Identification of Key Genes and Pathways Associated With Irradiation in Breast Cancer Tissue and Breast Cancer Cell Lines

Changchun Zhu, Chang Ge, Junbo He, Xueying Zhang, Guoxing Feng, and Saijun Fan

Abstract
Radiotherapy is mainly a traditional treatment for breast cancer; however, the key genes and pathways in breast cancer associated with irradiation are not clear. In this study, we aimed to explore the messenger RNA expression changes between preradiation and postradiation breast cancer. The gene expression data set (GSE59733) was downloaded from Gene Expression Omnibus database. According to |log₂FC (fold change)| ≥ 1 and with false discovery rate adjusted P value < .05, differentially expressed genes (DEGs) were screened and annotated by R programming software. The protein–protein interaction (PPI) network was conducted through STRING database, and subnetworks and hub genes were extracted by plug-in in Cytoscape. A total of 82 DEGs (74 upregulated and 8 downregulated genes) were identified. These DEGs mainly enriched in an intrinsic apoptotic signaling pathway and G-protein-coupled receptor binding. What’s more, tumor necrosis factor signaling pathway and interleukin 17 signaling pathway abnormally activated in postradiation tumor samples. Two characteristic subnetworks and 3 hub genes (FOS, CCL2, and CXCL12) were strongly distinguished in PPI network. Moreover, the expression level of the hub genes was confirmed in irradiated MCF-7 cell and SUM-159 cell using quantitative real-time polymerase chain reaction assay. These findings imply that these hub genes may play momentous function in breast cancer to irradiation.

Keywords
radiotherapy, breast cancer, GEO database, DEGs, bioinformatics

Introduction
Breast cancer is one of the most prevalent cancers in women worldwide. At present, radiotherapy has made remarkable achievements for solid tumor therapy, which especially acts as a positive role for most breast cancer. For example, postoperative radiotherapy can effectively reduce the mortality and recurrence of breast cancer. However, there is accumulating evidence suggesting that patients with breast cancer exhibit heterogeneous response to radiotherapy, seriously affecting the clinical efficacy and quality of life. Previous research studies have shown that there are beneficial response to radiotherapy in patients with breast cancer characterized by Rec-positive or HER2-negative diagnosis. Of note, the classification standard based on subtype is unable to fully describe clinically significant heterogeneity to radiotherapy among breast tumors. Therefore, it remains an urgent clinical problem to explore the difference in response to radiotherapy in breast cancer, which may effectively avoid overtreatment and undertreatment.

Recent studies suggested that abnormal gene expression was the principal contributor to cell behavior response to radiation. For example, Tob1 markedly mediated radiosensitivity in breast cancer cell MDA-MB-231 by targeting the JNK and p38 pathways. Radiation promotes the migratory and invasive behavior in lung cancer cell A549 through transforming growth factor β-mediated epithelial–mesenchymal transition.
Furthermore, radiation induced the upregulated expression of Sox2 and Oct3/4, which is responsible for spherogenesis in cancer cells. However, a recent study identified that vascular endothelial growth factor A and other factors secreted from irradiated MCF-7 cells contribute to angiogenic responses in endothelial cells. However, the exact molecular mechanism of cell behavior response to radiation regrettably hasn’t been well researched at present. Thus, we boldly speculated that there might be some undiscovered gene that plays a key role in mediating the breast cancer cell behavior to radiotherapy.

Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) is the largest and most comprehensive database of public gene expression, providing a large amount of high throughput data for gene research. It has been making a research hot spot to screen gene expression profiling to satisfy personalized breast cancer treatment for several years. For example, Pajic et al screened 11 potential microRNAs associated with local relapse following radiotherapy by gene expression profile and further confirmed that miR-139-5p may be an effective therapeutic target involved in DNA damage response to radiotherapy in breast cancer. We screened the data set related to radiotherapy for breast cancer at GEO. The GSE59733 was selected for our research. The reason for choosing GSE59733 is that the treatment condition of GSE59733 is simple and clear, only before and after irradiation. There are few interference factors which are helpful to explain the influence of radiotherapy on breast cancer more clearly.

In this study, we identified DEGs with a cutoff degree of 2-fold change by microarray data analysis. We analyzed total transcriptional changes through functional annotation of DEGs. The hub genes, which are the possession of the complex interactions with other protein nodes, and primary subnetworks were extracted from the protein–protein interaction (PPI) network. Meanwhile, we exactly validated the hub gene expression trend in irradiated breast cancer cells by quantitative real-time polymerase chain reaction (qRT-PCR). Taken together, these findings provide new insights into cell response to radiation to further understand the important roles of messenger RNA (mRNA) differential expression in cellular radiation response.

**Materials and Methods**

**Data Collection and Preprocessing**

The raw expression matrix file from GSE59733 microarray data set was downloaded from GEO database supported by National Center of Biotechnology Information. It consists of 19 samples with early-stage breast cancer that received preoperative radiotherapy, including 9 preradiation tumor samples and 10 postradiation tumor samples. According to the annotation information provided by platform GPL8990 [Xcel] Affymetrix Human Almac Xcel Array, probe ID and Gene Symbol achieved transformation in downloaded expression matrix. The unrecognized probes were removed, and the repeated gene probes were merged by averaging using basic functions in R (version 3.5.1, https://www.r-project.org/). Missing data were imputed using the k-nearest neighbor algorithm by impute package in R. The data normalization was also performed by normalizeBetweenArrays function in limma package in R.

**Identification and Functional Annotation of DEGs**

The DEGs in postradiation tumor samples, compared to preradiation tumor samples, were screened from preprocessed gene expression data by limma package. The t-test method was applied to calculate the adjusted P value of the DEGs. The adjusted P value <.05 and |log2FC (fold change)| ≥ 1 were selected as the cutoff criteria. To better describe the biological functions for DEGs involved in encoding protein, Gene Ontology (GO) functions analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed by clusterProfiler package in R. The P value <.01 and Q value <.05 were used to restrict the output results for functional annotation.

**The PPI Network Construction**

STRING (version10.5, https://string-db.org/) database is mainly used to describe the complex biological relationship between genes, containing a large amount of known and predicted PPI information. In this study, the DEGs involved in encoding proteins were retrieved in STRING database, and combined score ≥0.4 was identified as a major selection condition. Then, the PPI network was visualized by Cytoscape software. Meanwhile, 2 typical network topological parameters were utilized to filter out hub nodes in PPI network by plugin cytoHubba. In Cytoscape, including degree centrality and betweenness centrality. Subsequently, the plugin Molecular Complex Detection (MCODE) was applied to screening the primary modules of PPI network with the threshold set as follows: cutoff criteria degree = 2, K-core = 2, node score = 0.2, max depth = 100, and node number ≥5. What’s more, the extracted modules were analyzed by KEGG pathway analysis based on STRING database.

**The MRNA Expression of Hub Genes in Normal and Tumor Samples**

The UALCAN database (http://ualcan.path.uab.edu/) is available and is a powerful database. It contains amounts of biological information regarding patients with cancer, such as survival information based on gene expression and its level. It also may help users to obtain more cancer transcriptome data and identify potential biomarkers. In this study, it was applied to evaluating mRNA expression data of hub genes between normal samples and tumor samples.
Cell Culture and Radiation

The human normal breast cell MCF-10A and the human breast cancer cell lines MCF-7, ZR-75-1, MDA-MB-231, SUM-159, and BT474 were obtained from American Type Culture Collection in the United States. The cells were cultured in Dulbecco’s Modified Eagle’s medium or RPMI medium (Gibco) supplemented with 10% fetal bovine serum (Gibco), 100 U/mL penicillin, and 100 μg/mL streptomycin in 5% CO2 at 37°C. Besides, the MCF-7 or SUM-159 cells (1 x 10⁶) were seeded in 6-well plates. After 24 hours, MCF-7 or SUM-159 cells were exposed to 5 Gy irradiation dose using a ¹³⁷Cs-ray source at a dose rate of 0.99 Gy/min.

The QRT-PCR Assay

Total RNA was extracted using TRIzol reagent (Invitrogen) according to the manufacturer’s protocol. Reverse transcription was performed using a PrimeScript RT Reagent Kit (TaKaRa), according to the manufacturer’s instructions. SYBR Green Master (ROX) (Roche) was used to quantify hub genes in mRNA level, including FOS, CCL2, and CXCL12. Glyceraldehyde 3-phosphate dehydrogenase was amplified as control. The relative mRNA expression levels of hub genes were calculated using the 2⁻ΔΔCT method. The qRT-PCR cycles were as follows: 50 °C for 2 minutes, 95 °C for 10 minutes, and then 40 cycles of 95 °C for 15 seconds and 60 °C for 2 minutes. The primers are listed in Table 1. Each experiment was repeated at least 3 times.

Statistical Analysis

The microarray data were analyzed by R software (version 3.5.1). Data in the present study were analyzed using SPSS statistical software (version 19.0). Statistical significance was assessed by comparing mean ± SD using a Student t test for independent groups. A 2-tailed P < .05 was considered statistically significant.

Results

Identification of DEGs in Breast Cancer After Irradiation

After data preprocessing, a total of 12 487 genes were mapped to probes. The normalization result for gene expression data was visualized with box plot (Figure S1). The result showed that the median values of each set of data were consistent, and the nonexperimental error was eliminated after normalization. According to the selected criteria (adjusted P value <.05 and |log₂FC (fold change)| ≥ 1), a total of 82 DEGs, including 74 upregulated genes and 8 downregulated genes, were obtained after data processing in postradiation tumor samples, compared to the preradiation tumor samples. These DEGs are shown in Table S1. The volcano plot was applied to distribution for DEGs (Figure S2).

Gene Ontology and KEGG Analysis of DEGs

To explore the functional roles of the DEGs, we analyzed these candidate genes and found that a total of 62 GO terms were eligible for screening conditions (P value <.01 and Q value <.05) by clusterProfiler package in R programming. Fifty-seven GO terms were enriched in biological process (BP), and 5 GO terms were enriched in molecular function (MF). The GO enrichment analysis for BP ontology showed that these DEGs were mainly enriched in response to extracellular stimulus, nutrient levels, and intrinsic apoptotic signaling pathway in response to DNA damage (Figure 1A). In MF ontology, these DEGs were closely associated with G-protein-coupled receptor binding, chemokine receptor binding, and CXCR chemokine receptor binding (Figure 1B). Additionally, the results of KEGG pathway analysis revealed that these DEGs were involved in interleukin 17 signaling pathway and tumor necrosis factor (TNF) signaling pathway. This result is shown in Table 2.

The PPI Network Analysis of DEGs

To further explore the potential biological functions of DEGs, we constructed the PPI network by retrieving DEGs in STRING public database. The results showed that there were 73 nodes and 101 interactions in this network (Figure 2A). Top 6 hub genes were determined by calculating degree centrality and betweenness centrality, respectively. This result is shown in Table 3. The results suggested that a total of 3 hub genes were obtained by overlapping, including FOS, CCL2, and CXCL12. What’s more, 2 representative network modules were extracted from PPI network by MCODE plugin in

Table 1. List of Primers in this Study.

| Genes                        | Primers | Primer sequence (5’-3’)       | Product length (bp) |
|------------------------------|---------|-------------------------------|---------------------|
| Glyceraldehyde 3-phosphate dehydrogenase | Forward | GGAGCGGATCCCTCCAAAT           | 197                 |
|                              | Reverse | GGCTGTGTCACTTTCTCATGG         | 126                 |
| FOS                          | Forward | GGGGCAAGGTTGAAAGTTAT          | 236                 |
|                              | Reverse | CCGGTTGAGGTGATATGCCAG         |                     |
| CCL2                         | Forward | CTTTCTATTCCCAAGGCTC           |                     |
|                              | Reverse | GTTTTGCTTTGCCTCAAGGTG         |                     |
| CXCL12                       | Forward | ATGCCCCATGCGGATTCTT           | 126                 |
|                              | Reverse | CACACTTGTCTGTTGTTTCTT         |                     |

Zhu et al
Cytoscape 3.6.0. And the results of KEGG pathway analysis for each module suggested that these DEGs were highly involved in inflammation-associated signaling pathways, where DEGs in 2 modules are enriched in TNF signaling pathway, respectively (Figure 2B and C).

Validation of the mRNA Expression of Hub Genes

Next, we found that the 3 hub genes (FOS, CCL2, and CXCL12) were downregulated in tumor samples, compared to normal samples based on UALCAN database (Figure 3A). Similarly, these results were confirmed in vitro. The mRNA expression level of the 3 hub genes in breast cancer cell lines were downregulated compared to normal breast cell MCF-10A (Figure 3B). To further verify the accuracy of microarray data analysis and explore the 3 hub genes response to radiation, we detected the relative mRNA expression trend of hub genes using qRT-PCR assay in irradiated breast cancer cells MCF-7 and SUM-159. As expected, our data demonstrated that the mRNA expression of these 3 hub genes were significantly increased after γ-irradiation (5 Gy) in 2 breast cancer cells which were consistent with results of data analysis (Figure 3C).

Discussion

In this study, we retrieved gene expression profiling in response to radiation treatment in breast cancer from GEO database. It is found that the hidden networks of interaction regarding radiation-associated genes in GSE59733 data set haven’t yet been fully explored using bioinformatics. Interestingly, we identified 82 DEGs, including 74 upregulated genes and 8 downregulated genes, based on 19 tumor samples using algorithm in limma package. Moreover, the results from GO analysis and KEGG pathway analysis implied that these DEGs were mainly involved in extracellular stimulus, cell apoptosis, inflammation, immune responses, and DNA damage response. It has been evidenced that these BPs are closely related to radiation-induced response.25-28
Three outstanding hub proteins were filtered out from the PPI network, including FOS (AP-1 transcription factor), CCL2 (C-C motif chemokine ligand 2), and CXCL12 (C-XC motif chemokine ligand 12). And we validated the mRNA expression changes of FOS, CCL2, and CXCL12 in breast cancer cells after radiation. The results showed that the expression trend of these 3 genes was consistent with gene expression profiling. It implied that the 3 hub genes are closely associated with radiation-induced response through mediating a series of cascade reaction.

Table 3. The Hub Genes Obtained by Cytohubba.

| Gene name | Degree | Gene name | Betweenness |
|-----------|--------|-----------|-------------|
| FOS       | 18     | CCL2      | 707.65      |
| CCL2      | 16     | FOS       | 556.01      |
| EGR1      | 12     | MAS4A6A   | 364.58      |
| CXCL12    | 10     | FGR2A     | 361.74      |
| ATF3      | 9      | FPR1      | 358.63      |
| FPR1      | 8      | CXCL12    | 236.36      |

Figure 2. Construction and analysis of the protein–protein interaction (PPI) networks. A, The PPI networks. B, Module 1 and enriched pathways. C, Module 2 and enriched pathways.
**FOS** is one of 3 hub genes which plays a critical role in the development and progression of breast cancer by mediating the transcription of AP-1 target genes. However, there are fewer studies to uncover FOS response to radiation in breast cancer. Recent studies indicated that silencing FOS could sensitize glioma cells to radiation by disturbing DNA damage repair and cell cycle arrest. In our analysis, FOS was one of the hub proteins with the highest interaction in the PPI network. Thus, whether altering its gene expression may affect the response to radiation of breast cancer cells is still an urgent need of in-depth study.

**Figure 3.** The expression levels of the 3 hub genes. A. The messenger mRNA (mRNA) expression data of the 3 hub genes were obtained from UALCAN database. B. The mRNA expression levels of the 3 hub genes in normal breast cell and breast cancer cell lines by quantitative real-time polymerase chain reaction (qRT-PCR). C. The mRNA expression levels of the 3 hub genes in irradiated MCF-7 and SUM-159 cells by qRT-PCR. ***p < .001.
Additionally, CCL2 and CXCL12 as chemokines were mainly involved in immunoregulatory and inflammatory processes. In in vivo experiment, it has been reported that the expression of CCL2 and CXCL2 was significantly increased after radiation.\textsuperscript{31,32} In this study, the results of functional annotation for DEGs further strengthen the evidence that the signaling pathways associated with inflammation are abnormally activated such as chemokine signaling pathway and TNF signaling pathway. There are studies pointing out the binding of chemokines and its ligand is involved in the biological responses of tumors to radiation. For example, CXCR4 and its receptor CXCL12 may regulate the survival of prostate cancer cells after irradiation.\textsuperscript{33} And Kalbasi et al demonstrated chemokine CCL2 could prevent the recruitment of inflammatory monocytes to tumors and reduce the growth of tumors when combined with radiotherapy.\textsuperscript{34}

In conclusion, the 3 hub genes obtained from the PPI network are closely associated with radiation response, which may be expected to be therapeutic targets or diagnostic markers for radiotherapy patients in the future. Nevertheless, it still required a rigorous experiment to support in the following research. Anyway, this study provided a new insight into breast cancer response to radiation.

Authors’ Note

Changchen Zhu and Chang Ge equally contributed to this work.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by grants from the National Natural Science Foundation of China [grant number 81703042, 81572969 and 81730086] and CAMS Innovation Fund for Medical Sciences [CIFMS, 2017-I2M-1-016], and this project was funded by China Postdoctoral Science Foundation [grant number 2018T110071].

ORCID iD

Guoxing Feng https://orcid.org/0000-0002-5403-3698

Supplemental Material

Supplemental material for this article is available online.

References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67(1):7-30.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
3. Early Breast Cancer Trialists' Collaborative Group, Mc Gale P, Taylor C, et al. Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials. Lancet. 2014;383(9935):2127-2135.
4. Pajic M, Froio D, Daly S, et al. MiR-139-5p modulates radiotherapy resistance in breast cancer by repressing multiple gene networks of DNA repair and ROS defense. Cancer research. 2018;78(2):501-515.
5. Speers C, Pierce LJ. Postoperative radiotherapy after breast-conserving surgery for early-stage breast cancer: a review. JAMA Oncol. 2016;2(8):1075-1082.
6. Montero A, Ciervide R, Garcia-Aranda M, Rubio C. Postmastectomy radiation therapy in early breast cancer: utility or futility? Crit Rev Oncol Hematol. 2020;147:102887.
7. Speers C, Zhao S, Liu M, Bartelink H, Pierce LJ, Feng FY. Development and validation of a novel radiosensitivity signature in human breast cancer. Clin Cancer Res. 2015;21(16):3667-3677.
8. Kyndt M, Sorensen FB, Knudsen H, et al. Estrogen receptor, progesterone receptor, HER-2, and response to postmastectomy radiotherapy in high-risk breast cancer: the Danish Breast Cancer Cooperative Group. J Clin Oncol. 2008;26(9):1419-1426.
9. Wu D, Zhou W, Wang S, Zhou Z, Wang S, Chen L. Tob1 enhances radiosensitivity of breast cancer cells involving the JNK and p38 pathways. Cell Biol Int. 2015;39(12):1425-1430.
10. Zhou YC, Liu JY, Li J, et al. Ionizing radiation promotes migration and invasion of cancer cells through transforming growth factor-beta-mediated epithelial-mesenchymal transition. Int J Radiat Oncol Biol Phys. 2011;81(5):1530-1537.
11. Ghisolfi L, Keates AC, Hu X, Lee DK, Li CJ. Ionizing radiation induces stenness in cancer cells. PloS one. 2012;7(8):e43628.
12. Jabbari N, Nawaz M, Rezaie J. Bystander effects of ionizing radiation: conditioned media from X-ray irradiated MCF-7 cells increases the angiogenic ability of endothelial cells. Cell Commun Signal. 2019;17(1):165.
13. Clough E, Barrett T. The Gene Expression Omnibus database. Meth Mol Biol. 2016;1418:93-110.
14. Bras LP, Menezes JC. Improving cluster-based missing value estimation of DNA microarray data. Biomol Eng. 2007;24(2):273-282.
15. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
16. Thomas PD. The Gene Ontology and the meaning of biological function. Meth Mol Biol. 2017;1446:15-24.
17. Kanehisa M. The KEGG database. Novartis Found Symp. 2002;247:91-101.
18. Yu G, Wang LG, Han Y, He QY. Cluster profiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-287.
19. Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic acids research. 2017;45(D1):D362-D368.
20. Cline MS, Smoot M, Cerami E, et al. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007;2(10):2366-2382.
21. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. Cytohubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8(suppl 4):S11.
22. Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks. Nature. 2001, 411(6833):41-42.
23. Xiong W, Xie L, Zhou S, Liu H, Guan J. The centrality of cancer proteins in human protein-protein interaction network: a revisit. Int J Comput Biol Drug Des. 2014, 7(2-3):146-156.
24. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, Varanbally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia. 2017;19(8):649-658.
25. Khan KH, Blanco-Codesido M, Molife LR. Cancer therapeutics: targeting the apoptotic pathway. Critical Rev Oncol/Hematol. 2014;90(3):200-219.
26. Schaue D, Micewicz ED, Ratikan JA, Xie MW, Cheng G, McBride WH. Radiation and inflammation. Semin Radiat Oncol. 2015;25(1):4-10.
27. Multhoff G, Radons J. Radiation, inflammation, and immune responses in cancer. Front Oncol. 2012;2:58.
28. Jabbari N, Nawaz M, Rezaie J. Ionizing radiation increases the activity of exosomal secretory pathway in MCF-7 human breast cancer cells: a possible way to communicate resistance against radiotherapy. Int J Mol Sci. 2019;20(15):8.
29. Fu J, Cheng L, Wang Y, et al. The RNA-binding protein RBMS1 represses AP-1 signaling and regulates breast cancer cell proliferation and migration. Biochim Biophys Acta. 2015;1853(1):1-13.
30. Bland KI, Konstadoulakis MM, Vezeridis MP, Wanebo HJ. Oncogene protein co-expression. Value of Ha-ras, c-myc, c-fos, and p53 as prognostic discriminants for breast carcinoma. Ann Surg. 1995;221(6):706-718.
31. Malaviya R, Gow AJ, Francis M, Abramova EV, Laskin JD, Laskin DL. Radiation-induced lung injury and inflammation in mice: role of inducible nitric oxide synthase and surfactant protein D. Toxicol Sci. 2015;144(1):27-38.
32. Gong HY, Hu WG, Hu QY, Li XP, Song QB. Radiation-induced pulmonary injury accelerated pulmonary metastasis in a mouse model of breast cancer. Oncology Lett. 2015;14(1):27-38.
33. Domanska UM, Boer JC, Timmer-Bosscha H, et al. CXCR4 inhibition enhances radiosensitivity, while inducing cancer cell mobilization in a prostate cancer mouse model. Clin Exp Metastasis. 2014;31(7):829-839.
34. Kalbasi A, Komar C, Tooker GM, et al. Tumor-derived CCL2 mediates resistance to radiotherapy in pancreatic ductal adenocarcinoma. Clin Cancer Res. 2017;23(1):137-148.