Bioinformatics-based analysis of TCGA to identify the prognostic biomarkers in HR-positive/HER2-negative tamoxifen-treated breast cancer

Liyun Zeng

Second Xiangya Hospital

Dengjie Ouyang

Second Xiangya Hospital

Qiongyan Zou

Second Xiangya Hospital

Qitong Chen

Second Xiangya Hospital

Na Luo

Second Xiangya Hospital

Wenjun Yi (✉ yiwenjun@csu.edu.cn)

<https://orcid.org/0000-0002-1563-4925>

Research article

Keywords: breast cancer; tamoxifen; TCGA

Posted Date: September 12th, 2019

DOI: <https://doi.org/10.21203/rs.2.14316/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background Patients with hormone receptor-positive (HR+) breast cancer, which accounts for approximately 70% of all breast cancers, receive tamoxifen therapy to inhibit recurrence and improve overall survival. However, the cumulative risk of distant recurrence in the past 20 years is still as high as 22 to 52% after 5 years of endotherapy. TNM stage, histological grade and age are important clinical factors related to recurrence. Differential gene analysis based on matching clinical factors will be helpful for understanding the molecular mechanism of recurrence; however, there have been no reports yet. **Results** In our study, we identified 647 DEGs in a TCGA dataset that included 20 patients, of which 10 patients had relapsed after tamoxifen treatment and 10 did not. Ten genes were found by the CytoHubba APP in Cytoscape. After analysis of the GO terms and KEGG pathways, we found that the genes were mainly enriched in the inflammatory response and cytokine-receptor interaction, and then these results were verified in the GEPIA2 and KM-plot websites. Finally, we identified CXCL1, CXCL2 and CX3CL1 as prognostic factors, and their overexpression in HR-positive/HER2-negative breast cancer predicted a longer disease-free survival. **Conclusion** In conclusion, the high expression of CXCL1, CXCL2, and CX3CL1 can predict longer disease-free survival in breast cancer after tamoxifen treatment and may be a biomarker for tamoxifen therapy.

Background

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 (HR)-positive (HR+)/HER2-negative, HR+/HER2-positive, HR-negative/HER2-positive and HR-negative/HER2-negative breast cancers[1]. Over 2/3 of patients are estrogen receptor (ER)-positive or progesterone receptor-positive and thus benefit from endocrine therapy, including the selective ER antagonist tamoxifen and aromatase inhibitors, which inhibit the estrogen product. Patients with HR+ breast cancer, which accounts for approximately 70% of all breast cancers, benefit from tamoxifen therapy in terms of the inhibition of breast cancer recurrence and the improvement of overall survival[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 one-third 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 to 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], in turn altering the levels of transcriptional coregulators, upregulating signaling factors such as NF- κ B[9, 10], FGFR3[11], NOTCH[12], c-Myc, and cyclin D1[13-15], and enhancing the IL-6/STAT3 signaling pathway[16] and other signaling pathways. The ER signaling pathway is also regulated by membrane receptor tyrosine kinases, including epidermal growth factor receptor (EGFR), HER2, and insulin-like growth factor receptor (IGF1R)[17]. These membrane kinases activate signaling pathways that eventually result in the phosphorylation of ER as well as its coactivators and corepressors at multiple sites to influence their specific functions[18]. Evidence suggests that the overexpression of HER2 confers tamoxifen resistance in ER+ breast cancer through the MAPK signaling pathway[18]. Palbociclib[19], an inhibitor of CDK4/6, has been approved by the FDA as the first-line treatment for drug resistance in advanced endocrine therapy; however, unfortunately, its overall response rate (ORR) was only 42.5% and clinical benefit rate (CBR) was 84.5%, with some patients still progressing in the Paloma-2 trial[20]. The overexpression of NCOR1, a

corepressor of ER, is associated with enhanced sensitivity to tamoxifen[21]. Therefore, searching for other direct mechanisms of drug sensitivity in endocrine therapy is of great clinical and scientific value.

Chemokines are small molecular weight proteins (~8 to 10 kDa) that are released in response to insult; they play double roles of recruiting neutrophils from the vasculature to tissue and then activating neutrophils in tissue for microbial killing and/or initiating tissue repair[22]. There are 7 neutrophil-activating chemokines (NACs) expressed in humans—CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL7, and CXCL8. NACs exist as monomers and dimers and exert their functions by activating the CXCR1 and CXCR2 receptors. The role of chemokines in tumors has been controversial, and proinflammatory cytokines such as CXCL1, CXCL2, and CXCL3 (GRO- α , β , and γ) contribute to multiple protumorigenic activities and are involved in the occurrence, development, and metastasis of tumors. Studies have demonstrated the role of CXCL1 and 2 in treatment resistance and metastasis; CXCL1 and 2 are overexpressed in breast cancer cells and prime cells for survival in metastatic sites[23]. In contrast, several groups have reported a tumor-suppressive role for some of the factors secreted by senescent cells. One group showed that the proinflammatory cytokines and chemokines secreted by senescent cells also regulate senescent growth arrest and trigger an innate immune response that results in the clearance of senescent lesions relying on p53[24, 25]. This group also showed that the levels of CXCR2 are upregulated in preneoplastic lesions, and evidence suggests that its expression might subsequently decrease during cancer progression, as seems to be the case in prostate cancer[25].

Here, we aimed to find gene biomarkers for overexpressed HR patients and then assess the sensitivity to tamoxifen therapy, which can help evaluate the prognosis of patients who receive tamoxifen treatment and provide guidance for HR+/HER2-negative patients. TNM stage, histological grade and age are important clinical factors related to recurrence. Differential gene analysis based on clinical factor matching is helpful for understanding the molecular mechanism of recurrence, but there have been no reports yet. Herein, we screened The Cancer Genome Atlas (TCGA) RNA-seq datasets, including 1097 breast cancer samples, and matched 10 pairs of patients who were all HR+/HER2-negative who received tamoxifen treatment according to their age, clinical stage and histological grade to find differentially expressed genes (DEGs). We then analyzed the enriched Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms of 647 DEGs. We found that the DEGs are mainly enriched in the inflammatory response and cytokine–cytokine receptor interaction, indicating that the mechanism of recurrence after tamoxifen therapy was associated with fractalkine in the tumor microenvironment. Finally, after validation in the GEPIA2 and KM-plot websites, the expression of CXCL1, CXCL2 and CX3CL1 was shown to predict the prognosis of breast cancer patients who were HR+/HER2-negative who received tamoxifen treatment. In addition, the overexpression of these chemokines was associated with longer disease-free survival. In brief, our results showed that the expression of CXCL1, CXCL2, and CX3CL1 could be used to assess the prognosis of HR+/HER2-negative breast cancer patients who received tamoxifen treatment and provided clues for the mechanism of relapse after taking tamoxifen. However, the findings we provided are limited, and the deeper underlying mechanism of relapse after tamoxifen still needs to be further explored.

Methods

TCGA analysis of the expression of DEGs in breast cancer patients

Processed and deidentified patient breast cancer (BRCA) data (TCGA “level 3” data designation) were downloaded from TCGA (2016) (http://gdac.broadinstitute.org/runs/stddata__2016_01_28/data/BRCA/20160128/). Ten pairs of HR+/her2-negative invasive ductal breast cancers, including 10 recurrence and 10 nonrecurrence samples after tamoxifen

therapy, were matched according to the patient age and pathological characteristics. The clinical characteristics of those 20 patients are presented in Table 1. Differential expression analysis of two groups was performed using the DEG-seq R package. The p-value was adjusted using the q value. A q value <0.05 and $|\log_2FC| > 1$ was set as the threshold for significantly differential expression.

GO and KEGG analysis of DEGs

The above DEGs were all entered in the DAVID website (<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 bubble diagram was illustrated with the ggplot2 R package and the Hmisc R package.

PPI network analysis

The PPI 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 STRING website (https://string-db.org/cgi/input.pl?sessionId=L1vQl87ngile&input_page_show_search=on) to explore the association between these DEGs. Cytoscape was used to visualize the PPI network from the STRING website, find the hub genes (with the CytoHubba APP) and beautify the below diagrams.

Validation of the expression of the top 10 hub genes

After selecting the top 10 hub genes, we validated their expression in the GEPIA2 website (<http://gepia2.cancer-pku.cn/#index>), which provides multiple functions, including allowing users to compare their own data with TCGA and GTEx data, comparing the gene expression in different subtypes of cancers and predicting the patient survival.

Survival analysis

Only 10 pairs of patient data were selected from TCGA. The clinical data could not be used to analyze patient survival. Thus, we performed RFS analysis on the KM plotter website (<http://kmplot.com/analysis/index.php?p=background>), which is dedicated to constructing survival curves for breast, liver, lung, ovarian, and gastric cancer, and the data used were based on TCGA and GEO. Survival curves with $p < 0.05$ were regarded as statistically significant.

Results

1. Identification of DEGs between the relapsed and nonrelapsed groups after tamoxifen treatment in HR+/HER2-negative breast cancer patients To identify gene expression related to tamoxifen resistance, we matched 10 pairs of breast cancer samples from the TCGA database of patients who received tamoxifen therapy according to their age and pathological characteristics (Table 1); all of these patients relapsed in 5 years. Figure 1a shows the process of screening for suitable patients. In total, we identified 647 DEGs between recurrent and nonrecurrent patients with the selection criteria of q value < 0.05 and $|\log_2(\text{Fold Change (FC)})| > 1$. Among the DEGs, 506 genes were downregulated, and 141 were upregulated. Figure 1b shows a volcano plot to examine the differences in the expression level of the genes in the two groups of samples and the statistical significance of these differences. 2. GO and KEGG pathway analysis of the DEGs The DEGs were annotated in the GO and KEGG databases. GO enrichment analysis of the DEGs was implemented using the Database for Annotation, Visualization and Integrated

Discovery (DAVID) website. Figure 2a-c shows the results of the DEGs annotated in the GO database. In the category of biological processes, the DEGs were mainly enriched in epidermal development, cytoskeletal organization, cell-cell signaling, BMP signaling, etc. In the cellular component category, the DEGs were mainly enriched in the extracellular space, extracellular region, integral component of the plasma membrane, plasma membrane, etc. 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 inflammatory response. The KEGG pathway analysis (Figure 2d) shows that 50 genes were distributed in metabolic pathways, 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 estrogen metabolism or tamoxifen metabolism, needs further investigation.

3. Protein-protein interaction network construction and module analysis To analyze the genes associated with tamoxifen resistance, the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) website was used to investigate the protein-protein interaction (PPI) network associated with the ER pathway. The PPI network was drawn using Cytoscape, which is an open-source bioinformatics software platform for visualizing molecular interaction networks, and the most significant module in the PPI network was identified by using the cytoHubba plug-in of Cytoscape, which is an APP for clustering a given network based on a topology to find densely connected regions. Figure 3b shows the top ten hub genes—C3, PF4, SAA1, CXCL1, CX3CL1, CXCL13, GAL, GPER1, CXCL2, and CXCL—most of which are inflammatory factors, which implies that tamoxifen resistance may be associated with the inflammatory response, and the differentiated expression of these ten genes was shown in Table 2. A previous study has proven that mediators secreted by cancer-associated fibroblasts (CAFs)[26], in addition to components of the inflammatory response[27] such as cytokines and growth factors, exert an important role in drug resistance.

4. Validation of the expression of the top 10 hub genes Considering that there were few samples from TCGA, we validated the expression of the top 10 hub genes in a different subgroup of breast cancer patients on the GEPIA2 website (<http://gepia2.cancer-pku.cn/#index>). As shown in Figure 4, the expression of SAA1, CX3CL1, CXCL1, CXCL2, and GPER1 was higher in the adjacent normal breast tissue than in the luminal A subgroup of breast cancer, while the expression of CXCL13 was lower, and the differential expression was statistically significant. The expression of C3, PF4, GAL, and CXCL6 was not significant between the para-cancerous normal breast tissue and luminal A breast cancer tissue, but its general trend was in line with our results, especially for C3, PF4, SAA1, CXCL1, CX3CL1, CXCL13, GAL, GPER1, CXCL2, and CXCL6, most of which are inflammatory factors, thus indicating that tamoxifen resistance may be associated with the inflammatory response. The previous study has proven that mediators secreted by CAFs[26], in addition to components of the inflammatory response[27] such as cytokines and growth factors, exert an important role in drug resistance.

5. Survival analysis Survival analysis was performed to explore the correlation between the 10 DEGs and patients' recurrence-free survival (RFS) with the KM-plot website (<http://kmplot.com/analysis/>). Figure 5(a-j) shows that the increased expression of C3, CX3CL1, CXCL1, CXCL2, CXCL6, GPER1, and PF4 was closely related to improved RFS, while this result did not show that the other three genes (CXCL13, GAL, and SAA1) were associated with RFS in luminal A breast cancer.

Discussion

The development of recurrence after tamoxifen treatment is a serious challenge in breast cancer therapy. Thus, addressing this problem and determining the underlying mechanism are top priorities. Our purpose was to identify the DEGs associated with resistance after tamoxifen therapy and discover genes that can serve as biomarkers of resistance after tamoxifen treatment for assessing prognosis. First, we used the TCGA RNA-seq dataset and identified 647 DEGs. Then, we analyzed the GO and KEGG pathways on the DAVID website. According to the results,

the DEGs were mainly enriched in epidemic development, cytoskeleton organization, cell-cell signaling, and inflammatory responses. In addition, the top 10 hub genes were mainly enriched in cytokine-cytokine receptor interactions, such as C3, CXCL1, CX3CL1, CXCL13, CXCL2, and CXCL6. In our study, the log2FC of these genes was negative in the comparison of nonrelapse and relapse patients; in other words, these genes had low expression in relapse patients and high expression in nonrelapse patients. Hence, we checked the expression of these genes in GEPIA2, and the expression of SAA1, CX3CL1, CXCL1, CXCL2, and GPER1 coincided with our data from TCGA, while the other genes were not statistically significant between para-cancerous tissues and luminal A breast cancer tissues, though their expression trend was broadly consistent with our data. Then, we verified the prognosis of all top 10 hub genes on the KM plotter website (Figure 5) and found that seven of the 10 genes (C3, CX3CL1, CXCL1, CXCL2, CXCL6, GPER, and PF4) were closely related to the patient RFS; specifically, the overexpression of these genes could predict a long RFS, while the other 3 genes were not associated with the prognosis of luminal A breast cancer patients (p-value >0.05). Combining the expression and prognosis results, only three genes (CX3CL1, CXCL1, and CXCL2) were consistent with our results from TCGA. CXCL1 and CXCL2 are proinflammatory cytokines, whose receptor is CXCR2, that have been proven to be involved in the acquisition of many malignant features, such as cancer cell proliferation, invasion, epithelial-mesenchymal transition (EMT)[28], migration[29] and chemoresistance in many tumors, including those of ER α -positive breast cancer patients treated with tamoxifen[30]. 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 proved 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[25, 31]. Moreover, the chemokine CXCL1 dimer is a potent agonist for the CXCR2 receptor[32]. CX3CL1, also a controversial inflammatory factor, is a CX3C chemokine that can chemoattract T lymphocytes[33]. 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 that appears to recruit cytotoxic T cells and natural killer (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[34]. We hypothesized that the profile of CX3CL1 expression in breast carcinoma can also predict a good prognosis. Though 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[35, 36]. Researchers found that CX3CR1, a specific receptor for CX3CL1, was associated with metastasis to the brain. Because the human brain constitutively expresses CX3CL1, it could attract CX3CR1-expressing breast cancer cells. Current research validated that the high expression of CX3CL1 in tumor cells correlates with a good prognosis in breast cancer. Researchers proposed that the expression of CX3CL1 in tumor cells enhances the recruitment of CD⁺ T cells, dendritic cells (DCs), and NK cells, thus resulting in a better prognosis in breast carcinoma[37].

In conclusion, the present study was designed to identify DEGs that may be involved in recurrence after tamoxifen treatment for HR+/HER2-negative breast cancer. CXCL1, CXCL2, and CX3CL1 may be regarded as diagnostic biomarkers for tamoxifen therapy, and their high expression predicts a good prognosis. However, further studies are needed to elucidate the biological function of these genes in breast cancer.

Conclusions

In conclusion, our study identified biomarkers to assess patient prognosis. Our results show that CXCL1, CXCL2, and CX3CL1 may be regarded as diagnostic biomarkers for tamoxifen therapy and that their high expression predicts a good prognosis. Our study may have some limitations considering the small sample size.

Abbreviations

TCGA: The Cancer Genome Atlas

GO: Gene ontology

DEGs: Differentially Expressed Genes

KEGG: Kyoto Encyclopedia of Genes and Genomes

HR: hormone receptor

HER2: Human epidermal growth factor receptor -2

CXCL1: C-X-C Motif Chemokine Ligand 1

CXCL2: C-X-C Motif Chemokine Ligand 1

CX3CL1: C-X3-C Motif Chemokine Ligand 1

Declarations

Acknowledgement

Not applicable

Funding

This article supported by National Natural Science Foundation of China (Grant No.81873640).

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author information

Affiliations:Department of Breast and Thyroid, the Second Xiangya Hospital, Central South University, *Changsha, Hunan, China*.

Liyun Zeng¹, Dengjie Ouyang, Qiongyan Zou, Qitong Chen, Na Luo & Wenjun Yi*

Contributions: Study design: Wenjun Yi, Qiongyan Zou, Liyun Zeng. Data analysis: Wenjun Yi, Liyun Zeng, Dengjie Ouyang, Qitong Chen. Image editor: Liyun Zeng, Dengjie Ouyang, Na Luo. Paper Writing: Wenjun Yi, Liyun Zeng, Dengjie Ouyang, Qiongyan Zou, Qitong Chen, Na Luo.

Corresponding author: Wenjun Yi*

Department of Breast and Thyroid, the Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China.

Ethics declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA *et al*: **Molecular portraits of human breast tumours**. *Nature* 2000, **406**(6797):747-752.
2. Lobbezoo DJ, van Kampen RJ, Voogd AC, Dercksen MW, van den Berkmortel F, Smilde TJ, van de Wouw AJ, Peters FP, van Riel JM, Peters NA *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**(3):507-514.
3. Rao RD, Cobleigh MA: **Adjuvant endocrine therapy for breast cancer**. *Oncology (Williston Park, NY)* 2012, **26**(6):541-547, 550, 552 passim.
4. **Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials**. *Lancet (London, England)* 2015, **386**(10001):1341-1352.
5. Pan H, Gray R, Braybrooke J, Davies C, Taylor C, McGale P, Peto R, Pritchard KI, Bergh J, Dowsett M *et al*: **20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years**. *The New England journal of medicine* 2017, **377**(19):1836-1846.
6. Robinson DR, Wu YM, Vats P, Su F, Lonigro RJ, Cao X, Kalyana-Sundaram S, Wang R, Ning Y, Hodges L *et al*: **Activating ESR1 mutations in hormone-resistant metastatic breast cancer**. *Nat Genet* 2013, **45**(12):1446-1451.
7. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S, King TA *et al*: **ESR1 ligand-binding domain mutations in hormone-resistant breast cancer**. *Nat Genet* 2013, **45**(12):1439-1445.
8. Miller TW, Hennessy BT, González-Angulo AM, Fox EM, Mills GB, Chen H, Higham C, García-Echeverría C, Shyr Y, Arteaga CL: **Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in**

- estrogen receptor-positive human breast cancer.** *The Journal of clinical investigation* 2010, **120**(7):2406-2413.
9. Oida K, Matsuda A, Jung K, Xia Y, Jang H, Amagai Y, Ahn G, Nishikawa S, Ishizaka S, Jensen-Jarolim E *et al*: **Nuclear factor-kB plays a critical role in both intrinsic and acquired resistance against endocrine therapy in human breast cancer cells.** *Scientific reports* 2014, **4**:4057.
 10. Nakshatri H, Bhat-Nakshatri P, Martin DA, Goulet RJ, Jr., Sledge GW, Jr.: **Constitutive activation of NF-kappaB during progression of breast cancer to hormone-independent growth.** *Molecular and cellular biology* 1997, **17**(7):3629-3639.
 11. Tomlinson DC, Knowles MA, Speirs V: **Mechanisms of FGFR3 actions in endocrine resistant breast cancer.** *International journal of cancer* 2012, **130**(12):2857-2866.
 12. Magnani L, Stoeck A, Zhang X, Lanczky A, Mirabella AC, Wang TL, Gyorffy B, Lupien M: **Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer.** *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(16):E1490-1499.
 13. Cheng R, Liu YJ, Cui JW, Yang M, Liu XL, Li P, Wang Z, Zhu LZ, Lu SY, Zou L *et al*: **Aspirin regulation of c-myc and cyclinD1 proteins to overcome tamoxifen resistance in estrogen receptor-positive breast cancer cells.** *Oncotarget* 2017, **8**(18):30252-30264.
 14. Green AR, Aleskandarany MA, Agarwal D, Elsheikh S, Nolan CC, Diez-Rodriguez M, Macmillan RD, Ball GR, Caldas C, Madhusudan S *et al*: **MYC functions are specific in biological subtypes of breast cancer and confers resistance to endocrine therapy in luminal tumours.** *British journal of cancer* 2016, **114**(8):917-928.
 15. Ishii Y, Waxman S, Germain D: **Tamoxifen stimulates the growth of cyclin D1-overexpressing breast cancer cells by promoting the activation of signal transducer and activator of transcription 3.** *Cancer research* 2008, **68**(3):852-860.
 16. Kitajima S, Yoshida A, Kohno S, Li F, Suzuki S, Nagatani N, Nishimoto Y, Sasaki N, Muranaka H, Wan Y *et al*: **The RB-IL-6 axis controls self-renewal and endocrine therapy resistance by fine-tuning mitochondrial activity.** *Oncogene* 2017, **36**(36):5145-5157.
 17. Schiff R, Massarweh SA, Shou J, Bharwani L, Mohsin SK, Osborne CK: **Cross-talk between estrogen receptor and growth factor pathways as a molecular target for overcoming endocrine resistance.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004, **10**(1 Pt 2):331s-336s.
 18. Shou J, Massarweh S, Osborne CK, Wakeling AE, Ali S, Weiss H, Schiff R: **Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer.** *Journal of the National Cancer Institute* 2004, **96**(12):926-935.
 19. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO, Ettl J, Patel R, Pinter T, Schmidt M *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.** *The Lancet Oncology* 2015, **16**(1):25-35.
 20. McShane TM, Wolfe TA, Ryan JC: **Updates on managing advanced breast cancer with palbociclib combination therapy.** *Therapeutic advances in medical oncology* 2018, **10**:1758835918793849.
 21. Lu R, Hu X, Zhou J, Sun J, Zhu AZ, Xu X, Zheng H, Gao X, Wang X, Jin H *et al*: **COPS5 amplification and overexpression confers tamoxifen-resistance in ERalpha-positive breast cancer by degradation of NCoR.** *Nature communications* 2016, **7**:12044.
 22. Kobayashi Y: **The role of chemokines in neutrophil biology.** *Frontiers in bioscience : a journal and virtual library* 2008, **13**:2400-2407.

23. Acharyya S, Oskarsson T, Vanharanta S, Malladi S, Kim J, Morris PG, Manova-Todorova K, Leversha M, Hogg N, Seshan VE *et al*: **A CXCL1 paracrine network links cancer chemoresistance and metastasis.** *Cell* 2012, **150**(1):165-178.
24. Xue W, Zender L, Miething C, Dickins RA, Hernando E, Krizhanovsky V, Cordon-Cardo C, Lowe SW: **Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas.** *Nature* 2007, **445**(7128):656-660.
25. Acosta JC, Gil J: **A role for CXCR2 in senescence, but what about in cancer?** *Cancer research* 2009, **69**(6):2167-2170.
26. Brechbuhl HM, Finlay-Schultz J, Yamamoto TM, Gillen AE, Cittelly DM, Tan AC, Sams SB, Pillai MM, Elias AD, Robinson WA *et al*: **Fibroblast Subtypes Regulate Responsiveness of Luminal Breast Cancer to Estrogen.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2017, **23**(7):1710-1721.
27. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemant F, Tutt AM, Gillet C, Ellis P, Ryder K, Reid JF *et al*: **Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen.** *BMC genomics* 2008, **9**:239.
28. Chen MC, Baskaran R, Lee NH, Hsu HH, Ho TJ, Tu CC, Lin YM, Viswanadha VP, Kuo WW, Huang CY: **CXCL2/CXCR2 axis induces cancer stem cell characteristics in CPT-11-resistant LoVo colon cancer cells via Galphai-2 and Galphaq/11.** *Journal of cellular physiology* 2019, **234**(7):11822-11834.
29. Hsu YL, Chen YJ, Chang WA, Jian SF, Fan HL, Wang JY, Kuo PL: **Interaction between Tumor-Associated Dendritic Cells and Colon Cancer Cells Contributes to Tumor Progression via CXCL1.** *International journal of molecular sciences* 2018, **19**(8).
30. Bieche I, Chavey C, Andrieu C, Busson M, Vacher S, Le Corre L, Guinebretiere JM, Burlinchon S, Lidereau R, Lazennec G: **CXC chemokines located in the 4q21 region are up-regulated in breast cancer.** *Endocrine-related cancer* 2007, **14**(4):1039-1052.
31. Acosta JC, O'Loughlen A, Banito A, Raguz S, Gil J: **Control of senescence by CXCR2 and its ligands.** *Cell cycle (Georgetown, Tex)* 2008, **7**(19):2956-2959.
32. Ravindran A, Sawant KV, Sarmiento J, Navarro J, Rajarathnam K: **Chemokine CXCL1 dimer is a potent agonist for the CXCR2 receptor.** *J Biol Chem* 2013, **288**(17):12244-12252.
33. Imai T, Hieshima K, Haskell C, Baba M, Nagira M, Nishimura M, Kakizaki M, Takagi S, Nomiyama H, Schall TJ *et al*: **Identification and molecular characterization of fractalkine receptor CX3CR1, which mediates both leukocyte migration and adhesion.** *Cell* 1997, **91**(4):521-530.
34. Ohta M, Tanaka F, Yamaguchi H, Sadanaga N, Inoue H, Mori M: **The high expression of Fractalkine results in a better prognosis for colorectal cancer patients.** *Int J Oncol* 2005, **26**(1):41-47.
35. Andre F, Cabioglu N, Assi H, Sabourin JC, Delaloge S, Sahin A, Broglio K, Spano JP, Combadiere C, Bucana C *et al*: **Expression of chemokine receptors predicts the site of metastatic relapse in patients with axillary node positive primary breast cancer.** *Annals of oncology : official journal of the European Society for Medical Oncology* 2006, **17**(6):945-951.
36. Hatori K, Nagai A, Heisel R, Ryu JK, Kim SU: **Fractalkine and fractalkine receptors in human neurons and glial cells.** *Journal of neuroscience research* 2002, **69**(3):418-426.
37. 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.** *Journal of surgical oncology* 2012, **106**(4):386-392.

Tables

Table1. the two groups clinical information from TCGA websites.

Recurrence					Nonrecurrence			
Sample	age(years)	menstrual_status	histologic grading	TMN	age(years)	menstrual_status	histologic grading	TMN
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

Table 2. The expression of top hub 10 genes

Gene	Log ₂ FC	FDR
CXCL1	-5.49791	1.96E-05
CX3CL1	-2.45057	0.000102
CXCL6	-5.39252	0.000457
SAA1	-3.97267	0.000892
CXCL2	-3.22999	0.017
CXCL13	-3.28201	0.046553
GPER1	2.674776	0.017648
GAL	-2.81032	0.048692
C3	-1.57181	0.019901
PF4	-3.91016	0.035945

Figures

Figure 1

Identification of DEGs between the relapsed and nonrelapsed groups after tamoxifen treatment in HR-positive/her2-negative patients. (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 $\log_2FC > 1$ and $FDR < 0.05$. Green nodes represent DEGs with $\log_2FC < -1$ and $FDR < 0.05$.

Figure 2

GO and KEGG pathway analysis of DEGs. (a) The 25 biological processes (BPs). (b) The 17 cellular components (CCs). (c) The 6 molecular function (MFs). (d) The 15 KEGG pathways. The depth of the color reflects the value of the $-\log_{10}(p\text{-value})$, red presents the largest $-\log_{10}(p\text{-value})$, and the size of the dot reflects the number of genes.

a

b

Figure 3

Analysis of the protein-protein interaction network. PPI network: DEGs of the TCGA dataset were integrated using the STRING website and modified in Cytoscape.

Figure 4

Validation of the expression of the top 10 hub genes. This image shows the expression of the 10 hub genes (a. C3, b. CX3CL1, c. CXCL1, d. CXCL2, e. CXCL6, f. CXCL13, g. GAL, h. GPER, i. PF4, and j. SAA1) between cancer and para-cancerous tissues in four different subtypes of breast cancer; these images were downloaded from GEPIA2, with data based on TCGA.

Figure 5

Survival analysis. The Kaplan–Meier curves of the 10 hub genes were downloaded from the KM plotter website. (a-j: C3, CX3CL1, CXCL1, CXCL2, CXCL6, CXCL13, GAL, GPER, PF4, and SAA1).