Genetic polymorphisms in the estrogen receptor-α Gene codon 325 (CCC→CCG) and risk of breast cancer among Iranian women: a case control study

Cyrus Azimi¹, MD., PhD., Sakineh Abbasi², PhD.

Departments of Medical Genetics and Medical Laboratory Sciences, Tehran University of Medical Sciences, Tehran, Iran.

Abstract

Background: The Iranian breast cancer patients are relatively younger than their Western counterparts. Evidence suggests that alterations in estrogen signaling pathways, including estrogen receptor-α (ER-α), occur during breast cancer development in Caucasians. Epidemiologic studies have revealed that age-incidence patterns of breast cancer in Asians differ from those in Caucasians. Genomic data for ER-α in either population is therefore of value in the clinical setting for the Iranian breast cancer.

Methods: A case-control study was conducted to establish a database of ER-polymorphisms in Iranian women population in order to compare Western and Asian with Iranian (Asian-Caucasians) distributions and to evaluate ER- polymorphism as an indicator of clinical outcome. DNA samples were prepared from Iranian women with breast cancer referred to Imam Khomeini Hospital Complex clinical breast cancer group (150 patients) and in healthy individuals (147 healthy control individuals). PCR single-strand conformation polymorphism technology was performed.

Results: A site of silent single nucleotide polymorphism (SNP) was found, as reported previously in Western and Eastern studies, but at significantly different frequencies. The frequency of allele 1 in codon 325 (CCC→CCG) was significantly higher in the breast cancer patients (39.6%) than control individuals (28.9%; P = 0.007). The allele 1 had also significant association with the occurrence of lymph node metastasis.

Conclusion: Data suggested that ER-α polymorphisms in exon 4 codon 325 was correlated with various aspects of breast cancer in Iran. ER-α genotype, as determined during presurgical evaluation, might represent a surrogate marker for predicting the breast cancer lymph node metastasis.

Keywords
breast cancer, estrogen receptor-α, polymorphism, PCR-SSCP, Lymph node metastasis.

Introduction

Breast cancer is the leading cancer among women in most parts of the world, including Iran. Geographical variations in incidence and mortality rates of breast cancer suggest that the known risk factors for breast cancer may vary in different parts of the world and that environmental factors are of greater importance than genetic factors [1]. For instance, in Iran it has been shown that, even after adjusting for age, young women are at relatively higher risk for...
developing breast cancer than are their Western counterparts [2,3].

It is known that breast cancer typically arises in luminal epithelial cells of the mammary gland [4,5]. These cells contain estrogen receptors (ERs), which respond to ovarian estrogen in normal mammary gland development. How estrogens stimulate cell growth is not fully understood, but it is known that estrogen activation of ER results in transcription of various genes that are involved in cellular proliferation. It has been shown that exposure to estrogen correlates with risk for breast cancer, with the risk increasing with duration of exposure [6]. It has been found that ERs are variably present in breast tumors, and that patients expressing ERs are more responsive to hormone treatment [7], making immunohistological assay of ER expression in tumor tissue a widely applied clinical technique. The ER-α, which is expressed in luminal epithelium-derived normal or cancerous cells but not in any of the other stromal cell types within the human breast [8], has been proposed to participate in breast carcinogenesis.

The ER-α is one of the most important mediators of hormonal response in estrogen-sensitive tissues such as the breast [9] and plays a crucial role in breast growth and differentiation as well as in the development of cancer [10]. The human ER-α gene is localized on chromosome 6q24-q27 [11], it extends more than 140 kb and includes eight exons [12]. Mutation and polymorphism of cancer-associated genes have been found to predict tumor formation and prognosis [8]. At present the literature contains little information regarding ER-α gene expression, mutational frequency, and allelic variants in breast cancer among Asians-Caucasian. Thus, the aim of present study was to examine ER-α polymorphism among Iranian women with breast cancer in order to establish a genetic polymorphism database for the ER-α encoding region of the Iranian genome, and to test for any correlation between ER-α polymorphism and various clinically observable features of breast cancer in Iranian women, and to evaluate the effect of these polymorphisms on breast cancer risk.

Methods

Study population

A case-control study was conducted from April 2004 to January 2008 in Tehran, Iran. The breast cancer patients (n=150; median age 47.49 ± 11.43 years) were newly diagnosed and mostly living in Tehran. They were involved in the study if they had a confirmed pathological breast cancer diagnosis at the Imam Khomeini Hospital Complex. The control group (n =147; median age 40.75±10.54 years) included healthy women with no history of breast cancer or any other neoplastic diseases, and also none of their relatives had a history of breast cancer [13]. Women with hysterectomy and artificial menopause or exposed to any kind of radiation and chemotherapy in their life time were excluded from the study. All patients were provided with written informed consent to participate in the protocol before entering into the present study.

The demographical and risk factor data were collected using a short structured questionnaire, including information on age, weight, height, race, religion, marital status, number of pregnancies and children, age at the first child birth, average lactation term, family history of breast cancer (first-degree relatives), age at menarche, age at marriage, parity, age at first pregnancy, menopausal status, and age at menopause, ABO and Rhesus blood groups, race, age at onset of breast cancer, lymph node metastasis, cancer stage at the time of testing and ER expression in breast cancer tissue. An ongoing protocol to collect and store blood samples for future genomic tests had been approved by the institutional review board. Peripheral whole blood was collected and stored in -80°C until genotyping analysis.
Screening for ER-α variants by single strand conformation polymorphism analysis

In order to identify any mutation or variant sites in the Iranian population, all samples were screened for the coding region of ER-α using the PCR single-strand conformation polymorphism (SSCP) method. A total of 150 breast cancer patients were screened at this stage and compared with 147 control individuals in order to identify disease-associated variants and mutations. Genomic DNA was extracted from whole blood cells using DNGTM-Plus extraction solution kit in accordance with the manufacturer's instructions. Genomic DNA (50 ng) was used for each run of PCR-based genotyping.

The exon 4 of the ER-α gene was amplified by PCR methods, using a set of primers according to the oligonucleotide sequences in Hsiao et al article [14].

The DNA amplification was performed for 30 cycles of 30 seconds at 95°C, 30 seconds at 58°C and 40 seconds at 72°C. Optimal electrophoretic separation for SSCP was conducted in 8% Polyacrylamide gel in buffer (90 mmol l-1 Tris-borate and 2 mmol l-1 EDTA) at 200 V for 2 hours followed with 250 V for 24 hours at 16°C. After electrophoresis, the bands were visualized using 0.1% silver nitrate stain.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Case n=150</th>
<th>Control n=147</th>
<th>Total n=297</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>47.49±11.43</td>
<td>40.75±10.54</td>
<td>44.15±11.49</td>
</tr>
<tr>
<td>Weight</td>
<td>68.41±13.21</td>
<td>61.36±10.63</td>
<td>64.92±12.49</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>161.46±6.01</td>
<td>160.61±5.21</td>
<td>160.48±5.54</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.28±5.09</td>
<td>23.8±4.07</td>
<td>25.05±4.8</td>
</tr>
<tr>
<td>Age at menarche (years)</td>
<td>12.94±1.62</td>
<td>13.24±1.25</td>
<td>13.09±1.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Case n=140</th>
<th>Control n=99</th>
<th>Total n=239</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at marriage (years)</td>
<td>19.24±4.72</td>
<td>22.14±4.56</td>
<td>20.44±4.88</td>
</tr>
<tr>
<td>Number of deliveries</td>
<td>3.92±2.1</td>
<td>2.03±1.4</td>
<td>3.14±2.09</td>
</tr>
<tr>
<td>Number of children</td>
<td>3.44±1.74</td>
<td>1.95±1.26</td>
<td>2.82±1.72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Case n=134</th>
<th>Control n=96</th>
<th>Total n=230</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at first child delivery(years)</td>
<td>21.63±5.46</td>
<td>24.35±4.59</td>
<td>22.77±5.28</td>
</tr>
<tr>
<td>Average lactation term(years)</td>
<td>17.91±8.27</td>
<td>16.01±13.13</td>
<td>17.12±10.59</td>
</tr>
</tbody>
</table>

Table 1. The distributions of selected demographic characteristics and major risk factors for breast cancer patients.

<table>
<thead>
<tr>
<th>Codon 325</th>
<th>Normal a</th>
<th>Heterozygote b</th>
<th>Homozygote c</th>
<th>Total</th>
<th>Test result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. %</td>
<td>Freq. %</td>
<td>Freq. %</td>
<td>Freq. %</td>
<td>χ²=27.035</td>
</tr>
</tbody>
</table>

Case 42 28.0 85 56.7 23 15.3 150 100
Control 78 53.1 53 36.1 16 10.8 147 100
Total 120 40.4 138 46.5 39 13.1 297 100

a Genotype, normal or 00, CCC/CCC, b Genotype, heterozygote or 01 CCC/CCG, c Genotype, homozygote or 11 CCC/CCG

Table 2. Genotypic distribution frequencies of codon 325 in exon 4 of estrogen receptor-α gene in the study population: breast cancer versus control groups.
samples exhibiting varying banding patterns as the result of first sequencing with forward primer, re-purified on agarose gel using a DNA Extraction Kit, Fermentas # K0153, Germany. Then directly sequenced by big dye Terminator V3.1 Cycle Sequencing kit protocol, by the sequencer ABI 3130XL (16 capillaries)

The purification of PCR products was used in order to confirm sequencing by reverse primer. The products were purified using QIAquick PCR purification Kit (50), QIAGEN cat. # 28104, USA

Statistical analysis
\( \chi^2 \) test was employed to assess the influence of polymorphism status on features of breast cancer. Unconditional logistic regression analysis was also performed using SPSS. For calculation of odds ratios (ORs) with 95% confidence intervals (CIs) and to examine the predictive effect of each risk factor for breast cancer. The P < 0.05 was considered as a statistically significant.

Results
The distributions of selected demographic characteristics and major risk factors for study breast cancer of whole of population: breast cancer versus control groups are shown in Table 1. The breast cancer cases (n=150) had median age 47.49 ± 11.43 years and control group (n = 147) with median age 40.75 ± 10.54 years. The median start age of menarche is lower in case (12.94 ± 1.62 years) than control (13.24 ± 1.25 years) groups. Age at marriage another risk factor in developing breast cancer was lower in case (median age 19.24 ± 4.72 years) than control (median age 22.14 ± 4.56 years) groups.

The encoding region exon 4 of the ER-\( \alpha \) gene from samples of 150 breast cancer patients and 147 healthy individuals was screened for mutation or variant sites by the PCR-SSCP and DNA sequencing revealed the presence, of one silent SNP, in exon 4, codon 325 (CCC→CCG), rs 1801132 (dbSNP128), where nucleotide C was converted to G.

The observed numbers of individuals with different genotypes showed that this SNP fitted the Hardy-Weinberg equilibrium for both control and patient groups (P > 0.05). The frequencies of this polymorphism showed significant in both homozygotes (CCC/CCC and CCG/CCG), and heterozygote (CCC/CCG) in the cancer patients, in comparison with the control group (\( \chi^2=27.035, P=0.029 \)) (Table 2).

Among all these risk factors, family history of breast cancer, lymph node (LN) metastasis and age at menarche presented statistically significant differences between various genotypes. Assessing the risk factor of the positive family history of breast cancer, the results showed that among the patients with breast cancer and with the first-degree family affected, the frequency of 11 genotype was three-fold higher (36.7%) than those without any affected family (12.2%), and vice versa among the patients with breast cancer and with no history of cancer, the frequency of 00 normal genotype was six-fold higher (31.3%) than those with the first-degree affected family (5.3%) [\( \chi^2=10.587, P=0.005 \)]. For risk factor of lymph node metastasis, the results showed that among the patients with breast cancer and with lymph node metastasis, the frequency of 11 genotype was five-fold higher (47.8%) than those without LN metastasis (9.5%), and vice versa among the patients with breast cancer and with no lymph node metastasis, the frequency of 00 normal genotype was two and a half times more (30.7%) than those with the LN metastasis (13.0%) [\( \chi^2 =22.349, P=0.001 \)].

Our study revealed that the frequency of normal 00 genotype (CCC/CCC) was 28.0% in case and 53.1% in control groups as it was expected. Heterozygote 01genotype (CCC/CCG) frequencies were higher in case group than control group (56.7% and 36.1% respectively). The frequencies for homozygote genotype 11 (CCG/CCG) were also higher in case group (15.3%) than the control group (10.8%), and all these
differences were significant ($\chi^2=19.448, P=0.001$).

The frequency of allele 1 (CCG) was significantly higher among cancer patients (39.6%) than in the control individuals (28.9%), in comparison with the frequency of allele 0 (CCC) which showed 60.4% and 71.1% respectively ($\chi^2 = 7.345, P = 0.007$).

The frequency of allele 1 was significantly higher in cancer patients with lymph node metastasis (67.4%) than in those without LN metastasis (39.4%), in comparison with the frequency of allele 0 which was significantly lower in cancer patients with LN metastasis (32.6%) than in those without LN metastasis (60.6%), ($\chi^2 = 12.432, P = 0.001$).

The frequency of allele 1 was significantly higher among cancer patients with first-degree family affected (65.8%) than in those without family history (40.5%), in comparison with the frequency of allele 0 which showed 44.2% and 64.4% respectively ($\chi^2 = 8.657, P = 0.003$). The frequency of allele 1 was also significantly higher among cancer patients with age at menarche <=12 years (55.8%) than in those with age at menarche >12 years (35.6%), in comparison with the frequency of allele 0 which showed 44.2% and 64.4% respectively ($\chi^2 = 12.035, P = 0.001$).

When breast cancer was considered, genotype frequencies exhibited significantly different distributions between the case and control groups, for codon 325 ($P = 0.001$). The estimated risk was higher in cases (61.6%) than in controls (38.4%) for 01 heterozygote (OR 0.336, 95% CI 0.202-0.558), and also for 11 homozygote with 59.0% and 41.0% respectively (OR 0.078, 95% CI 0.015-0.404), when compared with the normal 00 homozygote with the frequencies of 4.8% and 95.2% (twenty-fold higher) respectively.

Genotype frequencies showed statistically significant variances in presence and absence of first-degree family history of breast cancer, for codon 325 ($P = 0.005$). The frequency of 01 heterozygote for individuals with the first-degree family history was 12.9% and for those without family history was 87.1% (seven-fold higher), (OR 0.164, 95% CI 0.020 - 1.316), and similar result was for 11 homozygote with 30.4% and 60.9% (two-fold higher) respectively (OR 0.056, 95% CI 0.006 - 0.490), in contrast with the normal 00 homozygote with the frequency of 0.4% and 97.6% (forty-fold higher) respectively.

Furthermore, statistically significant differences were achieved in the presence and absence of LN metastasis for the same codon ($P = 0.001$). The frequency of 01 heterozygote for patients with the presence of LN metastasis was 10.6% and for those with the absence of LN metastasis was 89.4% (eight-fold higher), (OR 0.650, 95% CI 0.166 - 2.537), and the figures for 11 homozygote were 47.8% and 52.2% respectively (OR 0.084, 95% CI 0.200 - 0.351), when compared with the normal 00 homozygote with the frequency of 7.1% and 92.9% (thirteen-fold higher) respectively.

**Discussion**

In the ER-α gene, several DNA sequence variations have been described that are of increasing interest because of their potential association with breast cancer and other hormone-related diseases [15], but no consistent effect over studies has been shown for the some variants with regard to breast cancer risk [16,17].
Substantial evidence indicates that ER participates in the mammary gland tumorigenesis, and thus is among the genes that affect breast cancer susceptibility. Breast cancer associated ER-\(\alpha\) polymorphisms were surveyed in previous studies [18-24], and it has been shown that ER-\(\alpha\) protein over expression is common in breast cancer [25]. Somatic mutation of the ER-gene has been identified [26], but mutation of ER- germ-line rarely occurs in breast cancer patients. In agreement with observed low mutation rates, the present study found no novel mutations. Unexplained differences between Asian and Western breast cancer symptomatology and demographics led us to consider whether unknown genetic factors within the Iranian genome are involved, prompting us to conduct the present PCR analysis of ER- polymorphism.

The frequency of allele 1 (CCC\(\rightarrow\)CCG) in codon 325 (CCC and CCG are codon which code for Proline, that have previously been reported in different study groups) was significantly, higher in cancer patients (39.6%) than control individuals (28.9%). The frequency of allele 1 among breast cancer individuals in Taiwan (52.1%) matches that in Korea (50.0%), and is much more common than found in USA, England, Australia and Portugal (approximately 20%) [14], but in Iranian was approximately Asian and Western populations. Although, the frequency of allele 1 in codon 325, was higher in patients, both for Iranian and Western populations but lower in Asian populations [14].

The frequency of allele 1 was significantly higher in cancer patients with the age at menarche \(<=12\) years old comparison with those with the age at menarche \(>12\) years old. Also, for the risk factors involving first-degree family history, the frequency of allele 1 was significantly higher in cancer patients with first-degree family history of breast cancer than those without family history. Therefore, in Iranian population, these factors could directly influence the accuracy of predicting breast cancer development. Also, the frequency of allele 1 was much greater in the Iranian population studied here than those in the Western populations; this finding, together with the relatively low incidence of breast cancer in Iran, suggests that this SNP could be a protective factor against breast cancer.

LN metastasis can be considered as a clinical indicator during presurgical evaluation, at least in the Iranian population. Such a test is of interest because of association of the lymphatic invasion is associated with local recurrence and disease progression, and LN metastasis is considered to be an important indicator when deciding whether chemotherapy should be administered [27-30]. Genotype frequencies exhibited statistically significant different distributions in the presence and absence of LN metastasis, with \((P=0.001)\). However, our data showed that there is a negative correlation between allele 1 and LN metastasis, indicating that presence of allele 0 and absence of allele 1 which may be independent parameters for node positivity. This observation was in agreement with a Taiwanese, Chinese and Portuguese studies based on ER-\(\alpha\) codon 325 PCR analysis of excised cancer tissue samples [14, 17, 22]. Therefore, it was noted that the greater the frequency of allele 1, the lesser the likelihood of LN metastasis in the Iranian breast cancer patients.

To our knowledge, the link between silent polymorphisms and phenotypes is unclear. One of the possibilities might be that the silent polymorphism might have linkage with another genetic mutation that directly affects breast cancer phenotype. The other possibility might be that the nucleotide composition at the silent polymorphic site could alter the gene expression level of ER-\(\alpha\), thus leading to the association of LN metastasis in breast cancer.

**Conclusion**

In conclusion, ER- polymorphisms in an Iranian clinical breast cancer group (150 breast cancer patients and 147 control individuals)
were established using PCR SSCP technology through analysis of peripheral blood cells. The SNP reported in Western studies were also found in the Iranian population studied, but at different frequencies than the Western counterparts. Statistically significant correlations were found between allele distributions among cancer patients and normal controls, and also between positive family history, LN metastasis and age at menarche. Because of the limited sample size in the present study, the observed correlations will require further confirmation. This is planned as part of our future work, since SNP determination from peripheral blood could represents a highly feasible and noninvasive option for preoperative evaluation.

**Abbreviations**

BMI = body mass index; CI = confidence interval; ESR = estrogen receptor; LN = lymph node; OR = odd ratio; PCR = polymerase chain reaction; Rh = Rhesus blood group system; SNP = single nucleotide polymorphism; SSCP = single-strand conformational polymorphism.

**Acknowledgment**

This research has been supported by Tehran University of Medical Sciences & Health Services grant (grant # 2850).

The authors would like to thank Ms. Elham Farazandeh and Ms. Maasumeh Jafari Eftekhar from Central Clinic No.1, Cancer Institute, Imam Khomeini Hospital Complex, which provided blood samples and clinical information from the patients. We are grateful to Ms. Roya Sharifian for her assistance in statistical analysis.

**References**


