RNA-Sequence Analysis Reveals Differentially Expressed Genes (DEGs) in Patients Exhibiting Different Risks of Tumor Metastasis

Li-Feng Dong
Shu-Ying Xu
Jing-Pei Long
Fang Wan
Yi-Ding Chen

Background: Breast cancer is one of the most common malignancies in women. In a previous study, we found that for two patients who had a high risk of lymphatic metastasis, lymphatic metastasis did not occur; whereas, for two patients who had a low risk of lymphatic metastasis, lymphatic metastasis did occur.

Material/Methods: We analyzed the differential gene expressions of these four patients by RNA-sequence. The data (HRNM_T versus HRNM_N, LRYM_T versus LRYM_N, and HRNM_T versus LRYM_T) was then processed using differentially expressed genes (DEGs) analysis, functional analysis for DEGs, and PPI network construct.

Results: For HRNM_T versus HRNM_N, there were 224 DEGs. There were 504 DEGs for LRYM_T versus LRYM_N, and 88 DEGs for LRYM_T versus LRYM_N. For HRNM_T versus HRNM_N, DEGs were up-regulated mainly in the cell cycle, the IL-17 signaling pathway, and the progesterone-mediated oocyte maturation; DEGs were down-regulated mainly in the IL-17 signaling pathway. For LRYM_T versus LRYM_N, DEGs were up-regulated mainly in protein digestion and absorption, and cytokine-cytokine receptor interaction; DEGs were down-regulated mainly in ECM-receptor interaction. For HRNM_T versus LRYM_T, DEGs were up-regulated mainly in the PPAR signaling pathway; DEGs were down-regulated mainly in the adipocytokine signaling pathway. The DEGs were screened to construct PPI networks.

Conclusions: The GO and KEGG functional enrichments of HRNM_T versus HRNM_N, and LRYM_T versus LRYM_N were consistent with earlier studies. For HRNM_T versus LRYM_T, DEGs were up-regulated mainly in PPAR signaling; DEGs were down-regulated mainly in the adipocytokine pathway.

MeSH Keywords: Gene Expression Profiling • Protein Interaction Maps • RNA 3’ Polyadenylation Signals

Full-text PDF: http://www.medscimonit.com/abstract/index/idArt/904789

Corresponding Author: Yi-Ding Chen, e-mail: ydchen@zju.edu.cn

Source of support: This work was supported by Natural Science Foundation of Zhejiang Province (LZ16H160002), which funded the study in collection, analysis, and interpretation of data.

1 Department of Breast, Women’s Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, P.R. China
2 Physical Examination Center, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, P.R. China
3 Department of Surgical Oncology, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, P.R. China

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]
Breast cancer is one of the most common malignancies in women, accounting for a quarter of new cases and 15% of deaths caused by cancer in females [1]. Being the leading cause of cancer death, breast cancer threatens the life and health of females and causes enormous impact on the economy, society, families, and psychological health of women [2]. In 2016, 1,677,000 women were newly diagnosed with breast cancer, representing 25.2% of all newly diagnosed cancers according to Cancer Statistics 2017 [3]. Among these new cases, more than 187,000 were from China and this number increased at twice the speed of the global growth rate [4]. Although the incidence of breast cancer in China is lower than the global rate, recently it has been rising sharply. According to the latest China Cancer Registration Report, the breast cancer incidence rate among Chinese women was 42.55/100,000 in 2012 [4]. Although this is a low incidence rate compared with the whole world, the growth rate has been twice the velocity of the global growth rate since 1990s, especially in urban areas. Currently, breast cancer is the most frequent cancer diagnosed in Chinese women and the sixth leading cause of death from cancer in Chinese women. The development of cancer involves multiple levels and stages. Metastasis plays a critical role in its progression, and ultimately leads to high fatality rates. In light of the complexity of metastasis itself, we need to more deeply explore its molecular mechanism including preclinical and clinical aspects. We need to identify early alarming signals and make timely interventions to improve therapeutic effects and quality of life, as well as relieve social and personal burdens. Tumor metastasis is very complex and involves many metastasis-associated molecules; there appears to be little effect on metastasis through single molecule-targeting interventions.

In this study, tumor tissue and the adjusted normal tissue from four patients were used for RNA-sequence analysis. Two of the patients had a high risk for lymphatic metastasis, but lymphatic metastasis had not occurred at the time of the study. The other two patients were of low risk for lymphatic metastasis, but lymphatic metastasis had occurred. We then analyzed the differential gene expression of these four patients by RNA-sequence in order to investigate elements of breast carcinoma metastasis.

**Material and Methods**

**Ethical statement**

This study was reviewed and approved by Ethics Committee of Department of Breast, Women’s Hospital, School of Medicine, Zhejiang University. Each patient involved in this study provided written informed consent.

**Background**

**Patient samples and clinicopathological characteristics**

Samples were acquired from breast cancer patients who underwent surgical resection between May 2012 and December 2013 at the Department of Breast, Women’s Hospital, School of Medicine, Zhejiang University. Clinicopathological characteristic of each patient were recorded. Adjacent normal tissues were resected at least 3 centimeters away from the tumor, while tumor samples were carefully extracted from the center of the breast cancer carcinoma. All of the samples were frozen in liquid nitrogen and stored at −80°C. Four matched pairs of normal tissue samples and breast cancer tissues were obtained. These samples were used for RNA sequencing analysis.

**RNA preparation**

Total RNA from breast cancer tissue and adjacent normal tissue was extracted by TRIzol (Invitrogen, Carlsbad, CA, USA), redissolved in DEPC-treated water and quantified using NanoVue Plus spectrophotometry (GE Healthcare, Fairfield, CT, USA). RNA integrity was evaluated using agarose gel electrophoresis, and DNA contamination was eliminated using gDNA Eraser (Takara, Tokyo, Japan) according to the manufacturer’s guidelines.

**RNA sequencing (RNA-Seq)**

Breast cancer and adjacent normal tissue samples were grounded in TRIzol and total RNA was isolated using the RNeasy Kit (Qiagen) according to the manufacturer’s guidelines. DNasel was added in order to ensure that the tissues were not mixed with genomic DNA. RNA purity was detected by Agilent BioAnalyzer 2100 (Agilent, Santa Clara, CA, USA). Transcriptome sequencing libraries were prepared using the TrueSeq RNA Access Library Prep Kit® (Illumina, CA, USA). Paired end sequencing (100 bp) was performed at an Illumina HiSeq 2000 instrument, which is supported by Beijing Novogene Biological Information Technology Co., LTD.

**Differentially expressed genes (DEGs) analysis**

The LIMMA package [5] in R language, a linear regression model, was used to identify the DEGs in breast tissue and adjacent tissue. A p value <0.05 and |log Fold Chance (FC)| >1.0 were set as thresholds to screen out differentially expressed genes (DEGs).

**Functional analysis for DEGs**

Database for Annotation, Visualization and Integrated Discovery (DAVID) [6] is considered an integrated biological knowledge-base and effective analytic tool to systematically extract biological information and explore the biological meaning behind lists of genes or proteins. GO [7] and KEGG [8] were used to...
perform the functional analysis of DEGs. We chose a \( p \) value less than 0.05 as cut-off criterion.

**PPI network constructing**

The PPI network corresponding to the screened DEGs was retrieved from the Search Tool for the Retrieval of Interacting Genes (STRING) database [9]. STRING is used for identifying known and predicted protein interactions. Proteins in the interaction network are represented with nodes, while the interactions between two proteins are represented with edges. The selection of the hub protein is based on the score of the nodes, which is calculated by the count of edges launching from a protein in the PPI network. The PPI networks in the present study were constructed using Cytoscape [10] when the confidence scores were more than 0.4.

**Results**

**DEGs in breast cancer tissues**

Patient clinicopathological characteristics are shown in Table 1, and four patients (Chenfx, Yuay, Dongyh, and Yumh) were involved in this study. In Table 2, we can see that Chenfx and Yuay showed high risk of metastasis, but there was no tumor
metastasis. Dongh and Yumh exhibited low risk of metastasis, but tumor metastasis occurred. Therefore, we analyzed the DEGs in breast cancer tissues (HRNM_T versus HRNM_N, LRYM_T versus LRYM_N, and HRNM_T versus LRYM_T). Table 3 shows that 107 genes were up-expressed and 117 genes were down-expressed in the comparison of HRNM_T versus HRNM_N. There were 220 genes up-expressed and 284 genes down-expressed in the comparison of LRYM_T versus LRYM_N. There were 61 genes up-expressed and 27 genes down-expressed in the comparison of HRNM_T versus LRYM_T.

**Significant function and pathways of DEGs**

GO functional and KEGG pathway enrichment analyses were performed to both up- and down-regulated genes, and the results are shown in Figure 1 and Tables 4–6. According to the...
result, the DEGs in HRNM_T versus HRNM_N were enriched in multiple GO categories, mainly the cell cycle, the IL-17 signaling pathway, and the progesterone-mediated oocyte maturation. The DEGs in LRYM_T versus LRYM_N were enriched mainly in protein digestion and absorption, cytokine-cytokine receptor interaction and ECM-receptor interaction. The DEGs in HRNM_T versus LRYM_T were enriched mainly in the PPAR signaling pathway and the adipocytokine signaling pathway.

**Analysis of PPI network and module**

The PPI network for the screened DEGs was constructed (Figures 2–4), and contained 129 nodes and 294 edges in HRNM_T versus HRNM_N (Figure 2), 321 nodes and 904 edges in LRYM_T versus LRYM_N (Figure 3), and 16 nodes and 13 edges in HRNM_T versus LRYM_T (Figure 4). Module analysis of PPI networks was performed by MCODE and the most significant module was selected, and was named subnetwork1. We also used DAVID to identify significant enrichment of these genes in multiple GO categories and KEGG pathways of subnetwork1. DEGs in subnetwork1 were mainly enriched in GO terms related to the cell cycle in HRNM_T versus HRNM_N, protein digestion and absorption in LRYM_T versus LRYM_N and the PPAR signaling pathway in LRYM_T versus LRYM_N.

**Table 4.** KEGG enrichment analyses for both up-regulated and down-regulated genes in HRNM_T vs. HRNM_N.

| Category                        | Genes                                      | p-Value   |
|---------------------------------|--------------------------------------------|-----------|
| **Up-regulated genes**         |                                            |           |
| KEGG_PATHWAY Cell cycle         | E2f1,2,3, CycA, Cdc6, Cdc25A, CDK1, CycB,  | 4.17E-06  |
|                                 | Bub1, Mad2, BubR1, Mps1                    |           |
| KEGG_PATHWAY IL-17 signaling pathway | Eotxin, CXCL10, S100A7, S100A8, S100A9,  | 7.49E-03  |
|                                 | MMP1, MMP13                                |           |
| KEGG_PATHWAY Progesterone-mediated oocyte maturation | PDE3, CycA, Cdc25, Bub1, Mad1/2/3, Cdc2, CdcB | 2.36E-04  |
| **Down-regulated genes**       |                                            |           |
| KEGG_PATHWAY IL-17 signaling pathway | IL-17B                                      | 7.49E-03  |

**Table 5.** KEGG enrichment analyses for both up-regulated and down-regulated genes in LRYM_T vs. LRYM_N.

| Category                        | Genes                                      | p-Value   |
|---------------------------------|--------------------------------------------|-----------|
| **Up-regulated genes**         |                                            |           |
| KEGG_PATHWAY Protein digestion and absorption | FATCD36, FABP, LPL, PGAR, Penlipin, aP2,  | 6.11E-08  |
|                                 | PECK, AQ7                                  |           |
| KEGG_PATHWAY Cytokine-cytokine receptor interaction | CXCL1, CXCL3, CXCL5, CCL28, LEP, CNTFR,  | 5.80E-06  |
|                                 | PDGFA, VEGFOP, KIT, VSF11A                 |           |
| **Down-regulated genes**       |                                            |           |
| KEGG_PATHWAY ECM-receptor interaction | Collagen, THBS, Fibronectin                | 1.38 E-02 |

**Table 6.** KEGG enrichment analyses for both up-regulated and down-regulated genes in HRNM_T vs. LRYM_T.

| Category                        | Genes                                      | p-Value   |
|---------------------------------|--------------------------------------------|-----------|
| **Up-regulated genes**         |                                            |           |
| KEGG_PATHWAY PPAR signaling pathway | Penlipin                                    | 5.82E-05  |
| KEGG_PATHWAY Adipocytokine signaling pathway | ADIPO                                       | 1.13E-03  |

KEGG – Kyoto encyclopedia of genes and genomes.
Discussion

Currently, breast cancer is the most frequent cancer in Chinese women and the sixth leading cause of death from cancer in Chinese women. The development of cancer involves multiply levels and stages. Previously, we analyzed clinicopathological characteristics of the study patients. We found Chenfx and Yuay showed high risk for metastasis, but no tumor metastasis occurred. We found that Dongyh and Yumh exhibited low risk of metastasis, but tumor metastasis occurred. Hence, we aimed to study the DEGs in HRNM_T versus HRNM_N, LRYM_T versus LRYM_N, and HRNM_T versus LRYM_T and clarify the differences.

The abnormality of cell proliferation is generally due to the abnormal expression of related genes, which gives rise to the occurrence and development of tumors [11,12]. In this study, we first identified the DEGs in HRNM_T versus HRNM_N, LRYM_T versus LRYM_N, and HRNM_T versus LRYM_T. The results are shown in Table 1. For HRNM_T versus HRNM_N, there were a total of 224 DEGs. There were a total 504 DEGs in LRYM_T versus LRYM_N, and a total of 88 DEGs in LRYM_T versus LRYM_N. DEGs in the three groups were then analyzed by GO (biological process, BP; cellular component, CC; molecular function, MF) and KEGG functional enrichment. For HRNM_T versus HRNM_N, the DEGs were up-regulated mainly in the cell cycle, the IL-17 signaling pathway and the progesterone-mediated oocyte maturation; and DEGs were down-regulated mainly in the IL-17 signaling pathway. Cdc6, Cdc25A, CDK1, and CycB were up-regulated; Cdc6, Cdc25A, CDK1, and CycB are known to be involved in the cell cycle [13,14], and up-regulation of Cdc6, Cdc25A, CDK1, and CycB can result in tumor initiation and progression. Numerous factors contribute to the progression of breast cancer, including inflammation. The IL17 family consist of six protein members, among them IL17B and its receptor; and the IL17RB signaling pathway plays a key role in development and progression of breast cancer [15,16]. Eotaxin, CXCL10, S100A7, S100A8, S100A9, MMP1, and MMP13 were up-regulated and IL-17B was down-regulated. For LRYM_T versus LRYM_N, DEGs were up-regulated mainly in protein digestion and absorption and cytokine-cytokine receptor interaction; and DEGs were down-regulated mainly in ECM-receptor interaction. CXCL1, CXCL3, and CXCL5 were up-regulated; CXCL1, CXCL3, and CXCL5 are known to contribute to the invasion and metastasis [17,18]. Collagen, THBS, and fibronectin are known to contribute to ECM-receptor interaction, which is down-regulated. ECM is an important organizational barrier of tumor metastasis [19–21]. For HRNM_T versus LRYM_T, DEGs were
Figure 3. PPI networks for DEGs in LRYM_T versus LRYM_N. The nodes represent proteins and edges represent pair-wise interactions. The red nodes represent the proteins encoded by up-regulated genes and the green represent the proteins encoded by down-regulated genes.

Figure 4. PPI networks for DEGs in HRNM_T versus LRYM_T. The nodes represent proteins and edges represent pair-wise interactions. The red nodes represent the proteins encoded by up-regulated genes and the green represent the proteins encoded by down-regulated genes.

up-regulated mainly in the PPAR signaling pathway and the adipocytokine signaling pathway. Peroxisome proliferator-activated receptor (PPAR) signaling pathways are nuclear hormone receptors that are activated by fatty acids and their derivatives and may be an important predictor of breast cancer response to neoadjuvant chemotherapy [22]. Expression of the nuclear receptor peroxisome proliferator activated receptor delta (PPARdelta) in breast cancer cells is negatively associated with patient survival [23]. ADIPO has been shown to be related to PPAR signaling pathway and AMPK signaling pathway, and up-regulated [24]. Cancer growth and metastasis
depends on the availability of energy. Energy-sensing systems are critical in maintaining a balance between the energy supply and utilization of energy for tumor growth. A central regulator in this process is AMPK [24]. In our study, the DEGs were screened to construct PPI networks and significant modules, and named subnetworks were identified. The functional analysis and pathways enrichment showed that DEGs in the subnetworks were enriched in pathways.

References:

1. Zhang M, Guo W, Qian J, Wang B et al: Negative regulation of CDC42 expression and cell cycle progression by miR-29a in breast cancer. Open Med (Wars), 2016; 11(1): 78–82
2. Falco G, Rocco N, Bordoni D et al: Contralateral risk reducing mastectomy in Non-BRCA-Mutated patients. Open Med (Wars), 2016; 11(1): 238–41
3. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2017. Cancer J Clin, 2017; 67(1): 7–30
4. Chen W, Zheng R, Baade PD et al: Cancer statistics in China, 2015. Cancer J Clin, 2016; 66(2): 115–32
5. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol, 2004; 3: Article3
6. Huang DW, Sherman BT, Tan Q et al: The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol, 2007; 8(9): R183
7. Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000; 25(1): 25–29
8. Kanelisa M, Goto S, Hattori M et al: From genomes to chemical genomes: New developments in KEGG. Nucleic Acids Res, 2013; 41(Database issue): D354–57
9. Franceschini A, Szklarczyk D, Frankild S et al: STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res, 2013; 41(Database issue): D808–15
10. Smoot ME, Ono K, Ruscheinski J et al: Cytoscape 2.8: New features for data integration and network visualization. Bioinformatics, 2011; 27(3): 431–32
11. Yang PY, Miao NF, Lin CW et al: Impact of maspin polymorphism rs2289520 G/C and its interaction with gene to gene, alcohol consumption increase susceptibility to oral cancer occurrence. PLoS One, 2016; 11(8): e0160841
12. Wang SY, Long JB, Killelea BK et al: Preoperative breast magnetic resonance imaging and contralateral breast cancer occurrence among older women with ductal carcinoma in situ. Breast Cancer Res Treat, 2016; 158(1): 139–48
13. Tsai CH, Hung AC, Chen YY et al: 3’-hydroxy-4’-methoxy-beta-methyl-beta-nitrostyrene inhibits tumorigenesis in colorectal cancer cells through ROS-mediated DNA damage and mitochondrial dysfunction. Oncotarget, 2017; 8(11): 18106–17
14. Su Z, Cai L, Lu J et al: Global expression profile of tumor stem-like cells isolated from MMQ rat prolactinoma cell. Cancer Cell Int, 2017; 17: 15
15. Klein JC, Wild CA, Lang S, Brandau S: Differential immunomodulatory activity of tumor cell death induced by cancer therapeutic toll-like receptor ligands. Cancer Immunol Immunother, 2016; 65(6): 689–700
16. Alinejad V, Dolati S, Motallebnezad M, Yousefi M: The role of IL17B-IL17RB signaling pathway in breast cancer. Biomed Pharmacother, 2017; 88: 795–803
17. Han KQ, He XQ, Ma MY et al: Targeted silencing of CXCL1 by siRNA inhibits tumor growth and apoptosis in hepatocellular carcinoma. Int J Oncol, 2015; 47(6): 2131–40
18. Wang M, Zhao J, Zhang L et al: Role of tumor microenvironment in tumorigenesis. J Cancer, 2017; 8(5): 761–73
19. Isik A, Gursul C, Peker K et al: Metalloproteinas and their inhibitors in patients with inguinal hernia. World J Surg, 2017; 41(5): 1259–66
20. Isik A, Karavas E, Peker K et al: Male Mondor’s disease is a rare entity. Breast J, 2016; 22(6): 700–1
21. Klein JC, Wild CA, Lang S, Brandau S: Differential immunomodulatory activity of tumor cell death induced by cancer therapeutic toll-like receptor ligands. Cancer Immunol Immunother, 2016; 65(6): 689–700
22. Chen YZ, Xue JY, Chen CM et al: PPAR signaling pathway may be an important predictor of breast cancer response to neoadjuvant chemotherapy. Cancer Chemother Pharmacol, 2012; 70(5): 637–44
23. Wang X, Wang G, Shi Y et al: PPAR-delta promotes survival of breast cancer cells in harsh metabolic conditions. Oncogenesis, 2016; 5(6): e232
24. Karim S, Merdad A, Schulten HJ et al: Low expression of leptin and its association with breast cancer: A transcriptomic study. Oncol Rep, 2016; 36(1): 43–48

Conclusions

This study analyzed the DEGs in HRNM_T versus HRNM_N, LRYM_T versus LRYM_N, and HRNM_T versus LRYM_T. The GO and KEGG functional enrichments of HRNM_T versus HRNM_N, and LRYM_T versus LRYM_N were consistent with earlier studies. For HRNM_T versus LRYM_T, DEGs were up-regulated mainly in the PPAR signaling pathway; DEGs were down-regulated mainly in the adipocytokine signaling pathway.