Identification of CDKN2A variants in breast cancer patients in Pakistan

Ranjha Khan¹, Ayesha Aftab², Sobia Tabassum², Hafiz Muhammad Jafar Hussain¹, Abdul Hameed³, Humera Mahmood⁴, Faiza Munir² and Ihtisham Bukhari⁵*

¹The CAS Key Laboratory of Innate Immunity and Chronic Diseases, School of Life Sciences, University of Science and Technology of China, Huangshan Road, Hefei, Anhui 230027, China
²Department of Biological Sciences, International Islamic University, Sector H10, Islamabad 44000, Pakistan
³Institute of Biomedical and Genetic Engineering (IBGE), Kashmir Highway, I, Service Road South, G9/1, Islamabad 44000, Pakistan
⁴Department of Oncology, Nuclear Medicine Oncology & Radiotherapy Institute (NORI) Hospital, Hanna Road, G8/3, Islamabad 44000, Pakistan
⁵Translational Research Institute, Henan Provincial People’s Hospital, School of Medicine, Henan University, Huanghe Road, Zhengzhou, Henan 450053, China

(Received 30 October 2018, accepted 25 February 2019; J-STAGE Advance published date: 21 June 2019)

The role of cyclin-dependent kinase inhibitor 2A gene (CDKN2A) variants in breast cancer is not well understood, here we investigated their possible effects on breast cancer in Pakistani women attending the NORI Hospital, Islamabad. Direct DNA sequencing of CDKN2A identified an already known polymorphism in the 3’ UTR, c.*29G>C (rs11515), in 5.88% patients and two novel variants. One, a deep intronic substitution (c.458-554T>G) in 1.96% patients, is also detected as a compound heterozygous form along with c.*29G>C in 1.96% patients (c.[458-554T>G; *29G>C]). The other is a novel deletion (c.458-82delG) occurring as a compound variant with two other identified variants c.[458-554T>G; 458-82delG; *29G>C] in 1.96% patients. In silico pathogenicity prediction analyses did not predict pathogenic effects on breast cancer for these individual variants. We conclude that variations in CDKN2A are not the major genetic cause of breast cancer in the enrolled Pakistani patients.

Key words: breast cancer, CDKN2A, p16INK4A, polymorphisms, mutations

INTRODUCTION

Breast cancer is one of the world’s leading causes of female mortality. Among Asian countries, Pakistan has the highest rate of breast cancer, possibly due to unawareness about breast health, low socioeconomic condition and genetic background (Menhas and Umer, 2015). Several genetic risk factors and causes of breast cancer have been identified, one of which is the cyclin-dependent kinase inhibitor (CDKN2A/p16INK4A) gene. CDKN2A (OMIM no. 600160) is a tumour suppressor gene located on the short arm of chromosome 9 (9q21) (Rayess et al., 2012), and encodes p16 protein, which plays a key role in cell division during early G1 phase by inhibiting the progression of the cell cycle from G1 to S phase (Shan et al., 2013). CDKN2A is a principal controller during cell division and most tumours, including breast cancer, show dysregulation in the p16 pathway (Shan et al., 2013).

Mutations in or functional inactivation of CDKN2A are known to be involved in numerous types of human cancers, such as leukaemia, lung cancer, oesophageal cancer, glioma, breast cancer and melanoma (Hinshelwood et al., 2009; McKenzie et al., 2010; Puig et al., 2016). It is primarily a melanoma-associated gene and mutations have been reported in 24% of melanoma-prone Latin America families (Puig et al., 2016). It is primarily a melanoma-associated gene and mutations have been reported in 24% of melanoma-prone Latin America families (Puig et al., 2016). Similarly, 14% of Spanish melanoma patients also harbour CDKN2A variants (Puig et al., 2016). However, CDKN2A has shown no clear association with breast cancer in several other populations or ethnic groups (Berns et al., 1995; Quesnel et al., 1995; Horcasitas et al., 2017; Aftab et al., 2019). To help resolve the ambiguity regarding pathogenic effects of CDKN2A...
on breast cancer, we recruited breast cancer patients from the Nuclear Medicine Oncology & Radiotherapy Institute (NORI) Hospital in Islamabad, Pakistan. Through DNA sequencing, we identified mutations/polymorphisms in CDKN2A and predicted the pathogenicity of each identified CDKN2A mutation/polymorphism in current breast cancer patients using various online bioinformatics tools.

**MATERIALS AND METHODS**

**Subjects** Fifty-one women diagnosed with breast cancer were recruited from the NORI Hospital. The ethical committee and institutional review board of the International Islamic University, Islamabad, Pakistan approved all sampling and experimental procedures, which were performed by strictly following the ethical standards of the 1964 Declaration of Helsinki and its latest amendments, or comparable ethical standards. Informed written consent was obtained from all participants.

**DNA extraction and genotyping** Venous blood samples (5–10 ml) were collected from each patient into ethylenediaminetetraacetic acid tubes. DNA was extracted by the phenol–chloroform method and CDKN2A was amplified using specially designed primers (Supplementary Table S1) with the Thermo Hybaid PXE 0.2 PCR system, in a 30-μl reaction containing 3.0 μl of 10% PCR buffer, 1.8 μl of 25 mM MgCl₂, 2.4 μl of 2.3 mM dNTPs, 0.6 μl of PrimeSTAR HS DNA polymerase and 1.5 μl of 5% dimethyl sulphoxide. PCR products were denatured at 95 °C for 5 min and analysed via single-strand conformation polymorphism (SSCP) on an 8% polyacrylamide gel run at 30 W for 4 h. The gel was stained with ethidium bromide and visualized with a UV transilluminator. Products with an altered electrophoretic mobility characteristic of single-stranded DNA were sequenced using an ABI 3130 Genetic Analyzer (Applied Biosystems PRISM 310, Foster City, CA, USA). The files generated were further inspected by Sequencing Analysis 5.2 software (Applied Biosystems).

**In silico analyses** To predict the impact of identified variants on splicing, we used different online tools. Human Splicing Finder (http://www.umd.be/HSF3/), SROOGLE (http://sroogle.tau.ac.il/) and NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) were used to evaluate the effect of intronic variants, while VarioWatch (http://grch38.genepipe.ncgm.sinica.edu.tw/variowatch/main.do), UTRScan (http://itbtools.ba.itb.cnr.it/utrsan) and the PolymiRTS database 3.0 were used for analysis of 3' UTR single-nucleotide polymorphisms (SNPs).

**RESULTS**

**Identification of variants** PCR–SSCP analysis of CDKN2A amplicons revealed band alterations on SSCP gels in six of the 51 samples (Supplementary Fig. S1). For further clarification, we performed DNA sequencing of these samples, which identified one already known polymorphism and three novel variants including an intronic substitution and two compound variants (Table 1). The known 3' UTR variant (c.*29G>C) was identified in three patients (5.88%), of whom one had the heterozygous form and two the homozygous form of the mutation (Fig. 1A). A deep intronic heterozygous substitution (c.458-554T>G) in intron 2 was detected in one patient (1.96%) (Fig. 1B), while another was found to have a deletion mutation (c.458-82delG) in intron 2 (Fig. 1C); this patient also carried two other mutations, c.*29G>C and c.458-554T>G, making her an exceptional case with three mutations, which is an extremely rare form of compound mutations. Another patient carried two variants, c.458-554T>G and c.*29G>C (Table 1).

**Predicting the functional impact of identified variants using in silico tools** Human Splicing Finder (HSF) analysis predicted that c.458-554T>G causes a new acceptor site, which would affect splicing. Variant c.458-82delG also showed a high positive variation percentage, predicted to create a new acceptor site, although HSF anticipated that this new site would have no effect on splicing. SROOGLE and NetGene2 analyses predicted little or no difference between wild-type and mutant scores, with no effect on splicing. Thus, overall, HSF analyses did not predict any substantial role for the newly identified variants in splicing patterns (Table 2).

VarioWatch predicted that the 3' UTR SNP (rs11515) has no functional impact on splicing. In the genome

<table>
<thead>
<tr>
<th>Variant</th>
<th>Type</th>
<th>Location</th>
<th>Nucleotide change</th>
<th>Frequency (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.458-554T&gt;G</td>
<td>Substitution</td>
<td>Intron 2</td>
<td>T&gt;G</td>
<td>1.96</td>
<td>This study</td>
</tr>
<tr>
<td>c.*29G&gt;C</td>
<td>SNP</td>
<td>3' UTR</td>
<td>G&gt;C</td>
<td>5.88</td>
<td>Kitts and Sherry, 2011</td>
</tr>
<tr>
<td>c.[458-554T&gt;G; *29G&gt;C]</td>
<td>Compound Mutation</td>
<td>Intron 2 &amp; 3' UTR</td>
<td>–</td>
<td>1.96</td>
<td>This study</td>
</tr>
<tr>
<td>c.[458-554T&gt;G; 458-82del; *29G&gt;C]</td>
<td>Compound Mutation</td>
<td>Intron 2 &amp; 3' UTR</td>
<td>–</td>
<td>1.96</td>
<td>This study</td>
</tr>
</tbody>
</table>
CDKN2A variants in breast cancer patients

A. Wild-type sequence given with heterozygous and homozygous SNP (c. 29G>C) rs11515 found in the 3' UTR. (A) A heterozygous deep intronic substitution (c.458-554T>G) in intron 2. (C) A homozygous deep intronic deletion (c.458-82delG) in intron 2.

Table 2. In silico analysis of intronic variants for predicting impact on splicing

<table>
<thead>
<tr>
<th>Variants</th>
<th>HSF a</th>
<th>SROOGLE b</th>
<th>NetGene2 c</th>
<th>Prediction of possible functional impact**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HSF metric score range (1-100)</td>
<td>Range (1-100)</td>
<td>Range (0-1)</td>
<td></td>
</tr>
<tr>
<td>c.458-554T &gt; G</td>
<td>39.78</td>
<td>68.72</td>
<td>+72.75</td>
<td>New A.S. a</td>
</tr>
<tr>
<td>c.458-82delG</td>
<td>65.13</td>
<td>41.27</td>
<td>+57.81</td>
<td>Probably no effect on splicing</td>
</tr>
</tbody>
</table>

b In SROOGLE, the effect on splicing is predicted through the difference between wild-type and mutant scores; scores close to 1 are signalled as strong splice sites, http://sroogle.tau.ac.il/.
c NetGene2 = neural network prediction server; predicts the acceptor site (A.S.) and donor site (D.S.); scores close to 1 are signalled as strong splice sites, http://www.cbs.dtu.dk/services/NetGene2/output.php.

** Prediction of overall possible impact.

view' results generated by this tool, a graphic representation of all chromosomes is given with highlighted colours for the position and risk level of a selected SNP (Fig. 2); the black colour associated with rs11515 indicates that this SNP has no functional impact (Cheng et al., 2012). The UTRScan server predicted that rs11515 has no damaging or other impact on the putative microRNA target site. Further analysis using the PolymiRTS database predicted that rs11515 has a changed motif pattern for miRNA binding. The high negative context + score...
indicated that a new target site has been created (Table 3). We were unable to predict the pathogenicity scores of the variants existing as double or triple compound forms because none of the tools used could analyse such complex variants.

**DISCUSSION**

Alternative splicing of CDKN2A generates two distinct transcripts, i.e., p16INK4A and p14ARF. Both transcripts have been extensively studied in various cancers. Although the role of transcripts cannot be neglected, p14ARF did not show any association with cancers (Rocco and Sidransky, 2001). In the current research, we focused on the identification and in silico prediction of CDKN2A variants in breast cancer patients recruited from the NORI Hospital in Islamabad, Pakistan.

The association of CDKN2A with different cancers has been studied extensively and CDKN2A is reported to be a melanoma-prone tumour suppressor (Agarwal et al., 2012; Helgadottir et al., 2016) that is mutated in different cancers (Oliveira et al., 2005). Various studies have investigated the association of CDKN2A variants with breast cancer. Some showed reasonable association (Silva et al., 2003; Phylloides Tumour Consortium, 2008; Jovanovic et al., 2010; Spitzwieser et al., 2017), but most have shown little or no clear association with breast cancer (Berns et al., 1995; Musgrove et al., 1995; Nagore et al., 2009; Aftab et al., 2019). Previously, the nonsense mutation p.R80X was found to be associated with ductal and metaplastic breast carcinoma (Guerini-Rocco et al., 2016; Nik-Zainal et al., 2016). Several studies have shown a role for this gene in breast cancer development by deleting CDKN2A altogether or suppressing its expression in breast cancer cell lines (Brenner and Aldaz, 1995; Tan et al., 2014; Dwyer and Clark, 2015). Surprisingly, some studies revealed that breast cancer cell lines have more mutations than the primary and germline breast cancer samples (Xu et al., 1994; Sonkin et al., 2013; Morris and Chan, 2015; de Oliveira et al., 2016).

Nowadays, great effort is being made to investigate the effect of intronic variants on gene splicing and expression, which ultimately lead to apparent phenotypic changes in human (Khurana et al., 2016). Previously, the two intronic mutations c.150+2T>C and c.151-1G>C showed a significant effect on breast cancer (Prowse et al., 2003;
Hollestelle et al., 2010). Not all intronic or exonic variants are pathogenic: in a recent study the two deep intronic variants c.458-82delG and c.458-554T>G showed no significant effect on splicing, and thus did not affect gene or protein expression. Furthermore, c.458-82delG was also predicted to be non-pathogenic as it does not cause any frameshift or change in protein conformation. However, we found that two patients had compound forms of mutations which are believed to be pathogenic. Liu et al. (2015) concluded in their work that multiple mutations in a single gene can lead to a higher severity of the disease than single mutations. We believe that the occurrence of a solo mutation in patients may not be pathogenic, but its occurrence as a compound with other mutation(s) may produce a significant effect on disease pathology. Compound mutations could cause severe disruption of splicing or translation, and ultimately lead to a severe phenotype (Kim et al., 2016). Likewise, Fu et al. (2018) have reported that a compound heterozygous mutation in the USH2A gene was pathogenic.

The 3’ UTR has a crucial role in gene regulation and expression (Kitts and Sherry, 2011). rs11515 has a higher minor allelic frequency, 0.16, in Africans (Kitts and Sherry, 2011). Several studies have shown a potential role for rs11515 in tumour development and a functional relationship between rs11515 and a microRNA binding site (Chen et al., 2008; Landi et al., 2008; Robertson et al., 2010; Turgut et al., 2017). It has been reported that rs11515 is commonly associated with aggressive breast cancer in patients from New Zealand (Royds et al., 2016) although, in contradiction, a recent meta-analysis did not show an association with breast cancer in an Asian population (Dong et al., 2017). Furthermore, the rs11515 region has a recognition motif for miR-601 (Chen et al., 2008; Landi et al., 2008; Robertson et al., 2010; Turgut et al., 2017). The interaction between miRNA and rs11515 thus needs further investigation, and large population-based studies are needed to assess the association of rs11515 and other polymorphisms in the Pakistani breast cancer population.

In summary, the variants identified in this research can assist breast cancer prognosis in the Pakistani population. The current findings are consistent with previous findings, i.e., CDKN2A is not frequently mutated in breast cancer patients (Horcasitas et al., 2017; Aftab et al., 2019), but here we present two novel variants (c.458-554T>G and c.458-82delG) in CDKN2A. The cross-sectional study design and low sample size are major limitations of the study. Further screening of additional breast cancer samples will clarify the pathogenic role of the identified variants.

**CONFLICT OF INTEREST**

All authors declare that they have no conflict of interest.

**REFERENCES**


Hollestelle, A., Nagel, J. H. A., Smid, M., Lam, S., Elstrodt,


