Non-Coding CK19 RNA in Peripheral Blood and Tissue of Breast Cancer Patients
Mana Oloomi, Najmeh Yardehnavi, Saeid Bouzari, and Neda Moazzezy
Department of Molecular Biology, Pasteur Institute of Iran, Tehran, Iran

Abstract- Breast carcinoma is the major cause of cancer-related death in women. The incidence of this carcinoma is rising and there are many attempts to decrease this problem. The aim of this study was detection of full-length cytokeratin 19 (CK19) mRNA, in peripheral blood and tissue of breast cancer patients in early stage of cancer. In this study, RT-PCR (reverse transcriptase-polymerase chain reaction) technique was used for detection of CK19 mRNA in peripheral blood and tissue of breast cancer patients. Primers were established to amplify the CK19 as a tumor marker. Moreover, CYFRA 21-1 subunit of CK19 protein was measured in the serum of patients. CK19 mRNA was detected and sequenced. It is shown that the most released CK19 mRNAs in blood and tissue of cancer patients are non-coding RNA. The mutated forms of mRNA are the incomplete transcripts of protein-coding gene as a long non-coding RNA (lncRNA) that could regulate gene expression. Moreover, small non-coding RNA (ncRNA) as fragments of CK19 is mostly observed in this experiment. They may play a role in tumorigenesis and their biologic exact function in breast cancer should be further elucidated.

© 2013 Tehran University of Medical Sciences. All rights reserved.

Keywords: Breast cancer; CK19; Non-Coding RNA; Tumor marker

Introduction

Breast cancer is known as a well characterized, rising and life threatening disease in the world (1-3). In the early stages of breast cancer, tumor cells are secreted into the blood (4). These circulating tumor cells (CTCs) are clinically important and they are hematogenous rout for metastasis (5). In this regard, finding an appropriate tumor marker is very crucial (6). Tumor marker is defined as a molecule that changes in non-cancerous (benign) or cancerous conditions qualitatively or quantitatively and this changes could be detected. Tumor markers usually are nucleic acids (DNA or RNA) or protein molecules. They can be applied in cancer diagnosis, risk estimation of recurrence or death, prediction and monitoring response to anticancer drugs (7). In an ideal condition, a diagnostic tumor marker must be highly sensitive and specific for the kind of tumor (8).

One of the most interesting tumor markers in breast cancer is CK19 (cytokeratin19) (9). CK19 is belonging to cytokeratin family (CKs) that owned string biopolymers in the cytoskeleton of eukaryotic cells (10). CK19 proteins is expressed on simple and stratified epithelium and separated from it in different cancer disease (11,12). It is reported that CK19 is a sensitive and specific marker for breast cancer (9). It can also be detected in both RNA and protein forms. RT-PCR (reverse transcriptase-polymerase chain reaction) is reported as a useful method for detection of CK19 RNA even in the early stage of breast cancer (13).

On the other hand, At the protein level, a piece of CK19 protein in the sera, called CYFRA 21-1, can be measured by a sandwich ELISA assay (CYFRA21-1 ELISA) using specific monoclonal antibodies for CK19 protein (11). Both fragment and full-length CK19 proteins have been detected in breast cancer cell lines and bone marrow of breast cancer patients with immunoassay tests (12).

In this study, CK19 molecular marker in blood and tissue of breast cancer patients was detected. Detection of CK19 marker, as a protein fragment, and also as small and long full-length RNA fragment was assessed.
Materials and Methods

Blood and tissue samples
From all patients and healthy donors, written consent letter was obtained for participation in the experiment. The local ethical committee of Pasteur Institute of Iran approved the protocol. Breast cancer patients were referred from Tehran, Milad hospital. Peripheral blood and tissue samples were obtained from 32 patients with primary breast cancer (stage I) before chemotherapy. For each patient 5 ml of the peripheral blood and tissue samples was collected. Normal blood samples from 32 healthy individuals were also obtained. The blood was aliquoted into two glass tubes. In one tube, 250 µl sodium citrate (0.032 g/ml) was added; this tube was used for RNA extraction. The second tube, which had no sodium citrate, was used for serum isolation.

CYFRA 21-1 ELISA
The blood samples were allowed to be clotted. They were centrifuged at 4000 rpm for 15 minutes for serum isolation. The serum was isolated and stored at -80 °C. CYFRA 21-1 was measured by ELISA (CanAg, FUJIREBIO, Diagnostic Inc.) according to the manufacturer’s instructions.

Total RNA extraction from blood and tissue
Total RNA was isolated by AccuZol reagent (Bioneer Inc, Seoul) according to the manufacturer’s instructions. The isolated RNA was dissolved in distilled water and stored at -80°C until use. The concentration of RNA was measured at 260 nm with Picodrop (Ltd Cambridge, UK). RNA integrity was controlled by RT-PCR using primers for Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) amplification.

One Step RT-PCR assay
Reverse transcription of RNA and PCR amplification was carried out using one-step AccuPower™ RT-PCR PreMix kit (Bioneer, Inc. Seoul). The RT-PCR reaction was carried out in an Eppendorf DNA thermal cycler. The sequences for each set of primers (GAPDH, CK19 full-length and fragment) are given in table 1. The RT-PCR reaction was done according to the manufacture instructions. In summary, 1 µg of total RNA template and 10 pmol of reverse primer were mixed and incubated for 5 min at 70 °C. The mixture was placed on ice. The incubated mixture and the forward primer were transferred to AccuPower™ RT-PCR PreMix tube and fill up with diethylpyrocarbonate (DEPC), Distilled water (DW) up to total volume of 20 µl. After vortex mixing, cDNA synthesis was performed at 42 °C for 60 min. Reverse transcriptase (RTase) inactivation was done at 94 °C for 5 min. The cycling protocol of PCR reaction consisted of a 94 °C for 1 min (denaturation), 56°C for 30 sec (annealing) and 72°C for 1 min (extension), repeated for 30 cycles. The final extension was carried out at 72 °C for 10 minutes.

To ensure that genomic DNA is not in RNA samples, total RNA extraction samples have passed a PCR reaction instead of RT-PCR reaction. They didn’t show any CK19 band.

Two step RT-PCR by Pfu polymerization

**Step 1- cDNA synthesis:** Extracted total RNA (1 µg) was used in Accupower® Cyclescript RT PreMix kit (Bioneer, Inc. Seoul). The reverse transcription reaction was composed of 3 steps: Step1 was performed at 25 °C, 30 sec, for primer annealing. Step 2 was performed at 45 °C, 4 min for cDNA synthesis. The step 3 was done at 55 °C, 30 sec for melting secondary structure of RNA template and cDNA synthesis. These steps repeated 12 times and then, the heat inactivation step performed at 95 °C, for 5 min.

**Step2- PCR amplification:** PCR amplification was performed by Accupower® Pfu PCR PreMix kit (Bioneer, Inc. Seoul) to secure high-fidelity PCR products. P5 and P6 primers were used and the total volume of the reaction mixture was 50 µl. In summary, 100 ng of cDNA template and 10 pmol of P5 and P6 primers were added to Accupower® Pfu PCR PreMix tube.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Nucleotide Sequence</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1: GAPDH-F</td>
<td>5'-GGTCGGAATCAACGGATTTG-3'</td>
<td>318bp</td>
</tr>
<tr>
<td>P2: GAPDH-R</td>
<td>5'-ATGAGGCCCAGCCTTCTCCAT-3'</td>
<td></td>
</tr>
<tr>
<td>P3: Fragment F</td>
<td>5'-ATGAAAAAGCTGCTTGGAAAGA-3'</td>
<td>136bp</td>
</tr>
<tr>
<td>P4: Fragment R</td>
<td>5'-TGATTCTGCGCCATCACTGCAG-3'</td>
<td></td>
</tr>
<tr>
<td>P5: Full-length F</td>
<td>5'-CGGAGATCCCATGACTTCTCTAGTATGC-3'</td>
<td>1201bp</td>
</tr>
<tr>
<td>P6: Full-length R</td>
<td>5'-GGAATTCCTCAGAGGACCTTGAGGAGCAG-3'</td>
<td></td>
</tr>
</tbody>
</table>

F: forward primer  R: reverse primer
Amplification was carried out with a Mastercycler personal eppendorf® AG, Hamburg, Germany. PCR reaction was initiated with a 5 min denaturation at 94 °C and terminated with a 15 min extension at 72 °C. The cycling protocol included of denaturation at 94 °C for 30 sec, annealing at 56 °C for 40 sec and extension at 72 °C for 1 min, repeated for 30 cycles.

Tumor marker detection was evaluated by both one and two step RT-PCR amplification to confirm the real polymerization of mRNA in blood and tissue of subjects.

Samples were also subjected to amplification for full-length CK19 marker, with the same primers. This PCR was done to avoid positive RT-PCR due to pseudogenes amplification.

**Sequencing**

The RT-PCR products from two-step assay were sequenced by Bioneer Inc. The nucleotide sequence of RT-PCR products was confirmed by sequencing and PCR amplification using Pfu polymerase (Fermentas) in both rounds was done to secure high-fidelity PCR products.

**Statistical analysis**

The relationship between positive results of CK19 in patients and normal subjects were subjected to ANOVA and student’s t-test for statistical analysis and a P-value less than 0.05 was considered significant.

**Results**

CYFRA 21-1 fragments of CK19 protein was measured in patient and normal sera. The results for CYFRA 21-1 fragment indicated that there was no significant difference between CYFRA 21-1 fragments of CK19 in sera of normal and patients (P=0.738) (Data not shown).

At first, total RNA extraction of subjects was used for GAPDH RT-PCR by P1-P2 primers (Figure 1a). Thereafter, the positive subjects are used for specific RT-PCR by specific primers for CK19 gene. CK19 fragment was detected in blood of patients and normal subjects by using P3 and P4 primers (Figure 1a). Expression of CK19 fragment in patients was statistically significant (P=0.019). The results are shown in table 2.

On the other hand, different RNA expression of CK19 in blood and tissue of patients were observed. Expression of CK19 from blood and tissue was compared. The results are summarized in table 3.

**Figure 1.** RT-PCR products

a) Lane 1= GAPDH band, M= DNA molecular size marker (50bp marker), Lane 2= CK19 fragment

b) Lane 1= CK19 full-length, M= DNA molecular size marker (1kb marker)
Table 2. Detection of CK19 fragment in blood of patients and normal subjects.

<table>
<thead>
<tr>
<th>Blood samples</th>
<th>CK19 Fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Normal</td>
<td>27</td>
</tr>
<tr>
<td>Patient</td>
<td>32</td>
</tr>
</tbody>
</table>

Chi-square=5.424, P-value = 0.0199 (significant)

Figure 2. Detection of CK19 fragment is shown in patients and normal blood.

Table 3. The comparison of CK19 marker in blood and tissue of patients.

<table>
<thead>
<tr>
<th>Patient samples</th>
<th>Fragment CK19 positive</th>
<th>Full-Length CK19 positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood</td>
<td>32</td>
<td>9</td>
</tr>
<tr>
<td>Tissue</td>
<td>32</td>
<td>22</td>
</tr>
</tbody>
</table>

Chi-square=3.743, P-value=0.053 (non-significant)

There was no statistically significant difference between CK19 expression in blood and tissue of breast cancer patients (P=0.053).

In figure 3, CK19 small and long full-length fragment expression in blood and tissue subjects is demonstrated.

Figure 3. Full length and fragment of CK19 detection in blood and tissue of patients.

The same expression of CK19 fragment was shown in blood and tissue of patients but CK19 full-length expression in tissue was higher than blood. This difference was not statistically significant; therefore the expression of full-length CK19 in tissue is the same as blood. There was no expression for full-length CK19 in normal patients.

The sequences of expressed CK19 gene in blood and tissue samples were also compared. Alignment was done by multiple alignments using vector NTi-11.5 Advance® software.

The CK19 gene was sequenced from blood of five patients; one tissue was also sequenced. The schematic of the sequenced mutated genes, aligned with CK19 gene is represented in figure 4.
<table>
<thead>
<tr>
<th></th>
<th>Blood3</th>
<th>Blood4</th>
<th>Blood5</th>
<th>Tissue</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>101-150</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGCGGCGGCTCCATGCGTTTTGGGGCAGGGGTCGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
</tr>
<tr>
<td>151-200</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
<td>Blood5</td>
</tr>
<tr>
<td>201-250</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>251-300</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>301-350</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
</tbody>
</table>

**Consensus**

<table>
<thead>
<tr>
<th></th>
<th>Blood3</th>
<th>Blood4</th>
<th>Blood5</th>
<th>Tissue</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>101-150</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGCGGCGGCTCCATGCGTTTTGGGGCAGGGGTCGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
</tr>
<tr>
<td>151-200</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
<td>Blood5</td>
</tr>
<tr>
<td>201-250</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>251-300</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>301-350</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
</tbody>
</table>

**Consensus**

<table>
<thead>
<tr>
<th></th>
<th>Blood3</th>
<th>Blood4</th>
<th>Blood5</th>
<th>Tissue</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>101-150</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGCGGCGGCTCCATGCGTTTTGGGGCAGGGGTCGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
<td>CCTGGGTGGTGGCTCCGTGAGTTTTGTGGCAGAGGTTGCCTTTCGCGCGC</td>
</tr>
<tr>
<td>151-200</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
<td>Blood5</td>
</tr>
<tr>
<td>201-250</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>251-300</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
<tr>
<td>301-350</td>
<td>CK19</td>
<td>Blood1</td>
<td>Blood2</td>
<td>Blood3</td>
<td>Blood4</td>
</tr>
</tbody>
</table>
Non-coding CK19 RNA in breast cancer

351
CK19 (351) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
Blood1 (343) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
Blood2 (343) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
Blood3 (343) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
Blood4 (343) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
blood5 (351) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
Tissue (343) GAAGCAGGGGCGCTGCGACTACAGCCACTACAGCAGCA
Consensus (351) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA

401
CK19 (401) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
Blood1 (392) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
Blood2 (392) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
Blood3 (392) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
Blood4 (392) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
blood5 (401) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
Tissue (392) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA
Consensus (401) TCCAGGACCTGGGACTACAGCCACTACAGCAGCA

451
CK19 (451) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
Blood1 (442) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
Blood2 (442) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
Blood3 (442) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
Blood4 (442) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
blood5 (451) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
Tissue (442) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA
Consensus (451) ATTGTCCTGGAGATCGACAAGCCCTGGTACAGCTACAGCAGCA

501
CK19 (501) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
Blood1 (492) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
Blood2 (492) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
Blood3 (492) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
Blood4 (492) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
blood5 (501) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
Tissue (492) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA
Consensus (501) CAAGAGTGAGACGGAGCAGGCTCTGCGCATGAGCGCTACAGCAGCA

551
CK19 (551) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
Blood1 (542) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
Blood2 (542) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
Blood3 (542) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
Blood4 (542) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
blood5 (550) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
Tissue (542) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA
Consensus (551) ACGGCCTGCAGAGGGTGCTGGATGAGCTGACTACCGACTACAGCAGCA

601
CK19 (601) GAGATGCAGATCGAAGGGCTGGATGAGCTGACTACCGACTACAGCAGCA
Blood1 (592) GAGATGCAGATCGAAGGGCTGGATGAGCTGACTACCGACTACAGCAGCA
Blood2 (592) GAGATGCAGATCGAAGGGCTGGATGAGCTGACTACCGACTACAGCAGCA

Blood3 (592) GAGATGCAGATCTAAGGCCTGAAGGAAGAGCTGGCCTACCTGAAGAAGAA
Blood4 (592) GAGATGCAGATCTAAGGCCTGAAGGAAGAGCTGGCCTACCTGAAGAAGAA
blood5 (597) AGAACGTTGATGTGGATCTCCAGGCTCAGCGGGGCCAGCTCCTCAAACT
Tissue (592) GAGATGCAGATCTAAGGCCTGAAGGAAGAGCTGGCCTACCTGAAGAAGAA
Consensus (601) GAGATGCAGATCTAAGGCCTGAAGGAAGAGCTGGCCTACCTGAAGAAGAA

701                                            750

CK19 (701) GTGTGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT
Blood1 (692) GTGGGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT
Blood2 (692) GTGGGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT
Blood3 (692) GTGGGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT
Blood4 (692) GTGGGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT
blood5 (696) ATCTCGAGAATGGCTGGCAGTGAAATCTTGTCCCATAGGTCCTGATG
Tissue (692) GTGGGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT
Consensus (701) GTGGGGAGGTGGATTCGGCTCAGGGCACCTATCTCGCCAAGATCCTGAGT

751                                            800

CK19 (751) GACATGCGAAGCCAATATGAGGTCATGGCCGAGCAGAACCGGAAGGATGC
Blood1 (742) TACATGCGAAGCCAATACGAGGTCATGGCGGAGCAGAACTGGAAGGATGC
Blood2 (742) TACATGCGAAGCCAATACGAGGTCATGGCGGAGCAGAACTGGAAGGATGC
Blood3 (742) TACATGCGAAGCCAATACGAGGTCATGGCGGAGCAGAACTGGAAGGATGC
Blood4 (742) TACATGCGAAGCCAATACGAGGTCATGGCGGAGCAGAACTGGAAGGATGC
blood5 (746) GGCTTGAGTAGTGGGTGTAAGCCCATGGCCGAGCAGAAACCAGGAGATGG
Tissue (742) TACATGCGAAGCCAATACGAGGTCATGGCGGAGCAGAACTGGAAGGATGC
Consensus (751) TACATGCGAAGCCAATACGAGGTCATGGCGGAGCAGAACTGGAAGGATGC

801                                            850

CK19 (801) TGAAGCCTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT
Blood1 (792) TGAAGGCTTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT
Blood2 (792) TGAAGGCTTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT
Blood3 (792) TGAAGGCTTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT
Blood4 (792) TGAAGGCTTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT
blood5 (796) TACAGCCTGTTCCCCACCTACAGTTGACCTCTTCTCGGCGTCTACCT
Tissue (792) TGAAGGCTTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT
Consensus (801) TGAAGGCTTGGTTCACCAGGCCGAGCAGACTGAAGATTTGAACCCGGAGGCTCGT

851                                            900

CK19 (851) GCCACACGAGCAGCTCAGAGTACGAGCGGCTAGTCGAACTGACCTGG
Blood1 (842) GCCACACAGATCACGTCCAGATGACCGCAGTGGCTAAGGCTGACTGCGG
Blood2 (842) GCCACACAGATGATCCTAGATGACCGCAGTGGCTAAGGCTGACTGCGG
Blood3 (842) GCCACACAGATGATCCTAGATGACCGCAGTGGCTAAGGCTGACTGCGG
Blood4 (842) GCCACACAGATGATCCTAGATGACCGCAGTGGCTAAGGCTGACTGCGG
blood5 (846) GCCAGCAGGAGGAGCAGTCCAGAGTACGAGCAGTGGCTAAGGCTGACTGCGG
Tissue (842) GCCACACAGATGATCCTAGATGACCGCAGTGGCTAAGGCTGACTGCGG
Consensus (851) GCCACACAGATGATCCTAGATGACCGCAGTGGCTAAGGCTGACTGCGG
Non-coding CK19 RNA in breast cancer

CK19 (901) CGCACCCTTCAGGGTCTTGAGATTGAGCTGCAGTCACAGCTGAGCATGAA
Blood1 (892) CGCACCCTTCAGGGTCTTGAG------CTGCAGTCACAGCTGAGCATGAA
Blood2 (892) CGCACCCTTCAGGGTCTTGAG------CTGCAGTCACAGCTGAGCATGAA
Blood3 (892) CGCACCCTTCAGGGTCTTGAG------CTGCAGTCACAGCTGAGCATGAA
Blood4 (892) CGCACCCTTCAGGGTCTTGAG------CTGCAGTCACAGCTGAGCATGAA
blood5 (896) C-CTACCAGCAGCCCATTGGACCTGGCCAGCATTCACAGCTGCCACTGTA
Tissue (892) CGCACCCTTCAGGGTCTTGAG------CTGCAGTCACAGCTGAGCATGAA
Consensus (901) CGCACCCTTCAGGGTCTTGAG CTGCAGTCACAGCTGAGCATGAA

951

CK19 (951) AGCTGCCTTGGAAGACACACTGGCAGAAACGGAGGCGCGCTTTGGAGCCC
Blood1 (936) AGCCGCCTTGGAAGCCACACTGGCAGAAACGGAGGCGCGCTTTGGAGTCC
Blood2 (936) AGCCGCCTTGGAAGNNNCACTGGCAGAAACGGAGGCGCGCTTTGGAGTCC
Blood3 (936) AGCCGCCTTGGAAGCCACACTGGCAGAAACGGAGGCGCGCTTTGGAGTCC
Blood4 (936) AGCCGCCTTGGAAGCCACACTGGCAGAAACGGAGGCGCGCTTTGGAGTCC
blood5 (945) AGCTGGCCTTGGAAGACACACTGGCAGAAATGGCAGATTAAATGCCAGCTGCC
Tissue (936) AGCCGCCTTGGAAGCCACACTGGCAGAAACGGAGGCGCGCTTTGGAGTCC
Consensus (951) AGCCGCCTTGGAAGCCACACTGGCAGAAACGGAGGCGCGCTTTGGAGTCC

1001

CK19 (1001) AGCTGGCGCATATCCAGGCGCTGATCAGCGGTATTGAAGCCCAGCTGGGC
Blood1 (986) AGCTGGCGCAGATCCAGCCGCTGATCAACTGTATTGAAGCCCAGCTGGGC
Blood2 (986) AGCTGGCGCAGATCCAGCCGCTGATCAACTGTATTGAAGCCCAGCTGGGC
Blood3 (986) AACTGGCGCAGATCCAGCCGCTGATCAACTGTATTGAAGCCCAGCTGGGC
Blood4 (986) AACTGGCGCAGATCCAGCCGCTGATCAACTGTATTGAAGCCCAGCTGGGC
blood5 (995) AACAGAGGCATATACTGAGCAGCAGTGGCCCATGGAAGTGTTGAGAT
Tissue (986) AGCTGGCGCAGATCCAGCCGCTGATCAACTGTATTGAAGCCCAGCTGGGC
Consensus (1001) AGCTGGCGCAGATCCAGCCGCTGATCAACTGTATTGAAGCCCAGCTGGGC

1051

CK19 (1051) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGAGTACCAGCGGCTCAT
Blood1 (1036) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC
Blood2 (1036) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC
Blood3 (1036) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC
Blood4 (1036) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC
blood5 (1045) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC
Tissue (1036) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC
Consensus (1051) GATGTGCGAGCTGATAGTGAGCGGCAGAATCAGGATTAACAGCAGTTTAC

1101

CK19 (1101) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
Blood1 (1086) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
Blood2 (1086) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
Blood3 (1086) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
Blood4 (1086) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
blood5 (1095) TGGCACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
Tissue (1085) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC
Consensus (1101) GGACATCAAGTCCGGGCCTGGGAGCAGAGATTTGCCACCTACCACGGCGGCCTGC

1151

CK19 (1151) TCGAGGGACAGGAAGATCACTACAACAAATTTTGCTCTCCAAAGCTTCTC
Blood1 (1136) TCGAGGGACAGGAAGATCACTACAACAAATTTTGCTCTCCAAAGCTTCTC
Blood2 (1136) TCGAGGGACAGGAAGATCACTACAACAAATTTTGCTCTCCAAAGCTTCTC
Blood3 (1136) TCGAGGGACAGGAAGATCACTACAACAAATTTTGCTCTCCAAAGCTTCTC
Blood4 (1136) TCGAGGGACAGGAAGATCACTACAACAAATTTTGCTCTCCAAAGCTTCTC
blood5 (1145) AGGAAGTACATGGGGATCCCGAATCGGACCGGGCCGACTGCAGAGGCCTC
Tissue (1135) TCGAGGGCCAGAAAGATCACTACAACAACCTGTCCCTCCAAGGTCTCTC
Consensus (1151) TCGAGGGCCAGAAAGATCACTACAACAACCTGTCTGCCTCCAAGGTCTCTC

1201
CK19 (1201) TGA
Blood1 (1186) TGA
Blood2 (1186) TGA
Blood3 (1186) TGA
Blood4 (1186) TGA
blood5 (1195) TGA
Tissue (1185) TGA
Consensus (1201) TGA

**Figure 4.** Schematic representation of newly reported non-coding *CK19* genes alignment (CK19-NM-002276.4). Residues in an alignment are colored according to vector-NTI-Advance-11 users Manual.

Blue on cyan: consensus residue derived from a block of similar residues at a given position
Yellow: consensus residue derived from a completely conserved residue at a given position

Furthermore, existence of two stop codons very nearly after the first codon (ATG) leading to no cytokeratin expression (Figure 5). Two nucleotide substitution, C A was observed. The existence of stop codons in the first part of CK19 RNA results non-coding CK19 RNA.

**Figure 5.** Chromatogram of non-coding CK19.
Non-coding CK19 RNA in breast cancer

Discussion

Cytokeratins, also known as keratins, make the cytoplasmic intermediate filaments. These filaments are existed in normal and malignant epithelial cells. There are more than twenty different cytokeratins in human epithelia. Cytokeratin 19 is type I of human cytokeratins. CK19 is known as a tumor marker and its mRNA indicates the presence of circulating tumor cells (CTCs) in some cancer especially breast cancer (5,14). Detectable CTCs, which are used as a novel prognostic marker, have been found by RT-PCR for CK19 in 20-40% of early breast cancer patients (5,15). CK19 has been diagnosed in tumor cells and bone marrow samples from breast cancer patients using immunoassay (12). CYFRA 21-1 protein, a known CK19 fragment protein, has also been used as a tumor marker in some cancers for example lung and breast cancer (16). It is supposed that full-length mRNA can be translated to protein (17).

The aim of this study was establishment of full-length CK19 mRNA as a tumor marker in blood and tissue of breast cancer. In this regard, patients were assessed for CYFRA 21-1 protein detection in the sera. The mechanism of the release of CK19 protein fragment (CYFRA 21-1) is a proteolytic process. During apoptosis, CK19 protein is cleaved by caspase 3 and produces soluble CYFRA21-1 fragment. This fragment could be measured by immunoassay method using two monoclonal antibodies (12, 18). Statistical analysis of our data indicates that there is no significant difference between CYFRA 21-1 fragment in patient and healthy serum samples. This means that CYFRA 21-1 detection is not important alone, in breast cancer. Probably, it is important when other markers and clinical history are evaluated (19). On the other hand, statistical analysis of RT-PCR results indicates that there are significant difference for CK19 fragment in blood of breast cancer patients and healthy subjects. However, other studies showed variable results for CK19 RT-PCR assays. In this study, there is no correlation between the level of CK19 mRNA fragment and CYFRA21-1 protein. That is similar to the results of Marrakchi et al. (20) because of post-transcriptional and post- translational mechanisms, mRNA level dose not always show the exact level of protein (21). On the contrary, Fujita and colleagues (22) found that there are very close relationship between the rate of CK19 mRNA and the amount of CYFRA 21-1 protein in lung cancer. However, at mRNA level, there is an incomplete mRNA unable to produce CYFRA21-1 protein. Incomplete RNA may interfere with RT-PCR results; underestimate the level of circulating tumor cells.

In addition, it has been shown that some mutations in promoter region of CK19 gene can cause down regulation of mRNA CK19 expression (17,22). Multiple alignment analysis of sequences shows there are mutations in full-length CK19 in comparison with mRNA CK19 gene (Homo sapiens keratin 19 [KRT19], NCBI Reference Sequence: NM-002276.4). It has also been observed that there are mutations in CK19 gene in lung cancer (17). It is likely that some mutations in CK19 gene at genomic level, leading to the production of defective keratin filaments (17). These sequence analysis also indicates that obtained sequences in this research are very identical to sequences of known CK19 pseudogenes (CK19a: accession No. M33101, CK19b: accession No. U85961). The presence of pseudogene has been reported previously by Ruud et al.; Submitted CK19 pseudogenes have high homology with CK19 mRNA and therefore it is supposed that they may interfere with RT-PCR leading to cause false-positive results (23). It is speculated that the source of these pseudogenes may be DNA contamination. The false positive signals had been remained even after using DNase in RNA extraction (23,24).

Moreover, some part of CK19 mutated gene in this research, has 88% identity with a known microRNA (Homo sapiens microRNA-492 (MIR492), NCBI Reference Sequence: NR-030171.1). It has been demonstrated that many of pseudogenes are transcribed to RNA as processed pseudogenes. The synthesizes of complete genes could regulate by these incomplete RNAs. On the other hand, pseudogenes may also adjust tumor suppressors and oncogenes (25). Recently, it has been reported that microRNA-492 can derive from CK19 gene and co-express with CK19 in metastatic hepatoblastoma tumor (26). Moreover, exRNAs is also identified as free nucleic acids, released spontaneously from tumor cells and the exact mechanism remains to be fully elucidated (27). It has also been proved that CK19 positive cells may have stem cell-like characteristics and CK19 is categorized as a putative stem cell marker (12,28,29). Stem cells have been considered as a model to study the role of microRNA (30). More additional tests are needed to prove the existence of CK19 microRNAs in breast cancer and to explain the mechanism of their action from converting healthy cells to tumor cells. In conclusion, CK19 biomarker increases significantly in breast cancer patients. This marker is detectable by RT-PCR assay in peripheral blood or tissue samples. Also, there is not any correlation.
between CYFRA21-1 protein and CK19 mRNA in women with breast cancer. The presence of newly non-coding CK19 is reported for the first time in this study. It might play a regulatory role in CK19 expression of the complete gene. However, it has already been suggested that the expressed processed pseudogenes could regulate coding gene expression as a non-coding function for mRNAs (31). Finally, regulation role of non-coding mRNAs in tumor biology is not clarified yet and needs to be more elucidated.

Acknowledgment

This work was financially supported by Pasteur Institute of Iran. We are grateful to Drs. Hashemian and Abasvandi at Milad Hospital (Tehran, Iran).

References

Non-coding CK19 RNA in breast cancer


