Abstract

Background/aim: Variation in the three prime untranslated regions (3'-UTR) affects the binding of microRNA (miRNA) to the Breast cancer susceptibility gene 1 (BRCA1) gene and thus regulates its expression. In this study, the consequences of variation in miRNA-Binding Site (rs8176318 G>T) in the 3'-UTR of BRCA1 and its association with the risk of breast cancer (BC) was investigated.

Materials and Methods: The selected variation (rs8176318 G>T) was genotyped in BC (n=300) cases and healthy controls (n=300) using allele-specific PCR (Tetra-Primer Amplification Refractory Mutation System-Polymerase Chain Reaction). The results of T-ARMS-PCR was further confirmed by Sanger sequencing through a random selection of 10% already analyzed samples by T-ARMS-PCR. Association of this variation with BC was tested by calculating the odds ratio (at 95% CI) and χ²-value using four different genetic models (co-dominant, dominant, recessive and additive models).

Results: Using fisher’s exact test, a significant association between variant rs8176318 (G>T) and BC was found in co-dominant [χ²-value=15.68, df:2 P=<0.0004], dominant [OR= 1.557 (1.082-2.241), P=<0.0213], recessive [OR=0.474 (0.3204-0.7017), P=0.0002] and additive models [OR=1.609 (1.282-2.018), P=<0.0001].

Conclusion: It is therefore concluded that there is a signification association between rs8176318 and BC risk in a case-control study in Pakistani population. Furthermore, an association study using a large sample size is required to further verify these findings.

Keywords: Case-control study, Single nucleotide polymorphism, un-translated regions, miRNAs
1. Introduction

Breast cancer is estimated to be the second leading cause of death worldwide [1]. There are several environmental and genetic factors contributing to BC. Among genetic factor, a number of protein and non-coding genes contribute to the etiology of BC. \textit{BRCA 1} is one of the main protein-coding gene related to the causes of BC. Mutations in \textit{BRCA1} account for about 40-45% of BC [2], Defects in \textit{BRCA1} effects cell cycle at various points, destabilizing the DNA and triggers DNA damage response hence increases the formation of tumors [3, 4]. \textit{BRCA1} is specifically involved in repairing breaks in double-stranded DNA, transcription, and recombination. The frequency and types of mutations in the \textit{BRCA1} gene varies among the different population worldwide. The 3′-UTR of \textit{BRCA1} plays a pivotal role in stability, localization, and transport of mRNA, hence affecting its expression [5, 6]. SNPs in 3′-UTRs affects the expression of genes, thus influencing the risk of cancer development. It has been reported that SNPs in \textit{BRCA1} 3′-UTR are disturbing the binding sites of miRNA. Therefore, it can be used as a biomarker for BC [7]. The previous study has shown that TT/TG variant is linked with increased BC risk [8].

MiRNAs are small non-coding RNAs molecules that control gene expression by regulating their corresponding mRNA. They act by binding to their target through miRNA:mRNA complementary base pairing at the untranslated regions of the genes. The capability of miRNAs binding to their target mRNA in the 3′-UTR is crucial for proper expression of proteins. However, the binding capacity can be adversely affected by mutations/Single nucleotide polymorphism in the sites to which miRNA binds in the gene [9].
These observations lead to the assumption that unidentified variations in the miRNA binding sites at the 3’-UTR of BRCA1 genes could provide an important insight into the etiology of BC risk factor.

Several techniques are used for the detection of mutation in the genome. However, T-ARMS PCR is flexible, fast and inexpensive SNP discovery strategy as compared to other genotyping tools [10]. It includes a single PCR and subsequent gel electrophoresis [11]. It utilizes four primers two inner and two outer primers. The inner two primers are allele-specific and result in allele-specific amplification. The inner primer amplification depends on the genotype used in the PCR. The outer primers act as an internal control and generate the outer fragment of the flanking locus of the SNP. Unequal positioning of the inner primers from the corresponding outer primer results in the generation of PCR product with diverse sizes hence easily identified in the gel [12].

In the current study, we sought to investigate SNPs in BRCA1 3’-UTR region and to explore its association with breast cancer in sporadic BC patients.

2. Materials and Methods

2.1. Study subjects and ethical approval
The study included female patients (n=300) ages ranged from 22-68 years with confirmed BC at the IRNM, Peshawar Pakistan and matching controls (n=300 with no history of cancer ages ranged from 20-70 years) as shown in Table 1. The present study followed the guidelines of the Helsinki declaration. The current research work was approved from the Advanced Study and Research Board (ASRB) on its 40th meeting held on March 29, 2017, University of Malakand, Pakistan.

2.2 Inclusion / exclusion criteria and genomic DNA extraction
The inclusion criteria for the study was 300 confirmed BC patients while women with no BC history was used as exclusion criteria. A total of 5ml whole blood was collected from BC patients and healthy controls in the EDTA tube and stored at refrigeration temperature until further processing. Genomic DNA was isolated from the collected blood samples using the phenol-chloroform method and stored at -20°C until further processing.

2.3 PCR based amplification and sanger sequencing of BRCA1 gene

Two sets of primers (Table 2A) were used to amplify the 3'-UTR region of BRCA1 gene following thermocycling conditions of initial denaturation at 95 ºC for 10 minutes, followed by 35 cycles of 95 ºC for 1 minute, primer annealing at 56ºC for 30 seconds, extension at 72 ºC for 1 minute and final single step extension at 72 ºC for 10 minutes. The PCR product was visualized on 1% agarose gel by staining with ethidium bromide.

The amplified region was commercially sequenced using Sanger sequencing method and the sequencing results were compared with the reference sequence obtained from NCBI. The analysis was carried out using Vector NTI software. Single Nucleotide variants were detected.

2.4 Allele-specific T-ARMS-PCR amplification

The SNP (rs8176318; G>T) in 3'-UTR region of the BRAC1 gene was identified through Sanger sequencing. Further screening was performed for genotyping of the selected SNP in cases and control using T-ARMS PCR technique. Primer 1 software available online (http://primer1.soton.ac.uk/primer1.html) was used to design allele-specific primers for the identified SNPs (Table 2B). The PCR reaction mixture of 30 µL in a 0.2 ml PCR tube containing approximately 100 ng/µL of DNA template, 17 µL of green-Taq PCR master mix
(Thermo Scientific), 4 µL of primers (1 µL forward inner, 1 µL inner reverse, 1 µL forward outer and 1 µL reverse outer primer (each primer of 10 pmol/µL) and up to 7 µL of nuclease-free water was prepared and processed. The amplification conditions used were initial denaturation at 95 ºC for 10 minutes followed by 35 cycles of denaturation, annealing and extension at 95 ºC for 1 minute, 56 ºC for 40 seconds, 72 ºC for 40 seconds. This was followed by the final extension step at 72 ºC for 10 minutes. The products of amplification were separated by subjecting to electrophoresis on 3% Agarose gel. Ethidium bromide was used for staining and visualization of PCR products. Results were confirmed by reproducing 15% of the samples and accuracy of reproducibility was 99%. Furthermore, the results of Sanger sequencing and T-ARMS PCR was 100% reproducible. *BRCA1*-3’-UTR wild (GG), as well as mutant genotype (TT), yielded 177 bp, 279 bp products while the outer primer product was 400bp respectively as depicted in Figure 1.

2.5 Statistical analysis

Hardy-Weinberg equilibrium was applied as quality control to check the observed and expected genotype frequencies in both cases and controls. Co-dominant, dominant, recessive and additive models were applied on the collected data. The odds ratio (OR) at 95% confidence intervals (CIs) was also calculated. A p-value < 0.05 was considered significant.

3 Results

Through the alignment of DNA sequences of the *BRCA1* gene against reference sequence from NCBI, using gene sequencing analysis software, a genetic variant (rs8176318; G>T) in the 3’UTR of BRCA1 was identified. The variant was present as homozygous GG, homozygous TT and heterozygous GT in the analyzed samples (Figure 1). Furthermore,
genotype data obtained through T-ARMS PCR were statistically analyzed for association with BC (Table 3). Selected samples were sequenced to confirm the result of T-ARMS PCR. The results obtained by T-ARMS PCR were concordant with sequencing (Figure 2). The distribution of genotype was found to be consistent with HWE in both BC cases ($P = 0.5707$) and controls ($p = 0.4672$). Co-dominant model (genotype frequency distribution) showed that genotype GG was 31.33% (n=94) in BC cases and 22.67% (n=68) in controls. Similarly, frequency of genotype GT was higher in BC cases than controls that is 52.0% (n=156) and 47.67% (n=143), respectively. Whereas, the frequency of genotype TT was significantly lower in cases than controls that are 16.67% (n=50) and 29.67% (n=89), respectively. Thus, a significant statistical difference in genotypes was determined through co-dominant model analysis ($\chi^2 = 15.68; P < 0.004$), indicating the role of this variant in BC risk.

However, through additive model (allele frequency distribution) analysis it was found that allele G was signifying higher in BC cases than controls that are 57.33% and 45.51%, respectively. Thus, a significant difference in the allele frequency was determined through additive model analysis in BC cases than controls at 95% CI [OR = 1.609 (1.282-2.018); $P < 0.0001$], indicating allele G as a susceptible allele while allele T is protective. Similarly, both dominant (GG vs. GT + TT) and recessive statistical models (TT vs GT + GG) showed significant association with BC risk at 95% CI ($P < 0.0213$ and $P < 0.0002$), indicating that individuals with genotype GG are at risk of breast cancer while those with TT genotype are protective of this risk.

4 Discussion

In the current study, we investigated the association of (rs8176318; G>T) in $BRCA1$ 3'-UTR SNP with the risk of BC. The risk of BC risk was found to be significantly increased in
individuals having GG or GT genotype as compared with those having TT genotype in a
Pakistani Population. The TT genotype was observed as to decrease the risk of BC as
compared to GG Genotype. The GG genotype with an odds ratio of 1.557 appeared to be the
risk allele while the TT allele with an odds ratio of 0.47 was found to provide protection
against breast cancer in Pakistani population analyzed in the current study. Previous studies
support our observations that variants within the miRNA binding sites of tumor suppressor
or oncogenes play a pivotal role in the development of BC [13]. Comparing the frequencies
of BRCA1-3’UTR-G>T heterozygosity in our study group with other studies, we found GT
heterozygosity generally comparable with Italy (TSI) 50%, Chinese in Colorado (CHD) 56%,
Gujarati Indians in Houston, Texas (GIH) 52%, European ancestry in Utah and 50% in Saudi
Arabian population [14]. Our frequencies of GG, GT and TT (31.33%), (52.0%) and
(16.67%) genotypes in patients are comparable to GIH (29.5%), (52.3%) and (18.2%). Our
G (57.3%) and T (42.6%) allele frequencies are comparable to GIH, Gujarati Indians in
Houston, Texas (55.7%) and (44.3%), CHD, Chinese individuals in Metropolitan Denver,
Colorado (56.5%) and (44.5%) [15]. Previously, studies in African-American women
reported rs3092995 in the 3’-UTR region of the BRCA1 gene to increase the risk of breast
cancer [16]. In support of our findings, another study showed that nrs8176318 GG genotype
had a significantly lower expression level (of BRCA1 mRNA) than the GT and TT
genotypes. Previous studies support our observations that variants within the miRNA binding
sites of tumor suppressor or oncogenes play a role in breast cancer development [13]. Other
studies have reported different alleles rs12516 and rs8176318 to increase the risk of BC [17].
The rs8176318 G/T variant was linked with increased BC risk in studies across different
population worldwide [15]. The rs8176318 has also been reported to be a risk factor in
African-American women with an odds ratio of 12.19 [18]. In contrast to our findings, some studies on variants inside \textit{BRCA1} 3'-UTR region have failed to show association with breast cancer risk [19]. Similarly, in another contrasting study, the \textit{BRCA1} rs8176318 (T) variant has been shown to lower the gene expression and is associated with advancing the breast cancer stage. According to this study a 4 fold increase in of shifting to stage V disease [8].

The study is carried out in a different population (Irish) as compared to ours (Asian). The study reported an increased risk of Triple-negative breast cancer patients. Our results reveal that the newly identified TT allele may be a new genetic marker for a decreased risk of sporadic BC. On the other hand, women with GG allele are at an increased risk of BC. It has been shown that the mutant (T) allele reduces the BRCA1 expression in response to decrease in estrogen amount as compared to the G allele showing that estrogen treatment dysregulating the T-allele expression, causing the down-regulation of \textit{BRCA1} gene [8]. The results from our study present variation in the non-coding region of \textit{BRCA1} that can influence its expression increasing the risk of BC in Pakistani females. More evidence is being produced supporting the notion that susceptibility to breast cancer is increased by variants inside the 3'-UTR [20]. The reason for the association of SNPs in 3'-UTR with increased BC risk may be due to the fact that miRNAs bind to mRNA in the 3'-UTR region thus regulating the levels of mRNA and hence its expression [21]. In support of our findings, it is indicated by previous work that 3'-UTR variants act as genetic markers of Cancer risk [22]. The major reason for discrepancies in our results with other study is a diverse ethnic group (study population).

5 Conclusion
It is concluded that females from with GG genotype invariant (rs8176318; G>T) have a higher risk of developing breast cancer and women with TT genotype in BRCA1 3'-UTR have less chance of developing breast cancer as compared to GG genotype.

**Conflict of interest**

The authors declared no conflict of interest.

**References**


17. Pongsavee M, Yamkamon V, Dakeng S, O-charoenrat P, Smith DR et al. The BRCA1 3′-UTR: 5711+ 421T/T_5711+ 1286T/T genotype is a possible breast and


**Tables**

**Table 1.** Categorization of pre-clinical data for each patient with their age and cancer stages (I-IV).

<table>
<thead>
<tr>
<th>Age group</th>
<th>Stage I</th>
<th>Stage II</th>
<th>Stage III</th>
<th>Stage IV</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>20s</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>30s</td>
<td>7</td>
<td>27</td>
<td>21</td>
<td>12</td>
<td>67</td>
</tr>
<tr>
<td>40s</td>
<td>7</td>
<td>34</td>
<td>22</td>
<td>16</td>
<td>79</td>
</tr>
<tr>
<td>50s</td>
<td>3</td>
<td>31</td>
<td>40</td>
<td>13</td>
<td>87</td>
</tr>
<tr>
<td>60s</td>
<td>5</td>
<td>8</td>
<td>19</td>
<td>5</td>
<td>37</td>
</tr>
<tr>
<td>70s</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>80s</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>25</strong></td>
<td><strong>109</strong></td>
<td><strong>114</strong></td>
<td><strong>52</strong></td>
<td><strong>300</strong></td>
</tr>
</tbody>
</table>

**Table 2A.** List of primers used for amplification of selected region of *BRCA1* gene.

<table>
<thead>
<tr>
<th>BRCA 1 primers</th>
<th>Primer sequences</th>
<th>Product size</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA1 3'-UTR 1 Forward</td>
<td>5'- AGCACTCTACCAGTGCCAG- 3</td>
<td>646 bp</td>
</tr>
<tr>
<td>BRCA1 3'-UTR 1 Reverse</td>
<td>5'- AGGTTCAGTTTCTTTCTTCA- 3</td>
<td></td>
</tr>
<tr>
<td>BRCA1 3'-UTR 2 Forward</td>
<td>5'- GAGTGCTTGGGATCGATTATGTGACT-3</td>
<td>678 bp</td>
</tr>
<tr>
<td>BRCA1 3'-UTR 2 Reverse</td>
<td>5'-GCAACAGCTTCCTCCTGGTGAGG-3</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2B** List of the outer and inner primers used for the genotyping of (rs8176318) in the miRNA-Binding site of *BRCA1* Gene.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Product Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward inner primer (G allele)</td>
<td>TTACTTCTCTAAAACCTGTGTTCCAGAG</td>
<td>177 bp</td>
</tr>
<tr>
<td>Reverse inner primer (T allele)</td>
<td>TCCATTGAGGGTGCTCTCTCTTACA</td>
<td>279 bp</td>
</tr>
<tr>
<td>Forward outer primer (5' - 3')</td>
<td>CTTAAAGATTTTCTGCTTGAAGTCTCCCT</td>
<td></td>
</tr>
<tr>
<td>Reverse outer primer (5' - 3')</td>
<td>TTCCTTTCATTTCTAATACCTGCCTCA</td>
<td>400 bp</td>
</tr>
</tbody>
</table>
Table 3. Genotype and allele frequencies between subject and control group at locus rs8176318) in miRNA-Binding Site of BRCA1 Gene using Co-Dominant, Dominant, Recessive and Additive statistical models

<table>
<thead>
<tr>
<th>Statistical Models</th>
<th>Genotypes</th>
<th>Cases (%age)</th>
<th>Controls (%age)</th>
<th>Fisher's exact test OR (95% CI)</th>
<th>$\chi^2$ value, df</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Dominant</td>
<td>GG</td>
<td>94 (31.33)</td>
<td>68 (22.67)</td>
<td>-</td>
<td>15.68, 2</td>
<td>&lt;0.0004</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>156 (52.0)</td>
<td>143 (47.67)</td>
<td>1.557 (1.082-2.241)</td>
<td></td>
<td>&lt;0.0213</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>50 (16.67)</td>
<td>89 (29.67)</td>
<td>0.4742 (0.3204-0.7017)</td>
<td></td>
<td>&lt;0.0002</td>
</tr>
<tr>
<td>Dominant</td>
<td>GG</td>
<td>94 (31.33)</td>
<td>68 (22.67)</td>
<td>1.557 (1.082-2.241)</td>
<td></td>
<td>&lt;0.0213</td>
</tr>
<tr>
<td></td>
<td>TT + GT</td>
<td>206 (68.67)</td>
<td>232 (77.33)</td>
<td>0.4742 (0.3204-0.7017)</td>
<td></td>
<td>&lt;0.0002</td>
</tr>
<tr>
<td>Recessive</td>
<td>TT</td>
<td>50 (16.67)</td>
<td>89 (29.67)</td>
<td>0.4742 (0.3204-0.7017)</td>
<td></td>
<td>&lt;0.0002</td>
</tr>
<tr>
<td></td>
<td>GG + GT</td>
<td>250 (83.33)</td>
<td>211 (70.33)</td>
<td>1.609 (1.282-2.018)</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Additive</td>
<td>G</td>
<td>344 (57.33)</td>
<td>279 (45.51)</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>256 (42.67)</td>
<td>334 (54.49)</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figures

Figure 1. Gel electrophoresis image of T-ARMS-PCR for the detection of polymorphism in $BRCA1$ gene. Genotypes are indicated (lane 1 & 8 = GT, 3 = TT, 2, 4, 5, 6, 7 = GG genotypes and M = Marker).

Figure 2. Validation assay to confirm T-ARMS-PCR result by using sequencing. Sequencing result of rs2910164 showed homozygote GG, heterozygous and CG genotypes using reverse primer, respectively.