. <https://doi.org/10.3389/fonc.2022.938502>.
- [45] Liu H, Zhang W, Zou B, Wang J, Deng Y, et al. DrugCombDB: a comprehensive database of drug combinations toward the discovery of combinatorial therapy. *Nucleic Acids Res* 2020;48(D1):D871–d881. <https://doi.org/10.1093/nar/gkz1007>.
- [46] Yang C, He B, Dai W, Zhang H, Zheng Y, et al. The role of caveolin-1 in the bioefficacy of anti-tumor drugs and their nano-drug delivery systems. *Acta Pharm Sin B* 2021;11(4):961–77. <https://doi.org/10.1016/j.apsb.2020.11.020>.
- [47] Pereira PMR, Sharma SK, Carter LM, Edwards JR, Pourat J, et al. Caveolin-1 mediates cellular distribution of HER2 and affects trastuzumab binding and therapeutic efficacy. *Nat Commun* 2018;9(1):5137. <https://doi.org/10.1038/s41467-018-07608-w>.
- [48] Sorkin A, Goh LK. Endocytosis and intracellular trafficking of ErbBs. *Exp Cell Res* 2008;314(17):3093–106. <https://doi.org/10.1016/j.yexcr.2008.08.013>.
- [49] Rao Y, Samuels Z, Carter LM, Monette S, Panikar SS, et al. Statins enhance the efficacy of HER2-targeting radioligand therapy in drug-resistant gastric cancers. *Proc Natl Acad Sci USA* 2023;120(14):e2220413120. <https://doi.org/10.1073/pnas.2220413120>.
- [50] Saura C, Oliveira M, Feng YH, Dai MS, Chen SW, et al. Neratinib plus capecitabine versus lapatinib plus capecitabine in HER2-positive metastatic breast cancer previously treated with ≥ 2 HER2-directed regimens: phase III NALA trial. *J Clin Oncol* 2020;38(27):3138–49. <https://doi.org/10.1200/jco.20.00147>.
- [51] Ma CX, Luo J, Freedman RA, Pluard TJ, Nangia JR, et al. The phase II MutHER study of neratinib alone and in combination with fulvestrant in HER2-mutated, non-amplified metastatic breast cancer. *Clin Cancer Res* 2022;28(7):1258–67. <https://doi.org/10.1158/1078-0432.Ccr-21-3418>.
- [52] Blackwell KL, Zaman K, Qin S, Tkaczuk KHR, Campone M, et al. Neratinib in combination with trastuzumab for the treatment of patients with advanced HER2-positive breast cancer: a phase I/II study. *Clin Breast Cancer* 2019;19(2):97–104.e4. <https://doi.org/10.1016/j.clbc.2018.12.011>.
- [53] Swain SM, Shastry M, Hamilton E. Targeting HER2-positive breast cancer: advances and future directions. *Nat Rev Drug Discov* 2023;22(2):101–26. <https://doi.org/10.1038/s41573-022-00579-0>.
- [54] Zhu K, Yang X, Tai H, Zhong X, Luo T, et al. HER2-targeted therapies in cancer: a systematic review. *Biomark Res* 2024;12(1):16. <https://doi.org/10.1186/s40364-024-00565-1>.
- [55] Feng Y, Spezia M, Huang S, Yuan C, Zeng Z, et al. Breast cancer development and progression: risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes Dis* 2018;5(2):77–106. <https://doi.org/10.1016/j.gendis.2018.05.001>.
- [56] Bombonati A, Sgroi DC. The molecular pathology of breast cancer progression. *J Pathol* 2011;223(2):307–17. <https://doi.org/10.1002/path.2808>.

- [57] Larsson SC, Giovannucci E, Wolk A. Folate and risk of breast cancer: a meta-analysis. *J Natl Cancer Inst* 2007;99(1):64–76. <https://doi.org/10.1093/jnci/djk006>.
- [58] Bansal A, Simon MC. Glutathione metabolism in cancer progression and treatment resistance. *J Cell Biol* 2018;217(7):2291–8. <https://doi.org/10.1083/jcb.201804161>.
- [59] Geck RC, Toker A. Nonessential amino acid metabolism in breast cancer. *Adv Biol Regul* 2016;62:11–7. <https://doi.org/10.1016/j.jbior.2016.01.001>.
- [60] Wu G, Fang YZ, Yang S, Lupton JR, Turner ND. Glutathione metabolism and its implications for health. *J Nutr* 2004;134(3):489–92. <https://doi.org/10.1093/jn/134.3.489>.
- [61] Bonifácio VDB, Pereira SA, Serpa J, Vicente JB. Cysteine metabolic circuitries: druggable targets in cancer. *Br J Cancer* 2021;124(5):862–79. <https://doi.org/10.1038/s41416-020-01156-1>.
- [62] Yu H, Yang C, Jian L, Guo S, Chen R, et al. Sulfasalazine-induced ferroptosis in breast cancer cells is reduced by the inhibitory effect of estrogen receptor on the transferrin receptor. *Oncol Rep* 2019;42(2):826–38. <https://doi.org/10.3892/or.2019.7189>.
- [63] Ding X, Zhang W, Li S, Yang H. The role of cholesterol metabolism in cancer. *Am J Cancer Res* 2019;9(2):219–27.
- [64] Liu D, Wong CC, Fu L, Chen H, Zhao L, et al. Squalene epoxidase drives NAFLD-induced hepatocellular carcinoma and is a pharmaceutical target. *Sci Transl Med* 2018;10(437). <https://doi.org/10.1126/scitranslmed.aap9840>.
- [65] Clendenen JW, Pandya A, Boutros PC, Ghamrasni SEI, Khosravi F, et al. Dysregulation of the mevalonate pathway promotes transformation. *Proc Natl Acad Sci USA* 2010;107(34):15051–6. <https://doi.org/10.1073/pnas.0910258107>.
- [66] Guo H, Malone KE, Heckbert SR, Li CI. Statin use and risks of breast cancer recurrence and mortality. *Cancer* 2024;130(18):3106–14. <https://doi.org/10.1002/cncr.35362>.
- [67] Modi S, Jacot W, Yamashita T, Sohn J, Vidal M, et al. Trastuzumab deruxtecan in previously treated HER2-low advanced breast cancer. *N Engl J Med* 2022;387(1):9–20. <https://doi.org/10.1056/NEJMoa2203690>.