



OPEN ACCESS

EDITED BY

Zhijie Zhao,
Shanghai Jiao Tong University, China

REVIEWED BY

Rong Guo,
Shanghai Cancer Center, Fudan University,
China
Yukui Gao,
First Affiliated Hospital of Wannan Medical
College, China

*CORRESPONDENCE

Yue Yu

✉ yuyue@tmu.edu.cn

Zhaohui Chen

✉ chen_zhaohui2021@tmu.edu.cn

Yao Tian

✉ tianyao@tmu.edu.cn

[†]These authors have contributed equally to this work

RECEIVED 30 August 2025

ACCEPTED 26 September 2025

PUBLISHED 10 October 2025

CITATION

Li Y, Cao J, Hai L, Zeng J, Tian D, Yu Y, Chen Z and Tian Y (2025) Multi-scale data reveal a CD24(+) MUCL1(+) tumor subgroup associated with unfavorable prognosis in ER+ breast cancer. *Front. Immunol.* 16:1695689. doi: 10.3389/fimmu.2025.1695689

COPYRIGHT

© 2025 Li, Cao, Hai, Zeng, Tian, Yu, Chen and Tian. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Multi-scale data reveal a CD24 (+) MUCL1(+) tumor subgroup associated with unfavorable prognosis in ER+ breast cancer

Yingxi Li^{1†}, Junming Cao^{2†}, Linyue Hai², Jing Zeng³, Dongchen Tian², Yue Yu^{2*}, Zhaohui Chen^{2*} and Yao Tian^{2,3*}

¹Health Science Center, Ningbo University, Ningbo, Zhejiang, China, ²The First Department of Breast Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, National Clinical Research Center for Cancer, Key Laboratory of Breast Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University, Tianjin, China, ³Department of Thoracic Surgery, The Affiliated LiHuiLi Hospital of Ningbo University, Ningbo, Zhejiang, China

Objective: Breast cancer remains the leading cause of cancer-associated death for women globally. For the group of ER+ breast cancer patients, there are still some problems of poor prognosis that need to be solved. This study aims to identify the poor prognostic tumor subgroups for the prognostic stratification of ER+ patients.

Methods: Through a comprehensive multi-omics strategy, we systematically characterized the biological and clinical significance of MUCL1+CD24+ cells in breast cancer, and we used multiplex immunohistochemistry to confirm the poor role of MUCL1+CD24+ cells.

Results: Single-cell transcriptomics unraveled the cellular ontogeny and immune microenvironment interactions of this subset, while bulk RNA sequencing exposed significant pathway heterogeneity and differential immunotherapy responses associated with varying cellular abundance levels. Genomic landscape analysis pinpointed specific somatic mutations correlated with MUCL1(+) CD24(+) cell infiltration patterns, findings that were subsequently validated through multiplex immunohistochemistry to demonstrate strong prognostic value. Crucially, we developed a clinically translatable radiomics approach that successfully correlated specific MRI features with cellular prevalence, establishing a foundation for noninvasive detection of this aggressive cellular subpopulation.

Conclusions: This integrative approach, spanning molecular to imaging analyses, provides novel insights into both the biological drivers and clinical implications of MUCL1(+) CD24(+) cells in breast cancer progression.

KEYWORDS

ER+ breast cancer, MUCL1(+) CD24(+) cells, tumor microenvironment, radiomics, single-cell RNA sequencing

Introduction

Breast cancer remains the predominant cause of cancer-associated mortality in the female population globally (1). The remarkable molecular heterogeneity of this malignancy results in significantly divergent clinical prognoses among affected individuals (2). Estrogen receptor-positive (ER+) breast cancer represents nearly 70% of all breast malignancies (3). Although the continuous research and development of endocrine therapy and CDK4/6 inhibitors have enabled ER+ breast cancer patients to have a better prognosis, the development of endocrine resistance remains a critical clinical challenge, especially in advanced disease stages (4, 5). Current data indicate that approximately 40%–50% of metastatic ER+ cases acquire treatment resistance within 24 months, frequently resulting in disease progression and diminished survival outcomes (6). Therefore, it is of vital importance to explore the factors of poor prognosis in the group of ER+ breast cancer patients, which is of great significance for the formulation of treatment strategies and the screening of people with poor prognosis.

The emergence of single-cell transcriptomic profiling has revolutionized our understanding of tumor heterogeneity, facilitating the discovery of rare cellular subpopulations that contribute to treatment refractoriness (7). This technological breakthrough enables the correlation of genomic alterations with cell-specific transcriptional programs, offering novel opportunities to decipher resistance mechanisms and discover actionable vulnerabilities (8). Besides that, advanced imaging modalities such as MRI, CT, and PET now play a pivotal role in predicting therapeutic response, molecular subtyping, and prognostic stratification in oncology (9). The integration of radiomic features with multi-omics data offers a transformative approach for discovering clinically relevant biomarkers and personalized treatment strategies.

In this study, we identified a tumor subgroup characterized by high CD24 and MUCL1 expression, which was linked to poor prognosis and invasive behavior. We also uncovered somatic mutations associated with the infiltration of CD24(+) MUCL1(+) cells, along with potential inhibitors for personalized treatment in ER+ breast cancer. Furthermore, a radiomic model effectively estimated the infiltration levels of these cells. Overall, our findings offer a new therapeutic target and a non-invasive strategy for immunotherapy and individualized treatment in ER+ breast cancer patients.

Materials and methods

Data collection and quality control

A total of 14 single-cell RNA sequencing data including seven ER+ breast cancer and seven paired lymph node metastatic tissues were downloaded from GSE161529 (10). The Seurat (v4.3.3) (11) package is used for single-cell sequencing quality control processes including standardization, clustering, and dimensionality reduction. Rigorous quality control was implemented, including (1) the removal of low-quality cells based on mitochondrial gene

content (<20%), unique molecular counts (>500 transcripts/cell), and detected genes (>200 genes/cell). Doubletfinder (v2.0.4) (12) and harmony packages were utilized to remove the mixed cells and batch effect. Bulk RNA expression profiles were also obtained from TCGA database. Patients with complete clinical information and expression profiles were included in the subsequent analysis. Finally, 474 breast cancer samples with ER status (+), PR status (+/-), and HER2 status (-) were included in the study. To reduce the effect of gene length and depth of sequencing, the format of the matrix was transformed into TPM.

Downstream analyses of scRNA-seq

A total of 40 cell clusters were identified after strict quality control procedures. The cell types were further annotated based on public research (13). Eight main cell types were annotated, and epithelial cells were isolated for malignant cell identification. The CNV correlation and score were calculated by infercnv (v1.14.2) (14) package. The details could be found in a referenced study (15). Subsequently, the same procedure as previously described was performed on the cluster resolution of malignant cells. The epitools (v0.5-10.1) package was used to investigate the tissue preference of malignant cell subgroups. The top marker genes of each tumor subgroup were analyzed by the “findallmarker” function. The trajectory inference and cell developmental direction of malignant cells were analyzed by using SCP (v0.4.7.9000) and vector packages. Additionally, the dynamic lineages of tumor cells were reconstructed, and gene clusters were further annotated based on biological process.

Estimation of cell abundance

Since single-cell data cannot directly reflect cell abundance, bulk RNA data was used to infer cell abundance. The single-cell RNA matrix was used as referenced matrix, while the TCGA-BRCA matrix was used as observed matrix. CIBERSORTX (<https://cibersortx.stanford.edu/>) was utilized to infer cell abundance. The high and low abundance of the C4 subgroup was divided based on the median value of the absolute score of C4. The statistics of survival analysis was investigated by log-rank test.

Gene set enrichment analysis

The differential genes between high and low abundance of the C4 subgroup was evaluated by using the limma (v3.54.2) (16) package. All significant genes were reordered and analyzed based on hallmark gene sets. *P*-value <0.05 was considered statistically significant.

Genomic mutation analysis

The TCGA-BRCA somatic mutation data was downloaded by using the tcgabioblinks (v2.26.0) (17) package. Maftools (v2.14.0)

(18) package was used to integrate genomic profiles. The differential mutations between the high- and low-C4 groups are identified by the mafCompare function. Fisher's exact test was used to detect statistical significance. Co-occurrence and co-exclusion patterns between genes were identified by the somatic interactions function.

TIDE and CMAP analysis

Tumor immune dysfunction and exclusion (TIDE) analysis was conducted by using an online database (<http://tide.dfci.harvard.edu/>). Concretely, the expression matrix was scaled and uploaded to the TIDE database. The TIDE score of each sample was obtained, and chi-square test was used to detect the statistical differences in immune responses between the high- and low-C4 subgroups. The potential drugs for high-C4 patients were identified by using the CMAP database (19). The details could be found in our previous study.

Non-invasive radiomics construction

A total of 11 ER+ breast cancers with both molecular subtype information and radiomic imaging profiles were included in the study. Pyradiomics (v3.0.1) was used to extract the radiomics features. Two experienced radiologists with over 10 years of experience in breast oncology together performed fully the manual segmentation of the tumors. All features underwent Z-score normalization, and Pearson correlation was used to filter candidate features correlated with C4 cluster infiltration ($p < 0.05$). LASSO regression algorithm was performed to establish a linear regression model to estimate the C4 cluster abundance. Pearson correlation and ROC curve were used to calculate the association between the abundance of C4 cluster and radiomic score and assess the discriminative efficiency of the model.

Breast cancer specimens

A total of 30 ER+ breast cancer specimens from patients who underwent surgery were collected at Tianjin Medical University Cancer Institute and Hospital. Informed written consent was obtained from the participants. The study was approved by the Ethical Committee of Tianjin Medical University Cancer Institute and Hospital and adhered to the ethical guidelines of the Helsinki Declaration.

Multiplexed fluorescent IHC staining and H&E staining

The consecutive ER+ breast cancer tissues were used to evaluate the expression level of EpCAM, MUCL1, and CD24. In brief, 5- μ m slides were deparaffinized and rehydrated through a graded series of ethanol solutions prior to antigen retrieval in heated citric acid

buffer (pH 6.0). Each slide was put through three sequential rounds of staining, each including a protein block with blocking buffer followed by primary antibody and corresponding secondary HRP-conjugated antibody. Each HRP-conjugated antibody mediated the covalent binding of a different fluorophore for signal amplification. This reaction was followed by additional antigen retrieval in heated citric acid buffer (pH 6.0) in microwave for 15 min to remove the bound antibodies before the next step. After three sequential reactions, the slides were counterstained with DAPI for 10 min and mounted with fluorescence mounting medium. Anti-EPCAM (GB12274, Servicebio), anti-MUCL1 (BS-17247R, Bioss), and anti-CD24 (BS-23867R, Bioss) were used. Images were acquired with a Nikon Eclipse C1 microscope.

For histological examination, a paraffin-embedded ear tissue was cut into 5- μ m sections with H&E staining (G1120, Solarbio) and other staining protocols accordingly (G3632 and G3670, Solarbio).

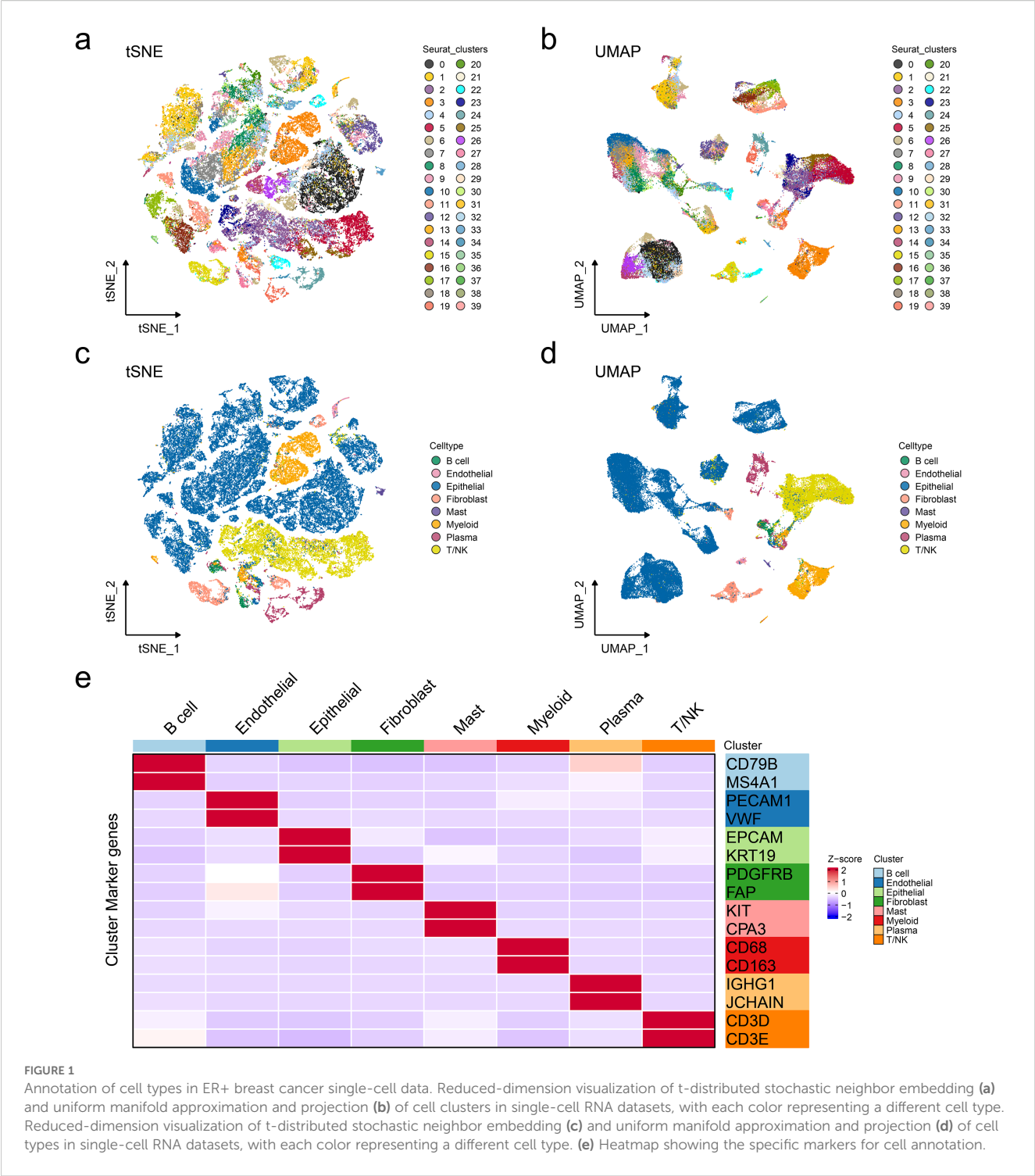
Statistical analysis

All bioinformatics analyses in this study were based on R studio (v4.2.2) and python 3.7. All statistical methods can be found in the corresponding methods sections.

Results

Annotation of cell types in ER+ breast cancer single-cell data

To investigate different cell types in ER+ breast cancer, single-cell transcriptomic analysis was performed on 14 ER+ breast cancer samples from GSE161529, including seven primary ER+ breast cancer tissues and seven paired lymph node metastatic tissues. Rigorous quality control and doublet removal were implemented as previously described. Following rigorous quality control and normalization, a total of 63,369 high-confidence cells were obtained, and 40 distinct cell clusters were identified (Figures 1A, B). Dimensionality reduction using uniform manifold approximation and projection and t-distributed stochastic neighbor embedding demonstrated the clear segregation of the annotated cell populations (Figures 1C, D). Then, cell type annotation was performed by integrating marker gene expression with reference datasets from online databases. We identified eight major cell populations: B cells ($N = 817$), endothelial cells ($N = 434$), epithelial cells ($N = 41,860$), fibroblasts ($N = 2444$), mast cells ($N = 340$), myeloid cells ($N = 4433$), plasma cells ($N = 2,520$), and T/NK cells ($N = 10,521$) (Figure 1E). Standard marker genes were employed for population annotation, including EPCAM (epithelial), PECAM1 (endothelial), and other well-characterized identifiers. In summary, eight main cell types in ER+ breast cancers were identified for the subsequent research. Subsequently, we performed DNA copy number variation analysis to recognize highly confident malignant cells. As shown in Figures 2A–G,



malignant cells presented obvious high CNV correlation coefficient and CNV score compared to normal epithelial cells and myeloid cells, indicating that the set threshold successfully distinguishes tumor cells from normal cells. The tumor cells were standardized and clustered, and 10 tumor subgroups were identified. Tissue preference analysis showed clusters 5–9 (C5–C9) to be more likely enriched in primary ER+ breast cancer tissues, while clusters 0–4 (C0–C4) were more inclined to be enriched in lymph node metastatic tissues (Figure 2H). Taken together, tumor cells

were successfully isolated, and 10 distinct tumor subgroups were firstly identified for subsequent analyses.

The cell trajectory and heterogeneity in tumor microenvironment

Deciphering the intrinsic lineage dynamics of cell clusters is crucial to elucidate their multifaceted roles in tumor

microenvironment (TME) remodeling. Thus, we analyzed the landmark genes with the highest expression in 10 different cell populations and presented them in the form of heat maps (Figure 3A). Cell trajectory analysis indicated six potential cell lineages that existed in these 10 clusters (Figure 3B), and each trajectory was displayed separately with the pseudo-time development (Figure 3C). Then, cell development analysis was conducted for validation, showing that lineage 4 and lineage 7 were the differentiation starting point of cell development, which is consistent with the results of the pseudo-temporal analysis.

Meanwhile, we presented the trends of cell differentiation and development (Figure 3D). We note that C4 is the branching point of all phylogenetic evolutionary trajectories, suggesting that this subgroup has a high degree of phylogenetic plasticity. Dynamic cell trajectory analysis was also performed to confirm the changes of biological function and markers (Figure 3E). Collectively, our findings delineate the cellular ontogeny of ER+ cells and their associated transcriptional reprogramming during tumor progression. The identified molecular signatures may serve as diagnostic biomarkers for ER+ patients' subpopulations and novel therapeutic targets.

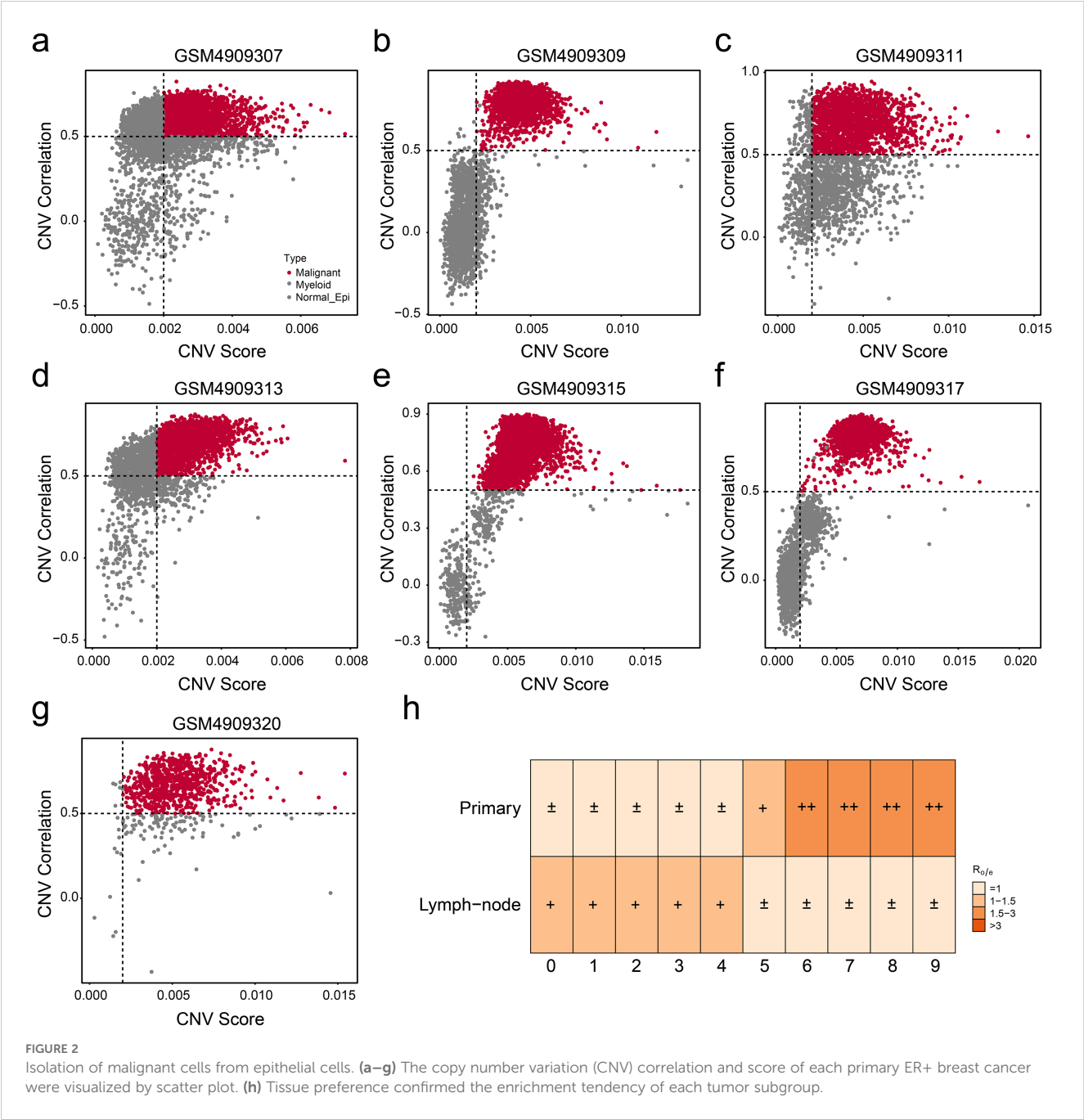


FIGURE 2 Isolation of malignant cells from epithelial cells. (a–g) The copy number variation (CNV) correlation and score of each primary ER+ breast cancer were visualized by scatter plot. (h) Tissue preference confirmed the enrichment tendency of each tumor subgroup.

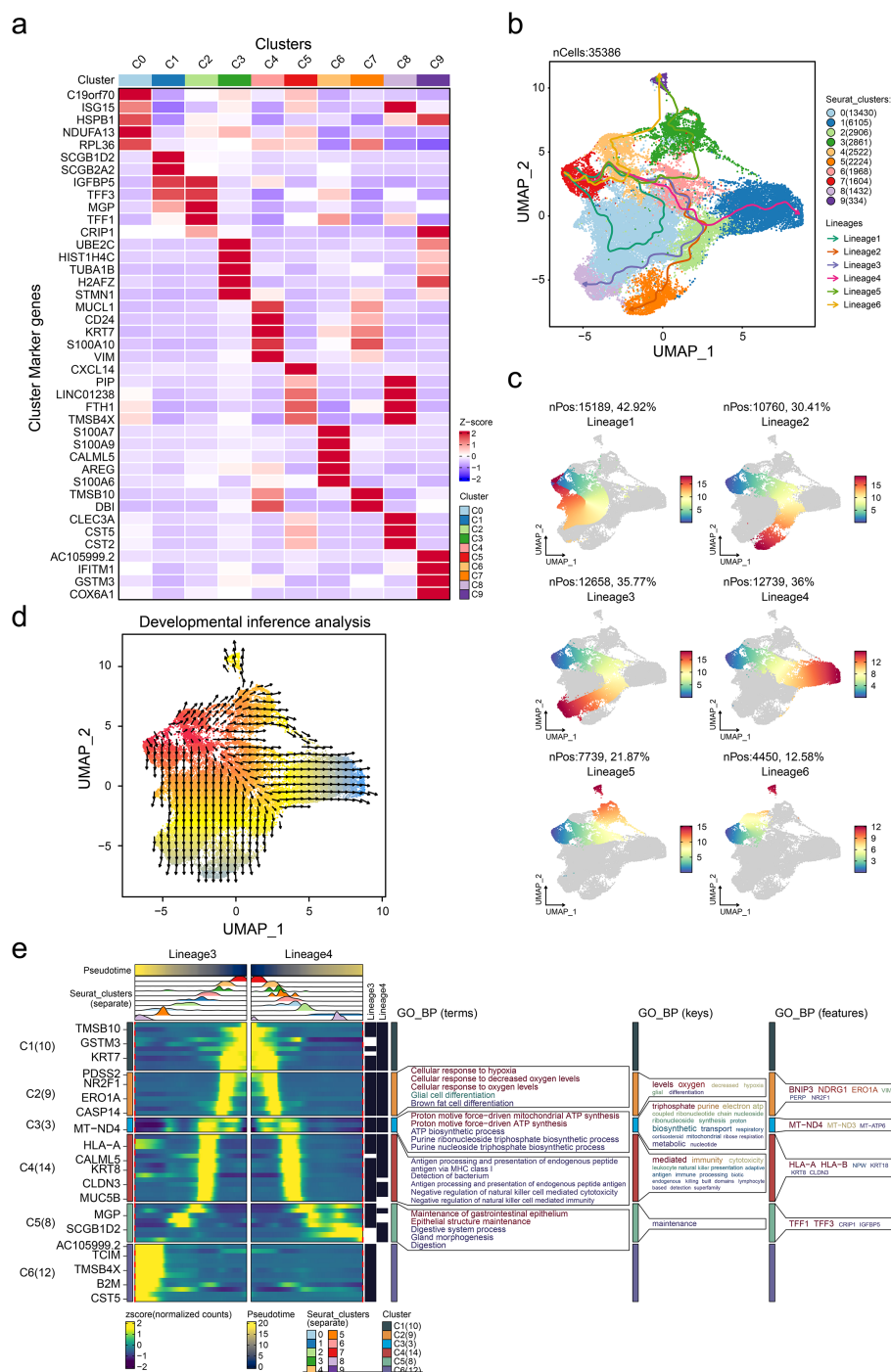


FIGURE 3

Characterization of cell trajectory of tumor cells. (a) Heatmap showing the overexpressed gene expression level of each tumor subgroup.

Pseudotime cell trajectory of tumor cells (b) and six potential cell lineages (c). (d) Cell developmental analysis confirmed the origin of cell states.

(e) Dynamic gene expression heatmap showing the gene expression tendency based on lineages 3 and 4. GO_BP, gene ontology biological process.

MUCL1(+) CD24(+) subcluster was correlated with prognostic outcomes of ER + breast cancer patients

Survival analysis using the KM-plotter platform revealed significant prognostic differences among breast cancer subclusters. Patients in the C4 subgroup exhibiting high invasiveness

demonstrated markedly reduced overall survival (OS), suggesting aggressive clinical behavior (Figure 4A). Based on the analysis results in the previous study (Figure 3A), the markers of the C4 subgroup were identified as MUCL1, CD24, KRT7, S100A10, and VIM (Supplementary Table S1). A multiplex immunohistochemistry (mIHC) was performed on tumor specimens from ER+ breast cancer patients and based on the expression levels of C4-specific

markers (MUCL1 and CD24); these cases were stratified into C4-high invasive and C4-low invasive subgroups (Figure 4B). Subsequently, we evaluated the differences in DFS between the two groups of people, the results of which indicated that the population with a high expression of MUCL1 and CD24 had a worse prognosis (Figure 4C). Besides that, the GSEA enrichment analysis results of the two groups of people showed that high expressions of MUCL1 and CD24 were positively associated with the EMT process (Figure 4D) as well as the TGF- β signaling pathway (Figure 4E). These findings demonstrate that the C4 subgroup, characterized by dual MUCL1(+) CD24(+)

expression, correlates with aggressive tumor behavior and inferior clinical outcomes.

The somatic mutations were associated with the upregulation of MUCL1(+) CD24(+) cells

Given the established role of somatic mutations in modulating cellular infiltration patterns, we propose that tumor-derived genetic

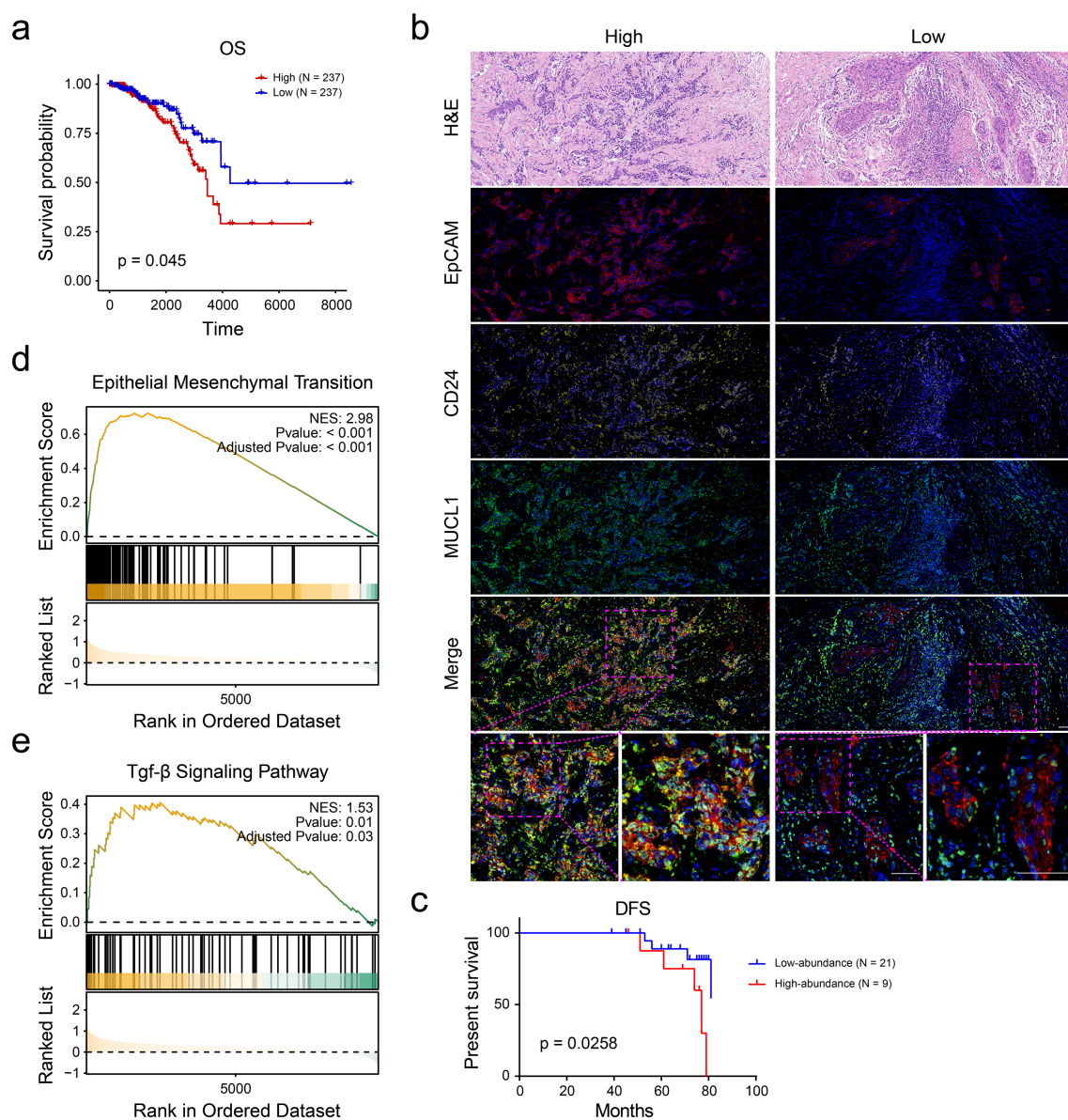


FIGURE 4

CD24(+) MUCL1(+) cells were associated with unfavorable survival of ER+ breast cancer patients. **(a)** Kaplan–Meier curve showing the survival probability of high and low abundance of C4 subgroup according to the median value of enrichment score. **(b)** Multiplex immunohistochemistry indicating the number of CD24(+) MUCL1(+) cells in ER+ breast cancer tissues. **(c)** Disease-free survival outcome of high- and low- abundance of CD24(+) MUCL1(+) cells. **(d, e)** Gene set enrichment analyses suggesting that the high abundance of CD24(+) MUCL1(+) cells was associated with epithelial–mesenchymal transition and TGF-beta signaling pathways.

alterations may similarly regulate MUCL1(+) CD24(+) cell expression within the tumor microenvironment. To identify the most important somatic mutations related to the infiltration of MUCL1(+) CD24(+) cells, we grouped the patients into high- and low-MUCL1(+) CD24(+) groups as previously described. The SNV analysis showed that the somatic mutations of PIK3CA ($P < 0.001$), TAF1 ($P < 0.01$), and AKT1 ($P < 0.05$) were closely associated with the high-MUCL1(+) CD24(+) group (Figure 5A). Notably, the MUCL1(+) CD24(+) high-expression group exhibited significant genetic interaction patterns, with strong co-occurrence and mutual exclusivity relationships among key alterations (Figure 5B). In contrast, these patterns were markedly attenuated in the low-expression cohort (Figure 5C). These results

reveal a subtle relationship between MUCL1(+) CD24(+) cells and somatic mutations.

MUCL1(+) CD24(+) subcluster demonstrated association with immune response modulation

The previous study expounded that C4 subgroup cells may have a prognostic predictive role in ER+ breast cancer. Subsequently, we explored the performance of this type of subgroup in immune response and breast cancer treatment. The immune escape ability of

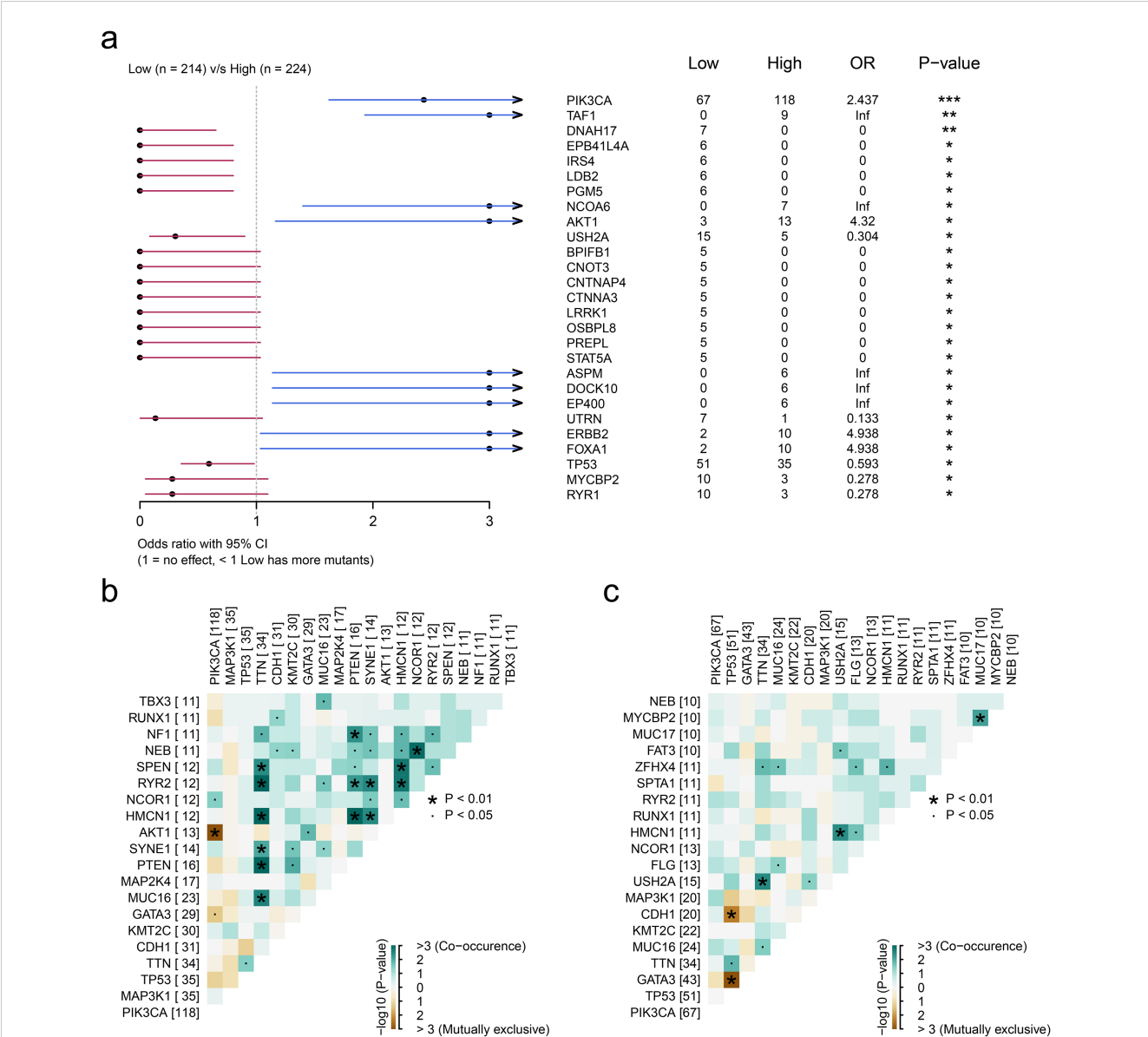


FIGURE 5 Genomic landscape of high and low infiltration levels of CD24(+) MUCL1(+) cells. **(a)** Forest plot illustrating the focal somatic single-nucleotide variant (SNV) being significantly different between patients with high and low CD24(+) MUCL1(+) cell infiltrations. OR > 1 indicates a higher SNV frequency in the high CD24(+) MUCL1(+) cell infiltration group. OR < 1 indicates a higher SNV frequency in the low infiltration group. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. **(b, c)** Landscape of gene co-mutations in patients with low and high abundance of CD24(+) MUCL1(+) cells.

the ER+ breast cancer cell population was evaluated by using the TIDE database (<http://tide.dfci.harvard.edu/>) (Figure 6A), and there were significant differences in the immune response ability in the groups with high and low infiltration of the C4 subpopulation, the results of which indicated that a higher percentage of C4 subpopulation exhibited a lower immune response (Figure 6B). Then, drugs with high sensitivity to the C4 subgroup were screened

out based on the scores. The top-ranked ones include 7b-cis, BMS-345541, and THM-I-94 (Figure 6C). Meanwhile, the included drugs and their potential mechanisms are listed in Figure 6D. These findings revealed that the C4 subgroup is also involved in immune response and drug sensitivity regulatory process. Further research is conducive to providing new targets and treatment strategies for clinical practice.

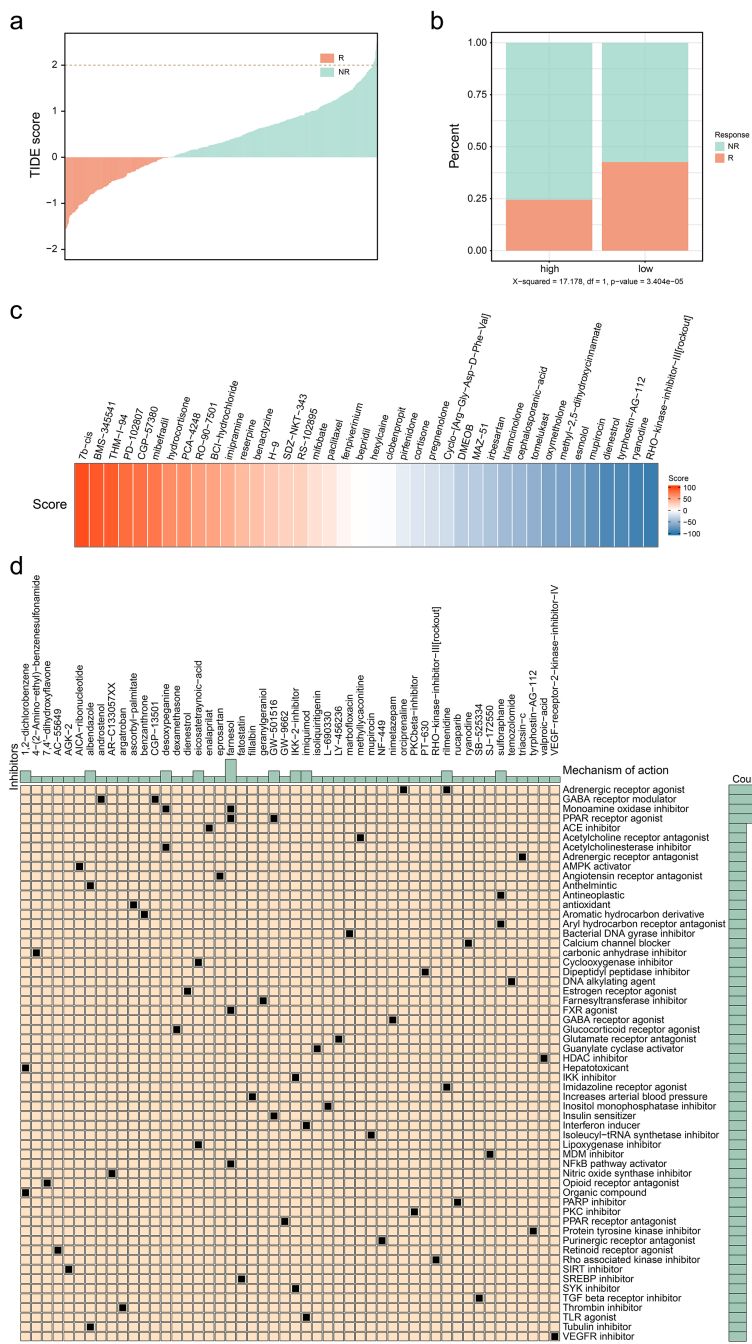


FIGURE 6 Exploration of potential therapies for patients with high infiltration level of CD24(+) MUC1(+). (a) Correlation between TIDE score and immune response status. (b) Correlation between different C4 groups and immune response status. (c) Heatmap showing several results of CMAP with specific scores of drugs. (d) Potential drugs identified for patients with a high infiltration level of C4.

Noninvasive MRI radiomics could be a promising tool to evaluate the infiltration of MUCL1(+) CD24(+) subcluster in ER+ breast cancer

Given the significant clinical associations of the MUCL1(+) CD24(+) population, we systematically evaluated its therapeutic potential applications. Our analysis included 61 ER+ breast cancer patients from the TCGA cohort with paired bulk RNA-seq and dynamic contrast-enhanced MRI (DCE-MRI) data. The tumor regions were segmented using the radiomics module in 3D Slicer software (Figure 7A). Pearson correlation analysis was performed to further identify significant radiomic features associated with the abundance of MUCL1(+) CD24(+) population (Supplementary Table S2). Then, we employed the LASSO regression algorithm based on 14 valuable features to develop a predictive model to estimate MUCL1(+) CD24(+) subpopulation abundances, balancing the feature selection with regularization to optimize model performance (Figures 7B, C, Supplementary Table S3). The cohort was randomly divided into the training set ($N = 43$) and the validation set ($N = 18$). Then, the analysis revealed that the infiltration level of the C4 cluster and the radiomics score showed a strong positive correlation in both the training set (Figure 7D) and the validation set (Figure 7E). Besides that, the exploratory analysis based on ROC curve also showed a potential association between radiomic score and the infiltration of C4 cluster infiltration in both the training set (Figure 7F) and the validation set (Figure 7G). In summary, we established a clinically applicable radiomics model that accurately predicts C4 cluster abundance, offering a non-invasive approach to personalize therapy for ER+ breast cancer patients.

Discussion

This investigation delineates a distinct C4 cellular subpopulation within ER+ breast cancer, identified via single-cell transcriptomic profiling and defined by the co-expression of MUCL1 and CD24. This phenotype demonstrated a significant correlation with adverse clinical outcomes, prompting further interrogation of its biological determinants. We also explored the dynamic evolution of different clusters to reveal the characteristics of each subtype. Additionally, we analyzed potential somatic mutations linked to the infiltration of these clusters. Finally, a radiomic model was established to estimate the abundance of target C4 cluster. To our knowledge, this is the first study to characterize the role of MUCL1(+) CD24(+) cells in ER+ breast cancer using multi-omics strategies, including spatial transcriptomics and radiomics. Our research provides pioneering insights into the pro-tumor effects and potential clinical applications of MUCL1(+) CD24(+) cells in ER+ breast cancer.

Single-cell RNA sequencing has opened up new avenues for the development of tumor markers, and a large number of studies have focused on the subpopulation analysis of the tumor

microenvironment—for instance, Ma et al. (20) identified a distinct luminal subgroup with a high expression level of HPN to diagnose and stratify early-stage prostate cancer by tissue-based single-cell RNA sequencing. Yang et al. (21) found a CEBPB+ tumor subcluster that specifically drives the formation of M2 tumor-associated macrophages to promote malignancy growth in glioblastoma. Additionally, Guo et al. (22) discovered that a metastasis-associated cell cluster overexpressed RAB13 in ovarian cancer by analyzing the primary and pair lymph metastatic node tissues. These studies demonstrate the great potential of single-cell sequencing in the development of tumor biomarkers. It is worth noting that studies analyzing subgroups of breast cancer have also been reported. Wang et al. (15) identified a tumor subgroup that overexpressed NENF, which is associated with distant metastasis of triple-negative breast cancer. Two distinct molecular subtypes of breast cancer stem cells have also been reported by analyzing single-cell RNA data (23). However, few studies have focused on the tumor heterogeneity of ER+ breast cancer. Our study reported a tumor cluster (C4 subgroup) with double-positive status of CD24 and MUCL1 in ER+ breast cancer, which was strongly associated with tumor metastasis. We noted that the C4 subgroup is at the differentiation bifurcation point between primary ER+ breast cancer and lymph node metastases. More interestingly, cell preference analysis showed that the C4 subgroup was enriched in primary tumor tissue, which suggests that the C4 subgroup may be the pre-differentiation state of lymph node metastatic tumor cells, showing a high degree of lineage plasticity. Furthermore, survival analysis and mIHC confirmed the unfavorable role of the C4 subgroup. The tumor metastasis-related signaling pathways including epithelial–mesenchymal transition (24) and TGF-beta signaling (25) were proved to be highly enriched in high abundance of C4 subgroup patients, supporting the pro-metastasis role of C4 subgroup. In summary, our findings reveal and define a class of tumor subpopulations that promote the metastasis of ER+ breast cancer, providing new biomarkers for the diagnosis and treatment of ER+ breast cancer.

The heterogeneity of tumor cell infiltration has been confirmed to be associated with focal somatic mutations (13). We observed that the high infiltration level of the C4 subgroup was associated with the somatic mutations of PIK3CA and FOXA1. PIK3CA-mutated ER+ metastatic breast cancer patients have been reported to demonstrate a poor outcome and resistance to chemotherapy (26). Meanwhile, FOXA1 mutations were confirmed to be associated with a lower response to aromatase inhibitors (27). These results reveal the source of infiltration heterogeneity in the C4 subpopulation and potential targeted therapeutic strategies for the C4 subpopulation. We next explore the novel therapy treatment for the C4 subgroup. We found that patients with high abundance of the C4 subgroup presented a lower proportion of immune responses by *in silico* analysis, indicating that immunotherapy may not be suitable for patients with a high infiltration of the C4 cluster. We utilized the CMAP database for the identification of potential inhibitors to target the C4 subgroup, which provides a theoretical basis for individual treatment for ER+ breast cancer.

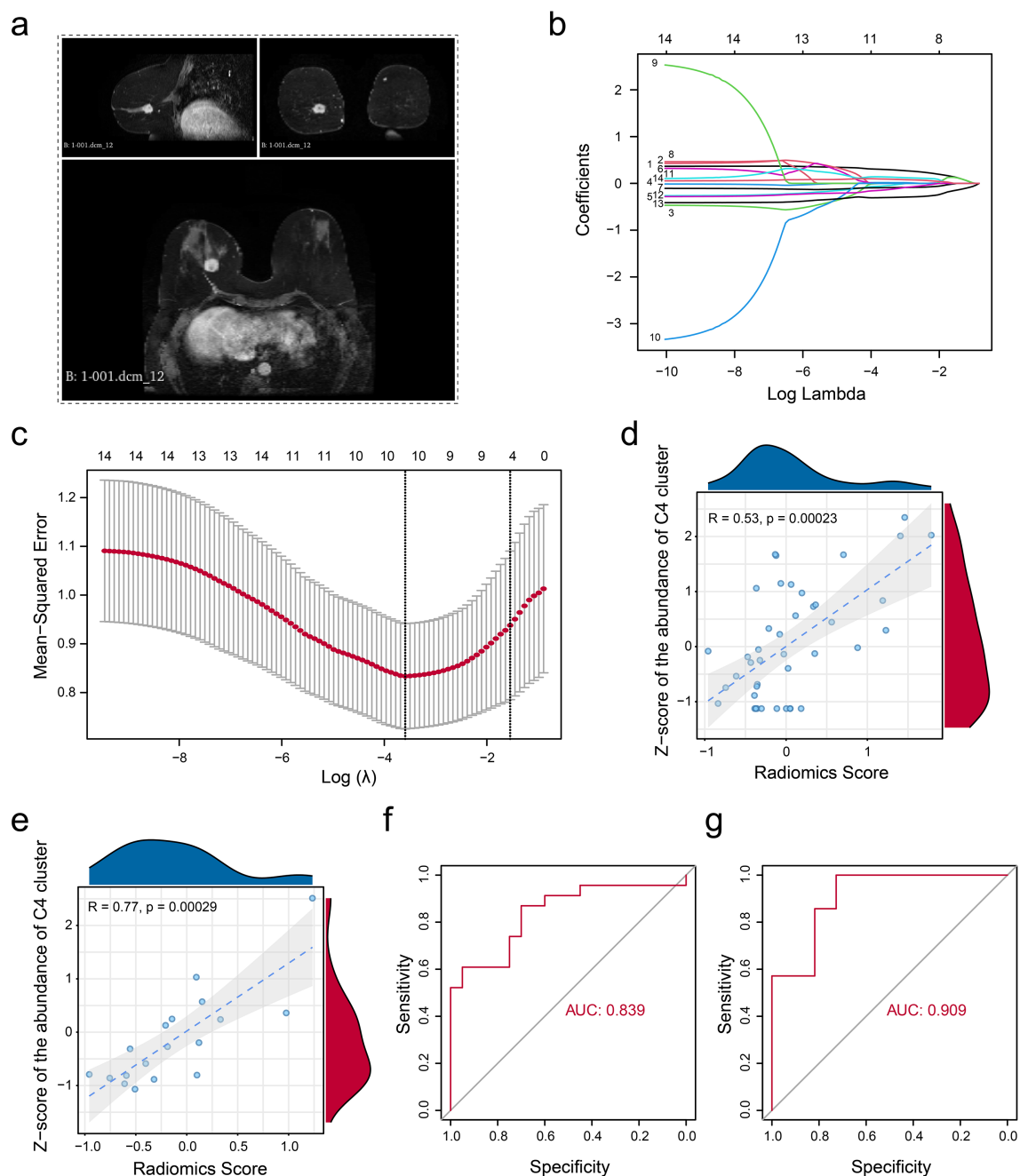


FIGURE 7

Non-invasive radiomic model construction. (a) Example indicating the segmentation of gross tumor volume on DCE-T1 MRI. (b) Parameter tuning plot for the LASSO regression analysis. (c) Distribution of coefficients for variables in the LASSO regression is presented, with each curve representing a radiomics feature filtered using Pearson's correlation. (d, e) Pearson's correlation was calculated in the training set (d) and the validation set (e) between the z-score-normalized abundance of CD24(+) MUCL1(+) cells assessed by the transcriptome and the fitted value obtained from the linear regression radiomics model. (f, g) ROC curve indicating the model's ability to discriminate the abundance of CD24(+) MUCL1(+) cells in both the training set (f) and the validation set (g).

Non-invasive assessment of radiomics has also been applied to a variety of tumors (28, 29)—for example, Wang et al. (13) used single-cell RNA to confirm the favorable role of gamma-delta T cells and developed a radiomic score to evaluate the infiltration level of gamma-delta T cells and the application of radiomics. In this study,

we constructed a radiomic model to estimate the abundance of the MUCL1(+) CD24(+) subcluster. Both the training set (AUC = 0.839) and the validation set (AUC = 0.909) demonstrated a good discriminatory ability in identifying the abundance of the C4 subgroup. Overall, we constructed a model for noninvasive

assessment of C4 subset abundance based on the imaging features of the C4 subset, which has good efficacy and may serve as a potential tool for future clinical translational applications.

It needs to be clarified that there are also deficiencies in our research. While our multi-omics integration provides comprehensive insights, cross-platform validation using alternative sequencing technologies (e.g., single-nuclei RNA-seq, spatial proteomics) would strengthen the findings. Additionally, *in vivo* and *in vitro* experiments need to be conducted to further explore the molecular function of the MUCL1(+) CD24(+) tumor cluster. Future work should incorporate functional validation through mechanistic studies and expand clinical correlation using independent cohorts. Although the current radiomic analysis serves as a proof of concept, prospective collection of multicenter MRI datasets will be essential for clinical translation.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The studies involving humans were approved by the Ethical Committee of Tianjin Medical University Cancer Institute and Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YL: Conceptualization, Writing – original draft, Writing – review & editing. JC: Conceptualization, Writing – original draft. LH: Writing – original draft. JZ: Investigation, Writing – original draft. DT: Writing – original draft. YY: Conceptualization, Writing – original draft. ZC: Conceptualization, Writing – original draft. YT: Conceptualization, Writing – original draft, Writing – review & editing.

References

1. Harbeck N, Gnant M. Breast cancer. *Lancet*. (2017) 389:1134–50. doi: 10.1016/S0140-6736(16)31891-8
2. Fahad UM. Breast cancer: current perspectives on the disease status. *Adv Exp Med Biol*. (2019) 1152:51–64. doi: 10.1007/978-3-030-20301-6_4
3. Roulstone V, Kyula-Currie J, Wright J, Patin EC, Dean I, Yu L, et al. Author Correction: Palbociclib and dsRNA sensor co-operate to enhance anti-cancer effects through ER stress and modulation of immune evasion. *Nat Commun*. (2025) 16:6510. doi: 10.1038/s41467-025-61212-3
4. Mamann A, Pradat Y, Bidard FC, Delaloge S, Cabel L, Faull I, et al. Prognostic significance of early on-treatment evolution of circulating tumor DNA in advanced ER+/HER2- breast cancer. *Ann Oncol*. (2025) 5:S0923–7534. doi: 10.1016/jannonc.2025.06.015
5. Broege A, Rossetti S, Sen A, de la Forest A, Davis L, Seibel M, et al. Functional analysis of the PI3K/AKT/mTOR pathway inhibitor, gedatolisib, plus fulvestrant with and without palbociclib in breast cancer models. *Int J Mol Sci*. (2025) 26:5844. doi: 10.3390/ijms26125844
6. Liu Y, Park S, Li Y. Breaking cancer's momentum: CDK4/6 inhibitors and the promise of combination therapy. *Cancers (Basel)*. (2025) 17:1941. doi: 10.3390/cancers17121941
7. Zafar H, Lin C, Bar-Joseph Z. Single-cell lineage tracing by integrating CRISPR-Cas9 mutations with transcriptomic data. *Nat Commun*. (2020) 11:3055. doi: 10.1038/s41467-020-16821-5
8. Wang Q, Sun K, Liu R, Song Y, Lv Y, Bi P, et al. Single-cell transcriptome sequencing of B-cell heterogeneity and tertiary lymphoid structure predicts breast

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by the National Natural Science Foundation of China (grant nos. 82304025, 82403020, and 82303857).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2025.1695689/full#supplementary-material>

cancer prognosis and neoadjuvant therapy efficacy. *Clin Transl Med.* (2023) 13:e1346. doi: 10.1002/ctm2.1346

9. Oikonomou EK, Sangha V, Vasisht SS, Coppi A, Krumholz HM, Nasir K, et al. Artificial intelligence-enabled electrocardiography and echocardiography to track preclinical progression of transthyretin amyloid cardiomyopathy. *Eur Heart J.* (2025) 46:3651–3662. doi: 10.1093/eurheartj/ehaf450

10. Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J.* (2021) 40:e107333. doi: 10.15252/embj.2020107333

11. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* (2015) 33:495–502. doi: 10.1038/nbt.3192

12. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst.* (2019) 8:329–337.e4. doi: 10.1016/j.cels.2019.03.003

13. Wang G, Wang S, Song W, Lu C, Chen Z, He L, et al. Integrating multi-omics data reveals the antitumor role and clinical benefits of gamma-delta T cells in triple-negative breast cancer. *BMC Cancer.* (2025) 25:623. doi: 10.1186/s12885-025-14029-8

14. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science.* (2014) 344:1396–401. doi: 10.1126/science.1254257

15. Wang G, Shi C, He L, Li Y, Song W, Chen Z, et al. Identification of the tumor metastasis-related tumor subgroups overexpressed NENF in triple-negative breast cancer by single-cell transcriptomics. *Cancer Cell Int.* (2024) 24:319. doi: 10.1186/s12935-024-03505-z

16. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007

17. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* (2016) 44:e71. doi: 10.1093/nar/gkv1507

18. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* (2018) 28:1747–56. doi: 10.1101/gr.239244.118

19. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell.* (2017) 171:1437–1452.e17. doi: 10.1016/j.cell.2017.10.049

20. Ma X, Guo J, Liu K, Chen L, Liu D, Dong S, et al. Identification of a distinct luminal subgroup diagnosing and stratifying early stage prostate cancer by tissue-based single-cell RNA sequencing. *Mol Cancer.* (2020) 19:147. doi: 10.1186/s12943-020-01264-9

21. Yang Y, Jin X, Xie Y, Ning C, Ai Y, Wei H, et al. The CEBPB(+) glioblastoma subcluster specifically drives the formation of M2 tumor-associated macrophages to promote Malignancy growth. *Theranostics.* (2024) 14:4107–26. doi: 10.7150/thno.93473

22. Guo J, Han X, Li J, Li Z, Yi J, Gao Y, et al. Single-cell transcriptomics in ovarian cancer identify a metastasis-associated cell cluster overexpressed RAB13. *J Transl Med.* (2023) 21:254. doi: 10.1186/s12967-023-04094-7

23. Wang G, Chen Z, Tian Y, Zhu Y, Wang S, Song W, et al. Multi-omics profiling identifies a high-risk subgroup of breast cancer stem cells for prognostic stratification and personalized treatment. *J Cancer.* (2025) 16:1860–72. doi: 10.7150/jca.109589

24. Huang Y, Hong W, Wei X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J Hematol Oncol.* (2022) 15:129. doi: 10.1186/s13045-022-01347-8

25. Tian Y, Yu Y, Hou LK, Chi JR, Mao JF, Xia L, et al. Serum deprivation response inhibits breast cancer progression by blocking transforming growth factor-beta signaling. *Cancer Sci.* (2016) 107:274–80. doi: 10.1111/cas.12879

26. Mosele F, Stefanovska B, Lusque A, Tran DA, Garberis I, Droin N, et al. Outcome and molecular landscape of patients with PIK3CA-mutated metastatic breast cancer. *Ann Oncol.* (2020) 31:377–86. doi: 10.1016/j.annonc.2019.11.006

27. Arruabarrena-Aristorena A, Maag J, Kittane S, Cai Y, Karthaus WR, Ladewig E, et al. FOXA1 mutations reveal distinct chromatin profiles and influence therapeutic response in breast cancer. *Cancer Cell.* (2020) 38:534–550.e9. doi: 10.1016/j.ccell.2020.08.003

28. Chen Z, Yang C, Wang J, Wang B, He L, Liu Z, et al. Integrating of radiomics and single-cell transcriptomics uncovers the crucial role of gammadelta T cells in lung adenocarcinoma. *J Cancer.* (2025) 16:1167–80. doi: 10.7150/jca.105586

29. Li J, Cao Y, Liu Y, Yu L, Zhang Z, Wang X, et al. Multiomics profiling reveals the benefits of gamma-delta (gammadelta) T lymphocytes for improving the tumor microenvironment, immunotherapy efficacy and prognosis in cervical cancer. *J Immunother Cancer.* (2024) 12:e008355. doi: 10.1136/jitc-2023-008355