



Prognostic and tumor-immune infiltration cell signatures in tamoxifen-resistant breast cancers

Zhenyu Cao^{1#}, Ziwei Jin^{1#}, Liyun Zeng¹, Hongye He¹, Qitong Chen¹, Qiongyan Zou¹, Dengjie Ouyang¹, Na Luo¹, Yulong Zhang², Yunchang Yuan^{3*}, Wenjun Yi^{1*}

¹Department of General Surgery, the Second Xiangya Hospital of Central South University, Changsha, China; ²Department of Glandular Surgery, Baise People's Hospital, Baise, China; ³Department of Thoracic Surgery, the Second Xiangya Hospital of Central South University, Changsha, China

Contributions: (I) Conception and design: W Yi, Y Yuan; (II) Administrative support: W Yi, Y Yuan, Z Cao, Z Jin; (III) Provision of study materials or patients: W Yi, Z Cao, Z Jin, L Zeng, H He, Q Zou, Q Chen; (IV) Collection and assembly of data: Z Cao, Z Jin, L Zeng, H He, Q Zou, N Luo; (V) Data analysis and interpretation: Z Cao, Z Jin, D Ouyang, Y Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

^{*}These authors contributed equally to this work as co-corresponding authors.

Correspondence to: Yunchang Yuan. Department of Thoracic Surgery, the Second Xiangya Hospital of Central South University, 139 Middle Renmin Road, Changsha 410011, China. Email: yuanyc@csu.edu.cn; Wenjun Yi. Department of General Surgery, the Second Xiangya Hospital of Central South University, 139 Middle Renmin Road, Changsha 410011, China. Email: yiwenjun@csu.edu.cn.

Background: The cumulative risk of distant recurrence of hormone receptor-positive (HR+) breast cancer in the past 20 years has ranged from 22% to 52% after 5 years of endo-therapy. The TNM stage, histological grade, and age are important clinical factors related to recurrence, however the exact mechanism of tamoxifen resistance is still unclear.

Methods: Differentially expressed genes (DEGs) were identified in 10 pairs of patients who had relapsed and non-relapsed after tamoxifen treatment based on matching their clinicopathological factors. After analysis of the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, 10 hub genes were identified using Cytoscape software. Next, real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database were used to verify the expression and overall survival (OS) of the 10 hub genes respectively, and GSE96058 and Kaplan-Meier Plotter website were used to further verify the OS of *C3*, *CX3CL1*, *CXCL2*, and *SAA1*. Finally, Immune Cell Abundance Identifier (ImmuCellAI) and the TIMER database were used to estimate immune cell infiltration and the expression of prognostic genes.

Results: The DEGs were mainly enriched in the inflammatory response and cytokine-receptor interaction. The expression and the survival analysis identified *CX3CL1*, *CXCL2*, and *SAA1* as prognostic factors, whose overexpression in HR+/human epidermal growth factor receptor 2 (HER-2) negative breast cancer possibly predicted a longer disease-free survival. The expression levels of these 3 genes are positively correlated with immune cell infiltration. Their high expression levels may predict longer disease-free survival in breast cancer after tamoxifen treatment and may be biomarkers for tamoxifen-resistant therapy.

Conclusions: In conclusion, the high expression of *CX3CL1*, *CXCL2*, and *SAA1* may predict longer disease-free survival in breast cancer after tamoxifen treatment and may be a biomarker for tamoxifen therapy.

Keywords: Tamoxifen; immune; breast cancer; The Cancer Genome Atlas (TCGA); resistance

Submitted Aug 05, 2021. Accepted for publication Sep 16, 2021.

doi: 10.21037/gs-21-566

View this article at: <https://dx.doi.org/10.21037/gs-21-566>

Introduction

Breast cancer, which is the most common malignant tumor in women worldwide, can be classified into 4 main subtypes according to the molecular profile of the tumor, including hormone receptor-positive/human epidermal growth factor receptor 2 (HR+/HER-2) negative, HR+/HER-2 positive, HR-/HER-2 positive, and HR-/HER-2 negative breast cancers (1). Over 2/3 of patients are estrogen receptor (ER)-positive or progesterone receptor (PR)-positive and thus benefit from endocrine therapy, including the selective ER antagonist tamoxifen and aromatase inhibitors, which inhibit the effects of estrogen in breast tissue (2). Tamoxifen, which is widely used in adjuvant therapy, has been indicated to substantially reduce the recurrence rate by approximately 40% and mortality by approximately 1/3 throughout the first 15 years in HR+ patients (3). However, 12.1% of patients still express high HR relapse even with 5 years of tamoxifen therapy, and the cumulative risk of recurrence in the past 20 years is still as high as 22–52% (4,5). Thus, recurrence after tamoxifen therapy is still the main and critical clinical challenge for ER+ patients.

Accumulating efforts have been made to investigate the mechanism of relapse after tamoxifen treatment. The failure of endocrine therapy has been reported to involve *ESR1* mutations (6,7), the downregulation of ER activation, and the amplification of the PI3K/AKT/mammalian target of rapamycin (mTOR) signaling pathway (8). Evidence suggests that the overexpression of HER-2 confers tamoxifen resistance in ER+ breast cancer through the MAPK signaling pathway (9). Also, the overexpression of *SPP1* gene in primary tumors was associated with higher risk of recurrence in ER+ breast cancer with endocrine treatment (10). What's more, the higher *PPF1A1* expression predicts poor response to endocrine therapy (11). Despite the mechanisms of endocrine therapy resistance are diverse and face great challenges. The current NCCN guidelines recommend successive endocrine therapy until three sequential endocrine therapy regimens could not offer clinical benefit or until symptomatic visceral disease progression occurs. In addition, new molecular targeted drugs, such as inhibitors of mTOR, CDK4/6, HDAC and PIK3CA, recently have become clinically available for treatments in combination with endocrine treatment for HR+/HER2- metastatic cancer patients (12,13). Palbociclib (14), an inhibitor of CDK4/6, has been approved by the Food and Drug Administration (FDA) as a first-line treatment for drug resistance in advanced

endocrine therapy; however, unfortunately, its overall response rate (ORR) was only 42.5%, and its clinical benefit rate (CBR) was 84.5%, with some patients still participating in the Paloma-2 trial (15).

Although little is known about the role of the immune system in endocrine therapy, tamoxifen has been reported to have multiple effects on immunity. For example, tamoxifen shifted CD4+ T cells from T helper 1 (Th1) to Th2 immunity (16) and thus can promote pro-tumor immunosuppression. Impaired anti-tumor immunity by tamoxifen may occur mostly in women with elevated Th2 immunity evident before treatment (17). However, it has also been reported that the effects of tamoxifen on immunity may explain why it is inferior to aromatase inhibitors in preventing disease recurrence in breast cancers (16). Therefore, searching for immune-related biomarkers of drug sensitivity in endocrine therapy is of great clinical and scientific value.

Our study aimed to find the markers of sensitivity to tamoxifen therapy by matching the clinical risk factors such as age, histological grading, tumor, node, metastasis (TNM) staging, and menstrual status in 10 pairs of patients who had relapsed and not relapsed after tamoxifen treatment. To sum up, our results suggested that the recurrence of tamoxifen was possibly associated with immune infiltration, and the high expression of *CX3CL1*, *CXCL2*, and *SAA1* may could predict longer prognosis after endocrine treatment, which may be identified as the biomarker for sensitivity to tamoxifen treatment, and the further mechanism of these biomarker remain to be investigated in the future studies. We present the following article in accordance with the REMARK reporting checklist (available at <https://dx.doi.org/10.21037/gs-21-566>).

Methods

Data acquisition and preprocessing

Processed and deidentified patient breast cancer (BRCA) data [The Cancer Genome Atlas (TCGA) “level 3” data designation] were downloaded from TCGA (2016) (http://gdac.broadinstitute.org/runs/stddata__2016_01_28/data/BRCA/20160128/). A total of 10 pairs of HR+/HER-2 negative invasive ductal breast cancers, including 10 recurrence and 10 non-recurrence samples after tamoxifen therapy, were matched according to the patients' age, grading, menstrual status, and TNM. The clinical characteristics of those 20 patients are presented in *Table 1*.

Table 1 Clinical information on the recurrence and non-recurrence HR+/HER-2 negative tamoxifen treated breast cancers from the TCGA website

Sample	Recurrence				Non-recurrence			
	Age (years)	Menstrual status	Histologic grading	TNM	Age (years)	Menstrual status	Histologic grading	TNM
A	61	Post	Stage iib	T2N1M0	63	Post	Stage iib	T2N1M0
B	59	Post	Stage iib	T2N1aM0	59	Post	Stage iib	T2N1aM0
C	47	Pre	Stage iiiia	T2N2M0	49	Pre	Stage iiiia	T2N2aM0
D	39	Pre	Stage iia	T2N0M0	39	Pre	Stage iia	T2N0M0
E	26	Pre	Stage iib	T2N1M0	28	Pre	Stage iib	T2N1M0
F	43	Pre	Stage iiiia	T2N2M0	39	Pre	Stage iiiia	T2N2M0
G	39	Pre	Stage iib	T1cN0M0	42	Pre	Stage iib	T1N0M0
H	82	Post	Stage iv	T2N2M0	81	Post	Stage iiiia	T1cN2aM0
I	61	Post	Stage iia	T2N0M0	61	Post	Stage iia	T2N0M0
J	69	Post	Stage i	T1cN0M0	63	Post	Stage ia	T1cN0M0

TCGA, The Cancer Genome Atlas; TNM, tumor, node, metastasis.

Microarray-based dataset GSE9893, a cohort of 132 primary tumors from tamoxifen-treated patients followed up for more than 5 years, with the study endpoint of relapse, was obtained from websites of Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Differentially expressed gene (DEG) analysis of the two groups was performed using the edgeR (18) package (version 3.24.3) in R (<https://www.rdocumentation.org/packages/edgeR/>). The P value was adjusted using the false discovery rate (FDR). An FDR <0.05 and $|\log_2\text{FC}| > 1$ was set as the threshold for significantly differential expression.

GO and KEGG analysis of DEGs

All DEGs were into the Database for Annotation, Visualization, and Integrated Discovery (DAVID) website (19,20) (<https://david.ncifcrf.gov/>) to explore the biological process, molecular function, and cellular component categories as well as KEGG pathways with the threshold of count >5 and P value <0.05. The results of the enrichment analysis were visualized by the ggplot2 package (21) (version 3.3.0) in R (<https://cran.r-project.org/web/packages/ggplot2/index.html>).

Protein-protein interaction (PPI) network analysis and identification of hub genes

The PPI (22) network explores the downstream relationship between proteins based on physical binding and genetic and functional relationships. Analyzing the functional interactions between proteins may provide insights into the mechanisms underlying the generation or development of diseases. We integrated 647 DEGs using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) website (23) (<https://string-db.org/>) to explore the association between these DEGs. The cut off value for STRING analysis is 0.04. Cytoscape (version 3.7.1; <https://cytoscape.org/>) was used to visualize the PPI network from the STRING website, recognize the hub genes with the CytoHubba (24) APP (version 0.1) and MCC algorithm, and produce the diagrams shown below.

The survival analysis of the hub genes

Overall survival (OS) analysis was conducted using Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) clinical datasets downloaded from cBioPortal (25,26) (<https://www.cbioportal.org/>). Patients were filtered for missing values, and 739 HR+/

HER-2 negative patients who had also received hormone therapy were selected. The Kaplan–Meier survival curves of 10 hub genes were generated by the Survival and survminer package in R (version 3.6.1; <https://cran.r-project.org/web/packages/survminer/index.html>). The prognosis of prognostic hub genes was further verified in the GSE96058. This database contained 3,678 breast cancers, of whom 3,273 patients included a median follow-up of 52 months, and 1,953 patients who were ER+/HER-2 negative and received endocrine therapy were selected to identify the OS. The Kaplan–Meier Plotter website (27) (<http://kmplot.com/analysis/index.php?p=background>) was also used to verify the OS of tamoxifen-treated patients.

Inference on the number of tumor-infiltrating immune cells and the correlation between prognostic genes and the immune cells

In the analysis of immune infiltration across a spectrum of cancer types, we analyzed the infiltration of the immune cells on the websites of Immune Cell Abundance Identifier (ImmuCellAI; <http://bioinfo.life.hust.edu.cn/ImmuCellAI#!/>) (28), which is a tool to estimate the abundance of 24 immune cell types, including 18 T-cell subsets and 6 other immune cells: B cell, natural killer (NK) cell, Monocyte cell, Macrophage cell, Neutrophil cell, and dendritic cell (DC), from gene expression data. Then, TIMER (29,30) (<https://cistrome.shinyapps.io/timer/>) was used to access the correlation between the tumor-infiltrating immune cells and the immune-related genes in the luminal breast cancers. As a comprehensive resource, TIMER applies a previously published statistical approach called deconvolution that uses gene expression profiles to produce an inference on the number of tumor-infiltrating immune cells (TIICs). These immune infiltrates include B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and DCs via gene modules.

Cell culture and quantitative real-time polymerase chain reaction (qRT-PCR)

We obtained MCF-7 and LCC-2 cells (a tamoxifen-resistant cell line) from American Type Culture Collection (Rockville, MD, USA) which were then grown in Dulbeccos; modified Eagle medium (DMEM; Mediatech, Inc., Herndon, VA, USA) supplemented with antibiotics, 10 mm HEPES, and 10% fetal calf serum (FCS; Gibco, Invitrogen Corporation, Carlsbad, CA, USA) at 37 °C in a humidified atmosphere of

5% CO₂. Total RNA was isolated from cells or tissues using Trizol reagent (Invitrogen, USA). Then, 1 µg of total RNA was reverse transcribed to complementary DNA (cDNA) using PrimeScript RT Reagent Kit (Takara, Kusatsu, Shiga, Japan) according to the manufacturer's instructions. Quantitative real-time polymerase chain reaction (qRT-PCR) was performed using the SYBR Green (Takara, Japan) to detect the messenger RNA (mRNA) expression levels. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as endogenous control, and all reactions were performed in triplicate. Results were calculated using the 2^{-ΔΔCt} method. The primers for the top 10 hub genes and GAPDH are shown in [Table S1](#).

Statistical analysis

All the data were analyzed by R (<https://www.r-project.org/>) and GraphPad Prism 7 (GraphPad Prism Software, Inc., San Diego, CA, USA). Quantitative data were expressed as mean ± standard deviation (SD) and analyzed by one way analysis of variance (ANOVA) test. The Kaplan–Meier method was used to estimate the significant difference in survival between the over-expression group and the low-expression group. The significance level of statistical difference was set at P<0.05.

Results

Identification of DEGs and the enrichment analysis

To identify gene expression related to tamoxifen resistance, we matched 10 pairs of breast cancer samples from TCGA database of patients who received tamoxifen therapy according to their age and pathological characteristics ([Table 1](#)); all of these patients relapsed within 5 years. The process of screening for suitable patients is shown in [Figure 1A](#). We identified 647 DEGs (Differentially Expressed Genes) between recurrent and non-recurrent patients with the 506 up-regulated genes and 141 down-regulated genes, according to the FDR <0.05 and |log₂[fold change (FC)]| >1 ([Figure 1B](#)). The DEGs were annotated in the GO and KEGG databases and used to perform the enrichment analysis by using the DAVID website. The results of the DEGs annotated in the GO database are shown in [Figure 1C–1E](#). In the category of biological processes, the DEGs were mainly enriched in cytoskeletal organization, cell-cell signaling, BMP signaling, inflammatory response, and so on. In the cellular

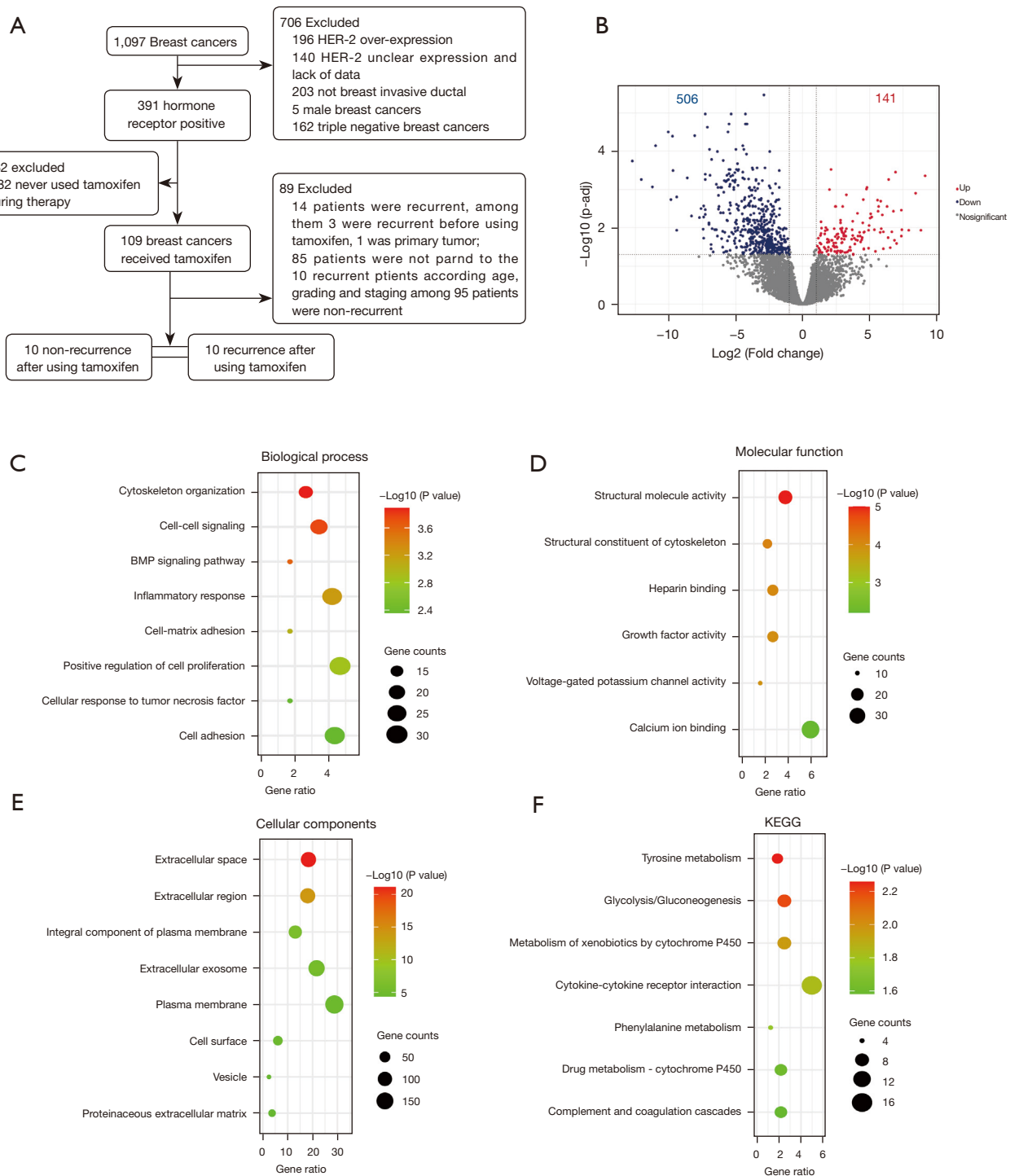


Figure 1 Identification of DEGs and the enrichment analysis. (A) The process of screening suitable patients. (B) The volcano plot of the TCGA datasets. Each point represents a gene. Red nodes represent DEGs with $\log_{2}\text{FC} > 1$ and $\text{FDR} < 0.05$. Blue nodes represent DEGs with $\log_{2}\text{FC} < -1$ and $\text{FDR} < 0.05$. (C-E) the GO terms of 647 DEGs in 3 categories (BP, MF, CC). (F) The KEGG enrichment pathways of 647 DEGs. The depth of the color reflects the value of the $-\log_{10}(\text{P value})$; red presents the largest $-\log_{10}(\text{P value})$, and the size of the dot reflects the number of genes. DEGs, differentially expressed genes; TCGA, The Cancer Genome Atlas; GO, Gene Ontology; BP, biological process; MF, molecular function; CC, cellular component.

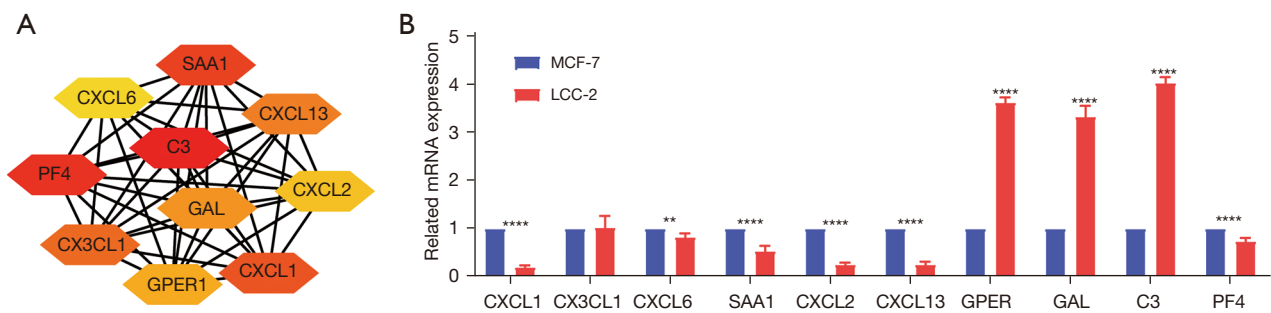


Figure 2 Identification of hub genes and validation of their expression *in vitro*. (A) The most significant module was selected by cytoHubba in Cytoscape using the MCC algorithm. (B) Validation the related mRNA expressions of the top 10 hub genes in MCF-7 and LCC-2 cell line. **, $P < 0.01$; ****, $P < 0.0001$. mRNA, messenger RNA.

Table 2 The expression of the top 10 hub genes

Gene	Log ₂ FC	FDR
<i>CXCL1</i>	-5.4979	<0.0001
<i>CX3CL1</i>	-2.4506	0.0001
<i>CXCL6</i>	-5.3925	0.0005
<i>SAA1</i>	-3.9727	0.0009
<i>CXCL2</i>	-3.2300	0.0170
<i>CXCL13</i>	-3.2820	0.0466
<i>GPER1</i>	2.6748	0.0176
<i>GAL</i>	-2.8103	0.0487
<i>C3</i>	-1.5718	0.0199
<i>PF4</i>	-3.9102	0.0359

Log₂FC, Log₂(Fold Change); FDR, false discovery rate.

component category, the DEGs were mainly enriched in the extracellular space, extracellular region, integral component of the plasma membrane, plasma membrane, and so on. In the category of molecular function, the DEGs were mainly enriched in structural molecule activity, structural constituents of the cytoskeleton, heparin-binding, growth factor activity, and so on. The KEGG pathway analysis (Figure 1F) showed that 16 genes were distributed in cytokine-cytokine receptor interactions, 7 genes were distributed in drug metabolism-cytochrome P450, and 6 genes were distributed in steroid hormone biosynthesis. Herein, the mechanism of tamoxifen therapy failure, whether associated with drug metabolism or inflammatory response, needs further investigation.

Identification of hub genes and validation of their expression *in vitro*

To find out the most important genes associated with tamoxifen resistance, the top 10 hub genes in the PPI network were identified by Cytoscape with the CytoHubba plugin. The top 10 hub genes—*C3*, *PF4*, *SAA1*, *CXCL1*, *CX3CL1*, *CXCL13*, *GAL*, *GPER1*, *CXCL2*, and *CXCL6* are shown in Figure 2A—all of which are inflammatory factors except *GPER1*, which implies that tamoxifen resistance may be closely related to the inflammatory response. The differentiated expression of these 10 genes is shown in Table 2. Previous studies have shown that mediators secreted by cancer-associated fibroblasts (CAFs) (31), in addition to components of the inflammatory response (32) such as cytokines and growth factors, exert an important role in drug resistance. Furthermore, the mRNA expression level of these 10 hub genes was confirmed by qRT-PCR *in vitro* (Figure 2B), and the results showed that the expression level of *PF4*, *SAA1*, *CXCL1*, *CXCL13*, and *CXCL2*, is lower in the LCC-2 than MCF-7, while the expression of *C3*, *GAL*, and *GPER1* is lower in the MCF-7 than LCC-2, and the expression of *CX3CL1* is not significantly differential.

The survival analysis of the hub genes

To further explore the effect of 10 hub genes on the OS of breast cancer patients, we then analyzed the 739 breast cancer patients who were HR+/HER-2 negative and hormone-therapy treated in the METABRIC database, and the outcome showed that upregulated *C3*, *SAA1*, *CXCL2*, and *CX3CL1* predicted a good prognosis (Figure 3A-3D). Moreover, the above 4 genes were also validated a longer

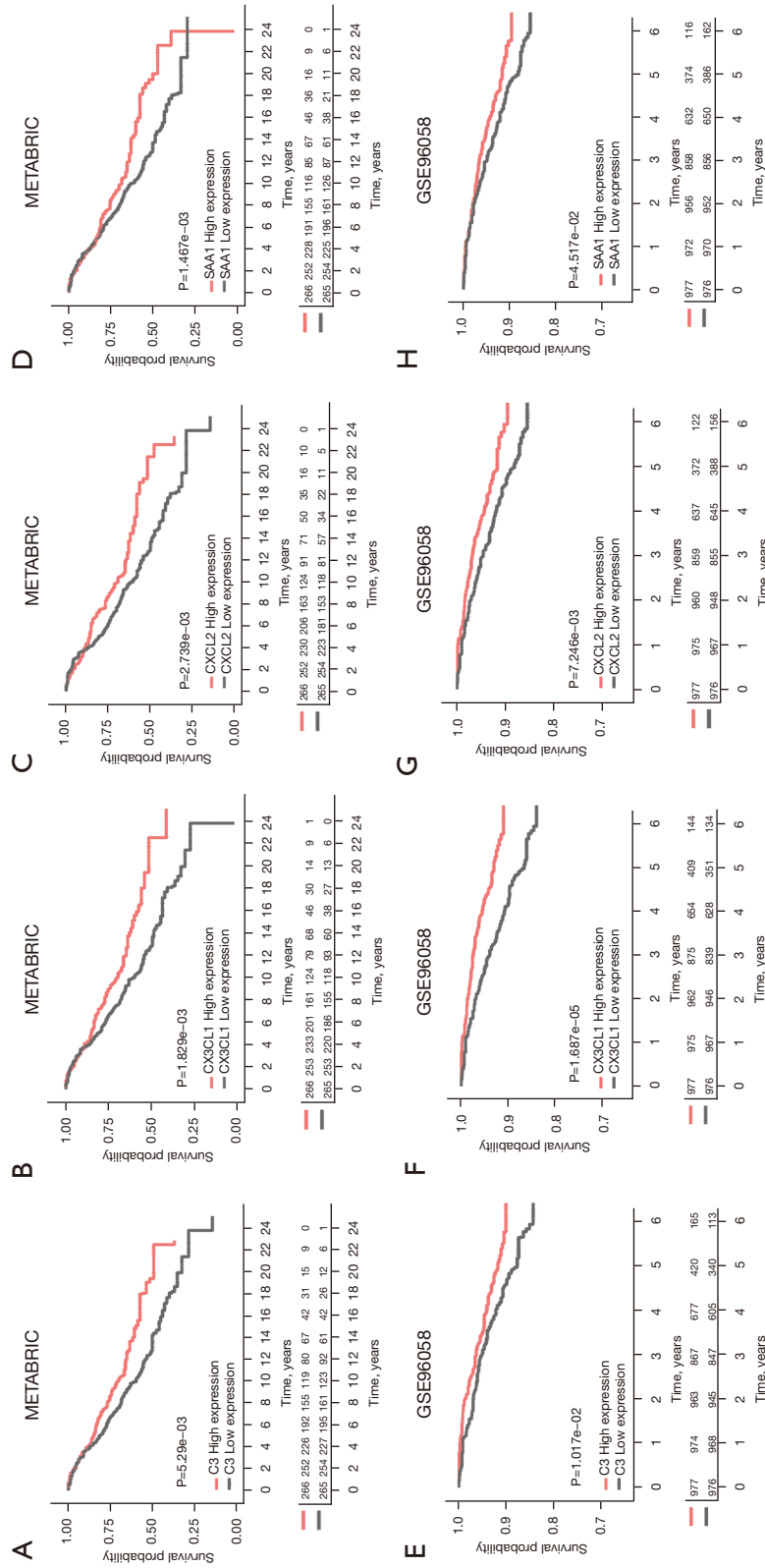


Figure 3 The survival analysis of *C3*, *CX3CL1*, *CXCL2*, and *SAA1*. (A-D) The OS value for the expression of *C3*, *CX3CL1*, *CXCL2*, and *SAA1* based on the METABRIC database. (E-H) The OS value for the expression of the *C3*, *CX3CL1*, *CXCL2*, and *SAA1* based on the GSE96058 database. OS, overall survival.

OS on the GSE96058 from which we selected 1,953 breast cancer patients who were ER+/HER-2 negative and received endocrine therapy (Figure 3E-3H). Additionally, the prognostic trend of these 4 genes is in line with our results on the Kaplan-Meier Plotter (<http://kmplot.com/analysis/index.php?p=background>) website with the ER+ and tamoxifen-treated breast cancers (Figure S1A-S1D).

The above findings indicated that sensitivity to tamoxifen treatment may be associated with the inflammatory response, and the overexpression of these immune genes is associated with a good prognosis. This finding is consistent with a previous study showing that tumor cells could excrete *CX3CL1* to recruit CD8+ T cells and NK cells, thereby inducing both innate and adaptive immunity, which yielded a better prognosis in gastric adenocarcinoma (33).

The differential immune cells infiltration in HR+/HER-2 negative patients

Considering that the immune-related genes and inflammatory response may involve in tamoxifen-resistance, we then compared the abundance of 24 differential immune cell types on the ImmuCellAI with gene expression dataset based on RNA-Seq data to further characterize immune microenvironment of tumors. The results showed that the abundance of Th2 cells, NK, Gamma_delta was higher in the recurrent HR+/HER-2 negative breast cancers than non-recurrent after tamoxifen treatment, while the abundance of CD4+ T cells and MAIT cells were higher in the 10 non-recurrent tamoxifen-treated breast cancers than 10 recurrent from TCGA (Figure 4A). We then verified the immune cell infiltration in the GSE9893 with 132 primary tumors from tamoxifen-treated patients followed up for more than 5 years. As shown in Figure 4B, compared with the recurrent patients, B_cell and CD4+ T cell showed a higher abundance, while the Th2 and Treg showed a lower abundance.

The expression and immune infiltration of *CX3CL1*, *CXCL2*, and *SAA1* in TIMER online datasets

Given the verification of qRT-PCR and the OS analysis of 10 hub genes, 3 (*SAA1*, *CXCL2*, *CX3CL1*) overexpressed genes demonstrated a significant prognostic role of immune genes in hormone-therapy-treated breast cancer patients, except for *C3*, of which the results of qRT-PCR were opposite to the differential analysis. We further investigated the relationships between these 3 genes and

immune infiltrating cells levels in the TIMER (<https://cistrome.shinyapps.io/timer/>) online database. The results showed that the expressions of *SAA1*, *CXCL2*, and *CX3CL1* are higher in normal breast tissue than in breast cancers (Figure S2A-S2C), and the expression levels of all 4 genes are negatively correlated with the tumor purity. Moreover, a positive correlation exists between the expression of *CX3CL1* and immune infiltrating levels of CD8+ T cells ($r=0.274$, $P=1.12e-10$), CD4+ T cells ($r=0.239$, $P=2.22e-08$), neutrophils ($r=0.167$, $P=1.02e-04$), and DCs ($r=0.166$, $P=1.19e-04$) in luminal breast cancers as well as between the expression levels of *CXCL2*, and *SAA1* and the immune cells infiltrating levels (Figure 5A). The results suggested that the higher the expression of *CX3CL1*, *CXCL2*, the lower the proportion of tumor cells, and the higher the proportion of 6 immune cell types. Expression of *SAA1* has a positive correlation with CD8+T cells ($r=0.092$, $P=3.36e-02$) and CD4+ T cells ($r=0.169$, $P=8.83e-05$). Thus, *CX3CL1*, *CXCL2*, and *SAA1* are associated with the recruitment of CD8+ T cells, CD57+T NK cells, and CD1a+ DCs and, therefore, inhibit the development of tumors. Besides, we further verified the correlation between the expression of *CX3CL1*, *CXCL2*, and *SAA1* and immune cell infiltration with 105 tamoxifen-treated breast cancers in TCGA. As shown in Figure 5B, *CX3CL1*, *CXCL2*, and *SAA1* are all positively correlated with the infiltration of CD8+ T cells, while negatively correlated with macrophages (M2) cells.

In general, our data analysis demonstrated a better OS in the HR+/HER2- negative breast cancer patients exhibiting overexpression of *CX3CL1*, *CXCL2*, and *SAA1*. Besides, the high expression of the aforementioned genes has a significant negative association with tumor purity and a positive association with major immune cells.

Discussion

This study aimed to identify the biomarker genes associated with tamoxifen-resistant therapy for assessing prognosis in HR+/HER-2 negative breast cancers. First, we found that the DEGs between relapse and non-relapse HR+/HER-2 negative breast cancers after tamoxifen treatment correlated with the inflammatory response and cytokine-cytokine receptor interaction through matching clinical factors such as TNM stage, histological grade, age, and pathological indicators from TCGA RNA-seq datasets. Then, the expression level of the top 10 hub genes were verified *in vitro* and the OS in METABRIC and GSE96058 datasets. It was suggested that both the expression

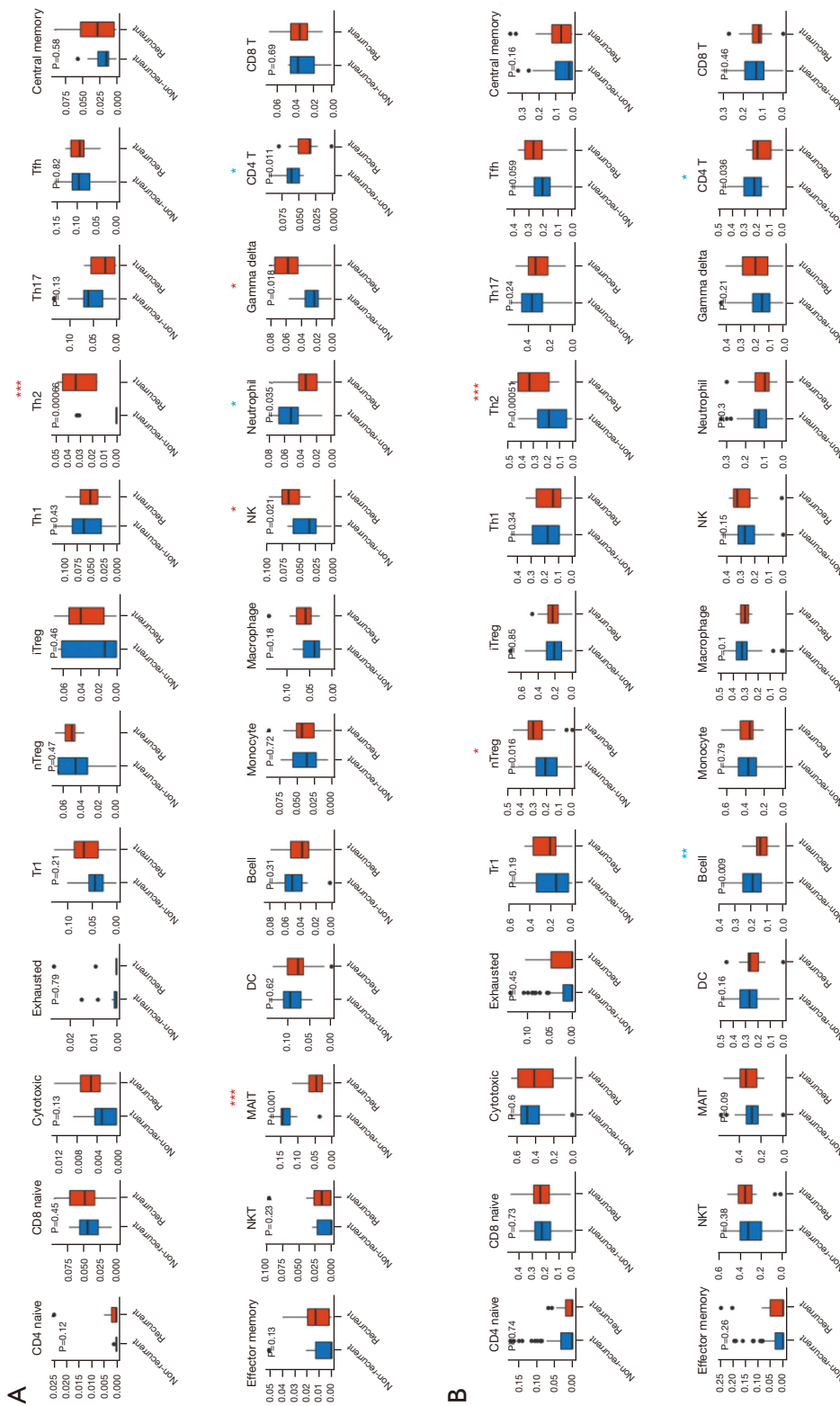


Figure 4 The differential immune cell infiltration in HR+/HER-2 negative patients. (A) The differential immune cell infiltration between the 10 non-recurrent tamoxifen-treated breast cancers and 10 recurrent from TCGA. (B) The validation of the differential immune cell infiltration between the non-recurrent tamoxifen-treated breast cancers and recurrent from GSE9893. The blue asterisk indicates that the number of immune cells in non-recurrence is higher than that in recurrence and is statistically significant. The red asterisk indicates that the number of immune cells in recurrence is higher than that in non-recurrence and is statistically significant. *, P<0.05; **, P<0.01, ***, P<0.001. DC, dendritic cells; Tr1, Type 1 T regulatory cells; Th1, T helper type 1 cell; Th2, T helper type 2 cell; Th17, T helper type 17 cell 17; Tfh, T helper cells; NKT, natural killer T cell; MAIT, mucosal-associated invariant T cell; TCGA, The Cancer Genome Atlas.

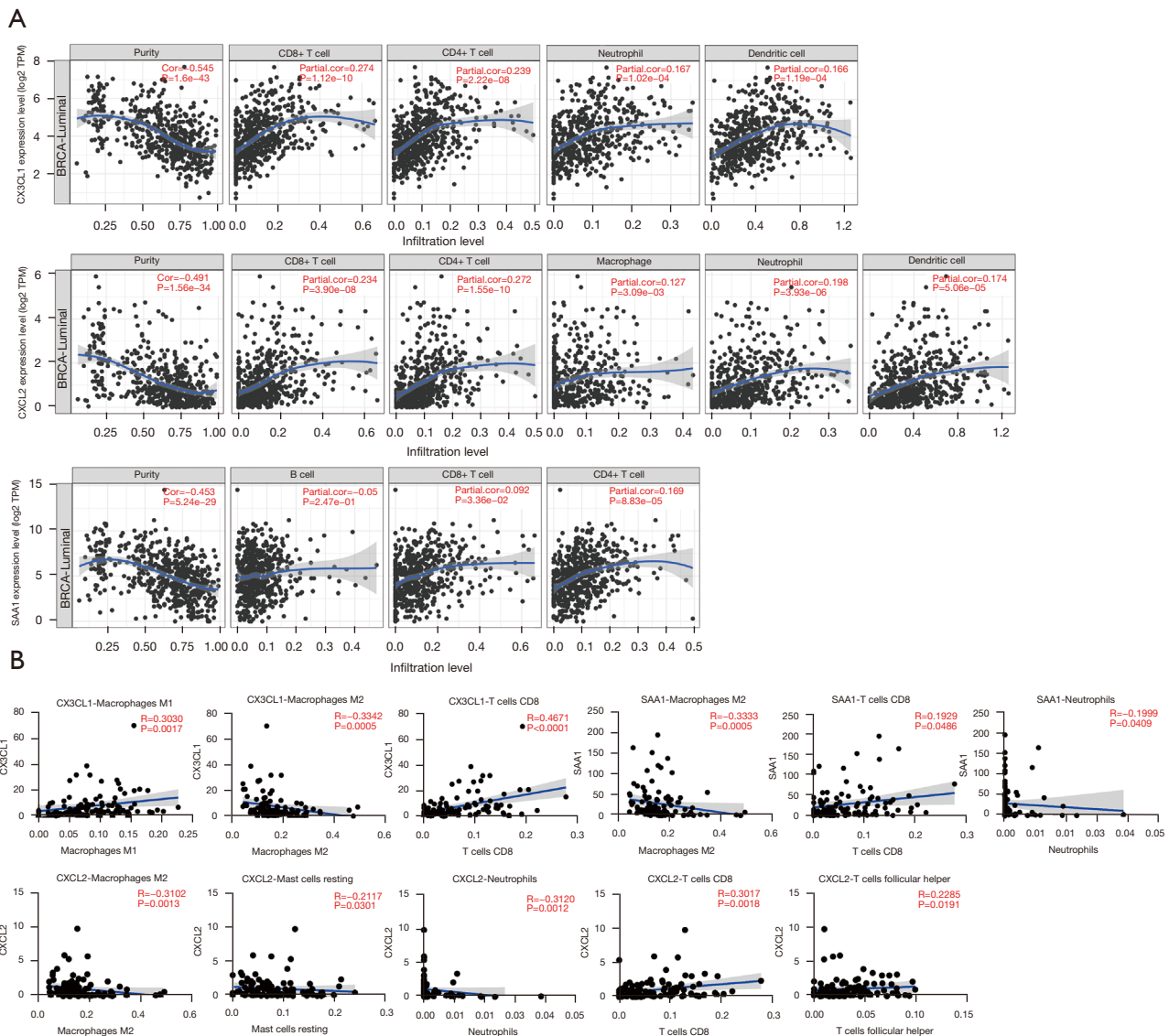


Figure 5 The correlation between the expression of *CX3CL1*, *CXCL2*, and *SAA1* and immune infiltration in luminal breast cancers. (A) *CX3CL1* expression was negatively correlated with immune purity but negatively correlated with CD8+ T cells, CD4+ T cells, neutrophils, and dendritic cells, *CXCL2* expression was negatively correlated with immune purity but positively correlated with CD4+ T cells, CD8+ T cells, macrophages, neutrophils, and DC, *SAA1* expression was negatively correlated with immune purity but positively correlated with CD4+ T cells and CD8+ T cells. (B) The correlation analysis between the expression of *CX3CL1*, *CXCL2*, and *SAA1* and immune cells (only statistically significant graph is shown here). DC, dendritic cells.

in vitro in line with the differential analysis and the survival prognosis of *CX3CL1*, *CXCL2*, and *SAA1* is significant. Besides, their expression levels are verified in the cBioPortal (<https://www.cbioportal.org/>) website. In brief, our results showed that the expression of *CX3CL1*, *CXCL2*, and *SAA1* could be used to predict a good prognosis of HR+/HER-2 negative breast cancer patients who received tamoxifen

treatment and provide clues for the mechanism of relapse after tamoxifen treatment. According to the results of the immune infiltration, the expression of these 3 genes had a negative correlation with immune purity and anti-immune cells (M2), while having a positive relationship with the proportion of pro-immune cells (CD8+ T cell), which suggests that the overexpression of *CX3CL1*, *CXCL2*, and

SAA1 could promote the occurrence of tumor immunity in tamoxifen-treated patients. However, our results only suggested a biomarker related to endocrine recurrence, also due to the few samples was selected in our research, further studies are needed to elucidate the biological function of these genes in breast cancers, and the deeper underlying mechanism of relapse after tamoxifen still needs to be further explored.

Reports on the clinical role of these 3 genes in tumors are contradictory, particularly in *CX3CL1*. Given the dual function of *CX3CL1* as a chemoattractant for leukocytes and an adhesion molecule for tumor cells, it unsurprisingly exerts both pro-tumor and anti-tumor effects. Although *CX3CL1* may have specific positive biological roles in breast cancer progression, the association between the chemokine receptors expressed in primary tumor cells and the site of metastatic relapse has been evaluated in patients with breast cancer (34,35). It has been reported that *CX3CL1* had no effects on proliferation of breast cancer cells, but it could promote the migration and invasion abilities of the CX3CR1-positive cells (36). It was also shown that *CX3CL1* has anti-tumor activity by chemoattraction and activation of T cells, NK cells, and DCs (37-40), and highly expressed *CX3CL1* had a significantly better prognosis for disease-free survival and OS (41). To date, accumulating studies have evaluated the clinicopathological significance of *CX3CL1* in human cancer. High *CX3CL1* expression in colon cancer has been reported to result in a better prognosis compared with low *CX3CL1* expression, which may be associated with fractalkine expressed in colon tumors, which appears to recruit cytotoxic T cells and NK cells to the tumor site. Thus, these cytotoxic cells result in a better prognosis mediated by tumor cell cytotoxicity using a perforin and granzyme B mechanism (42). Another anti-tumor mechanism of *CX3CL1* in tumor is associated with the mutation of the TP53 gene, interfering with the functioning of p53, which could regulate the promoter of *CX3CL1*. Therefore, the mutation of TP53 could reduce the expression of *CXCL3* (43). Therefore, blocking the chemotaxis of *CX3CR1* and *CX3CL1* may be a promising strategy for the anti-tumor.

The gene *CXCL2* is a proinflammatory cytokine, whose receptor is *CXCR2*, that has been shown to be involved in the acquisition of many malignant features, such as cancer cell proliferation, invasion, epithelial-mesenchymal transition (EMT) (44), migration (45), and chemoresistance in many tumors, including those of ER α -positive breast cancer patients treated with tamoxifen (46). However, these

chemokines and corresponding receptors have a role in either the promotion or inhibition of cancer, depending on their capacity to suppress or stimulate the action of the immune system, respectively, leading to their controversial role in cancer. Additionally, previous studies have shown that high *CXCR2* levels are associated with senescence in premalignant lesions and that the loss of *CXCR2* expression contributes to reducing the severity of the arrest via the activation of the p53 pathway (47,48).

The gene *SAA1* is reported to be a potent inflammatory factor, inducing chemotaxis and pro-inflammatory cytokine production by monocytes and neutrophils. It has been implicated in the Stat3-dependent transcription regulatory network controlling metastatic progression. Interestingly, *SAA1* could predict a good prognosis and was overexpressed in normal breast tissue versus tumor tissue, which was in line with a previous study showing that *SAA1* was downregulated during p130Cas-dependent invasion (49).

In summary, our results suggested that recurrence in tamoxifen-treated HR+/HER-2 negative breast cancer patients was may closely associated with immune infiltration. Meanwhile, *CX3CL1*, *CXCL2*, and *SAA1* may be regarded as diagnostic biomarkers for tamoxifen therapy, and their higher expression could predict a good prognosis in our results (50).

Acknowledgments

This manuscript has been released as a pre-print at Research Square.

Funding: This research was supported by the Natural Science Foundation of Hunan Province (Grant NO: 2020JJ4828).

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at <https://dx.doi.org/10.21037/gs-21-566>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/gs-21-566>). All authors reported that this work was supported by Natural Science Foundation of Hunan Province: Grant NO: 2020JJ4828. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all

aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747-52.
2. Lobbezoo DJ, van Kampen RJ, Voogd AC, et al. Prognosis of metastatic breast cancer subtypes: the hormone receptor/HER2-positive subtype is associated with the most favorable outcome. *Breast Cancer Res Treat* 2013;141:507-14.
3. Rangel-Méndez JA, Rubi-Castellanos R, Sánchez-Cruz JF, et al. Tamoxifen side effects: pharmacogenetic and clinical approach in Mexican mestizos. *Transl Cancer Res* 2019;8:23-34.
4. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015;386:1341-52.
5. Pan H, Gray R, Braybrooke J, et al. 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. *N Engl J Med* 2017;377:1836-46.
6. Robinson DR, Wu YM, Vats P, et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet* 2013;45:1446-51.
7. Toy W, Shen Y, Won H, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet* 2013;45:1439-45.
8. Miller TW, Hennessy BT, González-Angulo AM, et al. Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. *J Clin Invest* 2010;120:2406-13.
9. Shou J, Massarweh S, Osborne CK, et al. Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. *J Natl Cancer Inst* 2004;96:926-35.
10. Göthlin Eremo A, Lagergren K, Othman L, et al. Evaluation of SPP1/osteopontin expression as predictor of recurrence in tamoxifen treated breast cancer. *Sci Rep* 2020;10:1451.
11. Alfarsi LH, El Ansari R, Craze ML, et al. PPFIA1 expression associates with poor response to endocrine treatment in luminal breast cancer. *BMC Cancer* 2020;20:425.
12. Johnston SR. Enhancing Endocrine Therapy for Hormone Receptor-Positive Advanced Breast Cancer: Cotargeting Signaling Pathways. *J Natl Cancer Inst* 2015;107:djv212.
13. Perez EA. Treatment strategies for advanced hormone receptor-positive and human epidermal growth factor 2-negative breast cancer: the role of treatment order. *Drug Resist Updat* 2016;24:13-22.
14. Finn RS, Crown JP, Lang I, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. *Lancet Oncol* 2015;16:25-35.
15. McShane TM, Wolfe TA, Ryan JC. Updates on managing advanced breast cancer with palbociclib combination therapy. *Ther Adv Med Oncol* 2018;10:1758835918793849.
16. Behjati S, Frank MH. The effects of tamoxifen on immunity. *Curr Med Chem* 2009;16:3076-80.
17. Li B, Li Y, Wang XY, et al. Profile of differentially expressed intratumoral cytokines to predict the immunopolarizing side effects of tamoxifen in breast cancer treatment. *Am J Cancer Res* 2015;5:726-37.
18. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139-40.
19. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44-57.
20. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009;37:1-13.
21. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag, 2016.

22. He X, Zhang J. Why do hubs tend to be essential in protein networks? *PLoS Genet* 2006;2:e88.
23. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47:D607-13.
24. Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014;8 Suppl 4:S11.
25. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401-4.
26. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013;6:pl1.
27. Nagy Á, Lániczky A, Menyhárt O, et al. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep* 2018;8:9227.
28. Miao YR, Zhang Q, Lei Q, et al. ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. *Adv Sci (Weinh)* 2020;7:1902880.
29. Li T, Fan J, Wang B, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res* 2017;77:e108-10.
30. Li B, Severson E, Pignion JC, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol* 2016;17:174.
31. Brechbuhl HM, Finlay-Schultz J, Yamamoto TM, et al. Fibroblast Subtypes Regulate Responsiveness of Luminal Breast Cancer to Estrogen. *Clin Cancer Res* 2017;23:1710-21.
32. Loi S, Haibe-Kains B, Desmedt C, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics* 2008;9:239.
33. Hyakudomi M, Matsubara T, Hyakudomi R, et al. Increased expression of fractalkine is correlated with a better prognosis and an increased number of both CD8+ T cells and natural killer cells in gastric adenocarcinoma. *Ann Surg Oncol* 2008;15:1775-82.
34. Andre F, Cabioglu N, Assi H, et al. Expression of chemokine receptors predicts the site of metastatic relapse in patients with axillary node positive primary breast cancer. *Ann Oncol* 2006;17:945-51.
35. Hatori K, Nagai A, Heisel R, et al. Fractalkine and fractalkine receptors in human neurons and glial cells. *J Neurosci Res* 2002;69:418-26.
36. Liang Y, Yi L, Liu P, et al. CX3CL1 involves in breast cancer metastasizing to the spine via the Src/FAK signaling pathway. *J Cancer* 2018;9:3603-12.
37. Guo J, Zhang M, Wang B, et al. Fractalkine transgene induces T-cell-dependent antitumor immunity through chemoattraction and activation of dendritic cells. *Int J Cancer* 2003;103:212-20.
38. Lavergne E, Combadière B, Bonduelle O, et al. Fractalkine mediates natural killer-dependent antitumor responses in vivo. *Cancer Res* 2003;63:7468-74.
39. Xin H, Kikuchi T, Andarini S, et al. Antitumor immune response by CX3CL1 fractalkine gene transfer depends on both NK and T cells. *Eur J Immunol* 2005;35:1371-80.
40. Nukiwa M, Andarini S, Zaini J, et al. Dendritic cells modified to express fractalkine/CX3CL1 in the treatment of preexisting tumors. *Eur J Immunol* 2006;36:1019-27.
41. Park MH, Lee JS, Yoon JH. High expression of CX3CL1 by tumor cells correlates with a good prognosis and increased tumor-infiltrating CD8+ T cells, natural killer cells, and dendritic cells in breast carcinoma. *J Surg Oncol* 2012;106:386-92.
42. Ohta M, Tanaka F, Yamaguchi H, et al. The high expression of Fractalkine results in a better prognosis for colorectal cancer patients. *Int J Oncol* 2005;26:41-7.
43. Korbecki J, Simińska D, Kojder K, et al. Fractalkine/CX3CL1 in Neoplastic Processes. *Int J Mol Sci* 2020;21:3723.
44. Chen MC, Baskaran R, Lee NH, et al. CXCL2/CXCR2 axis induces cancer stem cell characteristics in CPT-11-resistant LoVo colon cancer cells via Gai-2 and Gαq/11. *J Cell Physiol* 2019;234:11822-34.
45. Hsu YL, Chen YJ, Chang WA, et al. Interaction between Tumor-Associated Dendritic Cells and Colon Cancer Cells Contributes to Tumor Progression via CXCL1. *Int J Mol Sci* 2018;19:2427.
46. Bièche I, Chavey C, Andrieu C, et al. CXC chemokines located in the 4q21 region are up-regulated in breast cancer. *Endocr Relat Cancer* 2007;14:1039-52.
47. Acosta JC, Gil J. A role for CXCR2 in senescence, but what about in cancer? *Cancer Res* 2009;69:2167-70.
48. Acosta JC, O'Loughlin A, Banito A, et al. Control of senescence by CXCR2 and its ligands. *Cell Cycle* 2008;7:2956-9.

49. Pincini A, Tornillo G, Orso F, et al. Identification of p130Cas/ErbB2-dependent invasive signatures in transformed mammary epithelial cells. *Cell Cycle* 2013;12:2409-22.
50. Zeng L, Ouyang D, Zou Q, et al. Bioinformatics-based analysis of TCGA to identify the prognostic biomarkers in HR-positive/HER2-negative tamoxifen-treated breast cancer. *Research Square* 2019.

Cite this article as: Cao Z, Jin Z, Zeng L, He H, Chen Q, Zou Q, Ouyang D, Luo N, Zhang Y, Yuan Y, Yi W. Prognostic and tumor-immune infiltration cell signatures in tamoxifen-resistant breast cancers. *Gland Surg* 2021;10(9):2766-2779. doi: 10.21037/gs-21-566

Table S1 PCR primers

Gene	Primer sequence
C3	Forward: 5'-CGGCCTTTGTTCTCATCTCG-3'
	Reverse: 5'-CAGAGCATAGCCAGCAATGG-3'
PF4	Forward: 5'-CAGTGCCTGTGTGTAAGAC-3'
	Reverse: 5'-CCTTCCATTCTTCAGCGTGG-3'
SAA1	Forward: 5'-CTTGGCGAGGCTTTTGATGG-3'
	Reverse: 5'-GGCATCGCTGATCACTTCTG-3'
CXCL1	Forward: 5'-GTCCGTGGCCACTGAACT-3'
	Reverse: 5'-CTATGACTTCGGTTTGGGCG-3'
CX3CL1	Forward: 5'-GAGCTTTTCCGAGTGCCTC-3'
	Reverse: 5'-GCCTTCAGACGGAGCATTCT-3'
CXCL13	Forward: 5'-CCTCTCTCCAGTCCAAGGTG-3'
	Reverse: 5'-TGAGGGTCCACACACACAAT-3'
GAL	Forward: 5'-AACCAGGAAGCTTTGACAGG-3'
	Reverse: 5'-CGCTCGATGTCTTCTGAGGA-3'
GPER1	Forward: 5'-TCACCTGGATGAGCTTCGAC-3'
	Reverse: 5'-GGACATCCGCGAAACAGAAG-3'
CXCL2	Forward: 5'-CCCCTGGCCACTGAACTG-3'
	Reverse: 5'-CTATGACTTCGGTTTGGGCG-3'
CXCL6	Forward: 5'-CTGGTCCTGTCTCTGCTGTG-3'
	Reverse: 5'-CTTGCTTCCCGTTCTTCAGG-3'
GAPDH	Forward: 5'-TGGTATCGTGAAG GACTCA-3'
	Reverse: 5'-CCAGTAGAGGCAGG GATGAT-3'

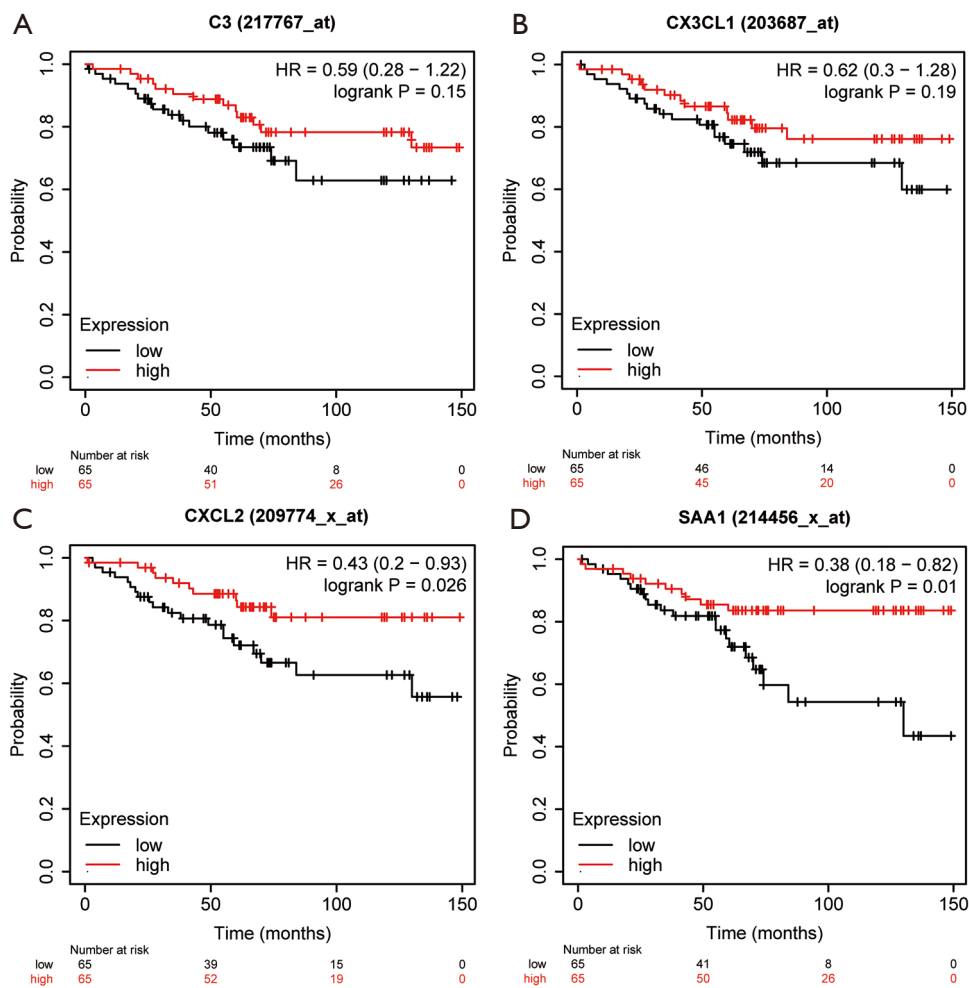


Figure S1 Validation of survival analysis of *C3*, *CX3CL1*, *CXCL2*, and *SAA1* on Kaplan-Meier Plotter website. (A-D) The overall survival value for the expression of the *C3*, *CX3CL1*, *CXCL2*, *SAA1* on the Kaplan-Meier Plotter with the condition of ER+ and only tamoxifen-treated breast cancers in the endocrine therapy.

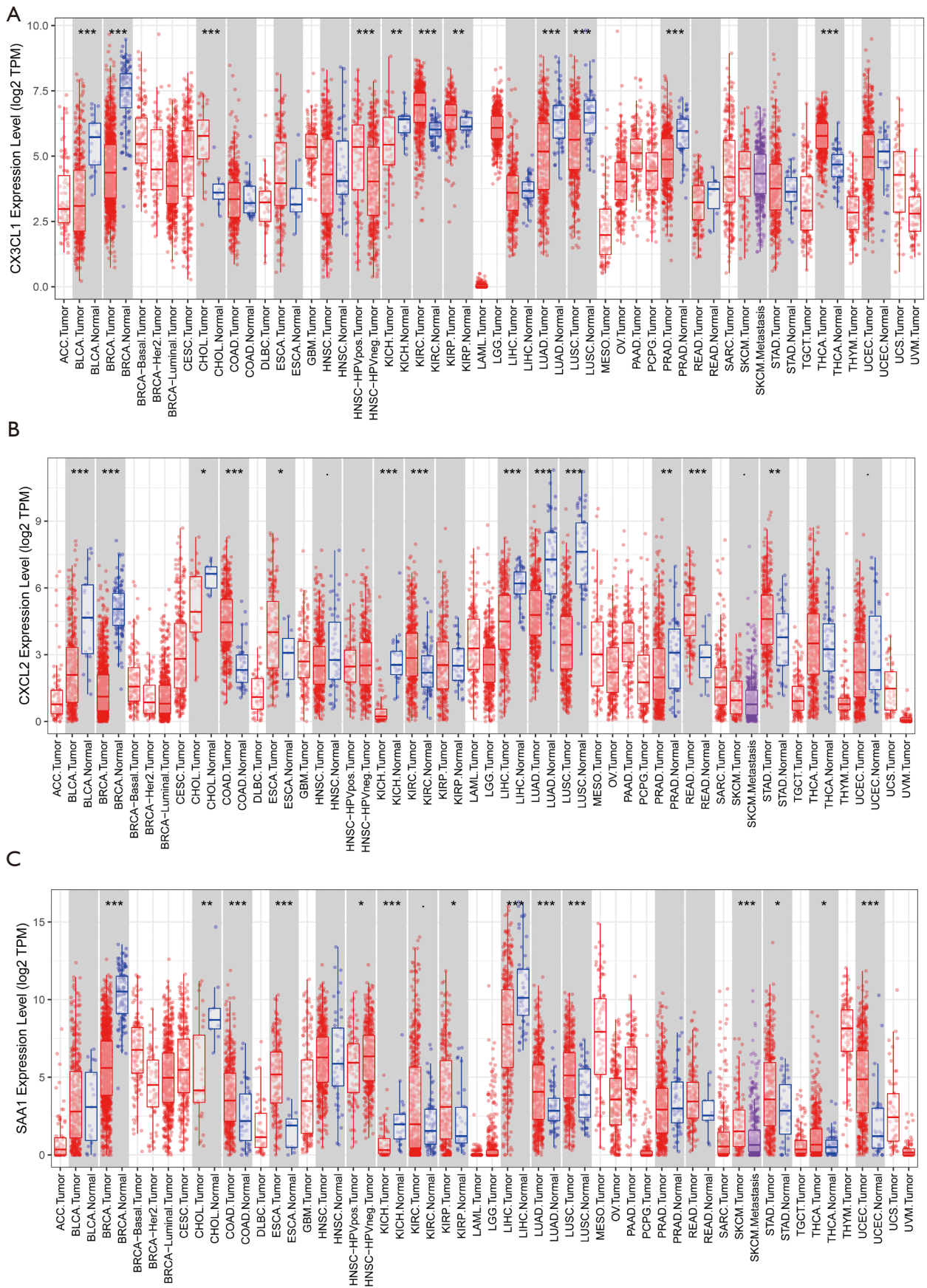


Figure S2 *CX3CL2*, *CXCL1*, and *SAA1* expression in different cancers. (A-C) *CX3CL1*, *CXCL1*, and *SAA1* expression levels in BRCA tumor tissue were lower than those in BRCA normal tissue in the TIMER database. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.