Pathology & Oncology Research (2020) 26:443–452
https://doi.org/10.1007/s12253-018-0478-5

ORIGINAL ARTICLE

The Overexpression of CD80 and ISG15 Are Associated with the Progression and Metastasis of Breast Cancer by a Meta-Analysis Integrating Three Microarray Datasets

Yuanhang Li[1] · Weijun Bai[1] · Linlin Zhang[1]

Received: 8 May 2018 / Accepted: 4 October 2018 / Published online: 9 November 2018

Abstract

Breast cancer is a common cancer and could result in a substantial mortality. The study aimed to screen gene signatures associated with the development and metastasis of breast cancer and explore their regulation mechanisms. Three datasets of GSE10797, GSE8977 and GSE3744 were downloaded from GEO (Gene Expression Omnibus) database, containing 55 breast cancer samples and 27 normal samples. After data preprocessing using limma software and RMA (robust multi-array average) algorithm, DEGs (differentially expressed genes) between breast tumor and normal tissues in three individual experiments were identified using MADAM package. Function and pathway enrichment analyses were performed for the DEGs. Transcription factors and TAGs (tumor associated genes) among the DEGs were recognized and the PPI (protein-protein-interaction) network for the DEGs was constructed using Cytoscape software. The mRNA expression was analyzed via real-time quantitative PCR and protein expression was measured by western blotting. Totally, 100 DEGs were identified, including 33 up-regulated genes and 67 down-regulated genes. Among them, up-regulated DEGs such as *CD80* was enriched in toll-like receptor (TLR) interaction pathway and the TAG, *ISG15* was related to RIG-I-like receptor signaling pathway, while *CXCL10* was involved in both of the two pathways. Whereas, the down-regulated DEG, *CXCL12* was significantly associated with axon guidance pathway. Additionally, these DEGs were also pivotal nodes in the PPI network with high degrees. Besides, *CXCL10* and *CD80* were both interacted with *IFNG*. The mRNA expression of *ISG15* was obviously enhanced in human breast cancer cells MCF-7, while no significant difference of *CXCL10* mRNA level was found between MCF10A and MCF-7 cells. Moreover, the proteins expression levels of CD80 and ISG15 were significantly increased in MCF-7, MDA-MB-468 and MDA-MB-231 breast cancer cells than in normal MCF10A cells. *CD80* might be responsible for the breast cancer's progression and metastasis via regulating innate immune system. In addition, *ISG15* is identified as a crucial gene signature associated with breast cancer development and metastasis via RIG-I-like receptor signaling pathway.

Keywords Microarray data · Meta-analysis · Breast cancer · Metastasis

✉ Yuanhang Li
yuanhangli619@hotmail.com

Weijun Bai
baiweijun008@126.com

Linlin Zhang
79579909@qq.com

[1] Medical Department, Cancer Hospital of China Medical University, No. 44 Xiaoheyan Road, Dadong District, Shenyang 110042, Liaoning Province, China

Background

Breast cancer is the most frequent cancer in women in developed countries [34]. Annually, approximately 1.2 million women are reported to suffer from breast cancer and 50,000 of them die from this disease [5]. Moreover, it is estimated that about 232,670 new cases would occur among women in 2014 in the United States [32]. Breast cancer is occasionally considered as a metastatic disease that spreads beyond the original organ. However, the ability of a tumor to metastasize is complicated and organ-specific [33]. Despite the advances in the operative and non-operative

managements, breast cancer metastasis remains a major clinical issue that affects numerous patients [1]. Besides, prognosis and survival rates for breast cancer vary greatly with cancer types, stages, treatments, and geographical locations of patient [30]. The accumulated gene alterations were implicated in the progression of breast cancer. Although considerable gene signatures such as *CD44* and *HER2* have been identified to be associated with tumorigenesis of breast cancer [24, 28], the comprehensive regulation mechanisms in the development and metastasis of breast cancer remain rudimentary.

In the present study, we performed a meta-analysis integrating three datasets of GSE10797 [6], GSE8977 [20] and GSE3744 [27], which were deposited in the public GEO (Gene Expression Omnibus) database to identify DEGs (differentially expressed genes) between breast cancer tissues and normal tissues. Additionally, the biological functions and pathways of these DEGs were also performed using GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analyses, thereby to uncover the underlying regulation mechanisms of these DEGs including their interaction relationships in the development and further metastasis of breast cancer.

Methods

Microarray Data Resources

Three datasets from gene expression profiles which were deposited in GEO database were downloaded, with the accession numbers of GSE10797 [6], GSE8977 [20] and GSE3744 [27]. The first dataset was based on GPL571 [HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array (Affymetrix Inc., Santa Clara, California, USA) platform, while the platform of the other two datasets was both GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix Inc., Santa Clara, California, USA). Combined with the three profiles, 55 samples from breast tumor tissues and 27 samples from normal tissues were collected. The detailed information of the platforms and samples were listed as Table 1. All subjects in GSE10797 were obtained informed consent meeting all federal, state and institutional guidelines. The studies of GSE8977 were performed under the supervision of MIT's Division of Comparative Medicine and were approved by the Institutional Animal Care and Use Committee. GSE3744 was conducted based on the monitor of Partners Hospital and Dana-Farber Cancer Institute Institutional Review Boards, and patient consent is obtained for all identified specimens.

Preprocessing of the Raw Gene Expression Data

The preprocessing of raw data in the gene expression was performed using limma software in affy Bioconductor R package (available at http://www.bioconductor.org/packages/release/bioc/html/limma.html) [15]. The extraction of probe level data, background correction and quantile normalization were conducted for samples from each expression profiles utilizing RMA (robust multi-array average) algorithm [18]. After obtaining the gene expression matrix, aggregate function were used to calculate the average value of gene expression.

Screening of DEGs and Meta-Analysis

The DEGs between breast tumor tissues and normal tissues in each individual experiment were identified using Bayes moderated t-test method based on limma package [15], with the threshold criteria of FDR (false discovery rate) < 0.05 and $|\log_2 FC \text{ (fold change)}| \geq 0.585$.

Thereafter, the MADAM package [23] was recruited to perform meta-analysis for the DEGs (P value <0.1) identified in three individual experiments. The statistical magnitude S was obtained based on the P value matrix of DEGs in each individual experiment. The formula is $S = -2 \sum \log p$, in which S obeys chi Square distribution with the degree of freedom of 2 k and the P value is calculated using the function of fisher.method. The genes with adjusted P value <0.01 was defined as DEGs. Meanwhile, the random statistical magnitude S_{random} [26] and the corresponding P value were also calculated. The P value was deemed as the ratio of S_{random} value ≥ the true S value and the times for calculating the statistical magnitude was set as B = 5000. The P value was calculated using the function of fisher.method.perm and genes with P value <0.05 were considered as DEGs. The method permutates the P values as reported by the single studies and calculates a new S with these random combinations. Then, the DEGs both fulfilled the threshold criteria under fisher.method and fisher.method.perm functions were identified as the final DEGs.

Function and Pathway Enrichment Analyses

GO [3] enrichment analysis was conducted for the identified DEGs to explore their functions mainly involves in biological process (BP), molecular function (MF) and cellular component (CC). Moreover, KEGG [19] pathway analysis was also conducted for the DEGs. The DAVID (Database for Annotation, Visualization and Integrated Discovery, http://david.abcc.ncifcrf.gov/) online software [12] was used and the threshold for significant GO terms and pathways were P values <0.01, based on the hypergeometric distribution method.

Table 1 Microarray data resources

Study	Platform	Samples from breast tumor tissues	Samples from normal breast tissues
Casey T, et al. [6]	GPL571 [HG-U133A_2]	8	5
Karnoub AE, et al. [20]	GPL570 [HG-U133_Plus_2]	7	15
Richardson AL, et al. [27]	GPL570 [HG-U133_Plus_2]	40	7

Functional Annotation

Combined with the information in the TRANSFAC database, the transcription factors (TFs) among these DEGs were selected. Besides, we also combined information from databases including TSGene (Tumor Suppressor gene) [35] and TAG (Tumor Associated Genes) [9] to detect all the oncogenes and tumor suppressor genes among the DEGs.

Construction of the Protein-Protein-Interaction (PPI) Network and Sub-Network

The DEGs were mapped in STRING database (http://string-db.org/) [14] to recognize their potential interaction relationships from protein level. The data were downloaded at May 9th, 2014. The selection criterion for the interaction pairs in the PPI network were combined score ≥ 0.9. It is required there was at least one DEG in the well-defined PPI interaction pairs. Only the interaction pairs which validated by experiment, text mining and co-expression analysis, as well as other recorded protein interactions, were utilized to establish modules. Finally, the PPI network was visualized by Cytoscape software (http://cytosca-pe.org/) [22]. The lines represented the corresponding interaction relationships between the nodes. Meanwhile, the nodes possessing a mass of connections were defined as hub nodes, which were selected based on the scale-free property of the network.

Besides, the enrichment of TFs and the target genes were performed and the regulatory network between DEGs and the related TFs was constructed, with the threshold criterion of FDR < 0.05.

Cell Culture

Human breast cancer cells (MCF-7, MDA-MB-468 and MDA-MB-231; 3 samples) were provided by Shanghai Iyunbio Co., Ltd., (Shanghai, China) and were maintained in medium containing 90% MEM (Gibco, Carlsbad, CA, USA) + 10% FBS (Gibco) + 1% PS (Gibco) at 37 °C in a humidified chamber supplemented with 5% CO_2. Normal breast cells (MCF10A, 3 samples) were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China) and were cultured with medium provided by the Chinese Academy of Sciences; this medium included 100 ml MEGM kit (Lonza/Clonetics, CC-3150) and cholera toxin (Sigma, C8052) at a concentration of 10 ng/ml at 37 °C in a humidified chamber supplemented with 5% CO_2.

Real-Time Quantitative PCR (RT-PCR)

Total RNA from the MCF-7 and MCF10A was extracted by TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA). Then, the mRNA reverse transcription kit (RR036A-1, TAKARA, Shiga, Japan) was used to synthesize first-strand cDNA. RT-PCR was performed using Power SYBR Green PCR Master (4,367,659; Thermo, USA) on an ABI 7500 FAST real-time PCR system kit (Applied Biosystems, Foster City, CA, USA). The procedures for the qPCR were as follows: 50 °C for 3 min, 95 °C for 3 min, 95 °C for 10 s, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s. The specificity of the PCR amplicons was examined by melting curve analysis. The comparative Ct method was employed for quantifying target mRNA expression, which was normalized to GAPDH. The primers used in this study were shown in Table 2.

Western Blotting

The MCF-7, MDA-MB-468, MDA-MB-231 and MCF10A cells were lysed by radio-immunoprecipitation assay (RIRA) lysis buffer (P0013B, Beyotime, China) containing 1 mM PMSF (ST506, Beyotime, China) and the lysates were centrifuged at 12,000 g for 10 min at 4 °C. Then, the proteins concentrations were detected by Bicinchoninic Acid (BCA)

Table 2 Primers of each gene

Name	Primers(5′-3′)
CXCL10-F	CATTCTGATTTGCTGCCTTATC
CXCL10-R	TGATGGCCTTCGATTCTGG
CD80-F	CTCCCATCCTGGGCCATTAC
CD80-R	CTCTCTGCATCTTGGGGCAA
ISG15-F	AGCAGCTCCATGTCGGTGT
ISG15-R	AGGTTCGTCGCATTTGTCC
CXCL12-F	ACACTCCAAACTGTGCCCTT
CXCL12-R	AGTGGGTCTAGCGGAAAGTC
GAPDH-F	TGACAACTTTGGTATCGTGGAAGG
GAPDH-R	AGGCAGGGATGATGTTCTGG AGAG

method with a BCA protein assay kit (PL212989, Thermo, USA). Following, total proteins with equal quality (20 μg/lane) from each sample were separated by 12% polyacrylamide gels and separated proteins were transferred into polyvinylidene fluoride membranes. The membranes were blocked with 5% skimmed milk diluted in phosphate buffer solution (PBS) at 37 °C for 2 h. After washing the membranes with 1 × PBS-T(1000 mL 1 × PBS + 1 mL Tween-20) (3 times), the membranes were incubated with the primary antibodies against CD80 (66406–1-Ig, 1:1000, Proteintech), ISG15 (15981–1-AP, 1:1000, Proteintech) or β-actin (Sc-47,778, 1:1000, Santa Cruz Biotechnology, INC) at 4 °C overnight. Subsequently, membranes were washed with TBST and incubated with M-lgG-kappa-BP-HRP (Sc-516,142, 1:1000, Santa Cruz Biotechnology, INC) or Mouse anti-rabbit lgG-HRP (Sc-2357, 1:1000, Santa Cruz Biotechnology, INC) secondary antibodies at 37 °C for 2 h. Immunoreactive bands were visualized by enhanced chemiluminescence (ECL) system.

Statistical Analysis

All data are expressed as mean ± SEM, and the differences between MCF-7 and MCF10A cells were compared using Student's *t* test. The statistical analyses were performed using SPSS 22.0 (SPSS Inc., Chicago, IL, USA), and GraphPad Prism 5 (GraphPad Software, San Diego, CA, USA) was used to map the results. A p value of <0.05 was considered to represent a significant difference.

Results

The Identified DEGs between Breast Tumor and Normal Tissues

The DEGs based on the cut-off criteria of FDR < 0.05 and $|\log_2 FC| \geq 0.585$ were identified in each individual experiment (Table 3). Then the genes with P value <0.05 were selected to perform meta-analysis, involving a total of 2895 down-regulated genes and 3888 up-regulated genes. Among them, 2 up-regulated genes and 15 down-regulated genes were detected in all three individual experiments, and two experiments (from Karnoub AL, et al. and Richardson AL, et al.) shared 300 up-regulated genes and 458 down-regulated genes. Thereafter, the DEGs (FDR < 0.1) in at least two individual experiments were selected to perform meta-analysis. Combining with the P values calculated by two kinds of functions, a total of 100 DEGs were screened out, in which 33 were up-regulated, while 67 were down-regulated.

Table 3 Differentially expressed genes identified in each individual experiment

	Casey T, et al.	Karnoub AE, et al.	Richardson AL, et al.
Up	5	580	935
Down	29	558	1217
Sum	34	1138	2152

Biological Functions and Pathways of the Identified DEGs

As presented in Table 4, the KEGG pathway analysis indicated that the down-regulated genes were mainly enriched in tumorigenesis and proliferation associated pathways (such as Focal adhesion, Hematopoietic cell lineage, mTOR signaling pathway and Melanoma) and the metabolism-related pathways (such as Fatty acid metabolism, Drug metabolism-cytochrome P450, Retinol metabolism, Glycolysis/Gluconeogenesis and Metabolism of xenobiotics by cytochrome P450), as well as pathways relevant to other diseases (including Acute myeloid leukemia and Melanoma). By contrast, the up-regulated DEGs were significantly enriched in ECM-receptor interaction pathway, Phagosome and Focal adhesion pathways, as well as signaling pathways (p53 signaling pathway, RIG-I-like receptor signaling pathway and toll-like receptor (TLR) signaling pathway), which is greatly correlated with carcinogenesis.

GO enrichment analysis revealed that the down-regulated DEGs were remarkably involved in Sugar Alcohols metabolism functions (such as glycosaminoglycan metabolic process, aminoglycan metabolic process and primary alcohol metabolic process) and functions related to cell movement (including cellular component movement, locomotion, cell migration and cell motility), as well as tumor spread and metastasis associated functions (such as vasculature development, anatomical structure morphogenesis and localization of cell). On the other hand, the up-regulated DEGs were mainly enriched in functions correlated with cell cycle, cell division, nuclear division and cellular component organization, which closely associated with tumor spread (Table 5).

Transcription Factors and Tumor Associated Genes of the Identified DEGs

Integrating the information in the relevant databases, the TFs and TAGs were recognized among these DEGs. As revealed in Table 6, there were 3 TFs including *NR3C2*, *EGR2* and *HLF* which were down-regulated and 16 TAGs, in which 10 were down-regulated and 6 were up-regulated.

Table 4 KEGG Pathway enrichment analysis of the DEGs

	KEGG Pathway	Counts	*P* value	geneSymbol
Down	ABC transporters	3	0.0009059	*ABCA8*, *ABCA9*, *ABCA6*
	Focal adhesion	5	0.0016013	*FIGF*, *PDGFD*, *COL6A6*, *RELN*, *IGF1*
	Axon guidance	4	0.002287	*SEMA5A*, *SEMA3G*, *SLIT3*, *CXCL12*
	Hematopoietic cell lineage	3	0.0065977	*KIT*, *ANPEP*, *CD34*
	Tyrosine metabolism	2	0.0139146	*ADH1C*, *ADH1B*
	Aldosterone-regulated sodium reabsorption	2	0.0145709	*IGF1*, *NR3C2*
	Fatty acid metabolism	2	0.0152403	*ADH1C*, *ADH1B*
	mTOR signaling pathway	2	0.0218416	*FIGF*, *IGF1*
	Acute myeloid leukemia	2	0.0259367	*RUNX1T1*, *KIT*
	Retinol metabolism	2	0.0321499	*ADH1C*, *ADH1B*
	Glycolysis / Gluconeogenesis	2	0.0330814	*ADH1C*, *ADH1B*
	Metabolism of xenobiotics by cytochrome P450	2	0.0388899	*ADH1C*, *ADH1B*
	Melanoma	2	0.0388899	*PDGFD*, *IGF1*
	Drug metabolism - cytochrome P450	2	0.0409074	*ADH1C*, *ADH1B*
Up	p53 signaling pathway	2	0.0056043	*RRM2*, *CCNB2*
	RIG-I-like receptor signaling pathway	2	0.0060969	*CXCL10*, *ISG15*
	ECM-receptor interaction	2	0.008648	*COL11A1*, *COMP*
	Toll-like receptor signaling pathway	2	0.0122862	*CD80*, *CXCL10*
	Phagosome	2	0.0264768	*OLR1*, *COMP*
	Focal adhesion	2	0.0434133	*COL11A1*, *COMP*

KEGG: Kyoto Encyclopedia of Genes and Genomes; DEGs: differentially expressed genes; Down: down-regulated DEGs; Up: up-regulated DEGs; Counts: The numbers of the DEGs enriched in the corresponding pathways

PPI Modules and Pathway Analysis of the Target Genes of TFs

In the established PPI network (Fig. 1), the top ten nodes with high degrees were revealed such as *TAC1*, *IGF1*, *RRM2*, *TOP2A*, *TPX2*, *CXCL10*, *CCNB2*, *MELK*, *NUSAP1* and *DLGAP5*. Based on the interaction information of TFs and target genes in corresponding databases, the integrated regulatory network of DEGs and relevant TFs was constructed (Fig. 2), involving 72 DEGs and 9 TFs. Further KEGG pathway analysis of these DEGs was carried out. As the results indicated, these DEGs were enriched in cell adhesion relevant pathways including focal adhesion and cell adhesion molecules (CAMs) pathways and tumor spread-related pathways such as hematopoietic cell lineage pathway, as well as signal pathways including mTOR signaling pathway, p53 signaling pathway and RIG-I-like receptor signaling pathway and metabolism-associated pathways such as Aldosterone-regulated sodium reabsorption and Glutathione metabolism.

CXCL10 and ISG15 mRNA Expressions

The mRNA expression of *ISG15* was obviously enhanced in human breast cancer cells MCF-7, while no significant difference of *CXCL10* mRNA level was found between MCF10A and MCF-7 cells (Fig. 3).

CD80 and ISG15 Protein Expressions

The protein levels of CD80 and ISG15 were detected in MCF10A, MCF-7, MDA-MB-468 and MDA-MB-231 cells, respectively. The results demonstrated that the protein level of CD80 was significantly increased in MCF-7, MDA-MB-468 and MDA-MB-231 breast cancer cells compared with normal MCF10A cells (Fig. 4). The levels of CD80 protein were highest in MDA-MB-468, lowest in MDA-MB-231 and intermediate in MCF7 cells.

In addition, the protein expression of ISG15 was also obviously enhanced in MCF-7, MDA-MB-468 and MDA-MB-231 breast cancer cells than in normal MCF10A cells (Fig. 4). However, due to the expression abundance of ISG15 in MCF10A cells was low, which was consistent with above mRNA expression result, thus the immunoreactive band of ISG15 in MCF10A cells was not detectable by current method. Interesting, the levels of ISG15 protein were highest in MCF7, lowest in MDA-MB-468 and intermediate in MDA-MB-231 cells.

Discussion

Breast cancer is a leading cause for death among women worldwide. In the present study, based on the meta-analysis

Table 5 GO enrichment analysis of DEGs

	GO ID	Term	Counts	*P*-value
Down	GO:0030203	glycosaminoglycan metabolic process	6	2.85E-05
	GO:0006022	aminoglycan metabolic process	6	3.91E-05
	GO:0034308	primary alcohol metabolic process	3	5.84E-05
	GO:0006928	cellular component movement	16	6.98E-05
	GO:0040011	locomotion	15	8.40E-05
	GO:0016477	cell migration	12	1.08E-04
	GO:0001944	vasculature development	9	2.00E-04
	GO:0009653	anatomical structure morphogenesis	19	2.06E-04
	GO:0048870	cell motility	12	2.37E-04
	GO:0051674	localization of cell	12	2.37 E-04
Up	GO:0000278	mitotic cell cycle	12	2.43E-08
	GO:0007049	cell cycle	15	2.81E-08
	GO:0022402	cell cycle process	13	8.27E-08
	GO:0051301	cell division	8	4.36E-06
	GO:0000280	nuclear division	7	4.51E-06
	GO:0007067	mitosis	7	4.51E-06
	GO:0000819	sister chromatid segregation	4	4.63E-06
	GO:0044763	single-organism cellular process	30	5.59E-06
	GO:0016043	cellular component organization	21	6.92E-06
	GO:0048285	organelle fission	7	7.39E-06

GO: Gene Ontology; DEGs: differentially expressed genes; Down: down-regulated DEGs; Up: up-regulated DEGs; Counts: The numbers of the DEGs enriched in the corresponding pathways

of three datasets of gene expression profiles, a total of 33 up-regulated and 67 down-regulated DEGs were recognized between breast cancer tissues and normal cancer tissues. Among them, predominant up-regulated DEGs such as *CD80* was mainly enriched in TLR interaction pathway and the TAG, *ISG15*, was greatly related to RIG-I-like receptor signaling pathway, while *CXCL10* was both involved in these two pathways. By contrast, the down-regulated DEG of *CXCL12* was significantly associated with axon guidance pathway. In addition, these DEGs were also revealed as pivotal nodes in the PPI network with high degrees. Besides, *CXCL10* and *CD80* were both interacted with *IFNG*. The mRNA expression of *ISG15* was obviously enhanced in human breast cancer cells MCF-7, while that of *CXCL10* was no difference in MCF-7 and MCF10A cells. Moreover, the proteins expression levels of CD80 and ISG15 were significantly increased in MCF-7, MDA-MB-468 and MDA-MB-231 breast cancer cells than in normal MCF10A cells.

Chemokines are small, cytokine-like, secreted proteins and have been suggested to play critical roles in breast cancer progression. *CXCL* (chemokine (C-X-C motif) ligand) *12* and *CXCL14*, which could bind to receptors on breast tumor epithelial cells to promote proliferation, migration and invasion, are overexpressed in tumor myoepithelial cells and myofibroblasts, respectively [11, 25]. Moreover, *CXCL12*, the corresponding ligand of the chemokine receptor CXCR4, is detected to express in metastatic organs for breast cancer and direct the migration of breast cancer cells into specific metastatic sites in the chemotactic and invasive responses [16]. It has been proposed that the interaction of *CXCL12* and CXCR4 facilitate the mediation of axon guidance in the developing nervous system [2]. The pathway analysis in the

Table 6 TFs and TAGs of the identified DEGs

	TF	TAGs
Down	NR3C2, EGR2, HLF	oncogene: *KIT*, *RUNX1T1*, tumor suppressor: *LTSCR2*, *SRPX*, *TGFBR3*, *EGR2*, *AKAP12*, *GPC3*, *SPRY2*, *TCEAL7*
Up	/	oncogene: *SALL4*, *CEP55*, *WISP1* tumor suppressor: *ISG15*, *MMP11* others: *TACC3*

TF: Transcription factor; TAG: Tumor associated gene; DEGs: differentially expressed genes; Down: Down-regulated DEGs; Up: Up-regulated DEGs

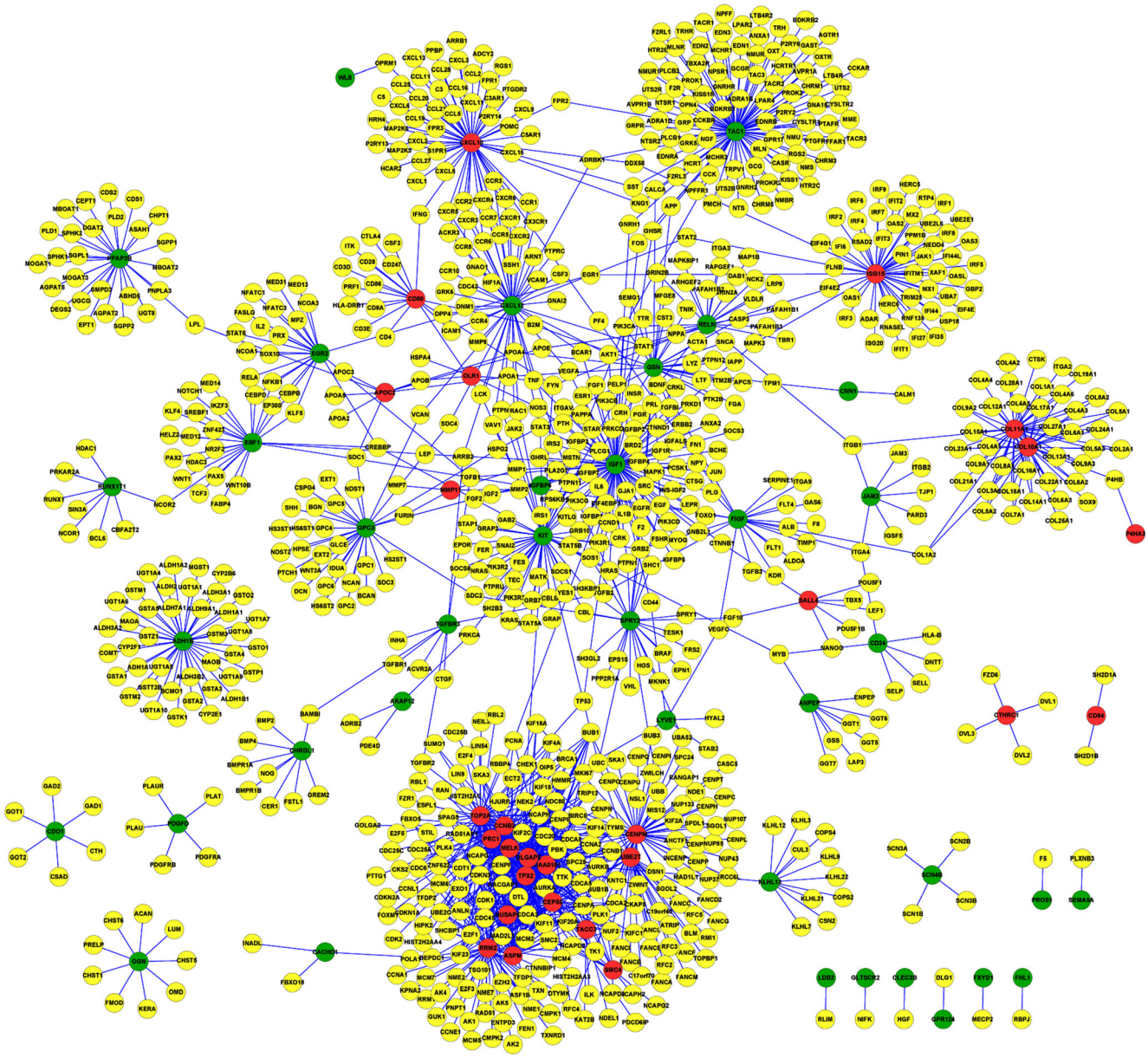

Fig. 1 PPI (protein-protein-interaction) network of DEGs (differentially expressed genes). The node in red represents up-regulated DEGs and the nodes in green represents down-regulated gene, while the node in yellow represents gene without differential expression. The line indicates the interaction relationship between the genes

current study also indicated the involvement of *CXCL12* in the axon guidance process, suggesting *CXCL12* played crucial roles in the regulation of axon guidance related pathways during the metastasis of breast cancer, and might also through the interaction with CXCR4. The protein encoded by *CXCL10* is a ligand for the receptor CXCR3, and the binding of the protein to CXCR3 has great potential to result in T cell migration [13]. Members of the TLR family are known to play key roles in both innate and adaptive immune responses and involve in the development of many cancer diseases. It is implicated that signaling cascade mediated by TLR3 and TLR4 may promote T cell infiltration in EAE (experimental autoimmune encephalitis). Meanwhile, the ligation of TLR3 and TLR4 by their cognate ligands has been detected to induce the secretion of *CXCL10*, which acts as a major T lymphocyte chemoattractant in the MS (multiple sclerosis) lesion [8], indicating that *CXCL10* was closely related to TLR signaling pathways, and it was consistent with our results of pathway enrichment analysis (Table 4). However, our finding revealed that in MCF-7 and MCF10A cells, *CXCL10* expression was not obviously different. Therefore, it might be inferred that TLR signaling-related pathways were not involved in the metastasis of breast cancer, but this result need more studies to verify it.

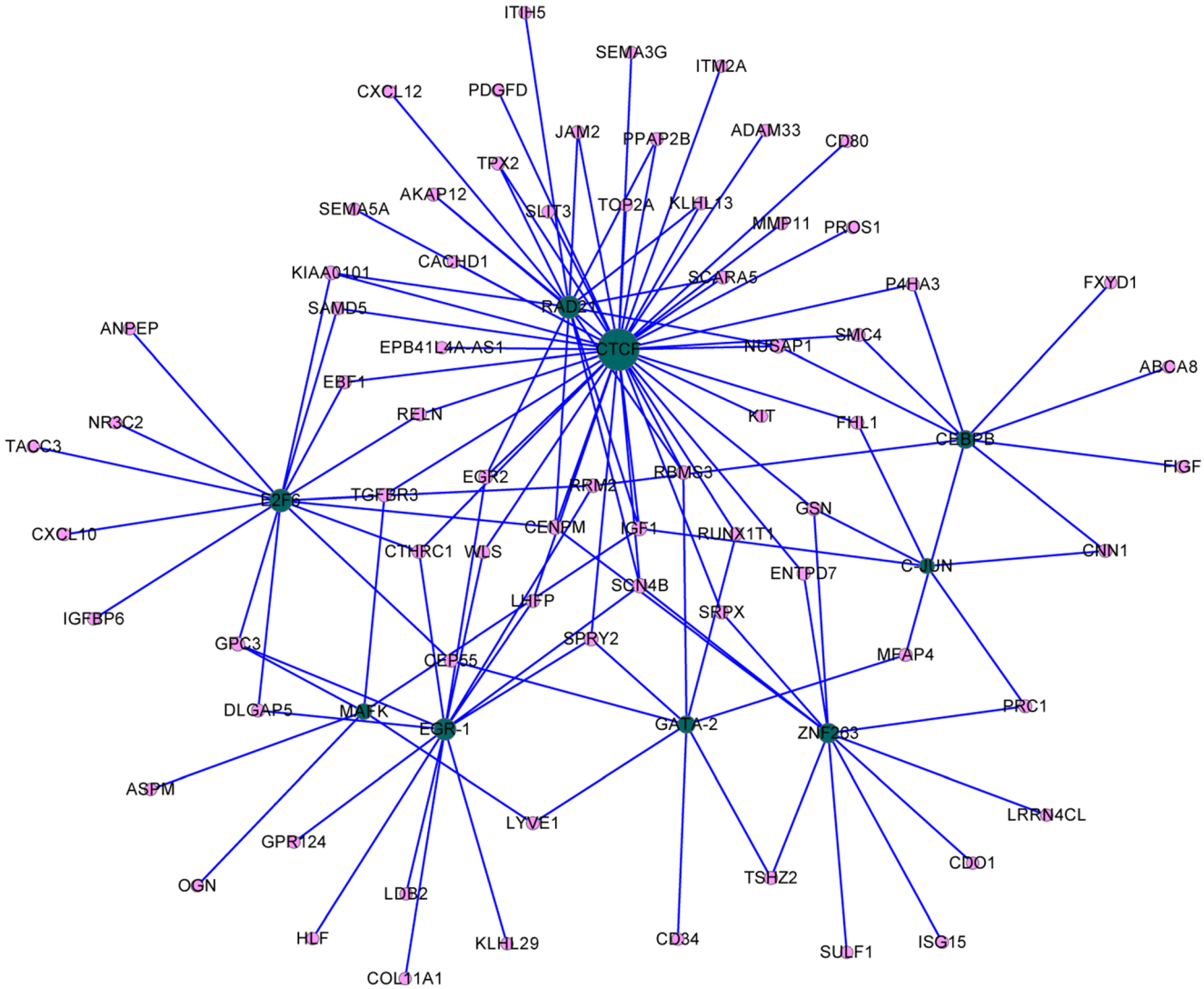

Fig. 2 Regulatory network of TF (transcription factor) and the target genes. Circle in pink denotes DEG (differentially expressed genes) and the blue denotes TF. The line indicates the interaction relationship between the genes

CD80 (Cluster of Differentiation 80, also known as B7–1) encodes a membrane receptor that is activated by the binding of CD28 or CTLA-4. This activated protein could induce T cell proliferation and cytokine production. TLRs could stimulate the innate immune system, which activates the adaptive immune system by the production of proinflammatory cytokines and the induction of key surface molecules such as *CD80* [21]. Additionally, the maturation of DC (dendritic

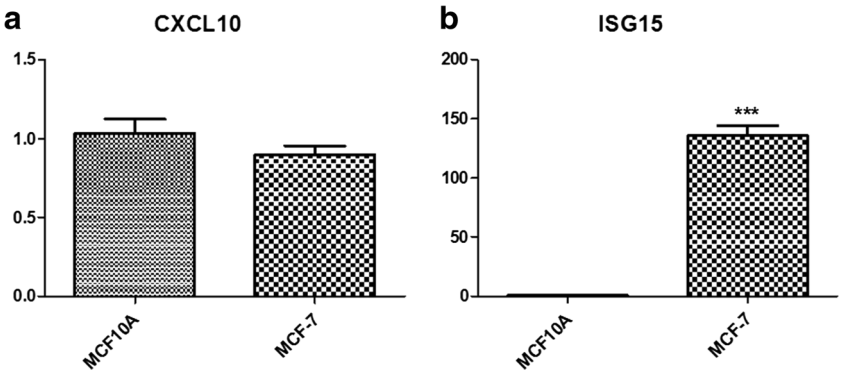

Fig. 3 mRNA expression levels of (a) CXCL10 and (b) ISG15 in normal breast cells (MCF-10A) and breast cancer cells (MCF-7) assessed using RT-qPCR. *** $p < 0.01$. RT-qPCR: real-time quantitative polymerase chain reaction

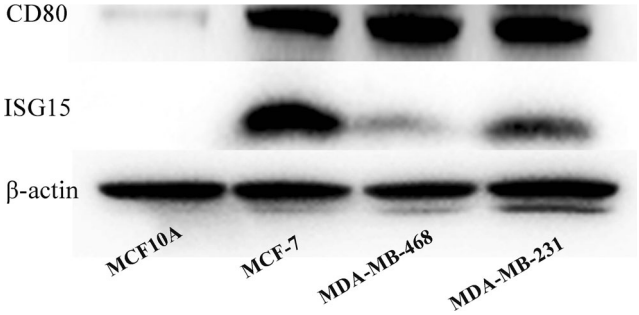

Fig. 4 Protein expression levels of (a) CD80 and (b) ISG15 in normal breast cells MCF-10A and breast cancer MCF-7, MDA-MB-468 and MDA-MB-231 cells measured by western blotting

cell), which plays vital roles in the immunological functions related to TLRs, involved a spectrum of events including up-regulation of costimulatory molecules such as *CD80* [31]. All these implied that *CD80* might also involve in the TLR signaling pathway, which was consistent with our pathway analysis. Notably, as presented in the PPI network (Fig. 1), *CD80* and *CXCL10* were both interacted with *IFNG* (interferon gamma, IFN-γ), which is a potent activator of macrophages and greatly related to immunoregulatory [7]. The inactivation of *CD80* caused a significant decrease in the production of interferon IFN-γ by T lymphocytes in nonatopic asthma patients [29], while the expression of *CXCL10* was induced by *IFNG* [17], which provided a hint that *CXCL10* might be in the downstream of *CD80*. In the present study, the proteins expression levels of CD80 were significantly increased in three breast cancer cells MCF-7, MDA-MB-468 and MDA-MB-231 than in normal MCF10A cells. Therefore, we suggested that *CD80* might be responsible for the breast cancer's development and metastasis via regulating innate immune system.

ISG15 was another critical DEG revealed in the PPI network and recognized as a TAG in our result of function annotation. The encoded protein of *ISG15* is an ubiquitin-like protein, which was associated with chemotactic activity towards neutrophils and direction of ligated target proteins to intermediate filaments. *ISG15* was overexpressed in breast carcinoma cells and proposed as a novel breast tumor marker with prognostic significance [4], in consistent with which, we also found the expression of *ISG15* was increased in human breast cancer cells. Furthermore, the overexpression of RIG-I (Retinoic acid-inducible gene-I) was detected to enhance the expression of *ISG15* [10], suggesting the involvement of *ISG15* in the RIG-I-like receptor signaling pathway, which also predicted in our pathway analysis. Thus, we speculated that *ISG15* mediated the development and metastasis of breast cancer via the involvement of RIG-I-like receptor signaling pathway.

Despite these profound findings, the limitations of the study should be discussed. Although the mRNAs expression levels of *ISG15*, *CXCL12* and *CXCL10*, and proteins expression levels of CD80 and ISG15 were analyzed, *CXCL12* mRNA expression and ISG15 protein expression in MCF10A cells were not detectable due to their low expression abundance in MCF10A.

Conclusions

In conclusion, *CD80* might be responsible for the breast cancer's development and metastasis via regulating innate immune system. *ISG15* might mediate the development and metastasis of breast cancer via the involvement of RIG-I-like receptor signaling pathway. Furthermore, *CXCL10* and *CXCL12* might be the crucial gene signatures remarkably associated with breast cancer development and metastasis. However, more experimental validations such as are needed for our results.

Authors' Contributions YL designed the research and provide the conception and drafted the manuscript. WB acquired and analyzed the data. WB and LZ interpreted the data. All authors revised and approved the final manuscript.

Compliance with Ethical Standards

Ethics Approval and Consent to Participate The studies of GSE8977 were performed under the supervision of MIT's Division of Comparative Medicine and were approved by the Institutional Animal Care and Use Committee. GSE3744 was conducted based on the monitor of Partners Hospital and Dana-Farber Cancer Institute Institutional Review Boards, and patient consent is obtained for all identified specimens.

Conflict of Interest The authors declare that they have no competing interests.

Abbreviations GEO, Gene Expression Omnibus; RMA, robust multi-array average; DEGs, differentially expressed genes; TAGs, tumor associated genes; PPI, protein-protein-interaction; TLR, toll-like receptor; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate; FC, fold change; BP, biological process; MF, molecular function\; CC, cellular component; DAVID, Database for Annotation, Visualization and Integrated Discovery; TFs, transcription factors; TSGene, Tumor Suppressor gene; RT-PCR, Real-time quantitative PCR; CAMs, cell adhesion molecules; EAE, experimental autoimmune encephalitis; MS, multiple sclerosis; IFNG, interferon gamma, IFN-γ; RIG-I, Retinoic acid-inducible gene-I; RIRA, radioimmunoprecipitation assay; BCA, Bicinchoninic Acid; PBS, phosphate buffer solution

References

1. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci 100:3983–3988
2. Arthur A, Shi S, Zannettino AC, Fujii N, Gronthos S, Koblar SA (2009) Implanted adult human dental pulp stem cells induce endogenous axon guidance. Stem Cells 27:2229–2237
3. Ashburner M et al (2000) Gene Ontology: tool for the unification of biology. Nat Genet 25:25–29

4. Bektas N et al (2008) The ubiquitin-like molecule interferon-stimulated gene 15 (ISG15) is a potential prognostic marker in human breast cancer. Breast Cancer Res 10:R58
5. Boyle P, Levin B (2008) World cancer report 2008. International Agency for Research on Cancer Lyon, France:330
6. Casey T et al (2009) Molecular signatures suggest a major role for stromal cells in development of invasive breast cancer. Breast Cancer Res Treat 114:47–62
7. Chen X, Armstrong MA, Li G (2006) Mesenchymal stem cells in immunoregulation. Immunol Cell Biol 84:413–421
8. Chen K, Huang J, Gong W, Iribarren P, Dunlop NM, Wang JM (2007) Toll-like receptors in inflammation, infection and cancer. Int Immunopharmacol 7:1271–1285
9. Chen J-S, Hung W-S, Chan H-H, Tsai S-J, Sun HS (2013) In silico identification of oncogenic potential of fyn-related kinase in hepatocellular carcinoma. Bioinformatics 29:420–427
10. Cui X-F, Imaizumi T, Yoshida H, Borden EC, Satoh K (2004) Retinoic acid-inducible gene-I is induced by interferon-γ and regulates the expression of interferon-γ stimulated gene 15 in MCF-7 cells. Biochem Cell Biol 82:401–405
11. Datta D et al (2006) Ras-induced modulation of CXCL10 and its receptor splice variant CXCR3-B in MDA-MB-435 and MCF-7 cells: relevance for the development of human breast cancer. Cancer Res 66:9509–9518
12. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA (2003) DAVID: database for annotation, visualization, and integrated discovery. Genome Biol 4:P3
13. Dufour JH, Dziejman M, Liu MT, Leung JH, Lane TE, Luster AD (2002) IFN-γ-inducible protein 10 (IP-10; CXCL10)-deficient mice reveal a role for IP-10 in effector T cell generation and trafficking. J Immunol 168:3195–3204
14. Franceschini A et al (2013) STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 41:D808–D815
15. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) Affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307–315
16. Goldberg-Bittman L, Neumark E, Sagi-Assif O, Azenshtein E, Meshel T, Witz IP, Ben-Baruch A (2004) The expression of the chemokine receptor CXCR3 and its ligand, CXCL10, in human breast adenocarcinoma cell lines. Immunol Lett 92:171–178
17. Hardaker EL et al (2004) Regulation of TNF-α-and IFN-γ-induced CXCL10 expression: participation of the airway smooth muscle in the pulmonary inflammatory response in chronic obstructive pulmonary disease. FASEB J 18:191–193
18. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249–264
19. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30
20. Karnoub AE et al (2007) Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. Nature 449:557–563
21. Kobayashi K, Hernandez LD, Galán JE, Janeway CA Jr, Medzhitov R, Flavell RA (2002) IRAK-M is a negative regulator of toll-like receptor signaling. Cell 110:191–202
22. Kohl M, Wiese S, Warscheid B (2011) Cytoscape: software for visualization and analysis of biological networks. In: Data Mining in Proteomics. Springer, pp 291–303
23. Kugler KG, Mueller LA, Graber A (2010) MADAM-an open source meta-analysis toolbox for R and bioconductor. Source Code Biol Med 5:3
24. Liu R et al (2007) The prognostic role of a gene signature from tumorigenic breast-cancer cells. N Engl J Med 356:217–226
25. Ma X-J, Dahiya S, Richardson E, Erlander M, Sgroi DC (2009) Gene expression profiling of the tumor microenvironment during breast cancer progression. Breast Cancer Res 11:R7
26. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM (2002) Meta-analysis of microarrays interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res 62:4427–4433
27. Richardson AL et al (2006) X chromosomal abnormalities in basal-like human breast cancer. Cancer Cell 9:121–132
28. Ross JS et al (2003) The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. Oncologist 8:307–325
29. Rutkowski R, Moniuszko T, Stasiak-Barmuta A, Kosztyla-Hojna B, Alifier M, Rutkowski K, Tatarczuk-Krawiel A (2003) CD80 and CD86 expression on LPS-stimulated monocytes and the effect of CD80 and CD86 blockade on IL-4 and IFN-gamma production in Nanotopic bronchial asthma. Arch Immunol Ther Exp 51:421–428
30. Sant M et al (2003) Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 106:416–422
31. Schreibelt G, Tel J, Sliepen KH, Benitez-Ribas D, Figdor CG, Adema GJ, de Vries IJM (2010) Toll-like receptor expression and function in human dendritic cell subsets: implications for dendritic cell-based anti-cancer immunotherapy. Cancer Immunol Immunother 59:1573–1582
32. Siegel R, Ma J, Zou Z, Jemal A (2014) Cancer statistics, 2014. CA Cancer J Clin 64:9–29
33. Smid M et al (2006) Genes associated with breast cancer metastatic to bone. J Clin Oncol 24:2261–2267
34. Turnbull C et al (2010) Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 42:504–507
35. Zhao M, Sun J, Zhao Z (2013) TSGene: a web resource for tumor suppressor genes. Nucleic Acids Res 41:D970–D976