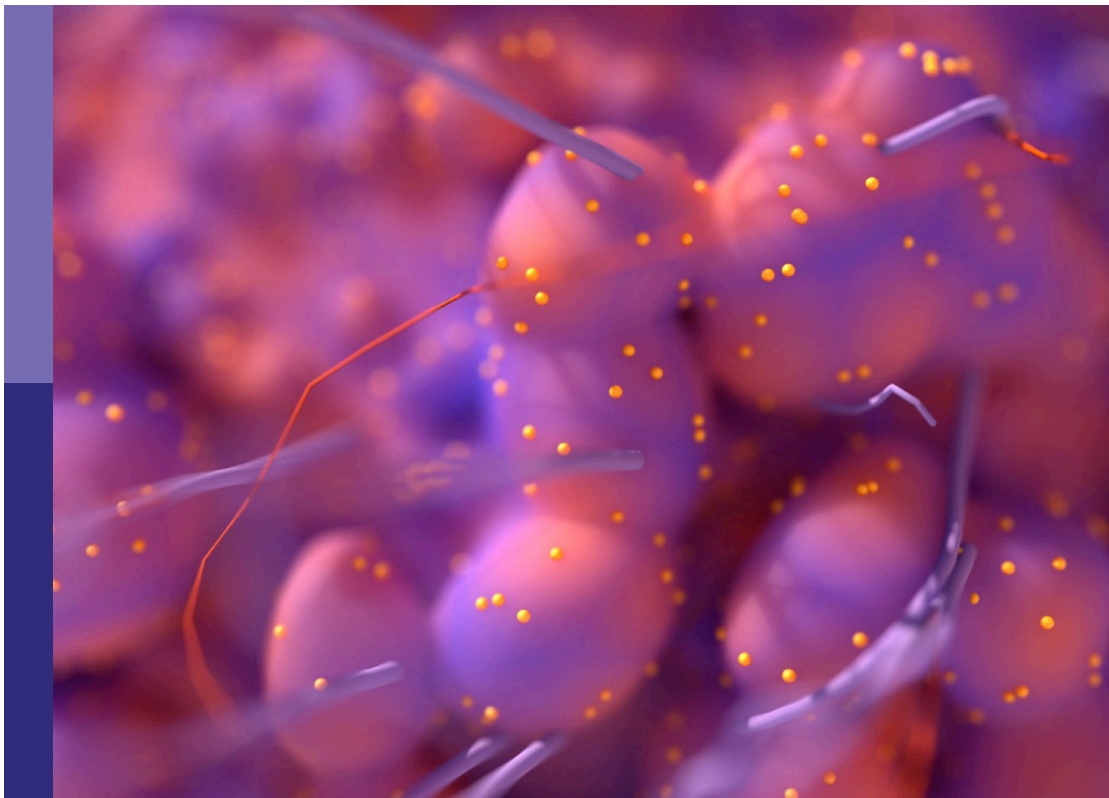
frontiers | Research Topics

Epigenetic and metabolic regulators of breast carcinogenesis

Edited by
Iman Mamdouh Talaat, Maha Mohamed Saber-Ayad, Hauke Busch and Noha Mousaad Elemam

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-4741-0
DOI 10.3389/978-2-8325-4741-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Epigenetic and metabolic regulators of breast carcinogenesis

Topic editors

Iman Mamdouh Talaat — University of Sharjah, United Arab Emirates
Maha Mohamed Saber-Ayad — University of Sharjah, United Arab Emirates
Hauke Busch — University of Lübeck, Germany
Noha Mousaad Elemam — University of Sharjah, United Arab Emirates

Citation

Talaat, I. M., Saber-Ayad, M. M., Busch, H., Elemam, N. M., eds. (2024). *Epigenetic and metabolic regulators of breast carcinogenesis*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4741-0

Table of contents

05 Editorial: Epigenetic and metabolic regulators of breast carcinogenesis
Iman Mamdouh Talaat, Maha Saber-Ayad, Hauke Busch and Noha Mousaad Elemam

08 Prognostic value of 12 m7G methylation-related miRNA markers and their correlation with immune infiltration in breast cancer
Wenchuan Zhang, Shuwan Zhang and Zhe Wang

23 Metabolic syndrome is a risk factor for breast cancer patients receiving neoadjuvant chemotherapy: A case-control study
Zhaoyue Zhou, Yue Zhang, Yue Li, Cong Jiang, Yang Wu, Lingmin Shang, Yuanxi Huang and Shaoqiang Cheng

38 Breast density and estradiol are associated with distinct different expression patterns of metabolic proteins in normal human breast tissue *in vivo*
Jimmy Ekstrand, Annelie Abrahamsson, Peter Lundberg and Charlotta Dabrosin

52 DNA methylation differences in noncoding regions in ER negative breast tumors between Black and White women
Jianhong Chen, Michael J. Higgins, Qiang Hu, Thaer Khoury, Song Liu, Christine B. Ambrosone and Zhihong Gong

63 DNA methylation profiling in Trisomy 21 females with and without breast cancer
Yosra Bejaoui, Sara Alresheq, Sophie Durand, Marie Vilaire-Meunier, Louise Maillebouis, Ayman Al Haj Zen, André Mégarbané and Nady El Hajj

71 Identification and comprehensive analysis of epithelial–mesenchymal transition related target genes of miR-222-3p in breast cancer
Yutong Fang, Qunchen Zhang, Chunfa Chen, Zexiao Chen, Rongji Zheng, Chuanghong She, Rendong Zhang and Jundong Wu

91 Next-generation sequencing-based detection in a breast MMPMN patient with EGFR T790M mutation: a rare case report and literature review
Huiyun Lv, Aijuan Tian, Shanshan Zhao, Jinbo Zhao and Chen Song

97 Molecular mechanisms of multi-omic regulation in breast cancer
Soledad Ochoa and Enrique Hernández-Lemus

114 Radiomic signatures based on pretreatment 18F-FDG PET/CT, combined with clinicopathological characteristics, as early prognostic biomarkers among patients with invasive breast cancer
Tongtong Jia, Qingfu Lv, Xiaowei Cai, Shushan Ge, Shibiao Sang, Bin Zhang, Chunjing Yu and Shengming Deng

127 Research progress of Claudin-low breast cancer
Chenglong Pan, Anqi Xu, Xiaoling Ma, Yanfei Yao, Youmei Zhao, Chunyan Wang and Ceshi Chen

139 Intra-tumoral microbial community profiling and associated metabolites alterations of TNBC
Yi Wang, Dingding Qu, Yali Zhang, Yiping Jin, Yu Feng, He Zhang and Qingxin Xia

150 Prognostic value analysis of cholesterol and cholesterol homeostasis related genes in breast cancer by Mendelian randomization and multi-omics machine learning
Haodong Wu, Zhixuan Wu, Daijiao Ye, Hongfeng Li, Yinwei Dai, Ziqiong Wang, Jingxia Bao, Yiying Xu, Xiaofei He, Xiaowu Wang and Xuanxuan Dai

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 28 March 2024
DOI 10.3389/fonc.2024.1383043

OPEN ACCESS

EDITED AND REVIEWED BY
Sharon R. Pine,
University of Colorado, United States

*CORRESPONDENCE
Iman Mamdouh Talaat
italaat@sharjah.ac.ae

RECEIVED 06 February 2024
ACCEPTED 21 March 2024
PUBLISHED 28 March 2024

CITATION
Talaat IM, Saber-Ayad M, Busch H and Elemam NM (2024) Editorial: Epigenetic and metabolic regulators of breast carcinogenesis. *Front. Oncol.* 14:1383043.
doi: 10.3389/fonc.2024.1383043

Editorial: Epigenetic and metabolic regulators of breast carcinogenesis

Iman Mamdouh Talaat[1,2]*, Maha Saber-Ayad[1,2], Hauke Busch[3,4] and Noha Mousaad Elemam[1,2]

[1]Department of Clinical Sciences, College of Medicine, University of Sharjah, Sharjah, United Arab Emirates, [2]Research Institute for Medical and Health Sciences, University of Sharjah, Sharjah, United Arab Emirates, [3]Division for Medical Systems Biology, Lübeck Institute of Experimental Dermatology, Universität zu Lübeck, Lübeck, Germany, [4]University Cancer Center Schleswig-Holstein, University Hospital Schleswig-Holstein, Lübeck, Germany

KEYWORDS

breast cancer, epigenetics, cancer metabolism, carcinogenesis, DNA methylation, miRNA

Editorial on the Research Topic
Epigenetic and metabolic regulators of breast carcinogenesis

1 Introduction

This editorial features the articles published in this Research Topic in Frontiers in Oncology, which aimed to uncover the different epigenetic aspects and metabolic processes involved in breast tumorigenesis. The article by Jia et al. assessed the predictive capacity of fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography (18 F-FDG PET/CT) in prognostic risk assessment of invasive breast cancer (BC) patients. In a retrospective analysis of 91 patients undergoing preoperative 18 F-FDG PET/CT, radiomic signatures (RSs) were identified, and a radiomic score (Rad-score) was computed. The Rad-score, along with other factors, was independently associated with progression-free survival and overall survival. The clinicopathologic-radiomic-based model outperformed single clinical or radiomic models, exhibiting good predictive performance and enhanced individualized prognosis estimation. Therefore, integrating clinicopathological risks with Rad-score provides a robust method for prognostic evaluation in invasive BC patients, enhancing the accuracy of outcome predictions.

In the study by Pan et al., the authors explored the origin, molecular and pathological characteristics, treatment, and prognosis of claudin-low BC (CLBC). They highlighted that CLBC displays a higher histological grade and a greater likelihood of spreading to lymph nodes compared to other subtypes. Moreover, it is often associated with increased invasiveness as well as a less favorable prognosis and a lower likelihood of complete remission for CLBC. Hence, this aims to contribute to a comprehensive understanding and lay the groundwork for personalized BC treatments. This could contribute to a comprehensive understanding and lay the groundwork for personalized BC treatments.

From an epigenetic perspective, a study explored another aspect associated with BC progression and epithelial-mesenchymal transition (EMT) by identifying EMT-associated target genes (ETGs) of miR-222-3p. Their bioinformatic analysis showed that miR-222-3p might be a specific biomarker of basal-like BC. Furthermore, 10 core ETGs of miR-222-3p

were identified where some of these genes might be useful diagnostic and prognostic biomarkers. The comprehensive analysis of these 10 ETGs and miR-222-3p indicated that they might be involved in the development of BC, shedding light on their potential as therapeutic targets for BC treatment (Fang et al.). On the other hand, the study by Zhang et al. developed a prognostic model for BC based on RNA guanine-7 methyltransferase (RNMT), FAM103A1, and 12 related microRNAs. Utilizing data from The Cancer Genome Atlas and TargetScan, a risk prognosis model accurately predicts 1-, 3-, 5-, and 10-year survival rates (>0.7 AUC). Such a model was linked to immune infiltration, suggesting potential immunotherapeutic targets for BC.

Down Syndrome (DS) patients present a unique cancer profile with a low risk of solid tumors but a higher risk of leukemia. A study by Bejaoui et al. explored DNA methylation and epigenetic aging in DS individuals with and without BC. Using the Infinium Methylation EPIC BeadChip array, differentially methylated sites in DS individuals with BC (T21-BC) were identified and linked to gene expression changes. Enriched processes included serine-type peptidase activity, epithelial cell development, GTPase activity, bicellular tight junction, and Ras protein signal transduction. Interestingly, epigenetic age acceleration analysis revealed no difference between T21-BC and DS individuals without BC (T21-BCF). This pioneering research illuminates DNA methylation variations in DS women, offering insights into potential protective factors against BC in DS.The prevalence of ER-negative (ER^-) BC is higher in African American/Black women than in other US ethnic groups. A study by Chen et al. explored genome-wide DNA methylation in ER^- tumors, initially focusing on protein-coding genes and later delving into 96 differentially methylated loci (DMLs) in intergenic and noncoding RNA regions. Using Illumina Infinium Human Methylation 450K array and RNA-seq data, 23 DMLs were found to significantly correlate with the expression of 36 genes within a 1Mb radius. One hypermethylated DML (cg20401567) in ER^- tumors from Black women mapped to a potential enhancer downstream of HOXB2, indicating reduced HOXB2 expression. Independent analysis of 207 ER^- BC from TCGA confirmed this, suggesting that epigenetic disparities may influence BC pathogenesis in ER^- tumors between Black and White women.An interesting study explored olfactory receptors, specifically G protein-coupled surface receptors, that are increasingly relevant in carcinogenesis and metastasis. Their ectopic expression, influenced by environmental factors, can lead to methylation aberrations. This study identified 68 differentially methylated olfactory receptors in BC. Notably, hypomethylation events included BC signatures. Network analysis suggests a pivotal role of those receptors in stimulating metastasis-related pathways. Phenotypic smell tests revealed a generalized impairment in BC patients, independent of chemotherapy, highlighting olfaction's crucial role in carcinogenesis. Olfaction receptors were shown as a potential factor of carcinogenesis in a well-characterized BC subset (Fessahaye et al.).

Recent advancements in genomics and other high-throughput biomolecular techniques, collectively referred to as "-omics," have provided valuable insights into the molecular processes driving the development and progression of BC. Numerous mechanisms involved in these processes operate at multiple regulatory levels. The review article by Ochoa and Hernández-Lemus aimed to present a comprehensive overview of the current understanding of how various omics, such as DNA methylation, non-coding RNA, and other epigenomic changes, contribute to the regulation of BC. The molecular intricacies of multi-omic regulation in BC hold significant promise and could guide the development of innovative therapeutic strategies for BC. Additionally, a case report by Lv et al. presented a patient with primary ovarian and breast cancers, a condition with rising incidence due to the advances in early cancer detection. Using the technology of next-generation sequencing, a rare EGFR T790M mutation was detected in the patient's primary BC tissue. A therapeutic recommendation with the targeted therapy "Osimertinib" was subsequently identified based on this mutation. In addition to the interesting case report, a mini-literature review was provided.

Several studies explored various angles of metabolic processes and their regulators in breast tumorigenesis. The case-control study by Zhou et al. investigated the impact of metabolic syndrome (MetS) on BC patients undergoing neoadjuvant chemotherapy (NAC). Among 221 female BC patients, 24.0% achieved pathologic complete response (pCR) after NAC. MetS was an independent predictor of lower pCR rates. Also, metabolic parameters, particularly blood lipid index, significantly worsened post-NAC. Over a 6-year follow-up, MetS was strongly linked to increased recurrence and mortality. The risk of death and disease progression rose with the number of MetS components. These findings suggest that MetS in BC patients undergoing NAC is associated with poorer outcomes. Another study focused on the challenges of triple-negative BC (TNBC) treatment, emphasizing the lack of therapeutic targets and poor prognosis. Using 16S rRNA MiSeq sequencing and metabolomic analysis on formalin-fixed, paraffin-embedded (FFPE) tissue samples, the research identifies Turicibacter's higher abundance in TNBC, along with distinct metabolites. Significant correlations were found between intra-tumoral microbiome, clinicopathological characteristics, and HER2 expression. Microbial taxa associated with tumor-infiltrating lymphocytes suggest potential markers for antitumor immunity. The study's innovative use of FFPE samples offers insights into diagnostic biomarkers, therapeutic strategies, and early TNBC clinical diagnosis (Wang et al.). Among various risk factors for BC, breast density and exposure to sex steroids are considered major ones. The research question posed by Ekstrand et al. was to explore whether those two key factors could affect extracellular space metabolism-regulating proteins, thus providing potential diagnostic and therapeutic markers. The investigators reported differentially expressed genes in both conditions and showed that two proteins, namely, pro-cathepsin H and galanin peptide, were similarly regulated in BC, dense- and estrogen-exposed breasts. The study underscores the potential role of metabolic proteins in better understanding the disease pathogenesis, diagnosis, and therapy. Furthermore, BC patients have been frequently observed with deranged lipid profiles and cholesterol metabolism. In the study conducted by Wu et al., an analysis of expression patterns of 73 cholesterol homeostasis-related genes was implemented on BC samples in the TCGA cohort with

consensus clustering analysis. They used machine learning to compare multi-omics of different samples, aiming to predict the disease prognosis in different risk groups. The study could decipher the signature of cholesterol homeostasis-related genes for several key processes, namely, angiogenesis, immune responses, and therapeutic response.

2 Conclusion

In conclusion, the collective articles presented in this editorial, "*Epigenetic and metabolic regulators of breast carcinogenesis*," underscore the intricate interplay between epigenetic mechanisms and metabolic dysregulation in breast cancer development. The findings from the contributed papers have illuminated various pathways and molecular mechanisms implicated in breast carcinogenesis. Moving forward, this comprehensive understanding offers promising avenues for future research endeavors, ranging from targeted therapeutic interventions to precision medicine approaches. This collection of articles not only enriches our current understanding of breast cancer pathogenesis but also serves as a catalyst for driving innovative research directions aimed at advancing diagnostics and prognostics and, ultimately, improving patient outcomes in this critical area of oncology.

Author contributions

IT: Writing – original draft, Writing – review & editing. MS-A: Writing – original draft, Writing – review & editing. HB: Writing – original draft, Writing – review & editing. NE: Writing – original draft, Writing – review & editing.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 05 August 2022
DOI 10.3389/fonc.2022.929363

OPEN ACCESS

EDITED BY
Xiaowei Qi,
Army Medical University, China

REVIEWED BY
Zheng Wang,
Shanghai Jiao Tong University, China
Bin Fu,
The First Affiliated Hospital of Nanchang University, China
Jin-Zhou Xu,
Huazhong University of Science and Technology, China
Wenjun Yi,
Central South University, China

*CORRESPONDENCE
Zhe Wang
wz_cmu@126.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

RECEIVED 26 April 2022
ACCEPTED 12 July 2022
PUBLISHED 05 August 2022

CITATION
Zhang W, Zhang S and Wang Z (2022) Prognostic value of 12 m7G methylation-related miRNA markers and their correlation with immune infiltration in breast cancer. *Front. Oncol.* 12:929363. doi: 10.3389/fonc.2022.929363

Prognostic value of 12 m7G methylation-related miRNA markers and their correlation with immune infiltration in breast cancer

Wenchuan Zhang†, Shuwan Zhang† and Zhe Wang*

Department of Pathology, Shengjing Hospital of China Medical University, Shenyang, China

RNA guanine-7 methyltransferase (RNMT), in complex with FAM103A1, plays an important role in tumorigenesis and development. The aim of this study was to establish a prognostic model of RNMT and FAM103A1-based upstream microRNAs and explore its correlation with immune cell infiltration in breast cancer (BC) while investigating its potential prognostic value and verify the model by quantitative real-time polymerase chain reaction (qRT-PCR). The miRNA expression data upstream of the m7G methyltransferase complex RNMT/FAM103A1 in BC was obtained from The Cancer Genome Atlas and TargetScan databases. We performed univariate Cox regression, LASSO regression, Kaplan-Meier survival, and principal component analyses, along with risk prognostic modelling. Based on multivariate Cox regression analysis, a total of 12 m7G methyltransferase-related miRNAs were found. The model showed good accuracy for predicting the 1-, 3-,5-, and 10-year survival rates, and the areas under the curve were almost >0.7. To characterize the risk-level model constructed from 12 miRNAs, 12 differentially expressed mRNAs related to prognosis and immune infiltration were obtained. The prognosis of BC patients is well predicted by the risk model we constructed. This model is also closely related to immune infiltration, and new immunotherapy targets can be explored from this field.

KEYWORDS

miRNA, RNMT, FAM103A1, m7G, immune infiltration, breast cancer

Introduction

Breast cancer (BC) is the most common cancer worldwide, surpassing lung cancer, and it has the highest incidence rate of malignancy. In 2020, there were approximately 2.3 million new cases worldwide (accounting for 11.7% of all cancer incidence rates). According to data from international research institutions, this number is expected to

increase to more than 3 million by 2040 (1). However, the side effects of the traditional treatment methods (including surgery, radiotherapy, chemotherapy, and endocrine therapy) are intolerable for the patients. In recent years, immunotherapy has achieved great success in treating melanoma, non-small-cell lung cancer, acute lymphoblastic leukaemia, and other tumors. BC patients traditionally considered to have "weak immunogenicity" are expected to benefit from immunotherapy (2). Compared with traditional treatment methods, immunotherapy is well tolerated, has no toxic drug accumulation, and can prevent adverse reactions caused by systemic therapy (3). Therefore, there is an urgent need to identify novel and effective prognostic markers and therapeutic targets. The occurrence and development of tumors are closely related to genetic and epigenetic changes. Recently, m7G has been shown to play a crucial role in various stages of RNA transcription, processing, degradation, and translation. Efficient ex-pression of genes in eukaryotes requires the addition of a 7-methylguanosine cap at the 5′ end of mRNA, which is an inverted 7-methylguanosine group connected to the first transcriptional nucleotide on RNA polymerase (Pol II) transcripts (4, 5). 7-Methylguanosine is attached to the transcript through a triphosphate from the 5′ hydroxyl group to generate a structure designated as m7G (5′)PPP(5′)X (where X is the first nucleotide transcribed). This unique molecular structure within the cell is thought to specifically target the 5′ end of RNA Pol II transcripts for several gene regulatory processes, including splicing, nuclear export of mRNA, and translation initiation (6, 7). The methyl cap also protects RNA from exonucleases until it is removed by a decapping enzymes (8). Enzymes that catalyse methyl cap synthesis are essential in organisms from yeast to humans. In mammals, these enzymes are RNA guanylyltransferase, 5′-phosphatase (RNGTT), and RNA guanine-7 methyltransferase (RNMT) (9). Among them, RNMT catalyses the methylation of the cap at the N 7 position to generate a methyl cap, resulting in the m7G(5′)PPP(5′)X (10–12). Recently, a study has shown that the proliferation rate of untransformed mammary epithelial cells does not change when cellular RNMT activity is reduced by 50%, whereas some BC cell lines show reduced proliferation and increased apoptosis. While the activity of RNMT is enhanced in most BC cell lines, PIK3CA, which encodes the p110a subunit PI3Ka, is oncogenically mutated. In contrast, all cell lines insensitive to RNMT depletion expressed wild-type PIK3CA. This indicates that inhibition of RNMT activity can inhibit oncogenic mutation of PIK3CA, thereby reducing the proliferation of cancer cells (13). Studies have shown that some cellular signalling pathways can regulate the formation of mRNA caps on specific target genes, thereby regulating their expression. For example, c-Myc and E2F1 increase the phosphorylation of RNA pol II, thereby promoting mRNA cap formation by re-recruiting methyl cap synthetic enzymes (14–16). During the cell cycle, mRNA cap formation is also regulated by CDK1-dependent phosphorylation (17). However, there is no research on the role of miRNAs upstream of RNMT in regulating target genes. Some studies have found that RNMT does not function as a monomer but that it forms a CAP methyltransferase complex with FAM103A1 to promote cap maturation and maintain mRNA levels for mRNA translation and cell survival (18). Therefore, this study aimed to investigate whether there is upstream miRNA regulation of the m7G methyltransferase complex RNMT/FAM103A1. Tumor initiation and progression is a complex process that requires interactions between cancer cells, the microenvironment, and the immune system (19, 20). The importance of the microenvironment and immunomodulatory factors in BC has been known for many years (21). Recent studies have found that in the tumor microenvironment, miRNA patterns associated with the molecular signatures of BC construct a complex immune regulatory network, revealing the biological functions of miRNAs in BC extracellular matrix and immune infiltration (22). Currently, in the study of immune checkpoint inhibitors, the anti-PD-1 antibody pembrolizumab is of great significance in the treatment of triple-negative breast cancer patients. Although there are an increasing number of studies on BC immunotherapy, they are still in the preclinical or clinical trial stage (21, 23, 24). Therefore, this study aimed to investigate the existence of miRNAs upstream of the m7G-modified methyltransferase complex RNMT/FAM103A1 and to identify new prognostic markers and immunotherapy drug targets in BC.

Results

12 important miRNAs are closely related to m7G methyltransferase RNMT/FAM103A1

A research flowchart is presented in Figure 1. A total of 1204 predicted miRNAs related to the m7G methyltransferase target gene RNMT/FAM103A1 in BC patients were analysed for differences (Figure 2A), and 201 miRNAs with differences were obtained, of which 136 were upregulated differentially expressed miRNAs (logFC ≥ 1, FDR < 0.05), and 65 downregulated differentially expressed miRNAs (logFC ≤ -1, FDR < 0.05) (Figure 2C). A heatmap of the top 20 most differentially expressed miRNAs among the 201 miRNAs is shown (Figure 2B). We randomly divided the dataset into two groups according to 0.5: training group and validation group. Univariate Cox analysis was performed on 201 miRNAs, where we set the P-value to < 0.05, to obtain 16 miRNAs that have an impact on prognosis. Among these, hsa-miR-3662, hsa-miR-2115-5p, hsa-miR-483-3p, hsa-miR-21-3p, hsa-miR-6844, hsa-miR-483-5p, hsa-miR-340-5p had a more significant effect on prognosis (P<0.01) (Supplementary Table S1). To further obtain meaningful miRNAs for prognosis, we performed least absolute shrinkage and selection operator (LASSO) regression to screen the 14 important miRNAs (Figures 2E, F). These miRNAs were analyzed by multivariate regression, and finally 12 important miRNAs were obtained. Hsa-

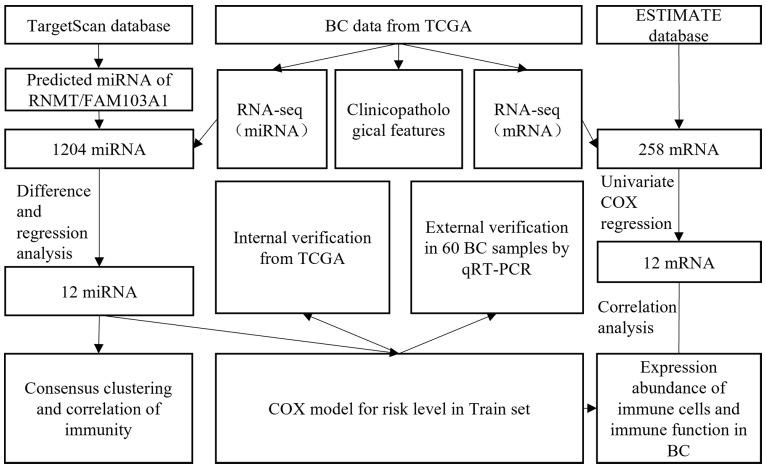


FIGURE 1
Flow chart of the present study.

miR-21-3p, hsa-miR-340-5p, hsa-miR-4501, hsa-miR-877-5p, hsa-miR-4675, hsa-miR-483-3p, and hsa-miR-6844 were screened as independent prognostic risk factors. Hsa-miR-629-3p was independent prognostic protective factor (Figure 2D).

Consensus clustering identified three clusters of BC patients with related to immunity

Based on the miRNA data of 511 BC patients, consensus clustering was performed on gene expression profiles by the ConsensusClusterPlus package to classify tumor tissues into 3 molecular subtypes (Figure 3A). Among these three molecular subtypes, survival was performed. On analysis, cluster 3 had the lowest survival rate (Figure 3B). The three types were compared in the expression of immune cell infiltration (Figures 3C–E), and the expression of the current mainstream immune checkpoints was compared between groups (Figures 3L–Q). Except that there is no difference in PD-L1 between BC patients and normal people, others were different and the difference in expression of PD-1, CTLA-4, TIM-3, and TIGIT were more significant in cluster three patients (Figures 3F–K).

Prognostic prediction ability and internal validation of the model constructed by 12 miRNAs

Kaplan-Meier univariate survival analysis was performed on the 12 miRNAs to study the effect of each factor on survival time. There was a significant correlation with patient outcomes (Figures 4A, B). The Cox model was constructed with these 12 miRNAs, which were divided into high- and low-risk groups. The calculating formula of risk score is:

$$Risk\ \ score = \sum_{i=1}^{12} Coef_i * x_i,$$

i = {1, 2, 3,…12}. $Coef_i$ means the coefficients of miRNAs in risk level model, x_i is the expression values of the miRNAs. First, principal component analysis (PCA) showed that the repeatability within the group was relatively good, the sample data were very similar, and there was a good difference between the groups (Figure 4C). Subsequently, to evaluate the ability of the model to predict prognosis, a receiver operating characteristic (ROC) curve was constructed. The results showed that the area under the curves (AUCs) of the 1-, 3-, 5-, and 10-year survival rates were 0.711, 0.694, 0.706, and 0.797, respectively. AUC > 0.7 indicated that the model had good accuracy in predicting 1-year, 5-year, and 10-year survival rates (Figure 4H). Kaplan-Meier survival analysis showed that the prognosis of the high-risk group was significantly worse than that of the low-risk group (P<0.001) (Figure 4J). According to the risk factor association diagram, the predicted risk value for each patient is presented in ascending order. The two groups were distinguished by the median risk value: the low-risk group (blue) and high-risk group (red) (Figures 4D, E). The relationship between patients sorted by predicted risk value and survival time showed that the survival time of low-risk groups was slightly longer than that of high-risk groups. Among them, blue dots represent living patients and red dots represent dead patients. The number of deaths in the high-risk group was significantly higher than that in the low-risk group (Figures 4F, G). In the validation set, the results showed that the AUCs of the 1-, 3-, 5, and 10-year survival rates were 0.643, 0.690, 0.627 and 0.578, respectively. Kaplan-Meier survival analysis also showed that the prognosis of the high-risk group was significantly worse (P=0.001) (Figures 4I, K).

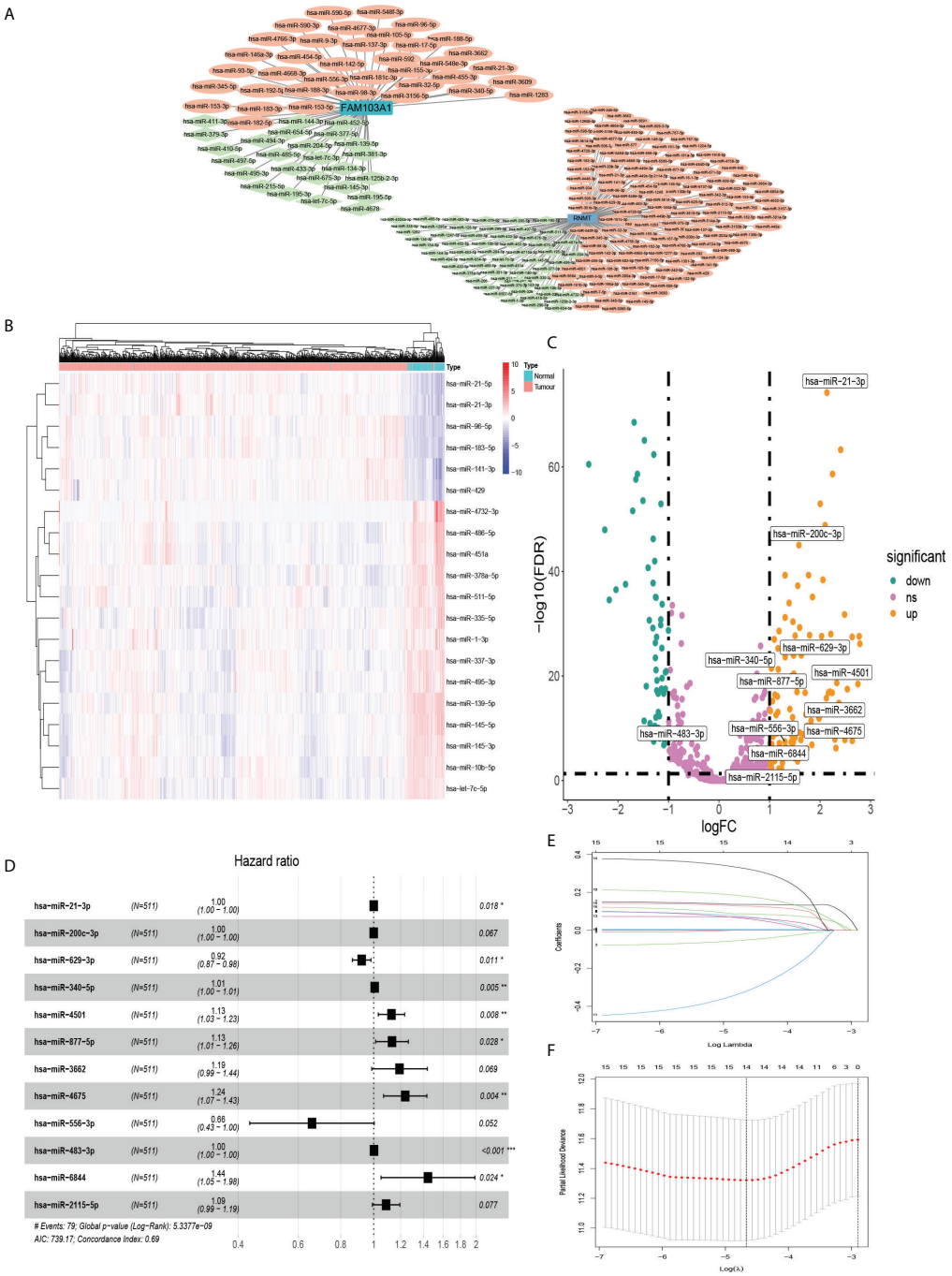


FIGURE 2
MiRNAs that were differentially expressed in BC and affected prognosis. **(A)** Co-expression network of RNMT/FAM103A1 and their upstream miRNAs. **(B)** Among the 201 upstream miRNAs of RNMT/FAM103A1 that can genetically modify by m7G, a heatmap of the top 20 most differentially expressed miRNAs in BC patients and healthy patients was drawn. Blue represents the healthy patients, and orange represents BC patients. **(C)** Volcano plot of miRNAs. **(D)** After multivariate Cox regression, a forest plot of 12 miRNAs. **(E, F)** Plots for LASSO regression coefficients. *p < 0.05, **p < 0.01, ***p < 0.001.

External verification of the model with 12 miRNAs in BC by qRT-PCR

To verify the model we constructed, we used qRT-PCR to evaluate the expression of 12 miRNAs. We found that the expression levels of 8 miRNA in BC tissues were significantly higher than that in normal tissues. However, there was no significant difference in has-miR-4675, has-miR-556-3p, has-miR-483-3p, and has-miR-2115-5p in BC tissues compared with normal tissues (Figure 5A). According to the formula of the

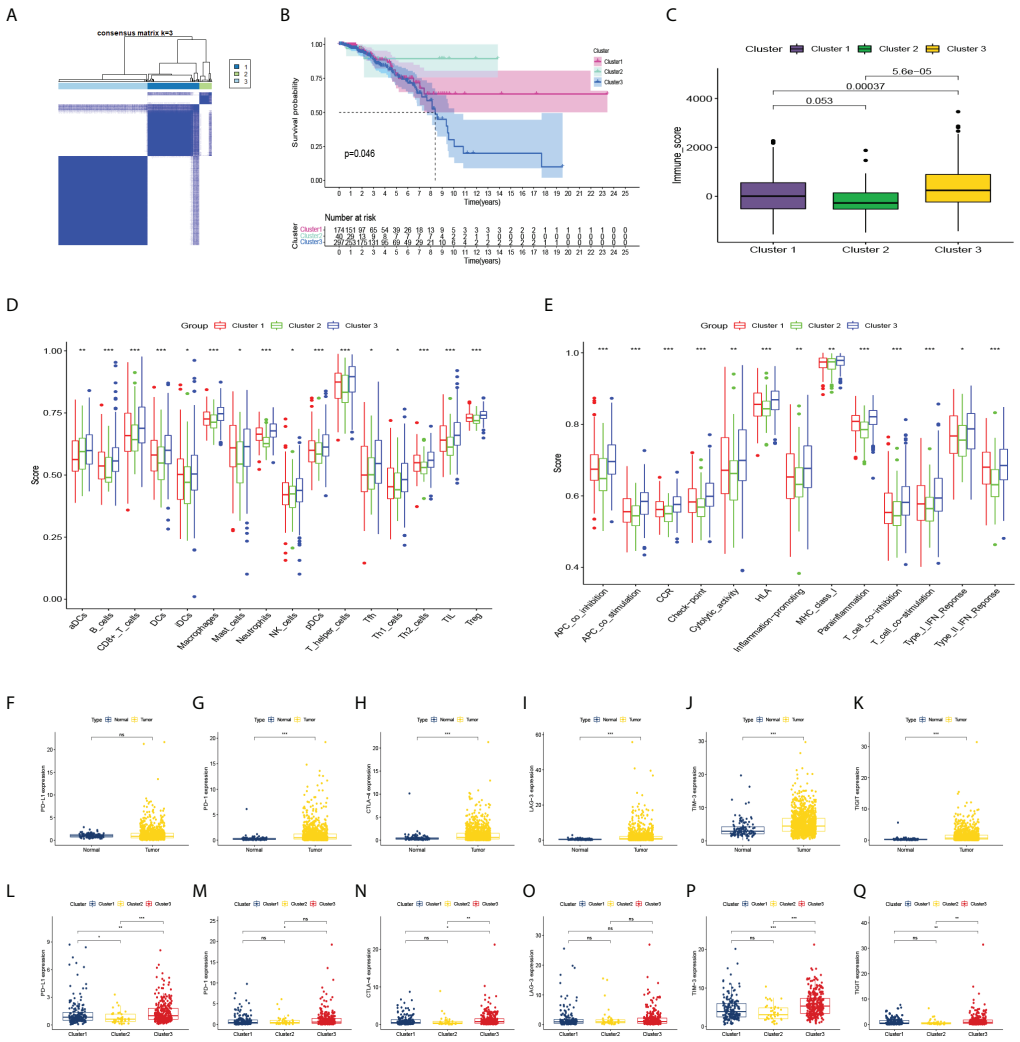


FIGURE 3
Consensus clustering and correlation of immunity. **(A)** Consensus clustering matrix for k = 3. **(B–E)** Kaplan-Meier curves, immune score, and the infiltrating levels of immune cell and immune function types of 3 clusters. **(F–K)** The expression levels of PD-L1, PD-1, CTLA-4, LAG-3, TIM-3, TIGIT in BC patients and normal people, and **(L–Q)** in 3 clusters. ns, not significant. *p < 0.05, **p < 0.01, ***p < 0.001.

COX model, with 0.975271 as the threshold, among the 60 samples, 23 were in the high-risk group and 37 were in the low-risk group. The AUCs of the model with 12 miRNAs expressions of qRT-PCR in 1-, 3-, 5-, and 10-year survival rates were 0.832, 0.802, 0.714, 0.839, respectively (Figure 5C). Kaplan-Meier survival analysis showed that the prognosis of high-risk group was significantly worse than the low-risk group (P<0.001), consistent with the results of the training set (Figure 5B).

Construction of risk-level model and clinicopathological features

To evaluate the predictive efficacy of the risk-level model in actual clinical practice, we combined the risk level with clinicopathological characteristics (including age, subtype, pathologic stage, and TNM stage). Univariate and multivariate Cox regression analyses revealed that risk level was an independent prognostic risk factor (P<0.001) (Figures 6A, B). The nomogram of the model was drawn and analysed for patient No. 1 (Figure 6C). Sankey diagram and heatmaps for model and clinicopathological features were also drawn. Patients with cluster 3 were more in the high-risk group (Figures 6E, F). To further evaluate whether the prediction model was in line with the actual situation, a calibration curve of the prediction model was drawn. The abscissa of the graph represents the prediction probability, and the prediction model predicts the possibility of event occurrence. The vertical axis represents the actual probability of the actual event rate of the patients. The green line is the fitted line for predicting 1-year overall survival, the

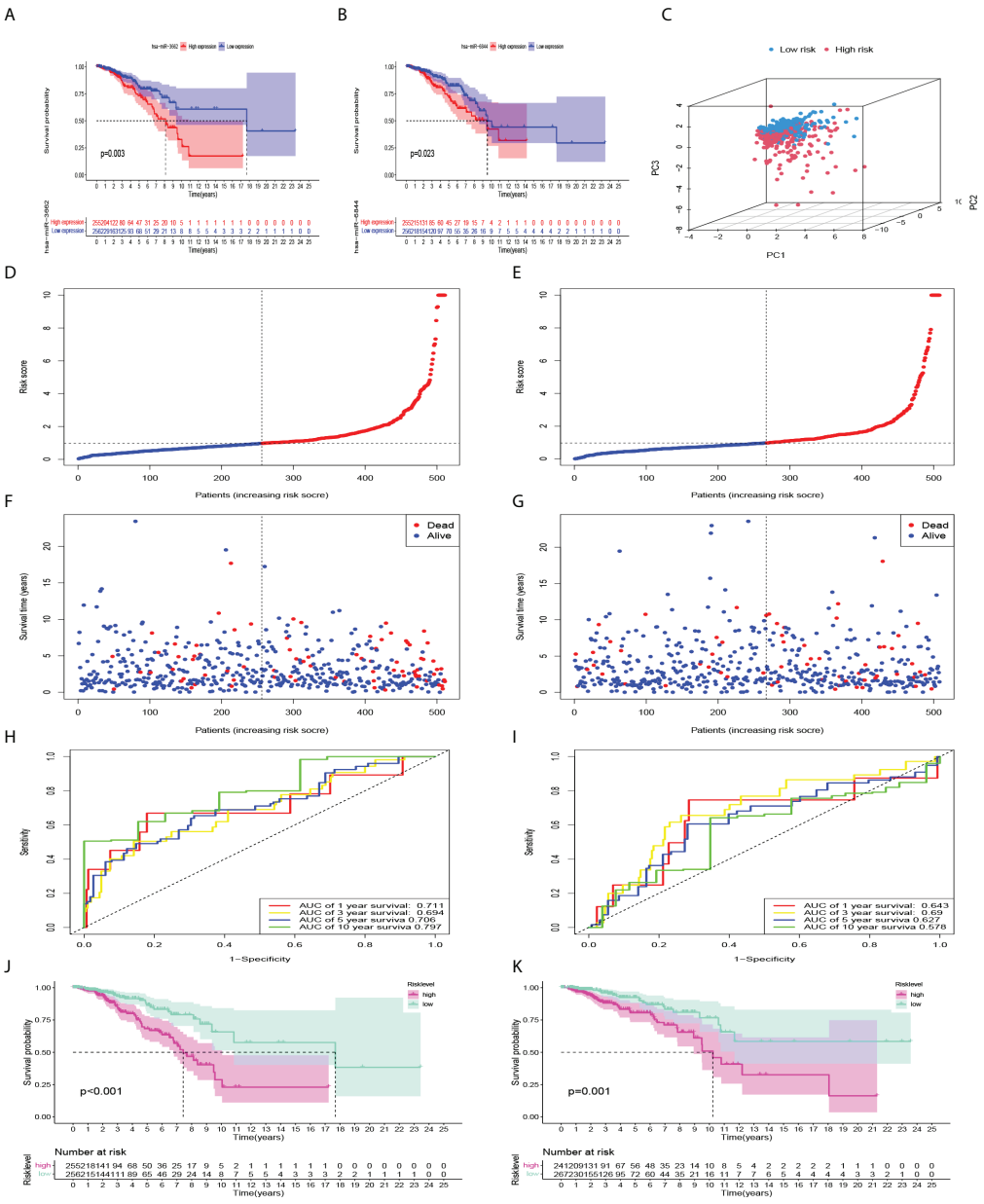


FIGURE 4
Internal verification and predictive efficiency of the Cox model consisting of 12 miRNAs. **(A, B)** Kaplan-Meier curves showed that high expression of miR-3662 and miR-6844 were significantly correlated with a poor prognosis. **(C)** PCA of the model. **(D, F)** Risk factor association map was made for each BC patient in train set, and **(E, G)** in validation set. **(H, J)** ROC curve and Kaplan-Meier curves of the model in train set, and **(I, K)** in validation set.

blue line for predicting 3-year overall survival, the red line for predicting 5-year overall survival, the orange line for predicting 10-year overall survival and the grey line is the reference line. The 1-, 3-, and 5-year three fitted lines almost completely coincide with the reference line, indicating that the predictive model has a high predictive efficacy. However, 10-year survival predictions suggested an underestimation of patient survival (Figure 6D). Importantly, the C-index indicates that the risk score was a very accurate indicator of the predictive ability of the model (Figure 6G). Therefore, we validated the risk model in clinical groups. The risk model was able to accurately estimate the survival rate of patients in age groups, luminal B and HER2 subtypes, early and late stages of tumor, and whether lymph nodes metastasis (Figure 6H).

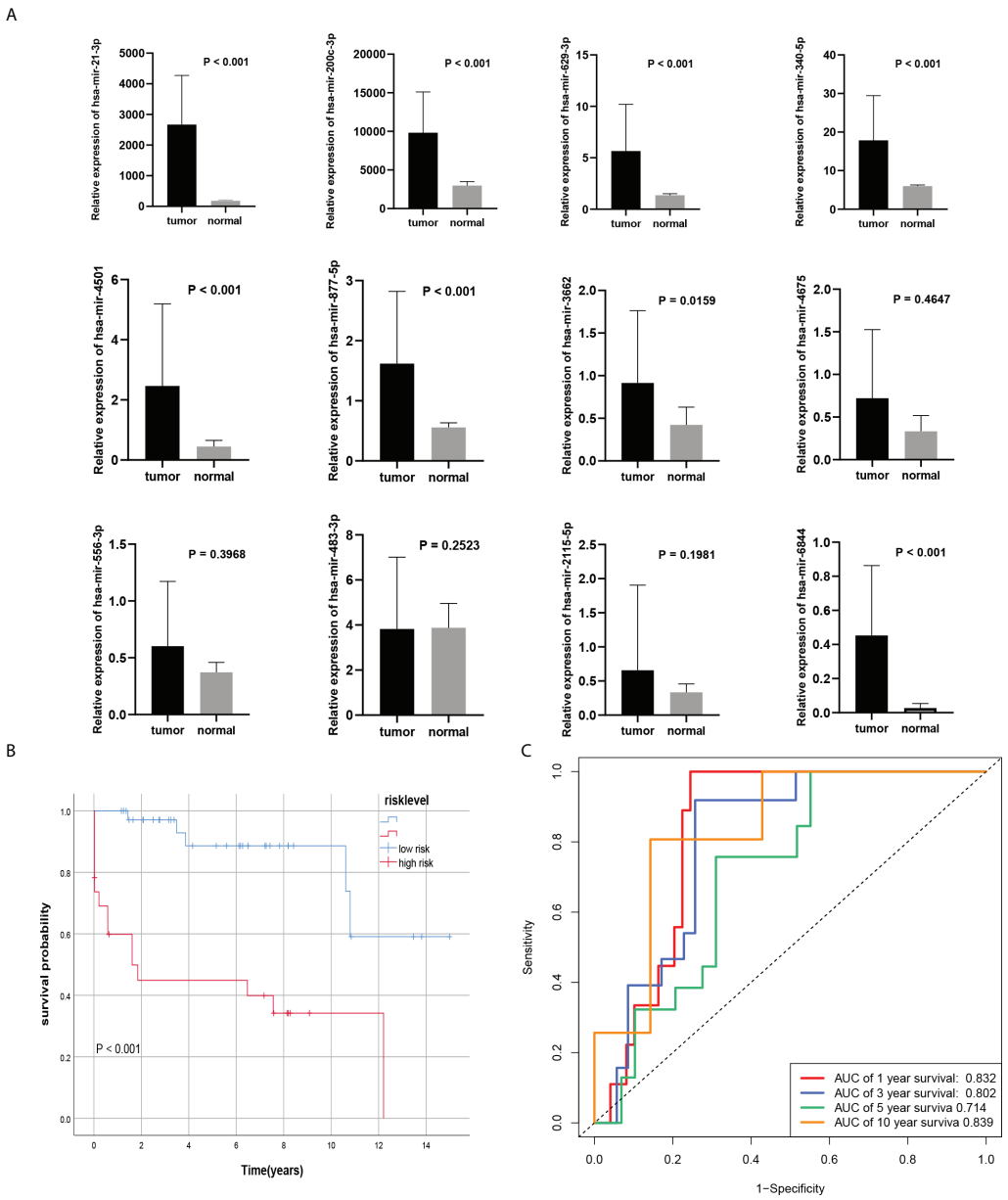


FIGURE 5
External verification of the Cox model. **(A)** 8 miRNAs were differentially expressed in BC tissues and normal tissues. **(B, C)** Kaplan-Meier curves and ROC curve of risk level model from own databases.

Correlation between risk-level models and immune infiltration

The single-sample gene set enrichment analysis (ssGSEA) algorithm was used to analyze the abundance of genes expressed by each immune cell and immune function signature in patients with BC. The heatmap of immune infiltration in BC patients revealed that the expression of T helper cells and major histocompatibility complex class 1 (MHC class I) was prominent in the tumor immune microenvironment (Figure 7A). Tumor-infiltrating lymphocytes (TILs) and plasmacytoid dendritic cells (pDCs) had the highest correlation in the BC immune microenvironment with an R value of 0.91; mast cells and activated dendritic cells (aDCs) were negatively correlated with an R value of -0.1 (Figure 7B). Immune checkpoints and T-cell co-stimulatory and co-inhibitory pathways showed the highest positive correlation with an R value of 0.96 (Figure 7C). In addition, immune infiltration was more pronounced in the

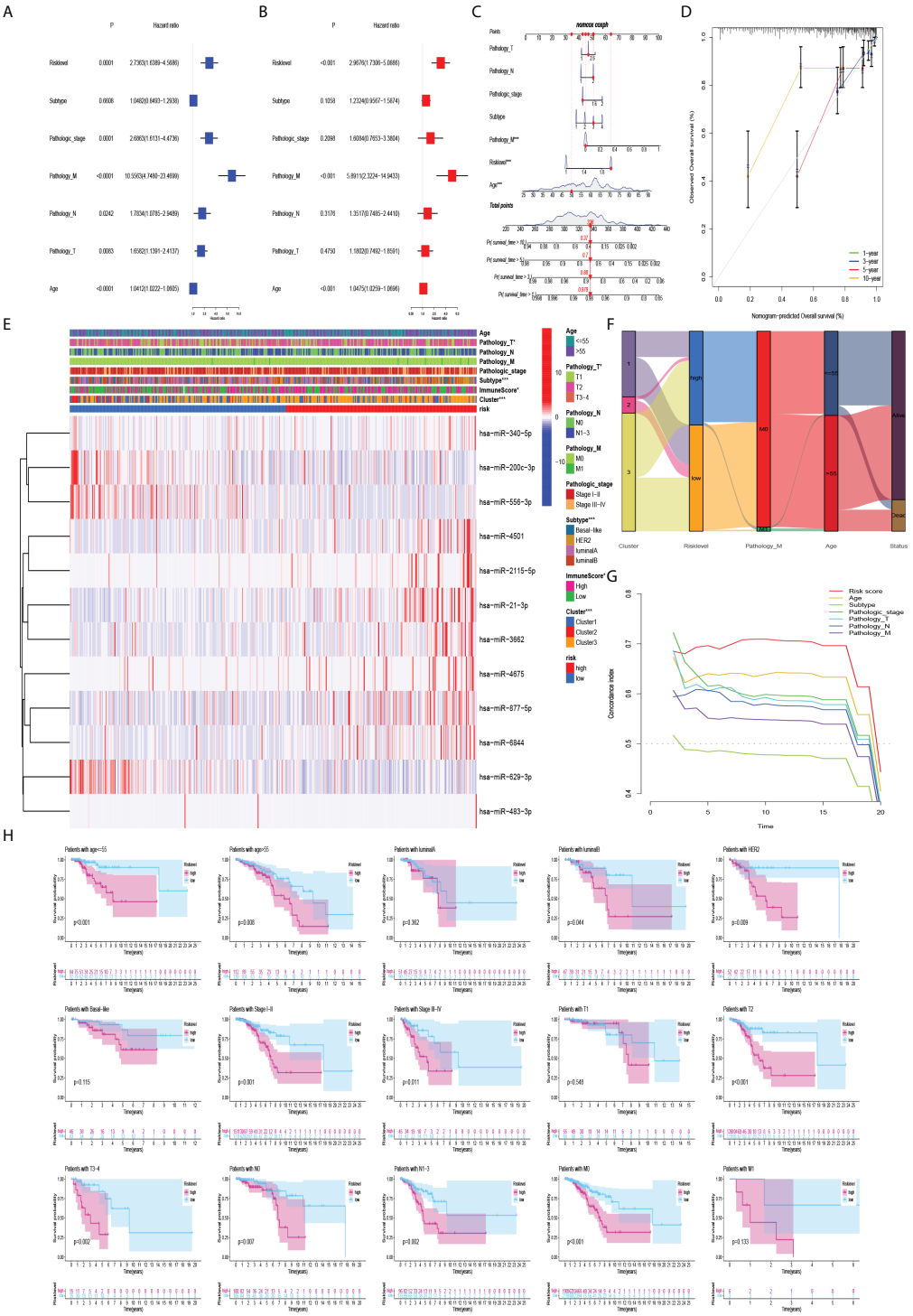


FIGURE 6
Association of model and clinicopathological features. **(A)** Univariate Cox regression analysis and **(B)** multivariate Cox regression analysis showed that risk level was an independent prognostic risk factor. **(C)** Nomogram of risk level and clinicopathological features for one of the patients. **(D)** Plot of the prediction model calibration curve. **(E–G)** Heatmaps, sankey diagram, and C-index for model and clinicopathological features. **(H)** Kaplan-Meier curves of the risk model in clinical groups.

high-risk group (Figure 7F). Interestingly, the expression of aDCs, B cells, dendritic cells (DCs), macrophages, neutrophils, NK cells, pDCs, T helper cells, follicular helper T cells (Tfh), helper T cells 1 (Th1 cells), helper T cells 2 (Th2 cells), TILs, regulatory T cells (Treg), APC co-stimulatory and co-inhibitory pathways, chemokine receptors (CCR), immune checkpoints, cytolytic activity, human leukocyte antigen (HLA), inflammation promotion, MHC class I, T cell co-stimulatory and co-inhibitory pathways, and parainflammation were significantly lower in the low-risk group ($P<0.05$) (Figure 7D, E). It has been suggested that BC patients with high-risk level may be candidates for immunotherapy. Interestingly, patients in the high-risk group may be more suitable for anti-LAG-3 immunotherapy (Figures 7K–P). Although PD-1 expression was higher in the low-risk group, patients in the low-risk group may not benefit more than those in the high-risk group because there was no difference in the immunotherapy score analysis (Figures 7G–J, L).

mRNAs associated with immune infiltration in risk-level models

To further analyse the factors affecting the risk level at the mRNA level, we found 629 differentially expressed mRNAs in the high-risk and low-risk groups in the risk-level model. The 1454 differentially expressed mRNAs from the immune score groups of the ESTIMATE database (with the median as the cut-off value) were intersected with 629 differentially expressed mRNAs to yield 258 mRNAs associated with risk level and immune infiltration (Figure 8A). They were enriched in epidermis development, sarcomere, actin binding, and neuroactive ligand-receptor interaction signalling pathways (Figures 8C, E). Univariate Cox regression analysis revealed that the 12 mRNAs were associated with prognosis (Figure 8B). ADD3-AS1, IGLJ6, OLFM4, PCSK1, IGLV1-36 positively correlated with immune infiltration. SYT4 genes was negatively associated with immune infiltration and correlated with poor prognosis (Figure 8D).

Discussion

N 7-methylguanosine (m7G) is an essential modification of the positively charged 5′ end of mRNA in mammals that regulates mRNA export, translation, and splicing (25). Abnormal m7G modifications are closely related to the occurrence and development of various cancers (26–30). RNMT was identified as a methyltransferase that installs a subset of m7G within mRNA and affects its translational capacity (30). In addition, FAM103A1 consists of an N-terminal RNMT activating domain and a C-terminal RNA-binding domain, which functions in the m7G methyltransferase complex with RNMT (18). It is now generally accepted that miRNAs play an important role in the occurrence and development of tumors, especially in epigenetic regulation, protein interactions, and RNA metabolism (31, 32). miRNAs are a group of highly conserved, single-stranded, short non-coding RNAs. They are a key regulator of mRNA expression in both normal and abnormal biological processes, including cancer (33). Dysregulated miRNA expression has also been implicated in cell survival and proliferation as well as in cell extravasation and metastasis (34). In BC, miRNAs represent an emerging group of molecules that play critical roles in disease development and are potential tools for improving treatment and impact diagnosis (35). Therapeutic strategies based on modulating the expression levels of miRNAs and identifying their targets are promising approaches for miRNA-based molecular therapy for BC (36). Detection of circulating miRNAs has also facilitated the formation of miRNA profiles in the blood of patients with BC, emphasizing that miRNAs are promising biomarkers for early disease screening, therapeutic targets, and prediction of prognosis (37). For example, miR-21 is associated with clinical stage, lymph node metastasis, and poor prognosis (38). High miR-21 expression is also associated with poor prognosis in Asian patients with BC (39). However, it remains unclear whether miRNA regulation exists upstream of the m7G methyltransferase complex RMNT/FAM103A1.

In our study, we obtained 12 important upstream miRNAs of m7G genes. Consensus clustering classified BC patients into 3 clusters. Patients with cluster 3 may benefit more from anti-PD-1, CTLA4, TIM-3, and TIGIT immunotherapy because patients with cluster 3 had higher expression of immune checkpoints compared to the other two groups. Subsequently, we found in the heatmap and Sankey diagram that patients with cluster 3 were mainly in the high-risk group. The high-risk group, like cluster 3, had higher expression of immune infiltrates. However, patients in the high-risk group benefited more from anti-LAG-3 immunotherapy. Which grouping method is more beneficial to BC patients still needs more comprehensive evaluation and further exploration. But to a certain extent, it can be shown that these 12 miRNAs have a certain hinting effect on immunotherapy. Kaplan–Meier survival analysis found that has-miR-3662 and has-miR-6844 were all highly expressed, suggesting a worse survival rate. Kaplan–Meier analysis also revealed a difference in survival probability between the high-risk and low-risk groups. The survival rate in the high-risk group was significantly lower than that in the low-risk group. In addition, the ROC curve showed that the AUC of the model was almost >0.7, which indicates more accurate prediction of prognosis. According to the risk factor association map, the number of deaths in the high-risk group was significantly higher than that in the low-risk group. Through PCA, the model can better distinguish between high-risk and low-risk groups. Therefore, this model could serve as a potential prognostic biomarker for BC. Li et al. found that the relative expression of miR-3662 in serum exosomes was significantly higher in BC patients than healthy controls, which was shown to be valuable biomarkers to monitor patient condition in the course of surgery

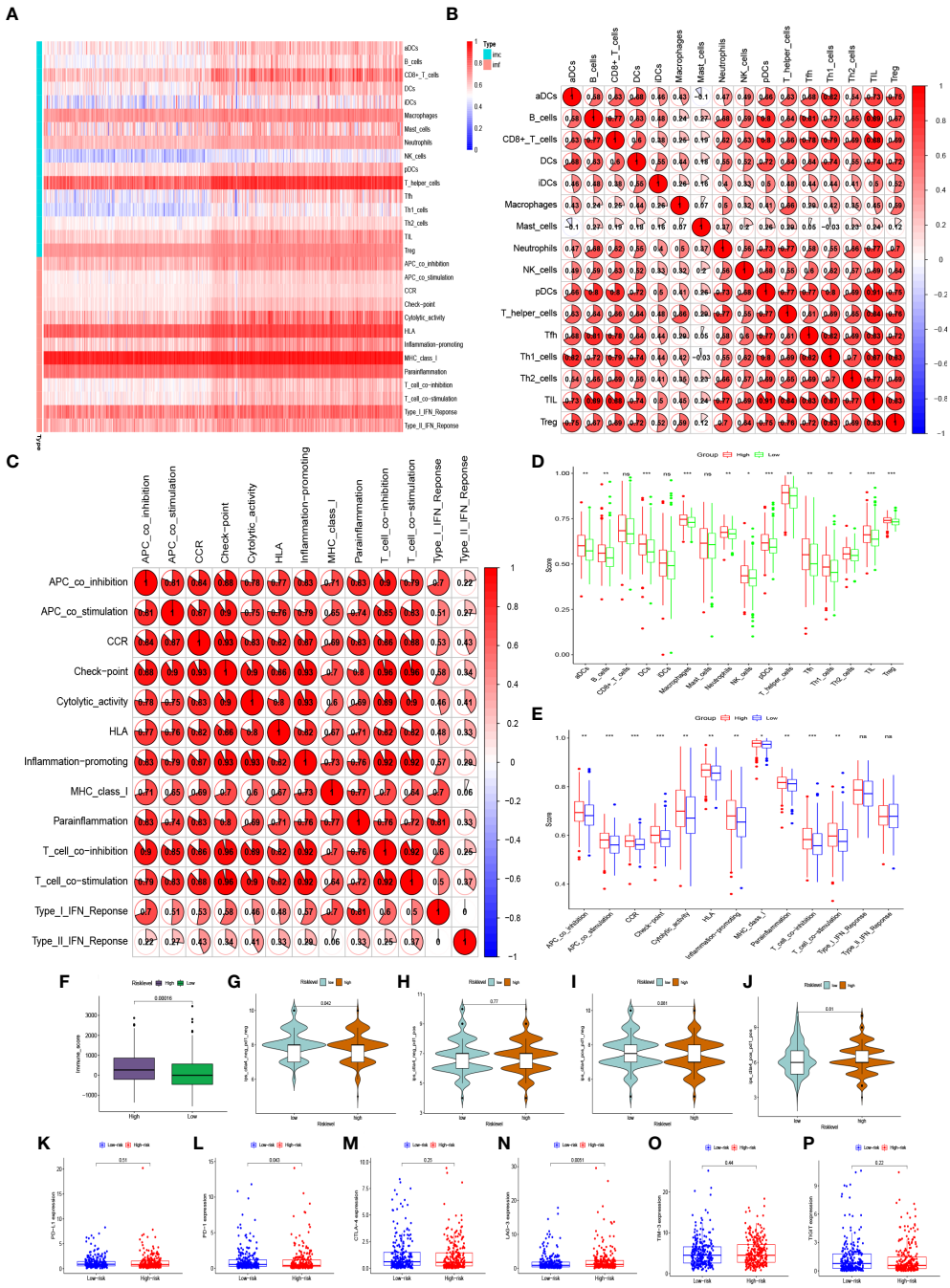


FIGURE 7
Abundant expressions of immune infiltration in BC patients. **(A)** Heatmap of expression abundances of 16 immune cells and 13 immune functions in BC patients. **(B)** Correlations between immune cells and **(C)** correlation between immune functions in BC patients. **(D)** Differences in immune cell abundance expression and **(E)** differences in abundance expression of immune function between high-risk and low-risk groups in BC patients. **(F, G–J)** Immune score and immunophenoscore of risk level groups. **(K–P)** The expression levels of PD-L1, PD-1, CTLA-4, LAG-3, TIM-3, TIGIT in high-risk and low-risk groups. ns, not significant. *p < 0.05, **p < 0.01, ***p < 0.001.

and chemotherapy (40). The data showed that miR-21-3p overexpression in BC was a hallmark of worse BC progression and it affected genes in pathways that drive breast cancer by down-regulating tumor suppressor genes (41). Curtaz et al. found the expression level of miR-340-5p was significantly correlated with the percentage of actively proliferating tumor cells (42). Elango et al. found, compared with patients with primary BC, the expression of miR-200-3p was decreased in BC

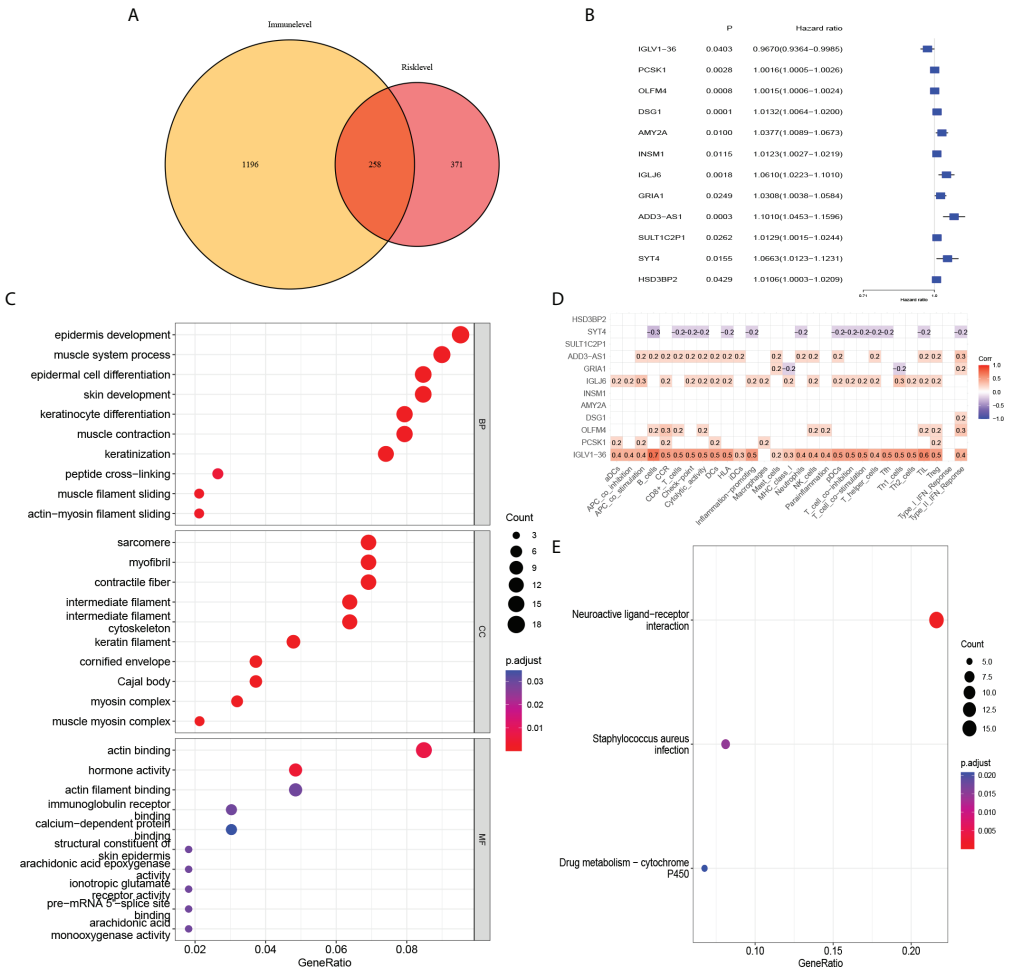


FIGURE 8
Association of risk level groups and mRNAs. **(A)** 258 mRNAs associated with risk level and immune infiltration were obtained. Subsequently, **(C)** GO enrichment analysis and **(E)** KEGG pathway enrichment analysis was performed. **(B)** By univariate Cox regression, 12 mRNAs were found to have effects on prognosis. **(D)** Correlation analysis of 12 mRNAs and immune infiltration.

patients with lymph node metastasis. This indicated that overexpression of hsa-miR-200-3p may inhibit BC progression and metastasis (43).In addition, eight miRNAs, miR-2115-5p, miR-483-3p, miR-6844, miR-4675, miR-877-5p, miR-4501, miR-629-3p, and miR-556-3p, have not yet been reported in PubMed for BC-related studies. Therefore, research in this area needs to be urgently conducted.

Furthermore, combining the risk level with clinicopathological characteristics (age, subtype, pathologic stage, TNM stage), we found that the risk level had a significant impact on prognosis ($P< 0.001$). The calibration curve showed that the predicted survival rates of the model at 1, 3, and 5 years were in good agreement with the actual situation. Interestingly, we found that the risk model had a more pronounced effect on prognosis in BC patients with HER2 subtype, so we collected paraffin samples from BC patients with HER2 subtype in our hospital for external validation. Interestingly, except that has-miR-4675, has-miR-556-3p, has-miR-483-3p, and has-miR-2115-5p were not statistically different between BC tissues and normal tissues, the expression of the other 8 miRNAs in BC tissues was significantly higher than that in normal tissues. In BC patients with HER2 subtype, patients in the high-risk group had a significantly worse prognosis than these in the low-risk group. The AUC of the model also indicated that the model had good accuracy in predicting 1-year, 5-year, and 10-year survival rates. Due to population differences, we found no differences in the expression of 4 miRNAs. In addition, because of the confounding of other molecular subtypes in the TCGA database, the power of the model to predict prognostic accuracy is lower than in our own database. And we found that the model was also very reliable in the internal validation set. Previous studies have found that miRNAs in the immune system play key roles in the developmental fate of lymphocytes and in innate and adaptive immunity (44, 45).

Abnormal expression of certain miRNAs in BC may be related to immune system dysfunction. For example, the BC-derived exosomal lncRNA SNHG16 can promote the expression of miR16-5p by targeting the TGF-β1/SMAD5 pathway, thereby inducing the differentiation of CD73+ γδ1 Treg cells (46). The elevated expression level of miR-182 in the tumor tissues of BC patients may exert an immunosuppressive effect by inducing Treg cell differentiation (47). Therefore, to explore whether miRNAs upstream of the m7G gene in the model are associated with immune infiltration in the BC tumor microenvironment, we first analyzed the expression abundance of immune cells and immune functions in BC patients. In addition to iDCs, NK cells, and Th1 cells with low expression, other immune cells and immune functions, especially T helper cells and MHC class I, are highly expressed in BC. Li et al. found that blocking TGF-β signalling in CD4+ Th cells can trigger vasculature reorganization, leading to tumor hypoxia and BC cell death. Thus, blocking TGF-β signalling in T helper cells could elicit an effective cancer defense response, thus offering the potential for BC immunotherapy (48). Therefore, in our BC samples with high expression of 11 miRNAs, reducing miRNA expression in BC patients may benefit not only targeted therapy drugs, but also immunotherapy. Furthermore, we analyzed the differences in BC immune infiltration between the high- and low-risk groups. Surprisingly, except for CD8+ T cells, iDCs, Mast cells, type I interferon response, and type II interferon response, other expressions were significantly higher in the high-risk group. Kaplan- Meier analysis of the model indicated that high immune infiltration in the high-risk group was associated with poor prognosis. This further may indicate that the high-risk group is more suitable for immunotherapy than low-risk groups, especially in HER2 BC subtypes. Therefore, the 12 important miRNAs identified in this study may provide new targets for BC immunotherapy.

To further characterize the risk-level model, we subsequently obtained 258 mRNAs based on the differences between the high and low immune component groups in the ESTIMATE database of BC patients. As with our immune-related model results, mRNAs representing these models were analyzed using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. These genes were enriched in epidermis development, sarcomere, actin binding, and neuroactive ligand-receptor interaction signalling pathways. Among the 12 mRNAs associated with prognosis, IGLV1-36 was positively correlated with immune infiltration and were associated with good prognosis. This suggest that this gene may be an immune-related gene and its high expression may inhibit tumor progression in BC patients. Existing studies report that IGLV1-36 is suitable for confirming the diagnosis of POEMS syndrome (49). While its role in BC has not been reported, this may be another research target for immune infiltration. However, further verification and in-depth research are needed to determine whether these markers can become targets for new immunotherapy and the related methylation mechanisms.

Materials and methods

Breast cancer data and acquisition of upstream miRNA of m7G methyltransferase RNMT/FAM103A1

We obtained transcriptome mRNA-seq, miRNA-seq gene expression data, and corresponding clinical information of BC patients from TCGA database (https://portal.gdc.cancer.gov/), including 1019 tumor samples and 103 healthy samples, and downloaded human miRNA target gene files from the TargetScan database (50). Then, using Perl software, we obtained mRNA, miRNA gene expression matrix, and clinical data of BC patients according to age, subtype, pathologic stage, TMN stage, survival time, and survival status. Using R language software, 2638 miRNAs of target genes RNMT and FAM103A1 were obtained and intersected with 2217 miRNAs co-expressed in BC patients and healthy patients to obtain 1204 upstream miRNAs of m7G methyltransferase target gene RNMT/FAM103A1 in BC patients and healthy persons.

Construction of a risk-level model for m7G-related miRNAs

Using the R software package (Limma package and edgeR package), 1204 miRNAs of m7G methyltransferase target gene RNMT/FAM103A1 in BC patients (n = 1019) and healthy controls (n = 103) were analyzed with logFC ≥ |1|, FDR < 0.05 for differential expression analysis. We randomly divided the dataset into two groups according to 0.5: training set and validation set. Then, the survival package was used to conduct univariate Cox regression analysis on 201 miRNAs, where the P value is set below 0.05, and 16 miRNAs related to prognosis were obtained. To further exclude unimportant variables and obtain less meaningful variables, we used the glmnet package to perform least absolute shrinkage and selection operator (LASSO) regression to screen out 14 important miRNAs and then carried out multivariate analysis and finally constructed a Cox model with 12 miRNAs. The high- and low-risk groups were divided into two groups. Consensus clustering identified three clusters of BC patients and explored the correlation between cluster and immunity. The Rtsne package was used to perform PCA analysis, the survival package to draw the Kaplan-Meier survival curve, and the timeROC package to evaluate the ability of the model to predict prognosis. In addition, we constructed a risk factor association map.

miRNA extraction and quantitative real-time polymerase chain reaction

We totally collected 60 BC samples from patients and 30 normal breast tissues who underwent surgical treatments in Shengjing Hospital of China Medical University from 2007 to 2021. In these 60 BC samples, they were HR-, HER2+++ on immunohistochemistry or HR-, HER2++ on immunohistochemistry with fluorescence *in situ* hybridization (FISH) indicating that HER2 was amplified. Formalin fixation and paraffin embedding (FFPE) were to preserve the specimens. The study was approved by the hospital institutional ethics review committee. For evaluating the expression levels of 12 miRNA, we deparaffinized these specimens using xylene and ethanol. According to the manufacturer's protocol, we extracted total RNA (including miRNAs) from FFPE tissue samples using TRIzol (Thermo Fisher Scientific, US), and cDNA synthesis was carrying out by using Mir-X miRNA qRT-PCR SYBR Kits (Takara Bio Inc., Kusatsu, Japan). Then, we performed real-time PCR reaction using One Step TB Green® PrimeScript™ RT-PCR Kit (Perfect Real Time) (Takara Bio Inc., Kusatsu, Japan) on The LightCycler 480 Real-Time PCR System. 12 miRNAs expression levels were calculated by the $2^{-\Delta\Delta Ct}$ method and the cycle threshold (CT) values of miRNAs were normalized to the level of U6 as internal reference. Primers sequences used in our study were shown in table (Supplementary Table S2).

Nomogram of risk level and clinicopathological features

The Cox model with risk level and clinicopathological characteristics (including age, subtype, pathologic stage, and TNM stage) was built using the survival package of the R language software. Through univariate and multivariate Cox regression, the impact of risk level on prognosis in clinical practice was evaluated. We then used the rms package to draw a nomogram for the Cox model, selected the fourth BC patient in the file to draw a nomogram, and used the calibration function to draw a calibration curve.

Correlation of differentially expressed mRNAs with immune infiltration in risk-level models

Differential analysis between the high-risk and low-risk groups of BC patients was performed using R software packages (Limma package and edgeR package). In total, 629 differentially expressed mRNAs were identified. Immune score data of BC patients were downloaded from the ESTIMATE database (https://bioinformatics.mdanderson.org/estimate/), which was divided into high and low immune component groups with the median as the cut-off. A total of 1454 differentially expressed mRNAs in the two immune component groups were intersected with the differentially expressed mRNAs in the risk level groups to obtain 258 differentially expressed mRNAs associated with immune infiltration and risk level in BC patients ($logFC \geq |1|$, $FDR < 0.05$). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed on 258 differentially expressed mRNAs. Using univariate Cox regression, we obtained 12 differentially expressed mRNAs that were associated with immune infiltration and risk levels in patients with BC which had an impact on prognosis. The ssGSEA algorithm in the GSVA package was used to calculate the abundance of genes expressed by each immune cell and immune function in BC patients and to draw a heat map and conduct correlation analysis between immune cells and immune functions. The expression of immune cells and immune function were explored in the high-risk and low-risk groups. Finally, we explored the association of immune cells and immune function with 12 differentially expressed mRNAs that were associated with immune infiltration and risk level in patients with BC and had an impact on prognosis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the institutional ethics review committee of Shengjing Hospital of China Medical University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

All authors contributed to the work presented in this paper. Conceptualization, WZ and ZW; TCGA resources, visualization, and analysis, SZ; writing—original draft preparation, WZ and S Z; writing—editing, ZW; supervision, ZW; project administration, WZ and ZW; funding acquisition, ZW. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81601692) and Technology Research from the Department of Education of Liaoning Province (No. JCZR2020013), and 345 Talent Project of Shengjing Hospital of China Medical University.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.929363/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Mediratta K, El-Sahli S, D'Costa V, Wang L. Current progresses and challenges of immunotherapy in triple-negative breast cancer. *Cancers (Basel)* (2020) 12(12):3529. doi: 10.3390/cancers12123529

3. Adams S, Loi S, Toppmeyer D, Cescon DW, De Laurentiis M, Nanda R, et al. Pembrolizumab monotherapy for previously untreated, pd-L1-Positive, metastatic triple-negative breast cancer: Cohort b of the phase ii keynote-086 study. *Ann Oncol* (2019) 30(3):405–11. doi: 10.1093/annonc/mdy518

4. Cowling VH. Regulation of mrna cap methylation. *Biochem J* (2009) 425 (2):295–302. doi: 10.1042/bj20091352

5. Shuman S. What messenger rna capping tells us about eukaryotic evolution. *Nat Rev Mol Cell Biol* (2002) 3(8):619–25. doi: 10.1038/nrm880

6. Bentley DL. Rules of engagement: Co-transcriptional recruitment of pre-mrna processing factors. *Curr Opin Cell Biol* (2005) 17(3):251–6. doi: 10.1016/j.ceb.2005.04.006

7. Moore MJ, Proudfoot NJ. Pre-mrna processing reaches back to transcription and ahead to translation. *Cell* (2009) 136(4):688–700. doi: 10.1016/j.cell.2009.02.001

8. Liu H, Kiledjian M. Decapping the message: A beginning or an end. *Biochem Soc Trans* (2006) 34(Pt 1):35–8. doi: 10.1042/bst20060035

9. Furuichi Y, Shatkin AJ. Viral and cellular mrna capping: Past and prospects. *Adv Virus Res* (2000) 55:135–84. doi: 10.1016/s0065-3527(00)55003-9

10. Chu C, Shatkin AJ. Apoptosis and autophagy induction in mammalian cells by small interfering rna knockdown of mrna capping enzymes. *Mol Cell Biol* (2008) 28(19):5829–36. doi: 10.1128/mcb.00021-08

11. Cowling VH. Enhanced mrna cap methylation increases cyclin D1 expression and promotes cell transformation. *Oncogene* (2010) 29(6):930–6. doi: 10.1038/onc.2009.368

12. Shafer B, Chu C, Shatkin AJ. Human mrna cap methyltransferase: Alternative nuclear localization signal motifs ensure nuclear localization required for viability. *Mol Cell Biol* (2005) 25(7):2644–9. doi: 10.1128/mcb.25.7.2644-2649.2005

13. Dunn S, Lombardi O, Lukoszek R, Cowling VH. Oncogenic Pik3ca mutations increase dependency on the mrna cap methyltransferase, rnmt, in breast cancer cells. *Open Biol* (2019) 9(4):190052. doi: 10.1098/rsob.190052

14. Cole MD, Cowling VH. Specific regulation of mrna cap methylation by the c-myc and E2f1 transcription factors. *Oncogene* (2009) 28(9):1169–75. doi: 10.1038/onc.2008.463

15. Inesta-Vaquera F, Chaugule VK, Galloway A, Chandler L, Rojas-Fernandez A, Weidlich S, et al. Dhx15 regulates Cmtr1-dependent gene expression and cell proliferation. *Life Sci Alliance* (2018) 1(3):e201800092. doi: 10.26508/lsa.201800092

16. Posternak V, Ung MH, Cheng C, Cole MD. Myc mediates mrna cap methylation of canonical Wnt/B-catenin signaling transcripts by recruiting Cdk7 and rna methyltransferase. *Mol Cancer Res* (2017) 15(2):213–24. doi: 10.1158/1541-7786.Mcr-16-0247

17. Aregger M, Kaskar A, Varshney D, Fernandez-Sanchez ME, Inesta-Vaquera FA, Weidlich S, et al. Cdk1-cyclin B1 activates rnmt, coordinating mrna cap methylation with G1 phase transcription. *Mol Cell* (2016) 61(5):734–46. doi: 10.1016/j.molcel.2016.02.008

18. Gonatopoulos-Pournatzis T, Dunn S, Bounds R, Cowling VH. Ram/Fam103a1 is required for mrna cap methylation. *Mol Cell* (2011) 44(4):585–96. doi: 10.1016/j.molcel.2011.08.041

19. Danaher P, Warren S, Lu R, Samayoa J, Sullivan A, Pekker I, et al. Pan-cancer adaptive immune resistance as defined by the tumor inflammation signature (Tis): Results from the cancer genome atlas (Tcga). *J Immunother Cancer* (2018) 6 (1):63. doi: 10.1186/s40425-018-0367-1

20. Toor SM, Sasidharan Nair V, Decock J, Elkord E. Immune checkpoints in the tumor microenvironment. *Semin Cancer Biol* (2020) 65:1–12. doi: 10.1016/j.semcancer.2019.06.021

21. Adams S, Gatti-Mays ME, Kalinsky K, Korde LA, Sharon E, Amiri-Kordestani L, et al. Current landscape of immunotherapy in breast cancer: A review. *JAMA Oncol* (2019) 5(8):1205–14. doi: 10.1001/jamaoncol.2018.7147

22. Dugo M, Huang X, Iorio MV, Cataldo A, Tagliabue E, Daidone MG, et al. Microrna Co-expression patterns unravel the relevance of extra cellular matrix and immunity in breast cancer. *Breast* (2018) 39:46–52. doi: 10.1016/j.breast.2018.03.008

23. Cortes J, Cescon DW, Rugo HS, Nowecki Z, Im SA, Yusof MM, et al. Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for previously untreated locally recurrent inoperable or metastatic triple-negative breast cancer (Keynote-355): A randomised, placebo-controlled, double-blind, phase 3 clinical trial. *Lancet* (2020) 396(10265):1817–28. doi: 10.1016/s0140-6736(20)32531-9

24. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discovery* (2019) 18 (3):197–218. doi: 10.1038/s41573-018-0007-y

25. Zhang LS, Liu C, Ma H, Dai Q, Sun HL, Luo G, et al. Transcriptome-wide mapping of internal N(7)-methylguanosine methylome in mammalian mrna. *Mol Cell* (2019) 74(6):1304–16.e8. doi: 10.1016/j.molcel.2019.03.036

26. Dai Z, Liu H, Liao J, Huang C, Ren X, Zhu W, et al. N(7)-methylguanosine trna modification enhances oncogenic mrna translation and promotes intrahepatic cholangiocarcinoma progression. *Mol Cell* (2021) 81(16):3339–55.e8. doi: 10.1016/j.molcel.2021.07.003

27. Ma J, Han H, Huang Y, Yang C, Zheng S, Cai T, et al. Mettl1/Wdr4-mediated M(7)G trna modifications and M(7)G codon usage promote mrna translation and lung cancer progression. *Mol Ther* (2021) 29(12):3422–35. doi: 10.1016/j.ymthe.2021.08.005

28. Liu Y, Zhang Y, Chi Q, Wang Z, Sun B. Methyltransferase-like 1 (Mettl1) served as a tumor suppressor in colon cancer by activating 7-methyguanosine (M7g) regulated let-7e Mirna/Hmga2 axis. *Life Sci* (2020) 249:117480. doi: 10.1016/j.lfs.2020.117480

29. Ying X, Liu B, Yuan Z, Huang Y, Chen C, Jiang X, et al. Mettl1-M(7) G-Egfr/Efemp1 axis promotes the bladder cancer development. *Clin Transl Med* (2021) 11 (12):e675. doi: 10.1002/ctm2.675

30. Galloway A, Kaskar A, Ditsova D, Atrih A, Yoshikawa H, Gomez-Moreira C, et al. Upregulation of rna cap methyltransferase rnmt drives ribosome biogenesis during T cell activation. *Nucleic Acids Res* (2021) 49(12):6722–38. doi: 10.1093/nar/gkab465

31. Okugawa Y, Grady WM, Goel A. Epigenetic alterations in colorectal cancer: Emerging biomarkers. *Gastroenterology* (2015) 149(5):1204–25.e12. doi: 10.1053/j.gastro.2015.07.011

32. Bartel DP. Micrornas: Genomics, biogenesis, mechanism, and function. *Cell* (2004) 116(2):281–97. doi: 10.1016/s0092-8674(04)00045-5

33. Mendell JT. Tumors line up for a letdown. *Nat Genet* (2009) 41(7):768–9. doi: 10.1038/ng0709-768

34. Valastyan S, Weinberg RA. Micrornas: Crucial multi-tasking components in the complex circuitry of tumor metastasis. *Cell Cycle* (2009) 8(21):3506–12. doi: 10.4161/cc.8.21.9802

35. Serpico D, Molino L, Di Cosimo S. Micrornas in breast cancer development and treatment. *Cancer Treat Rev* (2014) 40(5):595–604. doi: 10.1016/j.ctrv.2013.11.002

36. Iorio MV, Croce CM. Microrna dysregulation in cancer: Diagnostics, monitoring and therapeutics. a comprehensive review. *EMBO Mol Med* (2012) 4 (3):143–59. doi: 10.1002/emmm.201100209

37. Di Cosimo S, Appierto V, Pizzamiglio S, Tiberio P, Iorio MV, Hilbers F, et al. Plasma mirna levels for predicting therapeutic response to neoadjuvant treatment in Her2-positive breast cancer: Results from the neoaltto trial. *Clin Cancer Res* (2019) 25(13):3887–95. doi: 10.1158/1078-0432.Ccr-18-2507

38. Yan LX, Huang XF, Shao Q, Huang MY, Deng L, Wu QL, et al. Microrna mir-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *Rna* (2008) 14(11):2348–60. doi: 10.1261/rna.1034808

39. Wang Y, Zhang Y, Pan C, Ma F, Zhang S. Prediction of poor prognosis in breast cancer patients based on microrna-21 expression: A meta-analysis. *PloS One* (2015) 10(2):e0118647. doi: 10.1371/journal.pone.0118647

40. Li S, Zhang M, Xu F, Wang Y, Leng D. Detection significance of mir-3662, mir-146a, and mir-1290 in serum exosomes of breast cancer patients. *J Cancer Res Ther* (2021) 17(3):749–55. doi: 10.4103/jcrt.jcrt_280_21

41. Amirfallah A, Knutsdottir H, Arason A, Hilmarsdottir B, Johannsson OT, Agnarsson BA, et al. Hsa-Mir-21-3p associates with breast cancer patient survival and targets genes in tumor suppressive pathways. *PloS One* (2021) 16(11): e0260327. doi: 10.1371/journal.pone.0260327

42. Curtaz CJ, Reifschläger L, Strähle L, Feldheim J, Feldheim JJ, Schmitt C, et al. Analysis of micrornas in exosomes of breast cancer patients in search of molecular prognostic factors in brain metastases. *Int J Mol Sci* (2022) 23(7):3683. doi: 10.3390/ijms23073683

43. Elango R, Alsaleh KA, Vishnubalaji R, Manikandan M, Ali AM, Abd El-Aziz N, et al. Microrna expression profiling on paired primary and lymph node metastatic breast cancer revealed distinct microrna profile associated with lnm. *Front Oncol* (2020) 10:756. doi: 10.3389/fonc.2020.00756

44. O'Connell RM, Kahn D, Gibson WS, Round JL, Scholz RL, Chaudhuri AA, et al. Microrna-155 promotes autoimmune inflammation by enhancing inflammatory T cell development. *Immunity* (2010) 33(4):607–19. doi: 10.1016/j.immuni.2010.09.009

45. Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD. Micrornas: New regulators of immune cell development and function. *Nat Immunol* (2008) 9 (8):839–45. doi: 10.1038/ni.f.209

46. Ni C, Fang QQ, Chen WZ, Jiang JX, Jiang Z, Ye J, et al. Breast cancer-derived exosomes transmit lncrna Snhg16 to induce Cd73+Γδ1 treg cells. *Signal Transduct Tar Ther* (2020) 5(1):41. doi: 10.1038/s41392-020-0129-7

47. Soheilifar MH, Vaseghi H, Seif F, Ariana M, Ghorbanifar S, Habibi N, et al. Concomitant overexpression of mir-182-5p and mir-182-3p raises the possibility of il-17-Producing treg formation in breast cancer by targeting Cd3d, itk, Foxo1, and nfats: A meta-analysis and experimental study. *Cancer Sci* (2021) 112(2):589–603. doi: 10.1111/cas.14764

48. Li S, Liu M, Do MH, Chou C, Stamatiades EG, Nixon BG, et al. Cancer immunotherapy *Via* targeted tgf-B signalling blockade in T(H) cells. *Nature* (2020) 587(7832):121–5. doi: 10.1038/s41586-020-2850-3

49. Bender S, Javaugue V, Saintamand A, Ayala MV, Alizadeh M, Filloux M, et al. Immunoglobulin variable domain high-throughput sequencing reveals specific novel mutational patterns in poems syndrome. *Blood* (2020) 135 (20):1750–8. doi: 10.1182/blood.2019004197

50. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microrna target sites in mammalian mrnas. *Elife* (2015) 4:e05005. doi: 10.7554/eLife.05005

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fonc.2022.1080054

OPEN ACCESS

EDITED BY
Vittorio Gebbia,
University of Palermo, Italy

REVIEWED BY
Eva Valentina Klocker,
University Hospital Graz, Austria
Xinxin Hou,
Shanghai University of Traditional Chinese Medicine, China

*CORRESPONDENCE
Shaoqiang Cheng
shaoqiangcheng@Hotmail.com
Yuanxi Huang
rxwk@163.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

RECEIVED 25 October 2022
ACCEPTED 30 November 2022
PUBLISHED 04 January 2023

CITATION
Zhou Z, Zhang Y, Li Y, Jiang C, Wu Y, Shang L, Huang Y and Cheng S (2023) Metabolic syndrome is a risk factor for breast cancer patients receiving neoadjuvant chemotherapy: A case-control study.
Front. Oncol. 12:1080054.
doi: 10.3389/fonc.2022.1080054

Metabolic syndrome is a risk factor for breast cancer patients receiving neoadjuvant chemotherapy: A case-control study

Zhaoyue Zhou[1†], Yue Zhang[2†], Yue Li[1], Cong Jiang[1], Yang Wu[1], Lingmin Shang[1], Yuanxi Huang[1*] and Shaoqiang Cheng[1*]

[1]Department of Breast Surgery, Harbin Medical University Cancer Hospital, Harbin, China, [2]Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China

Purpose: To investigate the impact of metabolic syndrome (MetS) on pathologic complete response (pCR) and clinical outcomes in breast cancer (BC) patients who received neoadjuvant chemotherapy (NAC).

Methods: We analyzed 221 female BC patients at Harbin Medical University Cancer Hospital who received NAC and divided them into MetS and non-MetS groups according to National Cholesterol Education Program Adult Treatment Panel III (NCEP-ATP III) criteria to investigate the association between MetS and clinicopathological characteristics, pathologic response, and long-term survival and to observe the changes in metabolic parameters after NAC.

Results: A total of 53 (24.0%) BC patients achieved pCR after NAC in our study. MetS status was an independent predictor of pCR, and pCR was more difficult to obtain in the MetS group than the non-MetS group (P=0.028). All metabolic parameters deteriorated significantly after NAC, especially the blood lipid index (P<0.010). The median follow-up time was 6 years. After adjusting for other prognostic factors, MetS was found to be strongly associated with an increased risk of recurrence (P=0.007) and mortality (P=0.004) in BC patients receiving NAC. Compared to individuals without any MetS component, the risk of death and disease progression increased sharply as the number of MetS components increased.

Conclusions: In BC patients who received NAC, MetS was associated with poor outcomes, including a lower pCR rate and increased risks of recurrence and mortality.

KEYWORDS

breast cancer, metabolic syndrome, neoadjuvant chemotherapy, pathologic complete response, prognosis

Introduction

In 2020, 19.3 million new cancer cases were diagnosed worldwide, including 2.3 million cases (11.7%) of breast cancer (BC), which has now surpassed lung cancer as the most commonly diagnosed cancer (1). Based on improved and intensified treatments developed over the past few decades, including neoadjuvant chemotherapy (NAC), the BC survival rate has improved significantly (2, 3). NAC has been established as a standard treatment approach in BC patients with locally advanced disease. Currently, the role of NAC has expanded to conversion of inoperable tumors to operable tumors or facilitating breast-conserving therapy (BCT) instead of mastectomy (4, 5), which is also known as tumor downstaging. Moreover, the assessment of tumor response to NAC is a useful tool that provides information on the impact of systemic therapies on BC biology (6). Pathologic complete response (pCR) after NAC serves as a significant surrogate marker that predicts better long-term prognosis (7).

As a significant public health problem worldwide, metabolic syndrome (MetS) is a multifactorial metabolic disease with main components, including obesity, hyperglycemia, dyslipidemia, and hypertension, which was initially linked to cardiovascular diseases (CVDs) (8). Several studies have found that CVD surpasses BC and has become the leading cause of death for BC survivors (9, 10). Accumulating evidence reveals a strong association between MetS and BC (11). MetS and its components are associated with increased risks of BC (12), and in-depth research on the association between MetS and the pathogenesis and prognosis of BC is increasing. Extensive literature has reported that metabolic dysregulation may affect the risk for occurrence, recurrence, and mortality of BC and the onset of additional chronic disease (13, 14). Investigation into the relationship between systemic therapies and MetS in BC survivors also represents an area of research that needs to be urgently addressed. Multiple studies have indicated that metabolic disorders, including overweight, dyslipidemia, and hypertension, are associated with worse pCR to NAC (15–17). However, clinical research on how MetS influences BC patients who receive NAC is currently lacking. This article retrospectively analyzed the clinical data of BC patients who underwent NAC before surgery and observed metabolic changes after adjuvant treatment to investigate the relationship between MetS and the pCR and long-term prognosis of BC patients after NAC and to provide a reference for the treatment of BC.

Abbreviations: ABCA1, adenosine triphosphate-binding cassette transporter A1; AJCC, American Joint Committee on Cancer; ATP, adenosine triphosphate; BCNACT, breast cancer neoadjuvant chemotherapy; BMI, body mass index; CI, confidence interval; CVD, cardiovascular disease; DBP, diastolic blood pressure; DFS, disease-free survival; ER, estrogen receptor; FBG, fasting blood glucose; HDL-C, high-density lipoprotein cholesterol; HER2, human epidermal growth factor receptor 2; HER2-E, HER2-enriched; HR, hazard ratio; IDF, International Diabetes Federation; IGF-1, insulin-like growth factor 1; IGFBP, IGF-binding proteins; IHC, immunohistochemical; IR, insulin resistance; ISH, *in situ* hybridization; LDL-C, low-density lipoprotein cholesterol; LDLR, low-density lipoprotein receptor; MetS, metabolic syndrome; NAC, neoadjuvant chemotherapy; NCEP-ATP III, National Cholesterol Education Program Adult Treatment Panel III; OR, odds ratio; OS, overall survival; pCR, pathologic complete response; PR, progesterone receptor; RR, risk ratio; SBP, systolic blood pressure; SD, standard deviation; TC, total cholesterol; TG, triglycerides; TNBC, triple negative breast cancer.

Materials and methods

Patient selection

Our study retrospectively analyzed 221 female BC patients who received NAC and underwent surgery at Harbin Medical University Cancer Hospital between September 2012 and December 2017. Before each treatment, patients signed the "Informed Consent Form for Secondary Use of Medical History Data/Biological Specimens" in our hospital. All procedures involving participants in this study were performed in accordance with Research Committee standards and complied with the 1964 Declaration of Helsinki and other amendments to ethical standards. The following patient inclusion criteria were employed (1): histopathologically confirmed BC by core needle biopsy and (2) preoperative NAC and no radiotherapy or endocrine therapy before chemotherapy. The following patient exclusion criteria were used: (1) patients with distant metastasis; (2) patients with other previous tumors; and (3) patients suffering from other diseases that affect body mass index (BMI), blood pressure, sugar and lipid metabolism or serious physical disease.

Data collection and biochemical variable determination

Clinical data were collected twice before and after NAC, and all data were collected from electronic medical records by two independent investigators. General clinical data included age, menopausal state, number of births, height, weight, blood pressure, fasting blood glucose (FPG), triglycerides (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C). BMI was calculated as body weight (kg) divided by the squared height (m2). Venous blood was taken after 12 hours of fasting, and the blood samples were sent to the Biochemical Laboratory of Medical University Cancer Hospital to detect FPG, TG, TC, HDL-C, and LDL-C.

Definition of MetS

The diagnosis of MetS was based on the National Cholesterol Education Program Adult Treatment Panel III

(NCEP-ATP III) criteria (18). Specifically, MetS was diagnosed if three of the following five criteria were present: obesity (waist circumference>88 cm); FPG≥110 mg/dl (6.1 mmol/L); TG≥150 mg/dl (1.7 mmol/L); HDL-C<50 mg/dl (1.3 mmol/L); and blood pressure≥130/85 mmHg. However, waist circumference was not available in this retrospective review given that this factor was not recorded at screening, so a BMI≥25 kg/m^2 replaced a waist circumference of 88 cm or more in women. This substitution was validated in previous studies (8, 19, 20) and is consistent with the diagnostic criteria for MetS established by the Diabetes Branch of the Chinese Medical Association in 2004 (21). Patients who met the diagnostic criteria for MetS were included in the MetS group; otherwise, patients were included in the non-MetS group.

Treatment plan

All patients received NAC before surgery and chose chemotherapy regimens according to modern treatment guidelines and patients' preferences. The following treatment regimens were noted: 78 cases of AC-T; TA scheme in 48 cases; TAC scheme in 76 cases; TCbH scheme in 7 cases; AC-TH scheme in 3 cases; TH scheme in 7 cases and TCb scheme in 2 cases (A: anthracycline; C: cyclophosphamide; T: taxane, including albumin paclitaxel or docetaxel; Cb: carboplatin; H: trastuzumab). The chemotherapy dose was decided by treatment guidelines and body surface area. One cycle of the chosen regimen was repeated every 3 weeks. All patients received at least three cycles of NAC. Surgery was performed after a rest period of 2-4 weeks after the completion of NAC, depending on the patient's condition. After surgery, all the enrolled patients received necessary follow-up treatment at Harbin Medical University Tumor Hospital. A total of 71.1% (64 cases) of estrogen receptor (ER)+/progesterone receptor (PR)+ patients and 61.3% (19 cases) of ER+/PR- patients received adjuvant endocrine therapy, and a total of 122 (55.2%) patients received radiation therapy.

Pathological features, molecular subtypes and pCR

The TNM staging system is based on the eighth edition of the American Joint Committee on Cancer (AJCC). ER, PR, human epidermal growth factor receptor 2 (HER2) and Ki67 status were assessed by immunohistochemical (IHC) staining or *in situ* hybridization (ISH). Luminal A, luminal B, HER2-enriched (HER2-E), and triple-negative molecular subtypes were included in this study. In our study, pCR was defined as no residual invasive disease (with or without ductal carcinoma in situ) in the breast and lymph nodes (ypT0/isN0).

Follow-up

Patients were regularly followed up at Harbin Medical University Cancer Hospital. Examinations were performed every 6 months during the first 5 years of follow-up and every 12 months thereafter. All patients were followed up until death or the study deadline (May 1, 2022) based on clinical records review and telephone. We defined disease-free survival (DFS) as the time from diagnosis until local, contralateral, and distant disease recurrence as well as secondary primary tumors or death from any cause. Overall survival (OS) was defined as the time from diagnosis to death from any cause or the end of follow-up.

Statistical analysis

All analyses were conducted with SPSS 26.0 statistical software. Descriptive statistics were reported as frequencies and percentages for categorical variables and as the mean ± standard deviation (SD) or median (interquartile range) for continuous variables. Comparison of patient characteristics between the different groups was performed using the independent T-test or nonparametric test for continuous variables and the chi-squared test or Fisher's exact test for categorical variables as appropriate. Univariate and multivariate analyses and subgroup analyses of the relationship between clinicopathological features and pCR were performed using logistic regression models and log-linear regression. Univariate and multivariate analyses of the association of clinicopathological characteristics with patients' OS and DFS were performed using the Cox proportional hazards model. The latter was adjusted for prognostic factors, including age, menopausal state, number of births, T stage, N stage, hormone receptors status, HER2 status, Ki67, p53 status, molecular subtype, endocrine therapy and radiation therapy. Survival curves were drawn using the Kaplan-Meier method. All statistical tests were two-tailed, and P values<0.05 were considered statistically significant.

Results

Patients' baseline characteristics

The 221 patients included in this study were all women with a median age of 49 years. A total of 49 (22.2%) BC patients were included in the MetS group, and 172 (77.8%) BC patients were in the non-MetS group. Compared to the non-MetS group, MetS group patients were more likely to be older (P<0.001) and postmenopausal (P<0.001), and the MetS group included a higher proportion of Ki-67≤14 (P=0.024) patients and more childbirths (P=0.014). Body weight, BMI, FBG, TG, TC, LDL-C,

and blood pressure were higher and HDL-C levels were lower in the MetS group than in the non-MetS group, and all these differences were statistically significant. MetS status was not associated with clinical T stage, N stage, hormone receptors, HER2 status or p53 status, and no differences in the number of NAC dose reductions and treatment interruptions were noted between the two groups (all P>0.05) (Table 1).

Univariate and multivariate analysis of pCR

In this study, a total of 53 (24.0%) patients achieved pCR after NAC, including five patients in the MetS group and 48 in the non-MetS group. Univariate analysis showed that the non-MetS group was more likely to achieve pCR than the MetS group (P=0.015). Patients who were hormone receptors negative, HER2 positive or Ki67>14% were more likely to achieve pCR (Table 2). Multivariate analysis showed that compared with ER+/PR+ patients, ER+/PR- patients and ER-/PR- patients had a higher probability of pCR, and ER-/PR- patients were particularly associated with pCR (OR=3.941, 95% CI: 1.772~8.766, P=0.001), and this finding reached statistical significance. Compared with the non-MetS group, it was more difficult for the MetS group to obtain pCR (OR=0.316, 95% CI: 0.113~0.886, P=0.028), indicating MetS and hormone receptors status were independent predictors of pCR (Table 3). Subgroup analysis showed that the relationship between MetS and pCR was more significant in the PR (−), HER2 (−), p53(−), and triple negative breast cancer (TNBC) subgroups (Figure 1).

Changes in MetS after NAC

The average duration of NAC was 4.67 months. After NAC, all metabolic parameters deteriorated, and the number of MetS components increased significantly. Among them, blood lipid indices, including TG, TC, HDL-C, and LDL-C, showed statistical deterioration (P<0.010) (Table 4). There were 49 (22.2%) patients in the MetS group before NAC and 80 (36.2%) patients in the MetS group after NAC. Forty-two (24.4%) patients in the non-MetS group met the diagnostic criteria for MetS after NAC (Figure 2).

Survival analysis

The mean OS and DFS values of 221 patients to the follow-up deadline were 96.75 and 87.46 months, respectively. The five-year survival rate of the MetS group was 64.6%, whereas that of the non-MetS group was 85.3%. In univariate analysis, MetS was associated with a greater than twofold increased risk of breast cancer mortality and recurrence (OR=2.463, 95% CI 1.391-4.363, P=0.002) (OR=2.213, 95% CI 1.336-3.668, P=0.002). Compared with postmenopausal patients, premenopausal patients had a longer OS and DFS (OR=2.316, 95% CI 1.315-4.079, P=0.004) (OR=1.792, 95% CI 1.108-2.898, P=0.017) (Table 5).

In the multivariate analysis, hazard ratios were adjusted for age, menopausal state, number of births, T stage, N stage, hormone receptors status, HER2 status, Ki67, p53 status, molecular subtype, endocrine therapy and radiation therapy. High TG (≥1.7 mmol/L) and low HDL-C (<1.3 mmol/L) were individually associated with an increased risk of death (OR=2.452, CI 95% 1.271-4.731, P=0.007) (OR=2.069, 95% CI 1.073-3.988, P=0.030). High TG (≥1.7 mmol/L), low HDL-C (<1.3 mmol/L) and hypertension were individually associated with an increased risk of relapse (OR=1.855, 95% CI 1.046-3.291, P=0.035) (OR=1.883, 95% CI 1.066-3.327, P=0.029) (OR=1.802, 95% CI 1.042-3.116, P=0.035) in multivariable-adjusted models. However, MetS remained the most significant predictor of disease progression and death after adjustment. MetS patients had a 2.587-fold increased risk of death (OR=2.587, 95% CI 1.359-4.924, P=0.004) and a 2.228-fold increased risk of recurrence (OR=2.228, 95% CI 1.251-3.970, P=0.007) compared with patients who were not diagnosed with MetS. Compared to individuals without any component of MetS present, the risk of death and disease progression increased steeply as the number of MetS components increased. Patients with 1-2, 3, 4, and 5 components had a 1.763-, 2.865-, 6.304-, and 15.488-fold higher risk of death and a 1.951-, 2.995-, 4.584-, and 12.129-fold higher risk of relapse, respectively, than patients with 0 components (Table 6).

The follow-up time ranged from 12 to 115 months. The median follow-up time of 221 patients was 72.00 ± 2.44 months (6 years). Six years after diagnosis, the rates for OS and DFS were 84.4% vs. 59.1% (P=0.001) (Figure 3A) and 74.7% vs. 53.1% (P=0.001) (Figure 3B), respectively, in patients with non-MetS vs. MetS. Specifically, rates for OS and DFS were 85.9% vs. 77.9% vs. 59.1% (P=0.002) (Figure 4A) and 82.4% vs. 68.4% vs. 53.1% (P=0.001) (Figure 4B) in patients with 0 vs. 1-2 vs. 3-5 components of MetS. Kaplan−Meier survival analysis showed that BC patients receiving NAC with MetS before treatment had worse OS and DFS than those without MetS, and the difference was statistically significant.

Discussion

As a significant public health problem worldwide, MetS is a cluster of risk factors for CVD and various malignant tumors. Several cohort studies and meta-analyses have highlighted the link between MetS and the prevalence, recurrence, and mortality of BC (11, 22, 23). NAC is increasingly being utilized as the first-line therapy for BC (6). Some studies have found that metabolic dysregulation status has predictive value for NAC in BC;

TABLE 1 Patient clinicopathological characteristics by MetS status.

Variable	Total (n=221)	MetS group (n=49)	Non-MetS group (n=172)	P
Age (years)	49.190 ± 9.415	54.730 ± 8.129	47.610 ± 9.174	<0.001
Menopause				<0.001
No	128 (57.9%)	16 (32.7%)	112 (65.1%)	
Yes	93 (42.1%)	33 (67.3%)	60 (34.9%)	
Number of births				0.014
0	27 (12.2%)	2 (4.1%)	25 (14.5%)	
1	137 (62.0%)	28 (57.1%)	109 (63.4%)	
2	44 (19.9%)	17 (34.7%)	27 (15.7%)	
>2	13 (5.9%)	2 (4.1%)	11 (6.4%)	
T Stage				0.138
cT_1	31 (14.0%)	10 (20.4%)	21 (12.2%)	
cT_2	143 (64.7%)	28 (57.2%)	115 (66.9%)	
cT_3	41 (18.6%)	8 (16.3%)	33 (19.2%)	
cT_4	6 (2.7%)	3 (6.1%)	3 (1.7%)	
N Stage				0.134
N_0	11 (5.0%)	3 (6.1%)	8 (4.6%)	
N_1	36 (16.3%)	4 (8.2%)	32 (18.6%)	
N_2	75 (33.9%)	14 (28.6%)	61 (35.5%)	
N_3	99 (44.8%)	28 (57.1%)	71 (41.3%)	
Hormone receptors				0.612
ER+/PR+	90 (40.7%)	22 (44.9%)	68 (39.5%)	
ER+/PR-	31 (14.0%)	8(16.3%)	23 (13.4%)	
ER-/PR-	94 (42.5%)	18 (36.7%)	76 (44.2%)	
HER2				0.127
Negative	142 (64.3%)	36 (73.5%)	106 (61.6%)	
Positive	79 (35.7%)	13 (26.5%)	66 (38.4%)	
Ki-67(%)				0.024
≤14	58 (26.2%)	19 (38.8%)	39 (22.7%)	
>14	163 (73.8%)	30 (61.2%)	133 (77.3%)	
p53				0.954
Negative	130 (58.8%)	29 (59.2%)	101 (58.7%)	
Positive	91 (41.2%)	20 (40.8%)	71 (41.3%)	
Subtype				0.550
Luminal A	21 (9.5%)	7 (14.3%)	14 (8.1%)	
Luminal B	106 (48.0%)	24 (49.0%)	82 (47.7%)	
HER2-E	50 (22.6%)	9 (18.4%)	41 (23.8%)	
TNBC	44 (19.9%)	9 (18.4%)	35 (20.4%)	

(Continued)

TABLE 1 Continued

Variable	Total (n=221)	MetS group (n=49)	Non-MetS group (n=172)	P
NAC dose reduction				0.626
No	194 (87.8%)	44 (89.8%)	150 (87.2%)	
Yes	27 (12.2%)	5 (10.2%)	22 (12.8%)	
NAC treatment interruption				0.294
No	144 (65.2%)	30 (61.2%)	119 (69.2%)	
Yes	77 (34.8%)	19 (38.8%)	53 (30.8%)	
Height (cm)	160.102 ± 5.322	159.306 ± 6.249	160.328 ± 5.024	0.296
Weight (kg)	62.887 ± 9.260	67.602 ± 8.598	61.544 ± 9.021	<0.001
BMI (kg/m2)	24.524 ± 3.309	26.617 ± 2.818	23.927 ± 3.200	<0.001
FBG (mmol/L)	5.300 (4.800-5.800)	6.100 (5.400-6.950)	5.100 (4.700-5.600)	<0.001
TG (mmol/L)	1.140 (0.785-1.550)	1.850 (1.340-2.780)	0.955 (0.730-1.320)	<0.001
TC (mmol/L)	4.693 ± 0.961	5.027 ± 1.017	4.597 ± 0.925	0.005
HDL-C (mmol/L)	1.620 ± 0.401	1.359 ± 0.341	1.695 ± 0.0.386	<0.001
LDL-C (mmol/L)	3.135 ± 0.890	3.440 ± 0.917	3.048 ± 0.865	0.006
SBP (mmHg)	124.411 ± 20.237	140.787 ± 18.057	119.746 ± 18.347	<0.001
DBP (mmHg)	75.813 ± 11.901	83.435 ± 12.856	73.641 ± 10.697	<0.001

MetS, metabolic syndrome; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HER2-E, HER2-enriched; TNBC, triple negative breast cancer; NAC, neoadjuvant chemotherapy; BMI, body mass index; FBG, fasting blood glucose; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; SBP, systolic blood pressure; DBP, diastolic blood pressure.

TABLE 2 Univariate analysis between clinical characteristics and pCR.

Variable	Total (n=221)	pCR (n=53)	OR	CI (95%)	P
Age(years)					
≤49	115 (52.0%)	30 (56.6%)	Ref.	Ref.	Ref.
>49	106 (48.0%)	23 (43.4%)	0.785	0.422-1.462	0.446
Menopause					
No	128 (57.9%)	35 (66.0%)	Ref.	Ref.	Ref.
Yes	93 (42.1%)	18 (34.0%)	0.638	0.335-1.215	0.171
Number of births					
0	27 (12.2%)	7 (13.2%)	Ref.	Ref.	Ref.
1	137 (62.0%)	34 (64.1%)	0.943	0.367-2.424	0.903
2	44 (19.9%)	9 (17.0%)	0.735	0.237-2.275	0.593
>2	13 (5.9%)	3 (5.7%)	0.857	0.182-4.042	0.846
T Stage					
cT_1+cT_2	174 (78.7%)	46 (86.8%)	Ref.	Ref.	Ref.

(Continued)

10.3389/fonc.2022.1080054

TABLE 2 Continued

Variable	Total (n=221)	pCR (n=53)	OR	CI (95%)	P
cT_3+cT_4	47 (21.3%)	7 (13.2%)	0.487	0.204-1.163	0.105
N Stage					
N_0	11 (5.0%)	2 (3.8%)	Ref.	Ref.	Ref.
N_1+N_2+N_3	210 (95.0%)	51 (96.2%)	1.443	0.302-6.898	0.646
Hormone receptors					
ER+/PR+	90 (40.7%)	11 (20.8%)	Ref.	Ref.	Ref.
ER+/PR-	31 (14.0%)	5 (9.4%)	1.381	0.439-4.346	0.581
ER-/PR-	94 (42.5%)	37 (69.8%)	4.662	2.193-9.912	<0.001
HER2					
Negative	142 (64.3%)	25 (47.2%)	Ref.	Ref.	Ref.
Positive	79 (35.7%)	28 (52.8%)	2.569	1.366-4.832	0.003
Ki67(%)					
≤14	58 (26.2%)	7 (13.2%)	Ref.	Ref.	Ref.
>14	163 (73.8%)	46 (86.8%)	2.864	1.212-6.773	0.017
p53					
Negative	130 (58.8%)	32 (60.4%)	Ref.	Ref.	Ref.
Positive	91 (41.2%)	21 (39.6%)	0.919	0.489-1.725	0.792
MetS status					
No	172 (77.8%)	48 (90.6%)	Ref.	Ref.	Ref.
Yes	49 (22.2%)	5 (9.4%)	0.294	0.110-0.785	0.015

pCR, pathologic complete response; OR, odds ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; MetS, metabolic syndrome.

specifically, higher BMI was associated with worse pCR to NAC (15). Diabetes and high FPG levels may be predictive of nonresponse to neoadjuvant chemotherapy in patients with BC (17). We evaluated the predictive effect of MetS on pCR in BC patients who received NAC, as it could be used to select those patients who demonstrate the most benefit from neoadjuvant systemic therapy, and analyzed long-term prognostic characteristics in these patients. To the best of our knowledge, our study is the first to date to systematically address the effect of MetS and its components on BC patients who received NAC.

Our paper retrospectively analyzed 221 BC patients who received NAC at Harbin Medical University Cancer Hospital. Similar to that noted other reports (14, 23), our study found that the MetS group included more elderly and postmenopausal patients than the non-MetS group. Unlike previous studies showing that MetS was associated with adverse pathological features (24), we found that the MetS group had a higher proportion of Ki-67≤14 patients. This finding may be due to the notion that lower Ki67 expression is associated with decreased metabolic activity, but more research is needed to reveal specific mechanisms (25). The MetS group had more childbirths than the non-MetS group. This is probably due to pregnancy involving marked alterations in metabolic parameters, including reduced insulin sensitivity in peripheral tissues, increased production of insulin from the pancreas, and accumulation and redistribution of body fat (26, 27). Previous studies showed that an increased number of births was associated with type 2 diabetes (28, 29).

In our study, 53 (24.0%) patients achieved pCR after NAC. The multivariate analysis showed that MetS (P=0.028) and hormone receptors status were independent predictors of pCR after NAC in breast cancer. Compared with the non-MetS group, the MetS group had more difficulty obtaining pCR. ER-/PR- patients had a higher probability of pCR than ER+/PR+ patients. In the subgroup analysis, we found that in the PR (−), HER2 (−), p53(−) and TNBC subgroups, MetS

10.3389/fonc.2022.1080054

TABLE 3 Multivariate analysis between clinical characteristics and pCR.

Variable	OR	CI (95%)	P
Hormone receptors			
ER+/PR+	Ref.	Ref.	Ref.
ER+/PR-	1.334	0.404-4.403	0.636
ER-/PR-	3.941	1.772-8.766	0.001
HER2			
Negative	Ref.	Ref.	Ref.
Positive	1.545	0.770-3.099	0.221
Ki67(%)			
≤14	Ref.	Ref.	Ref.
>14	2.395	0.962-5.962	0.061
MetS status			
No	Ref.	Ref.	Ref.
Yes	0.316	0.113-0.886	0.028

pCR, pathologic complete response; OR, odds ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; MetS, metabolic syndrome.

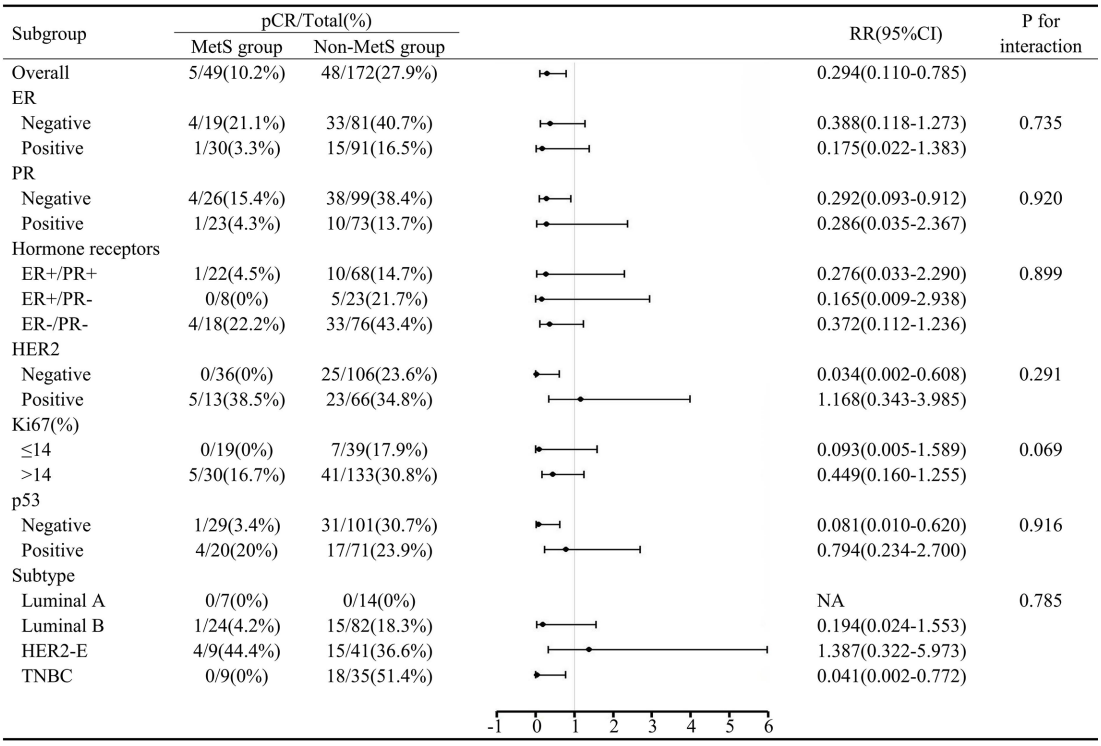


FIGURE 1
Subgroup analysis of MetS and pCR. MetS, metabolic syndrome; pCR, pathologic complete response; RR, risk ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HER2-E, HER2-enriched; TNBC, triple negative breast cancer.

10.3389/fonc.2022.1080054

TABLE 4 Changes in metabolic parameters before and after NAC.

Variable	Pre-NAC	Post-NAC	t/Z	P
Height (cm)	160.102 ± 5.322	–	–	–
Weight (kg)	62.887 ± 9.260	63.991 ± 9.389	-1.245	0.214
BMI (kg/m2)	24.524 ± 3.309	24.956 ± 3.370	-1.362	0.174
FBG (mmol/L)	5.300 (4.800-5.800)	5.400 (4.950-5.900)	-1.880	0.060
TG (mmol/L)	1.140 (0.785-1.550)	1.730 (1.245-2.490)	-8.054	<0.001
TC (mmol/L)	4.693 ± 0.961	5.172 ± 1.053	-5.002	<0.001
HDL-C (mmol/L)	1.620 ± 0.401	1.503 ± 0.392	3.101	0.002
LDL-C (mmol/L)	3.135 ± 0.890	3.572 ± 0.906	-5.121	<0.001
SBP (mmHg)	124.411 ± 20.237	126.883 ± 17.612	-1.369	0.172
DBP (mmHg)	75.813 ± 11.901	76.706 ± 12.441	-0.771	0.441
No. of MetS components	1.440 ± 1.308	2.030 ± 1.321	-4.704	<0.001

NAC, neoadjuvant chemotherapy; BMI, body mass index; FBG, fasting blood glucose; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C; low-density lipoprotein cholesterol; SBP, systolic blood pressure; DBP, diastolic blood pressure; MetS, metabolic syndrome.

intervention can improve the pCR rate more effectively. This information should be considered when selecting patients who are most likely to benefit from NAC. Given that multicollinearity exists between hormone receptors and subtypes, the latter was not included in logistic regression models for analysis. Consistent with the extremely low pCR rates (0.3%) reported in previous studies (30), no luminal A patients obtained pCR in our study. The relationship between MetS and pCR in BC patients who underwent NAC was not consistent in previous studies. A study of 150 breast cancer neoadjuvant chemotherapy (BCNACT) patients which adopted International Diabetes Federation (IDF) criteria to diagnose MetS reported that MetS before BCNACT predicted a lower pCR rate (P=0.003) (31). Tong et al. found that in HER2-positive BC patients receiving neoadjuvant therapy, MetS showed a tendency to interfere with NAC effcacy, but the difference was not statistically significant in multivariate analysis (32). Similarly, Alan et al. did not identify a relationship between MetS and pCR in a study of 55 patients (33).

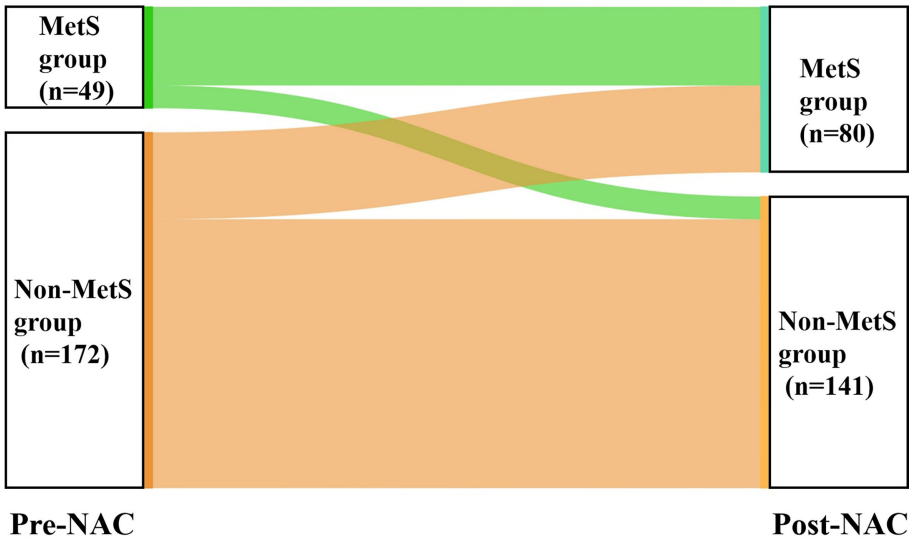


FIGURE 2
Changes in MetS status before and after NAC. MetS, metabolic syndrome; NAC, neoadjuvant chemotherapy.

10.3389/fonc.2022.1080054

TABLE 5 Univariate analysis of hazards ratios for OS and DFS by clinical characteristics and MetS status.

Variable	OS			DFS		
	HR	CI (95%)	P	HR	CI (95%)	P
Age(years)						
≤49	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
>49	1.509	0.863-2.638	0.149	1.355	0.838-2.190	0.216
Menopause						
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	2.316	1.315-4.079	0.004	1.792	1.108-2.898	0.017
Number of births						
0	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1	1.213	0.470-3.126	0.690	1.255	0.564-2.794	0.578
2	1.629	0.574-4.627	0.359	1.507	0.614-3.697	0.370
>2	1.299	0.310-5.436	0.720	0.912	0.236-3.528	0.894
T Stage						
cT_1+cT_2	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
cT_3+cT_4	1.301	0.691-2.448	0.415	1.182	0.674-2.073	0.560
N Stage						
N_0	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
N_1+N_2+N_3	2.786	0.385-20.181	0.311	0.615	0.247-1.529	0.296
Hormone receptors						
ER+/PR+	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
ER+/PR-	2.107	0.978-4.540	0.057	1.735	0.906-3.324	0.097
ER-/PR-	1.430	0.755-2.707	0.272	0.971	0.564-1.672	0.915
HER2						
Negative	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Positive	1.478	0.845-2.583	0.171	1.102	0.672-1.808	0.701
Ki67(%)						
≤14	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
>14	1.334	0.683-2.605	0.399	1.318	0.742-2.341	0.346
p53						
Negative	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Positive	1.409	0.809-2.456	0.226	1.194	0.737-1.932	0.472
Subtype						
Luminal A	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.

(Continued)

10.3389/fonc.2022.1080054

TABLE 5 Continued

Variable	OS			DFS		
	HR	CI (95%)	P	HR	CI (95%)	P
Luminal B	2.591	0.614-10.946	0.195	2.708	0.741-5.831	0.165
HER2-E	2.408	0.534-10.868	0.253	1.363	0.439-4.225	0.592
TNBC	3.118	0.698-13.933	0.137	1.868	0.615-5.676	0.270
Endocrine therapy						
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	0.818	0.462-1.448	0.491	1.345	0.833-2.172	0.225
Radiation therapy						
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	1.081	0.616-1.896	0.786	1.012	0.625-1.638	0.961
MetS status						
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	2.463	1.391-4.363	0.002	2.213	1.336-3.668	0.002

OS, overall survival; DFS, disease-free survival; HR, hazard ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HER2-E, HER2-enriched; TNBC, triple negative breast cancer; MetS, metabolic syndrome.

TABLE 6 Multivariate analysis of hazards ratios for OS and DFS by MetS components.

Variable	OS			DFS		
	HR	CI (95%)	P	HR	CI (95%)	P
MetS status						
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	2.587	1.359-4.924	0.004	2.228	1.251-3.970	0.007
BMI (kg/m2)						
<25	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
≥25	1.548	0.852-2.813	0.151	1.609	0.961-2.693	0.071
FBG (mmol/L)						
<6.1	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
≥6.1	1.902	0.981-3.687	0.057	1.687	0.905-3.147	0.100
TG (mmol/L)						
<1.7	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
≥1.7	2.452	1.271-4.731	0.007	1.855	1.046-3.291	0.035
HDL-C (mmol/L)						
≥1.3	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.

(Continued)

TABLE 6 Continued

Variable	OS			DFS		
	HR	CI (95%)	P	HR	CI (95%)	P
<1.3	2.069	1.073-3.988	0.030	1.883	1.066-3.327	0.029
Hypertension						
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	1.639	0.883-3.041	0.117	1.802	1.042-3.116	0.035
Number of components						
0 components	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1–2 components	1.763	0.750-4.144	0.194	1.951	0.949-4.009	0.069
3 components	2.865	1.036-7.922	0.042	2.995	1.250-7.176	0.014
4 components	6.304	1.692-23.492	0.006	4.584	1.335-15.734	0.016
5 components	15.488	3.282-73.083	0.001	12.129	2.833-51.921	0.001

OS, overall survival; DFS, disease-free survival; HR, hazard ratio; CI, confidence interval; MetS, metabolic syndrome; BMI, body mass index; FBG, fasting blood glucose; TG, triglycerides; HDL-C, high-density lipoprotein cholesterol.
Multivariate analysis adjusted for age, menopausal state, number of births, T Stage, N Stage, hormone receptors status, HER2 status, Ki67, p53 status, molecular subtype, endocrine therapy and radiation therapy.

After NAC, all metabolic parameters worsened to varying degrees, and the number of MetS components was significantly increased (P<0.001). We can learn how quickly metabolic changes occur during NAC in BC patients who do not have any severe comorbidities at the time of diagnosis. Consistent with Tong's study (32), the major metabolic disturbances observed were impaired lipid metabolism after NAC. We found that all blood lipid indices, including TG, TC, HDL-C, and LDL-C, were significantly worsened (P<0.010) to a greater extent than other metabolic biomarkers. Dyslipidemia, especially elevated LDL-C levels, is the most important independent risk factor for atherosclerotic CVD (34). The mechanism of dyslipidemia after NAC is unclear. Studies have shown that doxorubicin can regulate a series of genes involved in lipoprotein metabolism in liver cells, such as adenosine triphosphate (ATP)-binding cassette transporter A1 (ABCA1) and apoA1. In addition, doxorubicin and paclitaxel increase apoB protein levels, and paclitaxel decreases low-density lipoprotein receptor (LDLR) protein levels (35). This result suggests that long-term management of blood lipid profiles is necessary for BC patients who have received NAC, especially in patients who also require endocrine therapy, such as tamoxifen and aromatase inhibitors, which could alter lipid profiles in different ways (36, 37).

In survival analysis, we evaluated the association between MetS and its components with clinical outcomes in BC patients

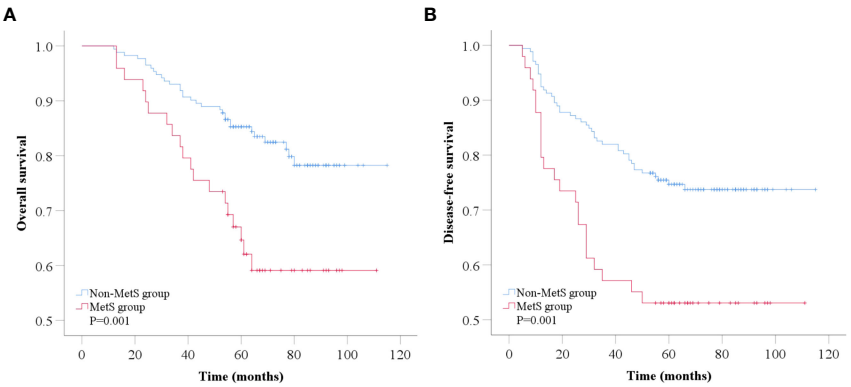


FIGURE 3
Kaplan-Meier analysis of overall survival **(A)** and disease-free survival **(B)** according to MetS status. MetS, metabolic syndrome.

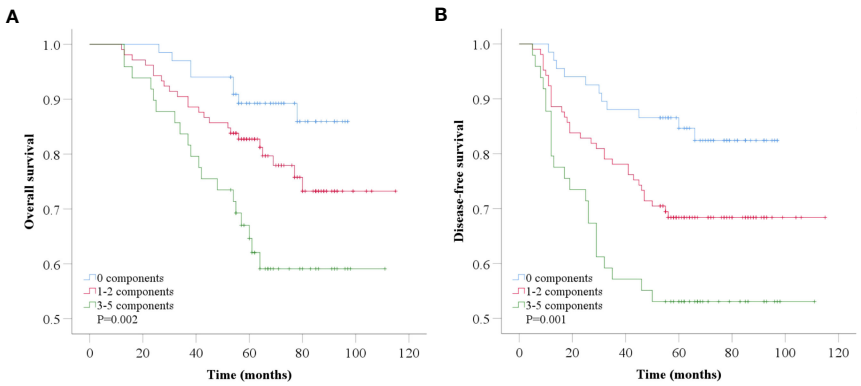


FIGURE 4
Kaplan-Meier analysis of overall survival **(A)** and disease-free survival **(B)** according to number of MetS components.

receiving NAC. By combining the results of multivariable adjusted data, our study showed that MetS was associated with higher overall mortality (P=0.004) and recurrence risk (P=0.007) in BC patients who received NAC, and this association was independent of some known prognostic factors, such as age, disease stage, and hormone receptors status. These results strongly indicated that MetS remains an independent predictor of poor prognosis in BC patients receiving NAC. In our study, even the presence of a single component of MetS was associated with an increased risk of recurrence and mortality in BC patients receiving NAC. In addition, as the number of MetS components increased, the risk of recurrence and mortality increased significantly. We observed that the risk of mortality increased from approximately twofold to greater than 15-fold among patients in whom the number of MetS components increased from 1 to 5 compared with those with no MetS components. Interestingly, among patients without MetS, the risk of recurrence mortality increased significantly as the number of MetS components increased. These results indicate that the greater the extent of metabolic dysregulation, the worse the outcomes in BC patients receiving NAC. This findings is consistent with Berrino's study in early-stage breast cancer (14). Our study also investigated the impact of individual MetS components on BC outcome with differing results. However, as a comprehensive indicator, MetS was a more precise indicator of prognosis than individual MetS components.

The potential mechanisms between MetS and poor prognosis in breast cancer are currently under exploration. MetS itself is not a disease but a series of interdependent abnormal metabolic factors. Each of the metabolic alterations may be associated with the more aggressive tumor biology of BC. Insulin resistance (IR) and hyperinsulinemia are essential to the pathogenesis of type 2 diabetes and obesity (38). Insulin directly promotes breast tissue and tumor cell proliferation, thus possibly promoting BC incidence. In addition, hyperinsulinemia increases insulin-like growth factor 1 (IGF-1) bioavailability by increasing hepatic growth hormone receptor expression and repressing hepatic production of IGF-binding proteins (IGFBP) (39), resulting in hyperactivation of the Ras-MAPK and PI3K/Akt pathways in malignant cells to promote cell proliferation (40). Chronic inflammation, another critical pathophysiological feature of MetS (41), is also involved in the development and aggression of many malignancies. This process is characterized by reduced levels of anti-inflammatory cytokines (such as adiponectin) and high levels of pro-inflammatory cytokines (42). Adiponectin promotes glucose and fatty acid metabolism and improves insulin sensitivity and resistance. Adiponectin induces cell cycle arrest and apoptosis, increases the expression of the proapoptotic genes BAD (BCL2-associated agonist of cell death) and TP53 (tumor protein p53), decreases the antiapoptotic gene BCL2, and reduces the expression of CCND1 (cyclin D1) and CCNE2 (cyclin E2) in breast cancer cells, thereby inhibiting growth, invasion, and migration and inducing apoptosis of cancer cells (43). As the aromatase enzyme synthesizes estrogens in adipose tissue from circulating androgens, obesity could promote estrogen production (40), especially estradiol. This process also reduces adiponectin production, thereby attenuating the antitumor effect of adiponectin (44). Similarly, adiponectin levels are reduced in patients with diabetes and coronary heart disease. Furthermore, cholesterol promotes tumor growth and metastasis in BC through the PI3K/Akt signaling pathway (45). The mechanisms of the different molecular pathways involved in MetS and poor prognosis in patients with BC deserve further investigation.

Several limitations of this study should be noted. First, as a single-center study, our samples are obtained from single provinces in China, which may increase the heterogeneity between samples. The conclusion from this study needs to be verified in a larger and racially diverse population. Second, the diagnosis of MetS in our study was based on NCEP-ATPIII criteria. However, as a retrospective study, we did not have waist circumference data of patients, so we replaced waist circumference with BMI, which is more consistent with the actual Chinese characteristics. Third, we

did not have medical treatment information for hyperglycemia, dyslipidemia, and hypertension in patients with MetS; thus, the number of patients with MetS was underestimated. In addition, this study did not exclude the interference from targeted therapy in the assessment of response to NAC in BC patients. In our study, only 17 (21.5%) of HER2-positive BC patients received targeted therapy with trastuzumab before surgery due to financial limitations. Trastuzumab has been covered by insurance in China only since 2017, and this information should be considered in further studies.

Conclusion

In BC patients who received NAC, MetS was associated with poor outcomes, including a lower pCR rate and increased risk of recurrence and mortality, suggesting that timely MetS intervention is needed for a better prognosis.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions. The data that support the findings of this study are available from Harbin Medical University Cancer Hospital, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics committee of Harbin Medical University Cancer Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ZZ, YZ, YH and SC contributed to the conception and design. ZZ, YL, CJ, YW and LS collected and analyzed the data and wrote the manuscript. All authors revised the manuscript critically. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Beijing Medical Award Foundation (YXJL-2020-1225-0323).

Acknowledgments

We are grateful to all participants and all coauthors in the study.

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.

Publisher's note

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Li T, Mello-Thoms C, Brennan PC. Descriptive epidemiology of breast cancer in China: Incidence, mortality, survival and prevalence. *Breast Cancer Res Treat* (2016) 159(3):395–406. doi: 10.1007/s10549-016-3947-0

3. Hendrick RE, Baker JA, Helvie MA. Breast cancer deaths averted over 3 decades. *Cancer* (2019) 125(9):1482–8. doi: 10.1002/cncr.31954

4. Shien T, Iwata H. Adjuvant and neoadjuvant therapy for breast cancer. *Jpn J Clin Oncol* (2020) 50(3):225–9. doi: 10.1093/jjco/hyz213

5. Sun C, Shi L, Gu Y, Hu Y, Wang J, Liu Y, et al. Clinical effects of neoadjuvant chemotherapy in treating breast cancer. *Cancer Biother Radiopharm* (2021) 36 (2):174–9. doi: 10.1089/cbr.2019.3545

6. Spring LM, Fell G, Arfe A, Sharma C, Greenup R, Reynolds KL, et al. Pathologic complete response after neoadjuvant chemotherapy and impact on breast cancer recurrence and survival: A comprehensive meta-analysis. *Clin Cancer Res* (2020) 26(12):2838–48. doi: 10.1158/1078-0432.CCR-19-3492

7. Wang H, Mao X. Evaluation of the efficacy of neoadjuvant chemotherapy for breast cancer. *Drug Des Devel Ther* (2020) 14:2423–33. doi: 10.2147/DDDT.S253961

8. Lemieux I, Després J-P. Metabolic syndrome: Past, present and future. *Nutrients* (2020) 12(11):3501. doi: 10.3390/nu12113501

9. Patnaik JL, Byers T, DiGuiseppi C, Dabelea D, Denberg TD. Cardiovascular disease competes with breast cancer as the leading cause of death for older females diagnosed with breast cancer: A retrospective cohort study. *Breast Cancer Res* (2011) 13(3):R64. doi: 10.1186/bcr2901

10. Abdel-Qadir H, Austin PC, Lee DS, Amir E, Tu JV, Thavendiranathan P, et al. A population-based study of cardiovascular mortality following early-stage breast cancer. *JAMA Cardiol* (2017) 2(1):88–93. doi: 10.1001/jamacardio.2016.3841

11. Esposito K, Chiodini P, Colao A, Lenzi A, Giugliano D. Metabolic syndrome and risk of cancer: A systematic review and meta-analysis. *Diabetes Care* (2012) 35 (11):2402–11. doi: 10.2337/dc12-0336

12. Agnoli C, Berrino F, Abagnato CA, Muti P, Panico S, Crosignani P, et al. Metabolic syndrome and postmenopausal breast cancer in the ordet cohort: A nested case-control study. *Nutr Metab Cardiovasc Dis* (2010) 20(1):41–8. doi: 10.1016/j.numecd.2009.02.006

13. Dong S, Wang Z, Shen K, Chen X. Metabolic syndrome and breast cancer: Prevalence, treatment response, and prognosis. *Front Oncol* (2021) 11:629666. doi: 10.3389/fonc.2021.629666

14. Berrino F, Villarini A, Traina A, Bonanni B, Panico S, Mano MP, et al. Metabolic syndrome and breast cancer prognosis. *Breast Cancer Res Treat* (2014) 147(1):159–65. doi: 10.1007/s10549-014-3076-6

15. Litton JK, Gonzalez-Angulo AM, Warneke CL, Buzdar AU, Kau S-W, Bondy M, et al. Relationship between obesity and pathologic response to neoadjuvant chemotherapy among women with operable breast cancer. *J Clin Oncol* (2008) 26(25):4072–7. doi: 10.1200/JCO.2007.14.4527

16. Hilvo M, Gade S, Hyötyläinen T, Nekljudova V, Seppänen-Laakso T, Sysi-Aho M, et al. Monounsaturated fatty acids in serum triacylglycerols are associated with response to neoadjuvant chemotherapy in breast cancer patients. *Int J Cancer* (2014) 134(7):1725–33. doi: 10.1002/ijc.28491

17. Arici S, Geredeli C, Secmeler S, Cekin R, Sakin A, Cihan S. The effects of diabetes and fasting plasma glucose on treatment of breast cancer with neoadjuvant chemotherapy. *Curr Probl In Cancer* (2020) 44(1):100485. doi: 10.1016/j.currproblcancer.2019.05.007

18. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the national cholesterol education program (Ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult treatment panel iii). *JAMA* (2001) 285(19):2486–97. doi: 10.1001/jama.285.19.2486

19. Abbasi F, Malhotra D, Mathur A, Reaven GM, Molina CR. Body mass index and waist circumference associate to a comparable degree with insulin resistance and related metabolic abnormalities in south Asian women and men. *Diabetes Vasc Dis Res* (2012) 9(4):296–300. doi: 10.1177/1479164111433578

20. Oda E, Kawai R. Comparison among body mass index (Bmi), waist circumference (Wc), and percent body fat (%Bf) as anthropometric markers for the clustering of metabolic risk factors in Japanese. *Intern Med* (2010) 49(15):1477–82. doi: 10.2169/internalmedicine.49.3363

21. Zhong Y, Hu M, Wang Q, Yang Z, Zhu N, Wang F, et al. The prevalence and related factors of metabolic syndrome in outpatients with first-episode drug-naive major depression comorbid with anxiety. *Sci Rep* (2021) 11(1):3324. doi: 10.1038/s41598-021-81653-2

22. Uzunlulu M, Telci Caklili O, Oguz A. Association between metabolic syndrome and cancer. *Ann Nutr Metab* (2016) 68(3):173–9. doi: 10.1159/000443743

23. Buono G, Crispo A, Giuliano M, De Angelis C, Schettini F, Forestieri V, et al. Metabolic syndrome and early stage breast cancer outcome: Results from a prospective observational study. *Breast Cancer Res Treat* (2020) 182(2):401–9. doi: 10.1007/s10549-020-05701-7

24. Healy LA, Ryan AM, Carroll P, Ennis D, Crowley V, Boyle T, et al. Metabolic syndrome, central obesity and insulin resistance are associated with adverse pathological features in postmenopausal breast cancer. *Clin Oncol (R Coll Radiol)* (2010) 22(4):281–8. doi: 10.1016/j.clon.2010.02.001

25. Jahani M, Shahlaei M, Norooznezhad F, Miraghaee SS, Hosseinzadeh L, Moasefi N, et al. Tsga10 over expression decreases metastasic and metabolic activity by inhibiting hif-1 in breast cancer cells. *Arch Med Res* (2020) 51(1):41–53. doi: 10.1016/j.arcmed.2019.12.002

26. Stuebe AM, Mantzoros C, Kleinman K, Gillman MW, Rifas-Shiman S, Seely EW, et al. Gestational glucose tolerance and maternal metabolic profile at 3 years postpartum. *Obstet Gynecol* (2011) 118(5):1065–73. doi: 10.1097/AOG.0b013e3182325f5a

27. Gunderson EP. Childbearing and obesity in women: Weight before, during, and after pregnancy. *Obstet Gynecol Clin North Am* (2009) 36(2):317–32. doi: 10.1016/j.ogc.2009.04.001

28. Araneta MRG, Barrett-Connor E. Grand multiparity is associated with type 2 diabetes in Filipino American women, independent of visceral fat and adiponectin. *Diabetes Care* (2010) 33(2):385–9. doi: 10.2337/dc09-1477

29. Mueller NT, Mueller NJ, Odegaard AO, Gross MD, Koh WP, Yuan JM, et al. Higher parity is associated with an increased risk of type-ii diabetes in Chinese women: The Singapore Chinese health study. *BJOG* (2013) 120(12):1483–9. doi: 10.1111/1471-0528.12364

30. Haque W, Verma V, Hatch S, Suzanne Klimberg V, Brian Butler E, Teh BS. Response rates and pathologic complete response by breast cancer molecular subtype following neoadjuvant chemotherapy. *Breast Cancer Res Treat* (2018) 170(3):559–67. doi: 10.1007/s10549-018-4801-3

31. Lu Y, Wang P, Lan N, Kong F, Abdumijit A, Tu S, et al. Metabolic syndrome predicts response to neoadjuvant chemotherapy in breast cancer. *Front Oncol* (2022) 12:899335. doi: 10.3389/fonc.2022.899335

32. Tong Y-W, Wang G, Wu J-Y, Huang O, He J-R, Zhu L, et al. Insulin-like growth factor-1, metabolic abnormalities, and pathological complete remission rate in Her2-positive breast cancer patients receiving neoadjuvant therapy. *Onco Targets Ther* (2019) 12:3977–89. doi: 10.2147/OTT.S194981

33. Alan O, Akin Telli T, Aktas B, Koca S, Ökten IN, Hasanov R, et al. Is insulin resistance a predictor for complete response in breast cancer patients who underwent neoadjuvant treatment? *World J Surg Oncol* (2020) 18(1):242. doi: 10.1186/s12957-020-02019-y

34. Manthravadi S, Shrestha A, Madhusudhana S. Impact of statin use on cancer recurrence and mortality in breast cancer: A systematic review and meta-analysis. *Int J Cancer* (2016) 139(6):1281–8. doi: 10.1002/ijc.30185

35. Sharma M, Tuaine J, McLaren B, Waters DL, Black K, Jones LM, et al. Chemotherapy agents alter plasma lipids in breast cancer patients and show differential effects on lipid metabolism genes in liver cells. *PloS One* (2016) 11(1):e0148049. doi: 10.1371/journal.pone.0148049

36. Wang X, Zhu A, Wang J, Ma F, Liu J, Fan Y, et al. Steroidal aromatase inhibitors have a more favorable effect on lipid profiles than nonsteroidal aromatase inhibitors in postmenopausal women with early breast cancer: A prospective cohort study. *Ther Adv Med Oncol* (2020) 12:1758835920925991. doi: 10.1177/1758835920925991

37. Alomar SA, Găman M-A, Prabahar K, Arafah OA, Almarshood F, Baradwan S, et al. The effect of tamoxifen on the lipid profile in women: A systematic review and meta-analysis of randomized controlled trials. *Exp Gerontol* (2022) 159:111680. doi: 10.1016/j.exger.2021.111680

38. Chen Y, Wen Y-y, Li Z-r, Luo D-l, Zhang X-h. The molecular mechanisms between metabolic syndrome and breast cancer. *Biochem Biophys Res Commun* (2016) 471(4):391–5. doi: 10.1016/j.bbrc.2016.02.034

39. Calle EE, Kaaks R. Overweight, obesity and cancer: Epidemiological evidence and proposed mechanisms. *Nat Rev Cancer* (2004) 4(8):579–91. doi: 10.1038/nrc1408

40. Khandekar MJ, Cohen P, Spiegelman BM. Molecular mechanisms of cancer development in obesity. *Nat Rev Cancer* (2011) 11(12):886–95. doi: 10.1038/nrc3174

41. Mendonça FM, de Sousa FR, Barbosa AL, Martins SC, Araújo RL, Soares R, et al. Metabolic syndrome and risk of cancer: Which link? *Metabolism* (2015) 64(2):182–9. doi: 10.1016/j.metabol.2014.10.008

42. Hauner D, Hauner H. Metabolic syndrome and breast cancer: Is there a link? *Breast Care (Basel)* (2014) 9(4):277–81. doi: 10.1159/000365951

43. Chung SJ, Nagaraju GP, Nagalingam A, Muniraj N, Kuppusamy P, Walker A, et al. Adipoq/Adiponectin induces cytotoxic autophagy in breast cancer cells through Stk11/Lkb1-mediated activation of the ampk-Ulk1 axis. *Autophagy* (2017) 13(8):1386–403. doi: 10.1080/15548627.2017.1332565

44. Zhao P, Xia N, Zhang H, Deng T. The metabolic syndrome is a risk factor for breast cancer: A systematic review and meta-analysis. *Obes Facts* (2020) 13(4):384–96. doi: 10.1159/000507554

45. Alikhani N, Ferguson RD, Novosyadlyy R, Gallagher EJ, Scheinman EJ, Yakar S, et al. Mammary tumor growth and pulmonary metastasis are enhanced in a hyperlipidemic mouse model. *Oncogene* (2013) 32(8):961–7. doi: 10.1038/onc.2012.113

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 29 March 2023
DOI 10.3389/fonc.2023.1128318

OPEN ACCESS

EDITED BY
Maha Mohamed Saber-Ayad,
University of Sharjah, United Arab Emirates

REVIEWED BY
Elizabeth Wellberg,
University of Oklahoma Health Sciences Center, United States
Suzanne Marie Ponik,
University of Wisconsin-Madison, United States

*CORRESPONDENCE
Charlotta Dabrosin
charlotta.dabrosin@liu.se

SPECIALTY SECTION
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

RECEIVED 20 December 2022
ACCEPTED 17 March 2023
PUBLISHED 29 March 2023

CITATION
Ekstrand J, Abrahamsson A, Lundberg P and Dabrosin C (2023) Breast density and estradiol are associated with distinct different expression patterns of metabolic proteins in normal human breast tissue *in vivo*. *Front. Oncol.* 13:1128318. doi: 10.3389/fonc.2023.1128318

Breast density and estradiol are associated with distinct different expression patterns of metabolic proteins in normal human breast tissue *in vivo*

Jimmy Ekstrand[1], Annelie Abrahamsson[1], Peter Lundberg[2,3] and Charlotta Dabrosin[1]*

[1]Department of Oncology and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden, [2]Department of Radiation Physics and Department of Medical and Health Sciences, Linköping University, Linköping, Sweden, [3]Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden

Background: Breast density and exposure to sex steroids are major risk factors for breast cancer. The local microenvironment plays an essential role in progression of breast cancer. Metabolic adaption is a major hallmark of cancer. Whether proteins from the extracellular space regulating metabolism are affected in breast cancer, dense breasts or by estrogen exposure are not yet fully elucidated.

Methods: Women with breast cancer, postmenopausal women with normal breast tissue with varying breast density or premenopausal women with breasts exposed to high levels of estradiol were included in the study. Microdialysis was used to collect proteins from the extracellular space *in vivo* in 73 women; 12 with breast cancer, 42 healthy postmenopausal women with different breast densities, and 19 healthy premenopausal women. Breast density was determined as lean tissue fraction (LTF) using magnetic resonance imaging. Data were evaluated in a murine breast cancer model. We quantified a panel of 92 key proteins regulating metabolism using proximity extension assay.

Results: We report that 29 proteins were upregulated in human breast cancer. In dense breasts 37 proteins were upregulated and 17 of these were similarly regulated as in breast cancer. 32 proteins correlated with LTF. In premenopausal breasts 19 proteins were up-regulated and 9 down-regulated. Of these, 27 correlated to estradiol, a result that was confirmed for most proteins in experimental breast cancer. Only two proteins, pro-cathepsin H and galanin peptide, were similarly regulated in breast cancer, dense- and estrogen exposed breasts.

Conclusions: Metabolic proteins may be targetable for breast cancer prevention. Depending on risk factor, this may, however, require different approaches as breast density and estradiol induce distinct different expression patterns in the

breast. Additionally, metabolic proteins from the extracellular space may indeed be further explored as therapeutic targets for breast cancer treatment.

KEYWORDS

mammography, microdialysis, sex steroids, estradiol, breast density

1 Introduction

Two major independent risk factors for breast cancer are mammographic dense breast tissue and exposure to sex steroids (1, 2). To date, the biological mechanisms that govern these processes remain elusive.

There is a 4-6-fold increased risk of breast cancer for women with dense breast as compared to women with nondense breasts and an inverse relationship with nondense area and risk of the disease has been shown (2). Women with > 50% dense area account for approximately 30% of all breast cancer cases (2). Dense breast tissue is characterized by high amounts of stroma, including collagen, in contrast to nondense breasts where fat is the major component (3). The proportion of breast epithelial cells in normal breast tissue is less than 10% and there are no conclusive data on differences on the quantity or proliferation rate of these cells depending on breast density (3–5).

Another major independent risk factor for breast cancer is exposure to sex steroids including estrogens (1). No association between circulating estrogen levels and breast density has been observed (2). Sex steroids, including estrogens, play a critical role in the regulation of the epithelial cell proliferation and apoptosis in normal breast tissue (6). However, only a minor fraction of the epithelial cells proliferates during the menstrual cycle (6). Additionally, the response in epithelial cells is highly dependent upon epithelial-stroma interactions in the microenvironment, which also are affected by estrogens (7). Estradiol has profound effects on the local immune microenvironment, angiogenesis, and fibroblasts function in the breast (8–12). Thus, it is rather the microenvironment surrounding the epithelial cells than the epithelial cells alone that determines the risk of breast cancer initiation and progression as an activated stroma is a prerequisite for tumor formation.

Metabolic reprogramming is included in the hallmarks of cancer (13). During progression cancer acquires various metabolic phenotypes in cooperation with stromal cells (14). These metabolic properties may enable cancer cells in dormant tumors to increase cell survival, which supports hyperplastic growth, increase invasion capacities, and evade immune surveillance (15). The metabolic phenotype in a tissue or in cancer is a result of complex interactions of intrinsic processes in epithelial or cancer cells and extrinsic factors in the microenvironment (14). The importance of cancer metabolism is supported by recent data suggesting that therapeutic targets in the microenvironment, including metabolism, may be more important than oncogenes (16).

To date it is unclear whether normal tissues with intrinsically increased risk of cancer express altered metabolic properties that support tumor initiation and early cancer progression. Whether proteins involved in metabolic pathways are affected in normal human breasts by two major risk factors for breast cancer, breast density and estradiol exposure, are previously not explored. Additionally, whether proteins sampled from the extracellular space involved in metabolism are altered in human breast cancer *in vivo* is previously not determined. Here we used microdialysis to sample proteins directly from live breast tissue. The advantage of this approach is that proteins can be quantified directly in the target organ, albeit its invasiveness, as compared to blood levels that will reflect circulating levels originating from many different organs in the body.

By using a panel of 92 key proteins involved in metabolism we explored whether levels were altered in the extracellular local microenvironment *in vivo* in breast cancer compared to normal breast tissue, in normal breast tissue depending on breast density or in breasts exposed to high levels of estradiol compared to breasts with low estradiol levels.

2 Materials and methods

2.1 Subjects

Previously collected and biobanked samples from different cohorts were used in the exploratory clinical study. The Regional Ethical Review Board of Linköping approved the collections which were carried out in accordance with the Declaration of Helsinki. All subjects gave informed consent. A total of 73 women were included.

Twelve postmenopausal women (ages 52-86 years) with breast cancer were investigated with microdialysis before surgery. All breast cancers were estrogen receptor (ER) positive and human epidermal growth factor 2 (HER-2) negative.

Healthy postmenopausal women from the mammography screening program at Linköping University Hospital that were categorized according to the Breast Imaging Reporting and Data System (BI-RADS) as either entirely fatty nondense (BI-RADS A) or extremely dense (BI-RADS D) were invited to the study (17). Forty-two healthy postmenopausal women (ages 55–74 years) were consecutively recruited for the study. The women were subjected to magnetic resonance imaging (MRI) (18, 19). On the MRI lean tissue fraction (LTF), as a continuous measure of breast density, was calculated in a volume selection of 20 x 20 x 20 mm in the upper

lateral quadrant of the left breast, as previously described (19, 20). In brief, 1.5 T Achieva MR scanner (Philips Healthcare, Best, Netherlands) using a dual breast seven-element breast coil was used. Water- and fat separated MR images were computed, as previously described (21), in summary: axial 3D 6-echo turbo field echo MRI images, anterior-posterior frequency encoding, first TE at 2.3 ms and ΔTE of 2.3 ms, TR 15.4 ms, flip angle 10°, 300×300×150 mm^3 field of view, 200×200 scan matrix and 3 mm slice thickness. LTF was computed as the ratio of lean tissue volume to total volume.

After a second review of the mammograms, it was noticed that two women had been miscategorized and intermediate with respect to breast density. These two women were not included in the analyses of dense vs. nondense breast but included in the correlation analyses.

For the premenopausal group, 19 nulliparous women (ages 20–32 years) with a history of regular menstrual cycles (cycle length, 27–34 days) were included. All of these were investigated with microdialysis in the luteal phase of the menstrual cycle.

None of the healthy volunteer women had a history of breast cancer or were currently using (or had used within the past 3 months) hormone replacement therapy, sex steroid-containing contraceptives, anti-estrogen therapies, including selective estrogen receptor modulators, or degraders.

2.2 Microdialysis procedure

Prior to insertion of the microdialysis catheters 0.5 mL lidocaine (10 mg/mL) was administrated intracutaneously. Microdialysis catheters (M Dialysis AB, Stockholm, Sweden), which consisted of a tubular dialysis membrane (diameter 0.52 mm, 100,000 atomic mass cut-off) glued to the end of a double-lumen tube were inserted *via* a splittable introducer (M Dialysis AB), connected to a microinfusion pump (M Dialysis AB) and perfused with 154 mmol/L NaCl and 60 g/L hydroxyethyl starch (Voluven®; Fresenius Kabi, Uppsala, Sweden), at 0.5 μL/min. The women with ongoing breast cancer were investigated with 10 mm long membranes; one catheter was inserted within the cancer tissue and the other into normal adjacent breast tissue. The healthy volunteer women were investigated with 20 mm long microdialysis membranes; one was placed in the upper lateral quadrant of the left breast and directed towards the nipple as previously described (9, 22–29). The premenopausal women were subjected to microdialysis in the luteal phase of the menstrual cycle. The microdialysis catheter was placed in the same quadrant where LTF was determined.

After a 60 min equilibration period, the outgoing perfusate was stored at -80°C for subsequent analysis.

2.3 Breast cancer model

The Institutional Animal Ethics Committee at Linköping University approved this study, which conformed to regulatory standards of animal care. Oophorectomized athymic mice (Balb/C-nu/nu, 6-8 weeks old, Scanbur, Sweden) were housed at Linköping University in ventilated cages with a light/dark cycle of 12/12 hours with rodent chow and water available *ad libitum*. Mice were anesthetized *via* intraperitoneal (i.p.) injection of ketamine/xylazine and implanted with a s.c. 3 mm pellet containing either 17β-estradiol (0.18 mg/60-day release, Innovative Research of America, Sarasota, FL, USA) or placebo. The active pellet releases serum concentrations of 150-250 pM estradiol. 5×10^6 MCF-7 cells were injected into the dorsal mammary fat pads in 200 μL PBS. MCF-7 cells require estrogen for tumor formation and growth in mice, therefore, a non-estrogen control group is not possible to achieve. When tumors reached ≈20 mm^2 in size the mice were treated with fulvestrant (5 mg/mouse twice per week, s.c.) in addition to the estradiol exposure.

2.4 Microdialysis in mice

Tumor-bearing mice with size-matched tumors were anesthetized with i.p. injections of ketamine/xylazine and maintained by repeated s.c. injections of ketamine/xylazine. Body temperatures were maintained using a heat lamp. Microdialysis probes with 4-mm membranes (CMA 20, 100 kDa cutoff; CMA Microdialysis AB, Kista, Sweden) were inserted into tumor tissue and connected to a microdialysis pump (CMA 102; CMA Microdialysis AB) perfused at 0.6 μl/min with 154 mmol/L NaCl and 60g/L hydroxyethyl starch (Voluven®; Fresenius Kabi, Uppsala, Sweden), as previously described (30, 31). After a 60 min equilibrium period, outgoing perfusates (*i.e.*, microdialysates) were collected and stored at -80°C for subsequent analysis.

2.5 Protein quantifications

The Metabolism panel (Olink Bioscience, Uppsala Sweden) was used for the microdialysis samples. Proteins included in the panel are listed in Supplementary Table 1. Although some of these proteins are considered to be of cellular origin, only one protein, ANXA4, was undetectable in human plasma during the development of the assay (Olink Bioscience, Uppsala Sweden). Thus 91 out of the 92 proteins were quantified circulating in plasma of healthy individuals during the validation of the assay; https://www.olink.com/content/uploads/2021/09/olink-metabolism-validation-data-v2.0.pdf. In the present study, microdialysis samples were analyzed with multiplex proximity extension assay (PEA, Olink Bioscience, Uppsala Sweden) as previously described (32–34). In brief, 1 μL microdialysis sample was incubated with proximity antibody pairs tagged with DNA reporter molecules. The DNA tails formed an amplicon by proximity extension, which was quantified by high-throughput real-time PCR (BioMark™ HD System; Fluidigm Corporation, South San Francisco, CA, USA). The generated fluorescent signal correlates with protein abundance by quantitation cycles (Cq)

produced by the BioMark Real-Time PCR Software. To minimize variation within and between runs, the data were normalized using both an internal control (extension control) and an interplate control and transformed using a predetermined correction factor. The pre-processed data were provided in the arbitrary unit normalized protein expression (NPX) on a $\log_2$ scale, which were then linearized by using the formula 2^{NPX}. A high NPX value corresponds to high protein concentrations. Values represented a relative quantification meaning that no comparison of absolute levels between different proteins could be made.

Protein-protein interactions were analyzed using the STRING data base.

2.6 Estradiol analysis

Estradiol levels were analyzed using a high sensitivity immunoassay kit (DRG International, Springfield Township, NJ, USA).

2.7 Statistical analyses

Statistical analyses were performed using nonparametric Wilcoxon matched-pairs signed rank tests or Kruskal Wallis tests followed by unpaired Mann-Whitney U tests when more than two groups were compared as the data was non-normally distributed. Spearman's correlation test was used for calculations of correlations. A $P<0.05$ was considered statistically significant. Statistics were performed with Prism 9.0 (GraphPad, San Diego, CA, USA).

3 Results

There were no statistically significant differences in BMI, age or local breast estradiol levels between dense and nondense group; BMI (mean ± SD) 24 ± 3.3 vs 25 ± 3.2, age (years, mean ± SD) 64 ± 5.8 vs 65 ± 5.2 respectively and local breast estradiol (pmol/L mean ± SD) 40 ± 14 vs 40 ± 12 respectively.

In the premenopausal group BMI (mean ± SD) was 24 ± 1.5 and local breast estradiol levels (pmol/L mean ± SD) were 196 ± 40.

3.1 Distinct patterns of proteins in breast cancer, postmenopausal dense- and premenopausal breasts

In the first set of analyses, we determined whether any of the proteins were significantly altered in breast cancer *in vivo* as a measure of the biological relevance for human disease. As shown in Figure 1A, 29 proteins, out of the panel of 92, were significantly up-regulated after FDR correction in breast cancers as compared to normal adjacent breast tissue. Thereafter we investigated whether dense breast tissue without any pathology as compared to normal nondense breasts in postmenopausal women exhibited any alterations of the 92 proteins. As shown in Figure 1B, 37 proteins were up-regulated in dense breasts. In premenopausal breast tissue, which per definition is dense, 28 proteins were significantly changed compared to postmenopausal dense breast, 19 were up-regulated and 9 down-regulated, Figure 1C. In Figure 2 all individual proteins that were changed in breast cancer are depicted. In Figures 3, 4 altered proteins in dense breasts and premenopausal breasts respectively are shown.

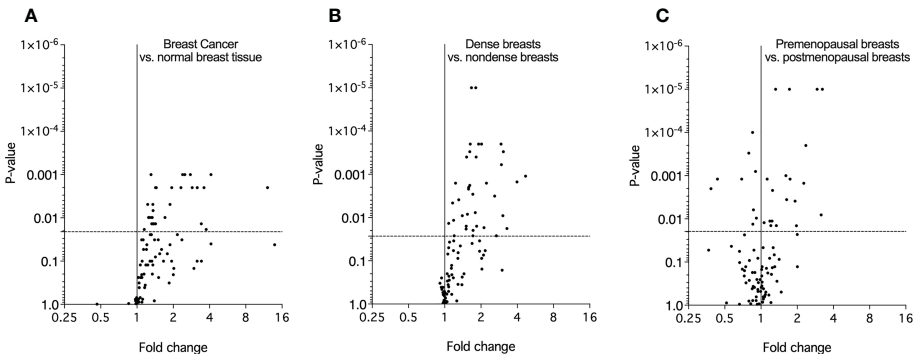


FIGURE 1
Molecular characterization of the extracellular microenvironment *in vivo* in breast cancer, postmenopausal dense breasts vs. nondense breasts, and in premenopausal breasts vs postmenopausal dense breasts. Microdialysis was performed for *in vivo* collection of proteins from the extracellular space that were quantified as described in the Materials and Methods. Volcano plots illustrate the $\log_{10}$ statistical significance (FDR-adjusted p-value) in relation to the $\log_2$ fold change of 92 proteins involved in metabolism. **(A)**. Twelve patients with breast cancer underwent microdialysis one day prior to their surgery. One catheter was inserted into the breast cancer and another catheter was inserted into normal adjacent breast tissue. Fold change of proteins in cancer tissue as compared to normal adjacent breast tissue. The dotted line indicates the FDR-adjusted *p*-value, <0.021. **(B)**. 40 healthy postmenopausal women with dense breast tissue (n=20) or nondense breast tissue (n=20) were subjected to microdialysis in the upper lateral quadrant of the left breast. Fold change was calculated from the median value of proteins in dense vs. nondense breasts. The dotted line indicates the FDR-adjusted *p*-value, <0.027. **(C)**. 19 premenopausal women were subjected to microdialysis in the upper lateral quadrant of the left breast in the luteal phase of the menstrual cycle. Fold change was calculated from median values of proteins in premenopausal breast vs. postmenopausal dense breasts (n=20). The dotted line indicates the FDR-adjusted *p*-value, <0.02.

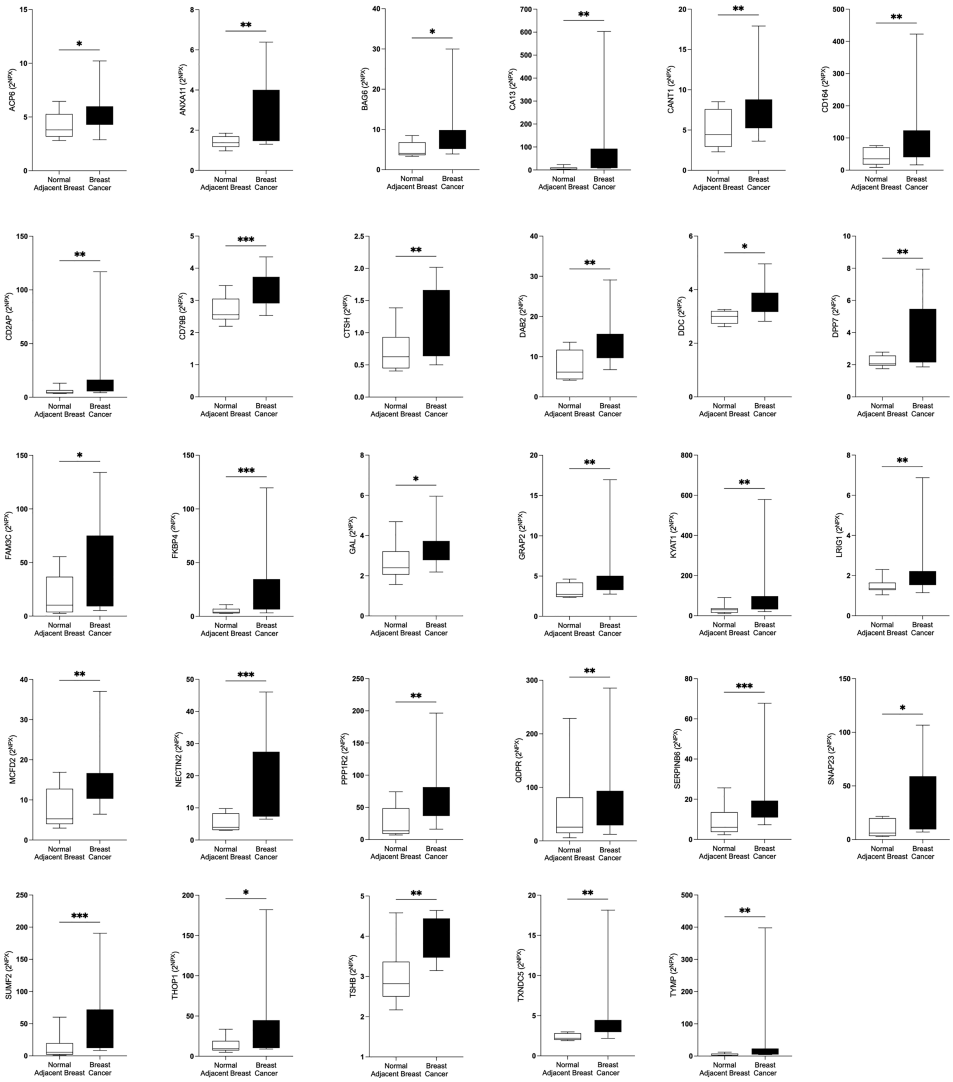


FIGURE 2
Significantly altered extracellular levels of proteins regulating metabolism in human estrogen receptor positive (ER+) breast cancer *in vivo*. Twelve patients with breast cancer underwent microdialysis one day prior to their surgery. One catheter was inserted into the breast cancer and another catheter was inserted into normal adjacent breast tissue for *in vivo* collection of proteins from the extracellular space. The levels of the significantly altered proteins that were depicted in Figure 1A are shown. Data are presented as box plots with whiskers of min and max values. **P*<0.05, ***P*<0.01, ****P*<0.001.

Protein-protein interactions are shown in Supplementary Figures 1A-C.

3.2 Similar patterns of altered proteins in dense breast as in breast cancer

Of the 29 significantly up-regulated proteins in breast cancer, 17 were also up-regulated in postmenopausal dense breasts as compared to postmenopausal nondense breasts. However, in premenopausal breast, who have by definition dense breasts, as compared to postmenopausal dense breast three of these proteins were up-regulated whereas two were down-regulated, Figure 5A. Only two proteins exhibited similar alteration in the three different cohorts; Pro-cathepsin H (CTSH) and Galanin peptides (GAL), which were up-regulated in all groups, Figure 5B.

3.3 Distinct different patterns of expression levels of protein in dense breast and premenopausal breast

In the next analysis we compared which proteins that were up-regulated in postmenopausal dense breast tissue as compared to postmenopausal nondense breasts. As shown in Figure 6A 37 proteins were, after FDR correction, significantly up-regulated in dense breast. In premenopausal breast 28 proteins were significantly altered; 19 were up-regulated and 9 significantly down-regulated, Figure 6A. Seven proteins were shared between the groups; CTSH, GAL, leukocyte immunoglobulin-like receptor subfamily A member 5 (LILRA5), paired immunoglobulin-like type 2 receptor beta (PILRB), and syndecan-4 (SDC4), which were up-regulated in both tissues. Two proteins were up-regulated in dense breast but down-regulated in premenopausal breasts: N-terminal prohormone

10.3389/fonc.2023.1128318

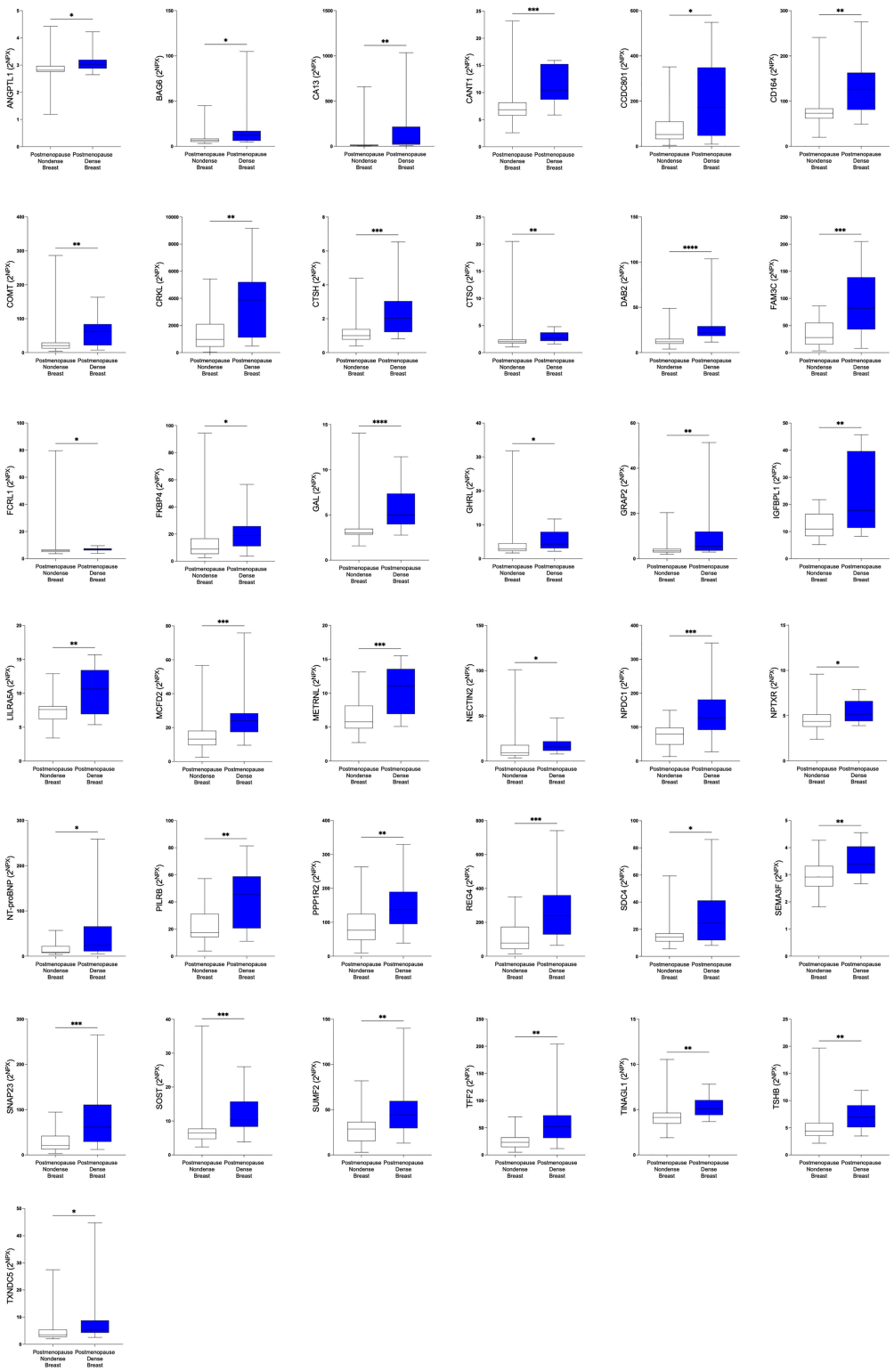


FIGURE 3
Significantly altered extracellular levels of proteins regulating metabolism in postmenopausal dense and nondense breast tissue. 40 healthy postmenopausal women with dense breast tissue (n=20) or nondense breast tissue (n=20) were subjected to microdialysis in the upper lateral quadrant of the left breast for *in vivo* collection of proteins from the extracellular space. The levels of the significantly altered proteins that were depicted in Figure 1B are shown. Data are presented as box plots with whiskers of min and max values. *P<0.05, **P<0.01, ***P<0.001.

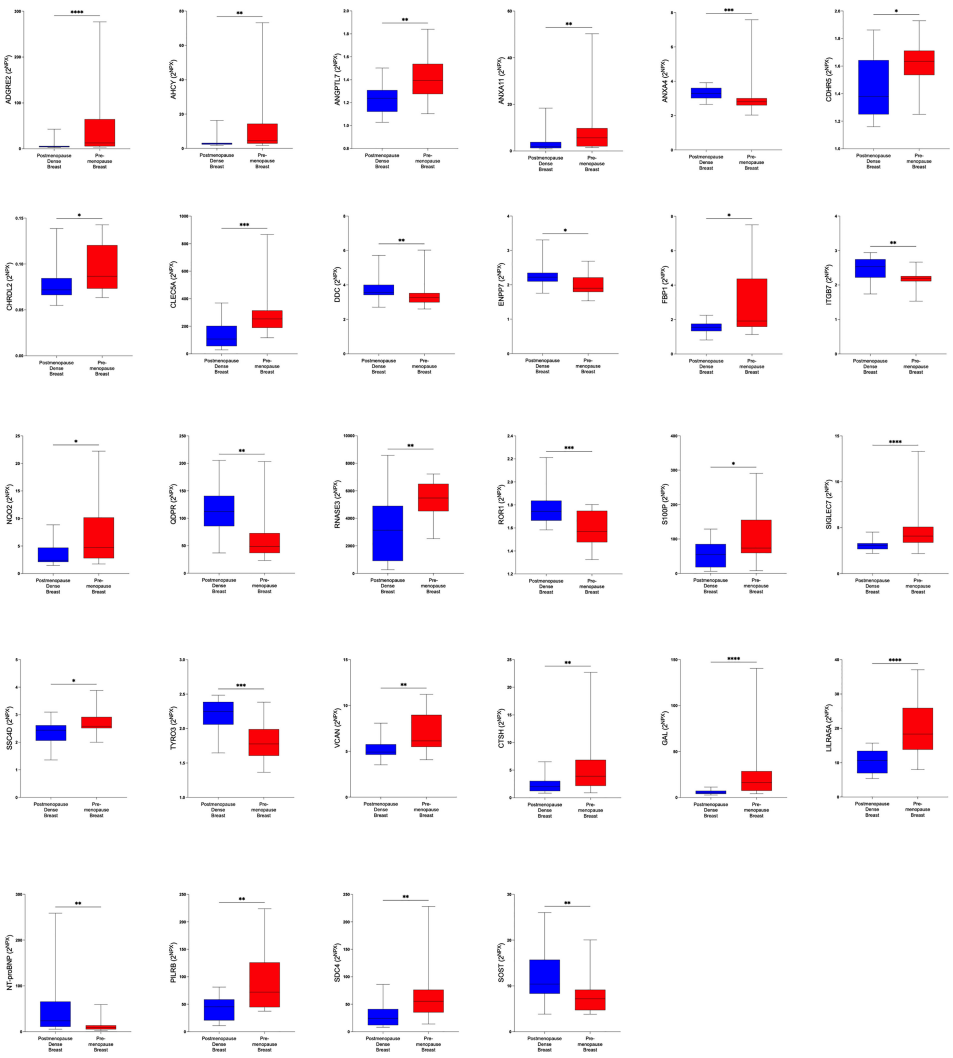


FIGURE 4
Significantly altered extracellular levels of proteins regulating metabolism in premenopausal breasts and postmenopausal dense breasts. 19 healthy premenopausal women in the luteal phase of the menstrual cycle and 20 healthy postmenopausal women with dense breast tissue were subjected to microdialysis in the upper lateral quadrant of the left breast. The levels of the significantly altered proteins that were depicted in Figure 1C are shown. Data are presented as box plots with whiskers of min and max values.*P<0.05, **P<0.01, ***P<0.001.

of brain natriuretic peptide (NT-ProBNP) and sclerostin (SOST), Figure 6B.

3.4 Correlations with breast density and estradiol

Thereafter we wanted to investigate whether breast density and estradiol correlated with the proteins that were significantly altered in the breasts as this would strengthen an involvement of these two measures on the regulation of proteins in the extracellular space. For breast density, we used the precise continuous measure of density calculated from MRI namely LTF, and for estradiol we used the local breast tissue levels. As shown in Figure 7A, 32 out of the 37 proteins that were significantly altered in dense breast also correlated significantly with LTF supporting a role of breast density in the regulation of these proteins. In Figure 7B, 27 out of 28 proteins that were significantly up- or down-regulated in premenopausal breast also exhibited significant correlations with estradiol levels in the breast.

As CTSH has been suggested to regulate GAL we tested if these two proteins correlated in our data set. Indeed, a strong positive correlation was found, Spearman's r=0.857, p<0.00001.

3.5 An E2 dependent regulation of proteins was corroborated in experimental breast cancer

Next, we wanted to explore whether the proteins that were associated with local estradiol levels in normal human breast tissue were estrogen regulated in experimental ER+ breast cancer in mice. 18 of the 27 proteins that correlated with estradiol in normal human breast tissue were quantifiable in the murine microdialysis samples.

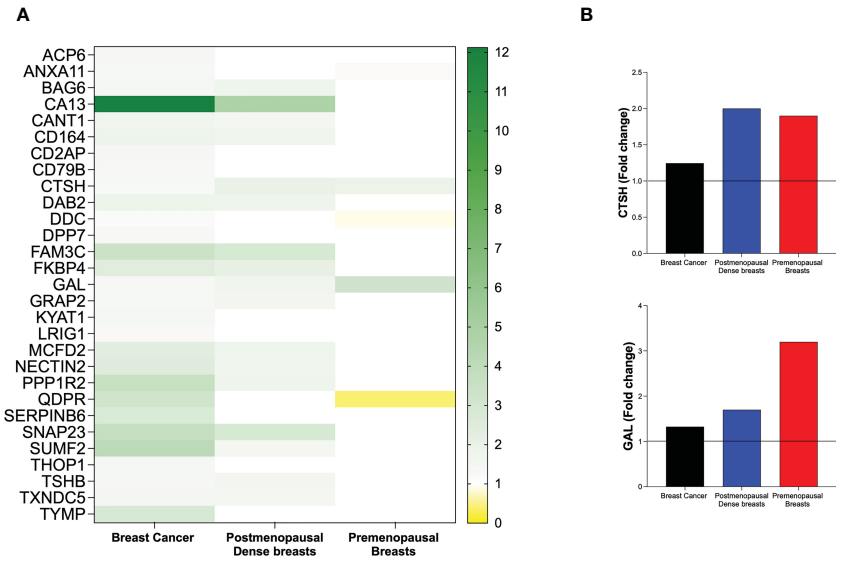


FIGURE 5
Heat map of extracellular *in vivo* metabolic proteins that were identified as significantly altered in human breast cancer. Microdialysis was performed for *in vivo* collection of proteins from the extracellular space that were quantified in as described in the Materials and Methods in three cohorts of women; 12 with ER+ breast cancer, 40 postmenopausal healthy women with dense (n=20) or nondense (n=20) breast tissue, and 19 premenopausal women investigated in the luteal phase of the menstrual cycle. **(A)** Left column depicts all 29 proteins that were identified as up-regulated in human estrogen receptor positive breast cancer patients as compared to normal adjacent breast tissue. Mid column depicts the regulation of the proteins in postmenopausal dense breast tissue vs. postmenopausal nondense breast. Right column depicts the regulation of the proteins in premenopausal breast tissue vs. postmenopausal dense breast tissue. **(B)** Fold change of the only two proteins that were similarly regulated in the three cohorts: Pro-cathepsin H (CTSH) and Galanin peptides (GAL).

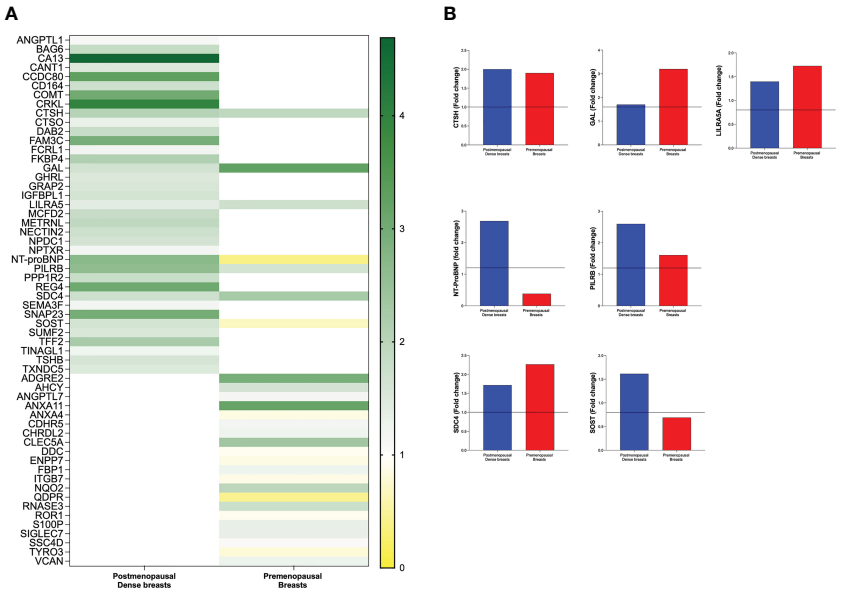


FIGURE 6
Distinct different patterns of expression levels of protein in dense breasts and premenopausal breasts. Microdialysis was performed for *in vivo* collection of proteins from the extracellular space that were quantified in as described in the Materials and Methods in 40 postmenopausal healthy women with dense (n=20) or nondense (n=20) breast tissue, and 19 premenopausal women investigated in the luteal phase of the menstrual cycle. **(A)** Left column depicts 37 proteins that were identified to be significantly altered in postmenopausal dense breast tissue vs. postmenopausal nondense breast and right column depicts the 28 proteins that were identified to be significantly altered in premenopausal breast tissue vs. postmenopausal dense breast tissue. **(B)** The seven that were significantly altered in both postmenopausal dense breasts vs. nondense breasts and in premenopausal breasts vs. postmenopausal dense breasts. Five proteins were up-regulated in both cohorts whereas two proteins were up-regulated in dense breasts and down-regulated in premenopausal breasts.

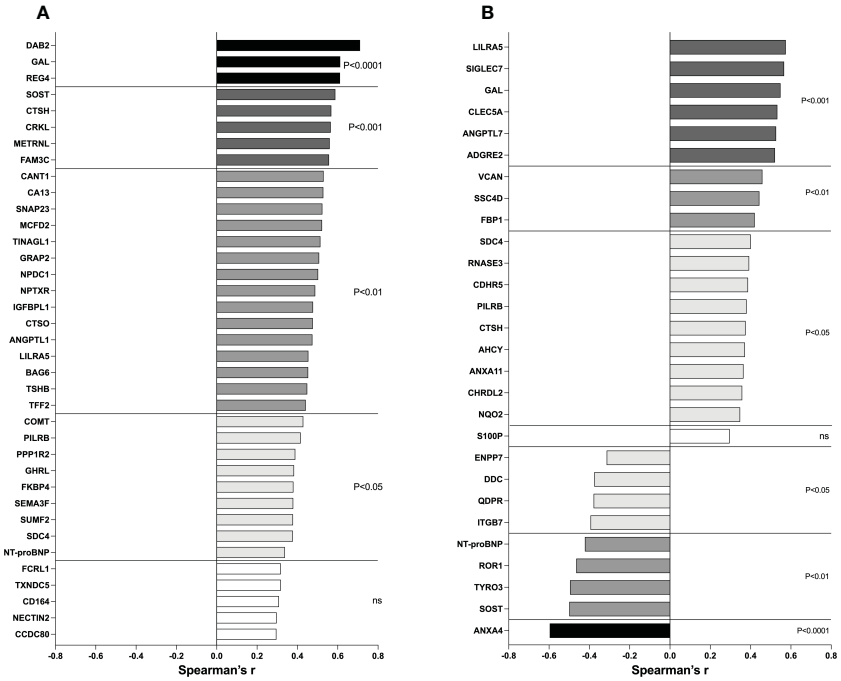


FIGURE 7
Correlations between metabolic proteins and LTF and estradiol. **(A)** Postmenopausal women underwent MRI and microdialysis for collection of proteins from the extracellular space as described in the Materials and Methods. Correlation analysis between proteins and breast density (LTF) was performed. **(B)** Postmenopausal women with dense breast tissue and premenopausal women underwent microdialysis as described in the Materials and Methods. Correlation analysis between proteins and local breast estradiol was performed. Bars represent Spearman's rank correlation coefficients. Colored bars indicate statistical significance, and white bars indicate ns=not significant.

The following proteins were below the lowest level of detection in all samples; sialic acid-binding Ig-like lectin 7 (SIGLEC7), c-type lectin domain family 5 member A (CLEC5A), eosinophil cationic protein (RNASE3), paired immunoglobulin-like type 2 receptor beta (PILRB), chordin-like protein 2 (CHRDL2), ribosyldihydronicotinamide dehydrogenase (NQO2), NT-proBNP, inactive tyrosine-protein kinase transmembrane receptor ROR1 (ROR1), SOST, annexin A4 (ANXA4) and thus, could not be analyzed for estrogen dependency.

As shown in Figure 8A, eight of the proteins that were significantly positive correlated with estradiol were downregulated by the anti-estrogen fulvestrant therapy in mice; GAL, versican core protein (VCAN), scavenger receptor cysteine-rich domain-containing group B protein (SSC4D), fructose-1,6-bisphosphatase 1 (FBP1), SDC4, cadherin-related family member 5 (CDHR5), adenosylhomocysteinase (AHCY), and annexin A11 (ANXA11). No changes were detected after fulvestrant treatment of LILRA5, angiopoietin-related protein 7 (ANGPTL7), adhesion G-protein coupled receptor G2 (ADGRG2), CTSH.

Protein S100-P (S100P) was up-regulated in premenopausal breast tissue but failed to show any correlation with local breast E2 levels. In the murine tumors S100P levels were unaffected by fulvestrant treatment, Figure 8B.

Of the proteins that correlated negatively with estradiol in human tissue *in vivo* up-regulations by fulvestrant were detected for ectonucleotide pyrophosphatase/phosphodiesterase family member 7 (ENPP7), aromatic-L-amino-acid decarboxylase (DDC), dihydropteridine reductase (QDPR), integrin beta-7 (ITGB7), and tyrosine-protein kinase receptor TYRO3TYRO3, Figure 8C. Thus, of the 18 detectable proteins in the murine samples, 14 corroborated the results from human breast tissue.

4 Discussion

Here we quantified *in situ* levels of a panel of 92 key proteins for metabolism, in estrogen receptor positive (ER+) breast cancer in women, in postmenopausal women with dense or nondense breast tissue and in premenopausal women. Our data suggest that dense breast tissue expresses a similar pattern of metabolic proteins as human breast cancer whereas estradiol induces a distinctly different pattern of these proteins in the breast. An estrogen dependent regulation of several of these proteins was corroborated in a murine model of ER+ breast cancer. Thus, preventive measures against breast cancer may require different approaches depending on the risk factor.

Metabolic adaptation is one major hallmark in cancer. In addition to energy metabolism, several other metabolic pathways are affected. Understanding these tissue-specific metabolic phenotypes is fundamental for the discovery of novel therapeutic targets. How proteins in the extracellular space involved in metabolism are affected in human breast cancer is less studied. Here we show that several of these proteins are indeed up-regulated in human ER+ breast cancer *in vivo*. Additionally, our data showed that dense normal breast tissue exhibits similar expression pattern

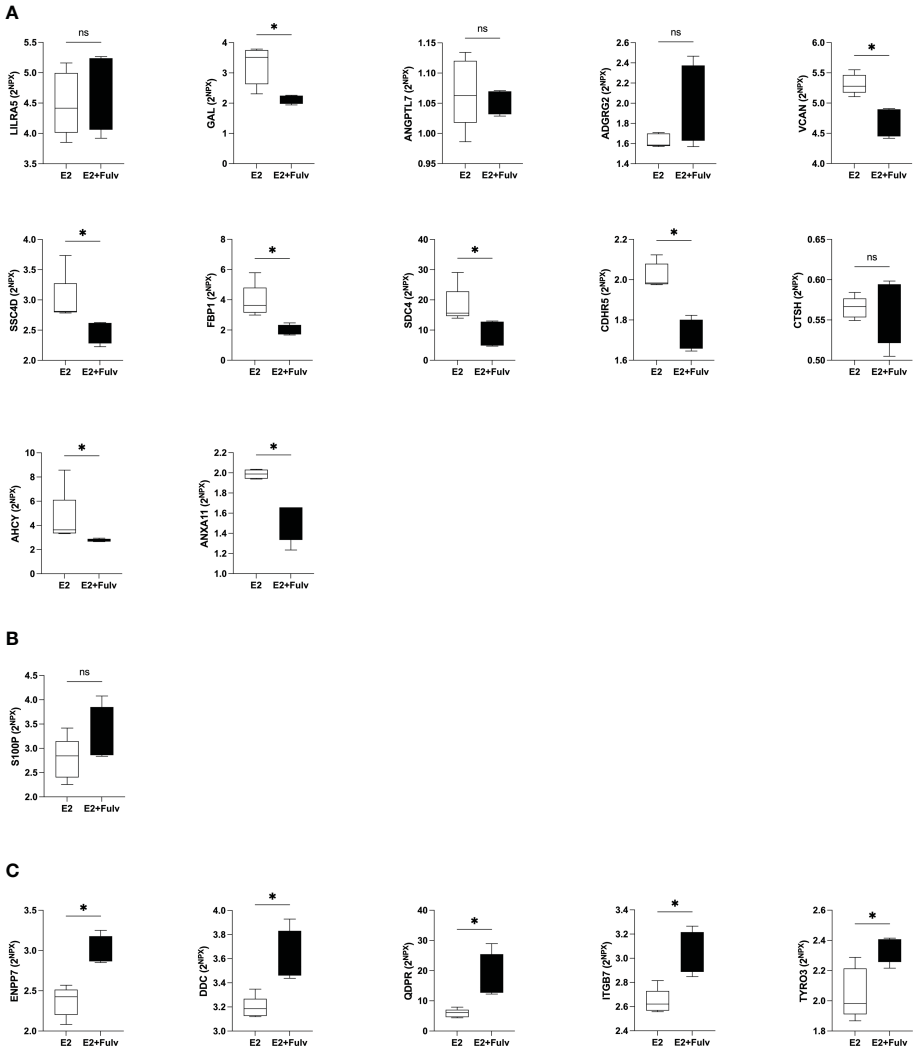


FIGURE 8
Hormonal regulation of extracellular metabolic proteins in estrogen receptor (ER+) experimental breast cancer. Oophorectomized athymic mice supplemented with physiological levels of estradiol (E2) were injected with MCF-7 cells into the dorsal mammary fat pads. At similar tumor sizes, mice either continued with E2 or were additionally treated with fulvestrant (E2+Fulv) (5 mg/mouse every 3 days, s.c.). Size-matched tumors from the different treatment groups underwent microdialysis for sampling of proteins from the extracellular space *in vivo*. **(A)** Proteins that were significantly positively correlated with estradiol in human breast tissue. **(B)** Protein that was up-regulated in premenopausal breast tissue but failed to show any correlation with local breast E2 levels. **(C)** Proteins that were significantly negatively correlated with estradiol in human tissue. *$P<0.05$, ns= not significant.

of these proteins as those found in breast cancer. Contrary, a distinct different pattern of affected proteins was shown to be estrogen dependent in normal human breast tissue.

Our data shows that the proteins with the highest fold change in breast cancer were carbonic anhydrase 13 (CA13), sulfatase-modifying factor 2 (SUMF2), and synaptosomal-associated protein 23 (SNAP23). All of these three proteins are involved in three entirely different pathways of primary metabolism; CA13, a carbonic anhydrase, contributes to cell respiration, lipogenesis and gluconeogenesis; SUMF2 inhibits the activity of SUMF1, which in turn is involved in cysteine conversions and modulations of cell metabolism; and SNAP23 a protein which contributes to membrane transport and vesicle trafficking important for insulin activities (35–37). These three proteins were also up-regulated in dense breast tissue with positive correlations with local LTF in postmenopausal women. In contrast, no changes in the regulation of these three proteins were detected in premenopausal breasts as compared to postmenopausal breast and no correlations to estradiol were revealed. The roles of these proteins in cancer progression remains elusive, but our data nevertheless suggest that they may have a role for breast cancer progression *per se* and possibly also in early stages of cancer development in dense breast tissue.

Of the 92 quantified proteins only two shared similar up-regulations in breast cancer, dense breast tissue, and in premenopausal breasts namely CTSH and GAL. CTSH is a lysosomal cysteine protease that also has been detected in secretory vesicles (38). Lysosomal and extracellular cathepsins are important for protein degradation and control of nutrient sensing and metabolic homeostasis (39). Cathepsins have also a vital role in controlling energy metabolism including the processing of

lipoproteins (40–42). The role of cathepsins, including CTSH, in cancer progression is somewhat contradictory. High serum levels of CTSH in lung- and thyroid cancer have been associated with good prognosis whereas high blood and tissue levels in lung- and gall bladder cancer are associated with decreased survival (43–46). In prostate cancer CTSH has been shown to increase metastases *via* modulations of integrin activation (47). The data of CTSH in breast cancer is sparse, but our results indeed show that this cathepsin may play a role in human breast cancer.

GAL was also up-regulated in all three different settings: ER+ breast cancer, postmenopausal dense breasts, and premenopausal breasts. GAL's role in metabolism is diverse including glucose uptake, insulin sensitivity, and growth hormone release (48). GAL expression alongside its receptors have been detected in several cancer forms including breast cancer and both pro- and anti-tumorigenic actions have been suggested (49). Previous studies have shown an estrogen dependent regulation of GAL (50). Our data support such regulation both in normal breast tissue and ER+ breast cancer as a significant positive correlation between GAL and estradiol was detected normal human breast tissue and a significantly decreased levels were found after fulvestrant therapy in experimental ER+ breast cancer in mice. Interestingly, CTSH in extracellular vesicles has been shown to be important for GAL production in brain cortex of mice (38). Our data suggest that this may indeed be the case also in human breast tissue as these two proteins correlated significantly *in vivo*.

There were 37 proteins that were significantly up-regulated in dense breast as compared to nondense breast. Of these 37, 32 correlated significantly with the continuous density measure LTF. Disabled homolog 2 (DAB2) and regenerating islet-derived protein 4 (REG4) were two proteins that were strongly associated with breast density. DAB2 is a multifunctional protein involved in many signaling pathways regulating homeostasis in cells including lipoprotein receptor regulations (51, 52). It has been suggested to be tumor suppressor but its effects on immune regulation may also indicate pro-tumorigenic effects (51).

DAB2 has also been shown to increase skin fibrosis in mice, which is in line with our results of increased levels in dense breasts, which contain elevated levels of collagen (53). REG4 is another multifunctional protein involved in cell cycle regulation, glycolytic metabolism, and diabetes (54, 55). REG4 up-regulation has been associated with cancer in the GI-tract including pancreas where it stimulates proliferation and inhibits apotosis (55). Low levels have been detected in breast cancer, which is corroborated by our data as no increase was detected in ER+ breast cancers. However, REG4 was three times higher in dense breasts as compared to nondense breast, indeed suggesting a role in normal breast physiology and possibly in breast cancer initiation and progression.

Estrogen exposure play an essential role in the control of metabolism throughout the body by affecting everything from food intake, fat cell function and distribution, peripheral insulin sensitivity and β-cell function in the pancreas, to lipid metabolism locally in blood vessels and in the liver (56). In the breast estrogen also affects the proliferation, which is related to metabolism. Even though there are limited data on hormonal regulations of the metabolic proteins analyzed in the present study it was expected that several of the proteins would correlate to estradiol. In premenopausal breast, 27 out of 28 proteins that were up-regulated as compared to postmenopausal breast correlated significantly with E2 levels. Of these, a strong negative correlation was found for ANXA4 whereas LILRA5 and SIGLEC7 were positively associated with E2. ANXA4 has an important function in membrane permeability and membrane trafficking (57). AXA4 is a negative regulator for adenylate cyclase type 5, which in turn is important for insulin secretion and cAMP production (58). Blood levels of ANXA4 have been shown to be increased in hepatocellular cancer and up-regulations in tumor tissue has been detected in colorectal- and ovarian cancer suggesting a tumor promoting effect of this protein (57). Studies of ANXA4 in breast cancer are sparse. In human endometrium a progesterone dependent up-regulation of ANXA4 has been revealed whereas estradiol did not change its levels (59). Surprisingly, our data indicates that estradiol down-regulates ANXA4 in normal human breast tissue. This suggests that ANXA4 may not be associated with estrogen dependent initiation of breast cancer. As ANXA4 was undetectable in murine tumors we could not confirm the human data. Further studies of its role in breast cancer are, however, warranted.

LILRA5, which was positively correlated with estradiol, may be important for the regulation of innate immune response as it is expressed on neutrophils. There are, however, no comprehensive analyses of its physiological function including metabolism, possible role in cancer or whether estradiol may be involved in its regulation (60). Our data suggests that estradiol may play a role in the regulation of this protein in normal breast, but this was not corroborated in ER+ breast cancer in mice. Further studies are warranted to conclude the role of estradiol in the regulation of LILRA5. The other protein that showed a strong positive correlation with estradiol was SIGLEC7. This protein may affect metabolism by interfering with the glycosylation pattern of proteins and lipids, which are important for normal cell homeostasis and cell turnover (61). In cancer, SIGLEC7 is expressed on most immune cells and can favor immune evasion in cancer, in addition to its contribution to tumor growth and progression (61). Regarding regulation of SIGLEC7 by estradiol little is known. Genomic data from rat brain has suggested a two-fold increased expression of SIGLEC1 by estradiol exposure suggesting that this protein family may be under hormonal control, which are in line with our data from breast tissue (62). SIGLEC7 was undetectable in the murine tumors.

Of the proteins that correlated with estradiol in human breast tissue, 18 were detectable in the experimental set up. Of these 18, 14 corroborated the human results suggesting that estradiol may indeed be a player of the regulation of metabolism locally in breast tissue. One limitation of the murine data is that the model is immune deficient.

The microenvironment is a rich milieu, a tissue ecosystem, in which all cells contribute to the total repertoire of proteins regulating the inter-cellular crosstalk (16, 63, 64). This intercellular crosstalk includes extracellular soluble proteins that may not be distinguished by standard molecular techniques for whole tissues such as biopsies. Microdialysis enables *in vivo* sampling of extracellular molecules that mirrors this crosstalk directly from the tissue of interest unlike blood samples that

reflect secreted proteins from all different organs combined. Microdialysis is therefore a useful tool for explorative clinical studies elucidating normal physiology or pathology for the discovery of novel targets that could be further investigated. The invasiveness of microdialysis is comparable to a core biopsy, and sampling is time consuming. Due to this, the technique is unsuitable for clinical use in healthy women, such as for screening purposes.

During the progression from premalignant lesions to locally invasive cancers, intrinsic cancer cell alterations alongside microenvironmental cues may induce metabolic changes that enable cancer progression. Interactions between cancer cells with the surrounding cells in the microenvironment shape the metabolic milieu that can affect cancer progression (14). Previously metabolic reprogramming has mostly been associated with aberrant glycolysis, the Warburg effect. However, as recently reviewed, metabolic reprogramming in cancer is a complex biological trait that includes many different pathways that evolve into different metabolic phenotypes during cancer progression (15). Understanding this complex reprogramming of metabolism is necessary for the discovery of actionable therapeutic targets.

We conclude that, out of 92 proteins related to metabolism, 29 were significantly altered in human breast cancer *in vivo*. In normal breast tissue, breast density and estradiol induced two distinct different patterns of these metabolic proteins. Surprisingly, tissue density seems to be more important than estradiol for the local control of the proteins as 37 proteins were associated with density whereas only 27 were associated with estradiol. Our data, which need to be confirmed in larger cohorts of patients, suggest that metabolic proteins may indeed represent targets that warrant further for treatment and prevention of breast cancer. However, preventive measures against breast cancer may require different approaches depending on risk factor.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by The Regional Ethical Review Board of Linköping. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by The Institutional Animal Ethics Committee at Linköping University.

Author contributions

CD designed the project and performed all microdialysis investigations and the animal study. AA participated in acquiring the microdialysis samples, carried out sample preparation, and participated in the animal study. PL design and analyzed the MRI part of the study. CD, JE, and PL analyzed the data and prepared the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants to CD from the Swedish Cancer Society (2018/464), the Swedish Research Council (2018–02584), LiU-Cancer, and ALF of Linköping University Hospital.

Acknowledgments

The authors would like to thank Ann-Christine Andersson of the Department of Oncology, Linköping University Hospital for providing excellent technical assistance and Anna Rzepecka MD and the staff of the Mammography Department, Linköping University Hospital, for identifying subjects with different mammographic densities.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1128318/full#supplementary-material

References

1. Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, et al. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. *Endocr Relat Cancer* (2005) 12(4):1071–82. doi: 10.1677/erc.1.01038

2. Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. *J Natl Cancer Inst* (2010) 102 (16):1224–37. doi: 10.1093/jnci/djq239

3. Alowami S, Troup S, Al-Haddad S, Kirkpatrick I, Watson PH. Mammographic density is related to stroma and stromal proteoglycan expression. *Breast Cancer Res* (2003) 5(5):R129–35. doi: 10.1186/bcr622

4. Hawes D, Downey S, Pearce CL, Bartow S, Wan P, Pike MC, et al. Dense breast stromal tissue shows greatly increased concentration of breast epithelium but no increase in its proliferative activity. *Breast Cancer Res* (2006) 8(2):R24. doi: 10.1186/bcr1408

5. Khan QJ, Kimler BF, O'Dea AP, Zalles CM, Sharma P, Fabian CJ. Mammographic density does not correlate with ki-67 expression or cytomorphology in benign breast cells obtained by random periareolar fine needle aspiration from women at high risk for breast cancer. *Breast Cancer Res* (2007) 9(3):R35. doi: 10.1186/bcr1683

6. Anderson TJ, Ferguson DJ, Raab GM. Cell turnover in the "resting" human breast: influence of parity, contraceptive pill, age and laterality. *Br J Cancer* (1982) 46(3):376–82. doi: 10.1038/bjc.1982.213

7. Wilson CL, Sims AH, Howell A, Miller CJ, Clarke RB. Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue. *Endocr Relat Cancer* (2006) 13(2):617–28. doi: 10.1677/erc.1.01165

8. Shekhar MP, Werdell J, Santner SJ, Pauley RJ, Tait L. Breast stroma plays a dominant regulatory role in breast epithelial growth and differentiation: implications for tumor development and progression. *Cancer Res* (2001) 61(4):1320–6.

9. Bendrik C, Dabrosin C. Estradiol increases IL-8 secretion of normal human breast tissue and breast cancer in vivo. *J Immunol* (2009) 182(1):371–8. doi: 10.4049/jimmunol.182.1.371

10. Garvin S, Nilsson UW, Huss FR, Kratz G, Dabrosin C. Estradiol increases VEGF in human breast studied by whole-tissue culture. *Cell Tissue Res* (2006) 325(2):245–51. doi: 10.1007/s00441-006-0159-7

11. Svensson S, Abrahamsson A, Rodriguez GV, Olsson AK, Jensen L, Cao Y, et al. CCL2 and CCL5 are novel therapeutic targets for estrogen-dependent breast cancer. *Clin Cancer Res* (2015) 21(16):3794–805. doi: 10.1158/1078-0432.CCR-15-0204

12. Vazquez Rodriguez G, Abrahamsson A, Jensen LD, Dabrosin C. Estradiol promotes breast cancer cell migration *via* recruitment and activation of neutrophils. *Cancer Immunol Res* (2017) 5(3):234–47. doi: 10.1158/2326-6066.CIR-16-0150

13. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013

14. Kim J, DeBerardinis RJ. Mechanisms and implications of metabolic heterogeneity in cancer. *Cell Metab* (2019) 30(3):434–46. doi: 10.1016/j.cmet.2019.08.013

15. Faubert B, Solmonson A, DeBerardinis RJ. Metabolic reprogramming and cancer progression. *Science* (2020) 368(6487). doi: 10.1126/science.aaw5473

16. Hahn WC, Bader JS, Braun TP, Califano A, Clemons PA, Druker BJ, et al. An expanded universe of cancer targets. *Cell* (2021) 184(5):1142–55. doi: 10.1016/j.cell.2021.02.020

17. Sickles EA, D'Orsi CJ, Bassett LW, Appleton CM, Berg WA, Burnside ES, et al. ACR BI-RADS® mammography. In: *ACR BI-RADS® atlas, breast imaging reporting and data system*. Reston, VA: American College of Radiology (2013).

18. Lundberg P, Forsgren MF, Tellman J, Kihlberg J, Rzepecka A, Dabrosin C. Breast density is strongly associated with multiparametric magnetic resonance imaging biomarkers and pro-tumorigenic proteins in situ. *Br J Cancer* (2022) 2025–2033. doi: 10.1038/s41416-022-01976-3

19. Abrahamsson A, Rzepecka A, Romu T, Borga M, Leinhard OD, Lundberg P, et al. Dense breast tissue in postmenopausal women is associated with a pro-inflammatory microenvironment in vivo. *Oncoimmunology* (2016) 5(10):e1229723. doi: 10.1080/2162402X.2016.1229723

20. Ekstrand J, Zemmler M, Abrahamsson A, Lundberg P, Forsgren M, Dabrosin C. Breast density and estradiol are major determinants for soluble TNF-TNF-R proteins *in vivo* in human breast tissue. *Front Immunol* (2022) 13:850240. doi: 10.3389/fimmu.2022.850240

21. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. *Magn Reson Med* (2008) 60(5):1122–34. doi: 10.1002/mrm.21737

22. Aberg UW, Saarinen N, Abrahamsson A, Nurmi T, Engblom S, Dabrosin C. Tamoxifen and flaxseed alter angiogenesis regulators in normal human breast tissue in vivo. *PloS One* (2011) 6(9):e25720. doi: 10.1371/journal.pone.0025720

23. Dabrosin C. Increase of free insulin-like growth factor-1 in normal human breast *in vivo* late in the menstrual cycle. *Breast Cancer Res Treat* (2003) 80(2):193–8. doi: 10.1023/A:1024575103524

24. Dabrosin C. Increased extracellular local levels of estradiol in normal breast *in vivo* during the luteal phase of the menstrual cycle. *J Endocrinol* (2005) 187(1):103–8. doi: 10.1677/joe.1.06163

25. Dabrosin C. Sex steroid regulation of angiogenesis in breast tissue. *Angiogenesis* (2005) 8(2):127–36. doi: 10.1007/s10456-005-9002-0

26. Garvin S, Dabrosin C. *In vivo* measurement of tumor estradiol and vascular endothelial growth factor in breast cancer patients. *BMC Cancer* (2008) 8:73. doi: 10.1186/1471-2407-8-73

27. Nilsson UW, Abrahamsson A, Dabrosin C. Angiogenin regulation by estradiol in breast tissue: tamoxifen inhibits angiogenin nuclear translocation and antiangiogenin therapy reduces breast cancer growth in vivo. *Clin Cancer Res* (2010) 16(14):3659–69. doi: 10.1158/1078-0432.CCR-10-0501

28. Abrahamsson A, Dabrosin C. Tissue specific expression of extracellular microRNA in human breast cancers and normal human breast tissue in vivo. *Oncotarget* (2015) 6(26):22959–69. doi: 10.18632/oncotarget.4038

29. Dabrosin C, Hallstrom A, Ungerstedt U, Hammar M. Microdialysis of human breast tissue during the menstrual cycle. *Clin Sci (Lond)* (1997) 92(5):493–6. doi: 10.1042/cs0920493

30. Garvin S, Dabrosin C. Tamoxifen inhibits secretion of vascular endothelial growth factor in breast cancer in vivo. *Cancer Res* (2003) 63(24):8742–8.

31. Lindahl G, Saarinen N, Abrahamsson A, Dabrosin C. Tamoxifen, flaxseed, and the lignan enterolactone increase stroma- and cancer cell-derived IL-1Ra and decrease tumor angiogenesis in estrogen-dependent breast cancer. *Cancer Res* (2011) 71(1):51–60. doi: 10.1158/0008-5472.CAN-10-2289

32. Abrahamsson A, Rodriguez GV, Dabrosin C. Fulvestrant-mediated attenuation of the innate immune response decreases ER(+) breast cancer growth *In vivo* more effectively than tamoxifen. *Cancer Res* (2020) 80(20):4487–99. doi: 10.1158/0008-5472.CAN-20-1705

33. Abrahamsson A, Rzepecka A, Dabrosin C. Equal pro-inflammatory profiles of CCLs, CXCLs, and matrix metalloproteinases in the extracellular microenvironment *In vivo* in human dense breast tissue and breast cancer. *Front Immunol* (2017) 8:1994. doi: 10.3389/fimmu.2017.01994

34. Mijic S, Dabrosin C. Platelet activation *In situ* in breasts at high risk of cancer: Relationship with mammographic density and estradiol. *J Clin Endocrinol Metab* (2021) 106(2):485–500. doi: 10.1210/clinem/dgaa820

35. Supuran CT. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nat Rev Drug Discovery* (2008) 7(2):168–81. doi: 10.1038/nrd2467

36. Buono M, Cosma MP. Sulfatase activities towards the regulation of cell metabolism and signaling in mammals. *Cell Mol Life Sci* (2010) 67(5):769–80. doi: 10.1007/s00018-009-0203-3

37. Foster LJ, Yaworsky K, Trimble WS, Klip A. SNAP23 promotes insulin-dependent glucose uptake in 3T3-L1 adipocytes: possible interaction with cytoskeleton. *Am J Physiol* (1999) 276(5):C1108–14. doi: 10.1152/ajpcell.1999.276.5.C1108

38. Lu WD, Funkelstein L, Toneff T, Reinheckel T, Peters C, Hook V. Cathepsin h functions as an aminopeptidase in secretory vesicles for production of enkephalin and galanin peptide neurotransmitters. *J Neurochem* (2012) 122(3):512–22. doi: 10.1111/j.1471-4159.2012.07788.x

39. Lim CY, Zoncu R. The lysosome as a command-and-control center for cellular metabolism. *J Cell Biol* (2016) 214(6):653–64. doi: 10.1083/jcb.201607005

40. Patel S, Homaei A, El-Seedi HR, Akhtar N. Cathepsins: Proteases that are vital for survival but can also be fatal. *BioMed Pharmacother* (2018) 105:526–32. doi: 10.1016/j.biopha.2018.05.148

41. Thibeaux S, Siddiqi S, Zhelyabovska O, Moinuddin F, Masternak MM, Siddiqi SA. Cathepsin b regulates hepatic lipid metabolism by cleaving liver fatty acid-binding protein. *J Biol Chem* (2018) 293(6):1910–23. doi: 10.1074/jbc.M117.778365

42. Mizunoe Y, Kobayashi M, Tagawa R, Nakagawa Y, Shimano H, Higami Y. Association between lysosomal dysfunction and obesity-related pathology: A key knowledge to prevent metabolic syndrome. *Int J Mol Sci* (2019) 20(15). doi: 10.3390/ijms20153688

43. Schweiger A, Christensen IJ, Nielsen HJ, Sorensen S, Brunner N, Kos J. Serum cathepsin h as a potential prognostic marker in patients with colorectal cancer. *Int J Biol Markers* (2004) 19(4):289–94. doi: 10.1177/172460080401900406

44. Peng P, Chen JY, Zheng K, Hu CH, Han YT. Favorable prognostic impact of cathepsin h (CTSH) high expression in thyroid carcinoma. *Int J Gen Med* (2021) 14:5287–99. doi: 10.2147/IJGM.S327689

45. Schweiger A, Staib A, Werle B, Krasovec M, Lah TT, Ebert W, et al. Cysteine proteinase cathepsin h in tumours and sera of lung cancer patients: relation to prognosis and cigarette smoking. *Br J Cancer* (2000) 82(4):782–8. doi: 10.1054/bjoc.1999.0999

46. Sahasrabuddhe NA, Barbhuiya MA, Bhunia S, Subbannayya T, Gowda H, Advani J, et al. Identification of prosaposin and transgelin as potential biomarkers for gallbladder cancer using quantitative proteomics. *Biochem Biophys Res Commun* (2014) 446(4):863–9. doi: 10.1016/j.bbrc.2014.03.017

47. Jevnikar Z, Rojnik M, Jamnik P, Doljak B, Fonovic UP, Kos J. Cathepsin h mediates the processing of talin and regulates migration of prostate cancer cells. *J Biol Chem* (2013) 288(4):2201–9. doi: 10.1074/jbc.M112.436394

48. Mills EG, Izzi-Engbeaya C, Abbara A, Comninos AN, Dhillo WS. Functions of galanin, spexin and kisspeptin in metabolism, mood and behaviour. *Nat Rev Endocrinology* (2021) 17(2):97–113. doi: 10.1038/s41574-020-00438-1

49. Rauch I, Kofler B. The galanin system in cancer. *Exp Suppl* (2010) 102:223–41. doi: 10.1007/978-3-0346-0228-0_16

50. Ormandy CJ, Lee CS, Ormandy HF, Fantl V, Shine J, Peters G, et al. Amplification, expression, and steroid regulation of the preprogalanin gene in human breast cancer. *Cancer Res* (1998) 58(7):1353–7.

51. Figliuolo da Paz V, Ghishan FK, Kiela PR. Emerging roles of disabled homolog 2 (DAB2) in immune regulation. *Front Immunol* (2020) 11:580302. doi: 10.3389/fimmu.2020.580302

52. Maurer ME, Cooper JA. The adaptor protein Dab2 sorts LDL receptors into coated pits independently of AP-2 and ARH. *J Cell Sci* (2006) 119(Pt 20):4235–46. doi: 10.1242/jcs.03217

53. Mei X, Zhao H, Huang Y, Tang Y, Shi X, Pu W, et al. Involvement of disabled-2 on skin fibrosis in systemic sclerosis. *J Dermatol Sci* (2020) 99(1):44–52. doi: 10.1016/j.jdermsci.2020.05.009

54. Bishnupuri KS, Sainathan SK, Ciorba MA, Houchen CW, Dieckgraefe BK. Reg4 interacts with CD44 to regulate proliferation and stemness of colorectal and pancreatic cancer cells. *Mol Cancer Res* (2022) 20(3):387–99. doi: 10.1158/1541-7786.MCR-21-0224

55. Chen Z, Downing S, Tzanakakis ES. Four decades after the discovery of regenerating islet-derived (Reg) proteins: Current understanding and challenges. *Front Cell Dev Biol* (2019) 7:235. doi: 10.3389/fcell.2019.00235

56. Mauvais-Jarvis F, Clegg DJ, Hevener AL. The role of estrogens in control of energy balance and glucose homeostasis. *Endocr Rev* (2013) 34(3):309–38. doi: 10.1210/er.2012-1055

57. Yao H, Sun C, Hu Z, Wang W. The role of annexin A4 in cancer. *Front Biosci (Landmark Ed)* (2016) 21(5):949–57. doi: 10.2741/4432

58. Heinick A, Pluteanu F, Hermes C, Klemme A, Domnik M, Husser X, et al. Annexin A4 n-terminal peptide inhibits adenylyl cyclase 5 and limits beta-adrenoceptor-mediated prolongation of cardiac action potential. *FASEB J* (2020) 34(8):10489–504. doi: 10.1096/fj.201902094RR

59. Ponnampalam AP, Rogers PA. Cyclic changes and hormonal regulation of annexin IV mRNA and protein in human endometrium. *Mol Hum Reprod* (2006) 12(11):661–9. doi: 10.1093/molehr/gal075

60. Lewis Marffy AL, McCarthy AJ. Leukocyte immunoglobulin-like receptors (LILRs) on human neutrophils: Modulators of infection and immunity. *Front Immunol* (2020) 11:857. doi: 10.3389/fimmu.2020.00857

61. van Houtum EJH, Bull C, Cornelissen LAM, Adema GJ. Siglec signaling in the tumor microenvironment. *Front Immunol* (2021) 12:790317. doi: 10.3389/fimmu.2021.790317

62. Humphreys GI, Ziegler YS, Nardulli AM. 17beta-estradiol modulates gene expression in the female mouse cerebral cortex. *PloS One* (2014) 9(11):e111975. doi: 10.1371/journal.pone.0111975

63. Kaymak I, Williams KS, Cantor JR, Jones RG. Immunometabolic interplay in the tumor microenvironment. *Cancer Cell* (2021) 39(1):28–37. doi: 10.1016/j.ccell.2020.09.004

64. Rosenbaum SR, Wilski NA, Aplin AE. Fueling the fire: Inflammatory forms of cell death and implications for cancer immunotherapy. *Cancer discovery* (2021) 11(2):266–81. doi: 10.1158/2159-8290.CD-20-0805

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 24 May 2023
DOI 10.3389/fonc.2023.1167815

OPEN ACCESS

EDITED BY
Iman Mamdouh Talaat,
University of Sharjah, United Arab Emirates

REVIEWED BY
Mohammed A. Al-Obaide,
Texas Tech University Health Science Center Amarillo, United States
Jyothi S. Prabhu,
St. John's Research Institute, India

*CORRESPONDENCE
Zhihong Gong
zhihong.gong@roswellpark.org

RECEIVED 16 February 2023
ACCEPTED 09 May 2023
PUBLISHED 24 May 2023

CITATION
Chen J, Higgins MJ, Hu Q, Khoury T, Liu S, Ambrosone CB and Gong Z (2023) DNA methylation differences in noncoding regions in ER negative breast tumors between Black and White women. *Front. Oncol.* 13:1167815. doi: 10.3389/fonc.2023.1167815

DNA methylation differences in noncoding regions in ER negative breast tumors between Black and White women

Jianhong Chen[1], Michael J. Higgins[2], Qiang Hu[3], Thaer Khoury[4], Song Liu[3], Christine B. Ambrosone[1] and Zhihong Gong[1*]

[1]Department of Cancer Prevention and Control, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States, [2]Department of Molecular and Cellular Biology, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States, [3]Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States, [4]Department of Pathology & Laboratory Medicine, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States

Introduction: Incidence of estrogen receptor (ER)-negative breast cancer, an aggressive tumor subtype associated with worse prognosis, is higher among African American/Black women than other US racial and ethnic groups. The reasons for this disparity remain poorly understood but may be partially explained by differences in the epigenetic landscape.

Methods: We previously conducted genome-wide DNA methylation profiling of ER- breast tumors from Black and White women and identified a large number of differentially methylated loci (DML) by race. Our initial analysis focused on DML mapping to protein-coding genes. In this study, motivated by increasing appreciation for the biological importance of the non-protein coding genome, we focused on 96 DMLs mapping to intergenic and noncoding RNA regions, using paired Illumina Infinium Human Methylation 450K array and RNA-seq data to assess the relationship between CpG methylation and RNA expression of genes located up to 1Mb away from the CpG site.

Results: Twenty-three (23) DMLs were significantly correlated with the expression of 36 genes (FDR<0.05), with some DMLs associated with the expression of single gene and others associated with more than one gene. One DML (cg20401567), hypermethylated in ER- tumors from Black versus White women, mapped to a putative enhancer/super-enhancer element located 1.3 Kb downstream of *HOXB2*. Increased methylation at this CpG correlated with decreased expression of *HOXB2* (Rho=-0.74, FDR<0.001) and other *HOXB/HOXB-AS* genes. Analysis of an independent set of 207 ER- breast cancers from TCGA similarly confirmed hypermethylation at cg20401567 and reduced *HOXB2* expression in tumors from Black versus White women (Rho=-0.75, FDR<0.001).

Discussion: Our findings indicate that epigenetic differences in ER- tumors between Black and White women are linked to altered gene expression and may hold functional significance in breast cancer pathogenesis.

KEYWORDS

breast cancer, DNA methylation, noncoding regions, ER negative tumor, Black and White women

1 Introduction

Evidence from both epidemiological and large-scale consortium studies supports the hypothesis that estrogen receptor positive (ER+) and negative (ER-) breast tumors derive from distinct etiologic pathways (1, 2). Compared to women diagnosed with ER+ breast cancer, those with ER- tumors in general have a poor prognosis, partly due to their aggressive phenotype and the lack of targeted therapy. ER- breast cancer is more common among Black women than White women (3, 4), but distinct reasons for these disparities remain to be elucidated.

DNA methylation, a major epigenetic mechanism, plays crucial roles in hormone-induced differentiation and tissue remodeling of the mammary gland through the life course (5). Aberrant DNA methylation patterns in breast cancer have been widely observed, and could contribute to differences in breast cancer risk between Black and White women (6–9). Studies have reported that hypermethylation at promoter regions of tumor suppressor genes, such as *RASSF1A* and *CDH13*, was inversely associated with gene expression, and that expression levels of these genes were lower in ER- tumors from Black women than those in White women, representing potential underlying tumor biological mechanisms explaining breast cancer racial disparities (10, 11). In addition, our previous research identified differentially methylated loci (DML) by tumor ER subtype and between races, with more DMLs by tumor ER subtype among Black women than that in White women; the number of race-related DMLs identified in ER- tumors were almost twice as those identified in ER+ tumors (6, 12).

Together, aberrant DNA methylation patterns have been used to dissect breast cancer risk by tumor subtype and racial groups, with potential diagnostic and prognostic applications (13, 14). However, previous studies have mainly focused on DNA methylation alterations associated with protein coding genes, with little attention on non-protein coding regions, which constitute more than 98% of the whole human genome.

It has long been known that a large portion of aberrant DMLs in breast cancer is located in intergenic regions (15), which constitute about 50% of the human genome (16). In a genome-wide expression-methylation quantitative loci (emQTL) analysis, Fleischer et al. reported several hundred regulatory elements not associated with protein coding genes whose methylation alterations were associated with different breast cancer lineages (17). Enhancers are critical *cis*-regulatory elements within non-coding regions, which contain the majority of cancer-associated variants based on genome-wide association studies (18). Studies have also revealed that enhancers are the most consistently differentially methylated regions and that their differential methylation is in a cell-type-specific manner, indicating the importance of enhancer methylation for epigenetic regulation of tumorigenesis (19). The underlying mechanisms could be that certain DMLs overlapping with enhancers can regulate tumor-associated genes and pathways, subsequently playing important roles in cancer (17, 20, 21). In addition, focused on small noncoding RNAs, microRNAs (miRNAs), we previously found that several hundred DMLs that mapped to miRNA genes were differentially methylated by tumor ER subtype and between Black and White women, and that their methylation levels were significantly correlated with corresponding miRNA gene expression (22). In summary, these studies highlight the importance of DMLs occurring within non-protein coding regions in relation to risk of breast cancer, especially their potential roles in explaining breast cancer racial disparities.

As described previously, we conducted genome-wide DNA methylation profiling on breast tumor tissue samples obtained from participants in the Women's Circle of Health Study (WCHS), a case-control study designed to investigate risk factors for aggressive breast cancer in Black and White women (12, 23). Our initial analysis focused on DMLs mapping within or near protein-coding genes, and revealed that a key pro-luminal transcription factor, *FOXA1*, was hypermethylated and repressed in tumors from Black women compared to White women (12). Herein, we focus on DMLs mapped to non-protein coding regions due to their biological importance and limited research in the area.

Motivated by the biological significance of the noncoding genome and our research interests in understanding the higher risk of ER- breast cancer in Black compared to White women, we aimed to identify DMLs by race within ER- tumors, with a focus on DMLs located in intergenic and noncoding RNA genomic regions. In addition, we integrated both DNA methylation (Illumina Infinium 450K array) and gene expression (RNA sequencing) data to examine whether DMLs were associated with altered gene expression. Our results were then validated using The Cancer Genome Altas (TCGA) dataset. We further investigated the epigenomic context and molecular features of the top DMLs confirmed in both our and TCGA datasets to determine their regulatory potential and biological functions.

2 Materials and methods

2.1 Study population and tissue samples

Data and breast tumor tissue samples were from participants enrolled in the WCHS. Details on the study design and participant recruitment have been described previously (12). The study protocol was approved by Institutional Review Boards at all participating institutes. In-home interviews were conducted to obtain data on known and suspected risk factors for breast cancer. As part of the informed consent, >95% participants signed a release for their pathology reports and archived specimens in form of formalin-fixed, paraffin-embedded (FFPE) tumor blocks, which were obtained from the pathology departments of the treating hospitals. Data on tumor pathological features, including ER status, were extracted from the pathology reports.

2.2 DNA extraction, DNA methylation profiling, and data processing

DNA was extracted from FFPE tumor tissues among 694 women enrolled in the WCHS (Table S1) as previously described (12). Briefly, FFPE samples were deparaffinized in xylene, lysed, and incubated at 56°C with constant rotation until completely digested. Lysates were then heated at 70°C for 20min to inactivate the Proteinase K and stored at 4°C. DNA from a 5ul aliquot of FFPE lysate was purified using the DNA Clean & Concentrator-5 kit. Genome-wide DNA methylation analysis was carried out at Roswell Park Genomics Shared Resource using the Illumina Infinium HumanMethylation450 BeadChip platform, which interrogates > 485,000 CpG dinucleotides per sample at single-nucleotide resolution and covers 99% of RefSeq genes. To minimize the impact of batch effects, DNA samples from tumors were randomized on plates according to age, race, and ER status. The raw intensities from the array were extracted using GenomeStudio, and the data summarized into BeadStudio IDAT files and processed by the minfi R package. The methylation level of each CpG site, calculated as a β value, ranged between (0, 1), with 0 for absent methylation and 1 for complete methylation. In brief, the 450K array data were subjected to rigorous sample and locus specific quality control criteria, SWAN normalization, and correction for batch effects using the ComBat algorithm (24). Low quality probes (probes with detection p value > 0.05 in more than half of samples) and samples with poor detection p values (samples with detection p values < 1 x 10^{-5} at more than 75% of CpG loci) were removed using the IMA package (25). We used Bowtie 2 for sequence alignment (26). Probes known to map ambiguously, exhibiting cross-reactivity, and that contain single nucleotide polymorphisms were also removed, leaving the final dataset containing 276,108 CpG loci in 694 tumor samples for final analyses (27, 28).

2.3 Data analysis pipeline

Unlike previous studies that focused primarily on CpGs located near a protein-coding gene, we investigated CpGs mapped to intergenic regions or noncoding RNAs in proximity, and validated results in the independent TCGA breast cancer cohort. We performed functional annotations for top CpGs to determine their potential roles in breast cancer racial disparity by exploring their integrative biologic context from multiple genome and epigenome databases, their associations with target gene expression, and differential gene expression patterns between Black and White women, as shown in Figure 1, the overall workflow.

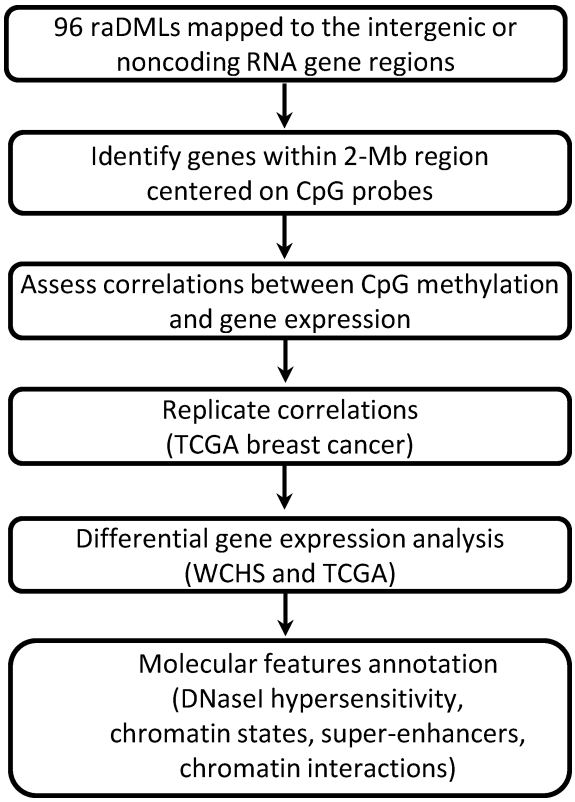


FIGURE 1
Study flowchart.

2.3.1 Differentially methylated loci in ER- tumors between Black and White women within noncoding regions

Differences in DNA methylation β-values for each probe were evaluated in ER- tumors by race. The Wilcoxon rank-sum test was used to evaluate the statistical significance for each probe in the comparison. To adjust for multiple comparisons, the false discovery rate (FDR) was computed using the Benjamini and Hochberg approach. Differentially methylated loci in ER- tumors between Black and White women (raDMLs) were defined as CpGs with an absolute mean β value difference (|delta β|) at least 0.10 between two race groups and *FDR*-adjusted p value <0.05. All analyses were performed using R package.

2.3.2 Association between DNA methylation and gene expression

Genes located within a 2 Mb window centered on a CpG site were considered as potential regulatory targets of DNA methylation (29–31).

Thus, for each raDML, we adopted a window width of 1Mb on either side of the CpG and assessed the correlations between CpG methylation level and relative RNA expression level of genes within this window. For each CpG/gene pair, the Spearman correlation between DNA methylation (β values) and relative gene expression levels (log counts per million, log CPM) was assessed. An independent collection of 50 fresh frozen breast tumor samples from Roswell Park Pathology Network Shared Resource (PNSR) was used for RNA extraction and relative gene expression analysis as previously described (6, 12). An association was considered significant if *FDR*-adjusted p value <0.05. Only significant associations were included for further functional annotations, which were then validated by repeating the same analysis using the TCGA breast cancer cohort.

2.3.3 Differential gene expression analysis

Genes with significant correlation between DNA methylation and relative gene expression (CpG/gene pairs) were analyzed for differential expression between Black and White women in ER- tumors using DESeq2 (32). P values were calculated using linear regression function of DESeq2, with adjustment of age at diagnosis and corrections for multiple testing.

2.3.4 TCGA data processing and analysis

To validate our findings, level 3 Illumina HM450 methylation data, Hiseq2 gene expression data, demographic and related clinical features (e.g., age, race, ER status, and tumor stage) of the 1,097 cases included in the TCGA breast cancer cohort were downloaded from the publicly available FireBrowse database (http://firebrowse.org). Validation of race-related methylation difference, association between methylation and associated gene expression, and differential gene expression analysis by race in ER- tumors were conducted following the same pipeline as described above.

2.3.5 Molecular feature annotation

For CpGs showing statistically significant correlation with gene expression confirmed in both our and TCGA dataset, we further investigated their epigenomic context and determined their regulatory potential. Specifically, the regional chromatin landscape in proximity to each CpG site was investigated using multiple publicly available databases. Chromatin state annotations were extracted from the Roadmap Epigenomics ChromHMM on ENCODE (E027 and E119). To facilitate functional interpretation, we focused on six ChromHMM states, including TssA (Active TSS), TssAFlnk (Flanking Active TSS), TssBiv (Bivalent/Poised TSS), EnhG (Genic enhancers), Enh (Enhancers), and EnhBiv (Bivalent Enhancer). DNaseI hypersensitive sites indicative of an open chromatin structure with potential transcriptional activity were similarly identified from the Roadmap/ENCODE reference epigenomes. GeneHancer module under the GeneCards Suite comprises a large collection of enhancer- gene association, which was used to annotate enhancer regions and their associated genes (33). We further inferred super-enhancer regions using a catalogue of super-enhancers in HMEC and MCF-7 cell lines published elsewhere (34). The UCSC Genome Browser was used to visualize genomic features centered on these CpGs. A circos plot summarizing global molecular features of these CpGs was generated using OmicCircos version 1.24.0. All statistical analyses were performed using the R statistical software.

3 Results

3.1 DMLs mapped within non-protein coding regions in ER- tumors between Black and White women

In our previous analysis (12), we identified a total of 396 raDMLs that exhibited significant differential methylation between Black and White women within the ER- breast cancer group. Of the 396 raDMLs, 276 CpGs were assigned with at least one protein-coding gene based on the Illumina self-manifestation file (https://support.illumina.com/downloads/infinium_humanmethylation450_product_files.html), leaving 120 CpGs uncharacterized. We further excluded 24 CpGs that were mapped in the protein-coding region based on an updated Ensemble gene annotation file (Ensemble Gene 104, http://www.ensembl.org/). In the end, the remaining 96 raDMLs mapped to intergenic regions or noncoding RNA regions, which were the focus of the current study (Table S2). Validation of observed methylation differences by race at each CpG site was then conducted using the TCGA breast cancer cohort.

Among the 96 raDMLs within non-protein coding regions in ER- tumors between Black and White women, 59 of these CpGs were located in intergenic regions and the remaining 37 CpGs mapped to at least one non-coding RNA gene. Consistent with our previous findings on protein-coding genes (12), there were more hypomethylated CpGs within the non-protein coding genome. Specifically, of the 96 raDMLs, 58 CpGs were hypomethylated and 38 CpGs were hypermethylated in ER- tumors from Black compared to White women.

Out of the 96 raDMLs, data on 59 CpGs were available in the TCGA dataset. Except for one CpG site (dot in red), all other probes showed consistent direction of methylation change by race (dot in black) in the TCGA dataset (Figure S1).

3.2 Associations between DNA methylation and gene expression

The 96 raDMLs were found to be associated with 1,998 unique genes, which correspond to a total of 2,357 unique CpG/gene pairs. Using previously described DNA methylation and RNA-seq data from analysis of an independent group of 50 fresh frozen breast tumor samples (6, 12), Spearman correlation between DNA methylation and gene expression was assessed for each CpG/gene pair. Among the 2,357 unique CpG/gene pairs, analysis identified significant correlations for 39 CpG/gene pairs, corresponding to 23 unique CpGs and 36 unique coding or noncoding RNA genes (FDR-adjusted p value < 0.05, Table S3). Of the 36 unique gene/RNAs, we found six long noncoding RNAs (lncRNAs), including three antisense lncRNAs (*SOX9-AS1*, *HOXB-*

AS1, and *HOXB-AS3*), a long intergenic noncoding RNA (*LINC01152*), and two uncharacterized lncRNAs (*LOC102723517*, *LOC283335*). Overall, most CpG/gene pairs (25/39, 64.1%) exhibited positive correlations between methylation and gene expression, while 14/39 (35.9%) pairs showed negative correlations.

Table 1 listed the top 10 CpG/gene pairs, including 5 unique CpGs with associated genes or lncRNAs. Notably, we identified several CpG sites at which methylation was correlated with the expression of multiple genes. Cg20401567 is the top CpG site in proximity to the *HOXB* gene cluster, with its methylation level inversely, highly associated with the expression levels of multiple members of the *HOXB* gene family, including *HOXB2* (correlation coefficient, rho=-0.74) and *HOXB3* (rho=-0.62). Previously, we reported methylation level at cg04932551, a raDML within the gene body of *FOXA1*, inversely correlated with *FOXA1* expression (12). In this study, we discovered another CpG, cg12212453 located at 5 kb downstream of *FOXA1*, whose methylation level was also strongly, inversely correlated with *FOXA1* gene expression (rho=-0.67). We identified a novel CpG, cg05322837, whose methylation level was correlated with expression of two lncRNAs, *LOC102723517* and *LINC01152* (rho=0.65 and 0.58, respectively), and a Solute Carrier Family 39 gene, *SLC39A11* (rho=-0.61). In addition, we found the methylation level of cg05199874 was positively correlated with expression of multiple genes, including the signal peptide-CUB-EGF domain-containing protein 2 (*SCUBE2*), a novel tumor suppressor gene, which showed inhibitory roles in breast tumor invasion and migration through concerted activities with *FOXA1* (35, 36). Moreover, methylation level of cg12821539 was positively correlated with expression of *ZIC5* (rho=0.62), which has been implicated as oncogenes in some cancers (37).

HOXB gene family, *FOXA1*, *SLC39A11*, and *SCUBE2* were found to play well-established roles in breast cancer development (12, 35, 38), and thus were the focus on our further analysis.

We further validated these findings by repeating the correlation analysis in the TCGA breast cancer cohort. Limited by data availability of HM450 DNA methylation and Hiseq gene expression in the TCGA breast cancer cohort, only 28 out of the 39 significant CpG/gene pairs identified in our analysis were available and thus included in the validation analysis. Nevertheless, 20 out of these 28 CpG/gene pairs were validated with respect to the magnitude and direction of the correlation coefficient and all reached statistical significance (Table S4). The remaining 8 CpG/gene pairs did not reach statistical significance in the TCGA cohort. As shown in Figure 2, we further showed in scatter plots for the top four highly correlated CpG/gene pairs that exhibited the most consistent correlation between methylation and gene expression in both WCHS and TCGA breast cancer cohort, respectively.

3.3 Molecular features accounting for regulatory effects of aberrant DNA methylation

Chromatin architecture in which DNA methylation occurs provide important clues as to how DNA methylation alterations mediate their effects in disease predisposition. For each of the 23 unique CpGs (out of the 39 CpG/gene pairs) exhibiting significant correlations with its paired gene, we annotated their molecular features globally, including genomic position, ChromHMM state, DNaseI hypersensitive sites, enhancer and super enhancer sites, and their correlations between CpG methylation and RNA expression. As shown in Figure 3, cg20401567, cg12211453, cg05322837, and cg05199874, paired with the *HOXB* gene cluster, *FOXA1*, *SLC39A11*, and *SCUBE2*, respectively, were highly enriched with candidate *cis*-regulatory elements (cCREs), characterized by hypersensitive DNase I sites, promoter/enhancer-related ChromHMM segments, H3K4m1/2/3 and H3K27ac histone modification, implicating them as subjects of intensive gene

TABLE 1 Top ten CpG/gene pairs based on methylation-gene expression correlation analysis.

CpG [a]	Gene [b]	Distance (Kb) [c]	deltaBeta[d]	rho[e]	FDR[f]
cg20401567	*HOXB2*	3.3	0.12	-0.74	2.90E-06
	HOXB3	-6.7	0.12	-0.62	7.40E-04
cg12212453	*FOXA1*	4.8	0.12	-0.67	1.10E-04
cg05322837	*LOC102723517*	117.5	0.13	0.65	2.30E-04
	LINC01152	121	0.13	0.58	3.00E-03
	SLC39A11	-494.3	0.13	-0.61	8.00E-04
cg05199874	*TMEM41B*	32.1	-0.13	0.59	2.50E-03
	RNF141	-910	-0.13	0.59	2.50E-03
	SCUBE2	582.1	-0.13	0.55	8.30E-03
cg12821539	*ZIC5*	27.6	0.11	0.62	7.40E-04

[a] Illumina 450K CpG probe.
[b] RefSeq genes located ≤1 Mb away from CpG.
[c] Genomic distance (kb) between CpG probe and transcriptional start site of the indicated gene.
[d] DNA methylation difference (delta Beta) at indicated CpGs comparing Black vs. White women.
[e] Spearman correlation coefficient (rho).
[f] False discovery rate (FDR) q value derived from correlation between beta values and RNA expression levels for an indicated CpG/gene pair.

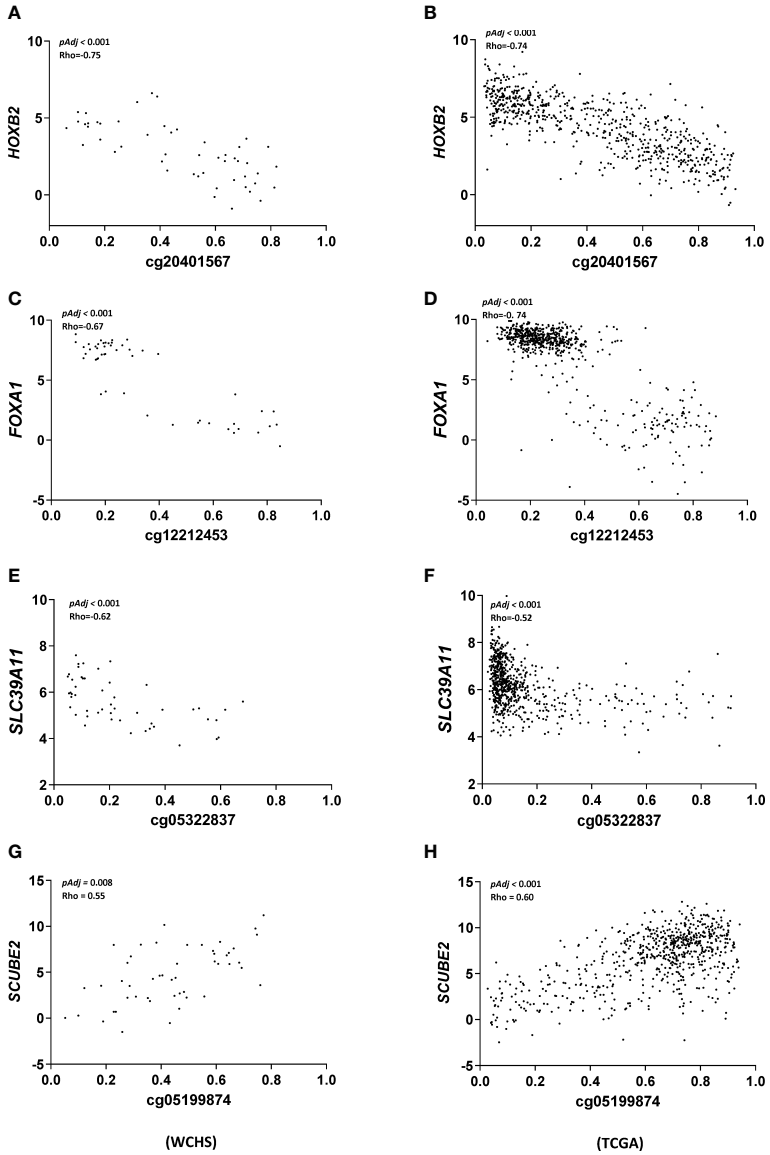


FIGURE 2
Scatter plot summarizing correlation analysis of top ranked CpG/gene pairs. X-axis denotes CpG site methylation levels in b value and Y-axis denotes relative gene expression of the associated genes in logCPM. For each pair [cg20401567 and HOXB2 **(A, B)**, cg12212453 and FOXA1 **(C, D)**, cg05322837 and SLC39A11 **(E, F)**, and cg05199874 and SCUBE2 **(G, H)**], Spearman correlation was used to test the relationship between DNA methylation and gene expression in WCHS (left panel) and the TCGA (right panel) breast cancer cohort with p values and correlation coefficients labeled in the inlet.

expression regulation through DNA methylation modifications. As we showed in supplementary Figure S2, the Genome Browser plot exhibits regional genomic features of enrichment of DNase I hypersensitive site, ChromHMM segments implicating promoter, and histone marks predictive of open chromatin region around the genomic position of cg20401567. Consistent with our results, the region around cg20401567 includes not only *HOXB* gene family members, but also *HOXB-AS* genes. The relative position of *cis*-elements, cg20401567, *HOXB2*, *HOXB3*, and *HOXB-AS1* on chromosome 17 was shown in Figure 4. Enhancer, super enhancer, and promoter annotations are obtained from various resources and their genomic positions are overlapped with each other. In addition to enrichment of *cis*-elements, we also observed high DNA sequence conservation across vertebrates for this region, indicative of important biological functions.

3.4 Expression differences on raDML-associated genes in ER- tumors between Black and White women

We further investigated whether the top CpG-correlated genes were differentially expressed in tumor tissues between Black and White women within the WCHS and TCGA breast cancer cohort. Except for *SLC39A11*, the other three genes (*HOXB2*, *FOXA1*, and *SCUBE2*) exhibited significantly lower expression levels in tumors

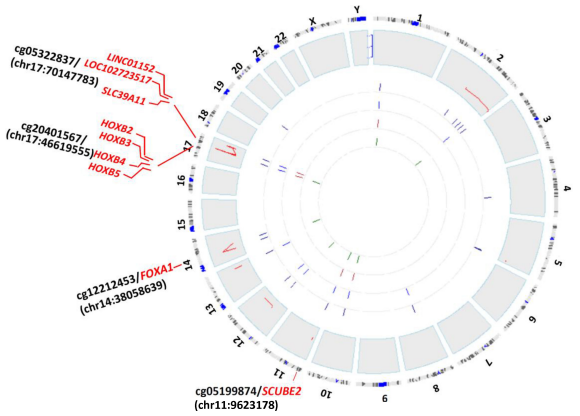


FIGURE 3
Visualization of genomic features of the 23 CpG sites. OmicCircos plot showed overall landscape of these CpG sites. Six tracks were displayed, from outside to inside: chromosome cytobands in black and blue, −log10 transformation of Spearman correlation p values (<0.05) between methylation and associated gene expression in red, the 23 CpG sites in navy blue, ChromHMM features of promoter and enhancer annotation in blue, hypersensitive DNase I sites annotation in brown, and super enhancer annotation of these probes in dark green. The top four CpGs and their associated gene (genes) were labeled in red.

from Black versus White women in both study cohorts after adjusting for age at diagnosis (Figure 5), which were consistent with observed CpG-gene expression correlation, with higher methylation at cg20401567 and cg12212453 (paired with *HOXB2* and *FOXA1*, respectively), and lower methylation at cg05199874 (paired with *SCUBE2*) in Black relative to White women. These results suggested that differential DNA methylation between races and their altered gene expression may contribute to breast cancer racial differences.

4 Discussion

Noncoding DNA regions comprise more than 98% of the human genome and play key roles in regulating gene expression through various mechanisms (39). For instance, aberrant

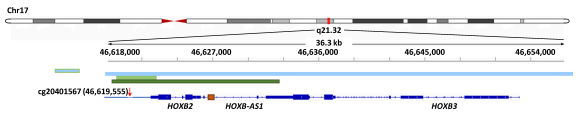


FIGURE 4
Relative genomic positions of cg20401567, *cis*-elements, *HOXB2*, *HOXB3*, and *HOXB-AS1* on chromosome 17. The chromosome ideogram was shown on top, with a region spanning this locus highlighted in red. Cg20401567, *cis*-elements, *HOXB* family members, and *HOXB-AS1* in this locus were displayed at the bottom of the diagram, with genomic coordinates labeled in the middle. *Cis*-element annotations from various resources around Cg20401567 were labeled, including two promoter/enhancers based on GeneCards (light blue), an enhancer based on ChromHMM (light green), and a super enhancer based on Hnisz D et al. (34) (dark green). In addition, at bottom of the plot, a promoter upstream of *HOXB2* was displayed (orange).

methylation in intergenic regions, such as enhancers, has been shown to be associated with altered expression of neighboring genes, including those involved in cell cycle processes, lymphocyte activation and apoptosis (17, 40). Very few prior studies have focused on identifying differentially methylated features, especially in noncoding genome regions, in breast tumors between Black and White women, which may contribute to race-related differential predisposition to aggressive tumor subtypes (6, 10–12). To our knowledge, we are one of the few to globally investigate the role of noncoding region DNA methylation patterns in relation to breast cancer racial disparities (22, 41).

Intergenic regions and lncRNA genes, which comprise most of the non-protein coding genome, are enriched in epigenetic modifications, and have elicited great efforts to systematically annotate the regulatory elements existing within these regions. In this study, we identified several CpGs located within the noncoding genome that were differentially methylated in ER- tumors between races, with some CpGs highly correlated with expression of specific protein coding and/or lncRNA genes. Using multiple publicly available epigenome databases, we further characterized molecular features of these CpGs, such as active histone modifications, chromatin accessibility, and enhancer/super-enhancer sequences. We found that most of the 96 raDMLs mapping to intergenic regions or lncRNA genes were in at least one of these functional elements, with some CpGs enriched by multiple functional elements. The rich content of putative regulatory elements located at intergenic CpGs suggests active regulation of transcriptional activity through DNA methylation modification in these regions. Consistent with our observation, Kamalakaran et al. reported that among featured DNA methylation sites efficiently distinguishing the five major breast cancer subtypes, 70% are within non-protein coding regions, while only 30% of the sites mapped to genes encoding proteins (15). It should be noted, that the high occurrence of these functional elements within intergenic noncoding regions might be also due to the fact that intergenic CpG loci of the original HM450 chip design were preferentially selected toward biologically significant/informative sites, DNase hypersensitive sites, and differentially methylated regions (42, 43). Nevertheless, our findings that these CpGs were differentially methylated between two racial groups and associated with altered gene expression, support their potential roles contributing to breast cancer differences between Black and White women.

One of the interesting observations is that aberrant DNA methylation occurring in CpG islands or within gene body generally correlates with expression levels of gene cluster members spanning an entire continuous region (44–46). Consistent with these findings, we observed that increased DNA methylation at cg20401567 correlated with reduced expression of *HOXB* gene cluster members and *HOXB-AS* members. *HOX* gene clusters are large superfamilies of genes whose members play fundamental roles in cell development (47). Boimel et al. reported that *HOXB2* knockdown promoted primary tumor growth in mammary adenocarcinoma cell lines, suggesting that in this context, it functions as a tumor suppressor (48). Intriguingly, there is evidence that in addition to the *HOX* gene coding regions, the *cis*-regulatory regions, including intergenic and

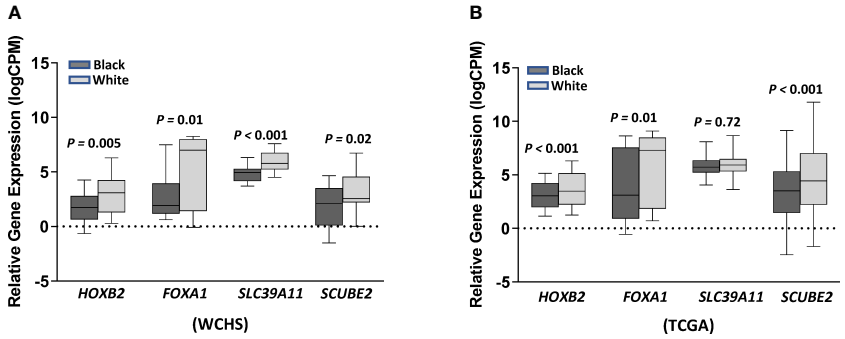


FIGURE 5
Gene expression levels of *HOXB2, FOXA1*, *SLC39A11*, and *SCUBE2* in ER- tumors between Black and White women from the WCHS cohort **(A)** and the TCGA **(B)** breast cancer cohort. Relative gene expression levels (logCPM) were presented as mean± standard errors (SE). Gene expression differences between Black and White women were tested using DESeq2 after adjustment of age at diagnosis, with adjusted p values labeled on top. Black and grey denotes relative gene expression levels in Black and in White women, respectively.

antisense transcribed regions, provide an extra layer of regulation of *HOX* expression (49). We observed that the genomic region around cg20401567 is highly conserved and enriched with various *cis*-regulatory elements in the integrative functional annotation map, which provides direct evidence supporting this model. This CpG locus is not in a gene promoter and its methylation affects at least 7 genes in the *HOXB* cluster spreading over 50 kb, thus, its mechanism of gene silencing is unlikely to be the same as canonical promoter repression (50). The current study aimed to provide further insights into why Black women are more likely than White women to be diagnosed with aggressive breast cancer, particularly ER- breast cancer. Reduced expression of *HOXB2* in breast tumors from Black versus White women represents a novel molecular feature which may be linked to racial differences in breast tumor biology and outcomes. Moreover, our findings support a novel regulatory site for *HOXB2* activity, which could be a potential therapeutic target in breast cancer treatment.

FOXA1 is an important transcription factor playing crucial roles in mammary gland development. We previously reported a *FOXA1* raDML (cg04932551), located in the gene body, which is annotated as a poised promoter. This site was hypermethylated in tumors from Black versus White women, and its methylation was inversely associated with *FOXA1* RNA and protein expression in ER- tumors among Black women (12, 51). We further found that methylation and expression of *FOXA1* is associated with parity and breastfeeding, suggesting a potential mechanism that links these reproductive exposures with ER- breast cancer among Black women (12). In the current study, we identified another raDML (cg12212453), located at 120 bp downstream of *FOXA1*, which was negatively correlated with *FOXA1* expression. Our analysis did not show associations (data not shown) of cg12212453 methylation with reproductive factors, suggesting it is less likely to mediate effects of reproductive exposures on cancer predisposition. Nevertheless, this CpG might represent a novel regulatory site for *FOXA1* expression and warrants further investigations.

SLC39A11 belongs to a member of a large family of membrane transport proteins participating in wide range of physiological processes (52). Significantly enhanced *SLC39A* family of proteins are expressed in multiple malignances including colorectal cancer, breast cancer, and esophageal cancer (53, 54). We observed a higher methylation of *SLC39A11* (cg05322837) correlated with lower *SLC39A11* expression levels in tumors from Black compared to that in White, implicating that differences in methylation and expression of *SLC39A11* may contribute to breast cancer racial disparities. Intriguingly, methylation at cg05322837 was positively correlated with expression of two overlapping lncRNAs, *LOC102723517* and *LINC01152*, as well as *SOX9-AS1*. The cg05322837 is in a putative enhancer region approximately 500 kb upstream of *SLC39A11* and about 100kb downstream of the two lncRNAs, thus it is possible that *SLC39A11* transcription is targeted by *LOC102723517* and *LINC01152*, with its expression down regulated by methylation at cg05322837.

SCUBE2 is another gene with its expression level positively correlated with DNA methylation at a CpG site (cg05199874) approximately 500kb upstream. It exhibited differential expression by race both in our study population and the TCGA breast cancer cohort. *SCUBE2* was reported to work synergistically with *FOXA1* as a novel breast-tumor suppressor, driving the reversal of epithelial–mesenchymal transition (EMT) (36). *SCUBE2* transcription was epigenetically inactivated by recruitment of DNA methyltransferase 1 onto its CpG islands during EMT while the exact CpGs remain unclear (36). We speculated that the downregulation of *SCUBE2* in concert with *FOXA1* is part of the EMT program that plays important roles in modulating breast-cancer cell migration and invasion. Our results implicated that EMT could also contribute to racial differences in ER- breast cancer predisposition and revealed that specific *cis* elements through which DNA methylation may influence two key regulators of EMT. Future studies are warranted to investigate this intriguing finding.

Transcription factors impact gene expression through binding to either positive (such as promoter and enhancer) or negative (such as silencer and insulator) regulatory elements under certain chromatin structure, which is drastically affected by epigenetic modifications including DNA methylation. Thus far, dysregulated DNA methylation in promoter region has attracted much of the research attention. How DNA methylation in other genome regions

affects gene expression remains largely unknown. In our study, while the same CpG sites correlated with two genes' expression levels in diverse direction, we also observed three CpG sites residing a small CpG island correlated with increased expression of the same genes, indicating these three consecutive CpG sites could belongs to the same regulatory element. The multiple correlation patterns between DNA methylation and gene expression suggested complex regulatory mechanisms of these *cis*-elements.

In summary, unlike previous studies that focused primarily on CpGs mapped to protein-coding genes, this study focuses on aberrant DML located in the noncoding genome regions, with the aim of interrogating biological mechanisms underlying the observed racial differences of high risk of ER- breast tumors in Black women compared to White women. We identified several important genes, being implicated in breast cancer pathology, with their expression correlated with aberrant DNA methylation of CpGs located in noncoding genome regions. The functional potentiality of these aberrantly methylated CpGs were further examined through integrative, molecular feature annotations. These results were subsequently validated in the independent TCGA breast cancer cohort. Our results provide new insights into the contribution of aberrant DNA methylation within the non-protein coding region to breast cancer racial disparities. Further experimental validations will be warranted to confirm these findings in future studies.

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

This study was approved by the institutional review boards of Roswell Park Comprehensive Cancer Center. The patients/participants provided their written informed consent to participate in this study.

Author contributions

CA, ZG and MH conceived and supervised the projects. QH, TK, and SL finished acquisition of all data. JC and QH did all statistical analysis and wrote the original draft. JC, MH, ZG, and CA reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Breast Cancer Research Foundation, the National Cancer Institute (R01 CA133264; R01 CA246688; P01 CA151135), and used Roswell Park Comprehensive Cancer Center's Data Bank and BioRepository, Genomics Shared Resource, and Bioinformatics Shared Resource (P30 CA016056).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1167815/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Validation of raDMLs identified in current (WCHS) study using the TCGA breast cancer cohort. Scatter plot of delta beta value of raDMLs from WCHS (X-axis) versus TCGA (Y-axis). Fifty-nine out of the 96 raDMLs were available in the TCGA breast cancer dataset. Fifty-eight raDMLs identified in both data, with a consistent direction of methylation changes (delta beta) and FDR-adjusted $P<0.05$, were plotted as black dots, whereas one CpG site showing an inconsistent methylation change was labeled in red. Cg20401567 was highlighted and labeled in green.

SUPPLEMENTARY FIGURE 2
An integrative annotation showing molecular features of the genomic regions of Cg20401567/*HOXB/HOXB-AS* gene clusters. From top to bottom, the tracks showing: *HOXB* and *HOXB-AS* gene clusters; the Cg20401567 site highlighted in vertical cyan line; measurements of evolutionary conservation from alignments of 100 vertebrate species; conserved transcriptional binding sites; the hypersensitive DNaseI sites profile of HMEC and MCF cell lines of the ENCODE project; the HMEC ChromHMM tracks indicating putative active (bright red) promoters, strong enhancer (orange), strong transcript (green), as well as putative weak enhancers (yellow); histone modifications of H3K27ac, H3K4me1/3 in various breast cell types.

References

1. Garcia-Closas M, Chanock S. Genetic susceptibility loci for breast cancer by estrogen receptor status. *Clin Cancer Res* (2008) 14(24):8000–9. doi: 10.1158/1078-0432.CCR-08-0975

2. Anderson WF, Rosenberg PS, Prat A, Perou CM, Sherman ME. How many etiological subtypes of breast cancer: two, three, four, or more? *J Natl Cancer Inst* (2014) 106(8):dju165. doi: 10.1093/jnci/dju165

3. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: convergence of incidence rates between black and white women. *CA Cancer J Clin* (2016) 66(1):31–42. doi: 10.3322/caac.21320

4. Amend K, Hicks D, Ambrosone CB. Breast cancer in African-American women: differences in tumor biology from European-American women. *Cancer Res* (2006) 66 (17):8327–30. doi: 10.1158/0008-5472.CAN-06-1927

5. Ivanova E, Le Guillou S, Hue-Beauvais C, Le Provost F. Epigenetics: new insights into mammary gland biology. *Genes (Basel)* (2021) 12(2):231. doi: 10.3390/genes12020231

6. Ambrosone CB, Young AC, Sucheston LE, Wang D, Yan L, Liu S, et al. Genome-wide methylation patterns provide insight into differences in breast tumor biology between American women of African and European ancestry. *Oncotarget* (2014) 5 (1):237–48. doi: 10.18632/oncotarget.1599

7. Davis MB, Newman LA. Breast cancer disparities: how can we leverage genomics to improve outcomes? *Surg Oncol Clin N Am* (2018) 27(1):217–34. doi: 10.1016/j.soc.2017.07.009

8. Lara OD, Wang Y, Asare A, Xu T, Chiu HS, Liu Y, et al. Pan-cancer clinical and molecular analysis of racial disparities. *Cancer* (2020) 126(4):800–7. doi: 10.1002/cncr.32598

9. Song MA, Brasky TM, Marian C, Weng DY, Taslim C, Dumitrescu RG, et al. Racial differences in genome-wide methylation profiling and gene expression in breast tissues from healthy women. *Epigenetics* (2015) 10(12):1177–87. doi: 10.1080/15592294.2015.1121362

10. Mehrotra J, Ganpat MM, Kanaan Y, Fackler MJ, McVeigh M, Lahti-Domenici J, et al. Estrogen receptor/progesterone receptor-negative breast cancers of young African-American women have a higher frequency of methylation of multiple genes than those of Caucasian women. *Clin Cancer Res* (2004) 10(6):2052–7. doi: 10.1158/1078-0432.CCR-03-0514

11. Wang S, Dorsey TH, Terunuma A, Kittles RA, Ambs S, Kwabi-Addo B. Relationship between tumor DNA methylation status and patient characteristics in African-American and European-American women with breast cancer. *PloS One* (2012) 7(5):e37928. doi: 10.1371/journal.pone.0037928

12. Espinal AC, Buas MF, Wang D, Cheng DT, Sucheston-Campbell L, Hu Q, et al. FOXA1 hypermethylation: link between parity and ER-negative breast cancer in African American women? *Breast Cancer Res Treat* (2017) 166(2):559–68. doi: 10.1007/s10549-017-4418-y

13. Chiu AM, Mitra M, Boymoushakian L, Coller HA. Integrative analysis of the inter-tumoral heterogeneity of triple-negative breast cancer. *Sci Rep* (2018) 8(1):11807. doi: 10.1038/s41598-018-29992-5

14. Stirzaker C, Zotenko E, Song JZ, Qu W, Nair SS, Locke WJ, et al. Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. *Nat Commun* (2015) 6:5899. doi: 10.1038/ncomms6899

15. Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, et al. DNA Methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables. *Mol Oncol* (2011) 5(1):77–92. doi: 10.1016/j.molonc.2010.11.002

16. Francis WR, Worheide G. Similar ratios of introns to intergenic sequence across animal genomes. *Genome Biol Evol* (2017) 9(6):1582–98. doi: 10.1093/gbe/evx103

17. Fleischer T, Tekpli X, Mathelier A, Wang S, Nebdal D, Dhakal HP, et al. DNA Methylation at enhancers identifies distinct breast cancer lineages. *Nat Commun* (2017) 8(1):1379. doi: 10.1038/s41467-017-00510-x

18. Sur I, Taipale J. The role of enhancers in cancer. *Nat Rev Cancer* (2016) 16 (8):483–93. doi: 10.1038/nrc.2016.62

19. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* (2009) 462(7271):315–22. doi: 10.1038/nature08514

20. Bell RE, Golan T, Sheinboim D, Malcov H, Amar D, Salamon A, et al. Enhancer methylation dynamics contribute to cancer plasticity and patient mortality. *Genome Res* (2016) 26(5):601–11. doi: 10.1101/gr.197194.115

21. Cho JW, Shim HS, Lee CY, Park SY, Hong MH, Lee I, et al. The importance of enhancer methylation for epigenetic regulation of tumorigenesis in squamous lung cancer. *Exp Mol Med* (2022) 54(1):12–22. doi: 10.1038/s12276-021-00718-4

22. Gong Z, Chen J, Wang J, Liu S, Ambrosone CB, Higgins MJ. Differential methylation and expression patterns of microRNAs in relation to breast cancer subtypes among American women of African and European ancestry. *PloS One* (2021) 16(3):e0249229. doi: 10.1371/journal.pone.0249229

23. Ambrosone CB, Ciupak GL, Bandera EV, Jandorf L, Bovbjerg DH, Zirpoli G, et al. Conducting molecular epidemiological research in the age of HIPAA: a multi-institutional case-control study of breast cancer in African-American and European-American women. *J Oncol* (2009) 2009:871250. doi: 10.1155/2009/871250

24. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* (2012) 28(6):882–3. doi: 10.1093/bioinformatics/bts034

25. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, et al. IMA: an r package for high-throughput analysis of illumina's 450K infinium methylation data. *Bioinformatics* (2012) 28(5):729–30. doi: 10.1093/bioinformatics/bts013

26. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods* (2012) 9(4):357–9. doi: 10.1038/nmeth.1923

27. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the illumina infinium HumanMethylation450 microarray. *Epigenetics* (2013) 8(2):203–9. doi: 10.4161/epi.23470

28. Zhang X, Mu W, Zhang W. On the analysis of the illumina 450k array data: probes ambiguously mapped to the human genome. *Front Genet* (2012) 3:73. doi: 10.3389/fgene.2012.00073

29. Jia Y, Chng WJ, Zhou J. Super-enhancers: critical roles and therapeutic targets in hematologic malignancies. *J Hematol Oncol* (2019) 12(1):77. doi: 10.1186/s13045-019-0757-y

30. Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, Oostra BA, et al. A long-range shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* (2003) 12(14):1725–35. doi: 10.1093/hmg/ddg180

31. Yip KY, Cheng C, Bhardwaj N, Brown JB, Leng J, Kundaje A, et al. Classification of human genomic regions based on experimentally determined binding sites of more than 100 transcription-related factors. *Genome Biol* (2012) 13(9):R48. doi: 10.1186/gb-2012-13-9-r48

32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8

33. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* (2017) 2017:bax028. doi: 10.1093/database/bax028

34. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. *Cell* (2013) 155(4):934–47. doi: 10.1016/j.cell.2013.09.053

35. Cheng CJ, Lin YC, Tsai MT, Chen CS, Hsieh MC, Chen CL, et al. SCUBE2 suppresses breast tumor cell proliferation and confers a favorable prognosis in invasive breast cancer. *Cancer Res* (2009) 69(8):3634–41. doi: 10.1158/0008-5472.CAN-08-3615

36. Lin YC, Lee YC, Li LH, Cheng CJ, Yang RB. Tumor suppressor SCUBE2 inhibits breast-cancer cell migration and invasion through the reversal of epithelial-mesenchymal transition. *J Cell Sci* (2014) 127(Pt 1):85–100. doi: 10.1158/1538-7445.AM2014-595

37. Satow R, Aiga Y, Watanabe T, Ishizuka N, Yoneda A, Fukami K. Zic family member 5 promotes survival in human pancreatic cancer and cholangiocarcinoma cells. *Biochem Biophys Rep* (2022) 31:101289. doi: 10.1016/j.bbrep.2022.101289

38. de Bessa Garcia SA, Araujo M, Pereira T, Mouta J, Freitas R. HOX genes function in breast cancer development. *Biochim Biophys Acta Rev Cancer* (2020) 1873 (2):188358. doi: 10.1016/j.bbcan.2020.188358

39. di Iulio J, Bartha I, Wong EHM, Yu HC, Lavrenko V, Yang D, et al. The human noncoding genome defined by genetic diversity. *Nat Genet* (2018) 50(3):333–7. doi: 10.1038/s41588-018-0062-7

40. Almamun M, Kholod O, Stuckel AJ, Levinson BT, Johnson NT, Arthur GL, et al. Inferring a role for methylation of intergenic DNA in the regulation of genes aberrantly expressed in precursor b-cell acute lymphoblastic leukemia. *Leuk Lymphoma* (2017) 58 (9):1–12. doi: 10.1080/10428194.2016.1272683

41. Gong Z, Wang J, Wang D, Buas MF, Ren X, Freudenheim JL, et al. Differences in microRNA expression in breast cancer between women of African and European ancestry. *Carcinogenesis* (2019) 40(1):61–9. doi: 10.1093/carcin/bgy134

42. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics* (2011) 98(4):288–95. doi: 10.1016/j.ygeno.2011.07.007

43. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* (2011) 6(6):692–702. doi: 10.4161/epi.6.6.16196

44. Kulis M, Esteller M. DNA Methylation and cancer. *Adv Genet* (2010) 70:27–56. doi: 10.1016/B978-0-12-380866-0.60002-2

45. Jadhav RR, Ye Z, Huang RL, Liu J, Hsu PY, Huang YW, et al. Genome-wide DNA methylation analysis reveals estrogen-mediated epigenetic repression of metallothionein-1 gene cluster in breast cancer. *Clin Epigenet* (2015) 7(1):13. doi: 10.1186/s13148-015-0045-9

46. Smith RG, Hannon E, De Jager PL, Chibnik L, Lott SJ, Condliffe D, et al. Elevated DNA methylation across a 48-kb region spanning the HOXA gene cluster is associated with alzheimer's disease neuropathology. *Alzheimers Dement* (2018) 14 (12):1580–8. doi: 10.1016/j.jalz.2018.01.017

47. Afzal Z, Krumlauf R. Transcriptional regulation and implications for controlling hox gene expression. *J Dev Biol* (2022) 10(1):4. doi: 10.3390/jdb10010004

48. Boimel PJ, Cruz C, Segall JE. A functional *in vivo* screen for regulators of tumor progression identifies HOXB2 as a regulator of tumor growth in breast cancer. *Genomics* (2011) 98(3):164–72. doi: 10.1016/j.ygeno.2011.05.011

49. De Kumar B, Krumlauf R. HOXs and lincRNAs: two sides of the same coin. *Sci Adv* (2016) 2(1):e1501402. doi: 10.1126/sciadv.1501402

50. Fukushige S, Horii A. DNA Methylation in cancer: a gene silencing mechanism and the clinical potential of its biomarkers. *Tohoku J Exp Med* (2013) 229(3):173–85. doi: 10.1620/tjem.229.173

51. Cheng TD, Yao S, Omilian AR, Khoury T, Buas MF, Payne-Ondracek R, et al. FOXA1 protein expression in ER(+) and ER(-) breast cancer in relation to parity and breastfeeding in black and white women. *Cancer Epidemiol Biomarkers Prev* (2020) 29 (2):379–85. doi: 10.1158/1055-9965.EPI-19-0787

52. Lin L, Yee SW, Kim RB, Giacomini KM. SLC transporters as therapeutic targets: emerging opportunities. *Nat Rev Drug Discov* (2015) 14(8):543–60. doi: 10.1038/nrd4626

53. Barman SK, Zaman MS, Veljanoski F, Malladi CS, Mahns DA, Wu MJ. Expression profiles of the genes associated with zinc homeostasis in normal and cancerous breast and prostate cells. *Metallomics* (2022) 14(8):mfac038. doi: 10.1093/mtomcs/mfac038

54. Liu L, Yang J, Wang C. Analysis of the prognostic significance of solute carrier (SLC) family 39 genes in breast cancer. *Biosci Rep* (2020) 40(8):BSR20200764. doi: 10.1042/BSR20200764

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 19 July 2023
DOI 10.3389/fonc.2023.1203483

OPEN ACCESS

EDITED BY
Iman Mamdouh Talaat,
University of Sharjah, United Arab Emirates

REVIEWED BY
Sung Hoon Sim,
National Cancer Center, Republic of Korea
Mohammed A. Al-Obaide,
Texas Tech University Health Science Center Amarillo, United States

*CORRESPONDENCE
Nady El Hajj
nelhajj@hbku.edu.qa

RECEIVED 10 April 2023
ACCEPTED 03 July 2023
PUBLISHED 19 July 2023

CITATION
Bejaoui Y, Alresheq S, Durand S, Vilaire-Meunier M, Maillebouis L, Zen AAH, Mégarbané A and Hajj NE (2023) DNA methylation profiling in Trisomy 21 females with and without breast cancer. *Front. Oncol.* 13:1203483. doi: 10.3389/fonc.2023.1203483

DNA methylation profiling in Trisomy 21 females with and without breast cancer

Yosra Bejaoui[1], Sara Alresheq[1], Sophie Durand[2], Marie Vilaire-Meunier[2], Louise Maillebouis[2], Ayman Al Haj Zen[1], André Mégarbané[2,3] and Nady El Hajj[1,4]*

[1]College of Health and Life Sciences, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar, [2]Institut Jérôme Lejeune, Paris, France, [3]Department of Human Genetics, Gilbert and Rose-Marie Chagoury School of Medicine, Lebanese American University, Byblos, Lebanon, [4]College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar

Background: Down Syndrome (DS) is the most common chromosome anomaly in humans and occurs due to an extra copy of chromosome 21. The malignancy profile in DS is unique, since DS patients have a low risk of developing solid tumors such as breast cancer however they are at higher risk of developing acute myeloid leukemia and acute lymphoblastic leukemia.

Methods: In this study, we investigated DNA methylation signatures and epigenetic aging in DS individuals with and without breast cancer. We analyzed DNA methylation patterns in Trisomy 21 (T21) individuals without breast cancer (T21-BCF) and DS individuals with breast cancer (T21-BC), using the Infinium Methylation EPIC BeadChip array.

Results: Our results revealed several differentially methylated sites and regions in the T21-BC patients that were associated with changes in gene expression. The differentially methylated CpG sites were enriched for processes related to serine-type peptidase activity, epithelial cell development, GTPase activity, bicellular tight junction, Ras protein signal transduction, etc. On the other hand, the epigenetic age acceleration analysis showed no difference between T21-BC and T21-BCF patients.

Conclusions: This is the first study to investigate DNA methylation changes in Down syndrome women with and without breast cancer and it could help shed light on factors that protect against breast cancer in DS.

KEYWORDS

Down syndrome, breast cancer, DNA methylation, epigenetics, Trisomy 21

Introduction

Down syndrome (DS) is a genetic disorder caused by an additional copy of all or part of chromosome 21 resulting in 47 chromosomes instead of the typical 46 chromosomes. The etiology of DS was identified following the discovery of karyotyping techniques when the French geneticist Jérôme Lejeune reported that an extra chromosome 21 results in the phenotypic features and intellectual disability associated with DS (1). DS is considered the most common chromosomal condition in humans occurring in 1 out of every 700 newborn babies (2). DS has three different forms including Trisomy 21 (nondisjunction), mosaicism, and translocation. Nondisjunction of chromosome 21, also called standard trisomy 21, is the most common DS type and accounts for ~ 95% of all cases. The cause of this chromosomal non-disjunction occurs mainly during maternal meiotic division (~88% of the cases). Whereas, ~ 5-10% of the cases are caused by non-disjunction during spermatogenesis and a small percentage of cases are due to mitotic error or occur during the first mitotic divisions of the embryo (3–5).

Trisomy 21 is associated with more than 100 features including intellectual disability, distinctive facial features, early aging, neurodegeneration, and muscle hypotonia during childhood (6). Intellectual disability is the most common feature in DS patients, where it usually ranges from mild to moderate. Besides, DS patients have a high incidence of congenital heart disease, early onset Alzheimer's disease, gastrointestinal and skeletal malformations, and a diversity of neurobehavioral abnormalities (7–9). Even though DS patients are predisposed to developing acute lymphoblastic and myeloblastic leukemia during childhood, solid tumors seem to be extremely rare in both children and adults (10–16). Several epidemiological studies suggested that the risk of developing solid tumors in DS patients is at least 12 times lower than that of the general population (16, 17). For example, breast cancer (BC) is almost non-existent in DS females, despite genetic instability, deficiencies in DNA repair, increased oxidative stress, sedentary lifestyle, higher obesity rates, and increased DNA damage. Environmental factors including decreased exposure to estrogens and low alcohol consumption are not sufficient by themselves to explain the low rate of BC in DS females (17–19). Therefore, it would be important to study possible molecular mechanisms that protect against the development of breast cancer in Down syndrome.

Epigenetic dysregulation in response to an additional copy of chromosome 21 has been reported to affect the entire genome and not only genes located on chromosome 21 (20–24). Those changes arise during development and systemically affect multiple tissues (21, 25). Epigenetic clocks based on DNA methylation measurements have been used to estimate a person's biological age and epigenetic aging acceleration. Epigenetic age acceleration has been reported to be associated with cancer risk, prognosis, and survival (26). Furthermore, patients with Down syndrome were reported to have drastic epigenetic age acceleration that was even higher than in certain progeroid syndromes (27, 28). Taking into account the occurrence of epigenetic alterations in most cancers and that they act as drivers to cancer progression, it would be important to study whether DNA methylation alterations affecting certain genes/pathways confers protection against breast cancer in DS. Therefore, we performed a genome-wide DNA methylation analysis in DS females with and without BC to determine epigenetically dysregulated regions linked to the lower BC frequency in DS. In addition, we compared epigenetic age acceleration in DS females with and without BC.

Materials and methods

Samples and data collection

A total of 5532 files were screened at the Jérôme Lejeune Institute (CRB BioJeL, Paris, France) to identify two DS females with homogeneous Trisomy 21 (T21) diagnosed with breast cancer (no mosaicism or translocation cases were included). Sequencing analysis revealed no pathogenic or likely pathogenic variants in genes associated with an increased risk of breast cancer in the selected samples. A total of 10 age matched DS females with homogeneous T21 and without breast cancer (or any mammary lesion) were selected as controls (Supplementary Table 1). All the recruited DS women were > 34 years old, without any chronic medications or social problems. No breast cancer was recorded in the families of the DS women in this study. Whole blood samples were collected from all the patients and human peripheral blood mononuclear cells (PBMCs) were isolated. DNA was extracted from both whole blood and PBMCs. Written informed consent was obtained from the parents or guardians for all participants included in the research study.

DNA methylation quantification using EPIC arrays

DNA methylation profiling was performed for two T21 females with breast cancer (referred to as T21-BC) and for 10 T21 females without breast cancer (T21-BCF) (n=10) using the Illumina Infinium Epic array. DNA samples were processed on Illumina Infinium Epic array according to the manufacturer's protocol. Briefly, 500 ng DNA for each sample was bisulfite converted using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA, USA). Afterwards, bisulfite converted DNA was whole-genome amplified, enzymatically fragmented, and hybridized to Infinium Methylation EPIC BeadChips. Array scanning was performed via the Illumina iScan. To avoid batch effects, all samples were processed simultaneously and measured samples were randomly hybridized on the arrays. Idat files were exported and analyzed with the R software package (version 3.2.2) and the BioConductor platform (version 3.2).

Differential DNA methylation analysis

The RnBeads package was used for differential methylation analysis (29). First, the data quality was assessed and probes mapping to multiple regions in the genome (Cross-reactive) or overlapping SNPs were removed. Furthermore, probes with

unreliable measurements were removed via greedycut prior to further analysis. Next, additional filtering of polymorphic probes in the European, admixed American, South and East Asian, and African was applied using "filtering.blacklist" option (30). Data was normalized using Dasen and probes located on the X chromosome were retained because only females were analyzed. A total of 534862 (whole blood) and 534049 (PBMC) probes were finally retained for differential DNA methylation analysis. Inference for blood cell composition was performed using the Houseman method (31). Next, a limma based approach was used to correct for cell type composition, age, and surrogate variables. Differential methylation analysis was performed at the single CpG site level and at the level of promoters, CpG islands, and tiling windows (5Kb). Combined p-values were calculated and adjusted for multiple testing using false discovery rate (FDR) correction. Gene Ontology (GO) enrichment analysis was performed via the methylglm function from the methylGSA package (32).

Calculating DNA methylation age and age acceleration

Epigenetic age acceleration was measured using several epigenetic clocks that utilize different CpG sites to estimate DNA methylation (DNAm) age using the DNAm age calculator (https://dnamage.genetics.ucla.edu/) with the normalization option selected.

DNA methylation data from breast cancer patients with normal karyotype

DNA methylation profiles of women with normal karyotype diagnosed with breast cancer (n=43) were downloaded from NCBI's Gene Expression Omnibus (GEO Series accession: GSE104942). The Raw (IDAT) files were processed as previously described in the "Differential DNA methylation analysis" section. In total, the studied dataset was comprised of blood DNA methylation data measured via the Illumina EPIC arrays on 43 Breast cancer patients and 12 controls.

Results

Differentially methylated sites in Down's syndrome females with breast cancer

To identify epigenetically altered regions associated with BC in DS, we measured DNA methylation levels in T21 breast cancer patients (T21-BC, n=2) vs T21 breast cancer-free patients (T21-BCF, n=10) using the Illumina EPIC arrays. DNA methylation was profiled in DNA isolated from both whole blood and from PBMCs. The number of T21-BC samples was limited because only two T21 females with breast cancer were identified after screening 5532 files at Jérôme Lejeune Institute. For this reason, we decided to measure DNA methylation in duplicates across both whole blood and PBMC samples. First, we compared the deconvoluted blood cell proportions in T21-BC vs. T21-BCF as estimated by the Houseman method, which revealed no change in immune blood cell proportions (Figure 1A).

Next, differential DNA methylation analysis was performed to compare T21-BC vs. T21-BCF. The differential methylation was assessed primarily at the CpG sites level in addition to the region level including promoter, CpG Island, and tiling regions using a 5Kb sliding window. We did not observe any significance at the CpG site or the region level after FDR adjustment when adjusting for age, gender, cell type composition, and surrogate variables. This

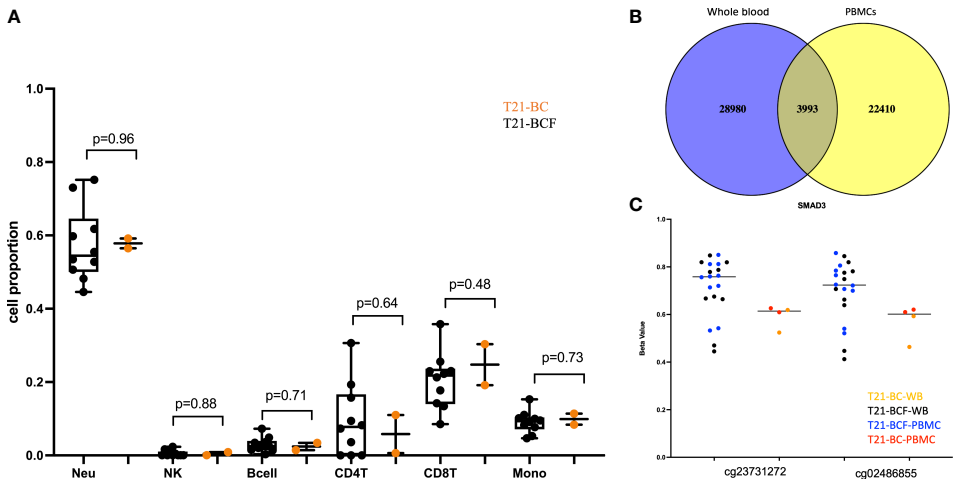


FIGURE 1
(A) Comparison of deconvoluted cell proportions measured via the Houseman method in whole blood of Trisomy 21 (T21) individuals with breast cancer (T21-BC) vs T21 without breast cancer (T21-BCF); **(B)** Ven-diagram of significant sites with unadjusted p-value <0.05 when comparing whole blood and PBMCs of T21-BC vs T21BCF; **(C)** differentially methylated probes s in *SMAD3* with more than 9% methylation differences in whole blood and PBMC and present in the list of differentially expressed genes in T21-BC performed on the same samples. T21-BC-WB: Trisomy 21 (T21) individuals with breast cancer whole blood analysis. T21-BCF-WB: Trisomy 21 (T21) individuals without breast cancer whole blood DNA methylation analysis. T21-BCF-PBMC and T21-BC-PBMC indicate the DNA methylation analysis in peripheral blood mononuclear cells.

could be related to the low sample number in the T21-BC group. For this reason, we looked at common significant sites/regions with unadjusted p-value <0.05 between T21-BC and T21-BCF. In total, 32973 and 26403 CpG sites were significant before FDR adjustment in WB and PBMCs, respectively. Out of which, 3993 CpG sites were common between the whole blood and PBMC samples (Figure 1B). Out of those, 3087 had a similar direction of DNA methylation change when comparing T21-BC and T21-BCF in both WB and PBMC samples. When we filtered for ≥ 3% methylation in both tissues, a total of 1601 CpG sites were retained. Next, we applied the methylGSA package to test for gene ontology (GO) and KEGG pathway enrichment in those CpG sites, after adjusting for probe bias distribution across genes in the EPIC arrays. The GO enrichment analysis revealed several significant terms including serine-type peptidase activity, exopeptidase activity, serine hydrolase activity, epithelial cell development, *etc* (Supplementary Table 2). On the other hand, the KEGG analysis did not reveal any pathway enrichment for the 1601 CpG sites. We additionally investigated epigenetic age acceleration in T21-BC vs T21-BCF using the Horvath clock, GrimAge and PhenoAge, which revealed no DNA methylation age acceleration difference T21-BC in whole blood (Figure 2) and PBMC samples.

Differentially methylated regions associated with breast cancer in Down's syndrome

Next, we looked at the promoter region where we could identify 832 significant promoters (unadjusted p-value <0.05) in whole blood and 744 significant promoters in PBMCs. A total of 78 promoters were significant in both analyses when comparing whole blood and PBMCs from T21-BC vs T21-BCF with the same direction of methylation change. Out of which, 22 promoters had > 2 CpG sites and ≥ 3% methylation in both whole blood and PBMC samples (Supplementary Table 3). For the CpG Island analysis, we could observe 131 common significant CGIs with the same direction of methylation change, including 43 with > 2 CpG sites and ≥ 3% methylation (Supplementary Table 4). For the tiling analysis, we could identify 677 regions (5Kb) differentially methylated in a similar direction in both datasets, however only 79 remained after filtering using the previously defined criteria (≥ 2 CpG sites, ≥ 3% methylation). Next, we tested whether the identified DMPs/DMRs are similarly epigenetically dysregulated in blood DNA of breast cancer patients. The studied dataset (GSE104942) contained blood DNA methylation data of 43 Breast cancer patients and 12 healthy controls. This analysis revealed no common significant DMRs between the T21-BC list and the differentially methylated genes in breast cancer patients. Two DMPs (cg05997779 and cg26845300) were similarly epigenetically altered in both datasets, however, they exhibited different direction of DNA methylation change.

Transcriptional changes in epigenetically dysregulated genes associated with breast cancer in Down's syndrome

Finally, we compared the differentially methylated sites/regions to the list of 183 differentially expressed genes in T21-BC identified following RNA-seq on the same samples (33). Here, we could observe 37 differentially methylated probes (DMPs) associated with differentially expressed genes and same direction of methylation change in both DNA methylation datasets. When we filtered for ≥ 3% methylation difference, we could only detect 12 CpG sites that fit this criteria including two close DMPs in *SMAD3* with more than 9% methylation differences in all comparisons (Figure 1C). In addition, there was a single CpG site located on chromosome 21 in the *BACH1* gene. Next, we checked the promoter and tiling differentially methylated regions (DMR), which revealed one DMR in the promoter analysis and two in the tiling region analysis. The common promoter was located in the gene TNFAIP3 Interacting Protein 1 (*TNIP1*) (Figure 3A), whereas the 5Kb tiling regions were located in the KRAB box domain

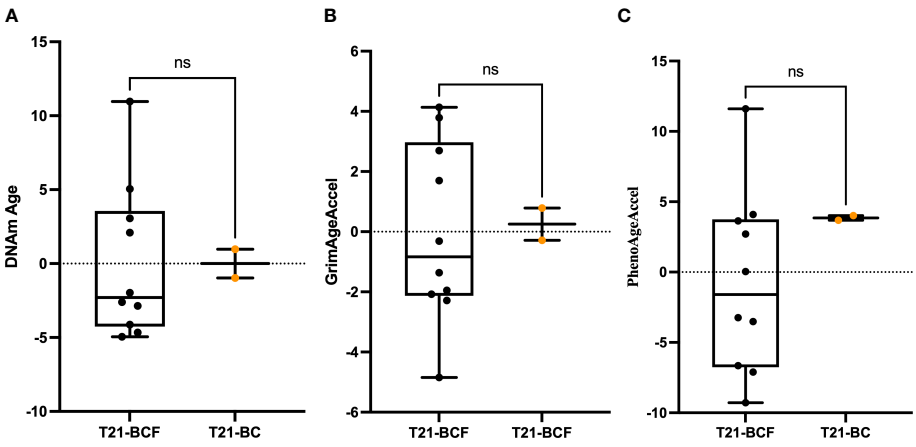


FIGURE 2
Epigenetic age acceleration in whole blood DNA of T21-BC vs T21-BCF using the **(A)** Horvath, **(B)** PhenoAge and, **(C)** GrimAge clocks. DNA methylation (DNAm) age acceleration, which represents the residual of regressing epigenetic age on chronological age is shown on the y-axes. ns, not significant.

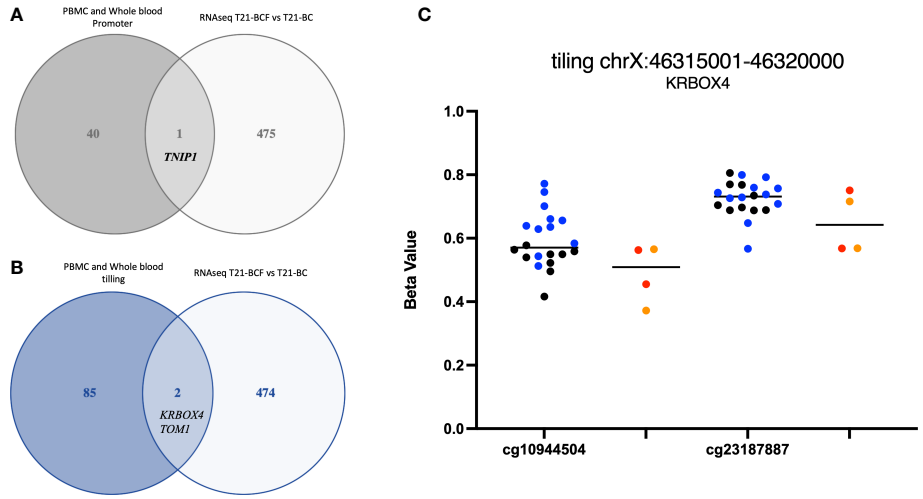


FIGURE 3
Ven-diagram displaying **(A)** differentially methylated promoters and **(B)** differentially methylated regions following a tiling analysis of differentially expressed genes in Trisomy 21 (T21) individuals with breast cancer (T21-BC) vs T21 without breast cancer (T21-BCF). **(C)** *KRBOX4* DNA methylation distribution in T21 BC vs T21 BCF. Genomic coordinates based on genome assembly GRCh37 (hg19).

containing 4 (*KRBOX4*) and Target Of Myb1 Membrane Trafficking Protein (*TOM1*) gene (Figures 3B, C). None of the CpG Island associated genes were differentially expressed in T21-BC.

Discussion

To understand whether epigenetic dysregulation might explain the lower frequency of breast cancer in DS, we profiled DNA methylation in 12 women with T21 including two with breast cancer and 10 without breast cancer. DNA methylation was measured in both whole blood DNA and PBMCs as replicates due to the small number of breast cancer patients with T21.

The differential methylation analysis at the single CpG site level revealed 1601 DMPs with the same direction in methylation change. The gene ontology analysis revealed enrichment for serine-type peptidase activity, exopeptidase activity, serine hydrolase activity, epithelial cell development, endothelium development, transcription coactivator activity, GTPase activity, GTP binding, bicellular tight junction, and Ras protein signal transduction. Cell surface anchored serine proteases are deregulated in cancer cells and contribute to tumour invasion and metastasis (34). Evidently, Ras protein signal transduction is extremely important in cancer where mutations in the RAS genes were the first mutations reported in human cancers (35–37). The expression and activity small GTPases subfamily of "Ras-homology" (Rho) GTPase are known to be linked with breast tumour progression, angiogenesis, and metastasis (38). Similarly, bicellular tight junctions play a role in the epithelial-mesenchymal transition, which is essential in cancer progression (39). Excessive angiogenesis is a crucial component of tumour growth, invasiveness, and metastasis (40). The individuals with DS showed an elevated expression of *DSCR1* on the extra copy of chromosome 21, which is known to inhibit the growth of new blood vessels "angiogenesis" by suppressing vascular endothelial growth factor (VEGF)-mediated angiogenic signalling (41). In the current study, GO was enriched for endothelium development in which many genes overlap with angiogenesis. In addition, the crosstalk between several Rho GTPases and VEGF signalling is essential to control the process of angiogenesis (42–45). The epigenetic dysregulation in DS due to dosage imbalance of an additional chromosome 21 has been reported to occur extensively throughout the genome and is not restricted to genes located on chromosome 21 (21). This might lead to DNA methylation alterations in genes associated with the previously mentioned GO terms. This epigenetic dysregulation might confer protection to DS patients from breast cancer, which might help explain its reduced risk.

Furthermore, the comparison of the differentially methylated sites to differentially expressed genes following RNA-seq on the same set of samples identified 12 DMPs including two CpG sites in *SMAD3* with > 9% methylation difference. *Smad3* is a major transcription factor mediating transforming growth factor-β (TGF-β) signaling (46). The TGF-β-Smad3 signaling has important roles in differentiation, apoptosis, and epithelial-mesenchymal transition (EMT) (46, 47). SMAD-dependent signaling mediated by TGF-β has two opposing roles in cancer, where it first acts as a tumor suppressor in the initial phase, however in more advanced stages it is involved in inducing invasion and metastasis (48, 49). The regulation of estrogen receptor signaling pathways via TGF-beta was shown to be mediated by SMAD3, which indicates a role of *SMAD3* in breast cancer progression (50). Furthermore, we identified a DMP located on chromosome 21 in the BTB and CNC homology1 (*BACH1*) gene. *BACH1* encodes a transcription factor that is upregulated in tumours from triple-negative breast cancer (TNBCs) patients (51). *BACH1* has been previously reported as a regulator of metastasis in TNBCs and its

gene signature was shown to predict poor outcomes in breast cancer (52). The promoter of *TNIP1* was differentially methylated and transcriptionally dysregulated in the T21-BC group. The tumor necrosis factor α–induced protein 3–interacting protein 1 (*TNIP1)* is part of the NF-κB and RAR signaling pathways (53, 54). *TNIP1* was one of the stromal genes exhibiting expression changes when comparing adenomas vs cancer-associated stroma (55). Two DMRs were identified in *KRBOX4* and *TOM1* in the tiling analysis located in promoter flanking regions. *KRBOX4* is located on the X chromosome and no studies so far have provided any link to breast cancer. *TOM1* is required for autophagosome maturation and endosomal trafficking (56). *TOM1* additionally represses Toll-like receptor signalling and plays a role in immune receptor recycling (57, 58). Mutations in *TOM1* have been recently shown to be associated with early-onset autoimmunity and combined immunodeficiency (59). In addition, it is important to mention that we did not observe any DNA methylation changes at the region level in GTPases of the immunity-associated proteins (GIMAPs) despite their recently identified tumour suppressive role against breast cancer in DS (33). Therefore, it seems that upregulation of *GIMAPs* in T21 women is not associated with changes in DNA methylation.

DS patients are known to exhibit strong epigenetic age acceleration and for this reason we tested T21-BC patients age acceleration in comparison to the T21-BCF group. Epigenetic age acceleration have been previously shown to occur in several diseases including cancer, and can be used as a potential biomarker for early disease detection (26). Furthermore, a longitudinal study reported epigenetic age acceleration measured on blood DNA to be associated with a higher risk of developing breast cancer (60). The PhenoAge clock was also shown to measure increased epigenetic age acceleration in breast tissue of from breast cancer patients. However, our analysis revealed no difference in epigenetic age acceleration between T21-BC and T21-BCF using various clocks. This might be related to the drastic increase in epigenetic age acceleration in DS, which masks the effects of breast cancer on DNA methylation age.

The limitations of our study are the small sample size, however this is related to the uniqueness of the condition since breast cancer is almost non-existent in T21 patients. Furthermore, we have only looked at blood DNA in this study, which might not reflect similar epigenetic changes to target tissues involved in disease pathogenesis. Therefore, future studies should include additional tissues to determine whether the observed epigenetic changes related to breast cancer in DS are systemic or only restricted to one tissue.

In conclusion, this is the first study to investigate the DNA methylation profile in Down syndrome women suffering from breast cancer. The identified differentially methylated genes/regions could help us better understand factors that protect against breast cancer, which can provide new avenues for potential therapeutic targets or preventive approaches.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board Statement. This research was conducted according to the guidelines of the Declaration of Helsinki. It was approved by the Institute Jérôme Lejeune Committee on Clinical Investigation and the Institutional Review Board (IRB) of the Qatar Biomedical Research Institute (reference number: QBRI-IRB 2021-07-098). Written informed consent was obtained from the parents or guardians prior to the participants being included in the research study. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AM and NH designed the study; YB and SA analyzed the data, performed experiments, and prepared figures; SD, MV, LM, and AM collected the samples and clinical data; NH and AA wrote the manuscript, NH, AM, and AA reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported with funding by Qatar Foundation to the College of Health and Life Sciences, Hamad Bin Khalifa (NH).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1203483/full#supplementary-material

References

1. Lejeune J, Gauthier M, Turpin R. [Human chromosomes in tissue cultures]. *C R Hebd Seances Acad Sci* (1959) 248:602–3.

2. Mai CT, Isenburg JL, Canfield MA, Meyer RE, Correa A, Alverson CJ, et al. National population-based estimates for major birth defect-2014. *Birth Defects Res* (2019) 111:1420–35. doi: 10.1002/bdr2.1589

3. Antonarakis SE. Parental origin of the extra chromosome in trisomy 21 as indicated by analysis of DNA polymorphisms. down syndrome collaborative group. *N Engl J Med* (1991) 324:872–6. doi: 10.1056/NEJM199103283241302

4. Sherman SL, Takaesu N, Freeman SB, Grantham M, Phillips C, Blackston RD, et al. Trisomy 21: association between reduced recombination and nondisjunction. *Am J Hum Genet* (1991) 49:608–20.

5. Mantikou E, Wong KM, Repping S, Mastenbroek S. Molecular origin of mitotic aneuploidies in preimplantation embryos. *Biochim Biophys Acta* (2012) 1822:1921–30. doi: 10.1016/j.bbadis.2012.06.013

6. Megarbane A, Ravel A, Mircher C, Sturtz F, Grattau Y, Rethore MO, et al. The 50th anniversary of the discovery of trisomy 21: the past, present, and future of research and treatment of down syndrome. *Genet Med* (2009) 11:611–6. doi: 10.1097/GIM.0b013e3181b2e34c

7. Epstein CJ. The consequences of chromosome imbalance. *Am J Med Genet Suppl* (1990) 7:31–7. doi: 10.1002/ajmg.1320370706

8. Nelson L, Johnson JK, Freedman M, Lott I, Groot J, Chang M, et al. Learning and memory as a function of age in down syndrome: a study using animal-based tasks. *Prog Neuropsychopharmacol Biol Psychiatry* (2005) 29:443–53. doi: 10.1016/j.pnpbp.2004.12.009

9. Dierssen M, Herault Y, Estivill X. Aneuploidy: from a physiological mechanism of variance to down syndrome. *Physiol Rev* (2009) 89:887–920. doi: 10.1152/physrev.00032.2007

10. Satge D, Sommelet D, Geneix A, Nishi M, Malet P, Vekemans M. A tumor profile in down syndrome. *Am J Med Genet* (1998) 78:207–16. doi: 10.1002/(SICI)1096-8628(19980707)78:3<207::AID-AJMG1>3.0.CO;2-M

11. Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with down's syndrome. *Lancet* (2000) 355:165–9. doi: 10.1016/S0140-6736(99)05264-2

12. Patja K, Pukkala E, Sund R, Iivanainen M, Kaski M. Cancer incidence of persons with down syndrome in Finland: a population-based study. *Int J Cancer* (2006) 118:1769–72. doi: 10.1002/ijc.21518

13. Sullivan SG, Hussain R, Glasson EJ, Bittles AH. The profile and incidence of cancer in down syndrome. *J Intellect Disabil Res* (2007) 51:228–31. doi: 10.1111/j.1365-2788.2006.00862.x

14. Nizetic D, Groet J. Tumorigenesis in down's syndrome: big lessons from a small chromosome. *Nat Rev Cancer* (2012) 12:721–32. doi: 10.1038/nrc3355

15. Hasle H, Friedman JM, Olsen JH, Rasmussen SA. Low risk of solid tumors in persons with down syndrome. *Genet Med* (2016) 18:1151–7. doi: 10.1038/gim.2016.23

16. Rethore MO, Rouesse J, Satge D. Cancer screening in adults with down syndrome, a proposal. *Eur J Med Genet* (2020) 63:103783. doi: 10.1016/j.ejmg.2019.103783

17. Satge D, Sasco AJ, Pujol H, Rethore MO. Breast cancer in women with trisomy 21. *Bull Acad Natl Med* (2001) 185:1239–52.

18. Ayed W, Gouas L, Penault-Llorca F, Amouri A, Tchirkov A, Vago P. [Trisomy 21 and cancers]. *Morphologie* (2012) 96:57–66. doi: 10.1016/j.morpho.2012.10.001

19. Osuna-Marco MP, Lopez-Barahona M, Lopez-Ibor B, Tejera AM. Ten reasons why people with down syndrome are protected from the development of most solid tumors -a review. *Front Genet* (2021) 12:749480. doi: 10.3389/fgene.2021.749480

20. Mendioroz M, Do C, Jiang X, Liu C, Darbary HK, Lang CF, et al. Trans effects of chromosome aneuploidies on DNA methylation patterns in human down syndrome and mouse models. *Genome Biol* (2015) 16:263. doi: 10.1186/s13059-015-0827-6

21. Do C, Xing Z, Yu YE, Tycko B. Trans-acting epigenetic effects of chromosomal aneuploidies: lessons from down syndrome and mouse models. *Epigenomics* (2017) 9:189–207. doi: 10.2217/epi-2016-0138

22. El Hajj N, Dittrich M, Haaf T. Epigenetic dysregulation of protocadherins in human disease. *Semin Cell Dev Biol* (2017) 69:172–82. doi: 10.1016/j.semcdb.2017.07.007

23. Almenar-Queralt A, Merkurjev D, Kim HS, Navarro M, Ma Q, Chaves RS, et al. Chromatin establishes an immature version of neuronal protocadherin selection during the naive-to-primed conversion of pluripotent stem cells. *Nat Genet* (2019) 51:1691–701. doi: 10.1038/s41588-019-0526-4

24. Haertle L, Muller T, Lardenoije R, Maierhofer A, Dittrich M, Riemens RJM, et al. Methylomic profiling in trisomy 21 identifies cognition- and alzheimer's disease-related dysregulation. *Clin Epigenet* (2019) 11:195. doi: 10.1186/s13148-019-0787-x

25. El Hajj N, Dittrich M, Bock J, Kraus TF, Nanda I, Muller T, et al. Epigenetic dysregulation in the developing down syndrome cortex. *Epigenetics* (2016) 11:563–78. doi: 10.1080/15592294.2016.1192736

26. Salameh Y, Bejaoui Y, El Hajj N. DNA Methylation biomarkers in aging and age-related diseases. *Front Genet* (2020) 11:171. doi: 10.3389/fgene.2020.00171

27. Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, et al. Accelerated epigenetic aging in down syndrome. *Aging Cell* (2015) 14:491–5. doi: 10.1111/acel.12325

28. Bejaoui Y, Razzaq A, Yousri NA, Oshima J, Megarbane A, Qannan A, et al. DNA Methylation signatures in blood DNA of Hutchinson-gilford progeria syndrome. *Aging Cell* (2022) 21:e13555. doi: 10.1111/acel.13555

29. Muller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. *Genome Biol* (2019) 20:55. doi: 10.1186/s13059-019-1664-9

30. Mccartney DL, Walker RM, Morris SW, Mcintosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the illumina infinium MethylationEPIC BeadChip. *Genom Data* (2016) 9:22–4. doi: 10.1016/j.gdata.2016.05.012

31. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA Methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinf* (2012) 13:86. doi: 10.1186/1471-2105-13-86

32. Ren X, Kuan PF. methylGSA: a bioconductor package and shiny app for DNA methylation data length bias adjustment in gene set testing. *Bioinformatics* (2019) 35:1958–9. doi: 10.1093/bioinformatics/bty892

33. Megarbane A, Piquemal D, Rebillat AS, Stora S, Pierrat F, Bruno R, et al. Transcriptomic study in women with trisomy 21 identifies a possible role of the GTPases of the immunity-associated proteins (GIMAP) in the protection of breast cancer. *Sci Rep* (2020) 10:9447. doi: 10.1038/s41598-020-66469-w

34. Martin CE, List K. Cell surface-anchored serine proteases in cancer progression and metastasis. *Cancer Metastasis Rev* (2019) 38:357–87. doi: 10.1007/s10555-019-09811-7

35. Der CJ, Krontiris TG, Cooper GM. Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and kirsten sarcoma viruses. *Proc Natl Acad Sci USA* (1982) 79:3637–40. doi: 10.1073/pnas.79.11.3637

36. Parada LF, Tabin CJ, Shih C, Weinberg RA. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. *Nature* (1982) 297:474–8. doi: 10.1038/297474a0

37. Stites EC, Ravichandran KS. A systems perspective of ras signaling in cancer. *Clin Cancer Res* (2009) 15:1510–3. doi: 10.1158/1078-0432.CCR-08-2753

38. Van Golen KL. Inflammatory breast cancer: relationship between growth factor signaling and motility in aggressive cancers. *Breast Cancer Res* (2003) 5:174–9. doi: 10.1186/bcr598

39. Kyuno D, Takasawa A, Kikuchi S, Takemasa I, Osanai M, Kojima T. Role of tight junctions in the epithelial-to-mesenchymal transition of cancer cells. *Biochim Biophys Acta Biomembr* (2021) 1863:183503. doi: 10.1016/j.bbamem.2020.183503

40. Folkman J. Role of angiogenesis in tumor growth and metastasis. *Semin Oncol* (2002) 29:15–8. doi: 10.1053/sonc.2002.37263

41. Baek KH, Zaslavsky A, Lynch RC, Britt C, Okada Y, Siarey RJ, et al. Down's syndrome suppression of tumour growth and the role of the calcineurin inhibitor DSCR1. *Nature* (2009) 459:1126–30. doi: 10.1038/nature08062

42. Adini I, Rabinovitz I, Sun JF, Prendergast GC, Benjamin LE. RhoB controls akt trafficking and stage-specific survival of endothelial cells during vascular development. *Genes Dev* (2003) 17:2721–32. doi: 10.1101/gad.1134603

43. Kim C, Yang H, Fukushima Y, Saw PE, Lee J, Park JS, et al. Vascular RhoJ is an effective and selective target for tumor angiogenesis and vascular disruption. *Cancer Cell* (2014) 25:102–17. doi: 10.1016/j.ccr.2013.12.010

44. Fantin A, Lampropoulou A, Gestri G, Raimondi C, Senatore V, Zachary I, et al. NRP1 regulates CDC42 activation to promote filopodia formation in endothelial tip cells. *Cell Rep* (2015) 11:1577–90. doi: 10.1016/j.celrep.2015.05.018

45. Hoeppner LH, Sinha S, Wang Y, Bhattacharya R, Dutta S, Gong X, et al. RhoC maintains vascular homeostasis by regulating VEGF-induced signaling in endothelial cells. *J Cell Sci* (2015) 128:3556–68. doi: 10.1242/jcs.167601

46. Millet C, Zhang YE. Roles of Smad3 in TGF-beta signaling during carcinogenesis. *Crit Rev Eukaryot Gene Expr* (2007) 17:281–93. doi: 10.1615/CritRevEukarGeneExpr.v17.i4.30

47. Hao Y, Baker D, Ten Dijke P. TGF-beta-Mediated epithelial-mesenchymal transition and cancer metastasis. *Int J Mol Sci* (2019) 20. doi: 10.3390/ijms20112767

48. Nawshad A, Lagamba D, Polad A, Hay ED. Transforming growth factor-beta signaling during epithelial-mesenchymal transformation: implications for embryogenesis and tumor metastasis. *Cells Tissues Organs* (2005) 179:11–23. doi: 10.1159/000084505

49. Moustakas A, Heldin CH. Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. *Cancer Sci* (2007) 98:1512–20. doi: 10.1111/j.1349-7006.2007.00550.x

50. Matsuda T, Yamamoto T, Muraguchi A, Saatcioglu F. Cross-talk between transforming growth factor-beta and estrogen receptor signaling through Smad3. *J Biol Chem* (2001) 276:42908–14. doi: 10.1074/jbc.M105316200

51. Lee J, Yesilkanal AE, Wynne JP, Frankenberger C, Liu J, Yan J, et al. Effective breast cancer combination therapy targeting BACH1 and mitochondrial metabolism. *Nature* (2019) 568:254–8. doi: 10.1038/s41586-019-1005-x

52. Elbasateeny SS, Yassin MA, Mokhtar MM, Ismail AM, Ebian HF, Hussein S, et al. Prognostic implications of MALAT1 and BACH1 expression and their correlation with CTCs and Mo-MDSCs in triple negative breast cancer and surgical management options. *Int J Breast Cancer* (2022) 2022:8096764. doi: 10.1155/2022/8096764

53. Heyninck K, De Valck D, Vanden Berghe W, Van Criekinge W, Contreras R, Fiers W, et al. The zinc finger protein A20 inhibits TNF-induced NF-kappaB-dependent gene expression by interfering with an RIP- or TRAF2-mediated transactivation signal and directly binds to a novel NF-kappaB-inhibiting protein ABIN. *J Cell Biol* (1999) 145:1471–82. doi: 10.1083/jcb.145.7.1471

54. Gurevich I, Aneskievich BJ. Liganded RARalpha and RARgamma interact with but are repressed by TNIP1. *Biochem Biophys Res Commun* (2009) 389:409–14. doi: 10.1016/j.bbrc.2009.08.159

55. Amini P, Nassiri S, Malbon A, Markkanen E. Differential stromal reprogramming in benign and malignant naturally occurring canine mammary tumours identifies disease-modulating stromal components. *Sci Rep* (2020) 10:5506. doi: 10.1038/s41598-020-62354-8

56. Tumbarello DA, Waxse BJ, Arden SD, Bright NA, Kendrick-Jones J, Buss F. Autophagy receptors link myosin VI to autophagosomes to mediate Tom1-dependent autophagosome maturation and fusion with the lysosome. *Nat Cell Biol* (2012) 14:1024–35. doi: 10.1038/ncb2589

57. Katoh Y, Shiba Y, Mitsuhashi H, Yanagida Y, Takatsu H, Nakayama K. Tollip and Tom1 form a complex and recruit ubiquitin-conjugated proteins onto early endosomes. *J Biol Chem* (2004) 279:24435–43. doi: 10.1074/jbc.M400059200

58. Xiao S, Brannon MK, Zhao X, Fread KI, Ellena JF, Bushweller JH, et al. Tom1 modulates binding of tollip to phosphatidylinositol 3-phosphate via a coupled folding and binding mechanism. *Structure* (2015) 23:1910–20. doi: 10.1016/j.str.2015.07.017

59. Keskitalo S, Haapaniemi EM, Glumoff V, Liu X, Lehtinen V, Fogarty C, et al. Dominant TOM1 mutation associated with combined immunodeficiency and autoimmune disease. *NPJ Genom Med* (2019) 4:14. doi: 10.1038/s41525-019-0088-5

60. Kresovich JK, Xu Z, O'brien KM, Weinberg CR, Sandler DP, Taylor JA. Methylation-based biological age and breast cancer risk. *J Natl Cancer Inst* (2019) 111:1051–8. doi: 10.1093/jnci/djz020

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 20 July 2023
DOI 10.3389/fonc.2023.1189635

OPEN ACCESS

EDITED BY
Noha Mousaad Elemam,
University of Sharjah, United Arab Emirates

REVIEWED BY
Rana A. Youness,
University of Hertfordshire,
United Kingdom
Adriane Feijo Evangelista,
Barretos Cancer Hospital, Brazil

*CORRESPONDENCE
Rendong Zhang
zhangrendong2021@163.com
Jundong Wu
wujun-dong@163.com

RECEIVED 19 March 2023
ACCEPTED 16 June 2023
PUBLISHED 20 July 2023

CITATION
Fang Y, Zhang Q, Chen C, Chen Z, Zheng R, She C, Zhang R and Wu J (2023) Identification and comprehensive analysis of epithelial–mesenchymal transition related target genes of miR-222-3p in breast cancer. *Front. Oncol.* 13:1189635. doi: 10.3389/fonc.2023.1189635

Identification and comprehensive analysis of epithelial–mesenchymal transition related target genes of miR-222-3p in breast cancer

Yutong Fang[1,2], Qunchen Zhang[1,2], Chunfa Chen[1,2], Zexiao Chen[1,2], Rongji Zheng[1,2], Chuanghong She[1,2], Rendong Zhang[1,2]* and Jundong Wu[1,2]*

[1]The Breast Center, Cancer Hospital of Shantou University Medical College, Shantou, Guangdong, China, [2]The Department of Central Laboratory, Cancer Hospital of Shantou University Medical College, Shantou, Guangdong, China

Background: Epithelial–mesenchymal transition (EMT) is a crucial mechanism that microRNA-222-3p (miR-222-3p) promotes breast cancer (BC) progression. Our study aimed to identify EMT-associated target genes (ETGs) of miR-222-3p for further analysis of their roles in BC based on bioinformatics tools.

Methods: Based on bioinformatics analysis, we identified 10 core ETGs of miR-222-3p. Then, we performed a comprehensive analysis of 10 ETGs and miR-222-3p, including pathway enrichment analysis of ETGs, differential expression, clinical significance, correlation with immune cell infiltration, immune checkpoint genes (ICGs) expression, tumor mutational burden (TMB), microsatellite instability (MSI), stemness, drug sensitivity, and genetic alteration.

Results: The expression of miR222-3p in basal-like BC was significantly higher than in other subtypes of BC and the normal adjacent tissue. Pathway analysis suggested that the ETGs might regulate the EMT process via the PI3K-Akt and HIF-1 signaling pathway. Six of the 10 core ETGs of miR-222-3p identified were down-expressed in BC, which were *EGFR, IL6, NRP1, NTRK2, LAMC2*, and *PIK3R1*, and *SERPINE1, MUC1, MMP11*, and *BIRC5* were up-expressed in BC, which also showed potential diagnostic values in BC. Prognosis analysis revealed that higher *NTRK2* and *PIK3R1* expressions were related to a better prognosis, and higher BIRC5 and miR-222-3p expressions were related to a worse prognosis. Most ETGs and miR-222-3p were positively correlated with various infiltration of various immune cells and ICGs expression. Lower TMB scores were correlated with higher expression of *MUC1* and *NTRK2*, and higher *BIRC5* was related to a higher TMB score. Lower expression of *MUC1*, *NTRK2*, and *PIK3R1* were associated with higher MSI scores. Higher expression of ETGs was associated with lower mRNAsi scores, except *BIRC5* and miR-222-3p conversely. Most ETGs and miR-222-3p expression were negatively correlated with the drug IC50 values. The analysis of the genetic alteration of the ETGs suggested that amplification was the main genetic alteration of eight ETGs except for *NTRK2* and *PIK3R1*.

Conclusion: MiR-222-3p might be a specific biomarker of basal-like BC. We successfully identify 10 core ETGs of miR-222-3p, some might be useful diagnostic and prognostic biomarkers. The comprehensive analysis of 10 ETGs and miR-222-3p indicated that they might be involved in the development of BC, which might be novel therapeutic targets for the treatment of BC.

KEYWORDS

breast cancer, miR-222-3p, epithelial-mesenchymal transition, target gene, diagnosis, prognosis, immune infiltration, drug sensitivity

1 Introduction

Breast cancer (BC) is a prevalent malignancy among women globally, with an annual incidence of over two million cases, posing a significant threat to women's health and life (1). The current therapeutic approach for BC patients involves a comprehensive therapeutic strategy comprising surgery, chemotherapy, radiotherapy, endocrine therapy, and targeted therapy (2). Despite the 5-year survival rate exceeding 90% for localized BC patients, the survival rate drops to <30% for those diagnosed with metastatic BC (3), which accounts for over 90% of cancer-related deaths in BC (4). Epithelial–mesenchymal transition (EMT) is a well-established process that plays a critical role in tumor-distant metastasis, whereby epithelial cells undergo a phenotypic switch to motile mesenchymal cells with heightened migratory and invasive capabilities (5). EMT has been linked to various biological properties of BC, including the acquisition of stem cell characteristics by BC cells (6). In addition, EMT has been shown to contribute to immunosuppression within the tumor microenvironment, thereby promoting tumor progression and resistance to immunotherapy (7). In addition, several lines of evidence have proven that the expression of the EMT-related gene was associated with therapeutic resistance; thus, it is necessary to evaluate the impact of EMT-related gene expression when developing a precise and individualized treatment plan for BC patients (8).

MicroRNAs (miRNAs) are a class of non-coding RNA molecules that are short and single-stranded. They function as post-transcriptional regulators of protein-coding gene expression and are involved in various cellular activities and the pathogenesis of numerous human diseases, including cancer (9, 10). Recent research has demonstrated that miRNAs can regulate the EMT process by targeting transcription factors such as Snail land Twist, thereby influencing tumor invasion and metastasis in different types of cancer (11). Certain EMT-related miRNAs have also been linked to cancer stemness and drug resistance (12).

MicroRNA-222-3p (miR-222-3p), a member of the miRNA family, is located on the human X chromosome and functions as a regulator of gene expression in the context of tumorigenesis. Specifically, miR-222-3p has been implicated in the progression of various types of cancers, acting as either a tumor suppressor or oncogene (13). In recent years, it has been established that miR-222-3p plays a significant role in promoting BC progression through the mechanism of EMT. A prior investigation has demonstrated that miR-222-3p can suppress the expression of Zinc finger E-box-binding homeobox 2 *(ZEB2)*, thereby inducing EMT (14). Moreover, recent evidence has identified *Notch3*, a member of the Notch receptor family, as a target of miR-222-3p that facilitates the EMT process in BC cells (15). Given that miR-222-3p is an EMT-associated miRNA, it is imperative to further explore the molecular mechanisms underlying its regulation of the EMT process in BC. Given the ability of a single miRNA to regulate the expression of multiple target genes concurrently (16), the objective of our investigation was to identify further potential target EMT-related genes of miR-222-3p and to explore their clinical significance, their relationship with tumor-infiltrating immune cells and drug sensitivity based on bioinformatics tools. Our findings may aid in the identification of novel drug therapy targets for patients with breast cancer.

2 Materials and methods

2.1 Data collection and analysis of differential expression and clinical significance of miR-222-3p

The level 3 normalized miRNA-sequencing data of miR-222-3p expression, including 104 normal samples and 1,103 BC tissue samples, and the clinical information of BC patients were downloaded from The Cancer Genome Atlas (TCGA) website (https://portal.gdc.cancer.gov/) (17). Then, the expression data of miRNA were then normalized to read per million (RPM) format and then were converted to log2(RPM+1). Samples lacking clinical information were excluded when analyzing the relation between expression and clinical significance in this study. The unpaired t-test was used to analyze the statistical difference between two groups of BC patients, and the Kruskal–Wallis test among more than two groups. Values of the expression level were displayed as means ± standard deviations. The differential expression of miR-222-3p was also validated by the GSE45666 dataset obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) and cell lines. The receiver operating

characteristic (ROC) analysis was performed by the pROC package in R to evaluate the power of miR-222-3p to differentiate BC subtypes. The Kaplan–Meier (KM) method with the log-rank test was used for analyzing the prognosis between high and low miR-222-3p expression groups with a cutoff set at the median expression level by the survival and survminer package in R. Plots were generated in R with the ggplot2 package.

2.2 Identification and enrichment analysis of EMT-related target genes

To identify the potential target genes of miR-222-3p, we used the miRWalk website of version 3.0 (http://mirwalk.umm.uni-heidelberg.de/) (Supplementary Table S1), which includes the prediction outcomes of various prediction databases (18). The EMT-related genes were obtained from the dbEMT 2.0 database (http://dbemt.bioinfo-minzhao.org/index.html) (Supplementary Table S2), including 1,184 genes (19). In addition, we identified the differentially expressed genes (DEGs) of the data from TCGA via the R software package limma package with the thresholds of $|\log FC| > 1$ and false discovery rate (FDR) < 0.05 (Supplementary Table S3). Then, the intersection among the potential target genes of miR-222-3p, EMT-related genes, and DEGs were selected as the possible EMT-related target genes (ETGs) of miR-222-3p. To further explore the related pathways involved in the process of BC and biological functions of the possible ETGs of miRR-222-3p, we performed the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses using the R clusterProfiler package and visualization via the R package ggplot2. To further select the hub ETGs to improve the precision of the study, we constructed the protein–protein interaction (PPI) network of the ETGs via the STRING database (https://cn.string-db.org/) (20) and visualized via the Cytoscape (version 3.8.2), with the cytoHubba tool of which we screened 10 top hub genes as the ETGs of miR-222-3p for further research.

2.3 Data collection and correlation analysis of ETGs

After identifying the ETGs of miR-222-3p, we further downloaded the level 3 RNA-sequencing data in fragments per kilobase million (FPKM) format of 10 ETGs from the TCGA website. The data were converted to the format of transcripts per million (TPM) as log2(TPM+1). The Spearman's correlation test was used to analyze the association between miR-222-3p and its ETGs and the pairwise correlation among the ETGs.

2.4 Differential expression and protein expression of the ETGs

The unpaired t-test was used to analyze the statistical difference of the differential expression between normal groups and BC groups of 10 ETGs with the data obtained from the TCGA and validated by the GSE45666 dataset obtained from the GEO database and cell lines. In addition, the immunohistochemistry images of BC tissues and paired adjacent normal tissues of 10 EGTs were downloaded from the Human Protein Atlas (https://www.proteinatlas.org/) to analyze the protein expression of the EGTs.

2.5 Cell lines and quantitative real-time PCR

MCF-7 is an ER-positive human BC cell line, and MDA-MB-231 is a human basal-like BC cell line with high invasiveness. MCF-10A is a normal breast epithelial cell line. The BC cell lines MCF-7 and MDA-MB-231, and the normal breast epithelial cell line MCF-10A were purchased from Procell (Wuhan, China) and cultured according to the manufacturer's recommendations. Total RNA was severely isolated from the cells using the RNAsimple total RNA kit (Tiangen, Beijing, China) according to the manufacturer's instructions. The quantitative real-time PCR (qRT-PCR) was performed using the PrimeScriptTM RT reagent kit (Takara, Japan) and the SYBR Premix Ex TaqTM II (Takara, Japan) according to the manufacturer's instructions. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal reference gene, and the relative expression levels were calculated by the $2^{-\triangle\triangle}$Ct method. All specific primers are shown in Table 1. The Student's t-test was used for pairwise comparison of the statistical difference between the MCF-10A cell line and BC cell lines. Plots were generated in GraphPad Prism (version 8.0).

2.6 Clinical significance and prognosis analysis of the EGTs

The ROC curves analysis was performed by the pROC package in R for potential diagnostic values evaluation of the up-expressed EMTs and validated by the GSE45666 dataset from GEO. To explore the clinical significance of the ETGs, we utilized the unpaired t-test to identify the association between the ETGs expression and clinical stages and PAM50 subtypes of BC. In addition, the KM method with the log-rank test was performed by the survival and survminer package in R for analyzing the prognosis of the ETGs expression, including the overall survival (OS) and the disease-specific survival (DSS), with a cutoff set at the median expression level between high and low expression groups. Plots were generated in R with the ggplot2 package.

2.7 Immune cell and immune checkpoint genes analysis

To explore the relationship between the ETGs expression and the immune cells in BC, we utilized the single-sample GSEA (ssGSEA) method (21) to present the infiltration enrichment of 24 common immune cells from the TCGA cohort, and the Spearman's test was used for correlation analysis. Additionally, we applied Spearman's correlation test to analyze the correlation

TABLE 1 Sequences of all primers.

Primers sequence (5′–3′)	
GAPDH F	GTCAAGGCTGAGAACGGGAA
GAPDH R	TGGACTCCACGACGTACTCA
EGFR F	TCAGCTAGTTAGGAGCCCATTTTT
EGFR R	TGTGACTGAACATAACTGTAGGCT
IL6 F	ACCTAGAGTACCTCCAGAACAGAT
IL6 R	CAGGGGTGGTTATTGCATCTAGAT
SERPINE1 F	AGATTCAAGCAGCTATGGGATTCA
SERPINE1 R	TGCTGATCTCATCCTTGTTCCATG
MUC1 F	GTGAGTGATGTGCCATTTCCTTTC
MUC1 R	CCAAGGCAATGAGATAGACAATGG
NRP1 F	TTGTCTGCCCTGGAGAACTATAAC
NRP1 R	TCATGCCTCCGAATAAGTACTCTG
MMP11 F	TCGACTATGATGAGACCTGGACTA
MMP11 R	GAAAGGTGTAGAAGGCGGACATC
NTRK2 F	GAGATTGGAGCCTAACAGTGTAGA
NTRK2 R	TTCTCAGTCCCACATAAGCTTCAA
LAMC2 F	TCACCAAGACTTACACATTCAGGT
LAMC2 R	GAGATTCCGCAGTAACCTTCGATA
PIK3R1 F	TAAACCAGACCTTATCCAGCTGAG
PIK3R1 R	TCTTCATCATCTTCCACCAGTGAA
BIRC5 F	TTGCGCTTTCCTTTCTGTCAAG
BIRC5 R	CCGCAGTTTCCTCAAATTCTTTCT
MiR-222-3p F	GTTCGTGGGAGCTACATCTGGC
MiR-222-3p R	GTGTCGTGGAGTCGGCAATTC
MiR-222-3p RT Primer	GTCGTATCCAGTGCAGGGTCCGAGG TATTCGCACTGGATACGACACCCAGTA

between eight ICGs and the ETGs and miR-222-3p expression. Plots were generated with the ggplot2 package.

2.8 Tumor mutational burden, microsatellite instability, and stemness analysis

The somatic mutation data for TMB analysis were downloaded from TCGA, and the TMB scores were calculated by the observed number of mutations divided by 38Mb (22). MSI scores of the BC samples from TCGA were obtained from a previous publication (23). The one-class logistic regression (OCLR) machine-learning algorithm (24) was used for calculating the mRNA expression-based stemness index (mRNAsi) score. The unpaired t-test was used to analyze the statistical difference of the TMB, MSI, and mRNAsi scores between the high-expression and the low-expression group of 10 ETGs and miR-222-3p. Plots were generated with the ggplot2 package.

2.9 Drug sensitivity analysis

The R pRRophetic package was used to predict the drug response of each sample from the TCGA, and the drug sensitivity (IC50) values of each sample were estimated using Ridge's regression with the data obtained from the Genomics of Drug Sensitivity in Cancer (GDSC) (25). Then, Spearman's correlation test was applied to analyze the correlation between the IC50 values and the expression levels. Plots were generated in R with the ggplot2 package.

2.10 Genetic alteration analysis

Genetic alteration of the ETGs in the BC cohort was analyzed by cBioPortal website (http://www.cbioportal.org). Additionally, we also analyzed the OS and DSS in altered and unaltered groups.

2.11 Statistical analysis

The R software (version 4.2.1) and GraphPad Prism (version 8.0) were used for all statistical analyses. The above section has described detailed statistical approaches for data processing. $p<0.05$ was considered statistically significant.

3 Result

3.1 Differential expression and clinical significance of miR-222-3p

The analysis of the differential expression of miR-222-3p from the TCGA suggested that the expression of the BC tissues was lower than that of the normal tissues ($p=0.002$), which were 5.168 ± 1.117 and 5.519 ± 0.63, respectively (Figure 1A). The same outcome was also validated in the GSE45666 dataset ($p<0.05$) (Figure 1B). Additionally, the expression of miR-222-3p in MCF-7 was downregulated but upregulated significantly in MDA-MB-231 (Figure 1C). The relation between the miR-222-3p expression and the clinical indicators is shown in Figure 1D. The expression of miR-222-3p was associated with the status of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), lymph node status, and the PAM50 subtypes (all $p<0.05$). From the results, we found that BC patients with negative expression of ER and PR had a higher expression level of miR-222-3p (both $p<0.001$). The negative status of HER2 was associated with the high expression of miR-222-3p ($p=0.031$). Additionally, patients with nodal status of N0 and N2 had higher expression of miR-222-3p than patients with lymph node metastasis of N3 ($p=0.021$, $p=0.040$, respectively). Notably, the expression of miR-222-3p was 6.176 ± 1.047, significantly higher than that of luminal A (4.848 ± 1.006, $p<0.001$), luminal B (5.099 ±

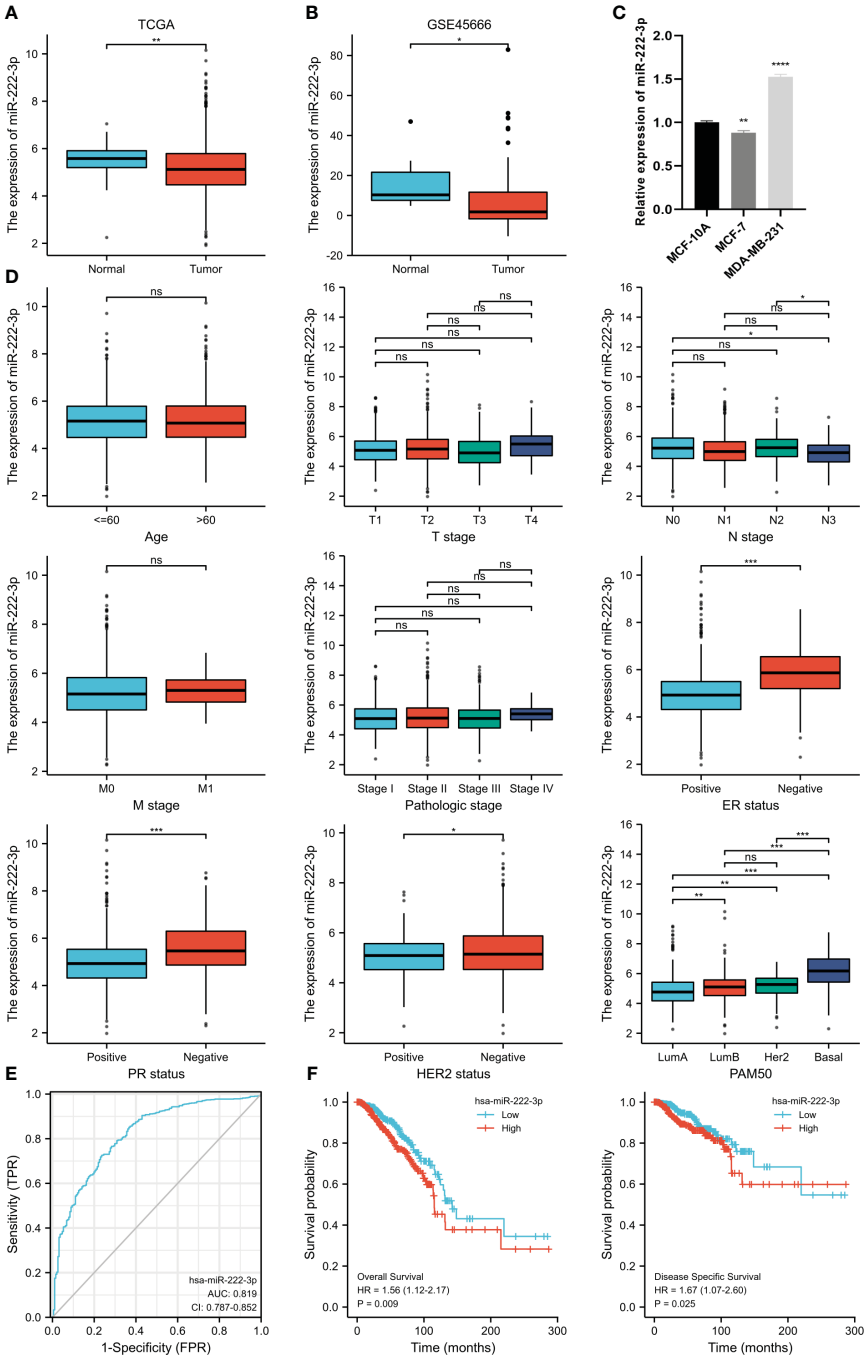


FIGURE 1
MiR-222-3p differential expression in BC and normal adjacent tissues based on TCGA database **(A)** and validated by the GSE45666 dataset **(B)** and cell lines **(C)**. The association between miR-222-3p expression and age, T stage, N stage, M stage, pathological stage, ER status, PR status, HER2 status, and PAM50 subtype **(D)**. The ROC curve shows the discriminative power between the basal-like subtype and other BC subtypes of miR-222-3p **(E)**. The KM survival curves show the OS and DSS of the high and low miR-222-3p expression groups in BC patients from TCGA **(F)**. NS, indicates no statistical difference, *p< 0.05, **p< 0.01, ***p< 0.001, ****P< 0.0001..

1.025, p<0.001) and HER2-enriched subtypes (5.116 ± 0.804, p<0.001). We further evaluate the discriminative power between the basal-like subtype and other BC subtypes of miR-222-3p, and the result showed that the area under the curve (AUC) of the ROC curve was 0.819 when the cutoff value was 5.423, with a specificity of 72.5% and a sensitivity of 76.4% (Figure 1E). Additionally, a higher expression of miR-222-3p was correlated with better OS (HR=1.56, p=0.009) and DSS (HR=1.67, p=0.025) (Figure 1F).

3.2 Identification and enrichment analysis of EMT-related target genes

As shown in the Wayne diagram in Figure 2A, a total of 2692 genes were predicted as the target of miR-222-3p via the miRWalk, and 2,401 differentially expressed genes and 1,184 EMT-related genes were identified. We selected the intersection and finally identified 38 genes as the possible *ETGs* of miR-222-3p. The GO and KEGG functional enrichment analyses were performed to further explore the potential biological mechanisms of the above 38 genes (Figure 2B, Supplementary Table S4). The GO analysis showed that the ETGs might regulate the biological process of cell-matrix adhesion and intracellular signal transduction. The KEGG pathway analysis suggested that the ETGs might regulate the EMT process via the PI3K-Akt and HIF-1 signaling pathway and were associated with the drug resistance of *EGFR* tyrosine kinase inhibitors in cancer treatment. We constructed a PPI network of 38 ETGs, with 27 nodes and 42 edges (Figure 2C). To improve the accuracy of prediction, we identified 10 top hub genes of the PPI network as the ETGs of miR-222-3p for further research, which were *EGFR, IL6, SERPINE1, MUC1, NRP1, MMP11, NTRK2, LAMC2, PIK3R1, BIRC5* (Figure 2D).

3.3 Correlation analysis

We analyzed the correlation between the miR-222-3p and its ETGs, and the results of Spearman's correlation test showed that 6 of 10 ETGs were significantly correlated with the expression of miR-222-3p. See Figure 3A for further details. The expression of *EGFR* (r=0.201, p<0.001), *IL6* (r=0.127, p<0.001), *LAMC2* (r=0.179, p<0.001), *BIRC5* (r=0.282, p<0.001) were positively correlated with miR-222-3p expression, and *MUC1* (r=−0.260, p<0.001) and *PIK3R1* (r=−0.096, p=0.001) negatively. Interestingly, the pairwise correlation among the ETGs showed that most ETGs had positive a correlation with others. Interestingly, the negative correlation mainly existed between *BIRC5* and other ETGs, such as *SERPINE1* (r=−0.136, p<0.001), *MUC1* (r=−0.417, p=0.003), *NRP1* (r=−0.263, p<0.001), *NTRK2* (r=−0.356, p<0.001), and *PIK3R1* (r=−0.314, p<0.001) (Figure 3B).

3.4 Differential expression of the ETGs

Based on the TCGA data, we analyzed the differential expression between the BC samples and the normal samples. As shown in

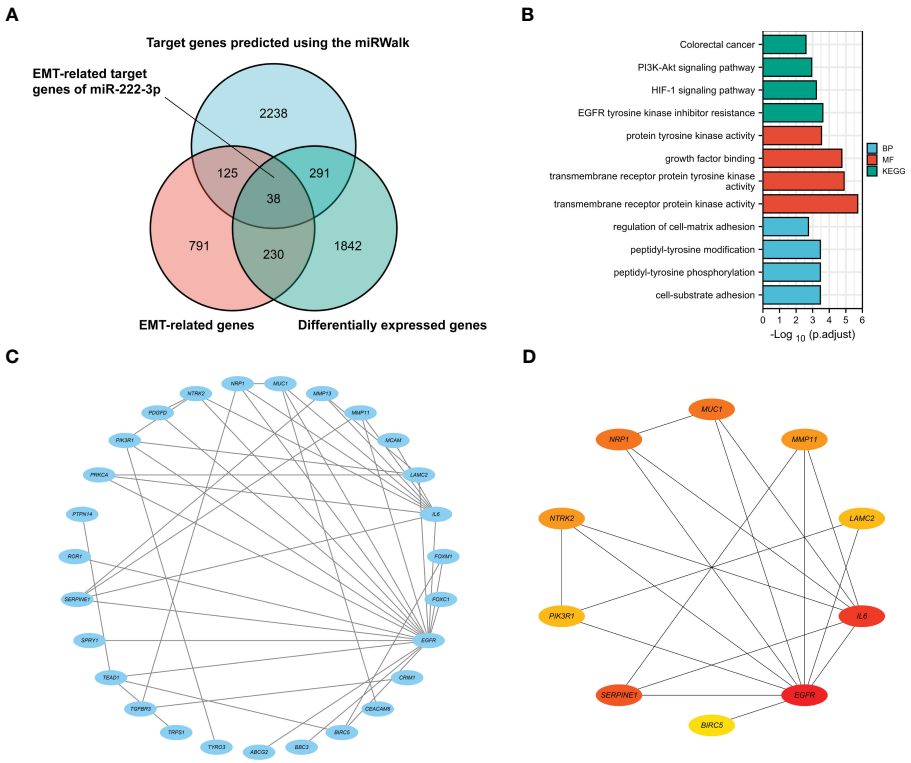


FIGURE 2
The Venn diagram shows 38 target genes as the possible ETGs of miR-222-3p **(A)**. GO and KEGG pathway enrichment analysis of 38 ETGs of miR-222-3p **(B)**. The PPI network of 38 ETGs of miR-222-3p **(C)**. Ten top hub genes of the PPI network were identified as the ETGs of miR-222-3p for further research **(D)**.

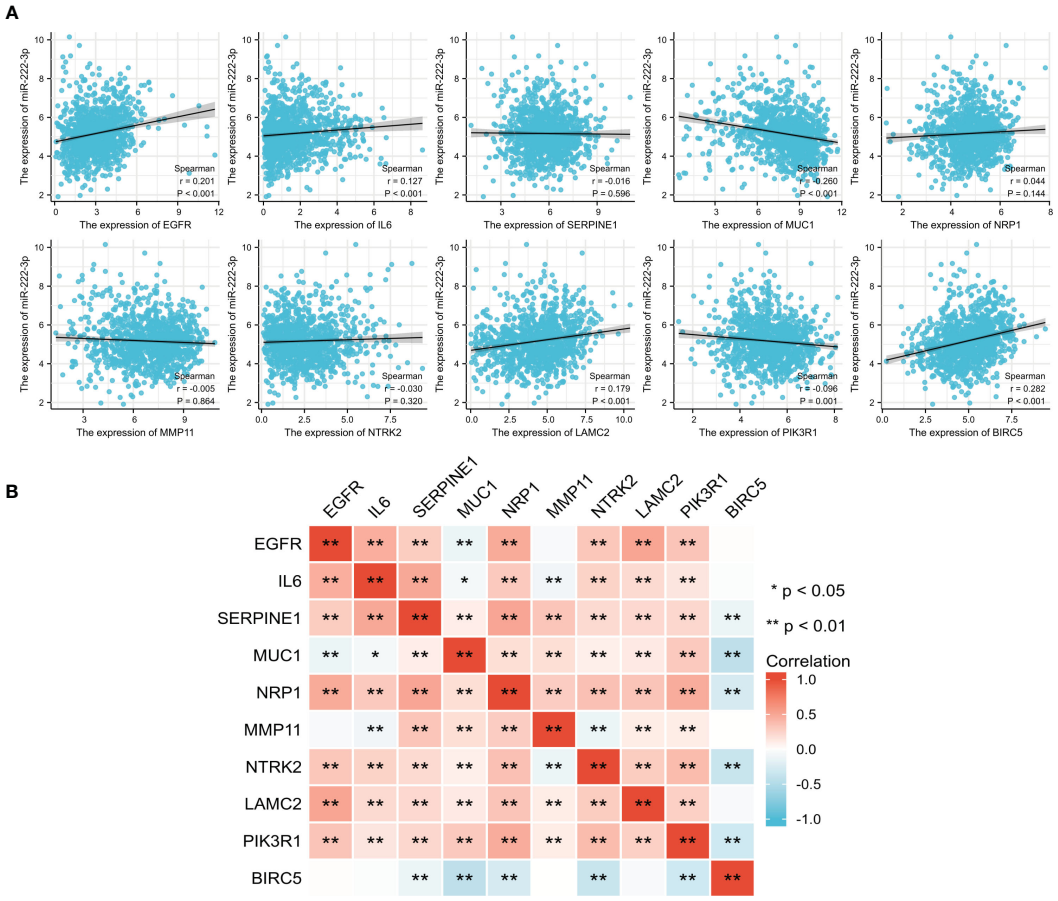


FIGURE 3
The correlation between the expression of miR-222-3p and its ETGs **(A)**. The pairwise correlation among the ETGs expression **(B)**. *p< 0.05, **p< 0.01.

Figure 4A, 7 of 10 ETGs exhibited lower expression levels in the tumor group than the normal group, which were *EGFR*, *IL6*, *NRP1*, *NTRK2*, *LAMC2*, and *PIK3R1*, and the remaining genes, *SERPINE1*, *MUC1*, *MMP11*, and *BIRC5* had higher expression level in the tumor group than the normal group (all p<0.05). The differential expression of the ETGs were also validated by the GSE45666 dataset (Figure 4B). In addition, the results of 10 ETGs and miR-222-3p expression were validated in two BC cells (MCF-7 and MDA-MB-231) and a breast epithelial cell line MCF-10A. The results showed that *EGFR*, *IL6*, *NRP1*, *NTRK2*, *LAMC2*, and *PIK3R1* were downregulated in BC cell lines, and *MUC1*, *MMP11*, and *BIRC5* were upregulated in BC cell lines (Figure 4C). The immunohistochemistry images of 10 ETGs were obtained from the HPA database to validate their protein expression. As shown in Figure 5, most ETGs had consistent protein expression with previous analyses in BC and normal samples of TCGA data. However, the protein expression of *SERPINE1*, *NTRK2*, and *BIRC5* showed no significant difference between BC tissue and normal tissue.

3.5 Diagnostic value of the up-expressed ETGs

We further evaluated the potential diagnostic values of the up-expressed ETGs for distinguishing the BC group and the normal group. As shown in Figure 6A, the ROC curve of *SERPINE1* had an AUC of 0.683, with a sensitivity of 42.5% and a specificity of 86.1% when the cutoff value was 4.355. The ROC curve of *MUC1* had an AUC of 0.819, with a sensitivity of 91.2% and a specificity of 67.3% when the cutoff value was 7.346. The AUC of *MMP11* was 0.993, which was the highest among the up-expressed ETGs; the sensitivity and specificity were 97.3% and 95.2%, respectively, when the cutoff was 3.461. The AUC of *BIRC5* was 0.955, with a sensitivity and specificity of 91.2% and 88.1%, respectively, when the cutoff was 3.379. The diagnostic values of the up-expressed ETGs were also validated in the GSE45666 dataset (Figure 6B).

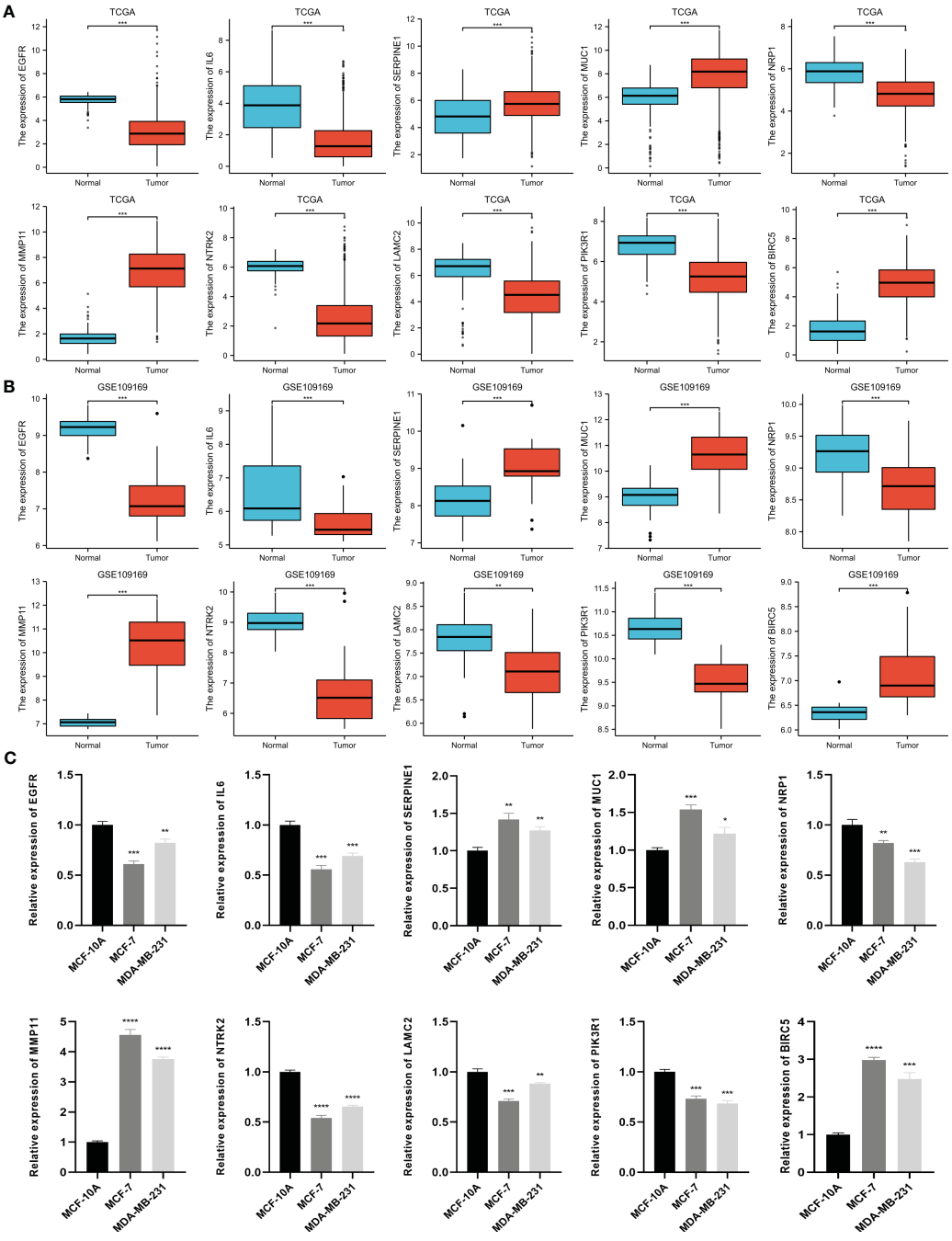


FIGURE 4
Differential expression of 10 ETGs in BC and normal adjacent tissues based on TCGA database **(A)**, which are validated by the GSE109169 dataset obtained from the GEO database **(B)** and cell lines **(C)**. *p< 0.05, **p< 0.01, ***p< 0.001, ****P< 0.0001.

3.6 Clinical significance of ETGs in BC

We analyzed the correlation of ETGs expression with clinical stages and PAM50 subtypes of BC. The results suggested that most ETGs showed no significant difference among clinical stages (Figure 7A). However, the analysis in Figure 7B suggested that the expression of most ETGs was associated with PAM50 subtypes, among which *EGFR*, *IL6*, and *LAMC2* tended to have a higher expression in the basal-like than others. Conversely, *MMP11* and *MUC1* tended to have a lower expression in the basal-like subtype. Additionally, *SERPINE1*, *MUC1*, *NRP11*, and *NTRK2* tended to have a higher expression in luminal A BC, and the expression of *BIRC5* was lower in the luminal A subtype. Interestingly, the expression of *PIK3R1* in the luminal A and HER2-enriched subtypes was higher than in the luminal B and basal-like subtypes.

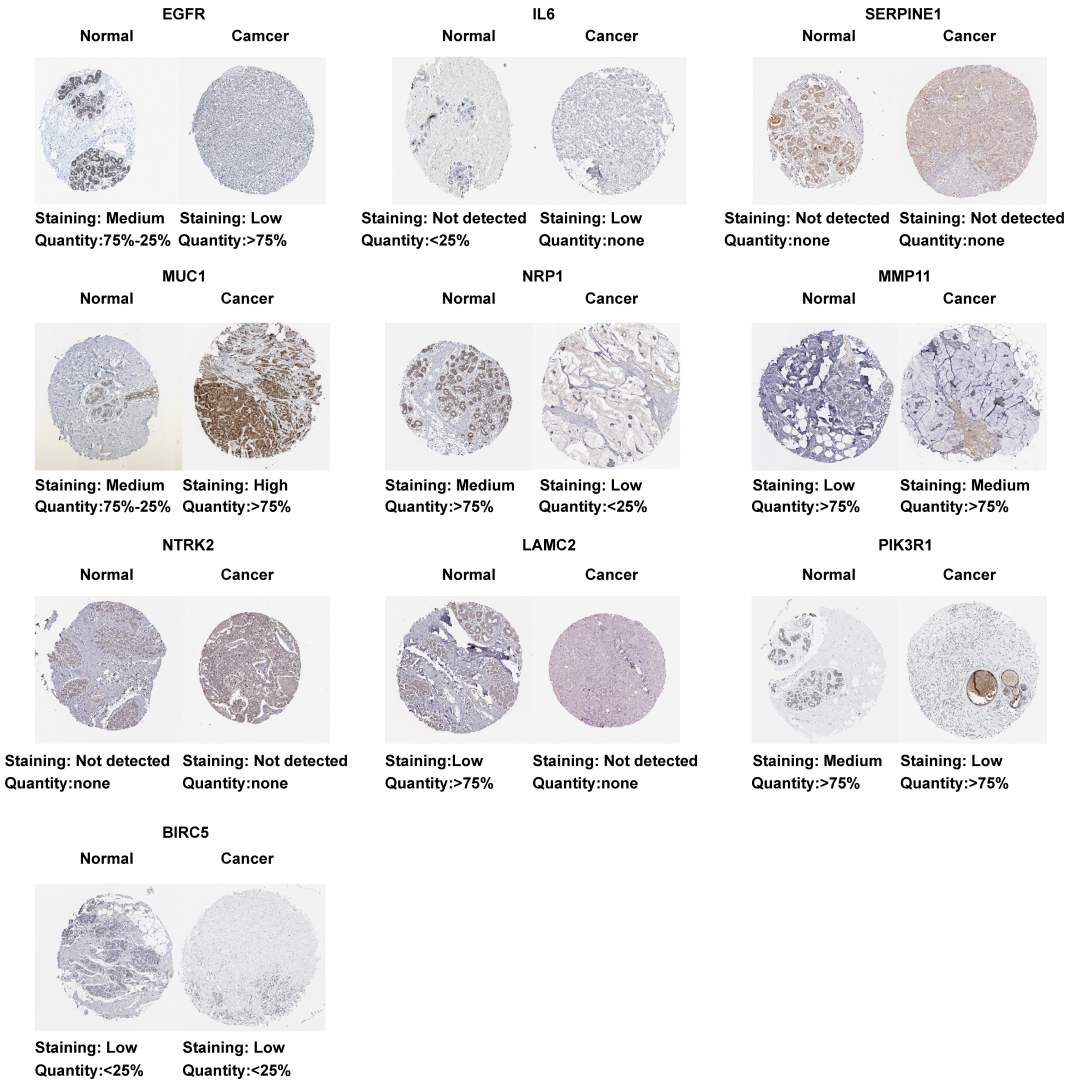


FIGURE 5
Protein expression of 10 ETGs. Images were obtained from the HPA database.

3.7 Prognosis analysis of the EGTs

The KM survival curves were drawn for prognosis analysis, and the results are shown in Figure 8. We found that BC patients with higher *NTRK2* expression had longer OS (HR=0.65, p=0.008) and DSS (HR=0.56, p=0.009). Interestingly, higher expression of *PIK3R1* was related to shorter DSS (HR=0.56, p=0.011), but OS was not significantly different. A higher expression of *BIRC5* was correlated with shorter DSS (HR=1.65, p=0.023) but not correlated with OS. The results of the KM survival curves of other ETGs were not statistically significant.

3.8 Immune infiltration analysis of ETGs and miR-222-3p

The immune infiltration levels of a total of 24 immune cells in BC were analyzed. The results shown in Figure 9A suggested that most ETGs, mainly *EGFR, IL6, SERPINE1, NRP1,* and *NTRK2*, were significantly positively correlated with various infiltration of various immune cells, among which *IL6* showed the highest positive correlation with activated DCs (aDCs) (r=0.322, p<0.001), B cells (r=0.463, p<0.001), $CD8^+$ T cells (r=0.419, p<0.001), cytotoxic cells (r=0.441, p<0.001), dendritic cells (DCs) (r=0.561, p<0.001), immature DCs (iDCs) (r=0.440, p<0.001), neutrophils (r=0.515, p<0.001), NK $CD56^-$ cells (r=0.328, p<0.001), plasmacytoid DCs (pDCs) (r=0.353, p<0.001), T cells (r=0.438, p<0.001), T effector memory (Tem) cells (r=0.313, p<0.001), T follicular helper (TFH) cells (r=0.299, p<0.001), and type 1 Th (Th1) cells (r=0.505, p<0.001). However, the negative correlation between ETGs expression and immune infiltration mainly existed in *MUC1* and *BIRC5*. Additionally, as shown in Figure 9B, the expression of miR-222-3p was positively correlated with most types of immune cells significantly, especially aDC (r=0.379, p<0.001), B cells (r=0.167, p<0.001), and macrophages (r=0.363, p<0.001). miR-222-3p expression was significantly negatively correlated with eosinophils (r=−0.267, p<0.001) and mast cells (r=−0.277, p<0.001).

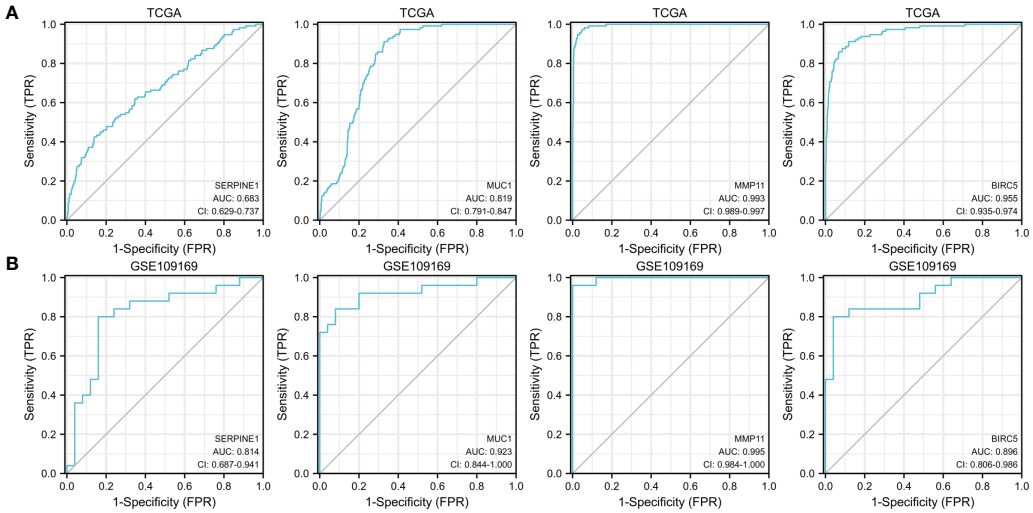


FIGURE 6
ROC curves show the diagnostic values of four up-expressed genes **(A)** and are validated by the GSE109169 dataset **(B)**.

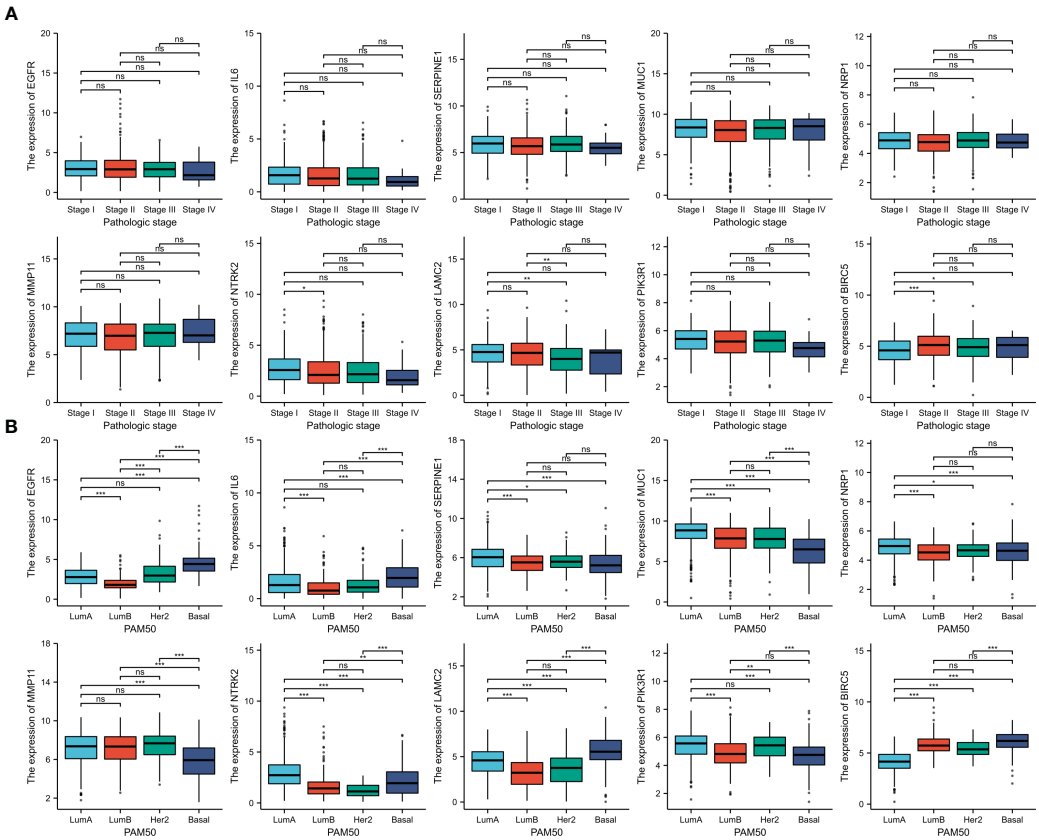


FIGURE 7
The association between 10 ETGs expression and pathologic stage **(A)** and PAM50 subtype **(B)**. NS, indicates no statistical difference, *p< 0.05, **p< 0.01, ***p< 0.001.

10.3389/fonc.2023.1189635

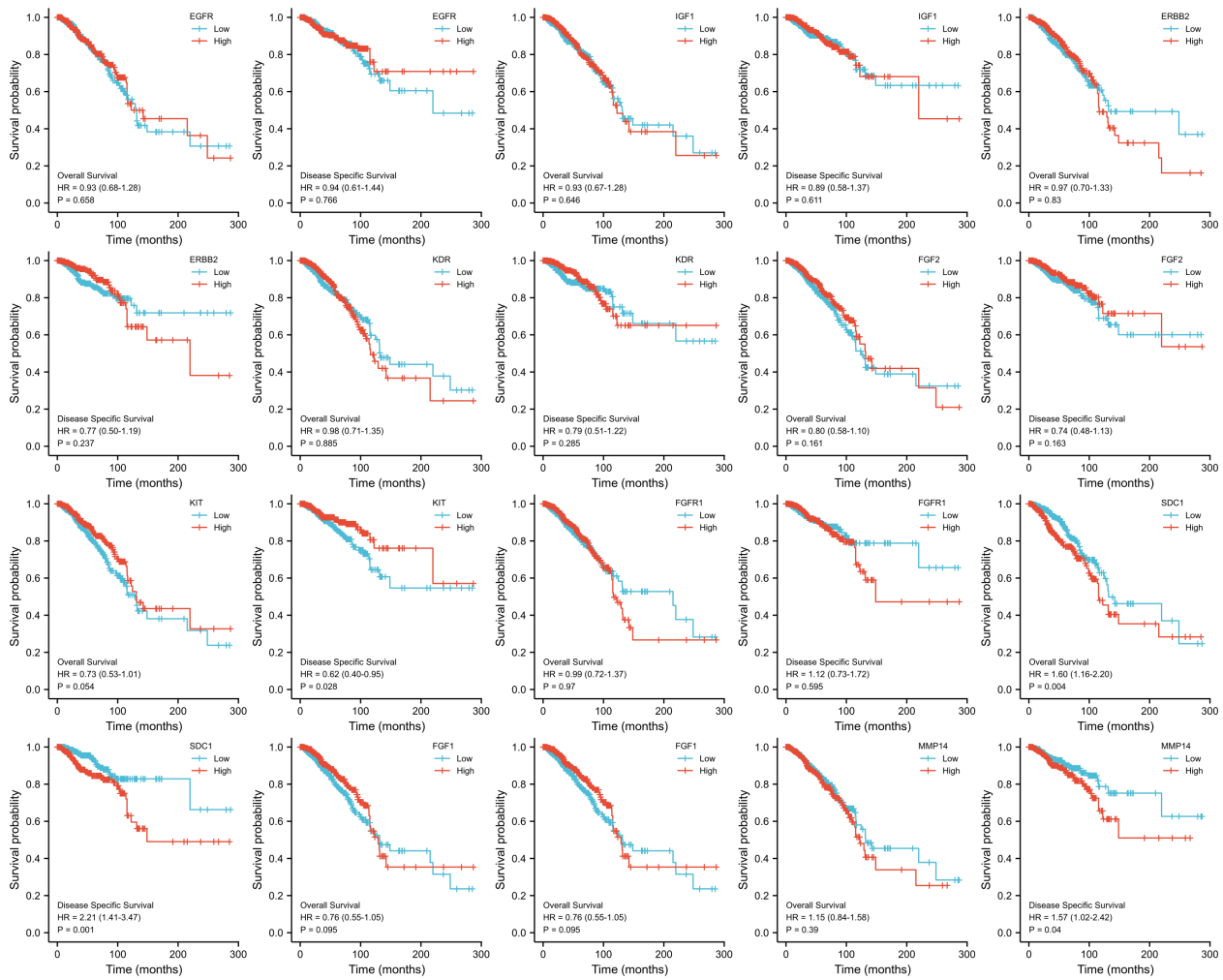


FIGURE 8
OS and DSS analysis of 10 ETGs based on TCGA.

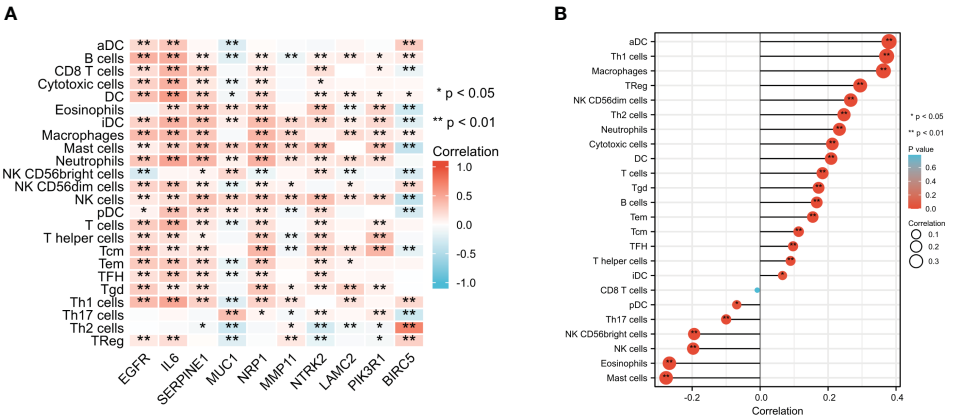


FIGURE 9
Comparison of infiltration levels in 24 common immune cells between low- and high-expression groups of 10 ETGs **(A)** and miR-222-3p **(B)**. *p< 0.05, **p< 0.01.

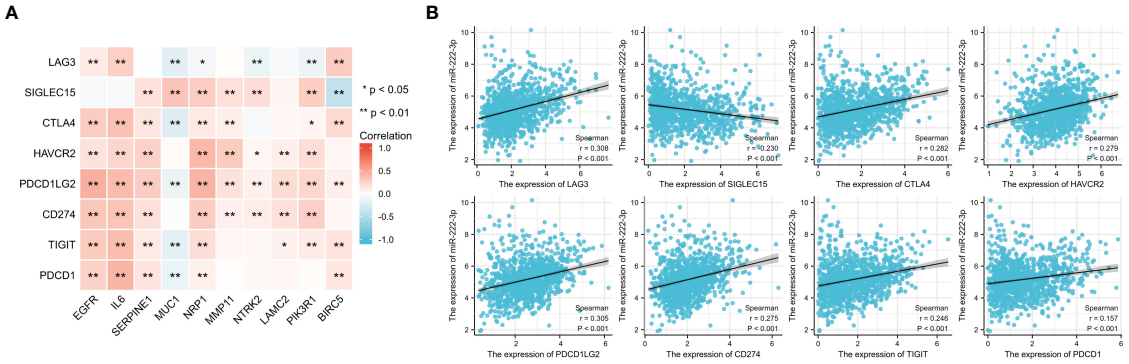


FIGURE 10
Correlation between 8 ICGs and 10 ETGs expression **(A)** and correlation between 8 ICGs and miR-222-3p expression **(B)**. *p< 0.05, **p< 0.01.

3.9 Correlation with the ICGs expression

From the results shown in Figure 10A, we found that most ETGs of miR-222-3p were positively correlated with the expression of the *ICGs*, especially *EGFR*, *IL6*, *SERPINE1*, and *NRP1*. Interestingly, the negative correlation mainly existed between *MUC1* and ICGs, which were *LAG3* (r=−0.202, p<0.001), *CTLA4* (r=−0.213, p<0.001), *PDCD1LG2* (r=−0.125, p<0.001), *TIGIT* (r=−0.179, p<0.001), and *PDCD1* (r=−0.157, p<0.001). Additionally, miR-222-3p showed a significant correlation with eight ICGs, which were only negatively correlated with *SIGLEC15* (r=−0.230, p<0.001) (Figure 10B).

3.10 TMB and MSI analysis

It has been suggested that patients with a high level of TMB tend to benefit from immunotherapy. Figure 11A showed the difference between high- and low-expression groups, and the results of the unpaired t-test showed that the lower TMB scores were correlated with higher expression of *MUC1* (p<0.001) and *NTRK2* (p<0.001), and high *BIRC5* expression was associated with a higher TMB score (p<0.001). We also evaluated the association of MSI scores between high- and low-expression groups of ETGs (Figure 11B), and we found that lower expression of *MUC1* (p=0.005), *NTRK2* (p=0.046), and *PIK3R1* (p=0.005) were associated with higher MSI scores, and higher *BIRC5* expression was related to higher MSI score (p=0.019). Additionally, we found that patients with higher miR-222-3p expression tend to have higher TMB scores (p=0.001) (Figure 11C).

3.11 Stemness analysis

We evaluated the difference in mRNAsi score between high- and low-expression groups. Figure 12 showed that higher expression of ETGs was associated with lower mRNAsi score, except *BIRC5* conversely (all p<0.001). Additionally, BC patients with higher expression of mir-222-3p tend to have a higher mRNAsi level (p<0.001).

3.12 Drug sensitivity analysis

The IC50 value was a value that was used to evaluate the sensitivity of drug treatment. Figure 13A showed the association between the IC50 of eight drugs and ETGs expression, from which we found that most ETGs were negatively correlated with the IC50 values. The positive correlation mainly existed between drug IC50 values and *MUC1*, which exhibited the highest correlation with IC50 of paclitaxel(r=0.170, p<0.001), cisplatin (r=0.313, p<0.001), and tamoxifen (r=0.203, p<0.001). In addition, in Figure 13B, we found that seven drug IC50 were negatively correlated with miR-222-3p expression (all p<0.001), and only the IC50 value of lapatinib was positively correlated with miR-222-3p expression (r=0.201, p<0.001).

3.13 Genetic alteration of ETGs

The analysis of genetic alteration of the ETGs shown in Figure 14A suggested that amplification was the main genetic alteration of nine ETGs except for *NTRK2* and *PIK3R1*. The genetic alteration rate of *MUC1* was highest among 10 ETGs, up to 10%. There was no difference between the ETGs altered group and the unaltered group in OS (Figure 14B), but the DSS of the unaltered group was longer than that of the altered group (p<0.05) (Figure 14C). The median months overall (95% CI) of *NRP1*, *MMP11*, *NTRK2*, and *BIRC5* were not applicable; thus, we analyzed the OS of the unaltered group and the ETGs-altered groups of *EGFR*, *IL6*, *SERPINE1*, *MUC1*, *LAMC2*, and *PIK3R1* (Figure 14D). The median months overall (95% CI) of the unaltered group was 146.50, which was longer than that of six ETG-altered groups.

4 Discussion

While the prognosis for patients diagnosed with early-stage BC is generally favorable, the treatment of metastatic breast cancer poses a significant challenge to public health due to its unfavorable

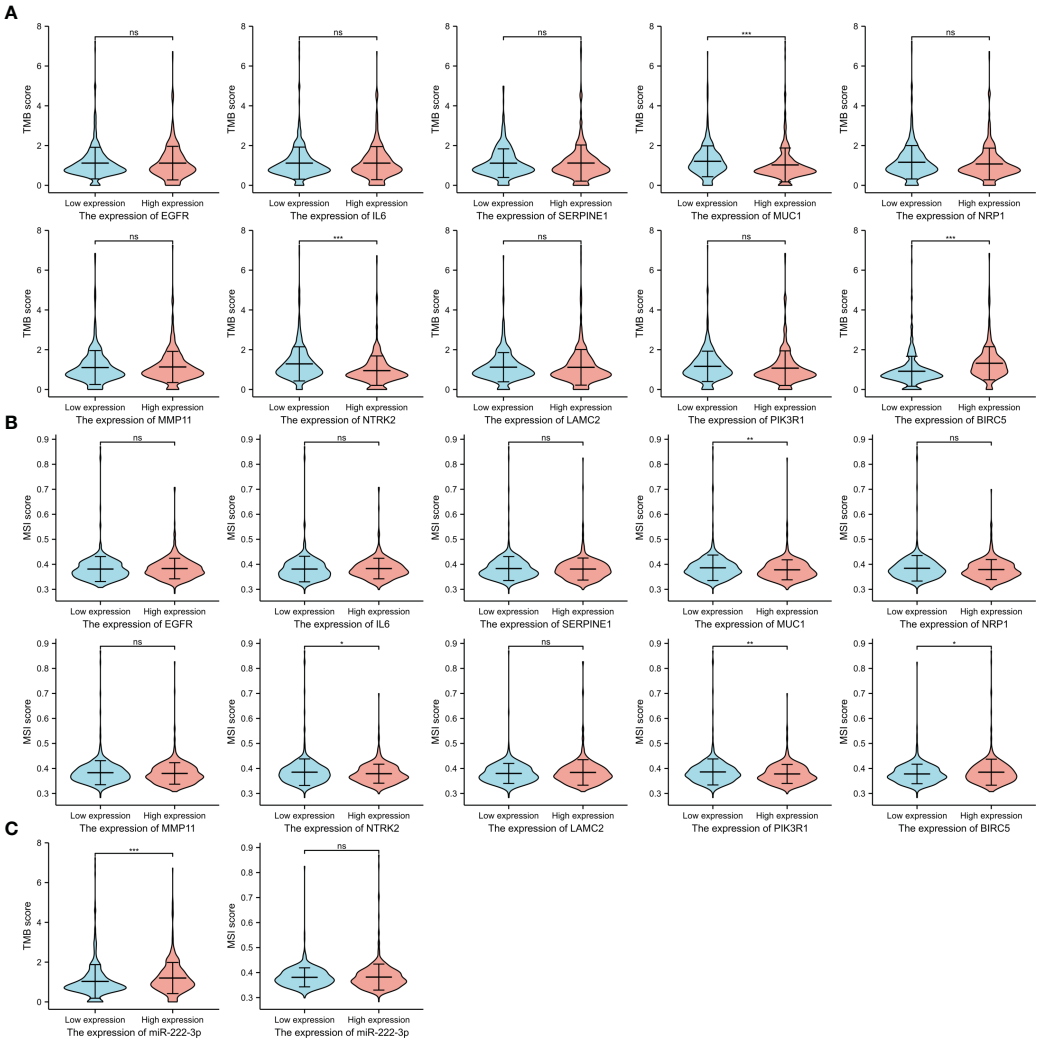


FIGURE 11
The TMB scores **(A)** and MSI scores **(B)** between the high- and low-expression groups of 10 ETGs. The TMB and MSI scores between the high- and low-expression groups of miR-222-3p **(C)**. NS, indicates no statistical difference, *p< 0.05, **p< 0.01, ***p< 0.001.

prognosis. The process of EMT plays a critical role in tumor metastasis (5) and has garnered increasing attention in recent years. Recent research has demonstrated that multiple microRNAs are involved in the progression of BC by regulating EMT through different signaling pathways mediated by various transcription factors, thereby disabling tumor-suppressing or tumor-promoting effects, which might also be served as therapeutic molecules for the treatment of BC (11, 26). Prior research has demonstrated that miR-222-3p functions as a regulator of EMT in BC (14, 15, 27). The present investigation employs bioinformatic analysis to identify 10 fundamental ETGs of miR-222-3p for further investigation of the underlying regulatory mechanisms.

The current investigation utilized the TCGA database to conduct an analysis, which revealed that miR-222-3p exhibited a comparatively reduced expression level in contrast to normal paracancerous tissues. This finding was subsequently confirmed through qRT-PCR in MCF-7 cell lines. Nevertheless, prior research has indicated that miR-222-3p tends to exhibit a relatively elevated expression in BC tissues (28, 29). The incongruity in the outcomes can be primarily attributed to the limited sample size and regional disparities in the selection of BC patients, predominantly in Asia, in the earlier studies. Our findings indicate a significant elevation in miR-222-3p expression in MDA-MB-231 cell lines. Furthermore, the PAM50 subtype analysis revealed that miR-222-3p exhibited significantly higher expression in the basal-like subtype compared to other BC subtypes and normal tissues, which is consistent with previous studies (14, 15, 30). The AUC of the ROC curve for distinguishing basal-like BC and other subtypes was 0.819, suggesting that miR-222-3p may be serve as a specific biomarker of basal-like BC. Furthermore, the clinical implications of miR-222-3p indicate that its heightened expression is inversely correlated with negative status of ER, PR, and HER2 statuses. Specifically, research has demonstrated that miR-222-3p overexpression directly inhibits ER translation, while ER can suppress miR-222-3p expression by enlisting the nuclear receptor corepressor *(NCoR)* and thyroid hormone receptor *(SMRT)* (31). However, the mechanism of interaction between miR-222-3p and PR or HER2 remains unexplored. In addition, our survival analysis revealed that

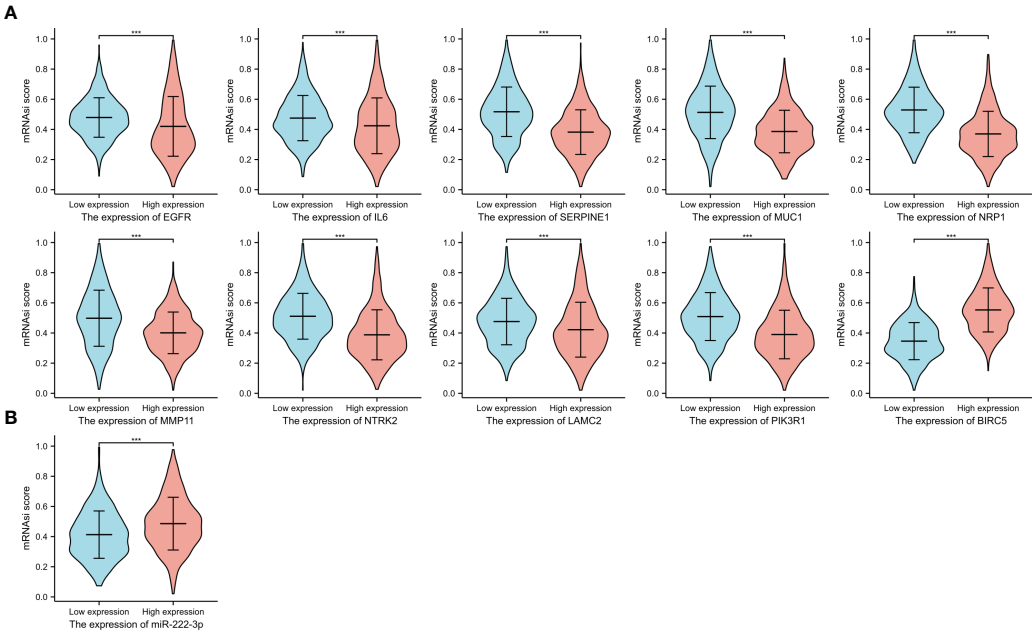


FIGURE 12
The mRNAsi scores between the high and low expression groups of 10 ETGs **(A)**, and the mRNAsi scores between the high and low expression groups of miR-222-3p **(B)**. ***, indicates P< 0.001.

breast cancer patients exhibiting elevated expression levels of miR-222-3p were more likely to experience a poorer prognosis, thus providing further evidence that heightened miR-222-3p expression is linked to increased invasion in BC. Notably, our study also demonstrated, for the first time, that miR-222-3p expression is correlated with a broad range of immune cell infiltration and ICGs, suggesting that it may play a role in regulating the immune microenvironment during the progression of BC. TMB was proposed for efficacy predictions of immunotherapy as a marker (32), and our analysis revealed that BC patients who exhibit elevated expression levels of miR-222-3p may experience greater advantages from immunotherapy. Additionally, our findings indicate a positive correlation between miR-222-3p expression and mRNAsi scores, which suggests that BC patients with high expression of miR-222-3p are more likely to have lower degrees of differentiation and higher levels of cell stemness.

In order to gain a deeper comprehension of the fundamental mechanisms governing the EMT process, a set of 38 genes associated with EMT were identified as the potential targets of miR-222-3p. Through pathway enrichment analysis reveal, it was determined that miR-222-3p may regulate EMT via the PI3K-Akt and HIF-1 signaling pathways. Previous research has indicated that the activated PI3K-Akt signaling pathway plays a direct role in inducing EMT by upregulating the expression of Snail and phosphorylated Twist and also collaborates with other signaling pathways to facilitate EMT either directly or indirectly during the progression of cancers (33). Moreover, it has been demonstrated that the upregulation of hypoxia-inducible factor 1 *(HIF-1)* in breast

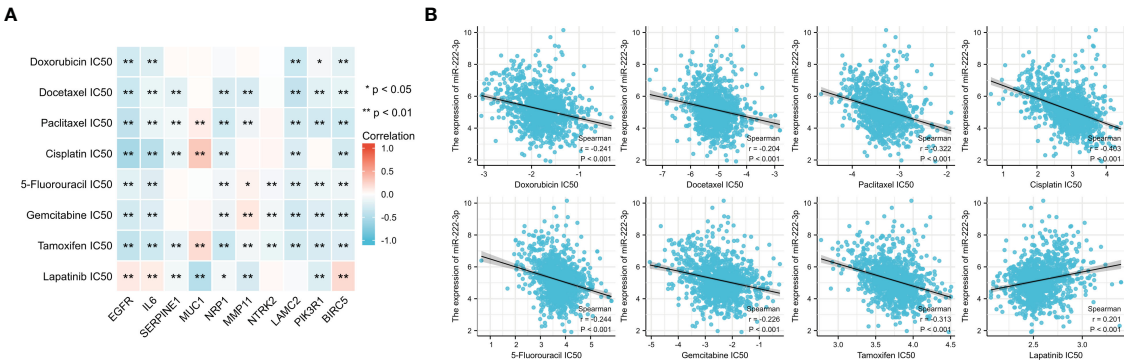


FIGURE 13
The association between the IC50 of eight drugs and ETGs expression **(A)**, and the association between the IC50 of eight drugs and miR-222-3p expression **(B)**. *p< 0.05, **p< 0.01.

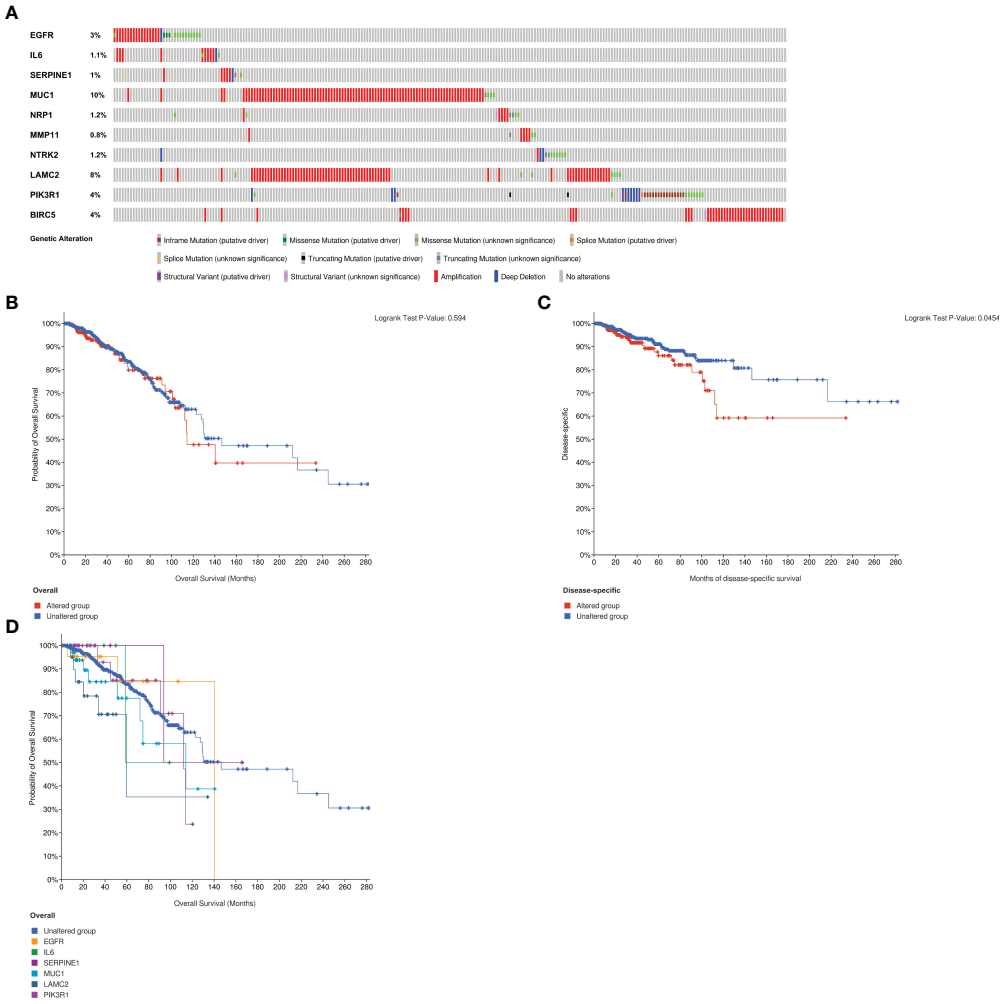


FIGURE 14
Analysis of genetic alteration of 10 ETGs **(A)**. The KM survival curves show the OS and DSS between the ETGs altered group and the unaltered group **(B, C)**. The KM survival curve shows the OS of the ETGs-unaltered group and the ETGs-altered groups of EGFR, IL6, SERPINE1, MUC1, LAMC2, and PIK3R1 **(D)**.

cancer is linked to metastasis as an EMT activator through the mediation of EMT-related signaling pathways, transcription factors, and inflammatory cytokines (34). In this study, we constructed the PPI network of 38 EMT-related genes and subsequently identified 10 top hub genes as the ETGs of miR-222-3p for further research.

Epidermal growth factor receptor *(EGFR)* is a transmembrane protein with tyrosine kinase activity that governs cellular functions, and miR-222-3p has been reported as a downstream modulator of the *EGFR* signaling pathway that regulates EMT as a promotor (35, 36). In this investigation, the top core ETG of miR-222-3p was identified as *EGFR*, and a positive correlation of expression between the two was established. This suggests the possibility of a positive feedback loop between activated EGFR and upregulated miR-222-3p, which may promote EMT in BC. Interestingly, our findings indicate that the *EGFR* expression in BC tissues is lower than that of normal adjacent tissues. Previous studies have reported that EGFR overexpression was detected in only 15%–30% of BC but at least half of basal-like BC (36, 37). Our investigation also found that *EGFR* expression in basal-like BC was significantly higher than in other subtypes, revealing that the overexpression of which is associated with BC invasion. Furthermore, our investigation revealed a positive correlation between elevated *EGFR* expression and a majority of immune cell infiltrates and ICGs, indicating that individuals with basal-like BC who exhibit high *EGFR* expression may derive greater therapeutic benefit from immune checkpoint inhibitor therapy. Recently, the advent of *EGFR*-targeted chimeric antigen receptor T cell therapy has shown promising results in treating Basal-like BC (38). These findings underscore the potential of *EGFR* as a viable therapeutic target for basal-like BC.

Interleukin-6 *(IL6)* is a pro-inflammatory cytokine that is secreted by different cell types and is known to modulate the growth and differentiation of BC (39, 40). Studies have shown that adipocyte-secreted *IL-6* can induce EMT in BC by triggering signal transducer and activated of transcription 3 *(STAT3)* (41). Interestingly, our analysis revealed that IL6 exhibited the strongest positive correlation with many types of immune cell infiltration across multiple ETGs, suggesting its potential role in immune modulation within the tumor microenvironment. Actually,

immune cells in the tumor microenvironment of most cancers were regulated by *IL6* to promote chronic inflammation to help angiogenesis for tumors (42).

Serine protease inhibitor clade E member 1 *(SERPINE1)*, also known as PAI1, is an inhibitor of the plasminogen/plasminase system, the upregulation of which was identified as a biomarker for predicting poor outcome and associated with the EMT process in BC (43, 44). Our findings indicate that the expression of *SERPINE1* was higher in BC tissues, particularly in luminal A BC. However, we did not observe a significant association between SERINE1 expression and prognosis, although the high-expression group tended to have a worse prognosis but without statistical significance. It has been reported that BC patients with high levels of *SERPINE1* may benefit from adjuvant chemotherapy (43). In our study, we found a negative correlation between *SERPINE1* expression and IC50 values of docetaxel, paclitaxel, cisplatin, tamoxifen, and lapatinib, suggesting that BC patients with high levels of *SERPINE1* might receive a better treatment effect of chemotherapy.

Mucin 1 *(MUC1)*, also known as CA15-3, is a transmembrane heterodimeric glycoprotein that is aberrantly overexpressed in BC and serves as a serum diagnostic biomarker for BC (45, 46). In the present study, we evaluated the diagnostic value of the elevated *MUC1* expression and found it to be favorable, with a sensitivity of 91.2% and a specificity of 67.3%. Previous research has demonstrated that *MUC1* can facilitate *IGF-1*-induced EMT in the MCF-7 breast cancer cell line and induce tamoxifen resistance in ER-positive BC patients (47), and we found the high expression of *MUC1* was related to high drug sensitivity of tamoxifen, cisplatin, and paclitaxel. Furthermore, the TCGA dataset revealed a negative correlation between the expression of *MUC1* and ICGs and TMB score in patients with breast cancer, indicating that those with low *MUC1* expression may derive greater benefit from immunotherapy. Despite extensive research on MUC1-based immunotherapy over the past few decades, its efficacy remains limited by factors such as the presence of diverse isoforms and immunosuppression (48).

Neuropilin 1 *(NRP1)* is a transmembrane glycoprotein that contributes to cancer development by inducing EMT through various signaling pathways (49, 50). Moreover, it has been reported that *NRP1* is expressed on human pDCs, which contributes to priming immune responses (51), and a positive correlation between *NRP1* expression and pDCs immune infiltrate was found in our study. *NRP1*-mediated immune modulation in cancer has garnered significant attention in recent years (52), and we found that its expression was associated with ICGs, suggesting that it was a promising checkpoint target in BC immunotherapy.

As a member of the matrix metalloproteinase (MMP) family, *MMP11*, also termed stromelysin 3, is overexpressed in BC tissues and BC cell lines, promoting tumor cell proliferation (53, 54). The available evidence suggests that miR-125a directly targets *MMP11*, resulting in downregulation and subsequent suppression of EMT and migration and invasion in osteosarcoma and hepatocellular carcinoma (55, 56). However, the mechanism by which *MMP11* is involved in EMT in BC has yet to be reported. Our study is the first to demonstrate a strong correlation between MMP11 expression and immune cells, ICGs expression, stemness, and drug sensitivity. Furthermore, we have identified MMP11 as a potentially valuable diagnostic marker, with an AUC of 0.993 in the diagnostic ROC curve, with a sensitivity and specificity of 97.3% and 95.2%, respectively. Prior studies have indicated that the combination of *MMP11* and Doppler ultrasound may enhance the diagnostic efficacy for early-stage BC patients (57). These findings propose that *MMP11* could serve as a potential biomarker for precise diagnosis in BC, which deserves our further in-depth study.

Neurotrophic receptor tyrosine kinase 2 *(NTRK2)*, also referred to as *KRKB*, is a constituent of the neurotrophic tyrosine kinase receptors family, serving as a regulator that facilitates EMT by activating PI3K/AKT and IL6/JAK2/STAT3 signaling pathways to foster tumor metastasis (58). Interestingly, while prior evidence has established a correlation between the upregulation of *NTRK2* and its main ligand brain-derived neurotrophic factor (BDNF) with the advancement of cancer progression (59), our current study has yielded contradictory results. Specifically, the expression of NTRK2 was observed to be lower in BC tissues, and patients exhibiting high levels of *NTRK2* expression demonstrated a more favorable prognosis, suggesting that *NTRK2* may play a defensive role in the BC progression. One plausible hypothesis is that *NTRK2* may be involved in the recruitment of immune cells within the tumor microenvironment. Our investigation revealed a positive correlation between *NTRK2* expression and heightened infiltration of $CD8^+$ T cells, NK cells, and DCs, all of which are critical components of anti-tumor immunity (60, 61). The precise mechanism underlying *NTRK2*-mediated anti-tumor immunity remains unreported and may differ from the signaling pathway activated by *NTRK2* and BDNF, warranting further investigation.

Laminin γ2 *(LAMC2)* is a subunit of the heterotrimeric glycoprotein laminin-332, which has been demonstrated to facilitate the proliferation and metastasis of cancer cells in basal-like BC (62). It has been observed that the secretion of *LAMC2* by intrahepatic cholangiocarcinoma cells can promote EMT (63). Additionally, *LAMC2* has been shown to regulate gemcitabine sensitivity in pancreatic ductal adenocarcinoma through EMT (64). In the current investigation, we found that *LAMC2* expression in basal-like BC was higher and positively correlated with miR-222-3p expression. Furthermore, we have found a link between higher *LAMC2* expression and reduced drug sensitivity and higher infiltration levels of most immune cells.

Phosphoinositide-3-kinase regulatory subunit 1 *(PIK3R1)* is a regulatory subunit of the PI3K-Akt signaling pathway, which is a tumor suppressor that is frequently downregulated in BC (65, 66). Previous research has shown that under-expression of *PIK3R1* is linked to poor prognosis in BC patients (67). Our study aligns with these findings, as we observed the downregulation of *PIK3R1* in BC and noted that patients with low *PIK3R1* expression had a shorter DSS. Additionally, it has been reported that *PIK3R1* is a target gene of miR-21 and that knockdown of miR-21 leads to upregulation of *PIK3R1*, which in turn inhibits EMT and the activation of the PI3K-Akt signaling pathway, ultimately suppressing BC development (68). Based on the findings of the present and previous studies, we supposed that miR-222-3p might target and downregulate *PIK3R1* expression to promote EMT via the activation of the PI3K-Akt signaling pathway in BC.

Baculoviral inhibitor of apoptosis repeat containing 5 *(BIRC5)* is a mitotic spindle checkpoint gene that belongs to the inhibitor of the apoptosis family, and a mitotic spindle checkpoint gene has been observed to be overexpressed in BC and linked to unfavorable clinical outcomes (69, 70). This study reveals that expression is notably elevated in BC tissues, particularly in basal-like and luminal B subtypes and that *BIRC5* may serve as a valuable biomarker for BC diagnosis, as evidenced by ROC analysis. Patients with high BIRC5 expression levels were found to have a poorer prognosis. Additionally, it has been reported that *BIRC5* can stimulate the expression of superoxide dismutase 1 (*SOD1*) in cancer-associated fibroblasts and transform them into myofibroblasts to promote EMT in BC (71). Furthermore, our investigation revealed a correlation between elevated *BIRC5* expression and reduced levels of NK cells, CD8+ cells, and increased miRNA scores, indicating a potential immunosuppressive function of *BIRC5* and its ability to promote BC cell stemness. Additionally, heightened *BIRC5* expression was linked to the expression of ICGs and TMB, highlighting the potential of *BIRC5* as a promising target for BC immunotherapy.

However, it is important to note that our study primarily relied on data obtained from the TCGA database, which needs further validation. To address the limitations, first, the luciferase reporter assay should be performed to validate the relationship between miR-222-3p and its identified ETGs. Subsequently, it is imperative to conduct further investigation into the EMT-regulated mechanism of the ETGs. Additionally, it is crucial to ascertain whether the ETGs have the potential to serve as therapeutic targets in the future. In conclusion, the findings of this study hold significant implications for future research endeavors aimed at exploring the EMT mechanism of BC, which could potentially yield promising treatment modalities for patients afflicted with BC.

5 Conclusion

In conclusion, our study has provided evidence indicating that miR-222-3p was overexpressed in basal-like BC and has the potential to serve as a specific biomarker of basal-like BC. Additionally, elevated expression levels of miR-222-3p in BC are associated with unfavorable prognosis. Through our investigation, we have identified 10 core ETGs of miR-222-3p, among which *MUC1*, *MMP11*, and BIR*BIRC5* may serve as useful diagnostic biomarkers for BC, and *NTRK2*, *PIK3R1*, and *BIRC5* may serve as biomarkers for predicting prognosis. Comprehensive analysis of the association between the expression of 10 ETGs and immune cells, ICGs, TMB, MSI, stemness, and drug sensitivity indicates their potential association with the tumor microenvironment in the progression of breast cancer. These findings suggest that these ETGs may serve as novel therapeutic targets for the treatment of BC.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

YF organized the article writing and critically modified the manuscript. JW and RDZ modified the manuscript. QZ drafted the manuscript and were responsible for the acquisition of data. CC and ZC participated in the data analysis. RJZ contributed to the literature search. CS checked and corrected language expression. All authors read and approved the manuscript and agreed to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

This work was supported by funds from the Foundation of Basic and Applied Basic Research of Guangdong Province, China (Grant No. 2022A1515220202), and the Youth Science Foundation of the Cancer Hospital of Shantou University Medical College (Grant No. 2023A002).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1189635/full#supplementary-material

References

1. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: an overview. *Int J Cancer* (2021). doi: 10.1002/ijc.33588

2. Cui G, Wu J, Lin J, Liu W, Chen P, Yu M, et al. Graphene-based nanomaterials for breast cancer treatment: promising therapeutic strategies. *J Nanobiotechnol* (2021) 19 (1):211. doi: 10.1186/s12951-021-00902-8

3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70 (1):7–30. doi: 10.3322/caac.21590

4. Xu X, Zhang M, Xu F, Jiang S. Wnt signaling in breast cancer: biological mechanisms, challenges and opportunities. *Mol Cancer* (2020) 19(1):165. doi: 10.1186/s12943-020-01276-5

5. Bakir B, Chiarella AM, Pitarresi JR, Rustgi AK. EMT, MET, plasticity, and tumor metastasis. *Trends Cell Biol* (2020) 30(10):764–76. doi: 10.1016/j.tcb.2020.07.003

6. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* (2008) 133 (4):704–15. doi: 10.1016/j.cell.2008.03.027

7. Taki M, Abiko K, Ukita M, Murakami R, Yamanoi K, Yamaguchi K, et al. Tumor immune microenvironment during epithelial-mesenchymal transition. *Clin Cancer Res* (2021) 27(17):4669–79. doi: 10.1158/1078-0432.CCR-20-4459

8. Shibue T, Weinberg RA. EMT. CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol* (2017) 14(10):611–29. doi: 10.1038/nrclinonc.2017.44

9. Saliminejad K, Khorram Khorshid HR, Soleymani Fard S, Ghaffari SH. An overview of microRNAs: biology, functions, therapeutics, and analysis methods. *J Cell Physiol* (2019) 234(5):5451–65. doi: 10.1002/jcp.27486

10. Shi Y, Liu Z, Lin Q, Luo Q, Cen Y, Li J, et al. MiRNAs and cancer: key link in diagnosis and therapy. *Genes (Basel)* (2021) 12(8):1289. doi: 10.3390/genes12081289

11. Feng J, Hu S, Liu K, Sun G, Zhang Y. The role of MicroRNA in the regulation of tumor epithelial-mesenchymal transition. *Cells* (2022) 11(13):1981. doi: 10.3390/cells11131981

12. Pan G, Liu Y, Shang L, Zhou F, Yang S. EMT-associated microRNAs and their roles in cancer stemness and drug resistance. *Cancer Commun (Lond)* (2021) 41 (3):199–217. doi: 10.1002/cac2.12138

13. Garofalo M, Quintavalle C, Romano G, Croce CM, Condorelli G. miR221/222 in cancer: their role in tumor progression and response to therapy. *Curr Mol Med* (2012) 12(1):27–33. doi: 10.2174/156652412798376170

14. Stinson S, Lackner MR, Adai AT, Yu N, Kim HJ, O'Brien C, et al. miR-221/222 targeting of trichorhinophalangeal 1 (TRPS1) promotes epithelial-to-mesenchymal transition in breast cancer. *Sci Signal* (2011) 4(186):pt5. doi: 10.1126/scisignal.2002258

15. Liang YK, Lin HY, Dou XW, Chen M, Wei XL, Zhang YQ, et al. MiR-221/222 promote epithelial-mesenchymal transition by targeting Notch3 in breast cancer cell lines. *NPJ Breast Cancer* (2018) 4:20. doi: 10.1038/s41523-018-0073-7

16. Hashimoto Y, Akiyama Y, Yuasa Y. Multiple-to-multiple relationships between microRNAs and target genes in gastric cancer. *PloS One* (2013) 8(5):e62589. doi: 10.1371/journal.pone.0062589

17. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* (2015) 19(1A):A68–77. doi: 10.5114/wo.2014.47136

18. Sticht C, de la Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. *PloS One* (2018) 13(10):e0206239. doi: 10.1371/journal.pone.0206239

19. Zhao M, Liu Y, Zheng C, Qu H. dbEMT 2.0: an updated database for epithelial-mesenchymal transition genes with experimentally verified information and precalculated regulation information for cancer metastasis. *J Genet Genomics* (2019) 46(12):595–7. doi: 10.1016/j.jgg.2019.11.010

20. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* (2019) 47(D1):D607–13. doi: 10.1093/nar/gky1131

21. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

22. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* (2017) 9(1):34. doi: 10.1186/s13073-017-0424-2

23. Bonneville R, Krook MA, Kautto EA, Miya J, Wing MR, Chen HZ, et al. Landscape of microsatellite instability across 39 cancer types. *JCO Precis Oncol* (2017) 2017:PO.17.00073. doi: 10.1200/PO.17.00073

24. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173(2):338–354.e15. doi: 10.1016/j.cell.2018.03.034

25. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2013) 41(Database issue):D955–61. doi: 10.1093/nar/gks1111

26. Kumar A, Golani A, Kumar LD. EMT in breast cancer metastasis: an interplay of microRNAs, signaling pathways and circulating tumor cells. *Front Biosci (Landmark Ed)* (2020) 25(5):979–1010. doi: 10.2741/4844

27. Hwang MS, Yu N, Stinson SY, Yue P, Newman RJ, Allan BB, et al. miR-221/222 targets adiponectin receptor 1 to promote the epithelial-to-mesenchymal transition in breast cancer. *PloS One* (2013) 8(6):e66502. doi: 10.1371/journal.pone.0066502

28. Amini S, Abak A, Estiar MA, Montazeri V, Abhari A, Sakhinia E. Expression analysis of MicroRNA-222 in breast cancer. *Clin Lab* (2018) 64(4):491–6. doi: 10.7754/Clin.Lab.2017.171002

29. Zong Y, Zhang Y, Sun X, Xu T, Cheng X, Qin Y. miR-221/222 promote tumor growth and suppress apoptosis by targeting lncRNA GAS5 in breast cancer. *Biosci Rep* (2019) 39(1):BSR20181859. doi: 10.1042/BSR20181859

30. Li Y, Liang C, Ma H, Zhao Q, Lu Y, Xiang Z, et al. miR-221/222 promotes s-phase entry and cellular migration in control of basal-like breast cancer. *Molecules* (2014) 19(6):7122–37. doi: 10.3390/molecules19067122

31. Di Leva G, Gasparini P, Piovan C, Ngankeu A, Garofalo M, Taccioli C, et al. MicroRNA cluster 221-222 and estrogen receptor alpha interactions in breast cancer. *J Natl Cancer Inst* (2010) 102(10):706–21. doi: 10.1093/jnci/djq102

32. Zhao F, Li Z, Dong Z, Wang Z, Guo P, Zhang D, et al. Exploring the potential of exosome-related LncRNA pairs as predictors for immune microenvironment, survival outcome, and microbiotain landscape in esophageal squamous cell carcinoma. *Front Immunol* (2022) 13:918154. doi: 10.3389/fimmu.2022.918154

33. Xu W, Yang Z, Lu N. A new role for the PI3K/Akt signaling pathway in the epithelial-mesenchymal transition. *Cell Adh Migr* (2015) 9(4):317–24. doi: 10.1080/19336918.2015.1016686

34. Liu ZJ, Semenza GL, Zhang HF. Hypoxia-inducible factor 1 and breast cancer metastasis. *J Zhejiang Univ Sci B* (2015) 16(1):32–43. doi: 10.1631/jzus.B1400221

35. Teixeira AL, Gomes M, Medeiros R. EGFR signaling pathway and related-miRNAs in age-related diseases: the example of miR-221 and miR-222. *Front Genet* (2012) 3:286. doi: 10.3389/fgene.2012.00286

36. Masuda H, Zhang D, Bartholomeusz C, Doihara H, Hortobagyi GN, Ueno NT. Role of epidermal growth factor receptor in breast cancer. *Breast Cancer Res Treat* (2012) 136(2):331–45. doi: 10.1007/s10549-012-2289-9

37. Hsu JL, Hung MC. The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer. *Cancer Metastasis Rev* (2016) 35(4):575–88. doi: 10.1007/s10555-016-9649-6

38. Xia L, Zheng ZZ, Liu JY, Chen YJ, Ding JC, Xia NS, et al. EGFR-targeted CAR-T cells are potent and specific in suppressing triple-negative breast cancer both *in vitro* and *in vivo*. *Clin Transl Immunol* (2020) 9(5):e01135. doi: 10.1002/cti2.1135

39. Masjedi A, Hashemi V, Hojjat-Farsangi M, Ghalamfarsa G, Azizi G, Yousefi M, et al. The significant role of interleukin-6 and its signaling pathway in the immunopathogenesis and treatment of breast cancer. *BioMed Pharmacother* (2018) 108:1415–24. doi: 10.1016/j.biopha.2018.09.177

40. Felcher CM, Bogni ES, Kordon EC. IL-6 cytokine family: a putative target for breast cancer prevention and treatment. *Int J Mol Sci* (2022) 23(3):1809. doi: 10.3390/ijms23031809

41. Gyamfi J, Lee YH, Eom M, Choi J. Interleukin-6/STAT3 signalling regulates adipocyte induced epithelial-mesenchymal transition in breast cancer cells. *Sci Rep* (2018) 8(1):8859. doi: 10.1038/s41598-018-27184-9

42. Jones SA, Jenkins BJ. Recent insights into targeting the IL-6 cytokine family in inflammatory diseases and cancer. *Nat Rev Immunol* (2018) 18(12):773–89. doi: 10.1038/s41577-018-0066-7

43. Duffy MJ, McGowan PM, Harbeck N, Thomssen C, Schmitt M. uPA and PAI-1 as biomarkers in breast cancer: validated for clinical use in level-of-evidence-1 studies. *Breast Cancer Res* (2014) 16(4):428. doi: 10.1186/s13058-014-0428-4

44. Xu J, Zhang W, Tang L, Chen W, Guan X. Epithelial-mesenchymal transition induced PAI-1 is associated with prognosis of triple-negative breast cancer patients. *Gene* (2018) 670:7–14. doi: 10.1016/j.gene.2018.05.089

45. Stergiou N, Nagel J, Pektor S, Heimes AS, Jäkel J, Brenner W, et al. Evaluation of a novel monoclonal antibody against tumor-associated MUC1 for diagnosis and prognosis of breast cancer. *Int J Med Sci* (2019) 16(9):1188–98. doi: 10.7150/ijms.35452

46. Liao G, Wang M, Ou Y, Zhao Y. IGF-1-induced epithelial-mesenchymal transition in MCF-7 cells is mediated by MUC1. *Cell Signal* (2014) 26(10):2131–7. doi: 10.1016/j.cellsig.2014.06.004

47. Merikhian P, Ghadirian R, Farahmand L, Mansouri S, Majidzadeh-A K. MUC1 induces tamoxifen resistance in estrogen receptor-positive breast cancer. *Expert Rev Anticancer Ther* (2017) 17(7):607–13. doi: 10.1080/14737140.2017.1340837

48. Li Z, Yang D, Guo T, Lin M. Advances in MUC1-mediated breast cancer immunotherapy. *Biomolecules* (2022) 12(7):952. doi: 10.3390/biom12070952

49. Li X, Zhou Y, Hu J, Bai Z, Meng W, Zhang L, et al. Loss of neuropilin1 inhibits liver cancer stem cells population and blocks metastasis in hepatocellular carcinoma via epithelial-mesenchymal transition. *Neoplasma* (2021) 68(2):325–33. doi: 10.4149/neo_2020_200914N982

50. Jin Q, Ren Q, Chang X, Yu H, Jin X, Lu X, et al. Neuropilin-1 predicts poor prognosis and promotes tumor metastasis through epithelial-mesenchymal transition in gastric cancer. *J Cancer* (2021) 12(12):3648–59. doi: 10.7150/jca.52851

51. Chaudhary B, Khaled YS, Ammori BJ, Elkord E. Neuropilin 1: function and therapeutic potential in cancer. *Cancer Immunol Immunother* (2014) 63(2):81–99. doi: 10.1007/s00262-013-1500-0

52. Chuckran CA, Liu C, Bruno TC, Workman CJ, Vignali DA. Neuropilin-1: a checkpoint target with unique implications for cancer immunology and immunotherapy. *J Immunother Cancer* (2020) 8(2):e000967. doi: 10.1136/jitc-2020-000967

53. Kasper G, Reule M, Tschirschmann M, Dankert N, Stout-Weider K, Lauster R, et al. Stromelysin-3 over-expression enhances tumourigenesis in MCF-7 and MDA-MB-231 breast cancer cell lines: involvement of the IGF-1 signalling pathway. *BMC Cancer* (2007) 7:12. doi: 10.1186/1471-2407-7-12

54. Zhuang Y, Li X, Zhan P, Pi G, Wen G. MMP11 promotes the proliferation and progression of breast cancer through stabilizing Smad2 protein. *Oncol Rep* (2021) 45 (4):16. doi: 10.3892/or.2021.7967

55. Bi Q, Tang S, Xia L, Du R, Fan R, Gao L, et al. Ectopic expression of MiR-125a inhibits the proliferation and metastasis of hepatocellular carcinoma by targeting MMP11 and VEGF. *PloS One* (2012) 7(6):e40169. doi: 10.1371/journal.pone.0040169

56. Waresijiang N, Sun J, Abuduaini R, Jiang T, Zhou W, Yuan H. The downregulation of miR–125a–5p functions as a tumor suppressor by directly targeting MMP–11 in osteosarcoma. *Mol Med Rep* (2016) 13(6):4859–64. doi: 10.3892/mmr.2016.5141

57. Ren H, Shen Z, Shen J, Zhang Y, Zhang Y. Diagnostic value of Doppler ultrasound parameters combined with MMP-11 in early breast cancer and benign breast diseases. *Oncol Lett* (2020) 20(2):1028–32. doi: 10.3892/ol.2020.11676

58. Kim MS, Lee WS, Jeong J, Kim SJ, Jin W. Induction of metastatic potential by TrkB via activation of IL6/JAK2/STAT3 and PI3K/AKT signaling in breast cancer. *Oncotarget* (2015) 6(37):40158–71. doi: 10.18632/oncotarget.5522

59. Serafim Junior V, Fernandes GMM, Oliveira-Cucolo JG, Pavarino EC, Goloni-Bertollo EM. Role of tropomyosin-related kinase b receptor and brain-derived neurotrophic factor in cancer. *Cytokine* (2020) 136:155270. doi: 10.1016/j.cyto.2020.155270

60. Fu C, Jiang A. Dendritic cells and CD8 T cell immunity in tumor microenvironment. *Front Immunol* (2018) 9:3059. doi: 10.3389/fimmu.2018.03059

61. Myers JA, Miller JS. Exploring the NK cell platform for cancer immunotherapy. *Nat Rev Clin Oncol* (2021) 18(2):85–100. doi: 10.1038/s41571-020-0426-7

62. He Y, Xiao B, Lei T, Xuan J, Zhu Y, Kuang Z, et al. LncRNA T376626 is a promising serum biomarker and promotes proliferation, migration, and invasion via binding to LAMC2 in triple-negative breast cancer. *Gene* (2023) 860:147227. doi: 10.1016/j.gene.2023.147227

63. Cen W, Li J, Tong C, Zhang W, Zhao Y, Lu B, et al. Intrahepatic cholangiocarcinoma cells promote epithelial-mesenchymal transition of hepatocellular carcinoma cells by secreting LAMC2. *J Cancer* (2021) 12(12):3448–57. doi: 10.7150/jca.55627

64. Okada Y, Takahashi N, Takayama T, Goel A. LAMC2 promotes cancer progression and gemcitabine resistance through modulation of EMT and ATP-binding cassette transporters in pancreatic ductal adenocarcinoma. *Carcinogenesis* (2021) 42(4):546–56. doi: 10.1093/carcin/bgab011

65. Turturro SB, Najor MS, Yung T, Portt L, Malarkey CS, Abukhdeir AM, et al. Somatic loss of PIK3R1 may sensitize breast cancer to inhibitors of the MAPK pathway. *Breast Cancer Res Treat* (2019) 177(2):325–33. doi: 10.1007/s10549-019-05320-x

66. Liu Y, Wang D, Li Z, Li X, Jin M, Jia N, et al. Pan-cancer analysis on the role of PIK3R1 and PIK3R2 in human tumors. *Sci Rep* (2022) 12(1):5924. doi: 10.1038/s41598-022-09889-0

67. Cizkova M, Vacher S, Meseure D, Trassard M, Susini A, Mlcuchova D, et al. PIK3R1 underexpression is an independent prognostic marker in breast cancer. *BMC Cancer* (2013) 13:545. doi: 10.1186/1471-2407-13-545

68. Yan LX, Liu YH, Xiang JW, Wu QN, Xu LB, Luo XL, et al. PIK3R1 targeting by miR-21 suppresses tumor cell migration and invasion by reducing PI3K/AKT signaling and reversing EMT, and predicts clinical outcome of breast cancer. *Int J Oncol* (2016) 48(2):471–84. doi: 10.3892/ijo.2015.3287

69. Li K, Liu T, Chen J, Ni H, Li W. Survivin in breast cancer-derived exosomes activates fibroblasts by up-regulating SOD1, whose feedback promotes cancer proliferation and metastasis. *J Biol Chem* (2020) 295(40):13737–52. doi: 10.1074/jbc.RA120.013805

70. Wang C, Zheng X, Shen C, Shi Y. MicroRNA-203 suppresses cell proliferation and migration by targeting BIRC5 and LASP1 in human triple-negative breast cancer cells. *J Exp Clin Cancer Res* (2012) 31(1):58. doi: 10.1186/1756-9966-31-58

71. Dai JB, Zhu B, Lin WJ, Gao HY, Dai H, Zheng L, et al. Identification of prognostic significance of BIRC5 in breast cancer using integrative bioinformatics analysis. *Biosci Rep* (2020) 40(2):BSR20193678. doi: 10.1042/BSR20193678

10.3389/fonc.2023.1189635

Glossary

BC	breast cancer
EMT	epithelial–mesenchymal transition
miRNAs	microRNAs
miR-222-3p	microRNA-222-3p
ZEB2	Zinc finger E-box-binding homeobox 2
TCGA	The Cancer Genome Atlas
RPM	read per million
GEO	Gene Expression Omnibus
DEGs	differentially expressed genes
ETGs	EMT-related target genes
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
PPI	protein–protein interaction
FPKM	fragments per kilobase million (FPKM)
TPM	transcripts per million
ROC	receiver operating characteristic
qRT-PCR	quantitative real-time PCR
KM	Kaplan–Meier
OS	overall survival
DSS	disease-specific survival
ssGSEA	single-sample GSEA
ICGs	immune checkpoint genes
TMB	tumor mutational burden
MSI	microsatellite instability
mRNAsi	mRNA expression-based stemness index
OCLR	one-class logistic regression
GDSC	Genomics of Drug Sensitivity in Cancer
HER2	human epidermal growth factor receptor 2
ER	estrogen receptor
PR	progesterone receptor
AUC	area under the curve
NCoR	nuclear receptor corepressor
SMRT	thyroid hormone receptor
HIF-1	hypoxia-inducible factor 1
EGFR	epidermal growth factor receptor
IL6	interleukin-6
STAT3	signal transducer and activated of transcription 3
SERINE1	serine protease inhibitor clade E member 1
NRP1	neuropilin 1
MMP	matrix metalloproteinase
NTRK2	neurotrophic receptor tyrosine kinase 2
BDNF	brain-derived neurotrophic factor
LAMC2	laminin γ2
PIK3R1	phosphoinositide-3-kinase regulatory subunit 1
BIRC5	baculoviral inhibitor of apoptosis repeat containing 5
SOD1	superoxide dismutase 1

frontiers | Frontiers in Oncology

TYPE Case Report
PUBLISHED 24 July 2023
DOI 10.3389/fonc.2023.1204041

OPEN ACCESS

EDITED BY
Maha Mohamed Saber-Ayad,
University of Sharjah, United Arab Emirates

REVIEWED BY
Alessandro De Luca,
Sapienza University of Rome, Italy
Marcus Vetter,
University Hospital of Basel, Switzerland

*CORRESPONDENCE
Chen Song
635320758@qq.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 11 April 2023
ACCEPTED 06 July 2023
PUBLISHED 24 July 2023

CITATION
Lv H, Tian A, Zhao S, Zhao J and Song C (2023) Next-generation sequencing-based detection in a breast MMPMN patient with EGFR T790M mutation: a rare case report and literature review.
Front. Oncol. 13:1204041.
doi: 10.3389/fonc.2023.1204041

Next-generation sequencing-based detection in a breast MMPMN patient with EGFR T790M mutation: a rare case report and literature review

Huiyun Lv[1†], Aijuan Tian[2†], Shanshan Zhao[1†], Jinbo Zhao[1] and Chen Song[1*]

[1]Department of Oncology, The Second Hospital of Dalian Medical University, Dalian, China,
[2]Department of Nuclear Medicine, The Second Hospital of Dalian Medical University, Dalian, China

Multiple primary malignant neoplasms (MPMNs) are difficult to identify from the metastasis or recurrence of malignant tumors. Additionally, the genetic mutations in each primary tumor vary from each other; therefore, it is critical to explore potential abnormal genes. Next-generation sequencing (NGS) technology has emerged as a reliable approach for detecting mutated genes in primary tumors and can provide several targeted therapeutic options for patients with MPMNs. Here, we report a case of metachronous multiple primary malignant neoplasm (MMPMN) patient with primary ovarian and breast cancer. Targeted NGS genetic profiling revealed a rare EGFR T790M mutation in this patient's primary breast tumor tissue, which has only been reported previously in breast cancer (BC). Based on the NGS results, osimertinib was recommended for this patient. Although this patient did not receive osimertinib because of gastrointestinal hemorrhage, this case highlights the significance of NGS technology in the diagnosis and treatment of MPMNs.

KEYWORDS

EGFR T790M mutation, multiple primary malignant neoplasms, next-generation sequencing, osimertinib, breast cancer

1 Introduction

Multiple primary malignant neoplasms (MPMNs) present an increasing incidence rate owing to the detection of early stages of cancer and the development of effective therapeutic strategies (1). MPMNs are defined as two or more unrelated primary malignant tumors that originate from different organs and occur simultaneously or one after the other (2). MPMNs are classified into two subtypes: synchronous multiple primary malignant neoplasms (SMPMNs) and metachronous multiple primary malignant neoplasms (MMPMNs). SMPMNs are defined as secondary and primary cancers that occur

simultaneously or within 6 months of the first primary cancer. If the interval time is more than 6 months, such tumors are called MMPMNs (3). To date, the prevalence of MPMNs is approximately 0.7%–11% and reports of MPMNs mainly focus on lung cancer and gastrointestinal tumors (4). There are few reports on female patients with multiple primary malignant neoplasms of breast cancer or genital malignancies.

The epidermal growth factor receptor (EGFR) is one of the major oncogenes identified in a variety of human cancers, including breast cancer (5), and is one of the most common driver genes in non-small cell lung cancer (NSCLC) (6). Although tyrosine kinase inhibitor (TKI) targeting EGFR have shown good initial response in NSCLC with mutated EGFR genes, the development of acquired resistance remains inevitable and has emerged as a major limitation of EGFR-targeted therapies with TKIs (7), with disease progression 10–12 months after treatment initiation in most patients (8). In approximately 60% of cases, the most frequent mechanism of acquired resistance is secondary T790M mutation in exon 20, and osimertinib is the standard second-line therapy (9). However, 2% of patients harbor either somatic or germline T790M mutations before any exposure to EGFR-TKIs, resulting in primary resistance (10). In contrast, EGFR mutations have been reported to be rare in human BC, whereas EGFR overexpression and/or amplification have been shown to occur frequently in human breast cancer (5, 11). EGFR T790M mutations are thought to be rare (12). In recent years, next-generation sequencing (NGS) has become available for distinguishing between multiple primary cancers and primary cancer metastasis in MPMNs and has facilitated the identification of targetable gene mutations in different primary tumors in patients. Various clinical studies have shown the promise of site-specific treatment and targeted therapy based on NGS testing results (13). Therefore, NGS containing related genes of great significance for the diagnosis and precise treatment of cancer is urgently needed and warrants further clinical investigation.

Here, we present a rare case of a 59-year-old female patient with MMPMNs harboring a pathogenic EGFR T790M mutation in breast cancer primary sites by NGS genetic testing.

2 Case presentation

A 59-year-old female without a genetic family history was diagnosed with high-grade left ovarian serous papillary carcinoma in June 2015. However, her detailed medical history revealed no family history of cancer or exposure to environmental risk factors. Subsequently, the patient underwent radical hysterectomy, bilateral salpingo-oophorectomy, partial rectotomy, and pelvic lymphadenectomy. Postoperative pathological evaluation of the resected tumoral tissues indicated high-grade serous papillary carcinoma of the left ovary, invading the abdomen and pelvic cavity. Metastatic lesions included the muscular wall of the uterine body, left fallopian tube, spleen, omentum greater, part of the diaphragm and mass, appendix, and part of the muscular layer of the rectum with the intestinal wall. Immunohistochemistry (IHC) was positive for estrogen receptor (ER), progesterone receptor (PR), high-grade cervical squamous intraepithelial lesion marker p16, and tumor marker p53, and the pathological stage was IIIC. After surgery, she received systemic chemotherapy with the TC protocol (taxol 300 mg (175 mg/m^2) in combination with carboplatin 450 mg (dosed by AUC), given every 21 days); however, the number of cycles was unknown. In April 2017, the patient underwent another abdominal tumor resection to relieve the symptoms caused by the tumors compressing the abdominal organs. Histopathology and immunochemistry confirmed metastatic adenocarcinoma of the abdominal cavity, consistent with high-grade serous carcinoma metastasis in the pelvic cavity.

Seven years after the diagnosis of ovarian cancer, the patient began to feel ill, and the main symptoms included abdominal distension and mild jaundice. She underwent whole-body Fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) in September 2022. High uptake of 18F-FDG was noted within the known lesions in the hepatopancreatic lesions with increasing radioactivity uptake (Figure 1A), which was new and increased in scope from March 2017. All of these were considered malignant, and metastatic tumors led to dilatation of the distal main pancreatic duct and the intrahepatic bile duct system (Figure 1B). Additionally, multiple lymph nodes with high FDG uptake were observed throughout the body, indicating metastasis. An avid enhancing soft tissue density lesion with the size of approximately 5.3 ∗ 4.3 cm was noted incidentally in the left breast, showing an uneven increase in radioactive uptake (SUVmax = 12.9) (Figure 1C). The adjacent skin was diffusively thickened, showing a slightly increased radioactive uptake (SUVmax = 3.3). Radioactive uptake in the right mammary gland is uniform. Multiple lymph nodes with increased radioactive uptake were observed in the left axillary region (Figure 1D). 18F-FDG PET/CT suggested that further pathological examination is warranted for breast lesions with high FDG uptake. The patient received the breast enhanced magnetic resonance imaging (MRI) examination at the same time. The results indicated that the left breast solid mass with skin thickening (BI-RADS grade 5) was 5.3 ∗ 6.0 ∗ 5.3 cm in size, and the left axillary lymph nodes were enlarged with a maximum of 2.1 ∗ 1.1 cm. In addition, there were some enhanced nodules in the right breast (BI-RADS 4A) (Supplementary Figure S1).

Considering a primary breast tumor with multiple metastases, ultrasound-guided core needle biopsy of the left breast mass and lymph nodes was performed in September 2022. The diagnosis of invasive ductal carcinoma of the left breast was confirmed by the pathological evaluation of the mass. To exclude breast metastasis, we compared the results with those of previous surgical pathology and confirmed primary breast cancer. Moreover, an invasive ductal carcinoma metastasis was observed in the left axillary lymph nodes. Finally, the pathological staging was determined to be IIB (cT2N1M0). The pathological results (Figure 2) were as follows: hematoxylin and eosin (H&E) staining revealed a histological pattern of adenocarcinoma. IHC staining showed that both ER and PR were negative, and human epidermal growth factor receptor

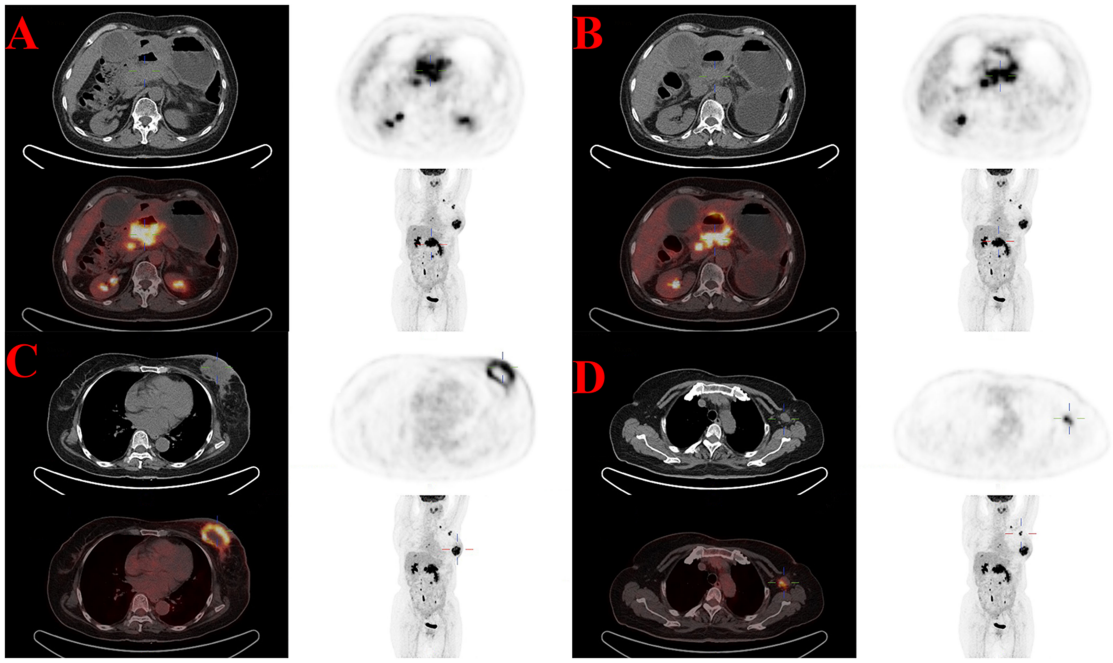


FIGURE 1
High uptake of 18F-FDG in the known lesions with increasing radioactivity uptake of 18F-FDG PET/CT in September 2022. **(A)** The hepatopancreatic lesions (SUVmax = 19.8). **(B)** The dilatation of the distal main pancreatic duct and the intrahepatic bile duct system. **(C)** The breast lesion (SUVmax = 12.9). **(D)** Multiple axillary lymph nodes lesions (SUVmax = 10.4).

2 (HER2) was moderately positive without amplification as detected by FISH. The tumor proliferation marker antigen Ki-67 was 35% positive, and others were tumor marker p53 (diffuse,+), intestinal adenocarcinoma marker cytokeratin7 (CK7,+), ovarian cancer marker Wilms tumor 1 (WT1, −), ovarian clear cell carcinoma marker, and Paired Box 8 (PAX8,−) marker for renal, Müllerian, and thyroid carcinomas. Computed tomography (CT) of the chest, abdomen, and pelvis was performed to rule out metastatic disease. Fortunately, no breast cancer-related metastases were detected. Finally, the patient was diagnosed with MMPMNs, including left breast invasive ductal carcinoma with left axillary lymph node metastasis and high-grade left ovarian serous papillary carcinoma with extensive abdominal and pelvic metastases. For a more detailed evaluation of the left breast lesion, contrast-enhanced breast MRI was performed, which displayed that left breast solid mass (BI-RADS Grade 5) 5.3 ∗ 6.0 ∗ 5.3 cm in size. In addition, multiple

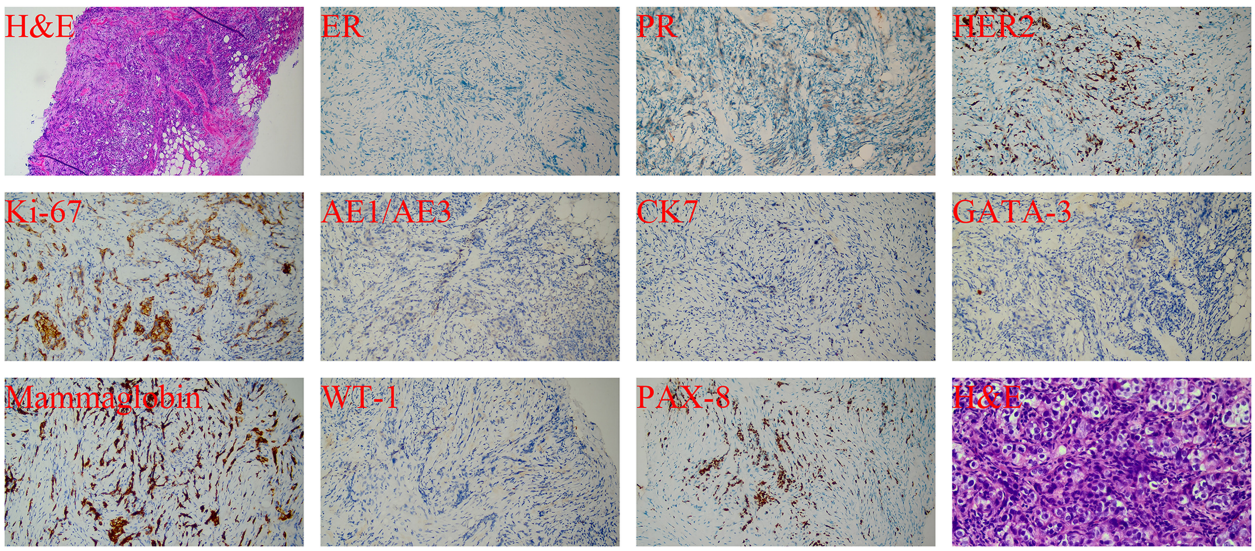


FIGURE 2
H&E(100× &400×) and IHC(200×) staining of the breast tissue in September 2022.

swollen lymph nodes in the left axilla were observed, with a maximum of 2.1 ∗ 1.1 cm indicating metastases.

To further explore a more efficient therapeutic strategy for this MMPMN patient, freshly collected plasma and formalin-fixed, paraffin-embedded (FFPE) primary ovarian and breast tumor tissues were subjected to NGS in October 2022. Targeted NGS of 425 cancer-related genes (Supplementary Table S1) was performed at Nanjing Geneseeq Technology Inc., approved by the College of American Pathologists (CAP) and Clinical Laboratory Improvement Amendments (CLIA). An EGFR p.T790M (c.2369C>T) mutation was revealed at a mutant allele frequency (MAF) of 1.59% in primary breast cancer tissue alone, compared to the primary ovary tissue and the plasma sample, which identified the EGFR T790M mutation as a new tumor-initial driver event in BC. Detailed results of the genetic alterations are shown (Supplementary Table S2). In particular, this patient did not have BRCA1–2 mutations, which, to some extent, ruled out a link between the patient's breast cancer and her previous ovarian cancer.

As such, our doctors communicated fully with the patient and her family, who were adequately informed and agreed to the treatment plan. In October 2022, the therapeutic regimen was planned as chemotherapy (docetaxel 120 mg/m^2 in combination with carboplatin 600 mg/m^2; administered every 21 days for six cycles), followed by osimertinib-targeted therapy, which is a promising, orally available, third-generation mutation-specific EGFR TKI for the treatment of EGFR T790M resistance mutation-positive NSCLC. However, no treatment options for breast cancer patients with the EFGR T790M mutation have been reported in the literature. Unfortunately, the patient died after the first cycle of chemotherapy (November 2022) due to worsening gastrointestinal hemorrhage and hypovolemic shock with disease progression.

3 Discussion

In recent years, identifying genetic mutations in EGFR has resulted in significant changes in the diagnosis and management of cancer (14). Mutations in the EGFR gene regulate cell proliferation and differentiation; hence, it is important to promote the development of cancer (15). EGFR gene is located on chromosome 7p11.2, contains 28 exons, and encodes including a cytoplasmic domain (also called the tyrosine kinase domain), which is responsible for the phosphorylation of its downstream targets and self-regulation (15). EGFR-TKIs are currently the first-line treatment for cancer patients with EGFR-sensitive mutations (16). EGFR mutations are present in solid tumors with a variety of mutation types that may affect the response to EGFR-TKIs. Different mutation types increase the activity of EGFR and activate different downstream signaling pathways (17); this cascade relates RTK activity to increased proliferation, motility, migration, survival, and anti-apoptotic cellular responses and facilitates the genesis and development of cancer (18). The most common alterations are deletions in exon 19 (ex19del, about 44%) and point mutations in exon 21 (L858R, about 41%), known as common mutations or classical mutations, which are considered sensitive to treatment with TKIs (19).

Patients with EGFR-TKIs as first-line treatment have an average progression-free survival (PFS) of 10–12 months (20); however, acquired resistance is inevitable. Secondary resistance mutations in the tyrosine kinase domain develop in up to 60% of patients with NSCLC receiving EGFR-TKIs, most commonly the secondary mutation of resistance p.Thr790Met (T790M), resulting from a gatekeeper mutation in exon20 of EGFR (21). This is referred to as a "gatekeeper" mutation because the 790 residue is in a key location at the entrance to the hydrophobic pocket of the adenosine triphosphate (ATP)-binding cleft (22). The T790M mutation results in a conformational change in the ATP-binding pocket and increases the affinity for ATP in the ATP-binding domain of EGFR. As a result, T790M causes steric hindrance of the binding to their connected ATP binding site on EGFR of an ATP-competitive kinase inhibitor (first- and second-generation TKIs), but irreversible inhibitors (third generation TKIs) overcome this resistance simply through covalent binding (23).

Compared to other selective third-generation mutation-specific EGFR TKIs inhibiting the T790M mutation, osimertinib has shown great superiority and was approved by the US Food and Drug Administration (FDA) as a competitive inhibitor of EGFR T790M-positive NSCLC patients progressing following EGFR-TKI therapy (24). More precisely, osimertinib irreversibly and covalently binds to the cysteine-797 residue in the ATP-binding pocket of EGFR, regardless of the hindering of T790M. Furthermore, because osimertinib creates an irreversible link with the ATP pocket of EGFR, it can overcome the increased affinity of ATP determined by the T790M mutation (23). EGFR T790M mutation in tumors is common and almost entirely presents as a secondary mutation, especially in NSCLC. However, it is extremely rare in breast tumors, and there are limited data supporting the role of EGFR T790M-directed agents in BC. To the best of our knowledge, only three existing BC cases with EGFR T790M positivity have been reported in a 2015 Norwegian clinical study, all of which were primary unilateral breast cancer (UBC) (5).

With the development of precision therapy, the model of targeted therapy guided by comprehensive gene testing is gradually mature, and NGS technology has been widely accepted in the aspects of disease diagnosis, targeted therapy, efficacy evaluation, drug resistance monitoring and other applications (13). Currently, the importance of NGS has been highlighted for discovering rare genetic alterations to guide disease prevention and to improve treatment decision-making and the use of targeted therapy. With the popularity of NGS, if a patient's economic condition permits, clinicians can recommend that the patient accepts NGS as early as possible, especially lung cancer (13–25). It can determine whether the patient has a target for follow-up immunotargeted therapy or immunochemotherapy combined therapy. In addition, at the critical point of disease recurrence or metastasis, NGS can effectively provide a new direction for therapy, thereby bringing new hope to patients (26). In terms of which type of NGS to perform, clinicians should advise the patient to make a discretionary choice based on each patient's own condition and

financial situation (27). In the present case, a pathogenic EGFR T790M mutation was identified in an MMPMN patient's primary breast sample alone by NGS, indicating the driving role of this mutation in BC. The benefit of NGS in identifying novel potential molecular targets for subsequent treatment has been confirmed (28). In addition, NGS can not only provide more detailed information for diagnosis and treatment decision making, but also reveal the efficacy and monitor drug resistance during targeted therapy.

The limitations of the single case presentation in this study should be noted. The patient died because of severe gastrointestinal bleeding and hypovolemic shock before receiving osimertinib. Therefore, the antitumor effect of ositinib could not be reflected in the treatment of this patient; however, our case proposed osimertinib as a reliable treatment option. To date, all studies of BC with the EGFR T790M mutation, including this one, are case reports based on the genomic status of individual patients. The clinical value of the population of breast cancer patients with EGFR T790M mutations should be systematically evaluated in larger cohorts. Nevertheless, additional preclinical studies and clinical evidence are needed to increase our understanding of this area, and multidisciplinary discussions on individualized management are also required (1).

4 Conclusion

In summary, we report a patient with MMPMNs harboring a primary EGFR T790M mutation in BC tissue. This rare case proposes a reliable treatment option for BC with EGFR T790M mutation and highlights the importance of clinical actionability derived from comprehensive genomic profiling results outside standard treatment strategies. The diagnosis and treatment of BC patients with rare genetic mutations remain challenging because of the lack of specific screening and well–established treatment guidelines.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Funding

This work was supported in part by the WU JIEPING Medical Foundation, Grant number(320.6750.2022-18-55 to CS.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1204041/full#supplementary-material

References

1. Wu D, Yu J, Guo L, Wei X, Tian Z, Duan X. Analysis of primary synchronous breast invasive ductal carcinoma and lung adenocarcinoma with next-generation sequencing: A case report. *Oncol Lett* (2023) 25(1):18. doi: 10.3892/ol.2022.13604

2. Moertel CG, Dockerty MB, Baggenstoss AH. Multiple primary malignant neoplasms. I. Introduction and presentation of data. *Cancer-Am Cancer Soc* (1961) 14:221–30. doi: 10.1002/1097-0142(196103/04)14:2<221::aid-cncr2820140202>3.0.co;2-6

3. Mariotto AB, Rowland JH, Ries LA, Scoppa S, Feuer EJ. Multiple cancer prevalence: a growing challenge in long-term survivorship. *Cancer Epidemiol Biomarkers Prev* (2007) 16(3):566–71. doi: 10.1158/1055-9965.EPI-06-0782

4. Aydiner A, Karadeniz A, Uygun K, Tas S, Tas F, Disci R, et al. Multiple primary neoplasms at a single institution: differences between synchronous and metachronous neoplasms. *Am J Clin Oncol* (2000) 23(4):364–70. doi: 10.1097/00000421-200008000-00011

5. Bemanian V, Sauer T, Touma J, Lindstedt AB, Chen Y, Odgard HP, et al. The epidermal growth factor receptor (EGFR / HER-1) gatekeeper mutation T790M is present in European patients with early breast cancer. *PloS One* (2015) 10(8):e134398. doi: 10.1371/journal.pone.0134398

6. Zhang K, Yuan Q. Current mechanism of acquired resistance to epidermal growth factor receptor-tyrosine kinase inhibitors and updated therapy strategies in

human nonsmall cell lung cancer. *J Cancer Res Ther* (2016) 12(Supplement):C131–7. doi: 10.4103/0973-1482.200613

7. Zhang N, Wang D, Li X, Yang Z, Zhang G, Wang Y, et al. A case report of EGFR mutant lung adenocarcinoma that acquired resistance to EGFR-tyrosine kinase inhibitors with T790M mutation and epithelial-to-mesenchymal transition. *Respir Med Case Rep* (2017) 22:183–6. doi: 10.1016/j.rmcr.2017.08.015

8. Blasi M, Kazdal D, Thomas M, Christopoulos P, Kriegsmann M, Brandt R, et al. Combination of crizotinib and osimertinib in T790M+ EGFR-mutant non-small cell lung cancer with emerging MET amplification post-osimertinib progression in a 10-year survivor: A case report. *Case Rep Oncol* (2021) 14(1):477–82. doi: 10.1159/000513904

9. Westover D, Zugazagoitia J, Cho BC, Lovly CM, Paz-Ares L. Mechanisms of acquired resistance to first-and second-generation EGFR tyrosine kinase inhibitors. *Ann Oncol* (2018) 29(suppl_1):i10–9. doi: 10.1093/annonc/mdx703

10. Ma L, Chen R, Wang F, Ma L, Yuan M, Chen R, et al. EGFR L718Q mutation occurs without T790M mutation in a lung adenocarcinoma patient with acquired resistance to osimertinib. *Ann Transl Med* (2019) 7(9):207. doi: 10.21037/atm.2019.04.37

11. Liu X, Adorno-Cruz V, Chang YF, Jia Y, Kawaguchi M, Dashzeveg NK, et al. EGFR inhibition blocks cancer stem cell clustering and lung metastasis of triple negative breast cancer. *Theranostics* (2021) 11(13):6632–43. doi: 10.7150/thno.57706

12. Li K, Zhang TT, Zhao CX, Wang F, Cui B, Yang ZN, et al. Faciogenital Dysplasia 5 supports cancer stem cell traits in basal-like breast cancer by enhancing EGFR stability. *Sci Transl Med* (2021) 13(586). doi: 10.1126/scitranslmed.abb2914

13. Zheng S, Wang X, Fu Y, Li B, Xu J, Wang H, et al. Targeted next-generation sequencing for cancer-associated gene mutation and copy number detection in 206 patients with non-small-cell lung cancer. *Bioengineered* (2021) 12(1):791–802. doi: 10.1080/21655979.2021.1890382

14. Wang X, Chen S, Emerson RE, Wu HH, Cramer HM, Curless K, et al. Molecular testing for EGFR mutations and ALK rearrangements in the cytological specimens from the patients with non-small cell lung cancer. *Appl Immunohistochem Mol Morphol* (2019) 27(2):119–24. doi: 10.1097/PAI.0000000000000701

15. Dolesova L, Konecny M, Markus J, Zavodna K, Bohmer D, Repiska V. Molecular analysis of EGFR gene in different types of tumor material from NSCLC patients. *Neoplasma* (2015) 62(3):439–48. doi: 10.4149/neo_2015_052

16. Schuler M, Paz-Ares L, Sequist LV, Hirsh V, Lee KH, Wu Y, et al. First-line afatinib for advanced EGFRm+ NSCLC: Analysis of long-term responders in the LUX-Lung 3, 6, and 7 trials. *Lung Cancer* (2019) 133:10–9. doi: 10.1016/j.lungcan.2019.04.006

17. Arteaga CL, Engelman JA. ERBB receptors: from oncogene discovery to basic science to mechanism-based cancer therapeutics. *Cancer Cell* (2014) 25(3):282–303. doi: 10.1016/j.ccr.2014.02.025

18. Wu L, Ke L, Zhang Z, Yu J, Meng X. Development of EGFR TKIs and options to manage resistance of third-generation EGFR TKI osimertinib: conventional ways and immune checkpoint inhibitors. *Front Oncol* (2020) 10:602762. doi: 10.3389/fonc.2020.602762

19. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol* (2012) 13(3):239–46. doi: 10.1016/S1470-2045(11)70393-X

20. Choo JR, Tan CS, Soo RA. Treatment of EGFR T790M-positive non-small cell lung cancer. *Target Oncol* (2018) 13(2):141–56. doi: 10.1007/s11523-018-0554-5

21. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* (2013) 19(8):2240–7. doi: 10.1158/1078-0432.CCR-12-2246

22. Yun CH, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci USA* (2008) 105(6):2070–5. doi: 10.1073/pnas.0709662105

23. Bertoli E, De Carlo E, Del CA, Stanzione B, Revelant A, Fassetta K, et al. Acquired resistance to osimertinib in EGFR-mutated non-small cell lung cancer: how do we overcome it? *Int J Mol Sci* (2022) 23(13). doi: 10.3390/ijms23136936

24. Papadimitrakopoulou VA, Mok TS, Han JY, Ahn MJ, Ramalingam SS, Kim SW, et al. Osimertinib versus platinum-pemetrexed for patients with EGFR T790M advanced NSCLC and progression on a prior EGFR-tyrosine kinase inhibitor: AURA3 overall survival analysis. *Ann Oncol* (2020) 31(11):1536–44. doi: 10.1016/j.annonc.2020.08.2100

25. Zhao J, Lin G, Zhuo M, Fan Z, Miao L, Chen L, et al. Next-generation sequencing based mutation profiling reveals heterogeneity of clinical response and resistance to osimertinib. *Lung Cancer* (2020) 141:114–8. doi: 10.1016/j.lungcan.2019.10.021

26. Halima A, Vuong W, Chan TA. Next-generation sequencing: unraveling genetic mechanisms that shape cancer immunotherapy efficacy. *J Clin Invest* (2022) 132(12). doi: 10.1172/JCI154945

27. Servetto A, Napolitano F, De Angelis C, Placido P, Giuliano M, Arpino G, et al. A review of the use of next generation sequencing methodologies to identify biomarkers of resistance to CDK4/6 inhibitors in ER+/HER2- breast cancer. *Crit Rev Oncol Hematol* (2021) 157:103191. doi: 10.1016/j.critrevonc.2020.103191

28. Wu T, Wan J, Xia K, Yang M, Feng L, Yin L, et al. Case report: next-generation sequencing-based detection in A patient with three synchronous primary tumors. *Front Oncol* (2022) 12:910264. doi: 10.3389/fonc.2022.910264

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 25 July 2023
DOI 10.3389/fonc.2023.1148861

OPEN ACCESS

EDITED BY
Iman Mamdouh Talaat,
University of Sharjah, United Arab Emirates

REVIEWED BY
Sangseon Lee,
Seoul National University,
Republic of Korea
Harikrishna Reddy Rallabandi,
Oklahoma Medical Research Foundation,
United States

*CORRESPONDENCE
Enrique Hernández-Lemus
ehernandez@inmegen.gob.mx

RECEIVED 20 January 2023
ACCEPTED 05 July 2023
PUBLISHED 25 July 2023

CITATION
Ochoa S and Hernández-Lemus E (2023) Molecular mechanisms of multi-omic regulation in breast cancer.
Front. Oncol. 13:1148861.
doi: 10.3389/fonc.2023.1148861

Molecular mechanisms of multi-omic regulation in breast cancer

Soledad Ochoa[1,2] and Enrique Hernández-Lemus[1,3]*

[1]Computational Genomics Division, National Institute of Genomic Medicine, Mexico City, Mexico,
[2]Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, Los Angeles, CA, United States,
[3]Center for Complexity Sciences, Universidad Nacional Autónoma de México, Mexico City, Mexico

Breast cancer is a complex disease that is influenced by the concurrent influence of multiple genetic and environmental factors. Recent advances in genomics and other high throughput biomolecular techniques (-omics) have provided numerous insights into the molecular mechanisms underlying breast cancer development and progression. A number of these mechanisms involve multiple layers of regulation. In this review, we summarize the current knowledge on the role of multiple omics in the regulation of breast cancer, including the effects of DNA methylation, non-coding RNA, and other epigenomic changes. We comment on how integrating such diverse mechanisms is envisioned as key to a more comprehensive understanding of breast carcinogenesis and cancer biology with relevance to prognostics, diagnostics and therapeutics. We also discuss the potential clinical implications of these findings and highlight areas for future research. Overall, our understanding of the molecular mechanisms of multi-omic regulation in breast cancer is rapidly increasing and has the potential to inform the development of novel therapeutic approaches for this disease.

KEYWORDS

multiomics, breast cancer, DNA methylation, epigenomic regulation, network biology

1 Molecular origins of breast cancer

Molecular heterogeneity is one of the archetypal features of breast cancer. This heterogeneity refers to the fact that breast tumors are composed of a mixture of cancer cells with different genetic and molecular characteristics. This diversity of features and origins *within a single tumor* can contribute to differences in tumor behavior, such as response to treatment and risk of recurrence. Research has identified several molecular subtypes of breast cancer, including estrogen receptor-positive, HER2-enriched, and triple-negative, each with its own unique set of genetic and molecular features. Additionally, within a given subtype, there can be further molecular heterogeneity, with different cancer cells possessing varying combinations of genetic and epigenetic alterations. This molecular heterogeneity can be driven by a variety of factors, such as inherited genetics, acquired mutations, and environmental exposures. Understanding the molecular diversity of breast cancer is important for developing personalized treatment approaches and for predicting patient-specific outcomes. However, the complexity of this heterogeneity presents a challenge for researchers studying the disease and for clinicians caring for breast cancer

patients. Here we will present an overview of some of the molecular (mostly genomic and epigenomic) factors behind, and will discuss some of the already synergistic mechanisms giving rise to these complex pathophenotypes.

Aside (but not independently) from the genomic and epigenomic background, one additional source of complexity and heterogeneity in breast tumors is the influence of metabolic reprogramming in general, and hormone signaling in particular. Estrogen, for instance promotes cell proliferation, and on the other, it is oxidized to reactive products that damage DNA (1). Exposure to estrogen is linked to the menstrual lifetime, with a higher risk for women with early menarche and late menopause, but also to prolonged use of contraceptives and obesity. Before menopause, most of the estrogen in the body comes from the ovaries and a small percentage from fat tissue, but after menopause, the main source of estrogen is fat tissue, and the more there is, the greater the risk of breast cancer. In addition, being overweight causes a higher level of insulin in the blood, which has also been associated with breast cancer (2).

Two related factors that reduce the risk are the age of the first birth and breastfeeding. Experiments in mice show that pregnancy causes the differentiation of mammary lobules into secretory units, with lower proliferative activity, which would reduce the subset of cells susceptible to carcinogenesis. The risk reduction from breastfeeding is independent of childbirth and menopausal status, without a strong functional explanation, but several hypotheses, which include: interruption of ovulatory cycles, lower estrogen production and terminal tissue differentiation (1).

Despite the variation between countries and stages, breast cancer has a relatively good recovery rate compared to other types of cancer. It is estimated that up to 15% of patients develop distant metastases, which are mostly detected in bones, liver, lung and brain, with an association between the site of metastasis and the subtype of breast cancer (2, 3). In this regard, patterns have been identified that allow tumors to be grouped in different ways (4–6), which affect the prognosis and treatment of the disease, as we will discuss in the rest of this review.

1.1 Ductal and lobular origins of breast cancer

Breast cancer is a possess complex histological origins, as it is able to develop from different types of cells within the breast. Two common origins are the ductal and lobular tissues. Ductal carcinoma starts in the cells lining the milk ducts. These are thin tubes able to carry milk from the lobules to the nipple. Ductal carcinoma constitutes the most common type of breast cancer, accounting for about 80% of all cases. Typically appears as a lump in the breast that can spread to nearby lymph nodes if left untreated. Lobular carcinoma, on the other hand, originates in the lobules, the milk-producing glands within the breast. Lobular carcinomas are about 15% of breast cancer cases. Unlike ductal carcinoma, lobular tumors may not form a distinct lump. Instead, it often appear as a subtle thickening in the breast tissue. Lobular carcinomas are also able to spread to other parts of the body, including the opposite breast, ovaries, and abdomen (See Figure 1).

Ductal and lobular breast tumors show a number of genetic and molecular differences. These can be summarized as follows:

1. Genetic alterations: Ductal carcinoma characteristically display a higher frequency of genetic alterations in the tumor suppressor gene TP53. As is known, TP53 mutations are associated with more aggressive tumor behavior, hence poorer prognosis. Lobular tumors, in contrast often show alterations in the E-cadherin gene (CDH1) involved in cell adhesion. Mutations in CDH1 can lead to the loss of cell adhesion, a hallmark of lobular carcinoma.

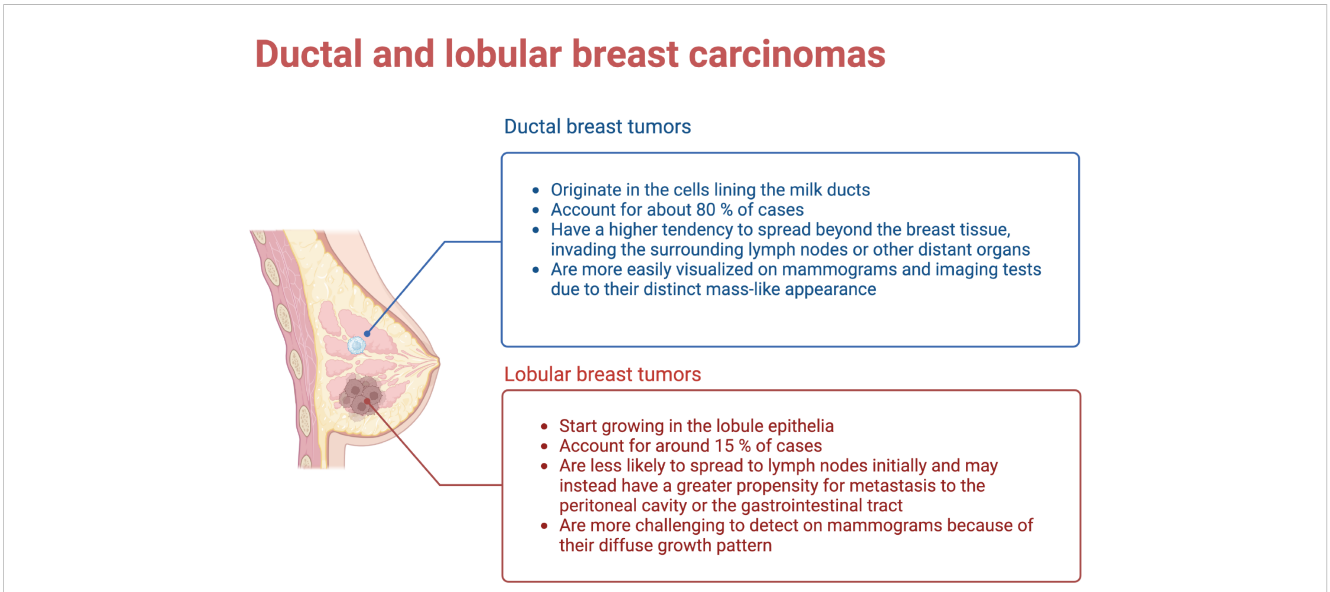


FIGURE 1
Ductal and Lobular breast tumors have different origins, development and outcomes. (Figure created using Biorender.com).

2. Molecular subtypes: Breast cancer can be classified into different molecular subtypes based on gene expression patterns (see the next subsection). Ductal carcinomas are more commonly associated with the basal-like subtype, characterized by aggressive behavior and a higher risk of recurrence. Lobular carcinoma, are instead often classified within the luminal subtypes, which are typically less aggressive and associated with hormone receptor-positive tumors.
3. Hormone receptor status: Hormone receptor status, including estrogen receptor (ER) and progesterone receptor (PR) expression, differs between ductal and lobular breast tumors. Ductal carcinoma tends to have a higher frequency of hormone receptor-positive tumors, meaning they respond to hormonal therapies targeting these receptors. Lobular carcinoma is also hormone receptor-positive in many cases, but it has a higher tendency to have loss or reduced expression of hormone receptors compared to ductal carcinoma.
4. Metastatic patterns: Ductal and lobular carcinomas are also different in their patterns of metastasis. Ductal carcinoma often spreads to the lymph nodes and distant organs such as the lungs, liver, and bones. Lobular carcinoma instead has a higher propensity for multi-focal and multi-centric growth within the breast and has a greater tendency to metastasize to the peritoneal cavity, ovaries, and gastrointestinal tract.
5. Cellular morphology: Ductal carcinoma is characterized by the formation of irregular glandular structures, while lobular carcinoma often shows a characteristic single-file pattern, where the tumor cells infiltrate the breast tissue in a linear fashion without forming distinct masses.

We should note that individual breast tumors may indeed present a combination of features from both the ductal and lobular cellular origins. Moreover, as we shall see in the next subsection, advancements in genomic and molecular profiling techniques are allowing us to uncover additional subtypes and molecular features that contribute to further refine our understanding of breast cancer heterogeneity.

1.2 Breast cancer molecular subtypes

Most breast tumors affect the epithelium of the mammary glands, a mesh of branching ducts, which extend radially from the nipple and end in lobules (7). Therefore, histologically, they are carcinomas, which can be further classified as ductal or lobular and be invasive or presented in situ. The preservation of gene expression patterns indicates that invasive carcinomas often arise from *in situ* lesions (8). Less than 1% of breast tumors are sarcomas, which develop from the stroma of the glands, including blood vessels and myofibroblasts (2).

In connection with hormone and other signaling pathways, estrogen receptors (ER), progesterone receptors (PR), and human epidermal growth factor receptor 2 (HER2) have been used as immunohistochemical markers for clinical classification (5). The presence of estrogen receptors in up to 1% of tumor cells indicates a tumor that is (commonly) responsive to hormonal therapy (9), well-differentiated, and less aggressive. Tumors that are positive for HER2 may respond to treatment with monoclonal antibodies and kinase inhibitors, but the prognosis depends on the status of other receptors, among other issues. Tumors that are triple-positive usually have a good prognosis, while those with the ER-PR-HER2- phenotype are more aggressive and poorly differentiated. As a result, tumors without these receptors do not have targeted therapies and are generally of worst prognosis (10).

The relevance of the receptors has a biological reason, as estrogen stimulates the proliferation of cells with the receptor and induces the progesterone receptor - a mitogenic hormone - making PR+ tumors commonly also ER+ (10). On the other hand, the binding of the growth factor causes the heterodimerization of Her2 and the activation of its intracellular domain, which then participates in multiple transduction pathways, such as MAPK and PI3K (1).

To the previous sub-divisions, we must add the classification by means of gene expression, also called *molecular subtypes*. The classification by gene expression comes from the transcriptional patterns shared between different samples of the same tumor, which identify the intrinsic subtypes: luminal A, luminal B, Her2-enriched, and basal (4). Originally, a subtype (Normal-like) similar to normal tissue was also identified, but the possibility that it was instead formed due to contamination by adjacent normal tissue has kept the existence of this subtype in controversy (11).

Although different molecular classifiers have been used, such as Mammaprint and BluePrint, and even the subtypification with immunohistochemical markers of proliferation and the aforementioned receptors (12) has been approximated, the use of the PAM50 classifier (Prediction Analysis of Microarray 50), predominates in the databases and genomic studies. This is based on an array that measures the expression of the 50 genes that best separate the intrinsic subtypes (13) and provides highly predictive information on recurrence and neoadjuvant response (11).

The improvement of high-performance techniques enriched the description of breast cancer subtypes, allowing the transition from a grouping of transcriptional signatures to sub-subtypes with their own multi-omic characteristics (see Figures 2, 3). In this way, the luminal subtypes can be clearly separated, as both are usually positive for hormonal receptors and negative for HER2; however, luminal B tumors have higher expression of genes associated with cell proliferation and lower expression of luminal tissue-related genes such as PR. A subset of luminal B tumors is characterized by hypermethylation of the Wnt pathway (5). Luminal A tumors have the lowest number of mutations but an increase in those affecting PIK3CA and MAP3K1 genes compared to the luminal B subtype. Interestingly, both subtypes have a good prognosis and high frequencies of around 30% of cases each, but luminal B tumors show greater chemosensitivity and the highest risk of recurrence in 10 years regardless of therapy. Therefore, this subtype has been proposed as the one to study, above others with a worse prognosis, to reduce mortality from breast cancer (13).

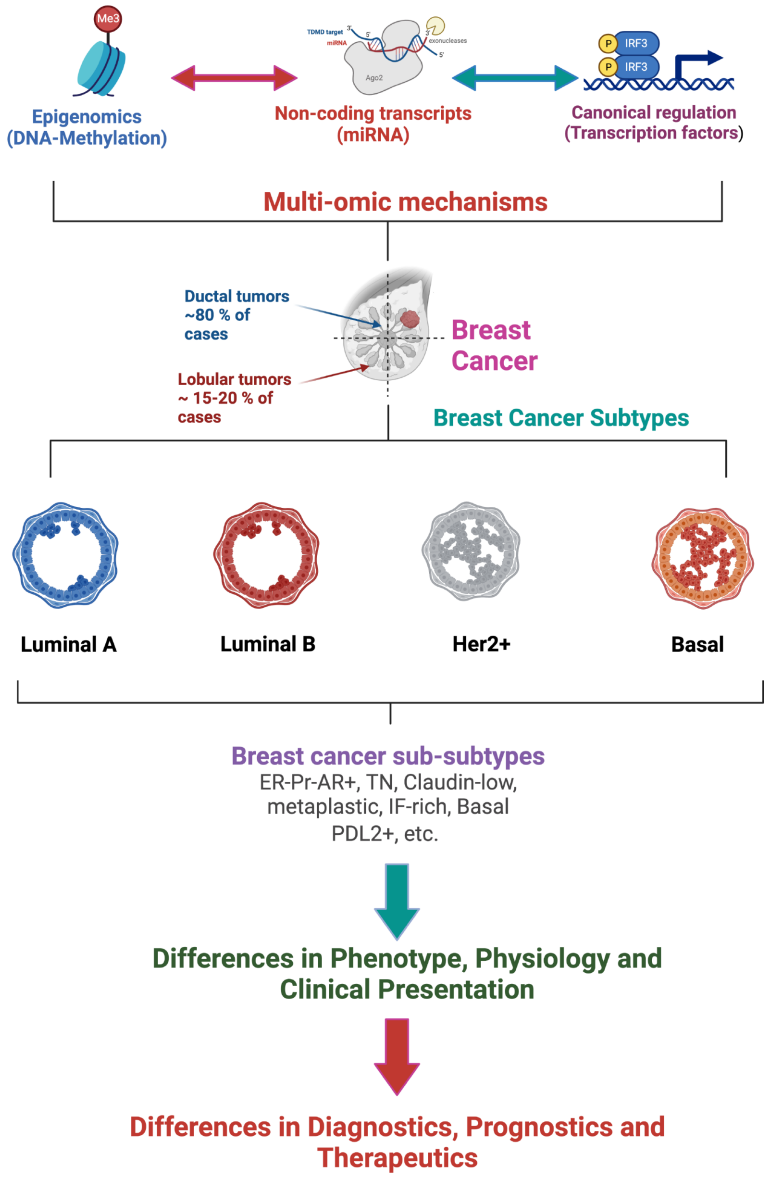


FIGURE 2
Multi-omic gene regulatory mechanisms influence breast cancer phenotypes affecting the classification, diagnostics, prognostics and therapeutics of breast tumors. (Figure created using Biorender.com).

Tumors of the Her2-enriched (HER2E) subtype are characterized by the overexpression of HER2 and nearby genes such as GRB7, both at the transcriptional and protein levels, and by presenting the highest number of mutations in general and on the APOBEC3B cytidine deaminase gene in particular (13). HER2 overexpression has been associated with amplification of the long arm of chromosome 17, which nearly-always contains the receptor but whose extent varies. However, this subtype remains somehow controversial, as nearly half of the tumors with amplification are classified as luminal, mostly, or even basal (5).

Furthermore, differential expression studies between tumors with and without amplification identify few genes outside of chromosome 17, with modest changes. In contrast, when comparing HER2E tumors against non-HER2E tumors, the androgen receptor (AR) and different ER targets stand out, which could be explained by the metabolic and molecular redundancy between ER and AR. Adding the cooperation between HER2 and AR, as well as the inverse relationship between HER2 expression and ER/PR (10), it has been speculated that amplification could be a driver event that masks the hormonal nature of the subtype, as mostly apocrine (ER-PR-AR+) (14).

Basal tumors overexpress genes associated with cell proliferation and breast basal tissue, are characteristically hypomethylated, have the highest frequency of alterations on TP53 and are associated with the inactivation of BRCA1. When compared to different types of cancer, this subtype is molecularly more similar to squamous cell lung cancer than to luminal breast subtypes, while its pattern of mutations brings it closer to serous ovarian tumors. Although basal tumors would correspond to the triple negative (TN) phenotype of immunohistochemical markers,

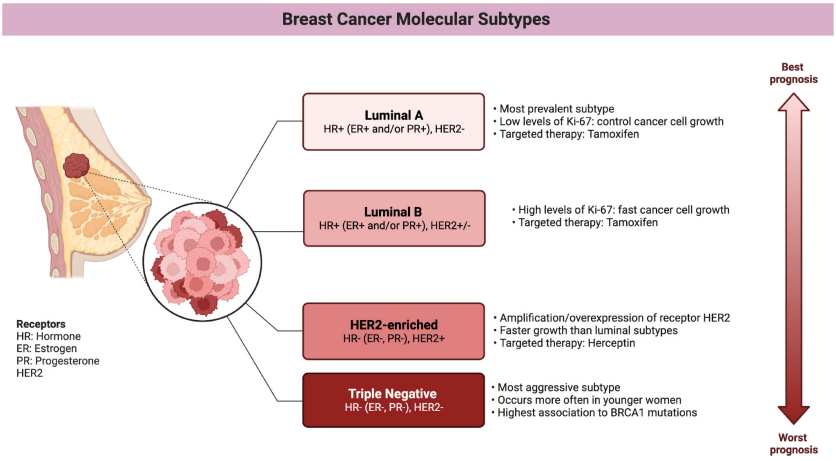


FIGURE 3
Main features of breast cancer molecular subtypes. (Figure created using Biorender.com).

only 75% of TN tumors have the basal subtype expression pattern (5). This expression pattern is associated with aggressive tumors that present at early ages, with greater susceptibility in African ancestry populations and the worst prognosis at 5 years (13). The correspondence with the TN phenotype implies that there are no targeted treatments, however, the use of PARP inhibitors in tumors with BRCA1 mutations has recently been approved.

Molecularly, the basal subtype can be further divided, although there is still no consensus on how many and which those sub-subtypes would be, the subtypes *claudin low*, *metaplastic* and *interferon rich* have been consistently mentioned though (10). ATAC-seq studies identify categories similar to basal, mesenchymal and ligated to the luminal androgen receptor. Each category has its own mutations and clinical characteristics, of which the higher age of diagnosis for tumors ligated to the luminal androgen receptor and the activation without amplification of HER2 stand out. For their part, basal tumors are further separated into two groups, BL1 and BL2, according to the risk of progression. Those classified as BL2 are associated with an intact G1/S checkpoint, while those identified as BL1 lose copies of RB, which affects protein expression. Finally, mesenchymal tumors are characterized by a high percentage of mutations on epigenetic modifiers and DNA repair genes, as well as frequent deletion of beta-2-microglobulin, which suggests a reduced antigen presentation. Mesenchymal breast tumors also exhibit DNA hypomethylation, which coincides with greater chromatin accessibility over various enhancers (15).

The classification of basal tumors is particularly interesting because recently differences in the immune response of each subtype have been observed. For a long time, breast cancer was considered as poorly immunogenic given its relatively low mutational burden. However, a high survival rate has been observed among patients of the basal subtype with PDL2 overexpression, suggesting that a subgroup of breast cancer patients could indeed benefit from immune therapy (16).

When characterizing the basal subtype microenvironment, three groups were defined by 1) the inability to attract innate immune system cells, 2) chemotaxis followed by inactivation of innate immunity, and 3) increased immune inhibitory factors.

The phenotype of the first group has been explained by the amplification of MYC, which induces the expression of various chemokines and PDL1, as well as the inactivation of dendritic cells and macrophages, limiting the recruitment of adaptive cells. The second phenotype would be justified by the high infiltration of cancer-associated fibroblasts (CAF) (17), which correlates negatively with the infiltration of T cells (16) and depends on the immunomodulator TGFB; in addition to the inhibitory effect that the frequent mutation of the PI3K-AKT pathway would allow (17). Because of the characteristics of the third phenotype, this would be the subgroup of patients that could benefit more directly from immune therapy

After considering the enormous differences between the subtypes and sub-subtypes, it has been postulated that cells of different origin are involved (15). In principle, the luminal-basal division reflects the normal epithelium of the mammary gland, which is composed of a two-layer of luminal cells that produce milk and basal cells that expel the milk (8). Thus, the basal or myoepithelial layer is composed of contractile cells that express KRT14, TP63, ACTA2/SMA, MME/CD10, and THY1/CD90; while the luminal layer is formed by cells that are able to respond to hormones and express EpCAM, KRT8, KRT18, and MUC1 in addition to receptors. However, the luminal layer can be further divided into luminal cells and luminal progenitor cells. While luminal cells are clearly distinguished by their ER and PR receptor status, luminal progenitor cells almost completely lack these receptors and instead express KRT5/6, a marker of the basal layer in many types of epithelium. Different genetic expression characteristics and chromatin structure suggest that luminal progenitors may be actually intermediate cells between basal and luminal cells (7).

When examining the growth capabilities of each cell type, it was observed that all three types can generate colonies, but only about 0.1% of the basal fraction can produce two-layered structures similar to the mammary gland when injected into mice and,

when appropriately stimulated, produce milk. Luminal progenitors only produce cells with luminal characteristics, with very short telomeres even when using samples from young women, and high levels of reactive oxygen species (7). In this way, according to the stem cell carcinogenesis model, poorly differentiated ER- tumors would arise from the most primitive cells - from the basal fraction -; Her2-enriched and luminal B tumors, which have been described as basoluminal, would come from an intermediate stem cell - the luminal progenitors -; finally, it is predicted that luminal A tumors would originate from the transformation of ER+ stem cells (8). Considering the limited division of luminal cells (7), the *clonal evolution carcinogenesis model* might be more appropriate for addressing the origin of luminal A tumors, as it proposes a population of genetically unstable cells that gain fitness by accumulating mutations and selection (8, 18). The origin of tumor cells is relevant because treatments typically eliminate proliferating cells, eliminating most of the tumor but often ignoring quiescent cells such as stem cells (8).

Beyond the origin of each subtype, it is clear that these are molecularly distinct entities and that these differences can affect their clinical behavior. Although this work delves into the transcriptional description, differences between subtypes can be observed at many other levels such as the rate of cis and trans interactions of the co-expression network (19) and the activation of metabolic pathways (20). Although it is not expected that intrinsic subtypes will replace immunohistochemical tests in the clinic soon, given the dependence on receptors for the assignment of treatments, nor should the tumor heterogeneity in these broad groups be oversimplified (14); molecular classification has been established as the unit of description of breast cancer and will be used throughout this work (See Figure 2).

2 Anomalous gene regulation in breast cancer

As data has been collected, the study of cancer has surpassed the reductionist approach that considered it simply *a disease of genes* (21). Thus, it has been considered a *disease of gene deregulation* (22, 23), *a disease of cellular processes* (24), a *disease of pathways* (25) and, when the origin of deregulation is considered, *a multi-scale disease*, where subcellular alterations affect the tissue, at the same time that the properties of the tissue -irrigation-, affect the phenotype and eventually the cellular genotype (26). In other words, a systems biology approach has been adopted, where interactions matter, whether they occur between genes or between scales. After all, it is not isolated genes that perform functions, but sets of proteins that have undergone regulated processes of transcription and translation and that need signals to enter into action or stop doing so.

The issue is that the regulation or deregulation of genes is already a multiscale problem, which at least involves regulatory sequences, transcription factors (TFs), histones, DNA methylation, non-coding RNAs, and chromatin conformation (27). The mentioned regulatory mechanisms can be organized into different categories, such as epigenetic, transcriptional, and post-transcriptional (as we will do in the rest of this review), but in reality they are interdependent and their simultaneous presence can be indeed identified in the same sample (28–30).

2.1 Epigenetic level: DNA methylation + transcriptomics

Epigenetic regulation involves modifications to chromatin that affect the binding of transcription factors. DNA methylation, specifically the addition of a methyl group to cytosine (5mC), is a well-studied mechanism in this process. Methylation primarily occurs in CpG dinucleotides, which are concentrated in CpG islands (CGIs) found in human genome promoters. Detection of DNA methylation can be done using sodium bisulfite treatment, sequencing, microarrays, methylation-sensitive restriction enzymes, or immunoprecipitation with antibodies against 5mC (22, 31–33). Microarrays, like the Illumina HumanMethylation450K BeadChip (HM450), have been widely used to characterize the methylome due to their cost-effectiveness and accuracy. Sequence-based methylation analysis such as the one carried out by sequencing bisulfite-converted DNA is a more comprehensive technique, able in principle to measure methylation at practically every cytosine in the genome (34). The method methods relies on bisulfite conversion of DNA to detect unmethylated cytosines. Bisulfite conversion changes unmethylated cytosines to uracil during library preparation. Converted bases are identified (following PCR) as thymine in the sequencing data, and sequencing reads are used to determine the fraction of methylated cytosines (35).

Methylation patterns generally correlate with CpG frequency, but CpG islands exhibit unique characteristics and play a role in transcriptional regulation (36–38). CGI promoters have distinct features and differ from other promoters in terms of transcription start regions, bidirectional transcription, and transcription factor binding sites (22, 39, 40).

DNA methylation plays a role in long-term genetic expression programming and cell type determination. After fertilization, the genome undergoes generalized de-methylation, followed by the establishment of *permanent* methylation patterns during embryogenesis (32, 41). *De novo* methylation occurs in early embryonic pluripotent cells, while maintenance methylation takes place during cell division, maintaining methylation patterns from the parental strand to the daughter strand. DNMT enzymes and S-adenosyl L-methionine are involved in DNA methylation, linking gene expression regulation to metabolism. To remove methylation, both passive and active mechanisms are proposed. The passive mechanism suggests that methylation is lost as cells divide, while the active mechanism involves TET enzymes. These mechanisms are associated with changes before implantation, with the maternal genome undergoing passive dilution of methylation and the paternal genome being influenced by Tet3. The gradual loss of DNA methylation observed with aging, particularly in monozygotic twins, may be attributed to the passive mechanism (42, 43).

The TET (Ten eleven translocation) protein family is a group of DNA hydroxylases responsible for oxidizing the methyl group of cytosine and its derivatives successively. The action of TET1, TET2

and TET3 catalyzes the conversion of 5-methylcytosine to 5-hydroxymethylcytosine (5hmC), which is converted to 5-formylcytosine (5fC), which in turn is oxidized to 5-carboxylcytosine (5caC). The 5fC and 5caC forms can be replaced by cytosines by the action of DNA glycosylase and base excision repair. The three derivatives are found simultaneously on the DNA, but cannot be specified by bisulfite sequencing, since 5hmC is read as 5mC, while 5fC and 5caC as cytosine. The identification of each form is relevant because, unlike 5mC, the derivatives do not allow the efficient binding of transcriptional regulators; but 5fC and 5caC favor the binding of proteins involved in DNA repair (44).

The binding of transcriptional regulators to methylated cytosines depends on proteins with MBD (methyl-CpG binding domain) domains, such as MeCP2, which also recruit histone deacetylases and methyltransferases and then reconfigure chromatin to its inactive form (41). Many TFs can bind to both methylated and unmethylated DNA, but with different affinities (45), such is the case of MYC, which binds to the CACGTG motif unless the central CpG has been methylated. Unlike MYC, methylation improves the binding of other transcription factors such as CEBPA and CEBPB (46).

The relationship between DNA methylation and transcription is complex and varied. While methylation of promoters generally inhibits transcription by blocking transcription factor binding, methylation of gene bodies can promote gene expression by facilitating transcriptional elongation (22, 31, 32, 41). However, there are diverse interactions between transcription and DNA methylation, including protection against methylation, promotion of methylation, and demethylation (46). Certain proteins, such as CFP1 and TET proteins, protect promoters from methylation by binding to non-methylated CpG sites. These proteins recruit methyltransferases or reverse *de novo* methylation. DNA-RNA loops resulting from active transcription have also been suggested to protect nearby promoters from methylation. On the other hand, transcriptionassociated proteins can promote DNA methylation by recruiting DNMTs. Examples include DNMT3B, MYC, and E2F6. The KRAB-ZNF family of transcription factors, characterized by an RH motif, can facilitate targeted methylation by interacting with DNMTs. Demethylation, on the other hand, involves the recruitment of TET proteins. Transcription factors like SPI1 and co-activators like PPARG can interact with TET proteins to induce demethylation or the conversion of 5mC to 5hmC in specific regions.

2.1.1 Methylation and cancer

Considering the importance of DNA methylation on defining cell type through transcriptional regulation, it is understandable its alteration in syndromes and diseases. Prader-Willi, Angelman, Beckwith-Wiedemann and Silver-Russell syndromes have been mapped to chromosomal aberrations, but also to imprinting defects due to altered methylation of the involved genes: UPD, ICR2, and ICR1 (42). In cancer, levels of DNMTs expression have been reported similar to those observed in embryos, while TET enzymes mutation has been recurrently identified in different liquid tumors (47). The alterations in DNA methylation described in cancer are not limited to specific point mutations or epimutations – Epimutations are changes in the epigenome relative to consensus, equivalent to mutations (42), but reversible and more frequent (36), but include simultaneous hypermethylation and hypomethylation of multiple regions of the genome (41).

The hypermethylation of DNA in cancer affects 5-10% of CGI promoters - which are normally not methylated - and has been associated with the silencing of tumor suppressor genes (TSGs) (22), responsible, for example, for inducing apoptosis and cell arrest. In addition to epigenetic silencing, tumor suppressors often suffer disruptive mutations such as indels and stop codon substitutions in both alleles, as according to Knudson's two-hit hypothesis, both copies of the gene must be inoperable for TSG inactivation (21). Promoter hypermethylation is usually the second impact of these genes and is thought to progress gradually, from the surrounding heterochromatin, to the transcription start site, subtly and heterogeneously reducing gene expression and favoring tumor plasticity (31). Thus, even the methylation coasts are differentially methylated in cancer. The number of affected CGIs also gradually increases as cell differentiation decreases (41).

Around half of the genes that cause familial forms of cancer can be found hypermethylated in sporadic tumors. In the case of breast cancer, 10-15% of women with sporadic tumors exhibit BRCA1 TSG hypermethylation, accompanied by an expression pattern consistent with hereditary tumors (31). Apart from tumor suppressors as such, hypermethylation causes harmful silencing of miRNAs and more complex deregulation, such as interference with ER-ERE binding (48) and loss of IGF2 imprinting. The expression of IGF2, involved in Beckwith-Wiedemann and Silver-Russell syndromes, is normally inhibited by the insulator H19, which prevents the action of a distal enhancer on the IGF2 promoter; however, in various types of cancer, H19 has been found to be hypermethylated, allowing the expression of the maternal IGF2 copy and causing excess growth factor. As with this, there are many examples of hypermethylation, to the point that filtering strategies are needed to identify their functional consequences (22).

Equivalently, hypomethylation causes the percentage of methylated CpG sites in the genome to drop from 80 to 60 or even 40% and progresses such that metastases have lower levels of methylation than primary tumors (49). The methylator phenotype identified in a subgroup of tumors is characterized by the coordinated methylation of a large number of CGIs, and has a low risk of metastasis and better survival rates. Taking advantage of these observations, agents have been found that reverse de-methylation, inhibiting the invasiveness and metastasis of breast cancer cell lines (41).

Unlike hypermethylation, hypomethylation does not occur in a focused manner on CGI promoters, but rather on a large scale, affecting repetitive elements that include transposons and oncogenes and mapping to late-replication regions associated with the nuclear lamina. Transcriptional activation of the repetitions predisposes the genome to recombination, as evidenced by the increase in the frequency of chromosomal alterations in cancer. Transposons are kept under control in basal tumors, due to compensation for the loss of methylation by trimethylation of lysine 27 of histone 3 (48). While hypomethylation promotes indels and translocations, methylation

alone increases the susceptibility of cytokines to mutagenesis, because it increases the hydrolytic deamination rate, which, due to the methyl group, converts the base into thymidine instead of uracil, as corresponds to cytosines, preventing efficient repair of damage (22).

Although consistently an excess of variability in methylation levels has been found in breast cancer compared to normal tissue (49), specific patterns are known, at least for the basal, luminal B and Her2-enriched subtypes. The basal subtype is the most hypomethylated and, as expected, also has a high genomic instability. Among luminal B samples, a hypermethylated subgroup has been recognized, where the affected CpGs are linked to the Wnt pathway (5). On the Her2-enriched subtype, a bias towards hypermethylation - over hypomethylation - compared to normal tissue has been reported, which is associated with Her2 amplification and particularly affects Hox genes (50).

Although regulation by methylation acts locally on genes, coordinated methylation between distant loci can reflect the same transcriptional program. In that sense, it has been reported that more than half of the pairs of highly co-methylated genes - with Pearson correlation coefficients above 0.75 - in breast cancer are on different chromosomes and tend to participate in similar functions, with enrichment in the pathways of adult onset diabetes, hematopoietic lineage, long-term depression and interaction between receptors and the extracellular matrix (40). Saving the differences between studies, a pan-cancer analysis, which includes breast cancer, reports tissue variability, but identifies 4 groups of genes consistently co-methylated, two of which allow discriminating between cancer and normal tissue samples, despite containing only six cancer-associated genes: CSF2, GALR1, IRF4, PTPRT, SOX11y NRG1 (51). However, the levels of methylation and co-methylation do not necessarily imply a functional change in the cell, there are more regulatory mechanisms at play and it is estimated that only 15% of differentially methylated genes also exhibit a change in expression (50).

Estrogen receptor-α (ER) drives tumor development in ER-positive (ER+) breast cancer. The transcription factor GATA3 has been closely linked to ER function. Epigenetic changes in GATA3 function may thus be relevant to breast cancer biology. It has been recently discussed how indirect changes in the activity of the transcription factor GATA3 by TET2 knockdown lead to epigenetic changes by significantly reducing 5-hydroxymethylcytosine (5hmC) levels without similar changes in methylated cytosine (5mC) (52). These changes are able to lead to global transcriptional deregulation (see Figure 4).

Other specific findings are opening new avenues of research in breast cancer biology. Such is the case of the recent discovery that overexpression of MAGI2-AS3 diminishes DNA methylation of MAGI2 in breast cancer cells (MCF-7) and thus would inhibit the Wnt/β-catenin pathway also diminishing cell proliferation and migration (53); the authors reported that MAGI2-AS3 may act as a cis-acting regulatory element down regulating the DNA methylation level of the MAGI2 promoter region.

In addition to providing information about the origin of tumors and potentially active genes, DNA methylation has gained clinical interest as a prognostic marker. DNA is a relatively resistant material that can be manipulated more easily than the RNA necessary to measure genetic expression (41) and that can be recovered from different bodily fluids depending on the type of cancer. As the tumor cells die, free DNA is released into the bloodstream, where it can be detected with high sensitivity (31). For example, from the levels of methylation in serum of women with metastatic breast cancer, a subgroup with higher disease-free survival could be distinguished, now recognizable by the methylation of SFN, hMLH1, HOXD13, PCDHGB7, RASSF1 and P16 (38). The hypermethylation of estrogen response elements is used to predict reduced response to endocrine therapy, with the methylation of PSAT1 as a specific indicator of response to tamoxifen. There are also numerous studies exploring the early detection of cancer using tests that measure DNA methylation. Both

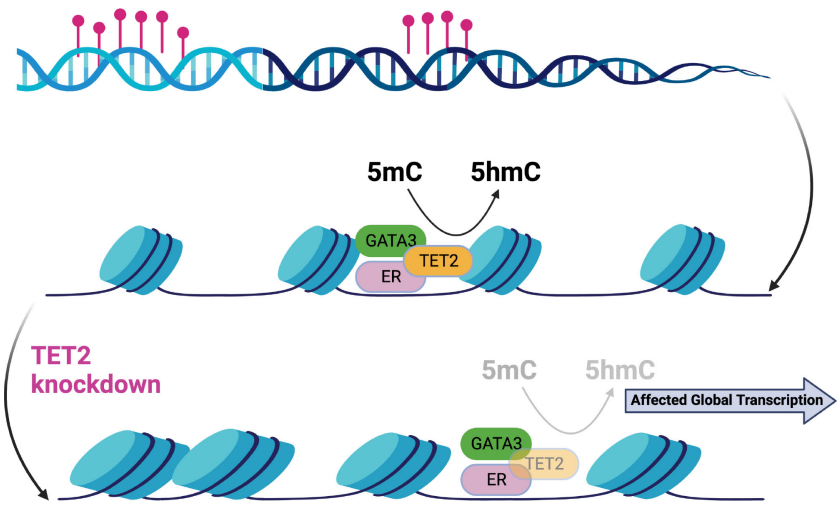


FIGURE 4
Epigenetic changes in the ER-complex may lead to global transcriptional deregulation in breast cancer [Figure created using Biorender.com, adapted from (52)].

their sensitivity and specificity exceed those reported for mammography screening and are even higher for advanced stages (48).

The other potential usefulness of DNA methylation is in the treatment of cancer. The use of DNMT inhibitors as a sensitization strategy to other treatments is promising for breast cancer, although it has not yet been approved for routine clinical use. The DNMT inhibitors decitabine and 5-azacitidine are used in the management of hematological malignancies and can inhibit tumor growth in ER + breast cancer models in combination with chemotherapy or immunotherapy. It is believed that these inhibitors activate the immune response by stopping the silencing of tumor antigens. What has been demonstrated is an increase in the expression of the immunomodulator PD-L1 in cell lines and xenografts treated with decitabine, which improves the recruitment of CD8+ cells and the effectiveness of immunotherapy. Additionally, a benefit has been reported in patients with BRCA1 methylation with the use of PARP inhibitors and it is believed that the epigenetic characterization of the response to CDK4/6 inhibitors could improve the management of patients with ER+ and metastatic breast cancer who receive this medication as first-line treatment, but do not always respond to treatment (48).

2.2 Transcriptional level: transcriptional factor analysis + transcriptomics

Transcription factors are a large family of proteins involved in the regulation of gene expression. They can be categorized into general factors involved in the transcription of most genes and sequence-specific factors that direct the spatial and temporal expression patterns of organisms. The ENCODE Factorbook database contains profiles of nearly 700 proteins related to transcription, including specific factors, cofactors, and members of the RNA polymerase II complex. Transcription factors can recruit RNA polymerase directly or rely on accessory factors for their function. Many eukaryotic TFs require co-activator or co-repressor complexes involved in chromatin remodeling. Some TFs interfere with the binding of other proteins (54–56). The HumanTFs database defines transcription factors as proteins that bind to DNA through a DNA binding domain (DBD) and regulate transcription. The database includes 1639 probable human transcription factors, with the C2H2-ZF and homeodomain DBDs being the most common. TFs often have multiple copies of a single DBD type and a combination of effector domains. The expression patterns of TFs largely depend on the DBD, with homeodomains showing tissue specificity. Transcription factors are often grouped based on the family of their DBD, which reflects the sequences they recognize. The largest families include C2H2-ZF, homeodomain, bHLH (basic helix-loop-helix), bZIP (basic leucine zipper), and NHR (nuclear hormone receptor), which were among the earliest described. This grouping by DBD family has its roots in homology and may have limited the identification of new factors, but it aligns with the evolutionary history of DBDs, which originated from a common ancestral set and underwent duplication and divergence (55, 57–59).

Transcription factors (TFs) can bind to specific sequences, called motifs, in the regulatory regions of target genes. These motifs are typically 6 to 20 base pairs long. Methods like ChIP-seq and HT-SELEX are used to identify binding sequences and determine motifs (56, 60). Weight matrices or hidden Markov models are then used to characterize the binding preferences of TFs, and databases like JASPAR store collections of motifs. However, the presence of a motif alone does not guarantee that a TF regulates a gene. Factors such as the accessibility of chromatin, DNA methylation state, nucleosome positioning, and interactions with other TFs also influence TF binding. Only the structural factor CTCF binds to almost 14,000 instances of its motif in the genome, while other TFs exhibit more complex binding patterns (61). TFs compete or interact with nucleosomes to access their motifs. The binding of TFs is associated with nucleosome repositioning, which is anti-correlated with DNA methylation levels. The absence of nucleosomes indicates a high and stable level of gene expression. Single-molecule tracking studies have shown that TFs transiently bind to DNA, and interactions between TFs can affect their diffusion dynamics (62).

Although transcription factors have been divided into activators and repressors, many TFs can recruit multiple cofactors with opposing effects, making it more appropriate to include the target and the condition under which a factor is operating. KRAB C2H2-ZF factors are repressors of transposable elements, by promoting their silencing (55); while HOXA5 functions as an activator of p53 in breast cancer cells (63, 64). Therefore, binding to the motif may be insufficient to determine the effect of the TF on the locus, and may simply reflect chromatin accessibility (54).

The binding of transcription factors (TFs) near the transcription start site can provide insights into gene expression levels. The binding of specific factors like E2F4 to non-methylated promoters explains a significant portion of gene expression variance, especially in CGI promoters. However, the predictive power decreases in promoters with low CpG density, suggesting the involvement of methylation in regulation. General factors contribute to a larger percentage of expression variance, and this percentage decreases when incorporating sequence-specific factors and histone modifications, indicating redundant regulatory mechanisms. Genes regulated post-transcriptionally, involved in cell cycle control and exhibiting tissue-specific expression differences, are particularly challenging to predict (54). In a study by Inoue and Harimoto, four expression patterns were identified among gene-TF pairs: no change, correlated expression, non-correlated expression due to constant TF levels, and lack of correlation due to variable genes. Correlated expression was associated with cell cycle and DNA replication genes, while various human diseases were linked to non-correlated expression. The lack of correlation revealed additional regulatory mechanisms. The third pattern, characterized by constant TF binding and gene degradation based regulation, was associated with genes whose expression is primarily determined by transcript degradation rather than synthesis. Disruption of degradation in such cases can have detrimental effects, as seen in the accumulation of the oncogene β-catenin (65).

2.2.1 TFs in breast cancer

Transcription factors regulate a large number of biological processes and are essential for maintaining homeostasis, so it is not surprising that their alteration is associated with different diseases. In particular, TFs represent almost 20% of the identified oncogenes (47). However, transcription factors are not only affected by direct mutations, but the mutation and methylation of regulatory regions can also disrupt their binding and function (55).

In addition, there are a large number of transcriptional cascades triggered by the action of a few factors, which act as master transcriptional regulators. Master regulators are the genes that control the specification of a lineage either by direct or indirect regulation, whose altered expression can change cell fate (66). In breast cancer, AGTR2, ZNF132, and TFDP3 have been identified as master regulators linked to the distinctive features of cancer. Focusing on the signal transduction pathways, TSHZ2, HOXA2, MEIS2, HOXA3, HAND2, HOXA5, TBX18, PEG3, GLI2 and CLOCK were also identified, with the latter being the only positive regulator. Regulators in both sets show some redundancy in their targets, suggesting robust regulation. In the case of signal transduction, the *Hedgehog* pathway stands out for its relationship with morphogenesis and the self-renewal of stem cells (64, 67).

This relationship between cancer and cell differentiation and morphogenesis fits with the oncogenic theory of cancer, according to which, the aberrant expression of development genes allows the reprogramming of somatic cells to an immortal stem cell line of cancer cells, and then to a new cell identity (68). The epithelial-mesenchymal transition is a good example of this theory, as it depends on the same transcription factors - Snail, Slug, Twist and FoxD3 - during development as well as during cancer progression. Eventually, metastasis also resembles embryonic development of different structures, as it depends on the same morphogens: Wnt and Hedgehog ligands, bone morphogenetic proteins (BMPs), and fibroblast growth factors (FGFs) (47).

On the other hand, while the alteration of transcription factors or their expression modifies complete processes, the alteration of binding motifs also has an effect, perhaps more limited, by affecting only the relationship between the TF and a target gene, but equally problematic. When analyzing the accessibility of DNA in 23 different types of cancer, hundreds of non-coding and somatic mutations were found that affect the binding of transcription factors, suggesting a ubiquitous mechanism for manipulating genetic expression. The grouping of cancer types by DNA accessibility agrees with the grouping by multi-omic expression - expression of transcripts, microRNAs and proteins, in addition to DNA methylation and copy number - suggesting functional relevance (69). Accessible and specific regions of a group are hypomethylated compared to other clusters, while exhibiting enrichment of SNPs and cancer-associated TF motifs better represented in the cluster. Approximately 65% of these SNPs do not have the nearest gene as a putative target. When focusing on breast cancer, 36% of accessible regions were also accessible in other types of cancer, establishing a division between basal and non-basal tumors, and, as a result, a survival difference dependent on the accessibility of ESR1 motifs (6).

Of the 294 oncogenic TFs (70), the androgen and estrogen receptors, the BRCA1 and BRCA2 genes, MYC and GATA3 stand out for their association with breast cancer subtypes. The androgen receptor has been associated with the Her2-enriched subtype (14), although it also has clinical relevance, and is in fact more common in ER+ tumors (71). The estrogen receptor, on the other hand, is the marker par excellence of the luminal subtypes; while germline or somatic mutations of the breast cancer susceptibility genes and MYC activation are frequent in the basal subtype. Finally, the transcriptional factor GATA3 is particularly mutated in luminal tumors, where it is also often overexpressed (5).

Given the relevance of the estrogen receptor in the classification of breast cancer, it is worth delving into its functioning. In addition to its role as a transcriptional factor, ER is a member of the nuclear hormone receptor superfamily, encoded by the paralogs ESR1, on 6q25.1 and ESR2, on 14q22-24. The receptors that result from each gene, ERα and ERβ, respectively, have tissue-specific expression and differences in terms of structure and DNA binding, which nevertheless allow the formation of homodimers and heterodimers with a similar affinity for DNA. Steroid hormones diffuse through the plasma membrane and once the ligand binding domain of the receptor receives estrogen, a stable dimer is formed, capable of interacting with specific sequences through the DNA binding domain. The estrogen response elements (EREs) are palindromes of 5 base pairs separated by 3 bps, whose consensus sequence is GGTCAnnnTGACC. When the activated receptor binds to the ERE, it is believed that a pre-initiation complex for RNA polymerase is formed, through the inactivation or dissociation of co-repressors and the recruitment of co-activators, which favors cell proliferation (1).

In addition to the nuclear ER, there are receptors on the plasma membrane and in the mitochondria. On the membrane, the ER associates with lipid vesicles, interacts with growth factor receptors such as EGFR and HER2 and participates in non-genomic responses to estrogen, which range from the activation of kinases to the modulation of cellular migration, survival and proliferation. In the mitochondria, the presence of ERβ affects metabolism and anti-apoptotic signals (72).

The activity of the receptor changes with the nature of the ligand, phosphorylation and interaction with other TFs. The ER can promote transcription without hormone, either by interacting with the transcription factor Sp1 and its response elements or because extracellular growth factors cause phosphorylation and activation of the ER, crossing steroid hormone signaling pathways and receptors. The interaction with other TFs explains the activation of genes without ERE, while the interaction of ER with cyclin D1 allows the receptor to bind to EREs, also without estrogen and additively when there is hormone.

In addition, the function of the ER depends on the expression of the receptor, which is subject to regulation at multiple levels. The receptor promoter contains the motifs of different transcription factors such as Sp1, FoxA1 and Ezh2; in addition to several incomplete EREs. For its part, the six known isoforms of the messenger encode the same protein, but exhibit tissue-specific expression patterns and include different 5'UTRs, which seem to fold with more or less stability and could alter the efficiency of

translation. On the other hand, the 3'UTR contains the seeds of 72 microRNAs, including miR-22, miR-206, miR-221 and miR-222, which are overexpressed in ER- tumors compared to ER+ and; the miR-17-92-miR-18a, miR-19b and miR-20b cluster, whose expression depends on ERα and cMYC, forming a negative feedback loop. Normally the 29 CpGs on ESR1 lack methylation, however extensive methylation has been documented in ER- cell lines (1).

The main alteration of the ER during the progression of breast cancer is in terms of its genetic expression. Although normal tissue only presents ERα, early ductal tumors have high levels of ERα and low levels of ERβ, while in advanced stages both receptors are lost. On the contrary, lobular tumors begin with high levels of both receptors and end up losing ERβ (73). Large disruptions and loss of heterozygosity rarely affect the receptor, so they cannot be used to explain ER- status. In other words, there are few documented mutations in primary tumors, which become frequent in metastatic lesions. For example, the Y537N mutant, which has been linked to bone metastasis and allows the constitutive activation of the TF, by abolishing the phosphorylation site. In addition, about 7% of tumors have mutations in the *enhancers* linked to ESR1 (7). Therefore, the alteration in breast cancer of the ER is more at the level of expression and has transcriptional effects.

Chromatin precipitation studies indicate between 5000 and 1000 EREs, which are reduced to approximately 1500 estrogen response genes (1). However, the effect of the TF is not solely local. Initially, it was described that ERα, FOXA1 and AP-2γmediated the long-distance interaction between GREB1 and TFF1, but thanks to ChIA-PET studies, 689 chromatin loops formed by the interaction between distal and proximal EREs are now known. The loops are formed both intrachromosomally and interchromosomally and are believed to form subcompartments in the nuclear space (1, 74).

2.3 Postranscriptional level: microRNA expression + transcriptomics

MiRNAs are small, non-coding RNAs that regulate gene expression post-transcriptionally (36). They inhibit translation through base complementarity and can positively influence translation (75). MiRNAs are evolutionarily conserved, tissue-specific, and crucial for various cellular processes like proliferation and apoptosis (76–78). They can regulate a large portion of coding genes, impacting the cell's gene expression profile (79). MiRNAs are abundant in somatic tissues and play a vital role in maintaining transcriptional network integrity (80). MiRNA production involves several steps: transcription of primary miRNAs (pri-miRNAs), recognition by the Drosha complex, formation of pre-miRNAs, export to the cytoplasm, processing by DICER, and transfer to the RISC complex (81, 82). Pri-miRNAs undergo cuts and modifications to become mature miRNAs, which play important roles in cellular processes (83). The production of mature miRNAs is efficient in healthy adult tissues (84).

MiRNA transcription can originate from their own promoters or coding gene promoters. They can be mono or polycistronic, with families sharing sequence similarity and functionality (75). For example, the miR-200 family is transcribed from different loci (81, 85). MiRNAs function as guides within the RISC complex, binding to messenger RNAs (mRNAs) in the 3'UTR region through miRNA response elements (MREs) (78). Binding leads to mRNA degradation mediated by argonaute proteins (77).

Predicting target messengers for miRNAs is challenging due to the size and low specificity of miRNAs. In addition to sequence information, conservation and thermodynamic stability play important roles. Various algorithms have been developed, including sequence-based and gene expression-based approaches that consider negative correlation or employ more complex methods (86). Databases like miRanda, TargetScan, and miRTarBase, which store predictions and validated cases, are valuable resources for miRNA target information (87).

2.3.1 miRNAs in breast cancer

Counterintuitively given its pleiotropic role, many microRNAs are found in fragile regions of the genome and suffer from alterations in copy number (88), as seen with miR-125b, let-7g, miR-21, and 72.8% of miRNAs associated with breast cancer (23). While mutations on specific microRNAs have a limited effect, alterations to the miRNA production process affect the cell on an even wider scale, as they simultaneously alter multiple pleiotropic regulators. As a result, mutations in DROSHA and DICER are linked to low survival in patients with ovarian, lung, and breast cancer. Genetic expression alteration has been attributed to the regulators MYC and ADARB1 in the case of DROSHA, and miR-103/107 and let-7 in the case of DICER. Under-expression of DICER is associated with the basal subtype of breast cancer (77). Interestingly, there are miRNAs that are over-expressed when DROSHA or DICER are under-expressed, suggesting an alternative mechanism. The binding of KSRP to the RISC complex along with some pre-miRNAs, such as miR-21, posits this splicing protein, which is induced in response to DNA damage, as a possible part of that mechanism (82).

Other components of the microprocessor complex that are altered in cancer are DGCR8 and the helicases p68 and p72, which connect the microprocessor complex to p53. In the next step in miRNA production, inactivating mutations of XPO5 in tumors with microsatellite instability in colon, gastric, and endometrial tumors have been identified. The mutation of XPO5 increases the risk of breast cancer. The phosphorylation of XPO5 by MAPK/ERK in liver cancer has the same result as inactivating mutations, by preventing the export of pre-miRNAs to the cytoplasm. Outside the nucleus, factors associated with DICER, such as TARBP2 and AGO2, also exhibit alterations. Mutations in TARBP2 identified in carcinomas with microsatellite instability change the reading frame of the gene; while its under-expression is associated with melanomas and metastatic tumors of the breast and prostate. Over-expression of AGO2 has been reported in breast, gastric, and head and neck tumors (89).

Dependent on transcription, miRNAs are also modulated by DNA methylation and transcription factor binding. It is estimated that about 33% of the de-regulated miRNAs in cancer have alterations in DNA methylation (84). In cell lines without

DNTM1 or DNTM3B, placentary miRNA expression is observed, normally silenced. In this regard, an important overlap has been reported between microRNAs marked by the Polycomb silencing complex in embryonic stem cells and those with CGI methylation in tumor cells (90). To mention a specific example, there is miR-205, whose sub-expression is associated with methylation of its promoter and resistance to treatment and epithelial-mesenchymal transformation (EMT) (75).

Another recent example is upregulation of miR-375 *via* EZH2 methylation leading to FOXO1 inhibition. Inactivation of FOXO1 in turn promotes deregulated responses of the p53 pathway associated with breast cancer oncogenesis (91). Thus mir-375 has been recognized as a epigenetically regulated oncomir in breast cancer (Figure 5).

Examples of transcriptional regulation of microRNAs include regulation of MYC over miR23a and of NFkB over miR-29b (79). The case of miR-29 is interesting, because both regulators and effectors of the miRNA are known. MYC binding seems to be the initial step of silencing, and is followed by recruitment of histone modifiers. As part of the so-called *epi-miRNAs*, the miR-29 family inhibits DNMT3A, DNMT3B and DNMT1 (75, 79).

Finally, the tumor microenvironment can also alter miRNA levels, as observed in hypoxic breast tumors, where hypoxia inhibits oxygen-dependent histone demethylases KDM6A and KDM6B. As a result, the methylation - at the histone level - of the DICER promoter increases and its expression decreases, which also decreases the processing of miRNAs. The miR-200 family is one of the main ones affected by the sub-expression of DICER (82). By regulating the expression of transcription factors ZEB1 and ZEB2, which inhibit the transcription of epithelial genes such as E-cadherin; the loss of miR-200 favors the epithelial-mesenchymal transformation and is associated with metaplastic and aggressive breast tumors (76). In parallel to EMT, the loss of miR-200 releases the transcription factor ETS1 from the repression of the miRNA. ETS1 regulates the expression of angiogenic factors and, together with ELK1, triggers the methylation - at the DNA level - of the DROSHA promoter, further reducing miRNA levels, which has been associated with poorly differentiated tumors (76, 82).

On the other hand, it is common to find circulating miRNAs in fluids such as plasma and saliva. MicroRNAs in blood serum can even be used as prognostic biomarkers in breast, prostate, colon, ovarian and lung cancer. Specifically, the detection of mir-21, miR-92a, miR-10b, miR-125b, miR-155, miR-191, miR-382 and miR-30a would allow early identification of breast cancer (36). These miRNAs are protected from the action of RNAses thanks to their binding with lipoproteins and ribonucleoproteins or by their packaging in microvesicles (77). Once they are endocytosed, the regulation of translation in receptor cells is altered, involving microRNAs as signaling molecules. In this sense, it has been shown that cancer-associated fibroblasts secrete a different spectrum of miRNAs than normal fibroblasts, and these are not the only components of the microenvironment releasing microRNAs (78).

Even without considering circulating miRNAs, there is a clear difference between the profiles of normal breast tissue and tumors, with miR-10b, miR-125b, miR-145, miR-21 and miR-155 showing the most significant differences (88). In addition, miRNA expression profiles can distinguish between subtypes of breast cancer (5) and between cell subpopulations, with luminal progenitors being the cells most similar to basal tumors and

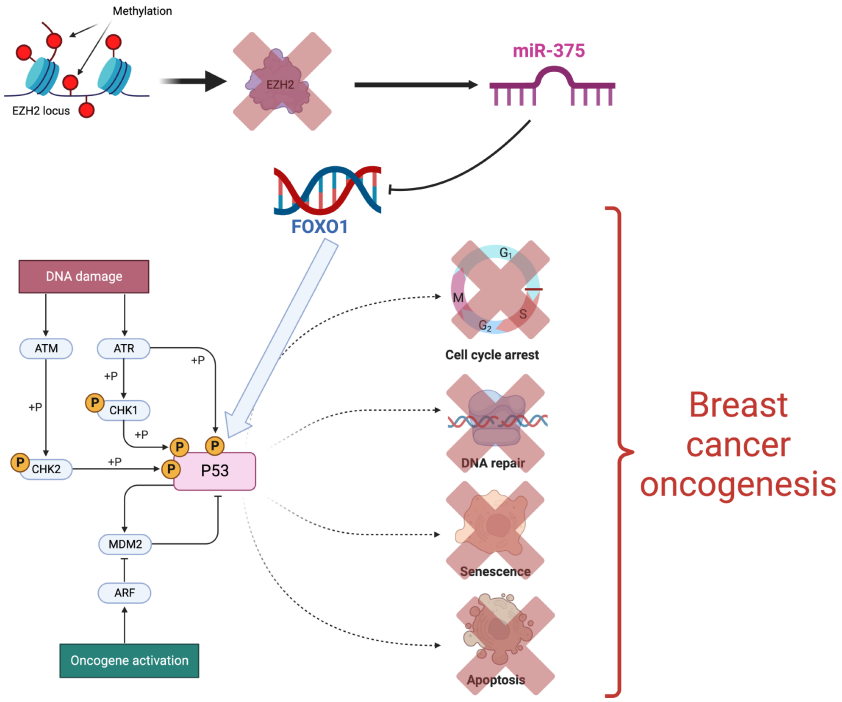


FIGURE 5
Epigenetic activation of the mir-375 oncogene [Figure created using Biorender.com, inspired from (91)].

mature luminal cells being the closest to luminal B subtype tumors. Luminal microRNAs regulate cell differentiation and development; while basal microRNAs regulate intracellular localization, organelle transport and biosynthesis, secretion and cell-cell interaction (92). Although the correspondence between intrinsic subtypes and miRNA profiles is noisy (5), the over-expression of miR-206 has been associated with ER- tumors and the under-expression of miR-125a/b with those enriched in Her2 (76).

MiR-206 inhibits the expression of ESR1; while its expression is favored by ERα and not by ERβ or progesterone, suggesting a negative feedback loop. Other miRNAs that regulate ESR1 are miR-18a/b, miR-193b and miR-302c, whose expression, along with that of miR-206, causes cell cycle arrest and inhibits estrogen-dependent proliferation. In addition, miR-17-5p has the same effect, due to an indirect regulation of ERα through AIB1. The miRNA profile of breast cancer stem cells is also different, being enriched with miRNAs associated with self-renewal, such as let-7 and miR-34. Let-7 regulates oncogenes such as HRAS, HMGA2, MYC and caspase-3. In breast cancer over-expression of miR-34 causes cell cycle arrest and its under-expression increases invasive capacity (77).

miRNAs with a role in cancer can function as oncogenes or as tumor suppressors, depending on their targets. Tumor suppressor microRNAs inhibit the expression of genes that promote tumor development, so their sub-expression is harmful, as is the case with miR-200. OncomiRs, on the other hand, regulate tumor suppressors and it is their over-expression that is harmful, as is the case with miR-21, which regulates promoters of apoptosis and cell migration (77). In addition, a subcategory of oncomiRs could be defined with miRNAs exclusively pro-metastatic, such as miR10b and the miR-373/520c family. It has been reported that miR-10b is over-expressed only in metastatic breast cancer cells and not in the primary tumor; miR-10b inhibits the transcriptional factor HOXD10 and, in doing so, triggers a cascade of changes that ultimately lead to pro-metastatic RHOC expression, cell migration, and invasion (76).

However, the role of a miRNA could depend on the cellular context, as miRNA-mRNA regulatory interactions may not necessarily exist in all types of cancer (86). In a pan-cancer, computational study of miRNAs that direct genetic expression, it was observed that miRNA-gene interactions are not conserved, even though there are 22 miRNAs that do function as *drivers* in different types of cancer. Except for miR-5001 in colorectal cancer and miR-2276 in endometrial cancer, in this study all miRNAs are classified as tumor suppressors and the let-7 family functions as a TSG and as an oncomiR at the same time (93).

Despite the fact that each miRNA can regulate hundreds of genes, miRNAs have been proposed as possible means to regulate cancer genes, either by introducing oligonucleotides similar to miRNA to restore miRNA expression and suppress oncogenes, or by introducing antagonists to inhibit the miRNA of interest. An example of antagonists or antagomiRs are miRNA sponges, synthetic messengers with multiple binding sites for a specific miRNA, which then capture it, preventing it from inhibiting TSGs. There are formulations of miRNA-like oligonucleotides, miRNA sponges, anti-miRNA oligonucleotides, and small molecules that are being studied in cancer models. For breast cancer, at least an antagomiR-10b and a miR-195-like oligo have been tested. The antagonist inhibits lung metastasis, but not the growth of the primary tumor in mice; whereas the miR-195-like oligo increases sensitivity to treatment and inhibits Raf1 and Bcl2 translation in cell lines (77, 78).

3 Single cell breast cancer multi-omics

In recent times, multi-omic approaches have been further advanced with the advent of single cell sequencing techniques that have allowed for the integration of transcriptome data, as well as, other omics such as ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) (94) to achieve a deeper understanding of molecular profiles and its functions at the level of a single cell or cell-type. These already outstanding methods are being further advanced by the integration of spatial multi-omics (95). The goal of *spatial* approaches is to be able to assign cell types (as identified by the mRNA and other omic sequencing readouts) to their locations in the histological sections of a given sample tissue. Spatial omics allow, for instance, to uncover cellular heterogeneity in tissues, tumors, immune cells as well as determine the subcellular distribution of biomolecules in diverse phenotypes.

Single cell and spatial multi-omics are thus becoming relevant tools and methods to analyze cancer biology from its basic principles (e.g. oncogenesis) (96) to the way these tumors evolve and their related outcomes by allowing to account for issues such as how dynamic processes and clonal selection manifest in cellular states, epigenetic profiles, spatial distributions and interactions with the microenvironment (97, 98).

Single cell studies in breast tumors, although quite recent, are starting to render fruits in the understanding of breast cancer heterogeneity as exemplified by the recent discovery of two lipid-associated macrophage states LAM1 and LAM2 (99) that are being established as biomarkers distinct clinical outcomes in several breast cancer datasets (100). Single cell multi-omics is also being used to develop strategies for clinical trial evaluation and drug discovery (101).

4 Applications of concerted multi-omic regulation analysis in breast cancer

The concurrent activity of several biological processes as measured by diverse omic technologies is paving the way towards advancing, both our knowledge about breast tumor biology and our therapeutic approaches. A number of these advances have ben summarized recently by Mehmood and collaborators (102). These authors have described the power of multi-omics to face the challenges of multidrug resistance (MDR) and relapse in breast carcinoma treatment. They emphasize the importance of elucidating multi-omic mechanisms to design therapies able to overcome drug resistance. Since breast carcinoma treatment decisions rely not only on prognosis factors but

also on the assessment of pathological and clinical factors, the integration of data from multiple factors through a multiomics approach can provide valuable insights for therapeutic decisions. Along the same lines Ektefaie et al. (103), describe the development of weakly supervised deep learning models for analyzing multiomics from breast cancer biopsy samples. These automated models developed for tumor detection and pathology subtype classification demonstrated high accuracy and were validated in independent cohorts.

Regarding the interplay of epigenomic and genomic features, it has been discussed (104) that the CT83 gene is frequently activated in triple negative breast carcinomas (TNBC) and several other cancers, while it remains silenced in non-TNBC, normal nontestis tissues, and blood cells. A significant correlation was found between hypomethylation on chromosome X and the abnormal activation of CT83 in breast cancer. Furthermore, the activated CT83 was associated with unfavorable overall survival in breast cancer and worse outcomes in other cancers. The authors argue that abnormal activation of CT83 is likely oncogenic by triggering cell cycle signaling. Also in the context of TNBC multi-omic studies (17) combined with immune profiling have revealed a classification of the microenvironment phenotypes in triple-negative breast cancer (TNBC) into three distinct clusters. Cluster 1, known as the *immunedesert* cluster, exhibits low infiltration of microenvironmental cells. Cluster 2, referred to as the *innate immune-inactivated* cluster, demonstrates the presence of resting innate immune cells and nonimmune stromal cells infiltration. Lastly, cluster 3, the *immune-inflamed* cluster, shows abundant infiltration of both adaptive and innate immune cells. The clustering results were validated internally using pathologic sections and externally using The Cancer Genome Atlas and METABRIC as independent cohorts. These microenvironment clusters also displayed significant prognostic efficacy. The authors describe potential immune escape mechanisms associated with each cluster. Cluster 1 is characterized by an inability to attract immune cells, with low immune infiltration correlated with MYC amplification. In cluster 2, chemotaxis but innate immune inactivation and low tumor antigen burden potentially contribute to immune escape, with mutations in the PI3K-AKT pathway possibly associated with this effect. Lastly, cluster 3 is distinguished by high expression of immune checkpoint molecules. A similar approach to classification was made by Coria-Rodriguez and coworkers (105) to infer epigenomic signatures to define TNBC classes with differential response to therapy with drug repurposing goals in mind.

Tumor metabolic reprogramming has been studied with a multi-omic strategy by Iqbal and his group of collaborators (106) to show that there are antagonistic roles of CBX2 and CBX7 in metabolic reprogramming of breast cancer. They identified significant roles of CBX2 and CBX7 in positive and negative regulation of glucose metabolism and provided functional evidence for the mTOR complex 1 signaling in mediating competing effects of CBX2 and CBX7 on breast cancer metabolism. Disease-specific survival and drug sensitivity analysis revealed that CBX2 and CBX7 predicted patient outcome and sensitivity to FDA approved/investigational drugs.

Muli-omic analysis have also provided relevant clues, for instance on the role of lipid metabolism for the development and outcomes on early breast carcinomas (107). Concurrent ultrahigh-performance liquid chromatography-mass spectrometry experiments along with transcriptomics, and genomics data led to the identification of 18 oxylipins, metabolites of omega-3 or omega-6 polyunsaturated fatty acids, that were differentially expressed in breast tumors versus healthy sample tissues, including anandamide, prostaglandins and hydroxydocosahexaenoic acids. The authors hypothesize that oxylipin signatures reflect the organism's level of response to the disease and may become markers of malignancy.

Tumor survival and drug-response predictions have been discussed at the light of breast cancer multi-omics (108) aimed on quantifying survival and drug response. The framework utilizes Neighborhood Component Analysis (NCA) for feature selection from multi-omics datasets obtained from The Cancer Genome Atlas (TCGA) and Genomics of Drug Sensitivity in Cancer (GDSC) databases. A Neural network framework, fed with the NCA selected features, is used to develop prediction models for survival and drug response in breast cancer patients. The results demonstrate a strong linear relationship between predicted and actual IC50 values outperforming previous approaches and highlighting the importance of multi-omics data integration.

The improved knowledge provided by multi-omic studies is also impacting on novel treatment designs such as immunotherapy, as has been recently summarized in the review work by Leung, et al. (109); a well as clues helping to advance druggable targets and autophagic modulators such as SF3B3 and SIRT3, that may improve the treatment of invasive breast carcinomas (110); and on exploiting the therapeutic and diagnostic value of IMMT in breast cancer as well as its immunological role (111). Immune infiltrate activity in breast tumors has been also further clarified by multi-omics as exemplified by the work of Tian and collaborators (112) about the relationship plasmacytoid dendritic cells and breast cancer.

Multi-omic strategies have also allowed to discern particular sets of biomolecular interactions relevant to certain aspects of breast cancer biology. Some of these interactions are indeed becoming interesting clues towards targeted therapy. Such is the case of the mechanisms by which the mitochondrial protease ClpP is activated by drugs that are able to breakdown essential mitochondrial pathways in triple-negative breast cancer (113). Similarly, the role of heat shock proteins (which may be either acting as oncogenes and onco-suppressor genes) has been recently discussed at the light of multi-omic analysis (114). Discerning the mechanisms of novel therapeutic drugs such as signaling inhibitors is crucial on our advance towards precision therapeutics of breast cancer. In this regard, Marczyk and collaborators (115) have studied the effects of navitoclax, a BCL2 family inhibitor, on the transcriptome, methylome, chromatin structure, and copy number variations of MDA-MB-231 triplenegative breast cancer (TNBC) cells. They were able to derive an 18-gene navitoclax resistance signature. Other pharmacological resistance mechanisms have been further elucidated. For instance, methylation events leading to HSD17B4 silencing have been identified as part of a predictive and response marker of HER2-positive breast cancer to HER2-directed therapy (116).

Breast cancer multi-omic integration tools have been recently developing at a fast pace. In order to better exploit the available and upcoming resources, researchers at the Chinese Academy of

Sciences implemented MOBCdb a database integrating multi-omics data on breast cancer (117). MOBCdb is a user-friendly and readily available database that combines genomic, transcriptomic, epigenomic, clinical, and drug response information from various subtypes of breast cancer. It offers a convenient platform for users to access simple nucleotide variations (SNV), gene expression, microRNA expression, DNA methylation, and specific drug response data through different search methods. Additionally, the genome-wide browser and navigation feature in MOBCdb enable simultaneous visualization of multi-omics data from multiple samples.

5 Conclusions

We have discussed how various types of gene regulatory phenomena in breast cancer arise from several omics data, such as gene and non-coding RNA transcriptomics, methylation and transcription factor activity, as reported in the recent literature. We have also discussed how this knowledge can be integrated to provide a more comprehensive understanding of gene regulation in breast cancer, highlighting the importance of considering the spatial and temporal context in which gene regulation occurs, as well as the role of regulatory elements such as non-coding RNA and epigenetic modifications. In this regard, recent advances in single cell approaches to breast cancer multi-omics have been also presented. Some applications to tumor sub-classification, prognosis and survival analysis, drug repurposing and personalized therapeutic designs were introduced.

For concreteness, other potentially relevant aspects of the complex regulatory patterns in breast cancer have been left out for future discussion. Such is the case of the role played by copy number variants, long non-coding RNAs and the multi-scale three dimensional structure of nuclear chromatin. However, by considering the levels discussed in this review article, we have tried to unveil the potential of multi-omic approaches to improve our understanding of the complex molecular processes underlying breast cancer that may hopefully help us identify new therapeutic targets.

Author contributions

Both authors performed research. SO drafted the first version of the manuscript, EH-L edited and rewrote the manuscript. All authors contributed to the article and approved the submitted version.

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.

Publisher's note

References

1. Parl FF. *The etiology of breast cancer: Endogenous and exogenous causes*. Seattle, WA, USA: Amazon Publishing. (2014).

2. Feng Q, Jiang M, Hannig J, Marron J. Angle-based joint and individual variation explained. *J multivariate Anal* (2018) 166:241–65. doi: 10.1016/j.jmva.2018.03.008

3. Wu C, Demerath EW, Pankow JS, Bressler J, Fornage M, Grove ML, et al. Imputation of missing covariate values in epigenome-wide analysis of dna methylation data. *Epigenetics* (2016) 11:132–9. doi: 10.1080/15592294.2016.1145328

4. Perou CM, Sørlie T, Eisen MB, Van De Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* (2000) 406:747. doi: 10.1038/35021093

5. Network CGA, Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, et al. Comprehensive molecular portraits of human breast tumours. *Nature* (2012) 490:61–70. doi: 10.1038/nature11412

6. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, et al. The chromatin accessibility landscape of primary human cancers. *Sci (New York N.Y.)* (2018) 362. doi: 10.1126/science.aav1898

7. Pellacani D, Tan S, Lefort S, Eaves CJ. Transcriptional regulation of normal human mammary cell heterogeneity and its perturbation in breast cancer. *EMBO J* (2019) 38:e100330. doi: 10.15252/embj.2018100330

8. Sims AH, Howell A, Howell SJ, Clarke RB. Origins of breast cancer subtypes and therapeutic implications. *Nat Clin practice Oncol* (2007) 4:516–25. doi: 10.1038/ncponc0908

9. Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. *Front Med* (2017) 4:227. doi: 10.3389/fmed.2017.00227

10. Dai X, Xiang L, Li T, Bai Z. Cancer hallmarks, biomarkers and breast cancer molecular subtypes. *J Cancer* (2016) 7:1281. doi: 10.7150/jca.13141

11. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* (2009) 27:1160–1167. doi: 10.1200/jco.2008.18.1370

12. Cárdenas-Sánchez J, Erazo Valle-Solís AA, Arce-Salinas C, Bargalló-Rocha JE, et al. *Consenso Mexicano sobre diagnosticó y tratamiento del cancer mamario (Mexican consense on the diagnostics and treatment of breast cancer)*. Colima, Mexico: Consenso Cancer Mamario. (2019). Available at: http://consensocancermamario.com/documentos/FOLLETO_CONSENSO_DE_CANCER_DE_MAMA_8aRev2019a.PDF.

13. Prat A, Pineda E, Adamo B, Galván P, Fernández A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* (2015) 24:S26–35. doi: 10.1016/j.breast.2015.07.008

14. Daemen A, Manning G. Her2 is not a cancer subtype but rather a pan-cancer event and is highly enriched in ar-driven breast tumors. *Breast Cancer Res* (2018) 20:1–16. doi: 10.1186/s13058-018-0933-y

15. Lehmann BD, Colaprico A, Silva TC, Chen J, An H, Ban Y, et al. Multi-omics analysis identifies therapeutic vulnerabilities in triple-negative breast cancer subtypes. *Nat Commun* (2021) 12:1–18. doi: 10.1038/s41467-021-26502-6

16. Li X, Zhou J, Xiao M, Zhao L, Zhao Y, Wang S, et al. Uncovering the subtype-specific molecular characteristics of breast cancer by multiomics analysis of prognosis-associated genes, driver genes, signaling pathways, and immune activity. *Front Cell Dev Biol* (2021) 9:2021. doi: 10.3389/fcell.2021.689028

17. Xiao Y, Ma D, Zhao S, Suo C, Shi J, Xue M-Z, et al. Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms

of triple-negative breast cancer. *Clin Cancer Res* (2019) 25:5002–14. doi: 10.1158/1078-0432.CCR-18-3524

18. Rich JN. Cancer stem cells: understanding tumor hierarchy and heterogeneity. *Medicine* (2016) 95:S2–S7. doi: 10.1097/MD.0000000000004764

19. García-Cortés D, Hernández-Lemus E, Espinal-Enríquez J. Luminal a breast cancer co-expression network: Structural and functional alterations. *Front Genet* (2021) 12:629475. doi: 10.3389/fgene.2021.629475

20. Serrano-Carbajal EA, Espinal-Enríquez J, Hernández-Lemus E. Targeting metabolic deregulation landscapes in breast cancer subtypes. *Front Oncol* (2020) 10:97. doi: 10.3389/fonc.2020.00097

21. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* (2004) 10:789–99. doi: 10.1038/nm1087

22. Baylin SB, Jones PA. Epigenetic determinants of cancer. *Cold Spring Harbor Perspect Biol* (2016) 8:a019505. doi: 10.1101/cshperspect.a019505

23. Bose B, Bozdag S. (2019). mirdriver: A tool to infer copy number derived mirna-gene networks in cancer, in: *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*. Association for computing machinery (ACM). pp. 366–75.

24. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discovery* (2022) 12:31–46. doi: 10.1158/2159-8290.CD-21-1059

25. Bell DW. Our changing view of the genomic landscape of cancer. *J Pathol* (2009) 220:231–43. doi: 10.1002/path.2645

26. Anderson AR, Maini PK. Mathematical oncology. *Bull Math Biol* (2018) 80:945–53. doi: 10.1007/s11538-018-0423-5

27. Rossi C, Cicalini I, Cufaro MC, Consalvo A, Upadhyaya P, Sala G, et al. Breast cancer in the era of integrating "omics". *approaches Oncogenesis* (2022) 11:17. doi: 10.1038/s41389-022-00393-8

28. Hernández-Lemus E, Reyes-Gopar H, Espinal-Enríquez J, Ochoa S. The many faces of gene regulation in cancer: a computational oncogenomics outlook. *Genes* (2019) 10:865. doi: 10.3390/genes10110865

29. de Anda-Jáuregui G, Hernández-Lemus E. Computational oncology in the multi-omics era: state of the art. *Front Oncol* (2020) 10:423. doi: 10.3389/fonc.2020.00423

30. Gómez-Cebrián N, Domingo-Ortí I, Poveda JL, Vicent MJ, Puchades-Carrasco L, Pineda Lucena A. Multi-omic approaches to breast cancer metabolic phenotyping: Applications in diagnosis, prognosis, and the development of novel treatments. *Cancers* (2021) 13:4544. doi: 10.3390/cancers13184544

31. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* (2002) 3:415–28. doi: 10.1038/nrg816

32. Li E, Zhang Y. Dna methylation in mammals. *Cold Spring Harbor Perspect Biol* (2014) 6:a019133. doi: 10.1101/cshperspect.a019133

33. Xiao C-L, Zhu S, He M, Chen D, Zhang Q, Chen Y, et al. N6-methyladenine dna modification in the human genome. *Mol Cell* (2018) 71:306–18. doi: 10.1016/j.molcel.2018.06.015

34. Barros-Silva D, Marques CJ, Henrique R, Jeronimo C. Profiling dna methylation based on next-generation sequencing approaches: new insights and clinical applications. *Genes* (2018) 9:429. doi: 10.3390/genes9090429

35. Gouil Q, Keniry A. Latest techniques to study dna methylation. *Essays Biochem* (2019) 63:639–48. doi: 10.1042/EBC20190027

36. Cava C, Bertoli G, Castiglioni I. Integrating genetics and epigenetics in breast cancer: biological insights, experimental, computational methods and therapeutic potential. *BMC Syst Biol* (2015) 9:62. doi: 10.1186/s12918-015-0211-x

37. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the illumina methylationepic beadchip microarray for whole-genome dna methylation profiling. *Genome Biol* (2016) 17:1–17. doi: 10.1186/s13059-016-1066-1

38. Leygo C, Williams M, Jin HC, Chan MW, Chu WK, Grusch M, et al. Dna methylation as a noninvasive epigenetic biomarker for the detection of cancer. *Dis Markers* (2017) 2017. doi: 10.1155/2017/3726595

39. Vavouri T, Lehner B. Human genes with cpg island promoters have a distinct transcriptionassociated chromatin organization. *Genome Biol* (2012) 13:R110. doi: 10.1186/gb-2012-13-11-r110

40. Akulenko R, Helms V. Dna co-methylation analysis suggests novel functional associations between gene pairs in breast cancer samples. *Hum Mol Genet* (2013) 22:3016–22. doi: 10.1093/hmg/ddt158

41. Szyf M. Dna methylation signatures for breast cancer classification and prognosis. *Genome Med* (2012) 4:26. doi: 10.1186/gm325

42. Zoghbi HY, Beaudet AL. Epigenetics and human disease. *Cold Spring Harbor Perspect Biol* (2016) 8:a019497. doi: 10.1101/cshperspect.a019497

43. Li J, Sun C, Cai W, Li J, Rosen BP, Chen J. Insights into s-adenosyl-lmethionine (sam)-dependent methyltransferase related diseases and genetic polymorphisms. *Mutat Research/Reviews Mutat Res* (2021) 788:108396. doi: 10.1016/j.mrrev.2021.108396

44. Kriaucionis S, Tahiliani M. Expanding the epigenetic landscape: novel modifications of cytosine in genomic dna. *Cold Spring Harbor Perspect Biol* (2014) 6: a018630. doi: 10.1101/cshperspect.a018630

45. Wang G, Luo X, Wang J, Wan J, Xia S, Zhu H, et al. Medreaders: a database for transcription factors that bind to methylated dna. *Nucleic Acids Res* (2018) 46:D146–51. doi: 10.1093/nar/gkx1096

46. Blattler A, Farnham PJ. Cross-talk between site-specific transcription factors and dna methylation states. *J Biol Chem* (2013) 288:34287–94. doi: 10.1074/jbc.R113.512517

47. Huilgol D, Venkataramani P, Nandi S, Bhattacharjee S. Transcription factors that govern development and disease: An achilles heel in cancer. *Genes* (2019) 10:794. doi: 10.3390/genes10100794

48. Brown LJ, Achinger-Kawecka J, Portman N, Clark S, Stirzaker C, Lim E. Epigenetic therapies and biomarkers in breast cancer. *Cancers* (2022) 14:474. doi: 10.3390/cancers14030474

49. Vidal Ocabo E, Sayols S, Moran S, Guillaumet-Adkins A, Schroeder MP, Royo R, et al. A dna methylation map of human cancer at single base-pair resolution. *Oncogene* (2017) 36(40):5648–57. doi: 10.1038/onc.2017.176

50. Lindqvist BM, Wingren S, Motlagh PB, Nilsson TK. Whole genome dna methylation signature of her2-positive breast cancer. *Epigenetics* (2014) 9:1149–62. doi: 10.4161/epi.29632

51. Zhang J, Huang K. Pan-cancer analysis of frequent dna co-methylation patterns reveals consistent epigenetic landscape changes in multiple cancers. *BMC Genomics* (2017) 18:1–14. doi: 10.1186/s12864-016-3259-0

52. Broome R, Chernukhin I, Jamieson S, Kishore K, Papachristou EK, Mao S-Q, et al. Tet2 is a component of the estrogen receptor complex and controls 5mc to 5hmc conversion at estrogen receptor cis-regulatory regions. *Cell Rep* (2021) 34:108776. doi: 10.1016/j.celrep.2021.108776

53. Xu X, Yuan X, Ni J, Guo J, Gao Y, Yin W, et al. Magi2-as3 inhibits breast cancer by downregulating dna methylation of magi2. *J Cell Physiol* (2021) 236:1116–30. doi: 10.1002/jcp.29922

54. Cheng C, Alexander R, Min R, Leng J, Yip KY, Rozowsky J, et al. Understanding transcriptional regulation by integrative analysis of transcription factor binding data. *Genome Res* (2012) 22:1658–67. doi: 10.1101/gr.136838.111

55. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The human transcription factors. *Cell* (2018) 172:650–65. doi: 10.1016/j.cell.2018.01.029

56. Pratt HE, Andrews GR, Phalke N, Huey JD, Purcaro MJ, van der Velde A, et al. Factorbook: an updated catalog of transcription factor motifs and candidate regulatory motif sites. *Nucleic Acids Res* (2022) 50:D141–9. doi: 10.1093/nar/gkab1039

57. Fulton DL, Sundararajan S, Badis G, Hughes TR, Wasserman WW, Roach JC, et al. Tfcat: the curated catalog of mouse and human transcription factors. *Genome Biol* (2009) 10:1–14. doi: 10.1186/gb-2009-10-3-r29

58. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* (2009) 10:252–63. doi: 10.1038/nrg2538

59. Frietze S, Farnham PJ. Transcription factor effector domains. In: *A handbook of transcription factors*. New York, USA: Springer. (2011) p. 261–77.

60. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. Jaspar 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* (2020) 48:D87–92. doi: 10.1093/nar/gkz1001

61. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the vertebrate insulator protein ctcf-binding sites in the human genome. *Cell* (2007) 128:1231–45. doi: 10.1016/j.cell.2006.12.048

62. de Jonge WJ, Patel HP, Meeussen JV, Lenstra TL. Following the tracks: how transcription factor binding dynamics control transcription. *Biophys J* (2022) 121:1583–92. doi: 10.1016/j.bpj.2022.03.026

63. Chen H, Chung S, Sukumar S. Hoxa5-induced apoptosis in breast cancer cells is mediated by caspases 2 and 8. *Mol Cell Biol* (2004) 24:924–35. doi: 10.1128/MCB.24.2.924-935.2004

64. Tapia-Carrillo D, Tovar H, Velazquez-Caldelas TE, Hernandez-Lemus E. Master regulators of signaling pathways: an application to the analysis of gene regulation in breast cancer. *Front Genet* (2019) 1180. doi: 10.3389/fgene.2019.01180

65. Inoue M, Horimoto K. Relationship between regulatory pattern of gene expression level and gene function. *PloS One* (2017) 12:e0177430. doi: 10.1371/journal.pone.0177430

66. Chan SS-K, Kyba M. What is a master regulator? *J Stem Cell Res Ther* (2013) 3. doi: 10.4172/2157-7633.1000e114

67. Tovar H, García-Herrera R, Espinal-Enríquez J, Hernández-Lemus E. Transcriptional master regulator analysis in breast cancer genetic networks. *Comput Biol Chem* (2015) 59:67–77. doi: 10.1016/j.compbiolchem.2015.08.007

68. Vinnitsky V. Oncogerminative hypothesis of tumor formation. *Med Hypotheses* (1993) 40:19–27. doi: 10.1016/0306-9877(93)90191-R

69. Ochoa S, Hernández-Lemus E. Functional impact of multi-omic interactions in breast cancer subtypes. *Front Genet* (2023) 13:1078609. doi: 10.3389/fgene.2022.1078609

70. Lambert M, Jambon S, Depauw S, David-Cordonnier M-H. Targeting transcription factors for cancer treatment. *Molecules* (2018) 23:1479. doi: 10.3390/molecules23061479

71. Chen M, Yang Y, Xu K, Li L, Huang J, Qiu F. Androgen receptor in breast cancer: from bench to bedside. *Front Endocrinol* (2020) 573. doi: 10.3389/fendo.2020.00573

72. Liu M-M, Albanese C, Anderson CM, Hilty K, Webb P, Uht RM, et al. Opposing action of estrogen receptors α and β on cyclin d1 gene expression. *J Biol Chem* (2002) 277:24353–60. doi: 10.1074/jbc.M201829200

73. Paterni I, Granchi C, Katzenellenbogen JA, Minutolo F. Estrogen receptors alpha (erα) and beta (erβ): subtype-selective ligands and clinical potential. *Steroids* (2014) 90:13–29. doi: 10.1016/j.steroids.2014.06.012

74. Jia R, Chai P, Zhang H, Fan X. Novel insights into chromosomal conformations in cancer. *Mol Cancer* (2017) 16:173. doi: 10.1186/s12943-017-0741-5

75. Singh PK, Campbell MJ. The interactions of microrna and epigenetic modifications in prostate cancer. *Cancers* (2013) 5:998–1019. doi: 10.3390/cancers5030998

76. O'Day E, Lal A. Micrornas and their target gene networks in breast cancer. *Breast Cancer Res* (2010) 12:201. doi: 10.1186/bcr2484

77. Bertoli G, Cava C, Castiglioni I. Micrornas: new biomarkers for diagnosis, prognosis, therapy prediction and therapeutic tools for breast cancer. *Theranostics* (2015) 5:1122. doi: 10.7150/thno.11543

78. Klinge CM. Non-coding rnas: long non-coding rnas and micrornas in endocrine-related cancers. *Endocrine-related Cancer* (2018) 25:R259–82. doi: 10.1530/ERC-17-0548

79. Liu X, Chen X, Yu X, Tao Y, Bode AM, Dong Z, et al. Regulation of micrornas by epigenetics and their interplay involved in cancer. *J Exp Clin Cancer Res* (2013) 32:96. doi: 10.1186/1756-9966-32-96

80. Drago-García D, Espinal-Enríquez J, Hernández-Lemus E. Network analysis of emt and met micro-rna regulation in breast cancer. *Sci Rep* (2017) 7:13534. doi: 10.1038/s41598-017-13903-1

81. Baer C, Claus R, Frenzel LP, Zucknick M, Park YJ, Gu L, et al. Extensive promoter dna hypermethylation and hypomethylation is associated with aberrant microrna expression in chronic lymphocytic leukemia. *Cancer Res* (2012) 72:3775–85. doi: 10.1158/0008-5472.CAN-12-0803

82. Rupaimoole R, Calin GA, Lopez-Berestein G, Sood AK. mirna deregulation in cancer cells and the tumor microenvironment. *Cancer Discovery* (2016) 6:235–46. doi: 10.1158/2159-8290.CD-15-0893

83. Martienssen R, Moazed D. Rnai and heterochromatin assembly. *Cold Spring Harbor Perspect Biol* (2015) 7:a019323. doi: 10.1101/cshperspect.a019323

84. Hulf T, Sibbritt T, Wiklund ED, Bert S, Strbenac D, Statham AL, et al. Discovery pipeline for epigenetically deregulated mirnas in cancer: integration of primary mirna transcription. *BMC Genomics* (2011) 12:54. doi: 10.1186/1471-2164-12-54

85. Humphries B, Yang C. The microrna-200 family: small molecules with novel roles in cancer development, progression and therapy. *Oncotarget* (2015) 6:6472. doi: 10.18632/oncotarget.3052

86. Pham VV, Zhang J, Liu L, Truong B, Xu T, Nguyen TT, et al. Identifying mirna-mrna regulatory relationships in breast cancer with invariant causal prediction. *BMC Bioinf* (2019) 20:1–12. doi: 10.1186/s12859-019-2668-x

87. Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, et al. The multimir r package and database: integration of microrna-target interactions along with their disease and drug associations. *Nucleic Acids Res* (2014) 42:e133. doi: 10.1093/nar/gku631

88. Cho WC. Oncomirs: the discovery and progress of micrornas in cancers. *Mol Cancer* (2007) 6:60. doi: 10.1186/1476-4598-6-60

89. Ali Syeda Z, Langden SSS, Munkhzul C, Lee M, Song SJ. Regulatory mechanism of microrna expression in cancer. *Int J Mol Sci* (2020) 21:1723. doi: 10.3390/ijms21051723

90. Suzuki H, Takatsuka S, Akashi H, Yamamoto E, Nojima M, Maruyama R, et al. Genomewide profiling of chromatin signatures reveals epigenetic regulation of microrna genes in colorectal cancer. *Cancer Res* (2011) 71:5646–8. doi: 10.1158/0008-5472.CAN-11-1076

91. Guan X, Shi A, Zou Y, Sun M, Zhan Y, Dong Y, et al. Ezh2-mediated microrna-375 upregulation promotes progression of breast cancer *via* the inhibition of foxo1 and the p53 signaling pathway. *Front Genet* (2021) 12:633756. doi: 10.3389/fgene.2021.633756

92. Pal B, Chen Y, Bert A, Hu Y, Sheridan JM, Beck T, et al. Integration of microrna signatures of distinct mammary epithelial cell types with their gene expression and epigenetic portraits. *Breast Cancer Res* (2015) 17:85. doi: 10.1186/s13058-015-0585-0

93. Bose B, Moravec M, Bozdag S. Computing microrna-gene interaction networks in pan-cancer using mirdriver. *Sci Rep* (2022) 12:1–17. doi: 10.1038/s41598-022-07628-z

94. Kashima Y, Sakamoto Y, Kaneko K, Seki M, Suzuki Y, Suzuki A. Single-cell sequencing techniques from individual to multiomics analyses. *Exp Mol Med* (2020) 52:1419–27. doi: 10.1038/s12276-020-00499-2

95. Vandereyken K, Sifrim A, Thienpont B, Voet T. Methods and applications for single-cell and spatial multi-omics. *Nat Rev Genet* (2023) 1–22. doi: 10.1038/s41576-023-00580-2

96. Peng A, Mao X, Zhong J, Fan S, Hu Y. Single-cell multi-omics and its prospective application in cancer biology. *Proteomics* (2020) 20:1900271. doi: 10.1002/pmic.201900271

97. Nam AS, Chaligne R, Landau DA. Integrating genetic and non-genetic determinants of cancer evolution by single-cell multi-omics. *Nat Rev Genet* (2021) 22:3–18. doi: 10.1038/s41576-020-0265-5

98. Pan D, Jia D. Application of single-cell multi-omics in dissecting cancer cell plasticity and tumor heterogeneity. *Front Mol Biosci* (2021) 8:757024. doi: 10.3389/fmolb.2021.757024

99. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet* (2021) 53:1334–47. doi: 10.1038/s41588-021-00911-1

100. Chen S, Teichmann SA. Completing the cancer jigsaw puzzle with single-cell multiomics. *Nat Cancer* (2021) 2:1260–2. doi: 10.1038/s43018-021-00306-5

101. Zielinski JM, Luke JJ, Guglietta S, Krieg C. High throughput multi-omics approaches for clinical trial evaluation and drug discovery. *Front Immunol* (2021) 12:590742. doi: 10.3389/fimmu.2021.590742

102. Mehmood S, Faheem M, Ismail H, Farhat SM, Ali M, Younis S, et al. Breast cancer resistance likelihood and personalized treatment through integrated multiomics. *Front Mol Biosci* (2022) 9:783494. doi: 10.3389/fmolb.2022.783494

103. Ektefaie Y, Yuan W, Dillon DA, Lin NU, Golden JA, Kohane IS, et al. Integrative multiomics-histopathology analysis for breast cancer classification. *NPJ Breast Cancer* (2021) 7:147. doi: 10.1038/s41523-021-00357-y

104. Chen C, Gao D, Huo J, Qu R, Guo Y, Hu X, et al. Multiomics analysis reveals ct83 is the most specific gene for triple negative breast cancer and its hypomethylation is oncogenic in breast cancer. *Sci Rep* (2021) 11:1–14. doi: 10.1038/s41598-021-91290-4

105. Coria-Rodríguez H, Ochoa S, de Anda-Jáuregui G, Hernández-Lemus E. Drug repurposing for basal breast cancer subpopulations using modular network signatures. *Comput Biol Chem* (2023) 105:107902. doi: 10.1016/j.compbiolchem.2023.107902

106. Iqbal MA, Siddiqui S, Ur Rehman A, Siddiqui FA, Singh P, Kumar B, et al. Multiomics integrative analysis reveals antagonistic roles of cbx2 and cbx7 in metabolic reprogramming of breast cancer. *Mol Oncol* (2021) 15:1450–65. doi: 10.1002/1878-0261.12894

107. Chistyakov DV, Guryleva MV, Stepanova ES, Makarenkova LM, Ptitsyna EV, Goriainov SV, et al. Multi-omics approach points to the importance of oxylipins metabolism in early-stage breast cancer. *Cancers* (2022) 14:2041. doi: 10.3390/cancers14082041

108. Malik V, Kalakoti Y, Sundar D. Deep learning assisted multi-omics integration for survival and drug-response prediction in breast cancer. *BMC Genomics* (2021) 22:1–11. doi: 10.1186/s12864-021-07524-2

109. Leung KL, Verma D, Azam YJ, Bakker E. The use of multi-omics data and approaches in breast cancer immunotherapy: A review. *Future Oncol* (2020) 16:2101–19. doi: 10.2217/fon-2020-0143

110. Zhang S, Zhang J, An Y, Zeng X, Qin Z, Zhao Y, et al. Multi-omics approaches identify sf3b3 and sirt3 as candidate autophagic regulators and druggable targets in invasive breast carcinoma. *Acta Pharm Sin B* (2021) 11:1227–45. doi: 10.1016/j.apsb.2020.12.013

111. Lin H-Y, Wu H-J, Chu P-Y. Multi-omics and experimental analysis unveil theragnostic value and immunological roles of inner membrane mitochondrial protein (immt) in breast cancer. *J Trans Med* (2023) 21:189. doi: 10.1186/s12967-023-04035-4

112. Tian S, Yan L, Fu L, Zhang Z, Zhang J, Meng G, et al. A comprehensive investigation to reveal the relationship between plasmacytoid dendritic cells and breast cancer by multiomics data analysis. *Front Cell Dev Biol* (2021) 9:640476. doi: 10.3389/fcell.2021.640476

113. Fennell EM, Aponte-Collazo LJ, Pathmasiri W, Rushing BR, Barker NK, Partridge MC, et al. Multi-omics analyses reveal clpp activators disrupt essential mitochondrial pathways in triple-negative breast cancer. *Front Pharmacol* (2023) 14:1136317. doi: 10.3389/fphar.2023.1136317

114. Buttacavoli M, Di Cara G, D'Amico C, Geraci F, Pucci-Minafra I, Feo S, et al. Prognostic and functional significant of heat shock proteins (hsps) in breast cancer unveiled by multi-omics approaches. *Biology* (2021) 10:247. doi: 10.3390/biology10030247

115. Marczyk M, Patwardhan GA, Zhao J, Qu R, Li X, Wali VB, et al. Multi-omics investigation of innate navitoclax resistance in triple-negative breast cancer cells. *Cancers* (2020) 12:2551. doi: 10.3390/cancers12092551

116. Yamashita S, Hattori N, Fujii S, Yamaguchi T, Takahashi M, Hozumi Y, et al. Multi-omics analyses identify hsd17b4 methylation-silencing as a predictive and response marker of her2-positive breast cancer to her2-directed therapy. *Sci Rep* (2020) 10:15530. doi: 10.1038/s41598-020-72661-9

117. Xie B, Yuan Z, Yang Y, Sun Z, Zhou S, Fang X. Mobcdb: a comprehensive database integrating multi-omics data on breast cancer for precision medicine. *Breast Cancer Res Treat* (2018) 169:625–32. doi: 10.1007/s10549-018-4708-z

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 27 July 2023
DOI 10.3389/fonc.2023.1210125

OPEN ACCESS

EDITED BY
Iman Mamdouh Talaat,
University of Sharjah, United Arab Emirates

REVIEWED BY
Xiaoliang Shao,
First People's Hospital of Changzhou, China
Carmelo Caldarella,
Fondazione Policlinico Universitario A. Gemelli IRCCS, Italy

*CORRESPONDENCE
Bin Zhang
zbnuclmd@126.com
Chunjing Yu
chunjingyu2022@163.com
Shengming Deng
dshming@163.com

†These authors have contributed equally to this work

RECEIVED 21 April 2023
ACCEPTED 06 July 2023
PUBLISHED 27 July 2023

CITATION
Jia T, Lv Q, Cai X, Ge S, Sang S, Zhang B, Yu C and Deng S (2023) Radiomic signatures based on pretreatment 18F-FDG PET/CT, combined with clinicopathological characteristics, as early prognostic biomarkers among patients with invasive breast cancer.
Front. Oncol. 13:1210125.
doi: 10.3389/fonc.2023.1210125

Radiomic signatures based on pretreatment 18F-FDG PET/CT, combined with clinicopathological characteristics, as early prognostic biomarkers among patients with invasive breast cancer

Tongtong Jia[1†], Qingfu Lv[2†], Xiaowei Cai[3†], Shushan Ge[1], Shibiao Sang[1], Bin Zhang[1*], Chunjing Yu[4*] and Shengming Deng[1*]

[1]Department of Nuclear Medicine, the First Affiliated Hospital of Soochow University, Suzhou, China, [2]Department of General Surgery, The First Affiliated Hospital of Soochow University, Suzhou, China, [3]Department of Nuclear Medicine, The Affiliated Suqian First People's Hospital of Nanjing Medical University, Suqian, China, [4]Department of Nuclear Medicine, Affiliated Hospital of Jiangnan University, Wuxi, China

Purpose: The aim of this study was to investigate the predictive role of fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) in the prognostic risk stratification of patients with invasive breast cancer (IBC). To achieve this, we developed a clinicopathologic-radiomic-based model (C-R model) and established a nomogram that could be utilized in clinical practice.

Methods: We retrospectively enrolled a total of 91 patients who underwent preoperative ^{18}F-FDG PET/CT and randomly divided them into training (n=63) and testing cohorts (n=28). Radiomic signatures (RSs) were identified using the least absolute shrinkage and selection operator (LASSO) regression algorithm and used to compute the radiomic score (Rad-score). Patients were assigned to high- and low-risk groups based on the optimal cut-off value of the receiver operating characteristic (ROC) curve analysis for both Rad-score and clinicopathological risk factors. Univariate and multivariate Cox regression analyses were performed to determine the association between these variables and progression-free survival (PFS) or overall survival (OS). We then plotted a nomogram integrating all these factors to validate the predictive performance of survival status.

Results: The Rad-score, age, clinical M stage, and minimum standardized uptake value (SUV_{min}) were identified as independent prognostic factors for predicting PFS, while only Rad-score, age, and clinical M stage were found to be prognostic factors for OS in the training cohort. In the testing cohort, the C-R model showed superior performance compared to single clinical or radiomic models. The

concordance index (C-index) values for the C-R model, clinical model, and radiomic model were 0.816, 0.772, and 0.647 for predicting PFS, and 0.882, 0.824, and 0.754 for OS, respectively. Furthermore, decision curve analysis (DCA) and calibration curves demonstrated that the C-R model had a good ability for both clinical net benefit and application.

Conclusion: The combination of clinicopathological risks and baseline PET/CT-derived Rad-score could be used to evaluate the prognosis in patients with IBC. The predictive nomogram based on the C-R model further enhanced individualized estimation and allowed for more accurate prediction of patient outcomes.

KEYWORDS

breast cancer, radiomic, PET/CT, prognosis, nomogram, biomarker

Introduction

Breast cancer (BC) is now the leading cause of malignancy incidence and tumor-related deaths among females worldwide, surpassing lung cancer (1). The comprehensive therapy of BC, including surgery, chemotherapy, radiotherapy, and targeted treatments, has been effective in reducing locoregional or distant recurrences and prolonging survival (2–4). However, the intrinsic intratumoral heterogeneity of BC has resulted in distinct patterns of tumor progression, metastasis formation, and therapy resistance (5). Despite active therapies, some patients still develop various forms of resistance, which has not altered mortality outcomes (6, 7). Clinicians have made initial prognostic predictions and individualized therapies based on the tumor-node-metastasis (TNM) staging system and molecular classification (8, 9). However, it remains difficult to precisely predict the prognosis of patients with advanced and inoperable BC. This can result in overtreatment or undertreatment due to the high heterogeneity of BC and the complexity of treatment strategies (10).

To improve risk stratification and monitor therapeutic efficacy in BC patients, it is crucial to develop robust image-driven biomarkers. Recently, fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) has become a common diagnostic tool for BC patients, as it combines functional metabolic quantification with morphological imaging. This technique is useful for initial staging, prognostic assessments, and response monitoring (11, 12). Certain studies have indicated that preoperative metabolic parameters, such as standardized uptake values (SUVs), metabolic tumor volume (MTV), and total lesion glycolysis (TLG), serve as reliable biomarkers associated with the prognosis of BC (13, 14). However, these traditional metabolic factors may not fully capture the spatial distribution between pairs of voxels (15, 16).

Radiomics, which extracts advanced texture features from medical images to non-invasively characterize tumor heterogeneity and predict prognostic response, has emerged as a promising research topic in BC (17–19). However, few studies have investigated the predictive value of baseline PET/CT radiomics in BC prognosis (20, 21). Moreover, combining clinicopathological characteristics with radiomic biomarkers to create predictive signatures and constructing a nomogram is a prevalent and effective approach for achieving prognosis prediction and individualized management (22, 23). In the present study, we aimed to develop a predictive nomogram using the C-R model that combined clinicopathological and radiomic signatures (RSs) based on pretreatment PET/CT to estimate the survival prognosis of BC patients.

Materials and methods

Patients and follow-up

This retrospective study was approved by the medical ethics committee of the First Affiliated Hospital of Soochow University and waived additional informed consent (Trial registration number: ChiCTR2300070309). The study was conducted in compliance with the Declaration of Helsinki, and no personal information was disclosed. The total cohort of consecutive patients who were initially diagnosed with BC using ^{18}F-FDG PET/CT and confirmed by pathology in our institution between September 2016 and April 2022 were further checked by the following criteria. Inclusion criteria were as follows: a) patients who did not receive any therapy prior to the standard examination of ^{18}F-FDG PET/CT; b) patients ultimately diagnosed with invasive carcinoma of BC, including invasive ductal, lobular, or papillary carcinomas, and confirmed by puncture biopsy or surgical specimen; c) patients with available clinical records and pathological data; and d) patients with immunochemistry (IHC) examination, including estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67. Exclusion criteria were as follows: a) patients with the suboptimal quality of ^{18}F-FDG PET/CT images due to motion artifacts or abnormal biodistribution of tracer; b) primary lesions with a too small size to be outlined the volume of interest (VOI) for measurement (short-axis diameter <1

cm); c) patients with other types of tumors; d) patients confirmed to have other specific histological types of BC (sarcomas, lymphomas and so on); and e) newly diagnosed patients with a follow-up time of less than 8 months. The enrolled patients were randomly divided into training and validation cohorts at a 7:3 ratio using computer-generated random numbers.

All patients were followed up from the confirmed time of primary diagnosis until the cut-off date of December 30, 2022, using outpatient review data and telephone follow-ups. Progression-free survival (PFS) was defined as the interval between the date of first diagnosis and the first relapse, tumor progression, death, or the last follow-up. Overall survival (OS) was defined as the interval between the date of first diagnosis and death from any cause or the last follow-up. The study endpoints were PFS and OS.

Image acquisition and reconstruction

According to guidelines from the European Association of Nuclear Medicine (EANM), patients must fast for at least 4 h and ensure that their plasma glucose level is lower than 11.0 mmol/L (about 200 mg/dL) prior to undergoing the 18F-FDG PET/CT procedure in clinical studies (24). Approximately 40-60 min after intravenous injection of ^{18}F-FDG (4.07-5.55 MBq/kg), patients were scanned using an integrated PET/CT scanner (Discovery STE, General Electric Medical Systems, Milwaukee, WI, USA) to acquire images from the base of the skull to the midthigh. Low-dose (140 kV, 120 mA) CT images were used for subsequent attenuation correction and anatomic localization of PET images, with acquisition parameters including a transaxial field of view of 70 cm, pitch of 1.75, rotation time of 0.8 s, and slice thickness of 3.75 mm. PET image acquisition was performed at 2-3 min per bed position, with a total of 8-10 bed positions. During image reconstruction, the ordered subset expectation maximization algorithm was used (two iterations and eight subsets) to ensure that reconstructed voxel sizes were within 3.0-4.0 mm in any direction.

Clinicopathological evaluation

The study collected several clinical factors, including age, menopausal status, tumor location, initial TNM stage, treatment strategies, and diagnosis time. Specimens obtained from core needle biopsy and excisional biopsy were fixed in formalin solution, embedded in paraffin, and stained with hematoxylin and eosin (H&E staining). Stained sections were evaluated by independently two experienced pathologists to confirm the histopathological type. The expressions of ER, PR, HER2, and Ki-67 were detected using IHC. ER and PR were considered positive if there was a proportion of at least 1% of nuclear staining. HER2 status was confirmed using a combination of IHC scores and fluorescence *in situ* hybridization (FISH), where a positive result was defined as IHC 2+ and FISH gene amplification or IHC 3+ (25). Ki-67 cell nuclear staining of ≥ 30% was considered a high expression.

VOI segmentation and radiomic feature extraction

The Local Image Features Extraction (LifeX) package (version 7.0.0, available at https://www.lifexsoft.org/) was used to automatically match and fuse PET and CT images in DICOM format for quantitative PET/CT analysis (26). Two experienced nuclear medicine physicians, who were blinded to the clinicopathological results, manually segmented the transaxial VOI layer by layer. The VOI was defined by integrating abnormal uptake of ^{18}F-FDG on PET and abnormal density on CT, which was optimized by setting a threshold of 40% of the SUVmax to ensure reproducibility (16). Once matched, the RSs (four or six conventional features, nine first-order features, and 32 high-order features) of the CT or PET images could be automatically extracted from the same VOI. To avoid overfitting, significant RSs of the training cohort were selected using correlation analysis, least absolute shrinkage and selection operator (LASSO) regression algorithm, and univariable Cox analysis before model construction (17, 27). Finally, 10-fold cross-validation was used to ensure the robustness of the optimal features.

Model construction and validation

The Rad-score was calculated using a linear fitting formula, which involved multiplying the remaining features with their respective weighted coefficients to create a radiomic model. Based on the optimal threshold value of the Rad-score, as determined through receiver operating characteristic (ROC) curve analysis, the cohorts were divided into high-risk and low-risk groups. In addition to clinicopathological factors, the Rad-score was further evaluated using Kaplan-Meier (KM) analysis and log-rank tests. All significant factors were entered into a multivariable Cox proportional hazards regression to identify the final subset of prognostic factors. Finally, the radiomic and clinical nomograms were evaluated in the training cohort and then validated in the testing cohort. To evaluate the discriminative ability, calibration, and clinical usefulness of the models, we employed the Harrell concordance index, calibration curves, and decision curve analysis (DCA), respectively (28).

Statistical analysis

Statistical data were calculated and analyzed using IBM SPSS Statistics (version 26.0, IBM Corp), Python (version 3.0, https://www.python.org), MedCalc software (MedCalc Software, Ostend, Belgium), and R (version 4.2.1, http://www.R-project.org). The normality and homogeneity of variance for continuous data were

evaluated using the Kolmogorov-Smirnov test and Levene's test, respectively. The independent t-test and Mann-Whitney U test were then used to evaluate any differences in baseline characteristics between the training and testing sets. Meanwhile, the Chi-square test and Fisher's exact test were applied to analyze categorical variables. For the final survival analysis, quantitative variables were transformed into dichotomous variables to conduct further univariate and multivariate Cox analyses, as well as to estimate hazard ratios (HRs). A two-sided p-value of less than 0.05 was considered statistically significant. Figure 1 provides an overview of the study's workflow.

Results

Patient population

In this retrospective study, 91 patients with BC pretreatment who met the inclusion criteria were selected (in Figure 2). Table 1 summarizes the clinical characteristics of the training and validation cohorts, and there were no statistically significant biases in patient distribution between the two cohorts (all p > 0.05). During the follow-up period from the time of primary diagnosis (median: 20 months, range: 4 - 95 months), 42 out of 91 patients (46.2%) developed PFS endpoints, and 27 out of 91 patients (29.7%) died. Patients without any PFS events who survived at least 8 months after cancer diagnosis were considered as the control group (n = 49). The event rates of PFS (46.0% and 46.4%, respectively) and OS (30.2% and 28.6%, respectively) were not significantly different between the two cohorts, indicating a balanced distribution.

RS selection and Rad-score construction

The intra-class correlation coefficient (ICC) of the extracted radiomic features was above 0.75 between the two experienced nuclear medicine physicians, and they reached a final agreement in consensus. Pearson correlation analysis between the RSs was visualized in Figure 3, and several strongly correlated clusters were circled by black boxes. After using LASSO for dimensionality reduction to remove redundant features (with zero coefficients), the most significant prognostic signatures were selected to calculate the Rad-score in the training cohort. Finally, a total of four RSs consisting of two CT RSs [SHAPE_Volume(mL)$_{CT}$ and GLZLM_GLNU$_{CT}$] and two PET RSs (NGLDM_Coarseness$_{PET}$ and GLZLM_GLNU$_{PET}$) were chosen for predicting PFS. With regard to OS, two CT RSs (NGLDM_Coarseness$_{CT}$ and NGLDM_Contrast$_{CT}$) and three PET RSs (SHAPE_Sphericity$_{PET}$, NGLDM_Coarseness$_{PET}$, and GLZLM_GLNU$_{PET}$) were involved in the predictive formula as follows: Rad-score$_{PFS}$= - 0.0941130 × SHAPE_Volume(mL)$_{CT}$- 0.033147 × GLZLM_GLNU$_{CT}$ + 0.057186 × NGLDM_Coarseness$_{PET}$- 0.105334 × GLZLM_GLNU$_{PET}$. Rad-score$_{OS}$ = 0.106539 × NGLDM_Coarseness$_{CT}$ + 0.050655×

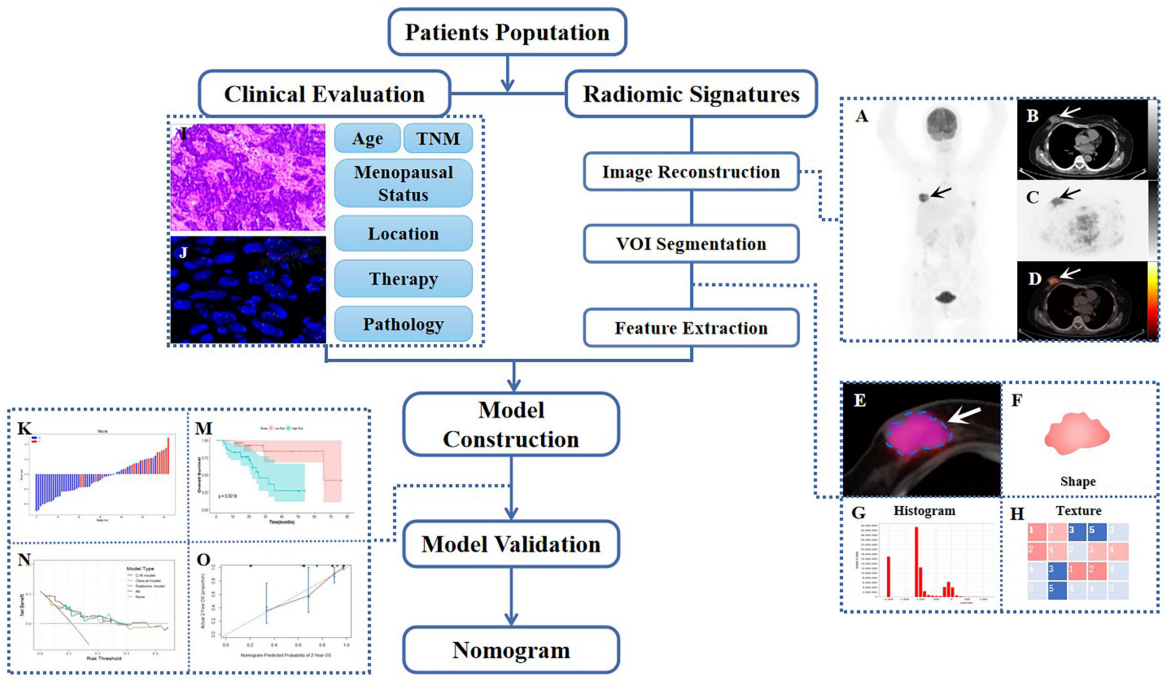


FIGURE 1
Workflow diagram of this retrospective study. The typical example of PET/CT image reconstruction in a patient with BC, whole-body maximum intensity projection (MIP) of PET image **(A)**, axial views of low-dose CT, PET or PET/CT infusion scans **(B–D)**. Arrows point to tumor lesions. The dotted line circles the VOI in magnified PET/CT image **(E)**, from which first-order **(F, G)** and high-order **(H)** radiomic features were extracted. The IHC and FISH of representatively pathological section **(I, J)**. Waterfall plot of radiomic score (Rad-score) and KM curve of survival analysis during model construction **(K, M)**, DCA, and calibration curve during model validation **(N, O)**.

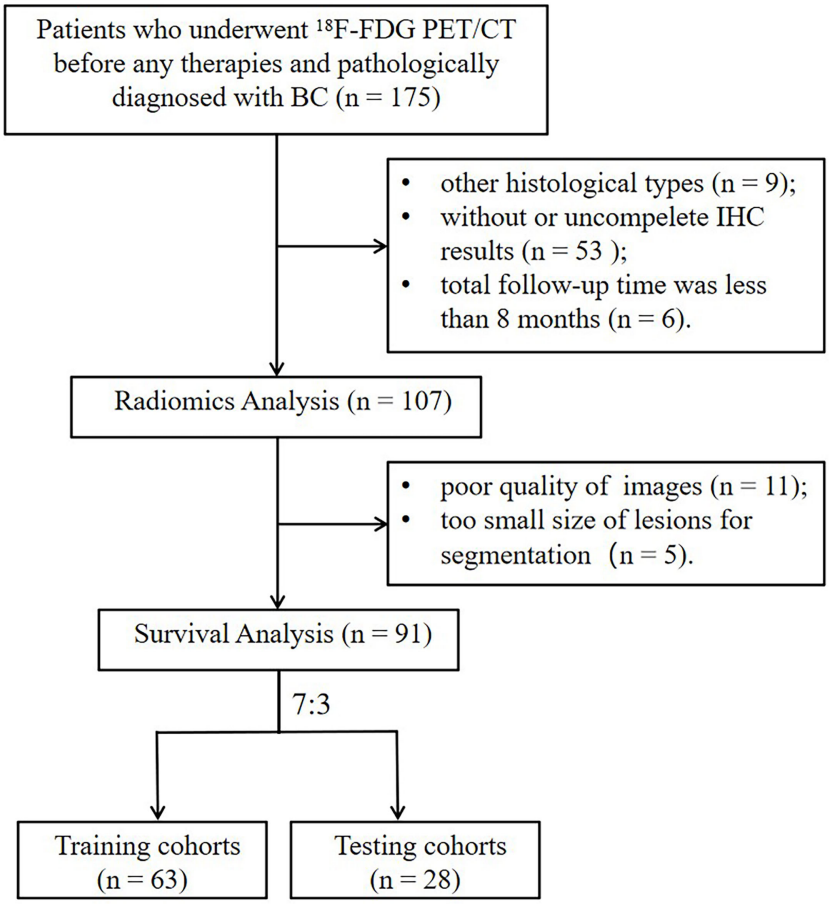


FIGURE 2
The flowchart of the selection process according to eligibility and exclusive criteria. BC, breast cancer; IHC, immunohistochemistry.

TABLE 1 The clinicopathological features of the enrolled population.

	Total (n=91)	Training (n=63)	Testing (n=28)	t/x^2	p
Age (years)	53.33 ± 13.46	53.05 ± 13.47	54.00 ± 13.66	-0.307	0.760
Menopausal Status				0.266	0.606
Premenopausal	34 (37.4%)	25 (39.1%)	9 (33.3%)		
Postmenopausal	57 (62.6%)	39 (60.9%)	18 (66.7%)		
Tumor Location				3.650	0.056
Left	51 (56.0%)	40 (62.5%)	11 (40.7%)		
Right	40 (44.0%)	24 (37.5%)	16 (59.3%)		
Histological Type				0.809	0.368
IDC	82 (90.1%)	56 (87.5%)	26 (96.3%)		
*Other	9 (9.9%)	8 (12.5%)	1 (3.7%)		
Clinical Stage				1.866	0.393
I-II	22 (24.2%)	18 (28.1%)	4 (14.8%)		
III	34 (37.4%)	23 (35.9%)	11 (40.7%)		

(Continued)

TABLE 1 Continued

	Total (n=91)	Training (n=63)	Testing (n=28)	t/x^2	p
IV	35 (38.4%)	23 (35.9%)	12 (44.4%)		
Clinical T Stage				0.646	0.886
cT1	15 (16.5%)	11(17.2%)	4 (14.8%)		
cT2	38 (41.8%)	28 (43.8%)	10 (37.0%)		
cT3	3 (3.3%)	2 (3.1%)	1 (3.7%)		
cT4	35 (38.4%)	23 (35.9%)	12 (44.4%)		
Clinical N Stage				2.949	0.223
cN0	16 (17.6%)	14 (21.9%)	2 (7.4%)		
cN1-2	39 (42.8%)	25 (39.1%)	14 (51.9%)		
cN3	36 (39.6%)	25 (39.1%)	11 (40.7%)		
Clinical M Stage				0.187	0.665
cM0	57 (62.6%)	41 (64.1%)	16 (59.3%)		
cM1	34 (37.4%)	23 (35.9%)	11 (40.7%)		
ER Status				0.033	0.856
Positive	56 (61.5%)	39 (60.9%)	17 (63.0%)		
Negative	35 (38.5%)	25 (39.1%)	10 (37.0%)		
PR Status				0.892	0.345
Positive	37 (40.7%)	24 (37.5%)	13 (48.1%)		
Negative	54 (59.3%)	40 (62.5%)	14 (51.9%)		
HER2 Status				1.133	0.287
Positive	31 (34.1%)	24 (37.5%)	7 (25.9%)		
Negative	60 (65.9%)	40 (62.5%)	20 (74.1%)		
Molecular Subtype				0.707	0.892
HR+/HER2-	43 (47.2%)	29 (45.3%)	14 (51.9%)		
HR+/HER2+	14 (15.4%)	10 (15.6%)	4 (14.8%)		
HER2+	18 (19.8%)	14 (21.9%)	4 (14.8%)		
TNBC	16 (17.6%)	11 (17.2%)	5 (18.5%)		
Ki-67				0.193	0.66
<30%	30 (33.0%)	22 (34.4%)	8 (29.6%)		
≥30%	61 (67.0%)	42 (65.6%)	19 (70.4%)		
Treatment				1.059	0.589
NAC	21 (23.0%)	13 (20.3%)	8 (29.6%)		
PCT	27 (29.7%)	19 (29.7%)	8 (29.6%)		
Other	43 (47.3%)	32 (50.0%)	11 (40.7%)		

Descriptive statistics were summarized with mean±standard deviation and analyzed by independent t-test. Categorical variables were shown as numbers and percentages and analyzed by Pearson's Chi-square test or Fisher's exact test. *Other: including invasive lobular or papillary carcinomas. Abbreviations: IDC, invasive ductal carcinoma; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; TNBC, triple-negative breast cancer; NAC, neoadjuvant chemotherapy; PCT, postoperative chemotherapy.

$NGLDM\_Contrast_{CT}$ + 0.049265×$SHAPE\_Sphericity_{PET}$ + 0.097559×$NGLDM\_Coarseness_{PET}$ - 0.013341×$GLZLM\_GLNU_{PET}$. The median Rad-score calculated using the above formula was 0.2923 (range: 0.0475 - 0.5508) for PFS and 0.4227 (range: 0.1636 - 0.9299) for OS. Moreover, the optimum threshold generated from the ROC analysis of PFS and OS was 0.3776 and 0.3197 in the training set, respectively. Table 2 shows that the area under the curve (AUC) was 0.670 (95% CI: 0.541 - 0.782) for PFS and 0.706 (95% CI: 0.579 -

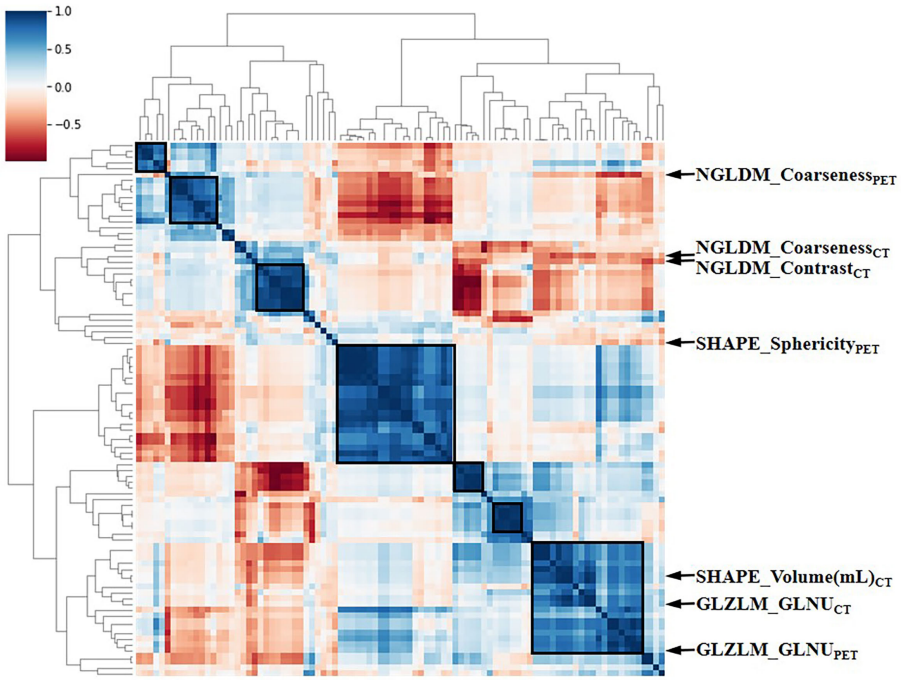


FIGURE 3
The heat map of Pearson correlation analysis among 97 radiomic features. Correlation coefficients were displayed by color scale. Boxes circled the clusters that were closely related. Representative RSs were marked by arrows. RSs, radiomic signatures; NGLDM, Neighboring Gray-level dependence matrix; GLZLM, Gray-Level Zone Length Matrix; GLNU, Gray-Level Non-Uniformity for zone.

0.814) for OS. Accordingly, patients were classified into low-score and high-score groups (Rad-score$_{PFS}$> 0.3776; Rad-score$_{OS}$> 0.3197). Then, Rad-score was incorporated into the subsequent survival analysis as a potential prognostic biomarker. Univariate Cox regression indicated that Rad-score was closely associated with both PFS (p=0.039) and OS (p=0.0085) and was shown by KM survival curves (Figures 4G, 5E).

Combined model construction

The log-rank test was used to perform a univariate analysis of clinicopathological and radiomic features in predicting the prognosis of BC. Additionally, characteristics with a p-value <0.05 were included in the final multivariate Cox regression (Table 3). In the univariate analysis of PFS, it was found that age, menopausal status, clinical stage, clinical N stage, clinical M stage, ER status, PR status, Ki-67, SUV parameters (SUV$_{max}$, SUV$_{min}$, SUV$_{mean}$, and SUV$_{peak}$), and Rad-score were potential biomarkers. Among them, age (HR = 3.532, P = 0.013), clinical M stage (HR = 2.977, P = 0.017), SUV$_{min}$ (HR = 4.240, P = 0.001), and Rad-score (HR = 2.660, P = 0.044) were independent factors for prognosis in the multivariate proportional hazards model (Figure 6). Meanwhile, age (HR = 5.644, P = 0.013), clinical M stage (HR = 3.499, P = 0.057), and Rad-score (HR = 3.627, P = 0.026) were selected from significant factors (age, menopausal status, clinical stage, clinical

TABLE 2 The Harrell's C-index and AUC results in the training and validation cohorts.

	Training cohort		Validation cohort	
PFS	C-index (95% CI)	AUC (95% CI)	C-index (95% CI)	AUC (95% CI)
Clinical Model	0.761 (0.666 - 0.857)	0.739 (0.614 - 0.841)	0.772 (0.656 - 0.888)	0.816 (0.620 - 0.938)
Radiomic Model	0.613 (0.492 - 0.735)	0.670 (0.541 - 0.782)	0.674 (0.532 - 0.815)	0.599 (0.394 - 0.781)
C-R model	0.786 (0.697 - 0.875)	0.787 (0.667 - 0.880)	0.816 (0.685 - 0.947)	0.830 (0.636 - 0.946)
OS				
Clinical model	0.794 (0.703 - 0.885)	0.731 (0.605 - 0.834)	0.824 (0.666-0.981)	0.780 (0.579 - 0.915)
Radiomic model	0.730 (0.640 - 0.819)	0.706 (0.579 - 0.814)	0.754 (0.584-0.923)	0.711 (0.505 - 0.867)
C-R model	0.878 (0.816 - 0.940)	0.789 (0.669 - 0.881)	0.882 (0.781-0.984)	0.859 (0.671 - 0.962)

C-index, concordance index; CI, confidence interval; AUC, area under the curve; C-R model, clinicopathologic-radiomic-based model.

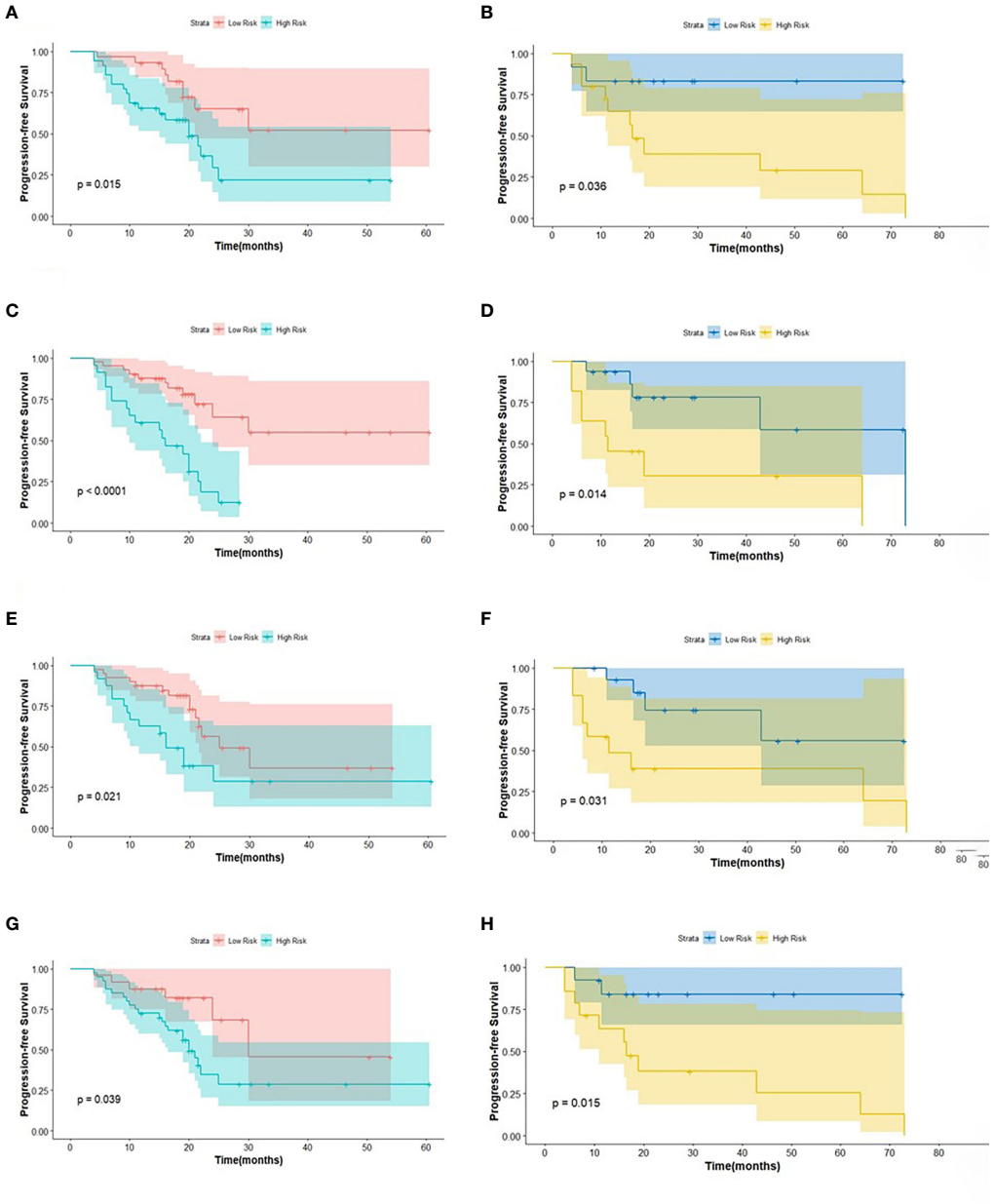


FIGURE 4
The KM analysis of independent factors for predicting PFS in the training cohort and in the testing cohort (**A, B**: age; **C, D**: clinical M stage; **E, F**: SUV_{min}; **G, H**: Rad-score).

N stage, clinical M stage, PR status, and Rad-score) to construct the integrated model for predicting OS. Furthermore, the remaining factors had a strong predictive value for PFS and OS in the training cohort, which was similar to the results observed in the validation cohort through KM analyses (Figures 4, 5).

Model validation and assessments

To evaluate the probability of 1-, 2-, and 3-year PFS and OS, we established nomograms for an integrated model that incorporated the most valuable clinical and imaging parameters (Figure 7). For the training cohort, the concordance index (C-index) and AUC of the C-R model were 0.786 (95% CI: 0.697 - 0.875) and 0.787 (95% CI: 0.667 - 0.880), respectively, for PFS prediction. These values were superior to those of the single clinical or radiomic models, demonstrating the good predictive accuracy of the C-R model (Table 2). Similar performance was observed for OS, where the C-index and AUC of the C-R model were 0.878 (95% CI: 0.816 - 0.940) and 0.789 (95% CI: 0.669 - 0.881), respectively, and were higher than those of other models. The C-R model was successfully validated in the testing set. In the validation cohort, the C-index of the C-R model was 0.816 (95% CI: 0.685 - 0.947) for predicting PFS and 0.882 (95% CI: 0.781 - 0.984) for predicting OS. Using ROC

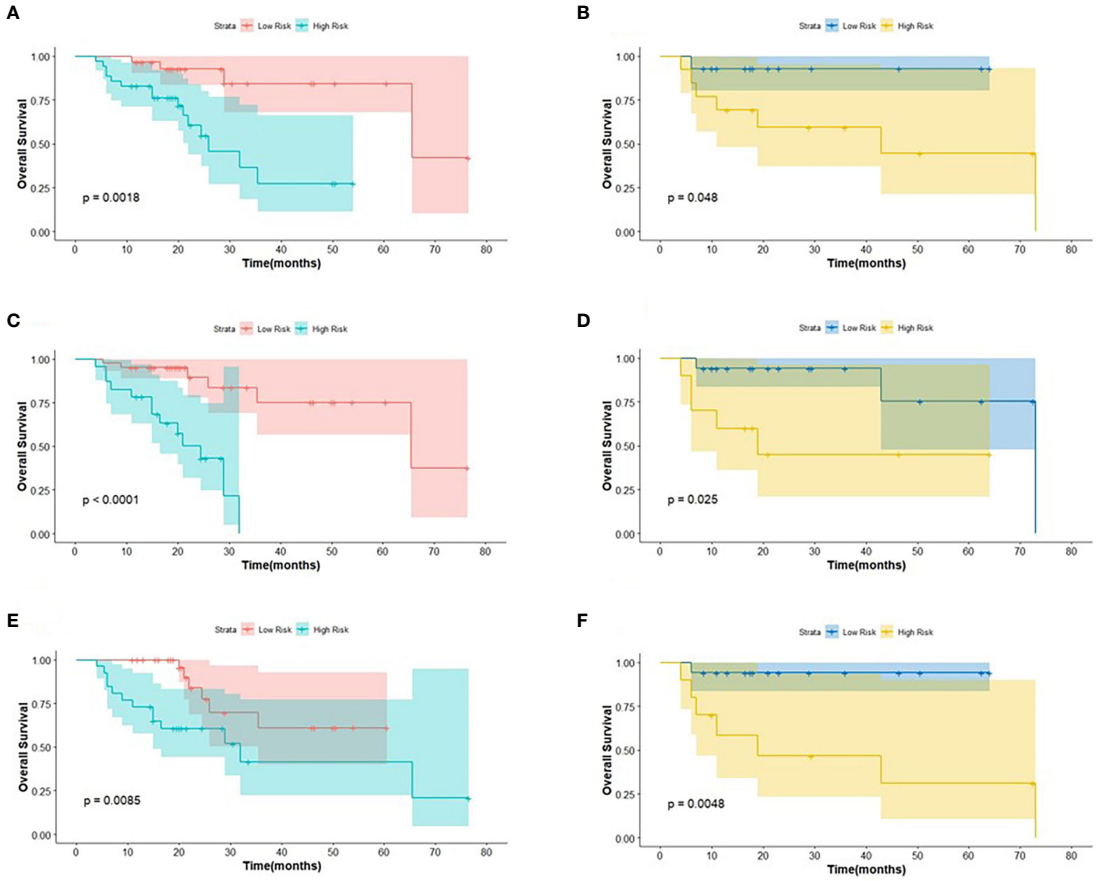


FIGURE 5
The KM analysis of meaningful biomarkers for predicting OS in the training cohort and in the testing cohort (**A, B**: age; **C, D**: clinical M stage; **E, F**: Rad-score).

curve analysis, we found that the C-R model had a higher AUC compared with the other two models for predicting PFS (AUC = 0.830, 95% CI: 0.636 - 0.946) and OS (AUC = 0.859, 95% CI: 0.671 - 0.962) (Table 2). Figure 8 presents the DCA and calibration curve of the nomogram (2-year predictive probability of the C-R model) for PFS and OS. The DCA for the PFS nomogram indicated that the net benefit of the clinical and C-R models was slightly higher compared with the radiomic model within reasonable threshold probabilities (Figure 8A). Regarding OS, while there was no significant difference observed among the three models, all of them provided overall net benefits in clinical application (Figure 8B). Both calibration curves of the nomograms for PFS and OS showed accurate discrimination between prediction and observation in the testing set (Figures 8C, D).

TABLE 3 The results of the univariate and multivariate Cox regression analysis.

Variable	PFS				OS			
	Log Rank		Cox Regression		Log Rank		Cox Regression	
	x2	p	HR (95%CI)	p	x2	p	HR (95%CI)	p
Age (years)	5.879	0.015*	3.532(1.301-9.593)	0.013	9.696	0.002*	5.644(1.430-22.277)	0.013
Menopausal Status	4.524	0.033*			4.644	0.031*		
Tumor Location	0.233	0.629			0.078	0.780		
Histologic Type	2.992	0.084			2.16	0.142		
Clinical Stage	16.067	<0.001*			19.806	<0.001*		

(Continued)

TABLE 3 Continued

Variable	PFS				OS			
	Log Rank		Cox Regression		Log Rank		Cox Regression	
	x2	p	HR (95%CI)	p	x2	p	HR (95%CI)	p
Clinical T Stage	1.273	0.736			6.472	0.091		
Clinical N Stage	9.238	0.010*			16.255	<0.001*		
Clinical M Stage	15.472	<0.001*	2.977(1.217-7.283)	0.017	19.46	<0.001*	3.499(0.962-12.727)	0.057
ER Status	3.902	0.048*			3.678	0.055		
PR Status	5.084	0.024*			5.357	0.021*		
HER2 Status	0.027	0.870			0.035	0.853		
Molecular Subtype	4.626	0.201			4.763	0.190		
Ki-67	4.174	0.041*			1.157	0.282		
Treatment	5.994	0.050			4.931	0.085		
SUV_{max}	7.058	0.008*			2.028	0.154		
SUV_{min}	5.359	0.021*	4.240(1.814-9.910)	0.001	2.598	0.107		
SUV_{mean}	3.950	0.047*			1.686	0.194		
SUV_{peak}	6.130	0.013*			1.685	0.194		
MTV	1.772	0.183			3.848	0.050		
TLG	2.891	0.089			3.758	0.053		
Rad-score	4.261	0.039*	2.660(1.029-6.880)	0.044	6.930	0.008*	3.627(1.171-11.241)	0.026

The clinicopathological factors and Rad-score were analyzed by the log-rank method in univariate analysis, and then characteristics with $p < 0.05$ were taken into the multivariate Cox regression to construct the final model. HR, hazard ratio; CI, confidence interval; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; SUV, standardized uptake value; MTV, metabolic tumor volume; TLG, total lesion glycolysis; Rad-score, radiomic score. *$P < 0.05$.

Discussion

Recent studies have revealed various clinical endpoints among individuals with BC, and precise prediction based on non-invasive machine learning methods has become increasingly prevalent in prognostic research (29–31). These promising results have encouraged the emergence of combination models based on clinical, histological, and imaging features to better meet clinical requirements (32, 33). In this study, we established novel nomograms that reflected the underlying role of clinicopathological and radiomic biomarkers extracted from ^{18}F-FDG PET/CT to estimate the outcomes of BC patients. Regarding the prediction of PFS, the C-index of the C-R model, clinical model, and radiomic model was 0.786, 0.761, and 0.613 in the training group and 0.878, 0.794, and 0.730 in the testing group, respectively. With regard to the endpoint of OS in the training test, the C-index of the three models was 0.816, 0.772, and 0.674, and it became 0.882, 0.731, and 0.706 in the testing test, respectively. These findings suggested the superior predictive performance of the combination model in both the training and validation cohorts compared to other single models.

Traditionally, clinicians have relied on TNM staging to make prognostic assessments based on physical examinations and clinical symptoms. With the emergence of PET/CT scans, these assessments have been improved. However, patients at the same stage still show varying outcomes, even when treated with similar strategies, as the stage changes. Therefore, there is an urgent clinical need for an accurate predictive method of prognosis, particularly for highly heterogeneous malignancies, such as BC. In our retrospective study, metabolic parameters of PET/CT played an invaluable role in predicting the risk of disease recurrence. Shingo Baba et al. have revealed a correlation between higher SUVs extracted from PET images and a poor prognosis of BC. On the other hand, Evangelista et al. have found that MTV and TLG are independent factors in predicting BC recurrence, while SUV_{max} demonstrates poor prognostic performance. However, it is important to note that metabolic parameters can be influenced by various physiological and technical factors. In the end, only SUV_{min} is identified as independently associated with PFS.

To comprehensively quantify tumor heterogeneity, radiomics can provide more detailed information on the tumor microenvironment beyond visual features, allowing for the reflection of multiple clinical endpoints. High-dimensional features, such as NGLDM and GLZLM, have been found to be associated with survival time in various tumor types and have been used in the construction of the Rad-score (34–36). These studies have also confirmed that Rad-score is an independent biomarker for predicting survival status. In our present study, $NGLDM_{Coarseness}$, $NGLDM_{Contrast}$, and $GLZLM_{GLNU}$ were selected using LASSO regression and used to calculate the Rad-score, which was found to be a meaningful predictor of both PFS and OS. As far as we know,

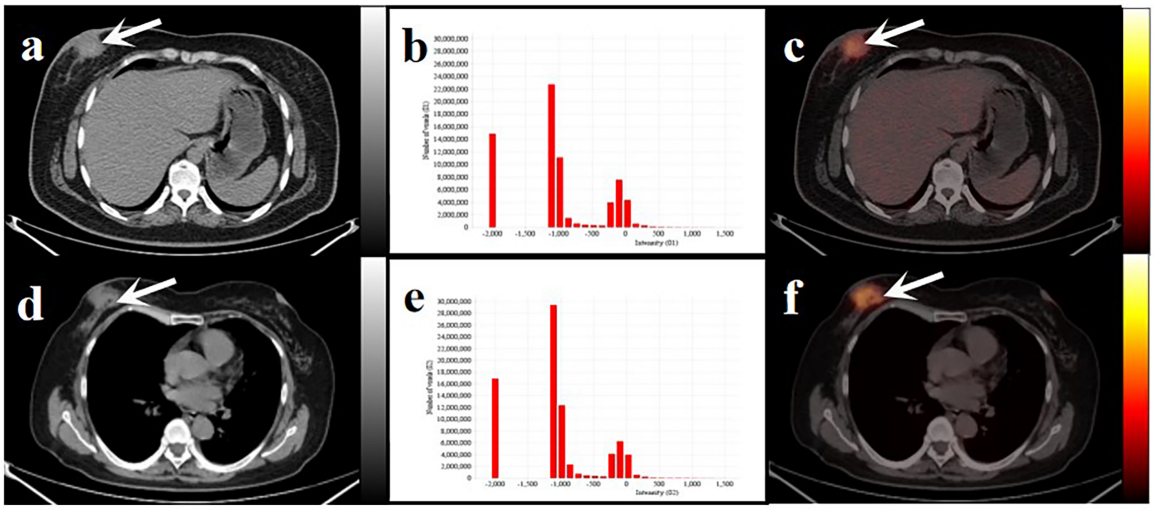


FIGURE 6
Two women with similar lesions were initially diagnosed with BC. Axial low-dose CT images **(A, D)** and infusion PET/CT images **(C, F)**. Black and white arrows point to the primary lesions located on the right breast. The histograms of CT intensity for the VOI **(B, E)**. The former (age: 32 years, IDC: grade 2, molecular type: HR+/HER2-, clinical stage: T4N1M0, Rad-score: 0.1647) who underwent neoadjuvant chemotherapy didn't have any obvious disease progression during the 19-month follow-up. The latter (age:61 years, IDC: grade 2, molecular type: TNBC, clinical stage: T4N3M1, Rad-score: 0.3089) underwent postoperative chemotherapy and died after 20 months. These primary lesions were similar on PET/CT images but showed significant differences in the histograms of the radiomic features and clinical outcomes.

our study is one of the few articles that have focused on predicting BC prognosis using PET/CT imaging and histology.

In addition, age and clinical M stage were also identified as prominent predictors for both PFS and OS in our study. Specifically, our study found that patients over the age of 50 and those with distant metastases (clinical M stage: M1) and higher Rad-score were more likely to experience earlier disease progression or even death. Moreover, recent research has confirmed that visually represented nomograms based on clinicopathologic risk factors and rad-score can significantly contribute to the prediction of prognosis (37). The C-index, along with its 95% CI, DCA, and calibration curve in the testing cohort, can provide a more comprehensive assessment, including discrimination, clinical applicability, and calibration of the nomogram for the predictive model (38). Therefore, we aimed to establish an integrated model visualized by a nomogram to assess the potential prognostic value of BC patients by combining PET/CT-based radiomics with clinical features.

However, our work has some limitations that need to be acknowledged. Firstly, our limited population needs to be taken into account, although it is homogeneous in terms of histology types. Our assumptions need to be further strengthened in multicenter prospective cohorts. Secondly, previous literature has shown that radiological signatures derived from PET/CT may be influenced by the equipment and software used for image acquisition, reconstruction, and analysis (39). All patients underwent PET/CT examination with consistent scanners in our study. Thirdly, although the inter-observer agreement was repeatable (ICC > 0.75), selection bias was inevitable. Lastly, the survival study highly depended on follow-up time, and adequate interviews will be necessary to validate our results.

Conclusion

In conclusion, we established a complex model that incorporated both clinicopathologic and radiomic factors, which

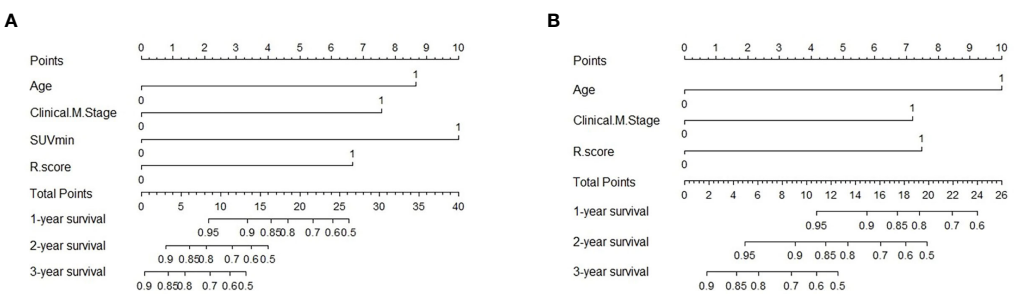


FIGURE 7
Predictive nomogram of C-R model for PFS **(A)** and OS **(B)** in the testing cohort. Summed by the points of every risk factor, the final points are located on the Total Point axis. R.score, radiomic score.

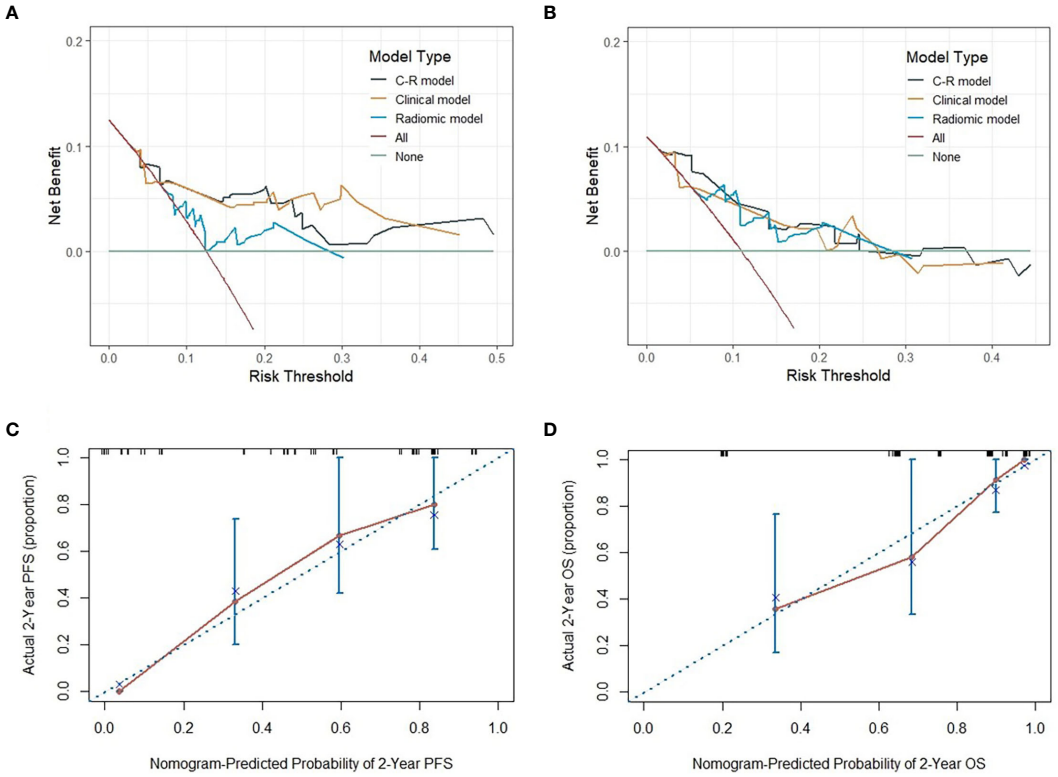


FIGURE 8
The DCA of three models for PFS and OS in the testing cohort **(A, B)**. The y-axis measures the net benefit, which is calculated by summing the benefits (true positive results) and subtracting the harms (false-positive results). The calibration curves of C-R models' nomograms for the 2-year probability of PFS and OS **(C, D)**. The dashed line indicated a perfect match between the actual probability (y-axis) and the nomogram-predicted probability of 2-year PFS and OS (x-axis).

could potentially serve as a biomarker for risk stratification of prognosis in patients with invasive BC. Our strategy demonstrated not only a great net benefit at a large range of threshold probabilities but also accurate discrimination in clinical applications.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by the First Affiliated Hospital of Soochow University. The ethics committee waived the requirement of written informed consent for participation.

Author contributions

TJ, QL, XC, and SD conceptualized and designed the study. SG and SS performed analysis. BZ and CY interpreted the data. TJ and QL drafted the manuscript. SS and SD revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the National Natural Science Foundation of China (Grant No. 81601522 and 61927801), Gusu Health Talent Program (No. GSWS2020013), Suzhou Science and Education Youth Project (No. KJXW2021004), Project of State Key Laboratory of Radiation Medicine and Protection, Soochow University (No. GZK1202127), NHC Key Laboratory of Nuclear Technology Medical Transformation (MIANYANG CENTRAL HOSPITAL) (No. 2021HYX023 and 2021HYX029), the Suqian Sci & Tech Program (Grant No.K202107), and the Natural Science Foundation of the First Affiliated Hospital of Soochow University (No. BXQN202114).

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.

Publisher's note

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660

2. Whelan TJ, Olivotto IA, Parulekar WR, Ackerman I, Chua BH, Nabid A, et al. MA.20 study investigators. Regional nodal irradiation in early-stage breast cancer. *N Engl J Med* (2015) 373:307–16. doi: 10.1056/NEJMoa1415340

3. Tryfonidis K, Senkus E, Cardoso MJ, Cardoso F. Management of locally advanced breast cancer-perspectives and future directions. *Nat Rev Clin Oncol* (2015) 12:147–62. doi: 10.1038/nrclinonc.2015.13

4. Early Breast Cancer Trialists' Collaborative group (EBCTCG). Trastuzumab for early-stage, HER2-positive breast cancer: a meta-analysis of 13 864 women in seven randomised trials. *Lancet Oncol* (2021) 22:1139–50. doi: 10.1016/S1470-2045(21)00288-6

5. Lüönd F, Tiede S, Christofori G. Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression. *Br J Cancer* (2021) 125:164–75. doi: 10.1038/s41416-021-01328-7

6. Harbeck N, Gnant M. Breast cancer. *Lancet* (2017) 389:1134–50. doi: 10.1016/S0140-6736(16)31891-8

7. Gradishar WJ, Anderson BO, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast cancer, version 3.2020, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* (2020) 18:452–78. doi: 10.6004/jnccn.2020.0016

8. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin* (2017) 67:93–9. doi: 10.3322/caac.21388

9. Cianfrocca M, Gradishar W. New molecular classifications of breast cancer. *CA Cancer J Clin* (2009) 59:303–13. doi: 10.3322/caac.20029

10. Wang X, Xie T, Luo J, Zhou Z, Yu X, Guo X. Radiomics predicts the prognosis of patients with locally advanced breast cancer by reflecting the heterogeneity of tumor cells and the tumor microenvironment. *Breast Cancer Res* (2022) 24:20. doi: 10.1186/s13058-022-01516-0

11. Groheux D, Cochet A, Humbert O, Alberini JL, Hindié E, Mankoff D. ^{18}F-FDG PET/CT for staging and restaging of breast cancer. *J Nucl Med* (2016) 57 Suppl;1:17S–26S. doi: 10.2967/jnumed.115.157859

12. Ulaner GA. PET/CT for patients with breast cancer: where is the clinical impact? *AJR Am J Roentgenol* (2019) 213:254–65. doi: 10.2214/AJR.19.21177

13. Baba S, Isoda T, Maruoka Y, Kitamura Y, Sasaki M, Yoshida T, et al. Diagnostic and prognostic value of pretreatment SUV in 18F-FDG/PET in breast cancer: comparison with apparent diffusion coefficient from diffusion-weighted MR imaging. *J Nucl Med* (2014) 55:736–42. doi: 10.2967/jnumed.113.129395

14. Evangelista L, Cervino AR, Ghiotto C, Saibene T, Michieletto S, Fernando B, et al. Could semiquantitative FDG analysis add information to the prognosis in patients with stage II/III breast cancer undergoing neoadjuvant treatment? *Eur J Nucl Med Mol Imaging* (2015) 42:1648–55. doi: 10.1007/s00259-015-3088-4

15. Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, et al. Impact of image reconstruction settings on texture features in 18F-FDG PET. *J Nucl Med* (2015) 56:1667–73. doi: 10.2967/jnumed.115.156927

16. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging* (2019) 46:2638–55. doi: 10.1007/s00259-019-04391-8

17. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141

18. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present… any future? *Eur J Nucl Med Mol Imaging* (2017) 44:151–65. doi: 10.1007/s00259-016-3427-0

19. Valdora F, Houssami N, Rossi F, Calabrese M, Tagliafico AS. Rapid review: radiomics and breast cancer. *Breast Cancer Res Treat* (2018) 169:217–29. doi: 10.1007/s10549-018-4675-4

20. Antunovic L, De Sanctis R, Cozzi L, Kirienko M, Sagona A, Torrisi R, et al. PET/CT radiomics in breast cancer: promising tool for prediction of pathological response to neoadjuvant chemotherapy. *Eur J Nucl Med Mol Imaging* (2019) 46:1468–77. doi: 10.1007/s00259-019-04313-8

21. Li P, Wang X, Xu C, Liu C, Zheng C, Fulham MJ, et al. 18F-FDG PET/CT radiomic predictors of pathologic complete response (pCR) to neoadjuvant chemotherapy in breast cancer patients. *Eur J Nucl Med Mol Imaging* (2020) 47:1116–26. doi: 10.1007/s00259-020-04684-3

22. Jiang Y, Yuan Q, Lv W, Xi S, Huang W, Sun Z, et al. Radiomic signature of 18F fluorodeoxyglucose PET/CT for prediction of gastric cancer survival and chemotherapeutic benefits. *Theranostics* (2018) 8:5915–28. doi: 10.7150/thno.28018

23. Li Y, Zhang Y, Fang Q, Zhang X, Hou P, Wu H, et al. Radiomics analysis of [18F]FDG PET/CT for microvascular invasion and prognosis prediction in very-early- and early-stage hepatocellular carcinoma. *Eur J Nucl Med Mol Imaging* (2021) 48:2599–614. doi: 10.1007/s00259-020-05119-9

24. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging* (2015) 42:328–54. doi: 10.1007/s00259-014-2961-x

25. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: american society of clinical oncology/college of american pathologists clinical practice guideline focused update. *J Clin Oncol* (2018) 36:2105–22. doi: 10.1200/JCO.2018.77.8738

26. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res* (2018) 78:4786–9. doi: 10.1158/0008-5472.CAN-18-0125

27. Wang S, Wei Y, Li Z, Xu J, Zhou Y. Development and validation of an MRI radiomics-based signature to predict histological grade in patients with invasive breast cancer. *Breast Cancer (Dove Med Press)* (2022) 14:335–42. doi: 10.2147/BCTT.S380651

28. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and calibration of clinical prediction models: users' Guides to the medical literature. *JAMA* (2017) 318:1377–84. doi: 10.1001/jama.2017.12126

29. Yu Y, Tan Y, Xie C, Hu Q, Ouyang J, Chen Y, et al. Development and validation of a preoperative magnetic resonance imaging radiomics-based signature to predict axillary lymph node metastasis and disease-free survival in patients with early-stage breast cancer. *JAMA Netw Open* (2020) 3:e2028086. doi: 10.1001/jamanetworkopen.2020.28086

30. Han X, Cao W, Wu L, Liang C. Radiomics assessment of the tumor immune microenvironment to predict outcomes in breast cancer. *Front Immunol* (2022) 12:773581. doi: 10.3389/fimmu.2021.773581

31. Huang SY, Franc BL, Harnish RJ, Liu G, Mitra D, Copeland TP, et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis. *NPJ Breast Cancer* (2018) 4:24. doi: 10.1038/s41523-018-0078-2

32. Wang C, Chen X, Luo H, Liu Y, Meng R, Wang M, et al. Development and internal validation of a preoperative prediction model for sentinel lymph node status in breast cancer: combining radiomics signature and clinical factors. *Front Oncol* (2021) 11:754843. doi: 10.3389/fonc.2021.754843

33. Mazurowski MA, Saha A, Harowicz MR, Cain EH, Marks JR, Marcom PK. Association of distant recurrence-free survival with algorithmically extracted MRI characteristics in breast cancer. *J Magn Reson Imaging* (2019) 497:e231–40. doi: 10.1002/jmri.26648

34. Brown PJ, Zhong J, Frood R, Currie S, Gilbert A, Appelt AL, et al. Prediction of outcome in anal squamous cell carcinoma using radiomic feature analysis of pretreatment FDG PET-CT. *Eur J Nucl Med Mol Imaging* (2019) 46:2790–9. doi: 10.1007/s00259-019-04495-1

35. Zhou Y, Li J, Zhang X, Jia T, Zhang B, Dai N, et al. Prognostic value of radiomic features of 18F-FDG PET/CT in patients with B-cell lymphoma treated with CD19/CD22 dual-targeted chimeric antigen receptor T cells. *Front Oncol* (2022) 12:834288. doi: 10.3389/fonc.2022.834288

36. Toyama Y, Hotta M, Motoi F, Takanami K, Minamimoto R, Takase K. Prognostic value of FDG-PET radiomics with machine learning in pancreatic cancer. *Sci Rep* (2020) 10:17024. doi: 10.1038/s41598-020-73237-3

37. Li Q, Xiao Q, Li J, Duan S, Wang H, Gu Y. MRI-based radiomic signature as a prognostic biomarker for HER2-positive invasive breast cancer treated with NAC. *Cancer Manag Res* (2020) 12:10603–13. doi: 10.2147/CMAR.S271876

38. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol* (2016) 34:2157–64. doi: 10.1200/JCO.2015.65.9128

39. Schöder H, Moskowitz C. Metabolic tumor volume in lymphoma: hype or hope? *J Clin Oncol* (2016) 34:3591–4. doi: 10.1200/JCO.2016.69.3747

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 11 October 2023
DOI 10.3389/fonc.2023.1226118

OPEN ACCESS

EDITED BY
Noha Mousaad Elemam,
University of Sharjah, United Arab Emirates

REVIEWED BY
Tomohiro Chiba,
Kyorin University, Japan
Wenbin Zhou,
Nanjing Medical University, China

*CORRESPONDENCE
Ceshi Chen
chenc@kmmu.edu.cn
Chunyan Wang
740729chunyan@sina.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 20 May 2023
ACCEPTED 26 September 2023
PUBLISHED 11 October 2023

CITATION
Pan C, Xu A, Ma X, Yao Y, Zhao Y, Wang C and Chen C (2023) Research progress of Claudin-low breast cancer.
Front. Oncol. 13:1226118.
doi: 10.3389/fonc.2023.1226118

Research progress of Claudin-low breast cancer

Chenglong Pan[1,2†], Anqi Xu[2,3†], Xiaoling Ma[1,2], Yanfei Yao[1,2], Youmei Zhao[1,2], Chunyan Wang[1*] and Ceshi Chen[4,5*]

[1]Department of Pathology, First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China, [2]Kunming Medical University, Kunming, Yunnan, China, [3]Department of Anesthesia, First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China, [4]Academy of Biomedical Engineering, Kunming Medical University, Kunming, Yunnan, China, [5]The Third Affiliated Hospital, Kunming Medical University, Kunming, Yunnan, China

Claudin-low breast cancer (CLBC) is a subgroup of breast cancer discovered at the molecular level in 2007. Claudin is one of the primary proteins that make up tight junctions, and it plays crucial roles in anti-inflammatory and antitumor responses as well as the maintenance of water and electrolyte balance. Decreased expression of claudin results in the disruption of tight junction structures and the activation of downstream signaling pathways, which can lead to tumor formation. The origin of Claudin-low breast cancer is still in dispute. Claudin-low breast cancer is characterized by low expression of Claudin3, 4, 7, E-cadherin, and HER2 and high expression of Vimentin, Snai 1/2, Twist 1/2, Zeb 1/2, and ALDH1, as well as stem cell characteristics. The clinical onset of claudin-low breast cancer is at menopause age, and its histological grade is higher. This subtype of breast cancer is more likely to spread to lymph nodes than other subtypes. Claudin-low breast cancer is frequently accompanied by increased invasiveness and a poor prognosis. According to a clinical retrospective analysis, claudin-low breast cancer can achieve low pathological complete remission. At present, although several therapeutic targets of claudin-low breast cancer have been identified, the effective treatment remains in basic research stages, and no animal studies or clinical trials have been designed. The origin, molecular biological characteristics, pathological characteristics, treatment, and prognosis of CLBC are extensively discussed in this article. This will contribute to a comprehensive understanding of CLBC and serve as the foundation for the individualization of breast cancer treatment.

KEYWORDS

breast cancer, Claudin-low, immunohistochemistry, mammary stem cells, epithelial-mesenchymal transformation

1 Introduction

Breast cancer is one of the three most common cancers in the world. Hierarchical cluster analysis of the genes that vary more between tumors than between repeated samples of the same tumor has revealed the existence of four major breast cancer intrinsic subtypes (luminal A, luminal B, HER2-enriched, and basal-like), as well as a normal breast-like group (1, 2). In

2007, Herschkowitz et al. (3) found a breast cancer subtype called the claudin-low subtype upon the examination of murine and human breast tumors. This newly discovered subtype of breast cancer has features of low expression of tight junction proteins and adhesion proteins (Claudin3, 4, 7, and E-cadherin) and luminal markers but high expression of basal-related genes and lymphocyte- and endothelial cell-related markers. Recently, scientists have paid more attention to this intrinsic molecular subtype of breast cancer, which is known as claudin-low breast cancer (CLBC). A comprehensive analysis of CLBC will help us to better understand this intrinsic subtype of breast cancer.

2 Structure and function of claudins

Furuse and Tsukita first discovered claudin in 1998. Claudin derives from the Latin word 'claudere', which means to close (4, 5). Claudins (CLDN) are cell–cell adhesion proteins that are expressed at tight junctions (TJs), which are the most prevalent apical cell–cell adhesions (6). Tight junctions, together with adherens junctions and desmosomes, form the apical junctional complex in epithelial and endothelial cellular sheets. Adherens junctions and desmosomes are responsible for the mechanical adhesion between adjacent cells, whereas tight junctions are essential for the tight sealing of the cellular sheets, thus controlling paracellular ion flux and therefore maintaining tissue homeostasis (4, 7, 8). The tight junction proteins are diverse and include occludins, claudins, tricellulin, cingulin, and junctional adhesion molecules (JAMs). These proteins interact within themselves and with the cellular cytoskeleton to form a complex architecture. Among these TJ proteins, claudins are key proteins that act as both pores and barriers, aiding the paracellular pathway between epithelial cells (9). In mammals, it is composed of 27 members, and specific CLDN combinations are expressed in specific cell and tissue types. Claudins can be divided into classical claudins and nonclassical claudins according to the homology of the claudin sequence and its function (10). Claudin has four transmembrane proteins; in mammals, 27 members are between 20 and 34 kDa in size (11). Claudin has four transmembrane helices, and its amino- and carboxy-termini penetrate deep into the cytoplasm. The N-terminal domain of Claudins is relatively short, contain 7 amino acid sequences followed by a large extracellular loop (ECL1) of approximately 50 amino acid sequences. A short inner loop is separated by a small extracellular loop (ECL2) of approximately 25 amino acid residues. The main role of ECL1 is paracellular transport with selective ion permeability. The role of ECL2 is to participate in the interaction between Claudins. ECL2 of Claudin3 and Claudin4 contains binding receptors for *Clostridium perfringens* enterotoxin (CPE) (5, 10, 12–14). Studies have found that CPE can be used as a target for the development of claudin-targeted drugs (15–17). The localization and function of claudin are also regulated by the phosphorylation of the C-terminus, a target of serine, threonine, and tyrosine kinases. Claudin regulates functional changes in TJ proteins through posttranslational modifications, such as phosphorylation, ubiquitination, palmitoylation, and glycosylation, to regulate claudin conformation, stability, transport, and function (10). The C-terminus of claudins contains a PDZ-binding domain, for signal transduction. Claudin binds to Zonula occludens (ZO1, ZO2, and ZO3 (18)), Pals1-associated tight junction protein (PATJ) (19) and Multi-PDZ domain protein 1 (MUPP1) (20) through a PDZ-interacting domain, which plays an important role in cell function. ZO-1, ZO-2, and ZO-3 are TJ-related proteins (21–23). ZO-1 and ZO-2 are key proteins for cell junction assembly and permeability, respectively (24, 25). When ZO-1 and ZO-2 are missing, cells cannot assemble TJs (26, 27). The structure diagram of claudin is shown in Figure 1. The PDZ domain-binding motif located at the C-terminal end of claudin cytoplasm can directly interact with ZO family proteins, which plays important roles in many cellular processes.

Whether different claudin proteins copolymerize to form tight junction chains as heteropolymers, and whether claudin interacts in a homogeneous manner, between two molecules of the same claudin member, or between two different claudin members remains unclear. According to a previous study, different claudin members can interact within and between tight junction strands, but these combinations were restricted to specific combinations of isoforms (28).

Claudin is widely expressed in many different tissues. The majority of claudin expression occurs in barrier-forming epithelial cells and endothelial cells. Not all claudins are expressed in all tissues at the same time, and these differences in how claudins are expressed control the functions of cells, especially paracellular barrier functions. Claudins play a key role in regulating the paracellular permeability selectivity. The overexpression of Claudin affects the resistance and permeability of epithelial cells to different ions, and ECL1 plays an important role in the selectivity of charge (29, 30). Various mouse models have shown that Claudin plays an important role in cellular barriers. For example, claudin-1-deficient mice are dehydrated due to increased cell permeability, leading to rapid death (31). Claudin shows abnormal expression in several cancers, which is related to the occurrence and progression of cancer. Claudin-1 is not expressed in breast cancer and colon cancer, which is related to the decomposition of TJs in tumor development (32, 33). Claudin is a double-edged sword. The loss or low expression of Claudin in some tumors is related to the progression, invasion and metastasis of cancer, such as gastric cancer (34), esophageal squamous cell carcinoma (35) and colorectal cancer (36). Claudin has the opposite effect in some cancers. Claudin can directly or indirectly activate various signal pathways or proteases, and promote the occurrence of tumors. Therefore, it has been observed that Claudin is highly expressed in some tumors. Claudin-3 and -4 have been found to be expressed in many types of tumors, and some studies have shown that the overexpression of Claudin in tumors is related to tumor growth and invasion (16, 37), such as ovarian cancer (38, 39), glioma (40) and pancreatic cancer (41, 42). The role of Claudin in cancer stem cell biology through the WNT pathway has attracted increasing attention. Claudin-1 and -2 transcription is regulated by WNT signaling, which is known to regulate the β-Catenin-T-cell factor/

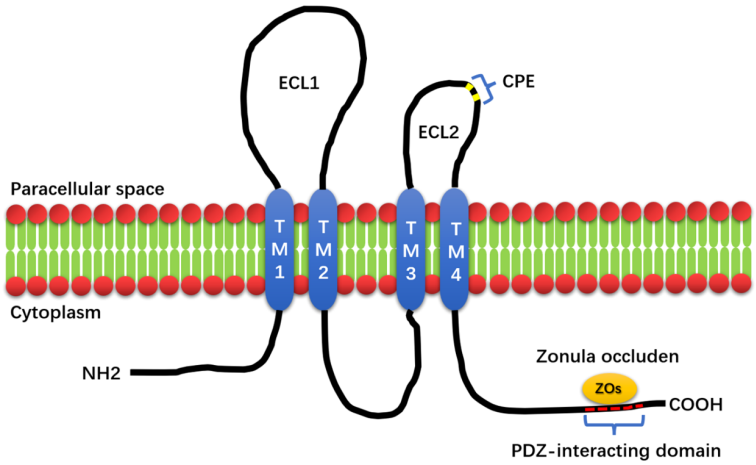


FIGURE 1
The structure of Claudin. The structure of Claudin mainly consists of four transmembrane structures (TM1, 2, 3, and 4), two extracellular loops, and one intracellular loop (ECL1, ECL2). The small extracellular loop contains the *Clostridium perfringens* enterotoxin-binding receptor (CPE), and the PDZ-interacting domain at the C-terminus of Claudin is capable of binding to Zonula occluden (Zos) to maintain cell homeostasis.

lymphoid enhancer-binding factor (TCF/LEF) signaling pathway and stemness (43, 44).

3 Definition of claudin-low breast cancer

CLBC was originally defined by a gene expression signature, which represents the low expression of claudin and cell-cell adhesion-related genes, including claudin 3, 4, and 7, and the high expression of EMT-related genes and breast cancer stem cell-related genes (45). Defining CLBC by genome is the foundation for most researchers to study CLBC. In addition, some researchers defined CLBC by immunohistochemistry. Most scholars define the low expression of claudins 3, 4, and 7 in the tumor cell membrane as CLBC. However, unlike HER2 in immunohistochemical staining, there is no unified standard to define the low expression and high expression of claudin in the interpretation of Claudin immunohistochemical (IHC) staining. Some scholars have multiplied the staining intensity and staining degree of claudin in tumor cells and delineated the corresponding boundary to define the high and low expression of claudin (46–50). Other researchers defined CLBC when less than 50% of tumor cells express claudins 3 and 4 and less than 5% of tumor cells express claudin 7 (51). However, some scholars have recently put forward a new view that Claudin-low is a biological characteristic that describes breast cancer and is not equivalent to the inherent molecular classification of breast cancer (52). Therefore, we believe that the definition of Claudin-low breast cancer should be a type of breast tumor characterized by low expression or absence of epithelial adhesion proteins (Claudin3, Claudin4, Claudin7 and E-cadherin, etc.). Compared with the intrinsic molecular classification of breast cancer, CLBC has its own unique characteristics in clinical pathology, molecular biology and other aspects.

4 The origin of claudin-low breast cancer

The origin of CLBC has been disputed. At present, there are three main theories about the origin of CLBC. Some scholars have found that CLBC has features of tumor-initiating cells or stem cells. They speculated that the intrinsic subtype of breast cancer may reflect different stages of epithelial development. CLBC represents the most primitive tumors that are similar to mammary stem cells and may originate from mammary stem cells (MaSCs) (53–55). Previous studies have found that the active form of RAS and various cytokines, including transforming growth factor-β (TGF-β), promote epithelial-mesenchymal transformation (EMT), which transforms breast epithelial cells into malignant cells, which have all the characteristics of CLBC (56–59). Many factors regulate the transformation of luminal epithelium into CLBC through EMT. For example, oncogenic RAS signaling drives the occurrence and progression of triple-negative breast cancer from the luminal epithelium (60). Absence of the Notch signaling regulators Lunatic Fringe (Lfng) and p53 leads to CLBC (61). TbL1 interacts with *ZEB1*, which inhibits the activation of the E-cadherin (CDH-1) gene, activates the *ZEB1* gene promoter, and promotes the transformation of mammary epithelial cells to CLBC (62). There is evidence that, under the influence of genetic mutations or environmental conditions, reactivation or dedifferentiation of luminal epithelial cells induces and accelerates the formation of more aggressive mammary tumors. Furthermore, the deletion of p53 can lead to clonal proliferation of the luminal epithelium, which promotes the development of breast tumors and the acquisition of MaSC characteristics, resulting in the formation of CLBC (63–65). In recent years, some scholars have deleted the *Pten* gene in the mouse mammary epithelium and induced the p53-R270H mutation, leading to CLBC (66). Other scholars have found that the synergistic effect of MET and p53 deletion can induce CLBC (67). However, according to

the study by Pommier et al. (68), CLBC shows remarkable diversity. Based on the analysis of genetics, gene methylation, and gene expression, they found that CLBC arises from three subgroups, two of which are related to luminal subtype breast cancer and basal-like subtype breast cancer, which are transformed by activating the EMT process during tumor development. The third subgroup is closely related to normal human breast stem cells (MaSCs), showing genome distortion and a low frequency of the TP53 mutation. In short, the origin of CLBC is a complicated process.

According to previous studies, claudin-low breast cancer has always been considered to fall into the category of triple-negative breast cancer. According to Pommier et al. (68), claudin-low breast cancer may originate from various stages of breast cancer development, and this is likely the best explanation for the high expression of breast cancer stem cell markers and EMT-related markers in claudin-low breast cancer as well as its overlap in gene and immune expression with luminal A/B, HER2 overexpression, and triple-negative breast cancer. The Shanghai Cancer Center of Fudan University (FUSCC) used androgen receptor (AR), CD8, FOXC1, and DCLK1 as immunohistochemical markers and classified TNBCs into five subtypes based on the staining results: (a) IHC-based luminal androgen receptor (IHC-LAR; AR+), (b) IHC-based immunomodulatory (IHC-IM; AR−, CD8+), (c) IHC-based basal-like immune-suppressed (IHC-BLIS; AR−, CD8−, FOXC1+), (d) IHC-based mesenchymal (IHC-MES; AR−, CD8−, FOXC1−, DCLK1+), and (e) IHC-based unclassifiable (AR−, CD8−, FOXC1−, DCLK1−) (69). Currently, there is no relevant research showing the relationship between claudin-low and FUSCC triple-negative breast cancer staging. Since claudin breast cancer is positive for EMT and stem cell-related immune markers, we believe it is more related to MES.

5 The mechanism of Claudin-low carcinogenesis

Claudin is highly expressed in normal tissues. Although the expression of Claudin varies across tumors, it is decreased in many tumors, indicating that Claudin can inhibit the function of tumors. The increased expression of Claudin significantly inhibits the proliferation of tumor cells and promotes EMT, migration and invasion of tumors. Low expression of claudins is associated with advanced disease, metastasis, and a poor prognosis (70). At present, the mechanism by which Claudin inhibits tumor occurrence is not clear, but some scholars have proposed the following three hypotheses: (A) Claudin is able to prevent microorganisms, toxins, and growth factors from passing through the paracellular channel, which are common oncogenic factors. In the absence of Claudin, these oncogenic factors enter the body and are likely to induce tumors (71–73). (B) Claudin directly binds or utilizes other scaffold proteins [mainly zonula occludens (ZOs)] to indirectly bind some key factors of signaling pathways, such as yes-associated protein/transcriptional coactivator (YAP/TAZ), pyruvate dehydrogenase kinase 1 (PDK-1), and β-catenin, which are all well-known carcinogenic factors. Claudin keeps them on the cell membrane and blocks carcinogenic signaling pathways, such as the PDK-AKT and YAP/TAZ-TEAD pathways (74–77). Reduced expression of ZO-1 in breast cancer destroys the structural integrity of the tight junction, leading to a loss of cell-cell adhesion. MUPP-1 expression is reduced in patients with poor prognosis and increased tumor grade. MUPP-1, like ZO-1, can be used as a crosslinker between claudin and tight-linked chains and tight-linked JAM oligomers, and other integral membrane proteins can be recruited into tight-linked claudin via MUPP-1. Studies have found that the transcription levels of ZO-1 and MUPP-1 in metastatic breast cancer are significantly reduced. The higher the expression of ZO-1 in breast cancer, the worse the prognosis of breast cancer patients (78). Normal tissues at the tight junction of mammary epithelial cells showed strong ZO-1 staining, and approximately 70% of breast cancers were found to have reduced or lost ZO-1. ZO-1 staining was positively correlated with tumor differentiation, and the reduction in ZO-1 staining was closely related to the reduction in E-cadherin staining. Therefore, ZO-1 may be directly involved in the malignant progression of breast cancer (79). A decrease in ZO-2 expression has also been observed in most breast cancers (80). (C) The basement membrane Claudin was found to colocalize with integrin β1 and form a protein complex to maintain epithelial cell attachment and inhibit cell proliferation. If claudin is depleted or expressed at low levels, epithelial cells become destabilized and proliferative (81, 82).

Apart from causing cancer in the above three aspects, Claudin also inhibits EMT. EMT is associated with tumorigenesis, progression and fibrosis (83). Transforming growth factor beta 1 (TGF-β1) has been shown to induce EMT during various stages of embryogenesis and progressive disease. Medici et al. (84) found that claudin protein expression was missing when TGF-β was used to induce EMT. SMAD3 and SMAD4 interact with SNAIL1, a transcriptional repressor and EMT promoter, and form a complex (SNAIL1-SMAD3/4) that targets the gene promoter of CAR, a tight junction protein, and E-cadherin. SNAIL1 and SMAD3/4 act as corepressors of the CAR, occludin, claudin-3, and E-cadherin promoters. In contrast, the combined silencing of SNAIL1 and SMAD4 with siRNA promoted the transcription of CAR and occludin during EMT (85). These studies indicate that Claudin is inhibited during EMT. Studies have found that the deletion of Claudin3 or Claudin4 leads to activation of the PI3K pathway, which is manifested as increased Akt phosphorylation, increased PIP3 content and PI3K activity, as well as up regulation of mRNA and protein levels of the transcription factor Twist. Claudin3 and Claudin4 maintain the epithelial phenotype, and their deletion promotes EMT (86).

Tumor metastasis is an important step in the process of tumor progression. The interaction between cancer cells and endothelial cells is the key for the distant metastasis of tumor cells. In the process of tumor metastasis, tightly connected Claudin participates in the defense of tumor cells, and tightly connected Claudin is the barrier of paracellular channels for epithelial and endothelial cell material exchange (87). Researchers believe that the loss of adhesion of tumor epithelial cells is a necessary condition for tumor invasion and metastasis. The expression of claudin in human cancer may decrease or increase in a tissue-specific manner. Claudin is

responsible for the structural and functional integrity of the tight junction of the epithelial cell layer. The decrease or loss of claudin expression is accompanied by cell-cell adhesion and polarity damage. Tumorigenesis is accompanied by the destruction of tight junctions, a process that plays an important role in the loss of adhesion and the enhancement of invasiveness in tumor cells. The loss of Claudins and other tight junction proteins in cancer is considered a mechanism of cell adhesion loss and an important step in metastasis (88). Claudin constructs a complete biological system based on the function of the cell-side barrier. Generally, these claudin functional deficits are associated with water imbalance, inflammation, cancer, and brain disease, depending on the tissues and organs involved (31, 74, 89–92). Many studies have found that Claudin is expressed at low levels in many breast cancer tissues (93–95).

The feedback pathway of ZEB1/miR-200a promotes the invasion and metastasis of breast cancer cells (96, 97). It was found that miR-200a can regulate YAP1 in the HIPPO pathway and promote the survival and metastasis of CLBC (98). The ectopic expression of miR-200a can weaken the migration and invasion of CLBC by reducing the expression level of *ELK3* mRNA (99). Compared with other subtypes of breast cancer, CLBC has the characteristics of a vascular system and endothelium and has higher vascular permeability, which promotes the metastasis of CLBC (100).

It has been proved that changes of protein glycosylation is involved in the regulation of EMT process of cancer cells (101). Claudin and low expression of Claudin play a role in carcinogenesis through multiple signaling pathways. Dedicator of cytokinesis 1 (DOCK1) down-regulates Claudin through ribosomal RNA processing 1B (RRP1B)/DNA methyltransferase (DNMT) to promote the growth and migration of breast cancer cells (102); Claudin through down-regulates the transforming growth factor-β(TGF-β)/Smad2/DNMT to promote EMT, metastasis and invasion of breast cancer cells (103). Low expression of Claudin promotes EMT by activating PI3K/AKt/mTOR, and regulates ERK/Sp1/CyclinD1 and ERK/IL-8 to promote the proliferation, migration and invasion of breast cancer cells (104). Claudin dysregulation promotes tumorigenesis by upregulating gp130/IL6/Stat3 signaling and activating the Wnt/β-catenin signaling pathway (74). Epithelial cell adhesion molecule (EpCAM) is a homophilic type I transmembrane glycoprotein belonging to the small GA733 protein family, EpCAM functions not only in physiological processes but also participates in the development and progression of cancer (105). The glycosylation of EpCAM is inhibited, which increases the expression of EpCAM in breast cancer. At the same time, the combination of Claudin and EpCAM increases, promotes the phosphorylation of PI3K/Akt and p38, promotes the activation of MAPK and PI3K/Akt pathways, promotes the process of EMT, reduces Apoptotic ability of breast cancer cells, promoting cell proliferation (106). In addition, Claudin and Integrin β1 inhibit the proliferation of tumor cells. If Claudin is missing, the combination of Claudin and Integrin β1 will be reduced, which will promote the tumor (81). The mechanism of claudin and claudin low-expression carcinogenic signal pathway is shown in Figure 2.

6 Molecular biological characteristics of claudin-low breast cancer

Currently, the following cell lines are used to study claudin-low breast cancer: MDA-MB157, MDA-MB436, BT549, MDA-MB231, HBL100, SUM159PT, Hs578T, MDA-MB435, and SUM1315. Some studies have extracted two cell lines, RM11A and RJ348, from breast tumors developed in MTB-IGFIR transgenic mice, which have the histological characteristics and gene expression pattern of claudin-low breast cancer and have been reinjected into the mouse breast fat

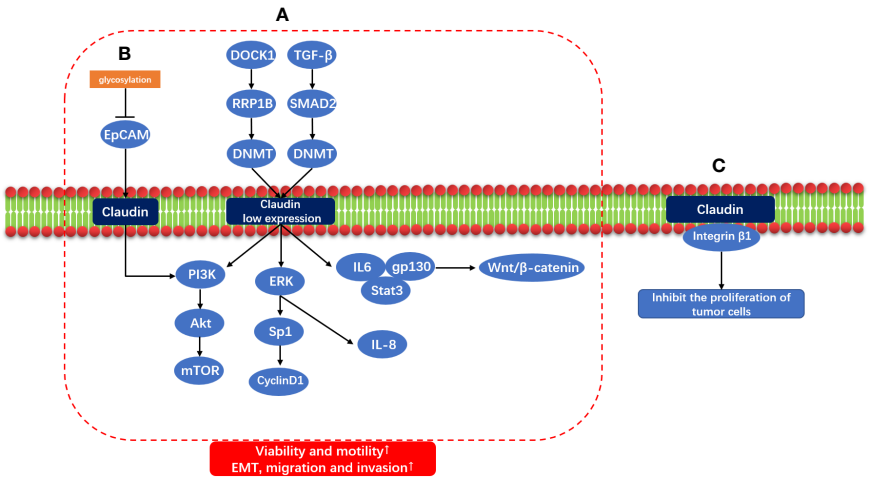


FIGURE 2
Claudin and Claudin low-expression carcinogenic signal pathway. **(A)** DOCK1/RRP1B/DNMT and TGF-β/SMAD2/DNMT induce Claudin low expression through methylation, and Claudin low expression through IL6/gp130/Stat3, Wnt/β-Catenin signal pathway; ERK/Sp1/CyclinD1 and ERK/IL8; PI3K/AKt/mTOR promotes EMT, migration and invasion of tumor. **(B)** Glycosylation of EPCAM was inhibited, and PI3K/AKt/mTOR was activated by combining with Claudin to promote EMT. **(C)** Claudin and Integrin β1 also inhibit the proliferation of tumor cells.

pad to obtain high tumorigenicity (107). Fougner et al. (108) found that claudin-low-like MPA/DMBA-induced mouse mammary tumors represent a transcriptionally accurate model for human claudin-low breast cancer.

6.1 Gene expression characteristics of claudin-low breast cancer

CLBC has the features of EMT and stem cells (109). Taube et al. (110) showed that tumor cells gain a certain degree of invasiveness and metastasize through EMT; at the same time, they also gain the potential for self-renewal, proliferating to form a macroscopically metastatic large cell population. Studies have shown that TGF-β induces EMT through the TGF-β/SMAD/LEF/PDGF pathway and the MAPK pathway (111). In addition, the homeobox proteins Goosecoid and the zinc finger proteins SNAIL and TWIST can also induce EMT. Another study found that miR-200 expression decreases during EMT (112). Surprisingly, Jones et al. (113) re-expressed miR-200c in murine claudin-low breast tumor cells to inhibit tumor cell proliferation and growth. Therefore, they believe that the expression of miR-200c can inhibit the growth of claudin-low breast tumor cells.

In addition, CLBC has a high degree of immune cell infiltration and high expression of T and B lymphoid cell markers (such as CD14 and CD79a) (45, 114, 115). Some scholars have compared CLBC with non-claudin-low breast cancer and found that CLBC has few gene mutations, low genetic instability, a low frequency of TP53 and PIK3CA mutations, and minimal MYC and MDM4 gain (52). Through Gene Ontology (GO) analysis, 165 genes were found to be differentially expressed in CLBC, among which 69 were upregulated and 64 downregulated. They analyzed 193 pathways and found that the inflammatory IL-13 signaling pathway was significantly enriched. Among them, five upregulated genes (IL6, CXCL8, VEGF-C, NRF1, and EREG) were mapped as hubs and may play an important role in CLBC (116). The comparison of Claudin-low breast cancer and other subtypes of breast cancer in immunohistochemistry and gene phenotype is shown in Table 1.

6.2 Claudin-low breast cancer is related to the MAPK pathway

Several studies have shown that RAS/MAPK kinases are highly activated in breast cancer (60, 68, 108). The dataset from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) and TCGA cohorts was studied, and RAS signals were found to be highly expressed in CLBC. The GDSC database was also used to observe the IC50 of three MEK inhibitors (trametinib, selumetinib, and refametinib) in various subtypes of breast cancer, and CLBC was sensitive to three MEK inhibitors compared with other subtypes. This finding indicates that CLBC is driven by the MAPK pathway. Unfortunately, these findings were only performed on the genetic claudin-low breast cancer cell line and the corresponding mouse model, and clinical trial studies are lacking.

7 Clinicopathological characteristics of claudin-low breast cancer

7.1 Patient age

At present, there is still controversy about the age of CLBC. Some scholars have found that patients with CLBC are younger than patients with other subtypes (114, 115, 117, 118). However, Xu et al. (119) found that the age of CLBC patients is not significantly different from other subtypes, which may be related to their small sample size. Fougner et al. (52) found that the onset age of patients with negative ER, PR and HER in CLCB is younger than that of patients without negative ER, PR and HER. However, when they compared the CLBC and non-CLBC groups, the ages of the two groups of patients were similar.

7.2 Tumor characteristics: tumor size, histological grade, and lymph node metastasis

The research results of Sabatier et al. (114) and Xu et al. (119) were consistent, in that CLBC was mainly larger than 2 cm (62%, 61.9%). In the tumor group larger than 2 cm, the proportion of CLBC was lower than that of the HER2-enriched subtype and the basal-like subtype and higher than that of the luminal subtype and the normal-like subtype. They also found that the proportion of histological grade 3 was higher than that of other subtypes (56%, 71.4%), except for the base-like subtype (85%, 56.6%). The rate of lymph node metastasis (46%, 38.1%) was higher than that of the base-like subtype (35%, 25%) and luminal A subtype (54%, 26.8%). This indicates that CLBC has a larger tumor diameter, a higher histological grade, and a higher rate of axillary lymph node metastasis. This may be related to the low expression of adhesion-related proteins in CLBC, which makes it easier to transfer.

7.3 Histopathology

Prat et al. (45) found that approximately 50% (8/16) of CLBCs are metaplastic carcinomas,

and Sabatier et al. (114) found that CLBCs are mainly nonspecialized breast cancer (77.7%), metaplastic carcinoma (71.4%), and medullary carcinoma (24%). Some scholars have also found that the histological types of CLBC are mainly invasive lobular carcinoma, metaplastic carcinoma, and medullary carcinoma (120). In short, the histological types of CLBC are mainly poorly adherent carcinomas, metaplastic carcinomas, and medullary carcinomas. Both CLBC and lobular carcinoma have down-regulation of E-cadherin, but CLBC does not always show the histological characteristics of lobular carcinoma. This may be along with the down-regulation of E-cadherin and other molecular changes, CLBC is more prone to ductal carcinoma. Characteristically, in contrast to lobular carcinoma, downregulation of E-cadherin in CLBC is associated with epigenetic or posttranscriptional dysregulation (121).

7.4 Immunohistochemistry

CLBC has low expression of tight junction proteins and adhesion proteins (Claudin3, 4, 7 and E-cadherin), and it also has low expression of cytokeratin (CK5, 8, 14, 17, 18, 19, and 24), HER-2 and luminal-related markers (45, 115). These features are similar to the basal-like subtype of breast cancer. CLBC has low expression of proliferation genes and Ki-67, which is inconsistent with the basal-like subtype of breast cancer (45, 115). Compared with other subtypes of breast cancer, CLBC has high expression of EMT-related markers (such as Vimentin, SNAI1 and SNAI2, TWIST1 and TWIST2, and ZEB1 and ZEB2) (114) and breast cancer stem cells/tumors with the initial cell markers ALDH1 and $CD44^{hi/}CD24^{-/low}$ (115, 117) and high expression of PD-L1 (122). Recently, caveolin-1, galectin-1, and SMYD3 were found to be highly expressed in CLBC (123–125).

Some studies have found that CLBC is heterogeneous, as is its immunohistochemical expression. Fougner et al. (52) believed that they listed CLBC for analysis without considering the six intrinsic

TABLE 1 Immunohistochemical and gene comparison of Claudin-low breast cancer and other subtypes of breast cancer.

IHC/Gene	Claudin-low	Luminal A/B	HER 2	TNBC
Claudin3	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
Claudin4	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
Claudin7	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
E-cadherin	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
ER	Negative	Low expression/Negative	Negative	Negative
PR	Negative	Low expression/Negative	Negative	Negative
HER2	Negative	Low expression/Negative	Hight expression	Negative
CK5	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
CK8	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
CK14	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
CK17	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
CK18	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
CK19	Negative	Low expression/Negative	Low expression/Negative	Low expression/Negative
Vimentin	Hight expression	Negative	Negative	Negative
SNAI1	Hight expression	Negative	Negative	Negative
SNAI2	Hight expression	Negative	Negative	Negative
TWIST1	Hight expression	Negative	Negative	Negative
TWIST2	Hight expression	Negative	Negative	Negative
ZEB1	Hight expression	Negative	Negative	Negative
ZEB2	Hight expression	Negative	Negative	Negative
ALDH1	Hight expression	Negative	Negative	Negative
CD44	Hight expression	Negative	Negative	Low expression
PD-L1	Hight expression	—	—	—
Caveolin-1	Hight expression	—	—	—
Galectin-1	Hight expression	—	—	—
SMYD3	Hight expression	—	—	—
CD24	Negative	Negative	Hight expression	Low expression
Ki-67	Low proliferation index	—	—	—
GATA3	Negative	Low expression	Negative	Negative
EPCAM	Negative	Negative	Low expression	Negative
CD10	Low expression	Negative	Negative	Negative

"-" means that these marker are not related to literature reports.

subtypes. This may have masked the intrinsic characteristics of CLBC. Pommier et al. (68) divided CLBC into three subgroups (CL-1, 2, and 3) according to FGA levels, which were derived from mammary stem cells (MaSC), mature luminal cells (mature luminal, mL), and luminal precursor cells (luminal progenitor, PL), respectively, which have the characteristics of expressing stem cell-related protein markers, luminal-related markers, and basal-like markers. This is in line with the findings of Fougner et al. (52) CLBC penetrates every subtype of breast cancer. Therefore, the immunohistochemical expression of CLBC is a complicated process, and different markers are expressed depending on the origin of the tumor.

8 The prognosis of claudin-low breast cancer

Prat et al. (114) analyzed UNC337 and two independent gene expression databases (NKI1295 and MDACC133) with nine cell lines (claudin-low and PAM50 subtype predictors) and found that the incidence of CLBC was 7%~14%. In terms of prognosis, the Kaplan-Meier analysis of the two databases showed that the prognosis of CLBC (HR 2.83, RFS and OS 5.66) was significantly lower than that of luminal A subtype breast cancer (HR 4.71, RFS and OS 17.98). A study by Dias et al. (115) in 2016 showed that the overall survival (OR) of CLBC at 3, 5, and 10 years was higher than that of the basal-like subtype and lower than that of other subtypes, and the local recurrence rate was higher than that of other subtypes. Disease-free survival (DFS) is lower than that of luminal A breast cancer but higher than that of other subtypes. Sabatier et al. (114) found that the five-year local recurrence survival rate (67%) of CLBC is similar to that of luminal B (64%) and basal-like (60%) but higher than that of the HER2-enriched type (55%) and lower than that of luminal A (79%). Lu et al. (50) found that claudin-low has worse disease-free survival than non-CLBC, which is consistent with the conclusion found by Xu et al. (119). Overall, the prognosis for CLBC is poor.

Fougner et al. (52) conducted a study on the METABRIC database and found that, when CLBC was studied as a single group, the results showed that CLBC was associated with a poor prognosis, which was consistent with the findings mentioned above. However, when they analyzed CLBC according to the intrinsic molecular classification of breast cancer, the results showed that the basal-like Claudin-low had the worst prognosis, which was consistent with the prognostic results of the intrinsic molecular classification. In addition, the analysis of the claudin-low and non-claudin-low subtypes of the same intrinsic subtype of breast cancer showed that the difference was not statistically significant, indicating that there is no evidence that claudin-low can affect the prognosis of breast cancer.

Several studies have found that many immunohistochemical markers are helpful in evaluating the prognosis of CLBC. TBL1 is necessary for the mesenchymal phenotype of transformed breast epithelial cell lines and claudin-low subtype breast cancer cell lines. The high expression of the *TBL1* gene is related to poor prognosis and an increased metastasis rate in breast cancer patients, indicating that the expression level of TBL1 can be used as a prognostic marker (62). According to the clinical data analysis of Fenizia C et al. (125), a higher level of SMYD3 is related to the poor prognosis of CLBC and the reduction in metastasis-free survival of breast cancer patients.

9 The pathological response to chemotherapy and the progressive treatment of claudin-low breast cancer

At present, research on CLBC treatment is still in the basic research stage, and most of the studies have not yet entered the animal experiment and clinical trial stages. At present, studies on the pathological complete remission of claudin-low breast cancer are based on the follow-up of the original cases. Prat et al. (45) followed up 133 patients with the MADCC Breast Cancer Database and found that the complete pathological remission rate (pCR rate, 38.9%) of CLBC was significantly lower than that of the basal subtype (pCR rate, 73.3%) but higher than that of luminal A and B. Sabatier et al. (110) studied 5,447 patients, and 1,294 patients received neoadjuvant chemotherapy with a pCR rate of 23%. The pCR rate of 228 CLBC patients was 32%, which was close to that of the basal sample (pCR rate, 33%), lower than that of the HER2-rich type (pCR rates, 37%), and much higher than that of the luminal A and luminal B subtypes (pCR rates, 7% and 18%). Both results showed that the pCR rate of CLBC is higher than that of other subtypes except basal-like breast cancer, indicating that CLBC is relatively sensitive to chemotherapeutic drugs.

Immunotherapy of CLBC. Fougner et al. (52) pointed out that CLBC showed high lymphocyte infiltration and high expression of PD-L1 and T lymphocyte reaction, which may provide opportunities for immunotherapy. At present, the means of immunotherapy mainly include immune checkpoint inhibitors and anti-PD1/PD-L1 therapy, etc. CLBC is rich in lymphocyte infiltration and drinks high expression of PD-L1 and IL6. These immunotherapy can provide new treatment schemes for CLBC patients (126). There is a lack of identification of CLBC from intrinsic molecular classification of breast cancer in clinical practice, so there is little research on immunotherapy of CLBC. In addition, some researchers have found that CLBC is rich in Tregs. Depletion of Tregs and suppression of immune checkpoints can weaken tumor growth and prolong survival, but they cannot lead to tumor regression (127).

Several novel targets have been proposed for the treatment of CLBC. Chang et al. (128) found that PIK3CG is a potential target for the treatment of CLBC. Inhibiting the activation of PIK3CG can enhance the therapeutic effect of PTX on CLBC. Stalker et al. (129) discovered that PDGFR inhibitors (sunitinib, regorafenib, and masitinib) can be used to inhibit the migration and metastasis of CLBC cells. Some scholars found that the combination of pyrazole

derivatives and doxorubicin can increase the death of MDA-MB-231 cells (130). Panobinostat combined with gefitinib can synergistically inhibit the proliferation of CLBC cells and promote apoptosis (131, 132). CLBC is rich in EMT features, which may provide a new therapeutic target for chemotherapy. ZEB1, AP-1, and TEAD/YAP, the effector of the HIPPO pathway, form a transactivation complex to activate oncogenes, disrupting their molecular interaction may provide a promising treatment for CLBC (133). Recently, some new therapeutic targets have been found. AgNPs can selectively induce lipid peroxidation and cause irreversible proteotoxicity in CLBCs (134). SLC20A1 siRNA knockdown leads to suppression the activity of CLBC, and high expression of SLC20A1 indicates a poor prognosis (135). Studies have identified PVT-1 as a long noncoding RNA can regulate the expression of Claudin-4 in CLBC, indicating that it may be an important target for treating CLBC (136). Although some therapeutic targets for CLBC have been discovered, there is still a long way to go before clinical application.

10 Conclusions and perspectives

In summary, CLBC is derived from breast cancer stem cells, luminal precursor cells, and luminal mature cells. The process of transforming to CLBC is related to EMT. We summarized the clinicopathological features of CLBC (Table 2), which is highly expressed in the immune response, breast cancer stem cells/tumor initiating cells, and EMT-related gene markers and is related to the MAPK pathway. CLBC is more aggressive because of its low expression of adhesion-related proteins, which show larger tumor size, a higher histological grade, and higher lymph node metastasis. The histological type is mainly invasive ductal carcinoma, and metaplastic carcinoma is associated with medullary carcinoma. CLBC is sensitive to chemotherapy, but its overall prognosis is poor. Currently, our knowledge of claudin-low breast cancer is inadequate due to the lack of uniformity in the classification of claudin-low breast cancer by previous investigators and the limited sample size of the study, thus leading to some bias in the results of the investigators. In the future, the origin of claudin-low breast cancer and the classification criteria by immunohistochemistry still need investigation. The potential therapeutic targets ans strategies for claudin-low breast cancer should be studied.

TABLE 2 The clinicopathological features of Claudin-low breast cancer.

The clinicopathological features of CLBC		
Patient age	younger than other subtypes	controversial
Tumor characteristics		
tumor size	larger	
histological grade	higher	
lymph node metastasis	easy	
Histopathology	nonspecialized type breast cancer and metaplastic carcinomas	
Immunohistochemistry		
low expression or no expression	claudin3, claudin4, claudin7, E-cadherin, HER-2	
high expression	Vimentin, SNAI1, SNAI2, TWIST1, TWIST2, ZEB1, ZEB2, ALDH1	
Treatment	at the basic medical research stage	
pCR	pCR rate of CLBC is higher than other subtypes	
Prognosis	poor	controversial

Author contributions

CP and AX wrote and edited the paper, they have made the same contribution to this work and share the identity of the first author. XM, YZ and YY generated the tables and figure. CC and CW revised the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from the National Natural Science Foundation of China (grant number 82160461) and the Science, Technology Innovation Team of Pathological Diagnosis of Triple Negative Breast Cancer in Kunming Medical University (grant number CXTD202208) and Yunnan Provincial Health Commission Medical Discipline Leader Training Plan (grant number D2018058).

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.

Publisher's note

References

1. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* (2012) 490(7418):61–70. doi: 10.1038/nature11412

2. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* (2000) 406(6797):747–52. doi: 10.1038/35021093

3. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* (2007) 8(5):R76. doi: 10.1186/gb-2007-8-5-r76

4. Farquhar MG, Palade GE. Junctional complexes in various epithelia. *J Cell Biol* (1963) 17(2):375–412. doi: 10.1083/jcb.17.2.375

5. Furuse M, Fujita K, Hiiragi T, Fujimoto K, Tsukita S. Claudin-1 and -2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. *J Cell Biol* (1998) 141(7):1539–50. doi: 10.1083/jcb.141.7.1539

6. Tsukita S, Furuse M. Pores in the wall: claudins constitute tight junction strands containing aqueous pores. *J Cell Biol* (2000) 149(1):13–6. doi: 10.1083/jcb.149.1.13

7. Schneeberger EE, Lynch RD. The tight junction: a multifunctional complex. *Am J Physiol Cell Physiol* (2004) 286(6):C1213–28. doi: 10.1152/ajpcell.00558.2003

8. Tsukita MF S, Itoh M. Multifunctional strands in tight junctions. *Nat Rev Mol Cell Biol* (2001) 2(4):285–93. doi: 10.1038/35067088

9. Tsukita S, Furuse M, Itoh M. Multifunctional strands in tight junctions. *Nat Rev Mol Cell Biol* (2001) 2(4):285–93. doi: 10.1038/35067088

10. Krause G, Winkler L, Mueller SL, Haseloff RF, Piontek J, Blasig IE. Structure and function of claudins. *Biochim Biophys Acta* (2008) 1778(3):631–45. doi: 10.1016/j.bbamem.2007.10.018

11. Itallie C, Anderson JM. Claudins and epithelial paracellular transport. *Annu Rev Physiol* (2006) 68(1):403. doi: 10.1146/annurev.physiol.68.040104.131404

12. Günzel D, Fromm M. Claudins and other tight junction proteins. *Compr Physiol* (2012) 2(3):1819–52. doi: 10.1002/cphy.c110045

13. Furuse M, Sasaki H, Fujimoto K, Tsukita S. A single gene product, claudin-1 or -2, reconstitutes tight junction strands and recruits occludin in fibroblasts. *J Cell Biol* (1998) 143(2):391–401. doi: 10.1083/jcb.143.2.391

14. Lal-Nag PJM M. The claudins. *Genome Biol* (2009) 10(8):235. doi: 10.1186/gb-2009-10-8-235

15. Kominsky SL, Tyler B, Sosnowski J, Brady K, Doucet M, Nell D, et al. Clostridium perfringens enterotoxin as a novel-targeted therapeutic for brain metastasis. *Cancer Res* (2007) 67(17):7977–82. doi: 10.1158/0008-5472.CAN-07-1314

16. Morin PJ. Claudin proteins in human cancer: promising new targets for diagnosis and therapy. *Cancer Res* (2005) 65(21):9603–6. doi: 10.1158/0008-5472.CAN-05-2782

17. Kominsky SL. Claudins: emerging targets for cancer therapy. *Expert Rev Mol Med* (2006) 8(18):1–11. doi: 10.1017/S1462399406000056

18. Itoh M, Furuse M, Morita K, Kubota K, Saitou M, Tsukita S. Direct binding of three tight junction-associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH termini of claudins. *J Cell Biol* (1999) 147(6):1351–63. doi: 10.1083/jcb.147.6.1351

19. Roh MH, Liu CJ, Laurinec S, Margolis B. The carboxyl terminus of zona occludens-3 binds and recruits a mammalian homologue of discs lost to tight junctions. *J Biol Chem* (2002) 277(30):27501–9. doi: 10.1074/jbc.M201177200

20. Hamazaki Y, Itoh M, Sasaki H, Furuse M, Tsukita S. Multi-PDZ domain protein 1 (MUPP1) is concentrated at tight junctions through its possible interaction with claudin-1 and junctional adhesion molecule. *J Biol Chem* (2002) 277(1):455–61. doi: 10.1074/jbc.M109005200

21. Balda MS, Garrett MD, Matter K. The ZO-1-associated Y-box factor ZONAB regulates epithelial cell proliferation and cell density. *J Cell Biol* (2003) 160(3):423–32. doi: 10.1083/jcb.200210020

22. Gumbiner B, Lowenkopf T, Apatira D. Identification of a 160-kDa polypeptide that binds to the tight junction protein ZO-1. *Proc Natl Acad Sci* (1991) 88(8):3460–4. doi: 10.1073/PNAS.88.8.3460

23. Haskins J, Gu L, Wittchen ES, Hibbard J, Stevenson BR. ZO-3, a novel member of the MAGUK protein family found at the tight junction, interacts with ZO-1 and occludin. *J Cell Biol* (1998) 141(1):199–208. doi: 10.1083/jcb.141.1.199

24. McNeil E, Capaldo CT, Macara IG. Zonula occludens-1 function in the assembly of tight junctions in Madin-Darby canine kidney epithelial cells. *Mol Biol Cell* (2006) 17(4):1922–32. doi: 10.1091/mbc.e05-07-0650

25. Umeda K, Matsui T, Nakayama M, Furuse K, Sasaki H, Furuse M, et al. Establishment and characterization of cultured epithelial cells lacking expression of ZO-1. *J Biol Chem* (2004) 279(43):44785–94. doi: 10.1074/jbc.M406563200

26. Umeda K, Ikenouchi J, Katahira-Tayama S, Furuse K, Sasaki H, Nakayama M, et al. ZO-1 and ZO-2 independently determine where claudins are polymerized in tight-junction strand formation. *Cell* (2006) 126(4):741–54. doi: 10.1016/j.cell.2006.06.043

27. Zihni C, Mills C, Matter K, Balda MS. Tight junctions: from simple barriers to multifunctional molecular gates. *Nat Rev Mol Cell Biol* (2016) 17(9):564–80. doi: 10.1038/nrm.2016.80

28. Furuse M, Sasaki H, Tsukita S. Manner of interaction of heterogeneous claudin species within and between tight junction strands. *J Cell Biol* (1999) 147(4):891–903. doi: 10.1083/jcb.147.4.891

29. Van Itallie JMA. Claudins CM. and epithelial paracellular transport. *Annu Rev Physiol* (2006) 68:403–29. doi: 10.1146/annurev.physiol.68.040104.131404

30. Colegio OR, Van Itallie CM, McCrea HJ, Rahner C, Anderson JM. Claudins create charge-selective channels in the paracellular pathway between epithelial cells. *Am J Physiol Cell Physiol* (2002) 283(1):C142–7. doi: 10.1152/ajpcell.00038.2002

31. Furuse M, Hata M, Furuse K, Yoshida Y, Haratake A, Sugitani Y, et al. Claudin-based tight junctions are crucial for the mammalian epidermal barrier: a lesson from claudin-1-deficient mice. *J Cell Biol* (2002) 156(6):1099–111. doi: 10.1083/jcb.200110122

32. Krämer F, White K, Kubbies M, Swisshelm K, Weber BH. Genomic organization of claudin-1 and its assessment in hereditary and sporadic breast cancer. *Hum Genet* (2000) 107(3):249–56. doi: 10.1007/s004390000375

33. Resnick MB, Konkin T, Routhier J, Sabo E, Pricolo VE. Claudin-1 is a strong prognostic indicator in stage II colonic cancer: a tissue microarray study. *Mod Pathol* (2005) 18(4):511–8. doi: 10.1038/modpathol.3800301

34. Dottermusch M, Krüger S, Behrens HM, Halske C, Röcken C. Expression of the potential therapeutic target claudin-18.2 is frequently decreased in gastric cancer: results from a large Caucasian cohort study. *Virchows Arch* (2019) 475(5):563–71. doi: 10.1007/s00428-019-02624-7

35. Sung CO, Han SY, Kim SH. Low expression of claudin-4 is associated with poor prognosis in esophageal squamous cell carcinoma. *Ann Surg Oncol* (2011) 18(1):273–81. doi: 10.1245/s10434-010-1289-4

36. Shibutani M, Noda E, Maeda K, Nagahara H, Ohtani H, Hirakawa K. Low expression of claudin-1 and presence of poorly-differentiated tumor clusters correlate with poor prognosis in colorectal cancer. *Anticancer Res* (2013) 33(8):3301–6.

37. Agarwal R, D'Souza T, Morin PJ. Claudin-3 and claudin-4 expression in ovarian epithelial cells enhances invasion and is associated with increased matrix metalloproteinase-2 activity. *Cancer Res* (2005) 65(16):7378–85. doi: 10.1158/0008-5472.CAN-05-1036

38. Romani C, Cocco E, Bignotti E, Moratto D, Bugatti A, Todeschini P, et al. Evaluation of a novel human IgG1 anti-claudin3 antibody that specifically recognizes its aberrantly localized antigen in ovarian cancer cells and that is suitable for selective drug delivery. *Oncotarget* (2015) 6(33):34617–28. doi: 10.18632/oncotarget.5315

39. Uthayanan L, El-Bahrawy M. Potential roles of claudin-3 and claudin-4 in ovarian cancer management. *J Egypt Natl Canc Inst* (2022) 34(1):24. doi: 10.1186/s43046-022-00125-4

40. Yang F, Xu W, Tang X, Li Q, Hou X, Hui X. Claudin 4 enhances the Malignancy of glioma cells via NNAT/Wnt signaling. *Am J Cancer Res* (2023) 13(6):2530–9.

41. Wang C, Wu N, Pei B, Ma X, Yang W. Claudin and pancreatic cancer. *Front Oncol* (2023) 13:1136227. doi: 10.3389/fonc.2023.1136227

42. Kyuno D, Kojima T, Yamaguchi H, Ito T, Kimura Y, Imamura M, et al. Protein kinase Cα inhibitor protects against downregulation of claudin-1 during epithelial-mesenchymal transition of pancreatic cancer. *Carcinogenesis* (2013) 34(6):1232–43. doi: 10.1093/carcin/bgt057

43. Miwa N, Furuse M, Tsukita S, Niikawa N, Nakamura Y, Furukawa Y. Involvement of claudin-1 in the beta-catenin/Tcf signaling pathway and its frequent upregulation in human colorectal cancers. *Oncol Res* (2001) 12(11-12):469–76. doi: 10.3727/096504001108747477

44. Gowrikumar S, Ahmad R, Uppada SB, Washington MK, Shi C, Singh AB, et al. Upregulated claudin-1 expression promotes colitis-associated cancer by promoting β-catenin phosphorylation and activation in Notch/p-AKT-dependent manner. *Oncogene* (2019) 38(26):5321–37. doi: 10.1038/s41388-019-0795-5

45. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* (2010) 12(5):R68. doi: 10.1186/bcr2635

46. Gerhard R, Ricardo S, Albergaria A, Gomes M, Silva AR, Logullo ÂF, et al. Immunohistochemical features of claudin-low intrinsic subtype in metaplastic breast carcinomas. *Breast* (2012) 21(3):354–60. doi: 10.1016/j.breast.2012.03.001

47. Duarte GM, Almeida NR, Tocchet F, Espinola J, Barreto C, Pinto GA, et al. Claudin-4 expression is associated with disease-free survival in breast carcinoma-in-situ: mean follow-up of 8.2 years. *Clin Breast Cancer* (2018) 18(5):e1111–1111e1116. doi: 10.1016/j.clbc.2018.06.005

48. Logullo AF, Pasini FS, Nonogaki S, Rocha RM, Soares FA, Brentani MM. Immunoexpression of claudins 4 and 7 among invasive breast carcinoma subtypes: A large diagnostic study using tissue microarray. *Mol Clin Oncol* (2018) 9(4):377–88. doi: 10.3892/mco.2018.1685

49. Katayama A, Handa T, Komatsu K, Togo M, Horiguchi J, Nishiyama M, et al. Expression patterns of claudins in patients with triple-negative breast cancer are associated with nodal metastasis and worse outcome. *Pathol Int* (2017) 67(8):404–13. doi: 10.1111/pin.12560

50. Lu S, Singh K, Mangray S, Tavares R, Noble L, Resnick MB, et al. Claudin expression in high-grade invasive ductal carcinoma of the breast: correlation with the molecular subtype. *Mod Pathol* (2013) 26(4):485–95. doi: 10.1038/modpathol.2012.187

51. Jääskeläinen A, Soini Y, Jukkola-Vuorinen A, Auvinen P, Haapasaari KM, Karihtala P. High-level cytoplasmic claudin 3 expression is an independent predictor of poor survival in triple-negative breast cancer. *BMC Cancer* (2018) 18(1):223. doi: 10.1186/s12885-018-4141-z

52. Fougner C, Bergholtz H, Norum JH, Sørlie T. Re-definition of claudin-low as a breast cancer phenotype. *Nat Commun* (2020) 11(1):1787. doi: 10.1038/s41467-020-15574-5

53. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, et al. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc Natl Acad Sci U.S.A.* (2009) 106(33):13820–5. doi: 10.1073/pnas.0905718106

54. Perou CM. Molecular stratification of triple-negative breast cancers. *Oncologist* (2010) 15 Suppl 5:39–48. doi: 10.1634/theoncologist.2010-S5-39

55. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* (2011) 5(1):5–23. doi: 10.1016/j.molonc.2010.11.003

56. Morel AP, Lièvre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS One* (2008) 3 (8):e2888. doi: 10.1371/journal.pone.0002888

57. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* (2008) 133 (4):704–15. doi: 10.1016/j.cell.2008.03.027

58. Morel AP, Hinkal GW, Thomas C, Fauvet F, Courtois-Cox S, Wierinckx A, et al. EMT inducers catalyze Malignant transformation of mammary epithelial cells and drive tumorigenesis towards claudin-low tumors in transgenic mice. *PloS Genet* (2012) 8(5):e1002723. doi: 10.1371/journal.pgen.1002723

59. Asiedu MK, Ingle JN, Behrens MD, Radisky DC, Knutson KL. TGFbeta/TNF (alpha)-mediated epithelial-mesenchymal transition generates breast cancer stem cells with a claudin-low phenotype. *Cancer Res* (2011) 71(13):4707–19. doi: 10.1158/0008-5472.CAN-10-4554

60. Rädler PD, Wehde BL, Triplett AA, Shrestha H, Shepherd JH, Pfefferle AD, et al. Highly metastatic claudin-low mammary cancers can originate from luminal epithelial cells. *Nat Commun* (2021) 12(1):3742. doi: 10.1038/s41467-021-23957-5

61. Chung WC, Zhang S, Challagundla L, Zhou Y, Xu K. Lunatic fringe and p53 cooperatively suppress mesenchymal stem-like breast cancer. *Neoplasia* (2017) 19 (11):885–95. doi: 10.1016/j.neo.2017.08.006

62. Rivero S, Gómez-Marín E, Guerrero-Martínez JA, García-Martínez J, Reyes JC. TBL1 is required for the mesenchymal phenotype of transformed breast cancer cells. *Cell Death Dis* (2019) 10(2):95. doi: 10.1038/s41419-019-1310-1

63. Van Keymeulen A, Lee MY, Ousset M, Brohée S, Rorive S, Giraddi RR, et al. Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* (2015) 525(7567):119–23. doi: 10.1038/nature14665

64. Suárez-Arriaga MC, Méndez-Tenorio A, Pérez-Koldenkova V, Fuentes-Pananá EM. Claudin-low breast cancer inflammatory signatures support polarization of M1-like macrophages with protumoral activity. *Cancers (Basel)* (2021) 13(9):2248. doi: 10.3390/cancers13092248

65. Tao L, Xiang D, Xie Y, Bronson RT, Li Z. Induced p53 loss in mouse luminal cells causes clonal expansion and development of mammary tumours. *Nat Commun* (2017) 8:14431. doi: 10.1038/ncomms14431

66. Wang S, Liu JC, Kim D, Datti A, Zacksenhaus E. Targeted Pten deletion plus p53-R270H mutation in mouse mammary epithelium induces aggressive claudin-low and basal-like breast cancer. *Breast Cancer Res* (2016) 18(1):9. doi: 10.1186/s13058-015-0668-y

67. Knight JF, Lesurf R, Zhao H, Pinnaduwage D, Davis RR, Saleh SM, et al. Met synergizes with p53 loss to induce mammary tumors that possess features of claudin-low breast cancer. *Proc Natl Acad Sci U.S.A.* (2013) 110(14):E1301–10. doi: 10.1073/pnas.1210353110

68. Pommier RM, Sanlaville A, Tonon L, Kielbassa J, Thomas E, Ferrari A, et al. Comprehensive characterization of claudin-low breast tumors reflects the impact of the cell-of-origin on cancer evolution. *Nat Commun* (2020) 11(1):3431. doi: 10.1038/s41467-020-17249-7

69. Zhao S, Ma D, Xiao Y, Li XM, Ma JL, Zhang H, et al. Molecular subtyping of triple-negative breast cancers by immunohistochemistry: molecular basis and clinical relevance. *Oncologist* (2020) 25(10):e1481–1481e1491. doi: 10.1634/theoncologist.2019-0982

70. Ding L, Lu Z, Lu Q, Chen YH. The claudin family of proteins in human Malignancy: a clinical perspective. *Cancer Manag Res* (2013) 5:367–75. doi: 10.2147/CMAR.S38294

71. Coussens LM, Werb Z. Inflammation and cancer. *Nature* (2002) 420(6917):860–7. doi: 10.1038/nature01322

72. Li J, Li YX, Chen MH, Li J, Du J, Shen B, et al. Changes in the phosphorylation of claudins during the course of experimental colitis. *Int J Clin Exp Pathol* (2015) 8 (10):12225–33.

73. Kim BG, Lee PH, Lee SH, Park CS, Jang AS. Impact of ozone on claudins and tight junctions in the lungs. *Environ Toxicol* (2018) 33(7):798–806. doi: 10.1002/tox.22566

74. Ahmad R, Kumar B, Chen Z, Chen X, Müller D, Lele SM, et al. Loss of claudin-3 expression induces IL6/gp130/Stat3 signaling to promote colon cancer Malignancy by hyperactivating Wnt/β-catenin signaling. *Oncogene* (2017) 36(47):6592–604. doi: 10.1038/onc.2017.259

75. Zhou B, Flodby P, Luo J, Castillo DR, Liu Y, Yu FX, et al. Claudin-18-mediated YAP activity regulates lung stem and progenitor cell homeostasis and tumorigenesis. *J Clin Invest* (2018) 128(3):970–84. doi: 10.1172/JCI90429

76. Shimobaba S, Taga S, Akizuki R, Hichino A, Endo S, Matsunaga T, et al. Claudin-18 inhibits cell proliferation and motility mediated by inhibition of phosphorylation of PDK1 and Akt in human lung adenocarcinoma A549 cells. *Biochim Biophys Acta* (2016) 1863(6 Pt A):1170–8. doi: 10.1016/j.bbamcr.2016.02.015

77. Akizuki R, Shimobaba S, Matsunaga T, Endo S, Ikari A. Claudin-5, -7, and -18 suppress proliferation mediated by inhibition of phosphorylation of Akt in human lung squamous cell carcinoma. *Biochim Biophys Acta Mol Cell Res* (2017) 1864(2):293–302. doi: 10.1016/j.bbamcr.2016.11.018

78. Martin TA, Watkins G, Mansel RE, Jiang WG. Loss of tight junction plaque molecules in breast cancer tissues is associated with a poor prognosis in patients with breast cancer. *Eur J Cancer* (2004) 40(18):2717–25. doi: 10.1016/j.ejca.2004.08.008

79. Hoover KB, Liao SY, Bryant PJ. Loss of the tight junction MAGUK ZO-1 in breast cancer: relationship to glandular differentiation and loss of heterozygosity. *Am J Pathol* (1998) 153(6):1767–73. doi: 10.1016/S0002-9440(10)65691-X

80. Chlenski A, Ketels KV, Korovaitseva GI, Talamonti MS, Oyasu R, Scarpelli DG. Organization and expression of the human zo-2 gene (tjp-2) in normal and neoplastic tissues. *Biochim Biophys Acta* (2000) 1493(3):319–24. doi: 10.1016/s0167-4781(00)00185-8

81. Lu Z, Kim DH, Fan J, Lu Q, Verbanac K, Ding L, et al. A non-tight junction function of claudin-7-Interaction with integrin signaling in suppressing lung cancer cell proliferation and detachment. *Mol Cancer* (2015) 14:120. doi: 10.1186/s12943-015-0387-0

82. Singh AB, Uppada SB, Dhawan P. Claudin proteins, outside-in signaling, and carcinogenesis. *Pflugers Arch* (2017) 469(1):69–75. doi: 10.1007/s00424-016-1919-1

83. Rastaldi MP. Epithelial-mesenchymal transition and its implications for the development of renal tubulointerstitial fibrosis. *J Nephrol* (2006) 19(4):407–12.

84. Medici D, Hay ED, Goodenough DA. Cooperation between snail and LEF-1 transcription factors is essential for TGF-beta1-induced epithelial-mesenchymal transition. *Mol Biol Cell* (2006) 17(4):1871–9. doi: 10.1091/mbc.e05-08-0767

85. Vincent T, Neve EP, Johnson JR, Kukalev A, Rojo F, Albanell J, et al. A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition. *Nat Cell Biol* (2009) 11(8):943–50. doi: 10.1038/ncb1905

86. Lin X, Shang X, Manorek G, Howell SB. Regulation of the epithelial-mesenchymal transition by claudin-3 and claudin-4. *PloS One* (2013) 8(6):e67496. doi: 10.1371/journal.pone.0067496

87. Tobioka H, Sawada N, Zhong Y, Mori M. Enhanced paracellular barrier function of rat mesothelial cells partially protects against cancer cell penetration. *Br J Cancer* (1996) 74(3):439–45. doi: 10.1038/bjc.1996.378

88. Martin TA, Jiang WG. Loss of tight junction barrier function and its role in cancer metastasis. *Biochim Biophys Acta* (2009) 1788(4):872–91. doi: 10.1016/j.bbamem.2008.11.005

89. Tokumasu R, Yamaga K, Yamazaki Y, Murota H, Suzuki K, Tamura A, et al. Dose-dependent role of claudin-1 in *vivo* in orchestrating features of atopic dermatitis. *Proc Natl Acad Sci U.S.A.* (2016) 113(28):E4061–8. doi: 10.1073/pnas.1525474113

90. Menard C, Pfau ML, Hodes GE, Kana V, Wang VX, Bouchard S, et al. Social stress induces neurovascular pathology promoting depression. *Nat Neurosci* (2017) 20 (12):1752–60. doi: 10.1038/s41593-017-0010-3

91. Ding L, Lu Z, Foreman O, Tatum R, Lu Q, Renegar R, et al. Inflammation and disruption of the mucosal architecture in claudin-7-deficient mice. *Gastroenterology* (2012) 142(2):305–15. doi: 10.1053/j.gastro.2011.10.025

92. Hagen SJ, Ang LH, Zheng Y, Karahan SN, Wu J, Wang YE, et al. Loss of tight junction protein claudin 18 promotes progressive neoplasia development in mouse stomach. *Gastroenterology* (2018) 155(6):1852–67. doi: 10.1053/j.gastro.2018.08.041

93. Hoevel T, Macek R, Mundigl O, Swisshelm K, Kubbies M. Expression and targeting of the tight junction protein CLDN1 in CLDN1-negative human breast tumor cells. *J Cell Physiol* (2002) 191(1):60–8. doi: 10.1002/jcp.10076

94. Kominsky SL, Argani P, Korz D, Evron E, Raman V, Garrett E, et al. Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in *situ* and invasive ductal carcinoma of the breast. *Oncogene* (2003) 22 (13):2021–33. doi: 10.1038/sj.onc.1206199

95. Swisshelm K, Machl A, Planitzer S, Robertson R, Kubbies M, Hosier S. SEMP1, a senescence-associated cDNA isolated from human mammary epithelial cells, is a member of an epithelial membrane protein superfamily. *Gene* (1999) 226(2):285–95. doi: 10.1016/s0378-1119(98)00553-8

96. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, et al. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep* (2008) 9(6):582–9. doi: 10.1038/embor.2008.74

97. Zou Q, Zhou E, Xu F, Zhang D, Yi W, Yao J. A TP73-AS1/miR-200a/ZEB1 regulating loop promotes breast cancer cell invasion and migration. *J Cell Biochem* (2018) 119(2):2189–99. doi: 10.1002/jcb.26380

98. Yu SJ, Hu JY, Kuang XY, Luo JM, Hou YF, Di GH, et al. MicroRNA-200a promotes anoikis resistance and metastasis by targeting YAP1 in human breast cancer. *Clin Cancer Res* (2013) 19(6):1389–99. doi: 10.1158/1078-0432.CCR-12-1959

99. Kim HK, Park JD, Choi SH, Shin DJ, Hwang S, Jung HY, et al. Functional link between miR-200a and ELK3 regulates the metastatic nature of breast cancer. *Cancers (Basel)* (2020) 12(5):1225. doi: 10.3390/cancers12051225

100. Harrell JC, Pfefferle AD, Zalles N, Prat A, Fan C, Khramtsov A, et al. Endothelial-like properties of claudin-low breast cancer cells promote tumor vascular permeability and metastasis. *Clin Exp Metastasis* (2014) 31(1):33–45. doi: 10.1007/s10585-013-9607-4

101. Greville G, Llop E, Howard J, Madden SF, Perry AS, Peracaula R, et al. 5-AZA-dC induces epigenetic changes associated with modified glycosylation of secreted glycoproteins and increased EMT and migration in chemo-sensitive cancer cells. *Clin Epigenet* (2021) 13(1):34. doi: 10.1186/s13148-021-01015-7

102. Chiang SK, Chang WC, Chen SE, Chang LC. DOCK1 regulates growth and motility through the RRP1B-claudin-1 pathway in claudin-low breast cancer cells. *Cancers (Basel)* (2019) 11(11):1762. doi: 10.3390/cancers11111762

103. Lu Y, Wang L, Li H, Li Y, Ruan Y, Lin D, et al. SMAD2 inactivation inhibits CLDN6 methylation to suppress migration and invasion of breast cancer cells. *Int J Mol Sci* (2017) 18(9):1863. doi: 10.3390/ijms18091863

104. Song P, Li Y, Dong Y, Liang Y, Qu H, Qi D, et al. Estrogen receptor β inhibits breast cancer cells migration and invasion through CLDN6-mediated autophagy. *J Exp Clin Cancer Res* (2019) 38(1):354. doi: 10.1186/s13046-019-1359-9

105. Liu Y, Wang Y, Sun S, Chen Z, Xiang S, Ding Z, et al. Understanding the versatile roles and applications of EpCAM in cancers: from bench to bedside. *Exp Hematol Oncol* (2022) 11(1):97. doi: 10.1186/s40164-022-00352-4

106. Wen R, Lin H, Li X, Lai X, Yang F. The regulatory mechanism of epCAM N-glycosylation-mediated MAPK and PI3K/akt pathways on epithelial-mesenchymal transition in breast cancer cells. *Cell Mol Biol (Noisy-le-grand)* (2022) 68(5):192–201. doi: 10.14715/cmb/2022.68.5.26

107. Campbell CI, Thompson DE, Siwicky MD, Moorehead RA. Murine mammary tumor cells with a claudin-low genotype. *Cancer Cell Int* (2011) 11:28. doi: 10.1186/1475-2867-11-28

108. Fougner C, Bergholtz H, Kuiper R, Norum JH, Sørlie T. Claudin-low-like mouse mammary tumors show distinct transcriptomic patterns uncoupled from genomic drivers. *Breast Cancer Res* (2019) 21(1):85. doi: 10.1186/s13058-019-1170-8

109. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, et al. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res* (2009) 69(10):4116–24. doi: 10.1158/0008-5472.CAN-08-3441

110. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci USA* (2010) 107(35):15449–54. doi: 10.1073/pnas.1004900107

111. Bhowmick NA, Ghiassi M, Bakin A, Aakre M, Lundquist CA, Engel ME, et al. Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. *Mol Biol Cell* (2001) 12(1):27–36. doi: 10.1091/mbc.12.1.27

112. Gregory PA, Bracken CP, Bert AG, Goodall GJ. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle* (2008) 7(20):3112–8. doi: 10.4161/cc.7.20.6851

113. Jones R, Watson K, Bruce A, Nersesian S, Kitz J, Moorehead R. Re-expression of miR-200c suppresses proliferation, colony formation and in *vivo* tumor growth of murine claudin-low mammary tumor cells. *Oncotarget* (2017) 8(14):23727–49. doi: 10.18632/oncotarget.15829

114. Sabatier R, Finetti P, Guille A, Adelaide J, Chaffanet M, Viens P, et al. Claudin-low breast cancers: clinical, pathological, molecular and prognostic characterization. *Mol Cancer* (2014) 13:228. doi: 10.1186/1476-4598-13-228

115. Dias K, Dvorkin-Gheva A, Hallett RM, Wu Y, Hassell J, Pond GR, et al. Claudin-low breast cancer; clinical & Pathological characteristics. *PloS One* (2017) 12(1):e0168669. doi: 10.1371/journal.pone.0168669

116. Zayed H. The identification of highly upregulated genes in claudin-low breast cancer through an integrative bioinformatics approach. *Comput Biol Med* (2020) 127:103806. doi: 10.1016/j.compbiomed.2020.103806

117. de Beça FF, Caetano P, Gerhard R, Alvarenga CA, Gomes M, Paredes J, et al. Cancer stem cells markers CD44, CD24 and ALDH1 in breast cancer special histological types. *J Clin Pathol* (2013) 66(3):187–91. doi: 10.1136/jclinpath-2012-201169

118. Okano M, Oshi M, Mukhopadhyay S, Qi Q, Yan L, Endo I, et al. Octogenarians' Breast cancer is associated with an unfavorable tumor immune microenvironment and worse disease-free survival. *Cancers (Basel)* (2021) 13(12):2933. doi: 10.3390/cancers13122933

119. Jing Xu, Ketao L, Tianhui Su, Zhenfeng Li, Yue W, Qinqin Gu. Clinicopathologic and prognostic features of Claudin-low type breast cancer. *Chin J Pathol* (2017) 46(9):634–9. doi: 10.3760/cma.j.issn.0529-5807.2017.09.009

120. Kim S, Moon BI, Lim W, Park S, Cho MS, Sung SH. Feasibility of classification of triple negative breast cancer by immunohistochemical surrogate markers. *Clin Breast Cancer* (2018) 18(5):e1123–1123e1132. doi: 10.1016/j.clbc.2018.03.012

121. Voutsadakis IA. Comparison of clinical subtypes of breast cancer within the claudin-low molecular cluster reveals distinct phenotypes. *Cancers (Basel)* (2023) 15(10):2689. doi: 10.3390/cancers15102689

122. Alsuliman A, Colak D, Al-Harazi O, Fitwi H, Tulbah A, Al-Tweigeri T, et al. Bidirectional crosstalk between PD-L1 expression and epithelial to mesenchymal transition: significance in claudin-low breast cancer cells. *Mol Cancer* (2015) 14:149. doi: 10.1186/s12943-015-0421-2

123. Thompson DE, Siwicky MD, Moorehead RA. Caveolin-1 expression is elevated in claudin-low mammary tumor cells. *Cancer Cell Int* (2012) 12:6. doi: 10.1186/1475-2867-12-6

124. Balestrieri K, Kew K, McDaniel M, Ramez M, Pittman HK, Murray G, et al. Proteomic identification of tumor- and metastasis-associated galectin-1 in claudin-low breast cancer. *Biochim Biophys Acta Gen Subj* (2021) 1865(2):129784. doi: 10.1016/j.bbagen.2020.129784

125. Fenizia C, Bottino C, Corbetta S, Fittipaldi R, Floris P, Gaudenzi G, et al. SMYD3 promotes the epithelial-mesenchymal transition in breast cancer. *Nucleic Acids Res* (2019) 47(3):1278–93. doi: 10.1093/nar/gky1221

126. Liu Y, Hu Y, Xue J, Li J, Yi J, Bu J, et al. Advances in immunotherapy for triple-negative breast cancer. *Mol Cancer* (2023) 22(1):145. doi: 10.1186/s12943-023-01850-7

127. Taylor NA, Vick SC, Iglesia MD, Brickey WJ, Midkiff BR, McKinnon KP, et al. Treg depletion potentiates checkpoint inhibition in claudin-low breast cancer. *J Clin Invest* (2017) 127(9):3472–83. doi: 10.1172/JCI90499

128. Chang J, Hong L, Liu Y, Pan Y, Yang H, Ye W, et al. Targeting PIK3CG in combination with paclitaxel as a potential therapeutic regimen in claudin-low breast cancer. *Cancer Manag Res* (2020) 12:2641–51. doi: 10.2147/CMAR.S250171

129. Stalker L, Pemberton J, Moorehead RA. Inhibition of proliferation and migration of luminal and claudin-low breast cancer cells by PDGFR inhibitors. *Cancer Cell Int* (2014) 14(1):89. doi: 10.1186/s12935-014-0089-5

130. Saueressig S, Tessmann J, Mastelari R, da Silva LP, Buss J, Segatto NV, et al. Synergistic effect of pyrazoles derivatives and doxorubicin in claudin-low breast cancer subtype. *BioMed Pharmacother* (2018) 98:390–8. doi: 10.1016/j.biopha.2017.12.062

131. Lyu H, Hou D, Liu H, Ruan S, Tan C, Wu J, et al. HER3 targeting augments the efficacy of panobinostat in claudin-low triple-negative breast cancer cells. *NPJ Precis Oncol* (2023) 7(1):72. doi: 10.1038/s41698-023-00422-8

132. Matossian MD, Burks HE, Elliott S, Hoang VT, Bowles AC, Sabol RA, et al. Panobinostat suppresses the mesenchymal phenotype in a novel claudin-low triple negative patient-derived breast cancer model. *Oncoscience* (2018) 5(3-4):99–108. doi: 10.18632/oncoscience.412

133. Feldker N, Ferrazzi F, Schuhwerk H, Widholz SA, Guenther K, Frisch I, et al. Genome-wide cooperation of EMT transcription factor ZEB1 with YAP and AP-1 in breast cancer. *EMBO J* (2020) 39(17):e103209. doi: 10.15252/embj.2019103209

134. Snyder CM, Rohde MM, Fahrenholtz CD, Swanner J, Sloop J, Donati GL, et al. Low doses of silver nanoparticles selectively induce lipid peroxidation and proteotoxic stress in mesenchymal subtypes of triple-negative breast cancer. *Cancers (Basel)* (2021) 13(16):4217. doi: 10.3390/cancers13164217

135. Onaga C, Tamori S, Motomura H, Ozaki A, Matsuda C, Matsuoka I, et al. High SLC20A1 expression is associated with poor prognoses in claudin-low and basal-like breast cancers. *Anticancer Res* (2021) 41(1):43–54. doi: 10.21873/anticanres.14750

136. Levine F, Ogunwobi OO. Targeting PVT1 exon 9 re-expresses claudin 4 protein and inhibits migration by claudin-low triple negative breast cancer cells. *Cancers (Basel)* (2021) 13(5):1046. doi: 10.3390/cancers13051046

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 12 October 2023
DOI 10.3389/fonc.2023.1143163

OPEN ACCESS

EDITED BY
Iman Mamdouh Talaat,
University of Sharjah, United Arab Emirates

REVIEWED BY
Mohammad T. Albataineh,
Khalifa University, United Arab Emirates
Alaa Bawaneh,
The University of Jordan, Jordan

*CORRESPONDENCE
Qingxin Xia
13838173710@139.com
He Zhang
zlyyzhanghe4202@zzu.edu.cn

RECEIVED 12 January 2023
ACCEPTED 28 September 2023
PUBLISHED 12 October 2023

CITATION
Wang Y, Qu D, Zhang Y, Jin Y, Feng Y, Zhang H and Xia Q (2023) Intra-tumoral microbial community profiling and associated metabolites alterations of TNBC.
Front. Oncol. 13:1143163.
doi: 10.3389/fonc.2023.1143163

Intra-tumoral microbial community profiling and associated metabolites alterations of TNBC

Yi Wang[1,2,3], Dingding Qu[1,2,3], Yali Zhang[1,2,3], Yiping Jin[1,2,3], Yu Feng[1], He Zhang[1,2,3]* and Qingxin Xia[1,2,3]*

[1]Department of Pathology, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, China, [2]Henan Medical Key Laboratory of Tumor Pathology and Artificial Intelligence Diagnosis, Zhengzhou, China, [3]Zhengzhou Key Laboratory of Accurate Pathological Diagnosis of Intractable Tumors, Zhengzhou, China

Triple-negative breast cancer (TNBC) presents significant challenges to female health owing to the lack of therapeutic targets and its poor prognosis. In recent years, in the field of molecular pathology, there has been a growing focus on the role of intra-tumoral microbial communities and metabolic alterations in tumor cells. However, the precise mechanism through which microbiota and their metabolites influence TNBC remains unclear and warrants further investigation. In this study, we analyzed the microbial community composition in various subtypes of breast cancer through 16S rRNA MiSeq sequencing of formalin-fixed, paraffin-embedded (FFPE) tissue samples. Notably, *Turicibacter*, a microbe associated with cancer response, exhibited a significantly higher abundance in TNBC. Similarly, mass spectrometry-based metabolomic analysis revealed substantial differences in specific metabolites, such as nutriacholic, pregnanetriol, and cortol. Furthermore, we observed significant correlations between the intra-tumoral microbiome, clinicopathological characteristics, and human epidermal growth factor receptor-2 expression(HER2). Three microbial taxa (*Cytophagaceae, Conexibacteraceae*, and *Flavobacteriaceae*) were associated with tumor-infiltrating lymphocytes(TILs), which are indicative of antitumor immunity. This study creatively utilized FFPE tissue samples to assess intra-tumoral microbial communities and their related metabolic correlations, presenting avenues for the identification of novel diagnostic biomarkers, the development of therapeutic strategies, and the early clinical diagnosis of TNBC.

KEYWORDS

TNBC, triple-negative breast cancer, FFPE, formalin-fixed paraffin-embedded, microbiota, metabolome, tumor-infiltrating lymphocytes

Introduction

Breast cancer (BC) is the most prevalent malignant tumor among females, and its incidence and mortality rates are on the rise, including that of triple-negative BC (TNBC) (1). TNBC is characterized by the absence of the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER2). Compared to non-TNBC, TNBC is the most aggressive subtype of BC and currently faces limited treatment options (2, 3). Owing to the differences in clinical manifestations between TNBC and non-TNBC, exploring the correlated mechanisms for developing novel therapeutic strategies and improving the prognosis of patients with TNBC is imperative.

Cancer progression is influenced by changes in various components of the tumor microenvironment (TME). Alterations in stromal composition, including the surrounding immune cells, lymphocytes, blood vessels, extracellular matrix, fibroblasts, and certain signaling molecules, can impact host metabolism, immune responses, and cancer-driving molecular alterations, thereby influencing tumor development and response to cancer therapy (4). Recent research has highlighted the complexity and significance of the relationship between the microbiome and cancer. Certain microbial species within tumors, as a host factor, can stimulate an inflammatory state or immune response, thereby promoting carcinogenesis and tumor progression (5). Furthermore, the diversity and composition of the bacterial community have been associated with different histological classifications of cancer, reflecting the distinct TME characteristics (6, 7). Consequently, the tumor microbiota may not only serve as a diagnostic tool for better cancer classification but also influence tumor behavior and patient prognosis based on the properties of the microbes themselves (8). Emerging evidence suggests a potential link between the microbiota and carcinogenic metabolites that contribute to tumor progression (9). As "tumor foragers," the tumor microbiota plays a crucial role in regulating the host metabolome by establishing a biological "digester" (10). Cancer metabolism can modulate the TME to facilitate cancer progression through the release of amino acids, nucleotides, organic acids, and lipids that fulfill the metabolic demands of the body.

Studies have provided evidence of a significant association between microbial composition and the development of BC. Urbaniak et al. demonstrated that patients with BC exhibited a relatively high abundance of *Lactobacillus, Hydrogenophaga*, and *Fusobacterium* compared to those with benign breast lesions or normal tissue (11). More importantly, the gastrointestinal microbiota plays a crucial role in regulating estrogen levels, and estrogen, in turn, influences BC development through host–microbe interactions (12). To investigate this further, we first analyzed the differences in microbial composition between non-TNBC and TNBC through 16S rRNA sequencing of formalin-fixed, paraffin-embedded (FFPE) BC tissue samples. Additionally, to explore the interactions between microbes and metabolites, we examined metabolite abundance in FFPE BC tissue samples through liquid chromatography–mass spectrometry (LC–MS) analysis, seeking to identify relevant metabolic signaling pathways that may potentially be involved in BC molecular mechanisms. Our findings may aid in the discovery of novel biomarkers in TNBC and pave the way for the development of effective therapeutic strategies.

Abbreviations: BC, breast cancer; TNBC, triple-negative breast cancer; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor-2; TME, tumor microenvironment; FFPE, formalin-fixed paraffin-embedded; HE, hematoxylin–eosin; TILs, tumor infiltrating lymphocytes; rRNA, ribosomal RNA; OTU, operational taxonomic unit; PCoA, principal coordinate analysis; db-RDA, distance-based redundancy analysis; LEfSe, linear discriminant analysis (LDA) coupled with effect size; LC-MS, liquid chromatography-mass spectrometry; PCA, principal component analysis; OPLS-DA, orthogonal partial least square-discriminant analysis; VIP, variable importance in projection.

Materials and methods

Human BC tissue samples

Surgical specimens of BC were collected from the Zhengzhou University Affiliated Cancer Hospital (Zhengzhou, China) between 2014 and 2016. Overall the clinical variables between TNBC group and non-TNBC group were comparable, with no significant difference in the age, BMI, tumor size and parity (Supplementary Table 1). FFPE tissue samples were sliced into 5-mm-thick sections and stained with hematoxylin and eosin (HE). Slices were evaluated by two or more pathologists. The study was approved by the ethics committee, and all patients provided informed consent. The patients from whom the samples were obtained did not undergo standard BC therapy, i.e., chemotherapy and/or radiotherapy.

Immunohistochemistry and tumor-infiltrating lymphocytes (TILs) evaluation

We assessed the protein expression levels of ER, PR, and HER2 in the BC tissue samples. The tissue sections were stained using the Ventana BenchMark ULTRA automatic immunohistochemical staining platform (Ventana Medical Systems Inc., Tucson, AZ, USA) and observed under a microscope (Olympus BX41). Rabbit monoclonal primary antibodies against ER (SP1 Roche), PR (1E2 Roche), HER2/NEU (Clone 4B5 Roche), and PD-L1 (SP142 Roche) were used, and the OptiView DAB immunohistochemistry Detection Kit and OptiView Amplification Kit (Ventana Medical Systems Inc.) were used for subsequent analysis. HER2 staining was scored according to the HER2 Testing Guidelines for Breast Cancer (2019 edition) (13). ER- and PR-positive staining was defined according to the ASCO/CAP guidelines (14).

In the HE slides, TILs were defined as a continuous parameter by two experienced pathologists. TILs on the boundaries of the cancer were included, while those in the tumor bed were excluded, and they were scored based on the area occupied over the entire region. The final percentage of TILs was calculated as the average of the specimens and was not restricted to hotspots. All evaluations were performed according to the criteria recommended by the

International TILs Working Group (2014) (15). TILs were assessed as a continuous parameter, and reported scores were rounded up to the nearest 10%.

DNA extraction and high-throughput 16S rRNA gene sequencing

We performed 16S rRNA sequencing on 22 BC samples, comprising 13 TNBC and 9 non-TNBC samples. For microbiota analysis, 10-μm thick sections were used. Total DNA was extracted using the QIAamp DNA FFPE Tissue Kit (QIAGEN, Redwood City, CA, USA) according to the manufacturer's protocol. The quality of the extracted DNA was assessed by subjecting each sample to 1% agarose gel electrophoresis at room temperature. DNA concentration and purity were determined using a NanoDrop 2000 spectrophotometer. To amplify the bacterial 16S rRNA gene V3–V4 region, we used the following primers: 338F: 5′-ACTCCTACGGGAGGCAGCAG-3′ and 806R: 5′-GGACTAC HVGGGTWTCTAAT-3′. The mixed PCR products were purified and quantified using a Quantus™ fluorometer (Promega). Subsequently, we constructed a database using the NEXTFLEX Rapid DNA-SEQ Kit and performed sequencing with 2 × 250 bp chemistry on the Illumina MiSeq PE300 platform (Illumina, San Diego, CA, USA) (16, 17).

Sequencing data analysis

The obtained gene sequences were attached to unique bar codes and clustered into operational taxonomic units (OTUs) with 97% identity, utilizing the USEARCH software (version 7.0, http://drive5.com/uparse/). Each sequence was then compared with the Silva database (SSU132) using the RDP classifier (http://rdp.cme.msu.edu/), and the comparison threshold was set to 70% to obtain the annotation results for species classification. To evaluate species richness based on OTU values, we performed dilution curve analysis (17). Principal coordinate analysis (PCoA) was conducted using the R package (http://www.r-project.org/) to assess the differences in microbiota between the groups (18). For exploring the extent to which certain the microbiota, distance-based redundancy analysis was performed at the OTU level, employing Bray–Curtis distances. Bacterial abundance and diversity were compared using an independent t-test. To evaluate differentially abundant taxa, we used linear discriminant analysis coupled with effect size (LEfSe) (19).

Metabolite extraction and LC–MS untargeted metabolomic analysis

Metabolites from the 22 FFPE samples were extracted by preparing 20-μm thick tissue sections. Details regarding sample preparation, LC–MS analysis, data quality management, and compound identification can be found in the Supplementary Information (20). For LC–MS analysis, we employed an ultra-high-performance liquid chromatography-triple time-of-flight mass spectrometry system (AB SCIEX LLC). The LC–MS data were then imported into the metabolomics processing software Progenesis QI (Waters Corporation, Milford, MA, USA).

Statistical analyses

Statistical analysis was conducted using SPSS software (version 22.0; SPSS Inc., Chicago, IL, USA). Clinical characteristics were assessed using the χ^2 test, while t-tests were utilized to determine differences between two groups. Pearson's correlation was used to analyze the correlation between microbial species at the phylum level and relevant environmental factors or metabolites, with the numerical matrix visually displayed using heatmaps. The color depth in the heatmap corresponds to the size of the data. Statistical significance was considered at $P < 0.05$.

Results

Abundance and diversity of microbiota in FFPE BC tissue samples

To investigate the microbial abundance and diversity in BC, all paraffin-embedded tissue samples were subjected to 16S rRNA sequencing. To eliminate any potential contaminants, simultaneous sample detection was performed after quality filtering. The rarefaction curves of all the samples (Figures S1A, B) validated the adequacy of the sampling efforts. A Venn diagram revealed 668 OTUs, including 211 overlapping OTUs (Figure 1A). Alpha diversity, based on Shannon, Sob, Simpson, and Chao indices, was significantly lower in patients with TNBC than in the other groups ($P < 0.0001$, 0.0032, 0.0190, and 0.0102, respectively) (Figure 1B). Beta-diversity was calculated using unweighted UniFrac at the OTU level, and PCoA showed that tumor microbial communities varied among the samples ($P = 0.046$) (Figure 1C). These results suggest that the diversity of the tumor microbiota varies across BC subtypes.

Alterations in tumor microbiota composition are associated with TNBC

To identify the composition of the intra-tumoral microbial community in each sample, we compared phylotypes with an abundance greater than 0.01% of the total OTU. *Proteobacteria* was the predominant phylum, accounting for 88.4% and 87.2% in the two groups. *Actinobacteria* (3.6% and 3.3%), *Firmicutes* (2.5% and 3.9%), and *Bacteroidetes* (0.97% and 1.67%) were enriched at the phylum level (Figure S1C). The abundance at the genus and OTU levels in the TNBC group differed from that in the non-TNBC group (Figure 2A; Figure S1D). Student's *t*-test was performed to analyze the differences in the microbial communities between the two groups. The abundance of *Firmicutes, Enterobacteriaceae*, and *Weekselllaceae* was lower in the TNBC group than in the non-

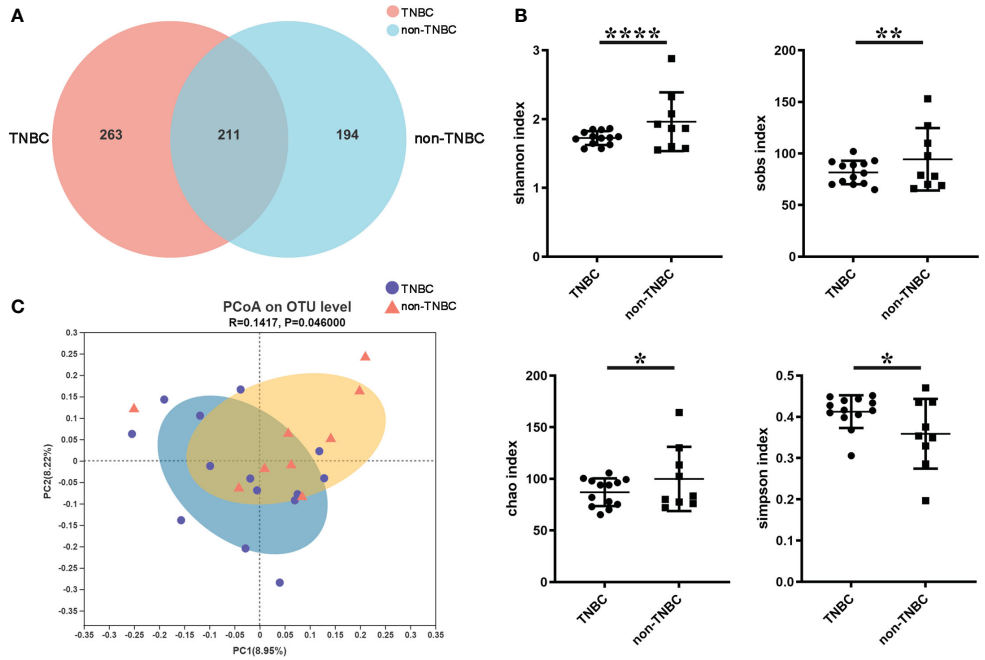


FIGURE 1
Tumor microbial diversity in TNBC and non-TNBC by FFPE. **(A)** A Venn diagram displaying the overlaps and unique OTUs between TNBC and non-TNBC groups. 263 of 474 OTUs were unique in TNBC. **(B)** Shannon index, Sobs index, Chao index, Simpson index estimated microbial diversity in the two groups. **(C)** Beta diversity was calculated based on unweighted unifrac by PCOA.

TNBC group (P = 0.039, 0.004, and 0.047, respectively), whereas the abundance of *Pseudonocardiaceae* was significantly increased (P = 0.005) (Figures 2B, C). Figure 2D shows similar results. LEfSe analysis revealed enriched microbial abundance in the two groups (Figures 2E, F). *Turicibacter* was significantly more abundant in the TNBC group than in the other group and may be a key factor associated with inflammatory and cancer responses in patients with TNBC. Therefore, the diversity and richness in the TNBC group appear to be much lower compared to those in the non-TNBC group.

Composition of detected metabolites in FFPE tissue samples of the TNBC and non-TNBC groups

We hypothesized that metabolites may be affected by the microbiota in tumors. To investigate this, we first assessed the component superclasses of the metabolites and determined their distribution in FFPE samples (Figure 3A). Figure 3B shows the proportions of 24 different steroids and steroid derivatives in the component classes. By comparison with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, we classified genes according to their functions (Figure 3C, D). Subsequently, we visualized the metabolite abundance between the two groups using a heatmap and found that both groups had similar metabolic abundances. Among these metabolites, steroids and steroid derivatives exhibited relative richness, suggesting that they might exert a notable influence on BC (Figure 3E).

Differentially abundant metabolites between TNBC and non-TNBC FFPE tissue samples

Next, we compared the metabolite abundance between the FFPE tissue samples of the TNBC and non-TNBC groups through LC–MS analysis. The differential abundance between the two groups was determined using a permutation t-test (Figure 4A). PLS-DA revealed differences between the two groups based on the first two principal components (PC1: 16%; PC2: 8.84%) (Figure 4B). Compared to the non-TNBC group, the TNBC group exhibited a lower abundance of metabolites. As shown in Figure S3, several specific metabolites differed between the TNBC and non-TNBC groups, and variables with higher VIP scores were considered important for classification.

To gain a comprehensive understanding of the detected metabolites in the FFPE tissue samples, we performed metabolite categorization (superclass, class, subclass, and metabolic pathway) based on the Human Metabolome Database (http://www.hmdb.ca/) and KEGG (http://www.genome.jp/kegg). Affected metabolic pathways were identified by KEGG topology analysis (Figure 4C). Enrichment analysis revealed that these metabolic pathways may influence the biological behavior of BC. Given the loss of hormone expression in the TNBC group, we specifically investigated the

10.3389/fonc.2023.1143163

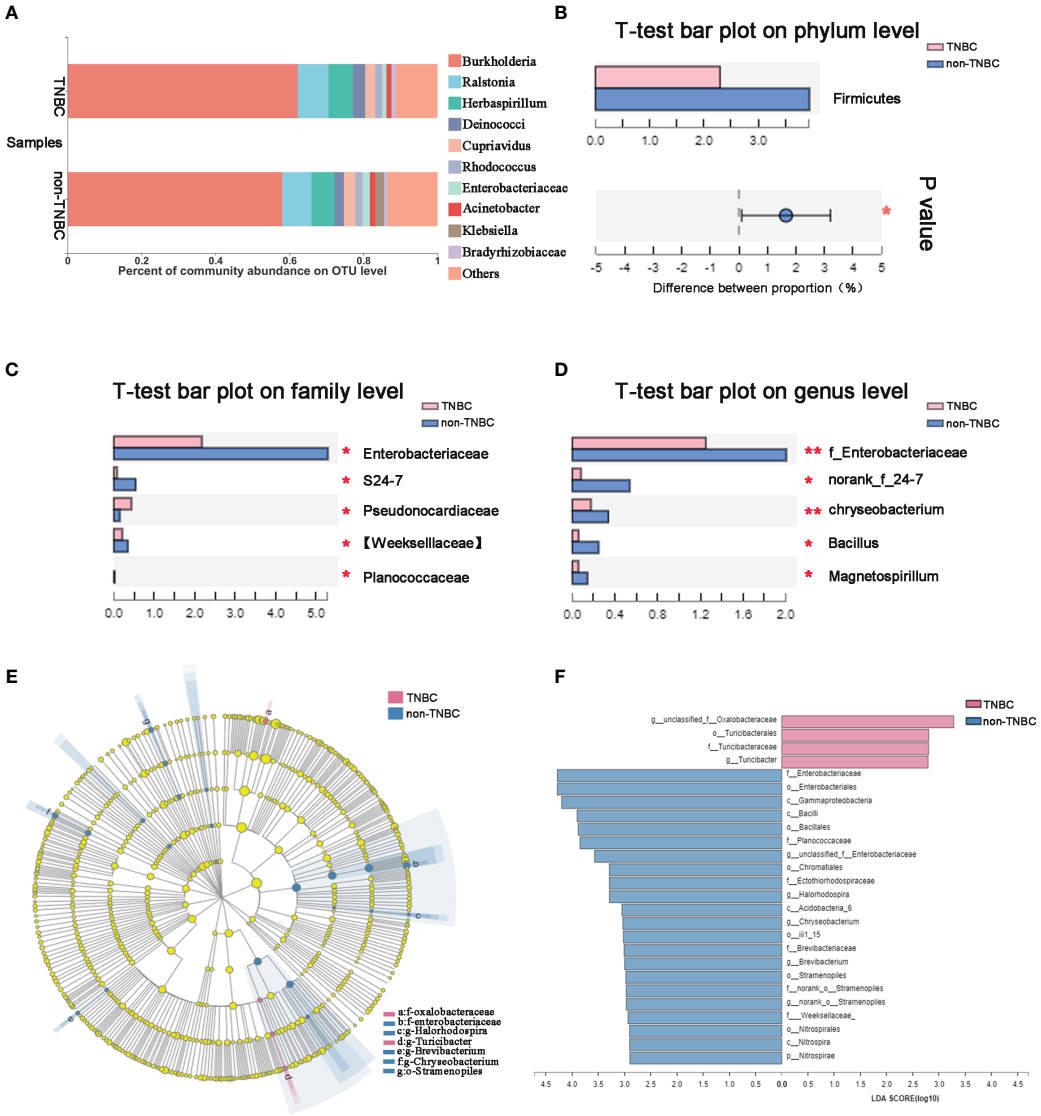


FIGURE 2
Tumor microbial communities are different between TNBC and non-TNBC. **(A)** Composition of microbiota at the OTU level between TNBC and non-TNBC. Significantly altered tumor microbiota between the TNBC and non-TNBC groups at **(B)** phylum, **(C)** family and **(D)** genus levels is represented by the Wilcoxon rank-sum test. f, family. *P<0.05; **P<0.01. **(E)** The specific characterization of tumor microbiota was analyzed by linear discriminant analysis (LDA) effect size (LEfSe) method (http://huttenhower.sph.harvard.edu/lefse/) between TNBC and non-TNBC. Each node represents a specific taxonomic type. Yellow nodes show there is no difference between non-TNBC and TNBC; red nodes show the taxonomic types with more abundance in TNBC group, while blue nodes represent the taxonomic features with more abundance in non-TNBC group. **(F)** LDA score computed from features differentially abundant between TNBC and non-TNBC. The criteria for feature selection is log LDA score > 2.

metabolite profiles corresponding to the steroid and steroid derivative classes between the two groups. Nutriacholic acid, 17beta-estradiol -2,3-quinone, and pregnanediol were found to be more abundant in the TNBC group than in the non-TNBC group, whereas cortol and androsterone glucuronide were more abundant in the non-TNBC group (Figure 4D). However, androstenol, ethisterone, and cortolone levels were not significantly different between the two groups. These metabolites exhibiting statistically different abundances may serve as robust markers and contribute to our understanding of the biological characteristics of patients with TNBC. Overall, these results strongly suggest a TNBC group-specific metabolomic abundance, signatures, and metabolic differences.

Correlations between tumor microbiota and differentially abundant metabolites in BC

To investigate the potential relationship between tumor microbiota and metabolites, we examined the correlations between several bacteria at the phylum level and certain metabolites (Figure 5A). The results revealed a positive correlation between several microbes (*Acidobacteria* and *Firmicutes*) and the levels of betaine. Conversely, we observed a negative correlation between several microbes (*Firmicutes* and *Bacteroidetes*) and the abundance of the metabolite cortol in the

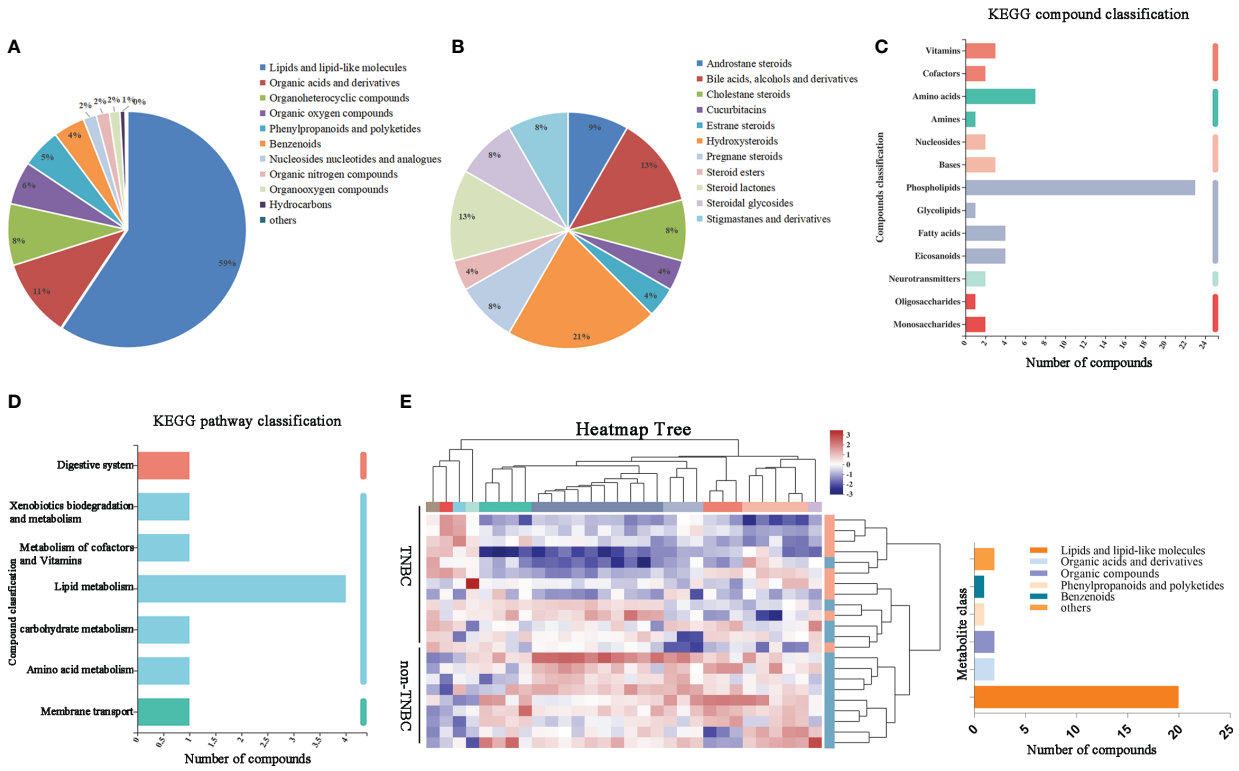


FIGURE 3
Identified tumor metabolites composition and diversity in BC by FFPE. **(A)** Pie chart based on counts of HMDB chemical taxonomy for different classes of metabolites in all breast cancer samples. **(B)** Pie chart based on counts of HMDB chemical taxonomy in Steroids and Steroid derivative content. **(C)** KEGG compound classification and **(D)** KEGG pathway classification: metabolites detected and annotated in breast cancer FFPE tissues. **(E)** Hierarchical clustering analysis for identification of different metabolites by comparison of the TNBC and non-TNBC group. Each column in the figure represents a sample, the expression level of the samples is indicated as a colored band on top of the heat map.

FFPE tissue samples. Similar results were obtained for several genera at the family and genus levels, which exhibited correlations with certain metabolites as determined by Pearson analysis (Figures S4A, B). Lipids and lipid-like molecules of metabolic species were more closely associated with microorganisms, implying that they may potentially influence molecular mechanisms and related pathways in BC.

Correlation between clinical indices and the BC microbiome

We also investigated the correlations between the microbiota and various clinical parameters, including pathological grade, tumor size, metastasis, lymphatic metastasis, survival, TNBC status, TILs, and the expression of HER2. These clinical parameters were correlated with the bacterial genus, as shown in the heatmap in Figure 6A. Notably, *Thermus* showed a positive correlation with pathological grade (r = 0.464, P = 0.029). Moreover, we observed a positive correlation between HER2 expression and *Klebsiella* (r = 0.500, P = 0.017) and *Staphylococcus* (r = 0.462, P = 0.030), and a negative correlation was observed between HER2 expression and *Burkholderia* (r = -0.469, P = 0.027). We also analyzed correlations between TILs and microbes. TILs were positively correlated with *Clostridiales* (r = 0.452, P = 0.034), *Bacteroidales* (r = 0.496, P = 0.018), and *Azospirillum* (r = 0.431, P = 0.044), but negatively correlated with *Streptophyta* (r = -0.161, P = 0.024). The correlation between PD-L1 expression, lymph node metastasis, TILs, distant metastasis, and microbes (Figures S1E–H) further emphasized the notable role of tumor microbes in patients with BC. As shown in Figure 6B, RDA at the OTU level revealed a relationship between the intra-tumoral microbial community and certain clinical indices. TILs were positively correlated with bacteria such as *Cytophagaceae*, *Conexibacteraceae*, and *Flavobacteriaceae* (Figures S5A–C). PCoA did not reveal significant differences in bacterial communities, short-term survival, long-term survival, TIL status, or lymph node metastasis between the groups (Figures S2C, E, G). Tumor microbial characteristics were analyzed through LEfSe, which revealed marked differences in the predominance of bacterial communities (Figures S2D, F, H). The samples were categorized into four groups based on the presence or absence of TNBC and TILs. The beta diversity among the groups displayed significant differences (P = 0.0002) (Figure S2A), and the abundance of the bacterial communities also showed marked differences (Figure S2B). These results further highlight the importance of microbial abundance and diversity between TNBC and non-TNBC, with TILs potentially serving as a prognostic indicator of microbial diversity in TNBC. To some extent, these results validate the critical communication between microbiota, metabolites in the TME, and clinical factors.

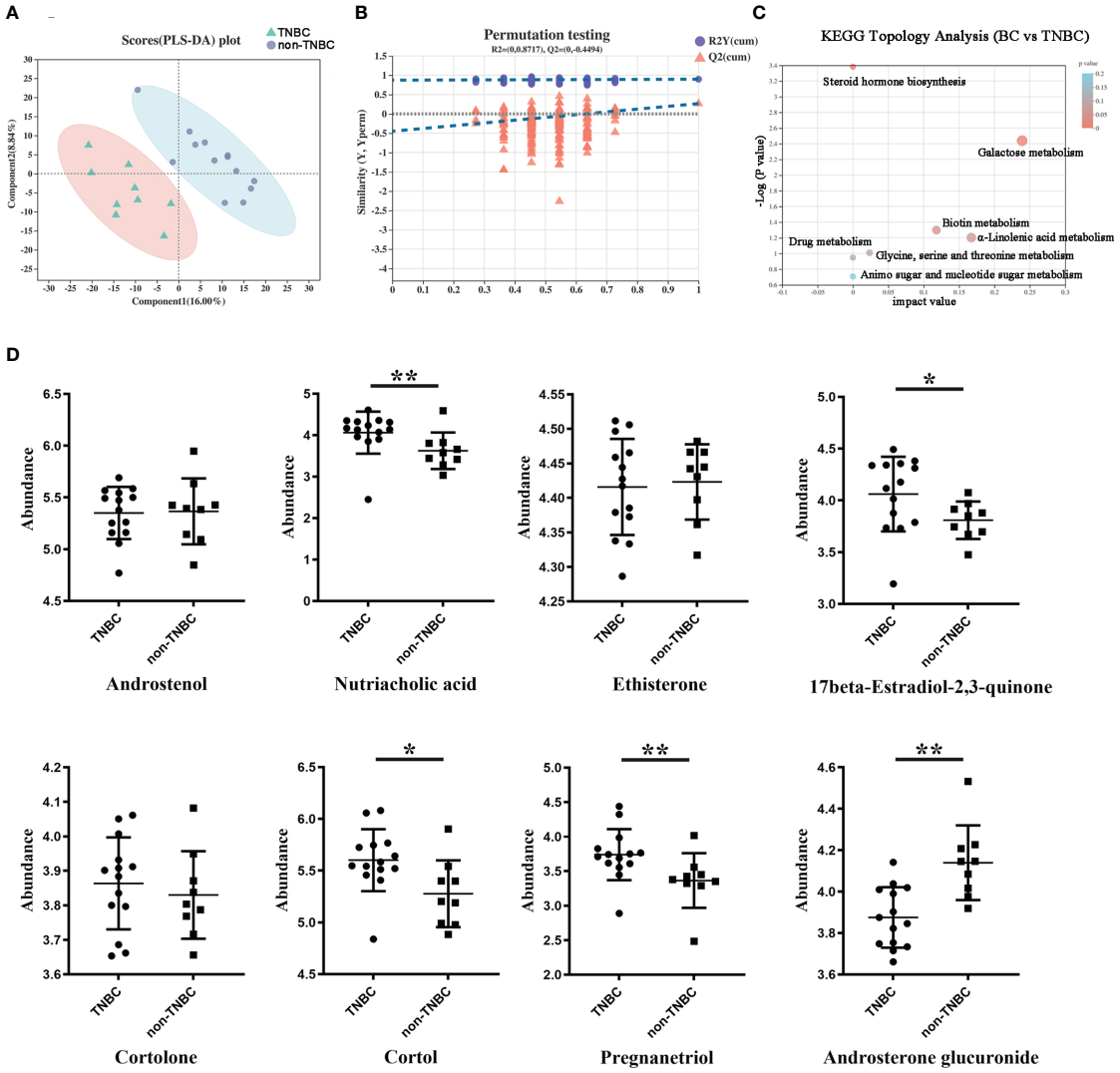


FIGURE 4
Differentially abundant metabolites between TNBC and non-TNBC FFPE samples. **(A)** PLS-DA analysis displaying the two group's classification comparison by the first two PCs. **(B)** PLS-DA model evaluation by permutation test. **(C)** EGG topology analysis shows that metabolites in the TNBC and non-TNBC groups have differentially accumulated [impact value on X-axis] and have significantly changed [-log10(p) on Y-axis]. Bubble size represents impact value; the bigger the bubble, the more important the pathway. **(D)** The metabolites difference from different subclass between the TNBC and non-TNBC groups. (*P < 0.05; **P < 0.01).

Discussion

Recent studies have underscored the critical importance of the relationship between tumor microbiota and the TME in understanding tumor development. BC tissues exhibit distinct microbiomes with an enrichment of specific species. In particular, TNBC demonstrates a unique microbial microenvironment that profoundly influences the biological behavior of tumors. Despite advancements in understanding the underlying tumor biology of TNBC, clinical outcomes for patients with TNBC unfortunately remain poor (21). Therefore, the establishment of a novel microbial typing system for TNBC is essential to providing valuable insights and assistance for patients with TNBC. Our study is the first to investigate the impact of intra-tumoral microbiota on TNBC using FFPE tissue samples and to validate the differences in microbiota between TNBC and non-TNBC. In our previous study, we explored the potential relationship between hepatocellular carcinoma and tumor microorganisms using paraffin-embedded tissue specimens and found that *Pseudomonas* is a differentially abundant microbe between cancer and adjacent tissues, thereby presenting potential avenues for the early clinical diagnosis and treatment of liver cancer (6). In this study, we analyzed a total of 22 BC FFPE tissue specimens through 16S rRNA MiSeq sequencing. The results showed that TNBC exhibited lower abundance of various microbes and that the microbiome diversity displayed significant differences in alpha and beta diversity owing to tumor heterogeneity

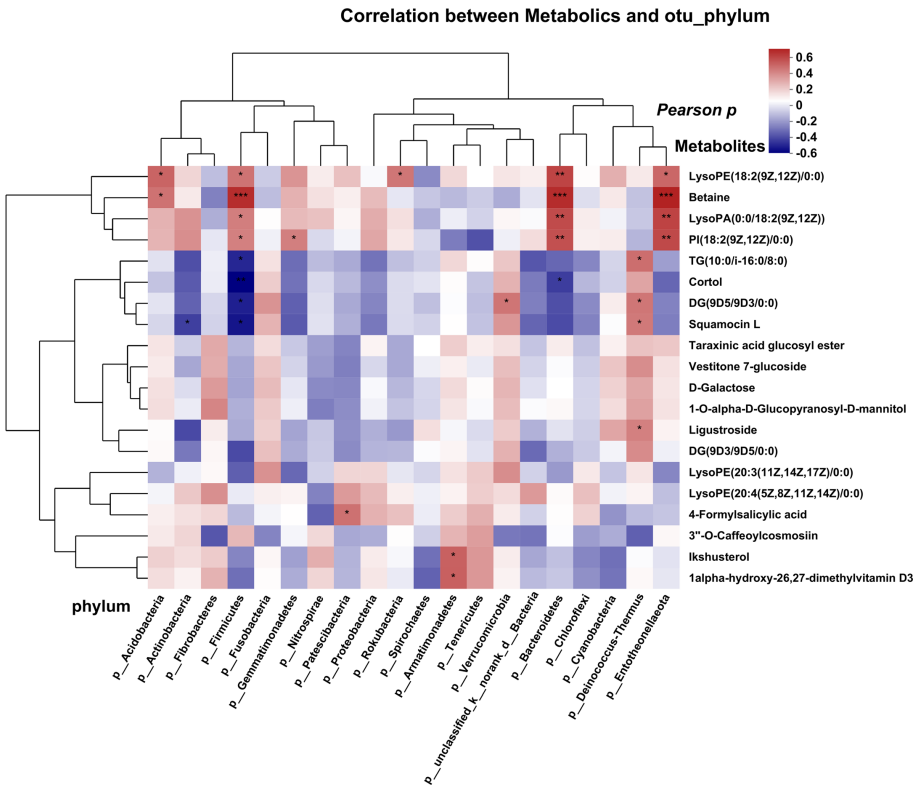


FIGURE 5
Correlation analysis of microbes and metabolites, each lattice represents a coefficient by Pearson's correlation analysis, each row represents a phylum, each column represents a metabolite. Red represents a positive correlation and blue represents a negative correlation. (*P < 0.05; **P < 0.01, ***P<0.001).

and distinct molecular subtypes of BC. Notably, specific bacteria, such as *Firmicutes* and *Enterobacteriaceae*, were significantly decreased in TNBC, and the microbial diversities were significantly different between TNBC and non-TNBC. Firmicutes, an adipocyte-derived bacteria, can directly or indirectly impact BC tissue through toxin or enzyme production and is associated with bacterial load and immune cell infiltration in BC (22, 23). Jeongshin et al. showed that the abundance of Firmicutes can influence

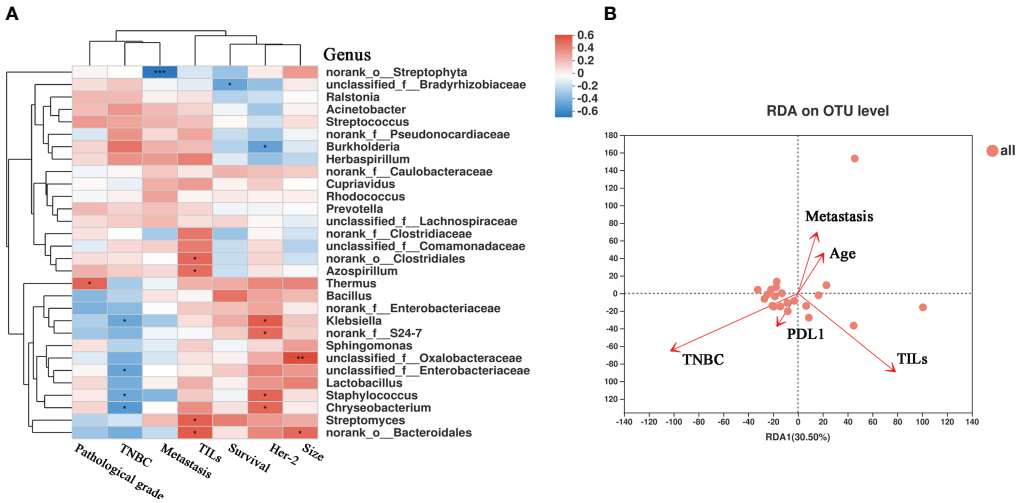


FIGURE 6
The correlation heatmap between clinical relative indices and the microbiome in breast cancer. **(A)** Clinical relative indices include pathological grade, TNBC or not, distant metastasis states, survival, HER-2 state and size. Red represents a positive correlation and blue represents a negative correlation. **(B)** Distance-based Redundancy Analysis(db-RDA) plot showing the relationship of age, metastasis, PD-L1 expression, TILs, and TNBC to the microbial community structure. (*P < 0.05; **P < 0.01).

diseases related to obesity in patients with BC and is associated with a poor prognosis, which aligns with our findings (24). Firmicutes thus hold potential as a diagnostic marker and risk factor for BC. Similarly, the family Enterobacteriaceae, which includes *Escherichia coli*, exhibited a relatively lower abundance in patients with TNBC, indicating decreased involvement of signaling pathways in the TNBC microenvironment (10). The differences in microbial communities among different BC subtypes suggest that microbial markers may serve as noninvasive diagnostic tools and guide the development of therapeutic strategies for patients with BC.

The microbiota can influence the regulation of various metabolic pathways associated with energy homeostasis, nutritional intake, and immune balance (22). Metabolites produced by these bacterial species can profoundly impact molecular events in BC (25). Recent studies have elucidated unique and common viruses, bacteria, and fungi, and various metabolic pathways exhibit distinct patterns in each type of BC (26). Wang et al. demonstrated that the crosstalk between microbiota, metabolites, and the immune system could serve as a novel therapeutic strategy for TNBC (27). They discovered a new metabolite, trimethylamine N-oxide, which affects the treatment of TNBC. Additionally, 17beta-estradiol-2,3-quinone, a reactive metabolite of estrogen, is considered responsible for estrogen-induced genotoxicity and serves as a significant predictor of BC. 17beta-estradiol-2,3-quinone can convert to estrogen catechols and undergo oxidation to form quinones. Accumulation of estrogen quinones along with the DNA damage contribute to estrogen-induced carcinogenesis. In our study, we found that 17beta-estradiol-2,3-quinone was more abundant in the TNBC group, suggesting that the disparity in estrogen disposition and the subsequent elevation of the cumulative quantity of 17beta-estradiol-2,3-quinone in the body may play a role in the development of BC (28, 29). Furthermore, lipid derivatives associated with microbes may effectively reduce BC risk. When changes occur in lipid pathways, it can affect the availability of structural lipids for membrane synthesis, lipid synthesis, and degradation that contribute to energy homeostasis and cell signaling functions (30). We observed a relationship between tumor microbiota and metabolites, with *Bacillus* showing a significant difference between the TNBC and non-TNBC groups and being associated with betaine metabolites. Regrettably, no other microorganism with differential abundance was found to be associated with metabolites. Moreover, the relationship between tumor microbiota and metabolites in different subtypes of BC was not adequately compared. These results suggest that a deeper understanding of the correlation between microbes and their metabolites in BC samples may offer valuable insights for the development of diagnostic, therapeutic, and preventive strategies.

In line with our results, TNBC exhibits unique clinicopathological features that can influence therapeutic decisions (Supplementary Table 2). Pathological grade and TILs exhibited statistically significant differences in the TNBC group (31). Additionally, we elucidated an association between clinicopathological factors and the BC microbiota. The results showed that certain microbes and metabolites were significantly correlated with survival, lymphatic metastasis, distant metastasis, TNM stage, pathology grade, HER2 expression, ER status, and PR status. For instance, *Klebsiella, Staphylococcus, Burkholderia*, and *Thermus* were found to be correlated with HER2 and pathological grade. Meng et al. also highlighted significant differences in tumor microbiota and metabolism among patients with BC with different histological grades (23). Importantly, amino acids and fatty acids displayed the most pronounced differences between TNBC and non-TNBC, consistent with the results of other studies (32). TILs provide novel insights into the crosstalk between microbiota and metabolites, which may potentially influence therapeutic strategies and benefit patients with TNBC. TILs play a crucial role in the response to immune checkpoint inhibitor therapy by increasing PD-L1 expression and are closely associated with the prognosis of BC, especially TNBC (33). The microbiota can regulate estrogen metabolism and tumor immune activation. TILs were significantly different between the TNBC and non-TNBC groups. This study is the first to elucidate that *Clostridiales, Bacteroidales, Azospirillum*, and *Streptophyta* are correlated with TILs. The association between TILs and PD-L1 could have implications for prognostics and may lead to the discovery of new targets for chemotherapy in BC (34, 35). Overall, our findings indicate an innate relationship among the ecological environment, immunity, and treatment efficiency in patients with BC. Additionally, besides metabolism, the microbiota could be utilized to monitor the efficiency of chemotherapy or chemotherapy resistance (36).

In conclusion, through FFPE tissue samples, our study highlights the antitumor role of intra-tumoral microbiota in TNBC and indicates a correlation between bacterial biomass in the tumor and clinical factors. Additionally, our results indicate that tumor microbiota possibly modulates the immune microenvironment and elicits an antitumor response through TILs. Furthermore, our study provides novel insights into the interplay between the microbiome and metabolome, which may pave the way for the discovery of clinical diagnostic indices for BC. Nevertheless, this study has several limitations. First, the sample size was relatively small, which might have hindered the detection of a potential relationship between the microbiome and the prognosis of patients with BC. While several positive results were obtained, a larger sample size will be necessary in future research to validate our findings. Second, considering the detailed study of differences across BC subtypes, another limitation of the study is the absence of a control group comprising adjacent normal breast tissues for a comparison of the microbial and metabolic differences between benign and cancerous breast tissues, which could corroborate the results of this study. In addition, this study primarily focused on the direct influence of microbial metabolites on BC tissues to modulate TME. However, validating whether microbial metabolites affect tumor cell function requires cytological and molecular mechanistic experiments. Despite these limitations, we propose that biomarkers targeted at tumor microbiota or metabolites could hold promise as diagnostic and therapeutic tools for TNBC.

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 Cancer Hospital of Zhengzhou University of the ethics committee. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from Surgical samples of inpatients in our hospital. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements. The studies were conducted in accordance with the local legislation and institutional requirements.

Author contributions

This work was designed by YW, HZ and QDD. ZH and DQ provided essential reagents and materials. YJ collected clinical samples. DQ and HZ conducted laboratory assays. YZ performed the statistical analyses and interpreted data. YW drafted the manuscript. QX and HZ revised the study. YZ and YF performed the statistical analyses and interpreted data. All authors contributed to the article and approved the submitted version.

Funding

This work was sponsored by grants obtained from the National Natural Science Foundation of China (Grant No.31901233), Medical Science and Technique Foundation of Henan Province (Grant No. LHGJ20200177), and the funding program for Medical Sciences and Technique Foundation of Henan Province (Grant No. LHGJ20190672), and Funding Program for the key project of Medical Sciences and Technique Foundation of Henan Province (Grant No. SBGJ202102068).

Acknowledgments

We are especially grateful to all the patients and their families who contributed to the data that made this study possible. We thank Shanghai Majorbio Bio-Pharm Technology Co., Ltd., China for technical support.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1143163/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Composition and differences of tumor microbiota in different clinical indices.

SUPPLEMENTARY FIGURE 2
The diversity and abundance of microbiota in other related elements.

SUPPLEMENTARY FIGURE 3
Variable importance in projection (VIP) scores of tumor metabolites in TNBC and non-TNBC groups.

SUPPLEMENTARY FIGURE 4
Correlation analysis of microbes and metabolites.

SUPPLEMENTARY FIGURE 5
The correlation between the intra-tumoral microbiota and clinicopathological characteristics.

References

1. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2015) 26 Suppl 5:v8–30. doi: 10.1093/annonc/mdv298

2. Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. *Asian Pac J Cancer Prev* (2016) 17 (S3):43–6. doi: 10.7314/APJCP.2016.17.S3.43

3. Vuong D, Simpson PT, Green B, Cummings MC, Lakhani SR. Molecular classification of breast cancer. *Virchows Arch* (2014) 465(1):1–14. doi: 10.1007/s00428-014-1593-7

4. Vitorino M, Almeida Baptista S, Costa Alpuim D, Faria A, Calhau C, Braga Azambuja S. Human microbiota and immunotherapy in breast cancer - A review of recent developments. *Front Oncol* (2021) 11:815772. doi: 10.3389/fonc.2021.815772

5. Riquelme E, Zhang Y, Zhang L, Montiel M, Zoltan M, Dong W, et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell* (2019) 178(4):795–806 e12. doi: 10.1016/j.cell.2019.07.008

6. Qu D, Wang Y, Xia Q, Chang J, Jiang X, Zhang H. Intratumoral microbiome of human primary liver cancer. *Hepatol Commun* (2022) 6(7):1741–52. doi: 10.1002/hep4.1908

7. Huang D, Su X, Yuan M, Zhang S, He J, Deng Q, et al. The characterization of lung microbiome in lung cancer patients with different clinicopathology. *Am J Cancer Res* (2019) 9(9):2047–63.

8. Salihoğlu R, Önal-Süzek T. Tissue microbiome associated with human diseases by whole transcriptome sequencing and 16S metagenomics. *Front Genet* (2021) 12:585556. doi: 10.3389/fgene.2021.585556

9. Alpuim Costa D, Nobre JG, Batista MV, Ribeiro C, Calle C, Cortes A, et al. Human microbiota and breast cancer-is there any relevant link?-A literature review and new horizons toward personalised medicine. *Front Microbiol* (2021) 12:584332. doi: 10.3389/fmicb.2021.584332

10. Yang Y, Misra BB, Liang L, Bi D, Weng W, Wu W, et al. Integrated microbiome and metabolome analysis reveals a novel interplay between commensal bacteria and metabolites in colorectal cancer. *Theranostics* (2019) 9(14):4101–14. doi: 10.7150/thno.35186

11. Urbaniak C, Gloor GB, Brackstone M, Scott L, Tangney M, Reid G. The microbiota of breast tissue and its association with breast cancer. *Appl Environ Microbiol* (2016) 82(16):5039–48. doi: 10.1128/AEM.01235-16

12. Parida S, Sharma D. The microbiome-estrogen connection and breast cancer risk. *Cells* (2019) 8(12):1642. doi: 10.3390/cells8121642

13. Zhang L, Wang WW, Jiang GZ, Zhang YP. [Impact and clinical value of the revised 2019 Chinese HER-2 testing guidelines on the detect result evaluation of invasive breast cancer cases with equivocal HER-2 immunostaining by using fluorescence in *situ* hybridization]. *Zhonghua Zhong Liu Za Zhi* (2021) 43(8):833–7.

14. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* (2010) 28(16):2784–95. doi: 10.1200/JCO.2009.25.6529

15. Salgado R, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol* (2015) 26(2):259–71. doi: 10.1093/annonc/mdu450

16. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* (2012) 6(8):1621–4. doi: 10.1038/ismej.2012.8

17. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* (2013) 10(10):996–8. doi: 10.1038/nmeth.2604

18. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* (2006) 72(7):5069–72. doi: 10.1128/AEM.03006-05

19. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* (2011) 12(6):R60. doi: 10.1186/gb-2011-12-6-r60

20. Yuan M, Breitkopf SB, Yang X, Asara JM. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat Protoc* (2012) 7(5):872–81. doi: 10.1038/nprot.2012.024

21. Tzeng A, Sangwan N, Jia M, Liu CC, Keslar KS, Downs-Kelly E, et al. Human breast microbiome correlates with prognostic features and immunological signatures in breast cancer. *Genome Med* (2021) 13(1):60. doi: 10.1186/s13073-021-00874-2

22. Hieken TJ, Chen J, Chen B, Johnson S, Hoskin TL, Degnim AC, et al. The breast tissue microbiome, stroma, immune cells and breast cancer. *Neoplasia* (2022) 27:100786. doi: 10.1016/j.neo.2022.100786

23. Meng S, Chen B, Yang J, Wang J, Zhu D, Meng Q, et al. Study of microbiomes in aseptically collected samples of human breast tissue using needle biopsy and the potential role of in *situ* tissue microbiomes for promoting Malignancy. *Front Oncol* (2018) 8:318. doi: 10.3389/fonc.2018.00318

24. An J, Kwon H, Kim YJ. The firmicutes/bacteroidetes ratio as a risk factor of breast cancer. *J Clin Med* (2023) 12(6):2216. doi: 10.3390/jcm12062216

25. Smith A, Pierre JF, Makowski L, Tolley E, Lyn-Cook B, Lu L, et al. Distinct microbial communities that differ by race, stage, or breast-tumor subtype in breast tissues of non-Hispanic Black and non-Hispanic White women. *Sci Rep* (2019) 9(1):11940. doi: 10.1038/s41598-019-48348-1

26. Poutahidis T, Erdman SE. Commensal bacteria modulate the tumor microenvironment. *Cancer Lett* (2016) 380(1):356–8. doi: 10.1016/j.canlet.2015.12.028

27. Wang H, Rong X, Zhao G, Zhou Y, Xiao Y, Ma D, et al. The microbial metabolite trimethylamine N-oxide promotes antitumor immunity in triple-negative breast cancer. *Cell Metab* (2022) 34(4):581–594 e8. doi: 10.1016/j.cmet.2022.02.010

28. Banerjee S, Tian T, Wei Z, Shih N, Feldman MD, Peck KN, et al. Distinct microbial signatures associated with different breast cancer types. *Front Microbiol* (2018) 9:951. doi: 10.3389/fmicb.2018.00951

29. Lin C, Chen DR, Kuo SJ, Feng CY, Chen DR, Hsieh WC, et al. Profiling of protein adducts of estrogen quinones in 5-year survivors of breast cancer without recurrence.*Cancer Control* (2022) 29:10732748221084196. doi: 10.1177/10732748221084196

30. Zipinotti Dos Santos D, de Souza JC, Pimenta TM, da Silva Martins B, Junior RSR, Butzene SMS, et al. The impact of lipid metabolism on breast cancer: a review about its role in tumorigenesis and immune escape. *Cell Commun Signal* (2023) 21(1):161. doi: 10.1186/s12964-023-01178-1

31. Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol* (2018) 19(1):40–50. doi: 10.1016/S1470-2045(17)30904-X

32. Chen DR, Chen ST, Wang TW, Tsai CH, Wei HH, Chen GJ, et al. Characterization of estrogen quinone-derived protein adducts and their identification in human serum albumin derived from breast cancer patients and healthy controls. *Toxicol Lett* (2011) 202(3):244–52. doi: 10.1016/j.toxlet.2011.02.010

33. Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *J Immunother Cancer* (2016) 4:59. doi: 10.1186/s40425-016-0165-6

34. Park IH, Kong SY, Ro JY, Kwon Y, Kang JH, Mo HJ, et al. Prognostic implications of tumor-infiltrating lymphocytes in association with programmed death ligand 1 expression in early-stage breast cancer. *Clin Breast Cancer* (2016) 16(1):51–8. doi: 10.1016/j.clbc.2015.07.006

35. Iwamoto T, Kajiwara Y, Zhu Y, Iha . Biomarkers of neoadjuvant/adjuvant chemotherapy for breast cancer. *Chin Clin Oncol* (2020) 9(3):27. doi: 10.21037/cco.2020.01.06

36. Delgir S, Bastami M, Ilkhani K, Safi A, Seif F, Alivand MR. The pathways related to glutamine metabolism, glutamine inhibitors and their implication for improving the efficiency of chemotherapy in triple-negative breast cancer. *Mutat Res Rev Mutat Res* (2021) 787:108366. doi: 10.1016/j.mrrev.2021.108366

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 07 November 2023
DOI 10.3389/fonc.2023.1246880

OPEN ACCESS

EDITED BY
Maha Mohamed Saber-Ayad,
University of Sharjah, United Arab Emirates

REVIEWED BY
Sangseon Lee,
Seoul National University, Republic of Korea
Veronika Holubekova,
Comenius University, Slovakia

*CORRESPONDENCE
Xiaowu Wang
wangxiaowu_email@163.com
Xuanxuan Dai
daoshidaixuanxuan@126.com

†These authors have contributed equally to this work

RECEIVED 24 June 2023
ACCEPTED 20 October 2023
PUBLISHED 07 November 2023

CITATION
Wu H, Wu Z, Ye D, Li H, Dai Y, Wang Z, Bao J, Xu Y, He X, Wang X and Dai X (2023) Prognostic value analysis of cholesterol and cholesterol homeostasis related genes in breast cancer by Mendelian randomization and multi-omics machine learning.
Front. Oncol. 13:1246880.
doi: 10.3389/fonc.2023.1246880

Prognostic value analysis of cholesterol and cholesterol homeostasis related genes in breast cancer by Mendelian randomization and multi-omics machine learning

Haodong Wu[1,2,3†], Zhixuan Wu[1,2†], Daijiao Ye[4†], Hongfeng Li[1], Yinwei Dai[1], Ziqiong Wang[1], Jingxia Bao[1], Yiying Xu[1], Xiaofei He[4], Xiaowu Wang[2*] and Xuanxuan Dai[1*]

[1]Department of Breast Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [2]Department of Burns and Skin Repair Surgery, The Third Affiliated Hospital of Wenzhou Medical University, Ruian, Zhejiang, China, [3]Key Laboratory of Clinical Laboratory Diagnostics (Ministry of Education), The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [4]Medical Research Center, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

Introduction: The high incidence of breast cancer (BC) prompted us to explore more factors that might affect its occurrence, development, treatment, and also recurrence. Dysregulation of cholesterol metabolism has been widely observed in BC; however, the detailed role of how cholesterol metabolism affects chemosensitivity, and immune response, as well as the clinical outcome of BC is unknown.

Methods: With Mendelian randomization (MR) analysis, the potential causal relationship between genetic variants of cholesterol and BC risk was assessed first. Then we analyzed 73 cholesterol homeostasis-related genes (CHGs) in BC samples and their expression patterns in the TCGA cohort with consensus clustering analysis, aiming to figure out the relationship between cholesterol homeostasis and BC prognosis. Based on the CHG analysis, we established a CAG_score used for predicting therapeutic response and overall survival (OS) of BC patients. Furthermore, a machine learning method was adopted to accurately predict the prognosis of BC patients by comparing multi-omics differences of different risk groups.

Results: We observed that the alterations in plasma cholesterol appear to be correlative with the venture of BC (MR Egger, OR: 0.54, 95% CI: 0.35-0.84, p<0.006). The expression patterns of CHGs were classified into two distinct groups(C1 and C2). Notably, the C1 group exhibited a favorable prognosis characterized by a suppressed immune response and enhanced cholesterol metabolism in comparison to the C2 group. In addition, high CHG score were accompanied by high performance of tumor angiogenesis genes. Interestingly, the expression of vascular genes (CDH5, CLDN5, TIE1, JAM2, TEK) is lower in patients with high expression of CHGs, which means that these patients have

poorer vascular stability. The CAG_score exhibits robust predictive capability for the immune microenvironment characteristics and prognosis of patients (AUC=0.79). It can also optimize the administration of various first-line drugs, including AKT inhibitors VIII Imatinib, Crizotinib, Saracatinib, Erlotinib, Dasatinib, Rapamycin, Roscovitine and Shikonin in BC patients. Finally, we employed machine learning techniques to construct a multi-omics prediction model (Risklight),with an area under the feature curve (AUC) of up to 0.89.

Conclusion: With the help of CAG_score and Risklight, we reveal the signature of cholesterol homeostasis-related genes for angiogenesis, immune responses, and the therapeutic response in breast cancer, which contributes to precision medicine and improved prognosis of BC.

KEYWORDS

Mendelian randomization, breast cancer, immune microenvironment, cholesterol homeostasis, prognosis prediction, machine learning method

1 Introduction

According to the World Health Organization (WHO) report in 2021, breast cancer (BC) has become the most prevalent tumor in the world with the increasing incidence (1). Attribute to the progress of surgical treatment and the application of immunotherapy, its survival rate is also higher than other tumors, but there is still a high recurrence rate, and the recurrence rate of patients who receive postoperative radiotherapy can reach 15% within 10 years (2). Therefore, it is particularly important to explore techniques and biomarkers for early identification and prevention of recurrence.

In addition to the effects at the genetic level, some studies have pointed out that the disruption of cellular cholesterol levels' dynamic balance can lead to cancer occurrence and a series of diseases (3). Elevated serum cholesterol is associated with the risk of melanoma, prostate cancer, endometrial cancer, non-Hodgkin's lymphoma, and breast cancer (3–5). Hypercholesterolemia has been identified as a comorbidity of obesity, becoming an independent risk factor for breast cancer in postmenopausal women. Dysregulation of cholesterol homeostasis can also lead to ferroptosis resistance, thereby increasing tumor tumorigenicity and metastatic capacity (6). However, most current studies have focused on determining the role of serum cholesterol or liver cholesterol in the progression and prognosis of BC (7, 8), while neglecting the involvement of cholesterol homeostasis-related genes (CHGs) in tumorigenesis.

Furthermore, the tumor microenvironment (TME) has garnered increasing attention (9). Tumor growth environment is a complex tissue environment, which is closely related to tumor growth, invasion, metastasis, and other functions. Under the induction of tumor cells, stromal cells in TME lead to increased angiogenesis and immune escape of tumor cells. The mechanism of immune cells such as T cells and tumor-associated macrophages (TAMs) involved in this process has attracted many scholars to explore, which means that TME can become a potential therapeutic target (10). At the same time, it has also been found that intracellular cholesterol metabolism has an important impact on the tumor-inhibitory effect of CD8+ T cells (11). However, the precise mechanisms underlying the interaction between TME and cholesterol metabolism as well as tumor immune evasion remain elusive.

Hence, we conducted a comprehensive analysis of the expression of CHGs and its impact on the tumor microenvironment (TME), disease progression, treatment response, and prognosis in breast cancer (BC) patients. Leveraging CAG_score and multi-omics machine learning techniques, we developed a robust model that accurately predicts both prognostic risk and immunotherapy efficacy for BC study will contribute to enhancing the rationalization of immunotherapeutic approaches in breast cancer.

2 Materials and methods

2.1 Mendelian randomization analysis

To assess the potential connection between cholesterol and the risk of breast cancer, genetic data on cholesterol (met-a-307, sample Size 7,813, number of SNPs 2545,608)and breast cancer(ieu-a-1132, ER+ Breast cancer (Oncoarray), sample size 833691, number of SNPs 10680275) were searched and obtained from the IEU Open GWAS project(https://gwas.mrcieu.ac.uk/). The data then were briefly collated and subjected to a two-sample Mendelian randomization (2-SMR) analysis. Mendelian randomization-Egger (MR-Egger) method analyses were the main way performed along with the inverse variance-weighted (IVW) method analysis, Weighted-median method analysis, Weighted mode method analysis and Simple mode method analysis (12).

2.2 Download of the BC dataset and acquisition of cholesterol homeostasis-associated genes

The basic information on breast cancer RNA sequencing transcriptome data, CNV files, somatic mutation data, and clinicopathologic data were acquired from the publicly available TCGA database (http://xena.ucsc.edu/). Microarray dataset GSE58812 was downloaded from the GEO database (https://www.gov/geo/). A total of 1324 breast cancer samples were analyzed in this study. 1097 patients with a survival time greater than 30 days and 120 normal tissue samples were selected from the TCGA-BRCA cohort. The GSE58812 cohort contains 107 samples of breast cancer patients. The 73 Cholesterol homeostasis genes (CHGs) and 36 Angiogenesis genes (AAGs) were retrieved from the MSigDB team (Hallmark Gene set) as indicated in Table S1.

2.3 Consensus clustering analysis of CHGs

9 CHGs were obtained with univariate Cox regression (UniCox) analysis. Consensus clustering was used to identify different cholesterol homeostasis-related patterns by the k-means algorithms with 1000 repetitions (13). The distinction in clinical characteristics between the C1 and C2 groups was assessed using a Chi-square test. Differences in the biological function of these patterns were investigated using Genetic Set Variable Analysis (GSVA) (14). OS time and OS state of various modes were compared using the Kaplan-Meier method (15). Additionally, we explored the association between molecular patterns of cholesterol homeostasis genes, clinical features, and survival differences.

2.4 Landscape of tumor immune environment in different subgroups of breast cancer

The "Estimation" R package was used to present the proportion of immune cells and stromal cells in BC by analyzing gene expression, which can further calculate the tumor purity (16). Abundance of 23 specific immune cell subtypes was measured in tumors with the CIBERSORT algorithm to reveal the infiltration of immune cells (17). We predicted the sensitivity of immunotherapy by comparing the expression levels of several immune checkpoints among different subgroups. Moreover, the degree of immune cell infiltration in tumor and normal samples was determined by single sample Gene Set Enrichment Analysis (ssGSEA analysis) (18).

2.5 Identification of DEGs and cholesterol homeostasis-related genes

Using the "limma" package, we acquired DEGs for breast cancer in the TCGA dataset. DEGs should comply with the | log2 fold change (FC) | ≥ 0.5, $p < 0.05$. Pearson correlation analysis was used to obtain genes that were related to Cholesterol homeostasis, with |cor|≥ 0.6.

2.6 Prognostic score of cholesterol homeostasis

A CAG_Score was established to quantitatively evaluate the state of cholesterol homeostasis for individual BC patients. Firstly, we performed uniCox analysis and multi-factor Cox analysis (mulCox) for CHG-related genes to search for which has significant prognostic value. Then, we integrated OS time, OS, and gene expression data with the "glmnet" package and developed the CAG_Score by the Lasso Regression Algorithm (19).

$$\text{CAG score} = \sum_{n} \text{Coefficient of gene(n)} \times \text{Expression of gene (n)}$$

The median CAG_score was adapted to classify breast cancer patients into low-risk and high-risk groups.

2.7 Construction of cholesterol homeostasis relevant nomograph

A CAGs-related nomograph was established to describe the clinical features and risk score of BC patients, as well as the clinical prediction of 3-year,4-year, and 5-year survival status. Calibration curves were generated to identify the accuracy of the predictive effect.

2.8 Drug sensitivity analysis and quantitative RT-PCR

The IC50 of commonly used clinical drugs was numerically analyzed by the "pRRophic" package in order to compare the chemotherapy effects of different risk groups (20). Total RNA of breast cancer cells (MDA-MB-231, MCF-7, SKBR-3) and normal breast cells (MCF-10A) were prepared by TRIzol reagent (Thermo Fisher Scientific, Waltham, USA). cDNA was synthesized with TOROIVD qRT Master Mix kit (TOROIVD, shanghai, China) according to the manufacturer's instructions. The qRT-PCR was performed using the TOROGreen qPCR Master Mix kit (TOROIVD, shanghai, China) on the ABI 7500 real-time fluorescence quantitative PCR system (Thermo Fisher Scientific). All sequences of primers used are shown in Table S2.

2.9 Development of a multi-omics machine learning model to predict the prognosis and microenvironment of breast cancer

The TCGA cohort was divided into a training cohort (n=824) and a test cohort (n=206) randomly. We defined BC prognostic risk markers as characteristic mRNA, lncRNA, and miRNA in the TCGA cohort. Screening for characteristic mRNAs, miRNAs, and lncRNAs based on high-risk score and low-risk score, for each type of data, the top 100 most relevant features were retained as BC-specific risk markers according to the P-value. Then, we performed

lasso regression for further feature filtering and reduced the number of markers to 20 for each type of data. With 20 makers per molecular layer, we created a risk predictor of each single molecule layer with three machine learning models, such as Light GBM, Logistic regression, and Random forest (21). Finally, based on 60 BC-specific markers from three data types, we developed a LightGBM model (RiskLight) to distinguish breast cancer patients with different prognostic risks associated with dysregulated cholesterol homeostasis.

2.10 Statistical analysis

In the statistical analysis, $p<0.05$ was considered statistically significant. The t-test is used for the analysis of normally distributed data, while the Wilcoxon rank sum test is used for the analysis of abnormally distributed data. In addition, Pearson correlation analysis or Spearman analysis was used to describe the relationship between two numerical variables. The above algorithms are all implemented in R Software (version 4.1.2).

3 Results

3.1 Clinical and mutations data of CHGs in BC

The flow chart of the research design is shown in Figure 1. To evaluate the role of cholesterol in the occurrence of breast cancer, the Mendelian randomisation-Egger (MR-Egger) method was used first in the main MR analysis, as the detailed results are presented in Figure 2A (MR Egger, OR: 0.54, 95% CI: 0.35-0.84, $p<0.006$). This means that cholesterol levels may be a risk factor for breast cancer. We obtained 73 genes for cholesterol homeostasis from the MSigDB database and verified the expression levels of 73 CHGs in tumor specimens and normal control in the TCGA-BC cohort (Figure 2B). 63 CHGs had differential expression (Figure 2C). Correlations between 73 CHGs were analyzed with the String website (Table S3). Protein interaction network (PPI) was constructed by Cytoscape software to explore the interactions between CHGs (22). And we identified SCD, PPARG, CTNNB1, FDPS, LDLR, ACSS2, FDFT1, FADS2, HMGCR, SREBF2, ACTG1 and HMGCS1 as the vital genes of cholesterol homeostasis (Figure S1). We calculated the CNV mutation rate of CHGs, Figure 2D shows the results. In addition, we determined the incidence of SNV of 73 CHGs in BC, and 142 out of 981 BC samples (14.46%) showed mutations, which indicated that the mutation rate of 73 CHGs was less than 1% (Figure S2).

3.2 Generation of cholesterol homeostasis subgroups in BC

Generation of a subset of genes related to cholesterol homeostasis regulation in BC to reveal the relationship between cholesterol homeostasis regulation and tumorigenesis. 1097 BC patients of TCGA-BC were included in this study, and uniCox analysis revealed 9 CHGs with prognostic significance (Figure 3A). To determine the relationship between CHGs expression patterns and BC subtypes, consensus cluster analysis was used to classify BC patients according to prognostic genes. When the clustering variable was 2, BC patients were well divided

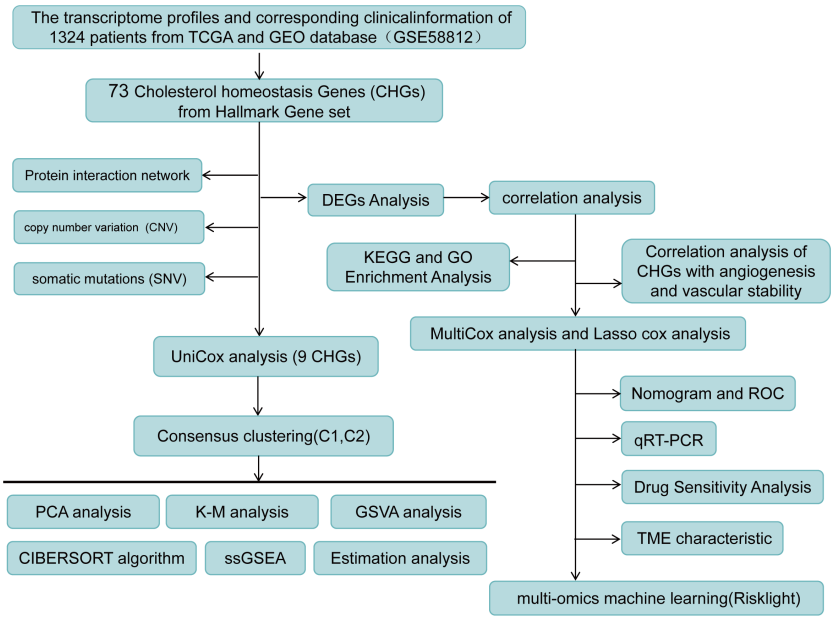


FIGURE 1
Flow chart of research design.

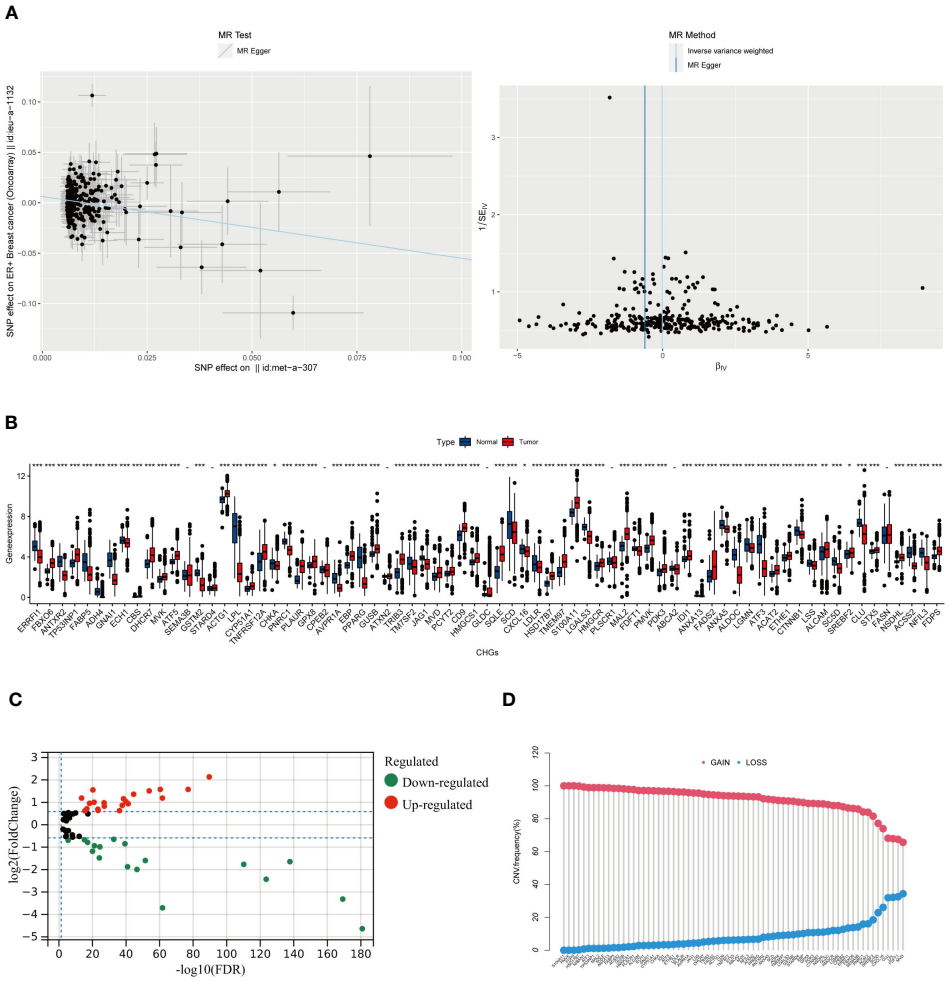


FIGURE 2
The result of Mendel randomized model and the Molecular Characteristics of CHGs in BC. **(A)** Association between cholesterol and ER+ breast cancer risk overall.MR-analyses are derived using random effect Ivw, MR-Egger, weighted median and mode.(MR Egger, OR: 0.54, 95% CI: 0.35-0.84, p<0.006). **(B)** Distribution of CHGs between BC and normal tissues. (p>0.05 -; p< 0.05 *; p< 0.01 **; p< 0.001 ***). **(C)** Volcano map of 63 DEGs (log2 fold change. (FC)|≥0.5, p-value<0.05). **(D)** Incidence rate of CNV gain, loss, and non-CNV among CHGs.

into the C1 group (n=510) and the C2 group (n=587) (Figure S3). PCA analysis showed significant subpopulation differentiation in samples (Figure 3B). KM analysis revealed that cluster C2 showed a worse prognostic status (Figure 3C).The clinical features distinguishing the C1 and C2 groups are presented in Table S4. In addition, the relationship between gene expression and clinical features of the two clusters was shown in Figure 3D. The heatmap indicated that the expression level of CHGs had a significant correlation with the clinical characteristics, and the genetic characteristics of the C2 subcluster were associated with distant tumor metastasis. The biological functions and signaling pathways of tumor cells were compared by the GSVA algorithm, and the findings showed that the C2 subcluster performed obvious immune pathway characteristics, lipid metabolism, and sterol metabolism-related pathways were down-regulated, and cancer metastasis-related pathways were significantly different as well (Figure 3E). This suggests that dysregulation of cholesterol metabolism is closely associated with tumor immunity and the development of tumors.

3.3 Characteristics of the TME in different subgroups

Investigating the infiltration extent of 23 human immune cells in both clusters by the CIBERSORT algorithm (Figure 4A), we found that the content of Macrophage M0, Macrophage M1, activated Dendritic cell and T cell include activated CD4 positive memory T cell, helper T cell, gamma and delta T cell were significantly higher in group C2, whereas Plasma cell, macrophage M2, resting dendritic cell mast cells behaved in an opposite way. Inter-individual differences in 23 immune cells were assessed by the "ssGSEA" algorithm and the number was generally higher in the C2 group (Figure 4B). The TME scores exhibited that patients in cluster C2 had a higher abundance of immune and matrix components (Figure 4C). In addition, PD-1, PD-L1, and CTLA-4 were shown a similar increase in cohort C2, which represents the critical expression status of the immune checkpoints (ICP) (Figure 4D). Meanwhile, the correlation analysis between CHGs and immune cells displayed that FBXO6,

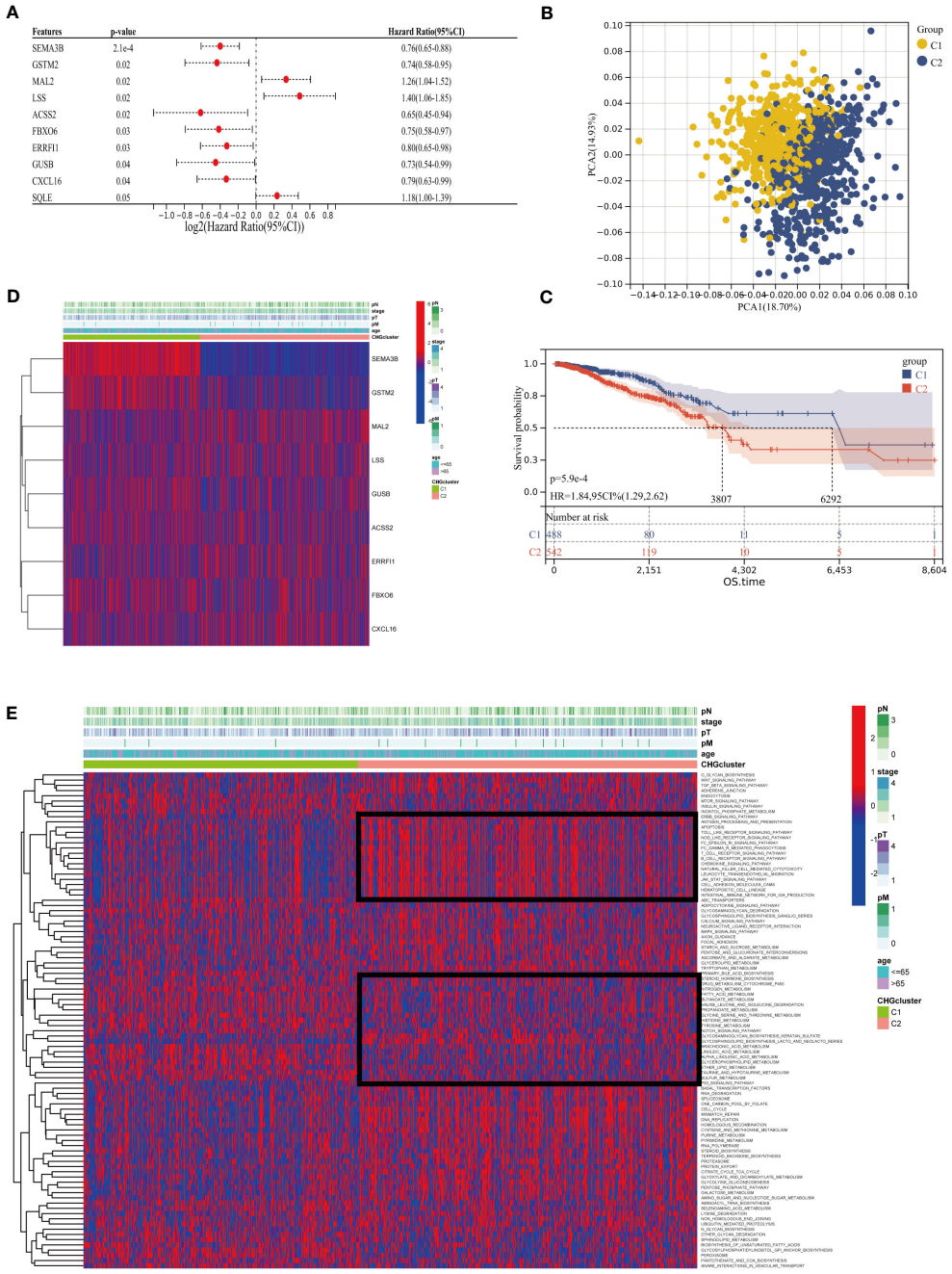


FIGURE 3
Cluster analysis of cholesterol homeostasis subgroups. **(A)** Univariate Cox regression (uniCox) analysis for CHGs (p<0.05 is considered significant). **(B)** PCA analysis showed the distribution of the two clusters. **(C)** Survival curve between different clusters. **(D)** Expression of prognostic genes and the presentation of clinical features in different clusters. **(E)** The heatmap of biological function and signaling pathway in two groups.

SEMA3B, GSTM2, and CXCL16 were correlated with immune cell abundance (Figure 4E).

3.4 Potential biological activity of cholesterol homeostasis gene, correlation analysis between CHGs and angiogenesis

The Pearson correlation algorithm was applied to analyze CHGs, resulting in 510 highly correlated DEGs (Figure 5A). Functional enrichment analysis of these DEGs was then performed to demonstrate the potential biological activity of cholesterol homeostasis genes. KEGG and GO analysis revealed an enrichment of cancer and metastasis-related pathways as well as blood vessel development and sterol metabolism, which suggested that cholesterol homeostasis is closely related to angiogenesis (Figures 5B, C). To reveal the association between cholesterol homeostasis and angiogenesis, we obtained 36 angiogenic genes (AAGs) from MsiGDB and explored the correlation between CHGs and AAGs. The results were as expected, especially when

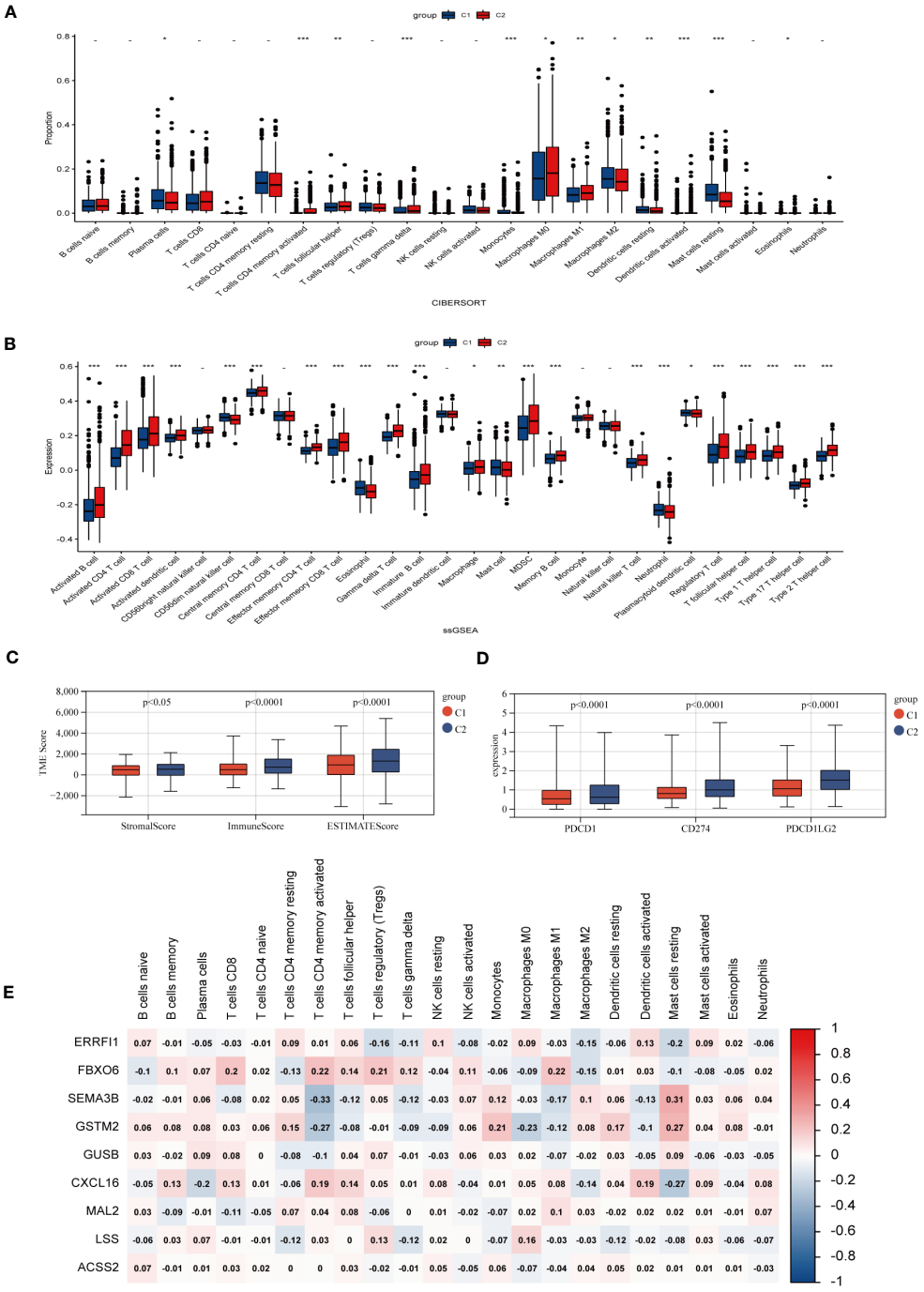


FIGURE 4
Characteristic of TME in two BC subgroups. **(A)** Abundances of 23 infiltrating immune cells in two BC subpopulations. **(B)** Enrichment score of 23 immune cells for each BC sample by ssGSEA analysis. (p>0.05 -; p< 0.05 *; p< 0.01 **; p< 0.001 ***). **(C)** Expression levels of immune checkpoints (PD-1, PD-L1, and CTLA-4) of different subgroups. **(D)** Immune infiltration scores for different groups. **(E)** Correlation of clustering genes with 23 immune cells.

ANTXR2, GPX8, and AVPR1A were strongly associated with angiogenesis (Figure 5D).

Subsequently, we examined the expression of AAGs in groups C1 and C2 (Figure 5E), as well as in tumor and normal tissues (Figure S4); however, a significant discrepancy exists. The GSVA algorithm was used to evaluate the cholesterol homeostasis score (CHG score) and angiogenesis score (AAG score) of TCGA BC samples based on 73 CHGs and 36 AAGs. And cholesterol homeostasis scores were positively correlated with angiogenesis scores in the TCGA-BC cohort (Figure 6A). Moreover, the cholesterol homeostasis score and angiogenesis score were compared between the C1 and C2 groups. We found that patients in the C2 group had a worse prognosis with higher cholesterol homeostasis score and angiogenesis score (Figure S5). The

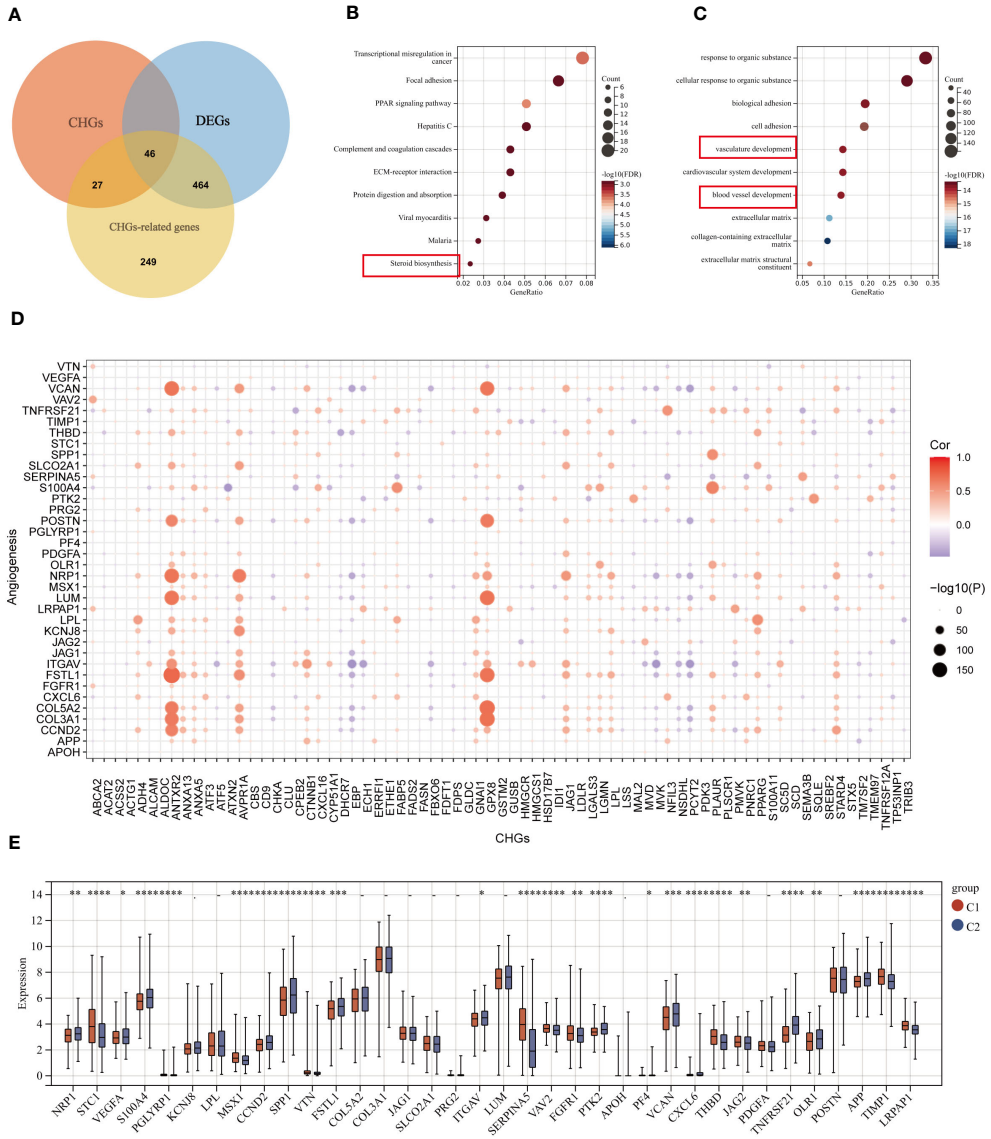


FIGURE 5
Correlation analysis of CHGs and angiogenesis. **(A)** Acquisition of cholesterol homeostasis-related DEGs (log2 fold change (FC)|≥0.5, p-value<0.05). **(B, C)** GO and KEGG enrichment analyses of cholesterol homeostasis-related DEGs among two subgroups. **(D)** Correlation analysis of CHGs and AAGs. **(E)** Expression levels of 36 AAGs between C1 and C2. *p < 0.05; **p < 0.01; ***p < 0.001; ****p< 0.0001.

correlation between vascular stability and cholesterol homeostasis score was also validated in addition. The abundance of genes related to vascular stability (CDH5, CLDN5, TIE1, JAM2, TEK) indicated that the group with a lower cholesterol score had higher vascular stability (Figures 6B-F), while low vascular stability often promotes cancer growth (23–27). All the findings were verified in the GSE58812 cohort (Figures 6G-L).

3.5 Development and validation of the prognostic CAG_score

Considering that cholesterol homeostasis is closely connected with angiogenesis, we developed a prognostic CAG_score based on genes related to cholesterol homeostasis. The BC patients were randomly assigned to the training cohort (n=731) or the test cohort (n=366). We performed UniCox analysis of 786 cholesterol-related genes, and 49 DEGs with prognostic significance (logFC>0.5, P<0.05). Subsequently, LASSO and multi-Cox analyses were performed on 49 DEGs to establish the most suitable prediction model. We set the Lambda value to 0.00298971135072249 and finally obtained 7 genes (Figures 7A, B).

$$\text{CAG\_score} = -0.21106035029347 * \text{ZMYND10} - 0.262724856118755 * \text{GBP1} - 0.522741360511683 * \text{DSCC1} + 0.465453655395411 * \text{MRPL13} + 0.16530191756177 * \text{YWHAZ} + 0.617851278765801 * \text{TCP1} + 0.147920101131816 * \text{TAGLN2}$$

In the scoring model established by CAGs, we found that higher scores were associated with a worse survival rate and higher

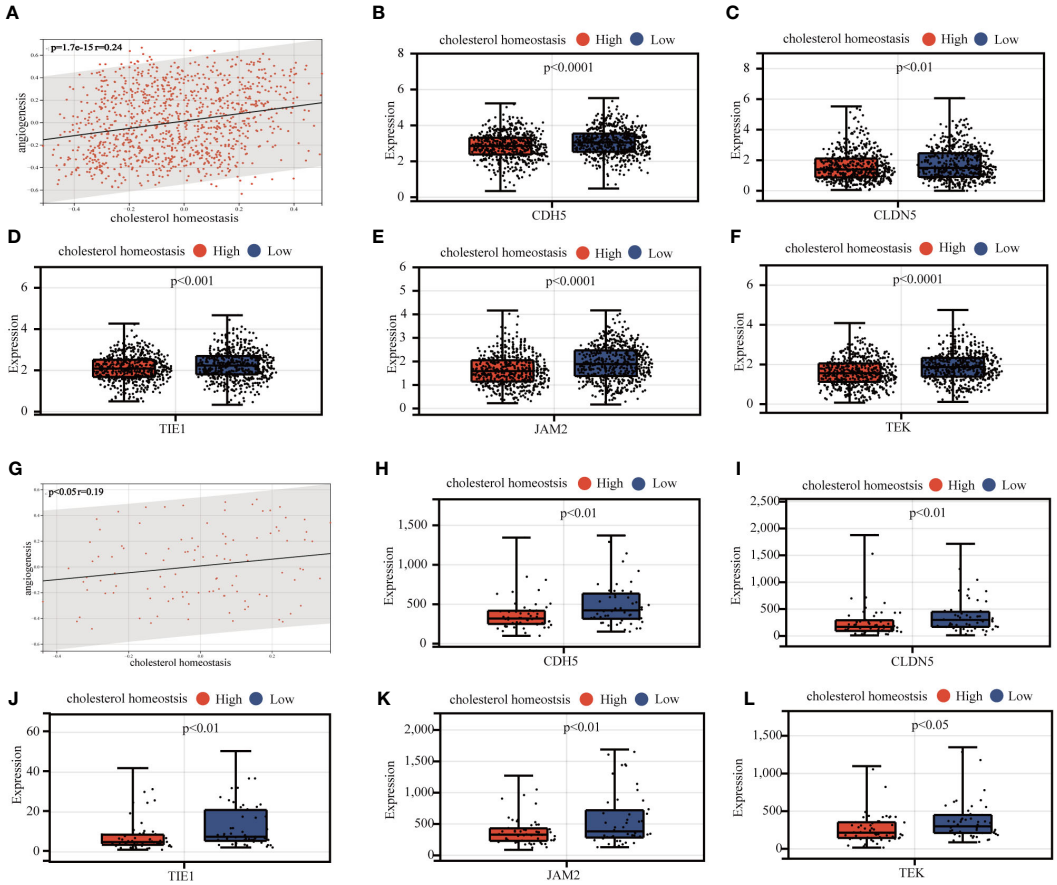


FIGURE 6
Analysis of the correlation between cholesterol homeostasis and vascular stability. **(A)** Association analysis of cholesterol homeostasis score and angiogenesis score. **(B-F)** Association between expression levels of vasostability genes and cholesterol homeostasis scores. **(G-L)** Validation of the above results in the GEO 58812 queue.

mortality rate (Figures 7C, D). The model genes also showed a trend with increasing CAG_score (Figure 7E). To evaluate the robustness of the CAG_score, we compared the CAG_score from the test to the whole cohort, and the results showed an excellent performance of the CAG_ score in assessing the prognosis of BC patients (Figures 7F, S6). Figure S7 shows the distribution of CHGs and AAGs in the two CAG_score clusters. We found significant differences in gene expression in both groups.

3.6 Construction of a nomogram to predict patient prognosis

Through the analysis of clinical indicators, we established a nomogram to predict 3, 4, and 5-year OS in BC patients (Figure 7G). The calibration curve shows that the method has a high forecasting accuracy (Figure 7H). Meanwhile, the R package "Rms" was conducted to integrate data on survival time, survival status, and 6 characteristics, and a nomogram was built using the Cox method to assess the prognostic value of these characteristics in a sample of 1030. The overall C-index of the model was: 0.783437290915762, 95%CI(0.740754644512089-0.826119937319435), p value=1.00165584281093e-38.

3.7 Assessment of TME characteristic in different groups

As mentioned above, CAG_score was positively correlated with the abundance of Macrophage M0, Macrophage M2, Plasma cell, and activated Dendritic cell, while CD8+ T cell, T cell gamma delta, activated or dormant CD4+ memory T cell, B memory cell, regulatory T cell, activated NK cell, macrophage M1 were negatively correlated with CAG_score (Figure 8A). In addition, there was a direct correlation between the CAG_ score and the TME score (Figure 8B). We explored the relevance between prognostic marker genes and 23 immune cells. We concluded that T cells and Macrophages are closely associated with the selected genes (Figure 8C). Furthermore, We evaluated the expression of ICPs in groups of different prognostic features. Figure 8D shows that the expression of 24 ICPs was inconsistent in both risk subgroups. The low-risk group showed a higher level of ICPs expression.

3.8 Drug sensitivity analysis

It is a meaningful research direction to select and guide the appropriate immunotherapy regimen for the patient (28). To

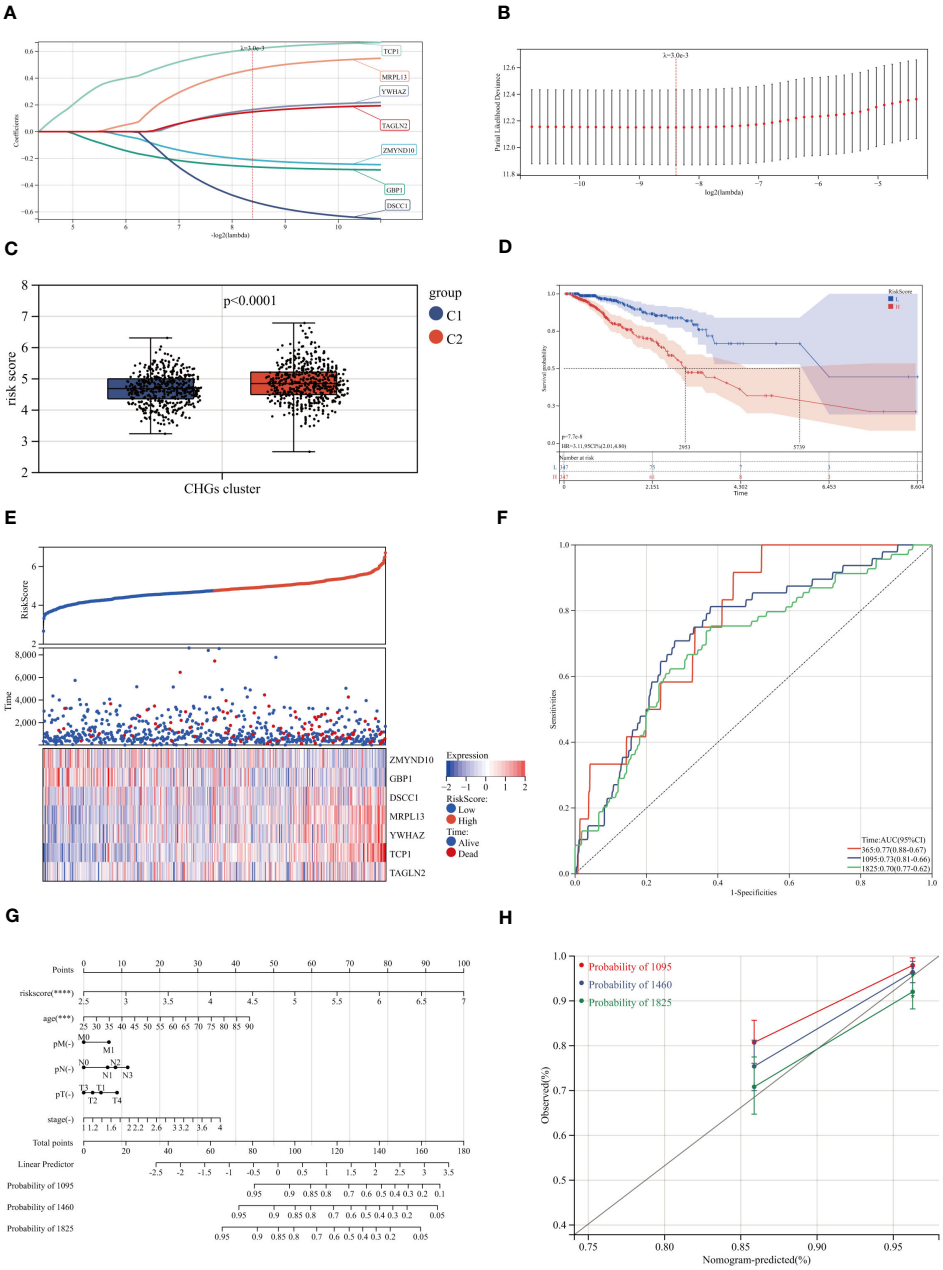


FIGURE 7
Construction of the CAG_score. **(A, B)** The LASSO analysis and determine the optimal LASSO settings. **(C, D)** Survival curves for groups and comparison of risk scores for different clusters. **(E)** Distribution of the risk score and BC patients. Dot plot of survival status. Heat maps of 7 cholesterol homeostasis-related gene expression of high- and low- risk groups. **(F)** ROC curve of the training group, the AUC values of 1, 3 and 5 years were 0.77, 0.73 and 0.70, respectively. **(G)** A nomogram for predicting the 3-, 4-, and 5-year OS for BC patients in TCGA cohort. **(H)** Calibration curves of the nomogram.

examine the role of CAG_scores in clinical diagnosis, we evaluated the IC50 for 138 Common drugs in TCGA-BC patients. The results showed that BC patients with higher CAG_ scores were more sensitive to the AKT inhibitors VIII and Imatinib, while patients with low CAG_ scores responded better to Crizotinib, Saracatinib, Erlotinib, Dasatinib, Rapamycin, Roscovitine and Shikonin (Figure 9A).

3.9 The results of qRT-PCR in several breast cancer cell lines

We detected the RNA expression of the CAG_score genes in breast cancer cell lines. Our results indicate that all genes were highly expressed in MDA-MB-231, MCF-7 and SKBR-3 cell lines (Figures 10A-H), which was consistent with our prediction.

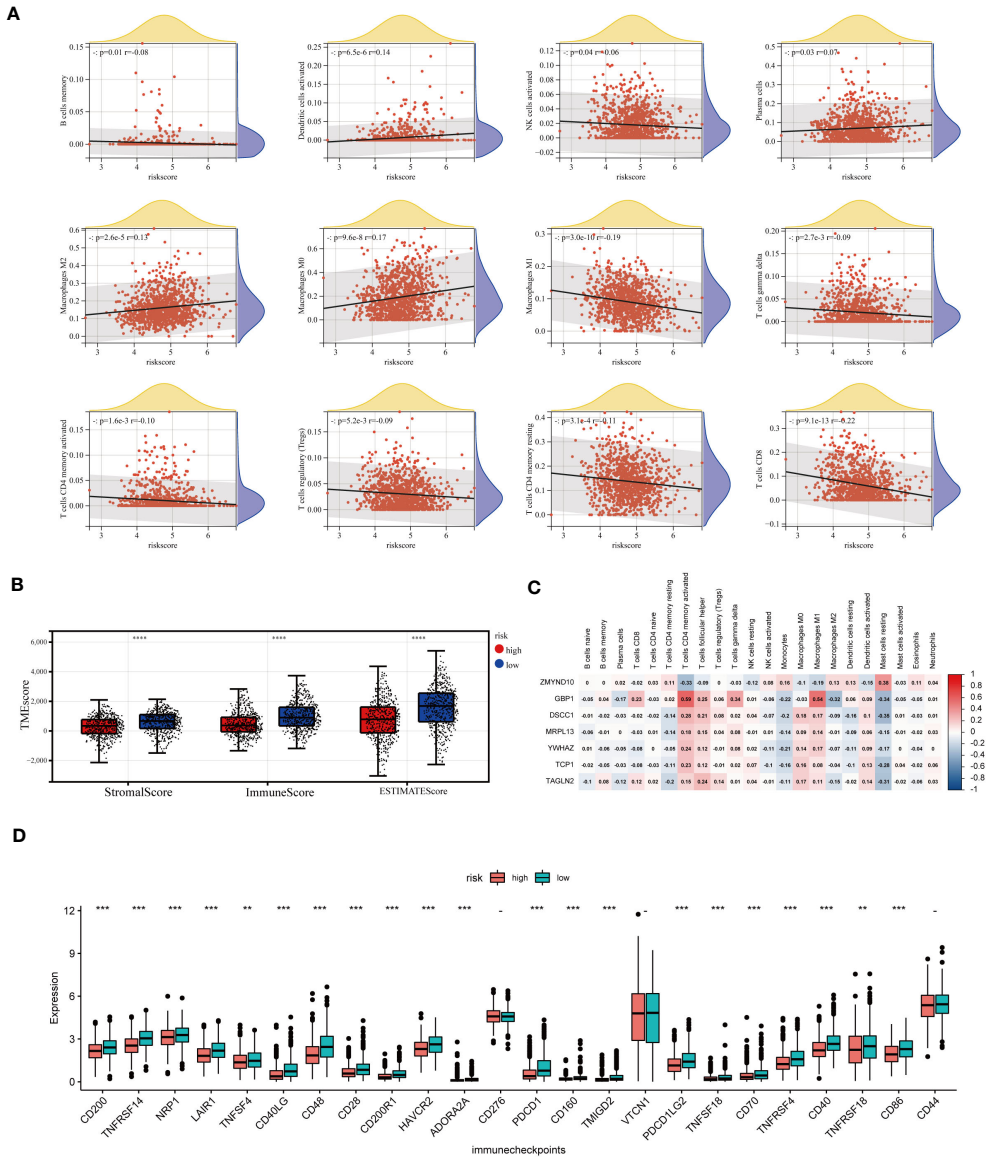


FIGURE 8
TME analysis of different risk score groups. **(A)** Correlations between CAG_score and immune cell types. **(B)** Immune infiltration scores for different groups. **(C)** Association of prognostic model genes with 23 immune cells. **(D)** Expression levels of 24 immune checkpoints in different subgroups. **p < 0.01; ***p < 0.001; ****p< 0.0001.

3.10 Multidimensional data features for different risk groups and multi-omics machine learning to build prognostic models

We have demonstrated the significance of metabolic regulation associated with cholesterol homeostasis as an immune micro-environmental factor in our study. To further identify molecular signatures associated with prognostic risk at the multi-omics level, we conducted an analysis of associations between three molecular layers (mRNA, miRNA, lncRNA) and high-low risk for each type of data, the top 100 most relevant features were retained as BC-specific risk markers according to the P-value (Figure S8A). We used the Light GBM framework to integrate multi-omics features to develop high- and low-risk prediction models as a way to emulate the tumor micro-environment in which cholesterol homeostasis is dysregulated. As a result, the three risk predictors based on the single molecular layer performed well in predicting high and low risk in the test cohort (AUC=0.8491 for the mRNA model, AUC=0.7939 for the lncRNA model, and AUC=0.7970 for the miRNA model) (Figures 11A, C, E). We also compared the superiority of random forest and logistic regression models with the Lightgbm model (Figures 11B, D, F). The results show that all three algorithms exhibit consistent results. Finally, we integrated 3 risk predictors, based on the LightGBM algorithm combined with multi-omics data to develop an integrated model (Risklight) for predicting cholesterol homeostasis-related risk patterns. Risklight is superior to all risk predictors based on single molecular layers (AUC=0.89) (Figures 11G, H).

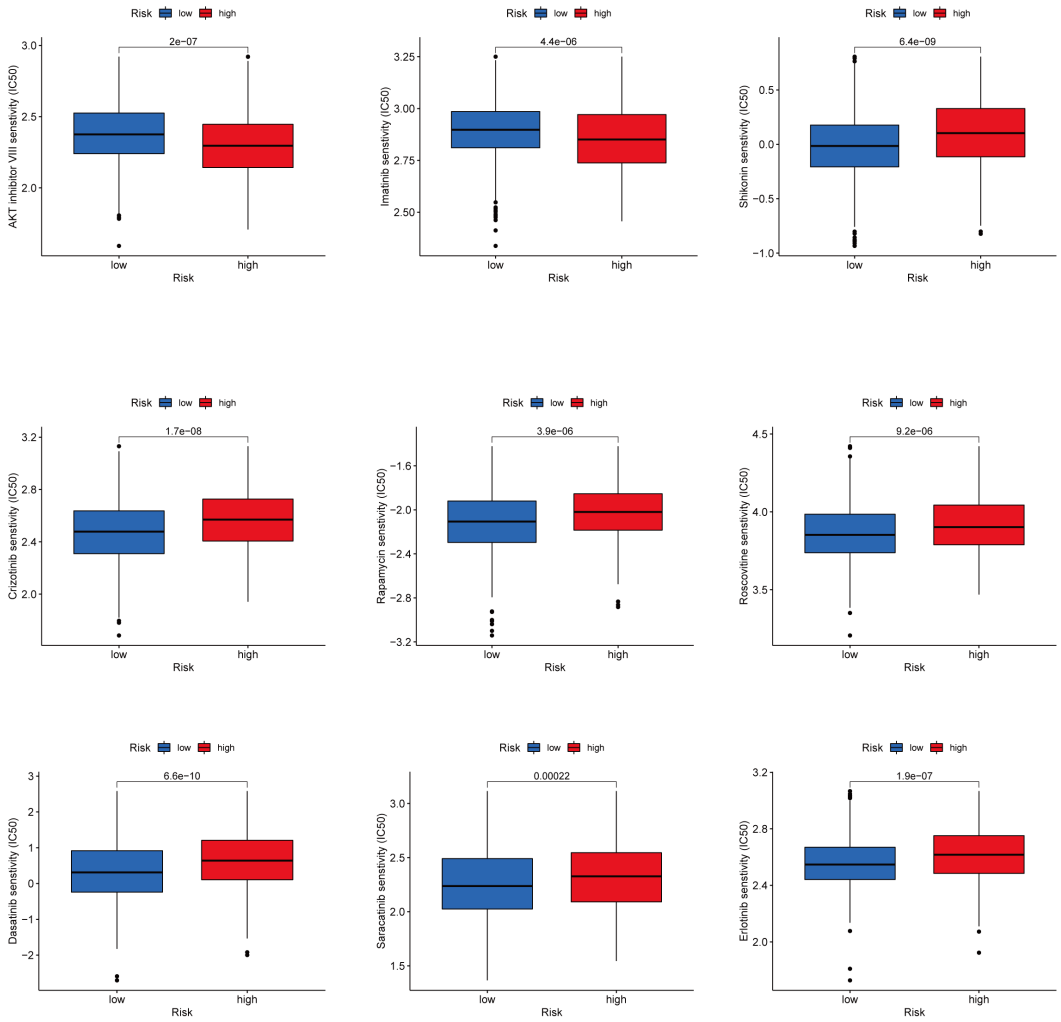


FIGURE 9
Drug Sensitivity Analysis Prediction of clinically common drug susceptibility in patients with different CAG scores.

4 Discussion

Locoregional and systemic therapies of BC have progressed substantially over the past years, at the same time, precision treatment has become a major focus on the treatment of BC. Since the importance of developing effective therapies has been noticed, it is still necessary to define the risk factors of BC and exploit this information to formulate chemopreventative strategies and improve lifestyles that can help to reduce the burden of BC. Although the results of our MR analysis suggest that cholesterol levels are a risk factor for BC, the exact mechanism of its occurrence remains unknown. It is still necessary to investigate the characteristics of cholesterol homeostasis genes and their potential biological activity in BC.

Our study quantified the cholesterol score of each BC patient's sample by utilizing a set of cholesterol homeostasis genes and evaluated different patterns of cholesterol homeostasis in BC. It showed significant differences in immune infiltration, functional enrichment, and clinical outcomes in different cholesterol gene expression pattern groups. ACT and ICI therapies, as we all know, are the success of cancer immunotherapy (29). There is no doubt about that that immune cells, particularly T cells, can be harnessed to eliminate tumor cells (30). The presence of TIL, especially CTL, is positively correlated with the survival rate of various cancer patients (31). Unexpectedly, despite having higher levels of CD8+T cells, including CTL, the C2 cluster exhibited a poorer prognosis and stronger features of distant tumor metastasis with downregulation of multiple metabolisms including sterol metabolism and fatty acid metabolism. Previous researches show that cholesterol metabolism plays a critical role in activation, proliferation, and effector function of CD8+ T cell (32). We imply that the downregulation of sterol and lipid metabolism reduces the effector function of CTL, making the C2 subgroup have a poorer prognosis with high levels of immunity levels (33). Not surprisingly, the C1 subpopulation has a well-prognostic with low levels of immune under high levels of sterol and lipid metabolism. Moreover, cholesterol homeostasis genes FBXO6, SEMA3B,

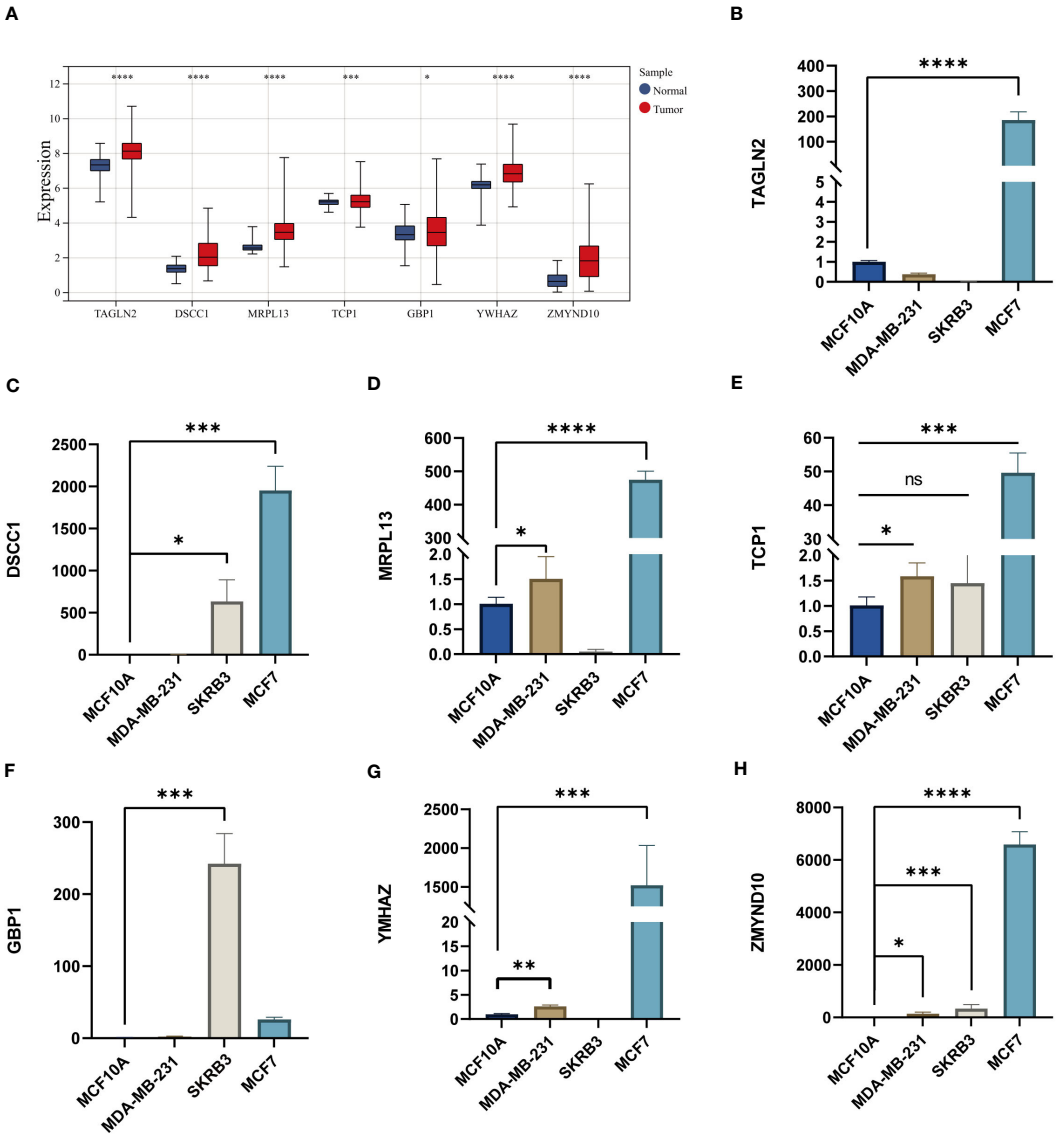


FIGURE 10
(A) The expression levels of CAG_score genes in TCGA BC cohort. **(B-H)** The mRNA levels of CAG_score genes in breast cancer cell lines. (*P <0.05, **P <0.01, *** P < 0.001, **** P < 0.0001; ns, nonsignificant).

GSTM2 and CXCL16 were associated with immune cell abundance. In previous studies, FBXO6 and CXCL16 have been shown to be directly related to immunity. Specifically, FBXO6 has been shown to impair the survival of alveolar macrophages by enhancing the degradation of NLRX1 (34), while CXCL16 serves as a critical ligand for CXCR6 that promotes the survival and local expansion of effector CTL in the TME (35). However, the exact role of SEMA3B and GSTM2 in the immune environment is unclear and requires further elucidated. These findings suggest that targeted regulation of cholesterol homeostasis may be a novel approach new immunotherapy in BC.

In addition, cholesterol homeostasis genes (CHGs) exhibit a strong correlation with the development of vasculature. Tumor growth necessitates neovascularization to adequately supply rapidly proliferating tumor cells with oxygen and nutrients (36).. Our study highlights the robust association between angiogenic genes and three specific CHGs: GPX8, ANTXR2, and AVPR1A. Glutathione peroxidase 8 (GPX8) has been demonstrated as crucial for maintaining the invasive phenotype in breast cancer (37); however, direct evidence linking GPX8 to breast cancer tumor angiogenesis is currently lacking. Hence, further investigation is warranted to explore the involvement of GPX8 in angiogenesis within breast cancer. ANTXR2 is a type I membrane protein participant in extracellular matrix homeostasis (38). Cholesterol depletion induces ANTXR2-dependent activation of MMP-2 in glioma cells (39). Down-regulation of ANTXR2 expression inhibits proliferation and capillary network formation in human umbilical vein endothelial cells (HUVECs) (40). Vasopressin receptor 1A (AVPR1A) serves as a pivotal receptor for vasoconstriction (41). Notably, patients with higher CHG expression scores showed reduced expression of the cluster vascular stability genes JAM2, CDH5 (VE calcineurin), CLDN5 (Claudin 5), TIE1, and TEK (TIE2).Under normal

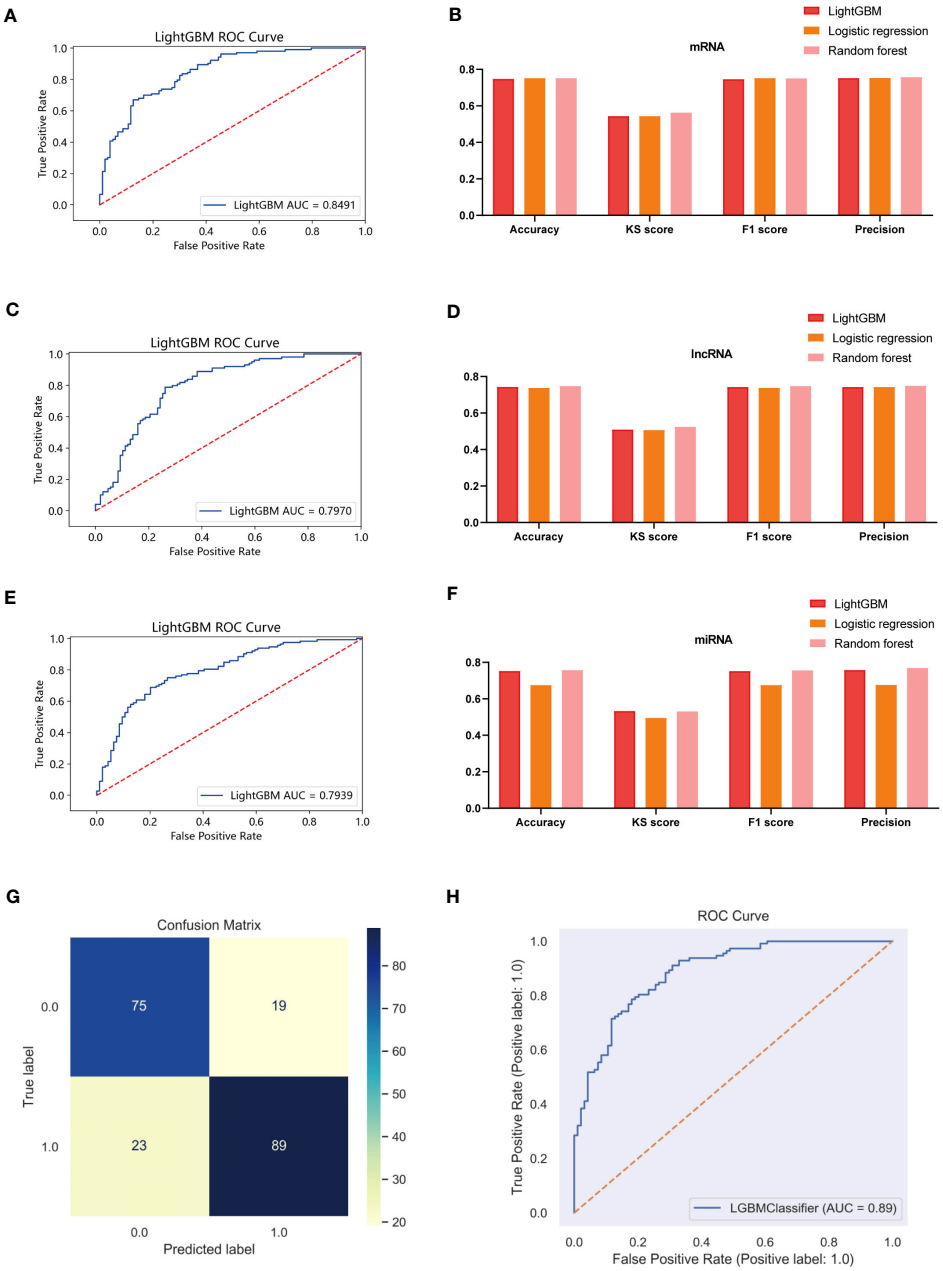


FIGURE 11
Multi-model comparison based on single molecular plane of characteristic mrna, mirna and lncrna. **(A)** ROC curve of mRNA risk model (test set, n=380). **(B)** Accuracy, KS score, F1 score and Precision of mRNA different risk model. **(C)** ROC curve of lncRNA risk model (test set, n=380). **(D)** Accuracy, KS score, F1 score and Precision of lncRNA different risk model. **(E)** ROC curve of miRNA risk model (test set, n=380). **(F)** Accuracy, KS score, F1 score and Precision of miRNA different risk model. **(G)** A confounding matrix for predicting patient prognostic risk with multi-omics data using the test set (N = 380) **(H)** ROC curve of Risklight (test set, n=380).

conditions, the endothelium of mature capillaries is quiescent, stable, and limits vascular leakage (42). Genetic deletion of JAM2, CLDN5, and CDH5 significantly increases vascular permeability and leads to vascular barrier dysfunction (43). Angiopoietin-1 (Ang-1) acts through its receptors TIE1 and TEK (44), and deletion of TIE1 and TEK ultimately leads to reduced vascular stability (45). A number of pathologic disorders can lead to destabilization of the vascular network, resulting in hyperendothelial permeability, excessive vascular outgrowth and angiogenesis. In turn, overgrowth or aberrant remodeling of blood vessels promotes many diseases, including cancers (46). Abnormalities in the vasculature and the resulting microenvironment accelerate tumor progression and lead to reduced efficacy of chemotherapy, radiotherapy, and immunotherapy (47).Therefore, we need to focus on the in-depth link between cholesterol homeostasis and tumor angiogenesis, which may be a potential node for the treatment of breast cancer.

The development of risk stratification tools concerning cancer survivorship has become a priority for research in clinical practice.

We developed a CAG_ score based on CHGs to predict prognostic risk and survival time in BC patients. In general, the higher the CAG_ score is, the worse the prognosis is, combined with a higher risk of death. The CAG_ score is related to the abundance of T lymphocytes ($CD4^+$, $CD8^+$ T) and Macrophages. $CD4^+$ T cells regulate the immune response by producing cytokines (48). $CD8^+$ T cells kill pathogens or produce inflammatory factors and cell division molecules (49). With the increase in risk score, the level of T cells showed a downward trend, which meant the level of immunity was decreased, and the patients were more vulnerable to tumors. In addition, the level of M1 macrophages decreased while the level of M2 increased with the increase of CAG_score. Tumor-associated macrophages (TAMs) are considered essential tumor-associated immune cells by promoting tumor growth, invasion, and metastasis, which contain two subtypes with separate functions (50). Typically activated M1 macrophages are known to reduce the survival of tumor cells through direct killing and antibody-dependent cell-mediated cytotoxicity (ADCC) (51).In contrast, M2 macrophages, as TAMs in a narrow sense, can inhibit the immune effect of T lymphocytes and promote tumor angiogenesis, leading to immune escape and tumor progression (52). The prediction of TAMs by the CAG_ score was completely consistent with tumor progression and clinical outcome, which means CAG_ score has a great ability to predict the status of TME in BC patients. Therefore, we recommend that risk stratification of cholesterol metabolism should be considered as a screening test for further investigation, intervention, and support of tumor.

Immune checkpoint inhibitors have been shown to be effective in the treatment of a variety of tumors (53). This can be observed a marked upregulation in the low CAG score group, suggesting that patients with low CAG scores may be more sensitive to immunotherapy. Currently, chemotherapy resistance in BC is getting worse (54). Our study also provides possible susceptibility drugs for patients with different CAG score groups, which could facilitate clinical selective medication.

Nevertheless, the molecular pathways involved in the development of BC have not been elucidated in detail. With multi-omics data exploration, molecular alterations in three different molecular layers (mRNA, miRNA, lncRNA) driven by the tumor microenvironment emerge in different patients. Combining with multi-omics features and the Risklight tool, we further developed a model of risk in BC patients associated with cholesterol homeostasis disequilibrium.

Overall, this study provides valuable insights on the prognosis of breast cancer patients. Reconsideration of alterations in cholesterol homeostasis as potential risk factors for tumor progression is warranted. The intricate relationship between functional changes in cholesterol homeostasis and tumor angiogenesis and immune response remains incompletelyated, necessitating further exploration of the association cholesterol homeostasis genes and angiogenesis as well as immune response. In especially, the roles of SEMA3B and GSTM2 in immunity should be further clarified. Notably, GPX8, a cholesterol homeostasis gene with unknown implications for angiogenesis but exhibiting a strong correlation with it, warrants thorough investigation. Additionally, a more robust prognostic model pertaining to cholesterol must be established, incorporating both the actual levels of patients' cholesterol and those of cholesterol homeostasis genes. This will significantly enhance the accuracy of breast cancer prognosis models related to cholesterol, bringing them closer to clinical research and practice. Finally, we hope that the secrets of cholesterol homeostasis in breast cancer will increasingly be revealed. That's why we started this study.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://xena.ucsc.edu/.

Ethics statement

Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

HW was the main contributor to the manuscript. All authors read and approved the final manuscript.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Young Talent Program [grant number qnyc108].

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1246880/full#supplementary-material

References

1. Woolston C. Breast cancer. *Nature* (2015) 527(7578):S101. doi: 10.1038/527S101a

2. Jokar N, Velikyan I, Ahmadzadehfar H, Rekabpour SJ, Jafari E, Ting HH, et al. Theranostic approach in breast cancer: A treasured tailor for future oncology. *Clin Nucl Med* (2021) 46(8):e410–20. doi: 10.1097/RLU.0000000000003678

3. Ahmad F, Sun Q, Patel D, Stommel JM. Cholesterol metabolism: a potential therapeutic target in glioblastoma. *Cancers (Basel)* (2019) 11(2). doi: 10.3390/cancers11020146

4. Umetani M, Domoto H, Gormley AK, Yuhanna IS, Cummins CL, Javitt NB, et al. 27-Hydroxycholesterol is an endogenous SERM that inhibits the cardiovascular effects of estrogen. *Nat Med* (2007) 13(10):1185–92. doi: 10.1038/nm1641

5. Wei W, Schwaid AG, Wang X, Wang X, Chen S, Chu Q, et al. Ligand activation of ERRα by cholesterol mediates statin and bisphosphonate effects. *Cell Metab* (2016) 23(3):479–91. doi: 10.1016/j.cmet.2015.12.010

6. Liu W, Chakraborty B, Safi R, Kazmin D, Chang CY, McDonnell DP, et al. Dysregulated cholesterol homeostasis results in resistance to ferroptosis increasing tumorigenicity and metastasis in cancer. *Nat Commun* (2021) 12(1):5103. doi: 10.1038/s41467-021-25354-4

7. Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol* (2015) 17(4):351–9. doi: 10.1038/ncb3124

8. Luo J, Yang H, Song BL. Mechanisms and regulation of cholesterol homeostasis. *Nat Rev Mol Cell Biol* (2020) 21(4):225–45. doi: 10.1038/s41580-019-0190-7

9. Keenan TE, Tolaney SM. Role of immunotherapy in triple-negative breast cancer. *J Natl Compr Canc Netw* (2020) 18(4):479–89. doi: 10.6004/jnccn.2020.7554

10. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* (2013) 19(11):1423–37. doi: 10.1038/nm.3394

11. Yuan J, Cai T, Zheng X, Ren Y, Qi J, Lu X, et al. Potentiating CD8(+) T cell antitumor activity by inhibiting PCSK9 to promote LDLR-mediated TCR recycling and signaling. *Protein Cell* (2021) 12(4):240–60. doi: 10.1007/s13238-021-00821-2

12. Birney E. Mendelian randomization. *Cold Spring Harb Perspect Med* (2022) 12(4). doi: 10.1101/cshperspect.a041302

13. Seiler M, Huang CC, Szalma S, Bhanot G. ConsensusCluster: a software tool for unsupervised cluster discovery in numerical data. *Omics* (2010) 14(1):109–13. doi: 10.1089/omi.2009.0083

14. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

15. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. *Otolaryngol Head Neck Surg* (2010) 143(3):331–6. doi: 10.1016/j.otohns.2010.05.007

16. Meng Z, Ren D, Zhang K, Zhao J, Jin X, Wu H. Using ESTIMATE algorithm to establish an 8-mRNA signature prognosis prediction system and identify immunocyte infiltration-related genes in Pancreatic adenocarcinoma. *Aging (Albany NY)* (2020) 12(6):5048–70. doi: 10.18632/aging.102931

17. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12

18. Liu Y, Hu J, Liu D, Zhou S, Liao J, Liao G, et al. Single-cell analysis reveals immune landscape in kidneys of patients with chronic transplant rejection. *Theranostics* (2020) 10(19):8851–62. doi: 10.7150/thno.48201

19. Garcia-Magariños M, Antoniadis A, Cao R, Gonzãlez-Manteiga W. Lasso logistic regression, GSoft and the cyclic coordinate descent algorithm: application to gene expression data. *Stat Appl Genet Mol Biol* (2010) 9:30. doi: 10.2202/1544-6115.1536

20. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS One* (2014) 9(9):e107468. doi: 10.1371/journal.pone.0107468

21. Wang X, Yang Y, Liu J, Wang G. The stacking strategy-based hybrid framework for identifying non-coding RNAs. *Brief Bioinform* (2021) 22(5). doi: 10.1093/bib/bbab023

22. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* (2015) 43(Database issue):D447–52. doi: 10.1093/nar/gku1003

23. Akwii RG, Sajib MS, Zahra FT, Mikelis CM. Role of angiopoietin-2 in vascular physiology and pathophysiology. *Cells* (2019) 8(5). doi: 10.3390/cells8050471

24. Bazzoni G. The JAM family of junctional adhesion molecules. *Curr Opin Cell Biol* (2003) 15(5):525–30. doi: 10.1016/S0955-0674(03)00104-2

25. Kakogiannos N, Ferrari L, Giampietro C, Scalise AA, Maderna C, Ravà M, et al. JAM-A acts *via* C/EBP-α to promote claudin-5 expression and enhance endothelial barrier function. *Circ Res* (2020) 127(8):1056–73. doi: 10.1161/CIRCRESAHA.120.316742

26. Korhonen EA, Lampinen A, Giri H, Anisimov A, Kim M, Allen B, et al. Tie1 controls angiopoietin function in vascular remodeling and inflammation. *J Clin Invest* (2016) 126(9):3495–510. doi: 10.1172/JCI84923

27. Mao XG, Xue XY, Wang L, Zhang X, Yan M, Tu YY, et al. CDH5 is specifically activated in glioblastoma stemlike cells and contributes to vasculogenic mimicry induced by hypoxia. *Neuro Oncol* (2013) 15(7):865–79. doi: 10.1093/neuonc/not029

28. Szeto GL, Finley SD. Integrative approaches to cancer immunotherapy. *Trends Cancer* (2019) 5(7):400–10. doi: 10.1016/j.trecan.2019.05.010

29. Jenkins RW, Barbie DA, Flaherty KT. Mechanisms of resistance to immune checkpoint inhibitors. *Br J Cancer* (2018) 118(1):9–16. doi: 10.1038/bjc.2017.434

30. Dong C. Cytokine regulation and function in T cells. *Annu Rev Immunol* (2021) 39:51–76. doi: 10.1146/annurev-immunol-061020-053702

31. Kennedy LB, Salama AKS. A review of cancer immunotherapy toxicity. *CA Cancer J Clin* (2020) 70(2):86–104. doi: 10.3322/caac.21596

32. Ganeshan K, Chawla A. Metabolic regulation of immune responses. *Annu Rev Immunol* (2014) 32:609–34. doi: 10.1146/annurev-immunol-032713-120236

33. Kidani Y, Elsaesser H, Hock MB, Vergnes L, Williams KJ, Argus JP, et al. Sterol regulatory element-binding proteins are essential for the metabolic programming of effector T cells and adaptive immunity. *Nat Immunol* (2013) 14(5):489–99. doi: 10.1038/ni.2570

34. Cen M, Ouyang W, Lin X, Du X, Hu H, Lu H, et al. FBXO6 regulates the antiviral immune responses *via* mediating alveolar macrophages survival. *J Med Virol* (2023) 95(1):e28203. doi: 10.1002/jmv.28203

35. Di Pilato M, Kfuri-Rubens R, Pruessmann JN, Ozga AJ, Messemaker M, Cadilha BL, et al. CXCR6 positions cytotoxic T cells to receive critical survival signals in the tumor microenvironment. *Cell* (2021) 184(17):4512–4530.e22. doi: 10.1016/j.cell.2021.07.015

36. Viallard C, Larrivée B. Tumor angiogenesis and vascular normalization: alternative therapeutic targets. *Angiogenesis* (2017) 20(4):409–26. doi: 10.1007/s10456-017-9562-9

37. Khatib A, Solaimuthu B, Ben Yosef M, Abu Rmaileh A, Tanna M, Oren G, et al. The glutathione peroxidase 8 (GPX8)/IL-6/STAT3 axis is essential in maintaining an aggressive breast cancer phenotype. *Proc Natl Acad Sci U.S.A.* (2020) 117(35):21420–31. doi: 10.1073/pnas.2010275117

38. Deuquet J, Lausch E, Superti-Furga A, van der Goot FG. The dark sides of capillary morphogenesis gene 2. *EMBO J* (2012) 31(1):3–13. doi: 10.1038/emboj.2011.442

39. Zou J, Xu L, Ju Y, Zhang P, Wang Y, Zhang B. Cholesterol depletion induces ANTXR2-dependent activation of MMP-2 *via* ERK1/2 phosphorylation in neuroglioma U251 cell. *Biochem Biophys Res Commun* (2014) 452(1):186–90. doi: 10.1016/j.bbrc.2014.06.001

40. Reeves CV, Dufraine J, Young JA, Kitajewski J. Anthrax toxin receptor 2 is expressed in murine and tumor vasculature and functions in endothelial proliferation and morphogenesis. *Oncogene* (2010) 29(6):789–801. doi: 10.1038/onc.2009.383

41. Russell JA. Vasopressor therapy in critically ill patients with shock. *Intensive Care Med* (2019) 45(11):1503–17. doi: 10.1007/s00134-019-05801-z

42. London NR, Whitehead KJ, Li DY. Endogenous endothelial cell signaling systems maintain vascular stability. *Angiogenesis* (2009) 12(2):149–58. doi: 10.1007/s10456-009-9130-z

43. Wang L, Lin L, Qi H, Chen J, Grossfeld P. Endothelial loss of ETS1 impairs coronary vascular development and leads to ventricular non-compaction. *Circ Res* (2022) 131(5):371–87. doi: 10.1161/CIRCRESAHA.121.319955

44. D'Amico G, Korhonen EA, Anisimov A, Zarkada G, Holopainen T, Hägerling R, et al. Tie1 deletion inhibits tumor growth and improves angiopoietin antagonist therapy. *J Clin Invest* (2014) 124(2):824–34. doi: 10.1172/JCI68897

45. La Porta S, Roth L, Singhal M, Mogler C, Spegg C, Schieb B, et al. Endothelial Tie1-mediated angiogenesis and vascular abnormalization promote tumor progression and metastasis. *J Clin Invest* (2018) 128(2):834–45. doi: 10.1172/JCI94674

46. Hu J, Frömel T, Fleming I. Angiogenesis and vascular stability in eicosanoids and cancer. *Cancer Metastasis Rev* (2018) 37(2-3):425–38. doi: 10.1007/s10555-018-9732-2

47. Goel S, Duda DG, Xu L, Munn LL, Boucher Y, Fukumura D, et al. Normalization of the vasculature for treatment of cancer and other diseases. *Physiol Rev* (2011) 91(3):1071–121. doi: 10.1152/physrev.00038.2010

48. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol* (2015) 15(8):486–99. doi: 10.1038/nri3862

49. Wherry EJ. T cell exhaustion. *Nat Immunol* (2011) 12(6):492–9. doi: 10.1038/ni.2035

50. Mantovani A, Sica A, Locati M. Macrophage polarization comes of age. *Immunity* (2005) 23(4):344–6. doi: 10.1016/j.immuni.2005.10.001

51. Atri C, Guerfali FZ, Laouini D. Role of human macrophage polarization in inflammation during infectious diseases. *Int J Mol Sci* (2018) 19(6). doi: 10.3390/ijms19061801

52. Anderson NR, Minutolo NG, Gill S, Klichinsky M. Macrophage-based approaches for cancer immunotherapy. *Cancer Res* (2021) 81(5):1201–8. doi: 10.1158/0008-5472.CAN-20-2990

53. Galluzzi L, Humeau J, Buqué A, Zitvogel L, Kroemer G. Immunostimulation with chemotherapy in the era of immune checkpoint inhibitors. *Nat Rev Clin Oncol* (2020) 17(12):725–41. doi: 10.1038/s41571-020-0413-z

54. Garcia-Martinez L, Zhang Y, Nakata Y, Chan HL, Morey L. Epigenetic mechanisms in breast cancer therapy and resistance. *Nat Commun* (2021) 12 (1):1786. doi: 10.1038/s41467-021-22024-3

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact