The identification of multifocal breast cancer-associated long non-coding RNAs

Z.-Q. LANG1,2, Y.-Q. WU3, X.-B. PAN2, G.-M. QU2, T.-G. ZHANG1

1Department of Pathology, Qilu Hospital, Shandong University, Jinan, China
2Department of Pathology, Yantai Yuhuangding Hospital Affiliated to Qingdao University, Yantai, China
3Department of Medicine, Yantai Yuhuangding Hospital Affiliated to Qingdao University, Yantai, China

Zhiqiang Lang and Yanqiu Wu contributed equally

Abstract. – OBJECTIVE: To identify long non-coding RNAs (lncRNAs), which may be associated with multifocal and multicentric breast cancer (MMBC) by analyzing the differential expression of non-coding RNA in MMBC and unifocal breast cancers (UBC).

PATIENTS AND METHODS: 156 cases of invasive MMBC patients (136 patients with 2 focuses and 20 patients with 3 focuses) and 130 cases of UBC were collected from January 2005 to December 2015 in Yantai Yuhuangding Hospital. The differentially expressed IncRNAs in MMBC and UBC were screened by gene chip. RT-PCR was used to verify the differentially expressed IncRNAs.

RESULTS: Significantly different expression was found in 1080 lncRNAs between MMBC and UBC (FC > 4, p < 0.05), among which, 458 were upregulated and 622 were down-regulated. In homologous lncRNAs, the expression levels of four lncRNAs including C19orf33, C3orf52, C15orf48 and C4orf19 in MMBC tissue were significantly different than those in UBC tissues. RT-PCR verified that the expression of C19orf33, C3orf52 and C15orf48 in MMBC tissue was significantly up-regulated and the expression of C4orf19 was significantly down-regulated, which was consistent with that of gene chip.

CONCLUSIONS: C19orf33, C3orf52, C15orf48 and C4orf19 may have important functions in MMBC, and may be used as markers for MMBC.

Key Words: Non-coding RNA, Multicentric breast cancer (MMBC), Unifocal breast cancers, Differential expression.

Introduction

Breast cancer is a malignant tumor, which develops from breast tissue1. Previous studies indicated that various populations were in poor health, suffering from it2. As the most common female cancers, breast cancer was reported to affect 1 out of 8 women worldwide and bring heavy psychological and economic burden to the patients and their families3,4. Due to the intrinsic resistance, previous clinical studies have shown that about 50% patients with breast cancer usually failed to respond to initial drug treatments, which in turn affect the outcomes5. Even worse, drug resistance, which was developed during the application of drugs in most of the cases, can significantly reduce the treatment efficacy and lead to the development of aggressive malignancies6,7. Multifocal and multicentric breast cancer, which is also called MMBC, is one type of breast cancer with multiple simultaneous foci8. Besides the above mentioned difficulties in the treatment of breast cancer, the treatment of MMBC is also challenged by the specific problems caused by its distinct pathogenesis features. Traditional treatments have been proved to be less effective in the treatment of MMBC compared with the treatment of UBC9. Due to the multiple invasive foci developed in MMBC, radiotherapy is less effective in stopping the tumor development and metastasis and the recurrence rate is high10. In addition, multiple wide local excisions caused by surgical treatment (either quadrantectomy or segmental resection) usually lead to less good cosmetic outcome11. Biomarkers have been widely used in the diagnosis and treatment of breast cancer. However, the specific biomarkers for MMBC still lack. Therefore, it will be of great clinical values to identify MMBC-specific biomarkers to improve the diagnosis and treatment of MMBC.

Long non-coding RNAs, which are also known as IncRNA, are a group of non-protein coding RNAs longer than 200 nucleotides. They are longer than short interfering RNAs (siRNAs),

Corresponding Author: Tingguo Zhang, MD; e-mail: zhangtingguoguo@163.com
microRNAs (miRNAs) and other short RNAs. lncRNAs play essential roles in a variety of physiological process, especially in the onset and progression human diseases, including liver diseases, heart diseases and may types of cancer. Yuan et al. reported that an lncRNA can be activated by transforming growth factor β (TGF-β) to promote the invasion and metastasis of hepatocellular carcinoma. In the progression of breast cancer, the nuclear factor NF-κB (NF-κB) can interact with long noncoding RNA to block the phosphorylation of inhibitor of NF-κB (IκB) to inhibit the tumor metastasis. It also found that lncRNA UCA1 can regulate the growth and apoptosis of breast cancer cells by downregulating tumor suppressive miR-143. In view of the diverse biological functions of lncRNAs and the pathogenesis of breast cancer and MMBC, it will be reasonable to speculate that specific lncRNAs may be involved in the onset and progression of MMBC.

In this study, Gene chip assay was used to identify differentially expressed lncRNAs in MMBC and UBC tissues. After validation with RT-PCR, we identified multiple MMBC-specific differentially expressed lncRNAs that can potentially serve as biomarker for the diagnosis and treatment of MMBC.

Patients and Methods

Clinical Data

This research has been approved by the Ethics Committee of the Qilu Hospital. All the patients signed the informed consent. A total of 3597 cases of invasive ductal carcinoma diagnosed between January 2005 and December 2015 in Yantai Yuhuangding Hospital were reviewed, and 156 cases of MMBC were selected by two experienced pathologists. Inclusion criteria are as follows; firstly, the presence of two or more invasive cancer lesions in the ipsilateral mammary gland, regardless of whether the lesion is located in the same segment. There is normal breast tissue between the lesions, but the distance between the lesions is not a concern. Secondly, only cases of invasive ductal carcinoma were selected for study; the cases with other types of invasive carcinoma were excluded. A total of 130 cases of UBC were also selected. All the classifications were determined according to the WHO Classification of Breast Tumors (2012 edition).

Gene Chip Assay

MMBC and UBC tissues (0.4 g) were subjected to RNA extraction using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The first strand cDNA probe was synthesized by PCR using a First Strand cDNA Synthesis Kit (Applied Biosystems, Foster City, CA, USA) and labeled with Cy3 fluorescence. After quantification and denaturation, the probe was hybridized with the gene chip (LncRNA + mRNA Human Gene Expression Microarray V4.0, capital biotech, Beijing, China) using a hybridization kit (Promega, Madison, WI, USA). After hybridization, ScanArray 4000 scanner (Packard Biochip Technologies, Billerica, CA, USA) was used to scan the gene chip. All the gene chip data analyses were performed by Capital Biotech (Beijing, China).

RT-PCR to Detect the Expression of Differentially Expressed lncRNAs

TRIzol reagent (Invitrogen, Carlsbad, CA, USA) was used to extract total RNA from the MMBC and UBC tissues according to the instruction. The reverse transcription reaction was then carried out using a kit (Promega, Madison, WI, USA) according to the instructions. The SYBR Green master kit (Applied Biosystems, Foster City, MA, USA) and 1 μl of cDNA were used to prepare the PCR reaction system.

The following primers were used: 5′-AAAGGGCCTCACAATTCAGGGG-3′ (sense) and 5′-GTGGGACTTCCATCCGTTG-3′ (anti-sense) for C19orf33; 5′-GCAGAGGTGCACGTCTTCT-3′ (sense) and 5′-TCACAGGAAACAAGGACGAC-3′ (anti-sense) for C3orf52; 5′-CGGCTCCAGATTTGGCAATG-3′ (sense) and 5′-GGGGCTGGTATCTGGAAGA-3′ (anti-sense) for C15orf48; 5′-ACTGTTCCCCCAACTCAACC-3′ (sense) and 5′-TGGGGCTGGTGATCTGGAAGA-3′ (anti-sense) for C4orf19, and 5′-GAAAGCCTGCCGGT-3′ (sense) and 5′-AGGAAAACCATCACCAGGA-3′ (anti-sense) for GAPDH. The reaction conditions: 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 45 s. The specificity of amplification was tested by melting curve analysis. The relative expression of RNA was calculated by 2^-ΔΔCt method. GAPDH was used as the endogenous control. Each sample was independently tested for three times. All the primers were synthesized by Sangon (Shanghai, China).

Statistical Analysis

The data were processed with Feature Extraction software (Santa Clara, CA, USA) and
the processed data were then analyzed by Gene Spring GX software (Santa Clara, CA, USA) to calculate the gene expression differences and p-values. The comparisons of the RT-PCR results between groups were performed by t-test and p < 0.05 was considered to be statistically significant.

**Results**

**The Differential Expression of IncRNAs in MMBC and UBC Tissues**

As shown in Scatter Plot (Figure 1A), the dots above the uppermost line and below the lowermost line represent the IncRNAs with more than 2 times' expression differences in MMBC and UBC tissues (p < 0.05). Cluster analysis (Figure 1B) showed that there were 1080 IncRNAs with significantly differential expression (FC > 4, p < 0.05), of which 458 were upregulated and 622 were down-regulated in MMBC tissue. Our data suggested that the expression of a huge number of IncRNAs was differentially regulated in MMBC and UBC tissues.

In homologous IncRNAs, the expression levels of four IncRNAs including C19orf33, C3orf52, C15orf48 and C4orf19 in MMBC tissue, were significantly different than those in UBC tissues. RT-PCR verified that the expression of C19orf33, C3orf52 and C15orf48 in MMBC tissue was significantly up-regulated and the expression of C4orf19 was significantly down-regulated, which was consistent with that of gene chip. Further analysis of IncRNAs with significantly differential expression in MMBC and UBC tissue found that, in homologous IncRNAs, four IncRNAs including C19orf33, C3orf52, C15orf48 and C4orf19, showed significantly differential expression. In MMBC tissue, C19orf33, C3orf52 and C15orf48 were up-regulated and C4orf19 was down-regulated (Table I). Gene ontology (GO) analysis showed that C19orf33 was associated with double DNA binding (GO: 0003690 double-stranded DNA binding), single strand DNA binding (GO: 0003697 single-stranded DNA binding), nucleus (GO: 0005634 nucleus) and biological process (GO: 0008150 biological process); C3orf52 is associated with the endoplasmic reticulum membrane (GO: 0016021 integral to membrane) and membrane composition (GO: 0016021 integral to membrane); C15orf48 was associated with the nucleus (GO: 0005634 nucleus), mitochondria (GO: 0005739 mitochondrion).

**RT-PCR Verification of Differential Expressed Homologous IncRNAs**

RT-PCR was used to detect the expression of IncRNA C19orf33, C3orf52, C15orf48 and C4orf19 in both MMBC and UBC tissues. Consistent with the results of the gene chip, the expression levels of C19orf33, C3orf52 and C15orf48 in MMBC tissue were 0.00039 ± 0.00004, 0.00046 ± 0.00006 and 0.00052 ± 0.00005, respectively, which were significantly higher than those in UBC tissue (0.00018 ± 0.00005, 0.00013 ± 0.00004 and 0.00016 ± 0.00003, respectively). The expression level of C4orf19 in MMBC was 0.00012 ± 0.00004, which was significantly lower than that in UBC tissue (0.00044 ± 0.00005, Figure 2). Our data indicate that the results of gene chip were reliable.
The identification of multifocal breast cancer-associated long non-coding RNAs

Discussion

Breast cancer is not a homogeneous disease but a group of heterogeneous diseases with different morphologic, clinical and molecular features. Various vulnerable groups suffered from it. Breast cancer affects 1 in 8 women and about 1/3 of the patients used to die from the disease. Multiple biochemical markers have been identified for the diagnosis and treatment of breast cancer. Estrogen receptor (ER) plays pivotal roles in estrogen-dependent growth, and the changes of ER can be used to accurately predict the outcomes of the hormone therapy and the prognosis of ER-positive breast cancer. The immunohistochemical testing of progesterone receptor (PgR), which is an estrogen-responsive gene, is also routinely performed in the treatment of ER-positive breast cancer. Besides those two routinely used hormone receptors, breast cancer 1 (BRCA1) and BRCA2 were also identified as novel maker for the prediction of patients with high risk of breast cancer. Previous studies have shown that women with mutation in BRCA1 or BRCA2 genes have a risk of breast cancer up to 45-80% during their whole life. The breast cancer cells with BRCA mutation usually lack the DNA-repair mechanism, which is responsible for the error-free DNA double-strand break repair. Therefore, tumor will occur due to the error prone replication. A recent study which carried out a genomic patterns resembling of breast cancers with BRCA1- and BRCA2-mutations using array comparative genomic hybridization (aCGH) have shown that aCGH can be used for the effective identification of breast cancer patients with

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Probe sequence</th>
<th>Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>C19orf33</td>
<td>Chromosome 19 open reading frame 33</td>
<td>AAGAAGAAGAAGGCTCCCCACTGAAGGGCCCTGGACAGGGCTCATTAAACCTTCCTCTTCT</td>
<td>Up</td>
</tr>
<tr>
<td>C3orf52</td>
<td>Chromosome 3 open reading frame 52</td>
<td>ACCAGGAGAAAGAGTTTTGTGCTGTATTGTGAGAGATCTCGCCTCTCAGTTAAATGAGCC</td>
<td>Up</td>
</tr>
<tr>
<td>C15orf48</td>
<td>Chromosome 15 open reading frame 48</td>
<td>AGTGTGCGGAAATGCTTCTGCTACATTTTTAGGGTTTGTCTACATTTTTITGGGCTCTGGA</td>
<td>Up</td>
</tr>
<tr>
<td>C4orf19</td>
<td>Chromosome 4 open reading frame 19</td>
<td>ACACCTTGACCTTAACTTTTTTACATTATTACTTTTAAATCTCCTTTGGGATTGGGGAGG</td>
<td>Down</td>
</tr>
</tbody>
</table>

Figure 2. The expression levels of lncRNA C19orf33, C3orf52, C15orf48 and C4orf19 detected by RT-PCR.
ER-positive, triple-negative and HER2-negative breast cancer tumors, which can be treated with the intensified, DSB-inducing chemotherapy, and the early diagnosis significantly improved the treatment outcomes. Aside from above-mentioned well-established biomarkers in the diagnosis and treatment of breast cancer and the prediction of outcomes, the specific markers for MMBC are still lack. The development of preoperative breast imaging techniques has greatly improved the identification of MMBC. However, those methods can only be used to identify visible tumors but lack the ability to identify MMBC before the tumor onset. Compared with UBC, more lymph nodes were involved in MMBC, and previous studies have shown that the treatment of MMBC with different therapies usually provides worse overall outcomes compared with the treatment of UBC. Therefore, the identification of early diagnosis biomarkers is urgently needed.

LncRNAs is a group of recently identified non-coding RNAs with diverse functions in various physiological and biochemical pathways. Ge et al. reported that lncRNA PCAT-1 is associated with the progression of colorectal cancer and the overexpression of this lncRNA is a biomarker for the poor prognosis of patients with colorectal cancer. In the study of pancreatic cancer, the abnormal higher expression level of lncRNA HULC was found to be positively correlated with the tumor size; therefore, lncRNA HULC is a novel biomarker of the poor prognosis of this disease. In the diagnosis of non-small cell lung cancer, the detection of lncRNA MALAT1 in blood was found to be able to serve as a reference. Recent studies have shown that IncRNAs are also involved in the regulation of tumor metastasis during the progression of breast cancer, indicating the potential of IncRNAs to serve as markers for the diagnosis of breast cancer. However, to our knowledge there are no studies on the functions of IncRNAs in MMBC have been reported.

In our paper, gene chip technique was utilized to identify the differentially expressed IncRNAs in MMBC and UBC tissues. Gene chip technique, which can provide a genome-wide view with a single experiment, is a powerful tool for genomic study. In previous researches, gene chip has been used to differentially expressed gene in different tissues to identify pathway-specific or disease-specific gene expression. Based on the gene chip data, 110 out of 14500 known genes were found to be associated with lymphatic metastasis in mouse. In our work, gene chip data analysis of 156 cases of invasive MMBC and 130 cases of UBC showed that 458 IncRNAs were upregulated and 622 were downregulated in MMBC tissue compared with UBC tissues. In addition, the differential expression of homologous IncRNA C19orf33, C3orf52, C15orf48 and C4orf19 were validated by RT-PCR. Our data suggested that the expression regulation of IncRNAs was related to the development of MMBC.

Conclusions

We showed that a huge number of LncRNAs including C19orf33, C3orf52, C15orf48 and C4orf19 were differentially expressed in MMBC and UBC tissue. So, those IncRNA may have important functions in the development of MMBC and may serve as markers to distinguish MMBC from UBC.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The Authors declare that they have no conflict of interests.

References

The identification of multifocal breast cancer-associated long non-coding RNAs


