Association of rs1042522 Polymorphism with Increased Risk of Prostate Adenocarcinoma in the Pakistani Population and its HuGE Review

Mohammad Haroon Khan1*, Hamid Rashid1, Qaiser Mansoor2, Abdul Hameed2, Muhammad Ismail2

Abstract

Prostate adenocarcinoma is one of the leading causes of cancer related mortality in men but still limited knowledge is available about its associated functional SNPs including rs1042522 (Pro72Arg). The present study was undertaken to explore the association of this SNP with susceptibility to prostate adenocarcinoma along with its structural and functional impacts in the Pakistani population in a case-control study. Three-dimensional structure of human TP53 with Pro72Arg polymorphism was predicted through homology modeling, refined and validated for detailed structure-based assessment. We also carried out a HuGE review of the previous available data for this polymorphism. Different genetic models were used to evaluate the genotypes association with the increased risk of PCa (Allelic contrast: OR=0.0.34, 95%CI 0.24-0.50, p=0.000; GG vs CC: OR=0.17, 95%CI 0.08-0.38, p=0.000; Homozygous: OR=0.08, 95%CI 0.04-0.15, p=0.000; GC vs CC: OR=2.14, 95%CI 1.01-4.51, p=0.046; Recessive model: OR=0.10, 95%CI 0.05-0.18, p=0.000; Log Additive: OR=3.54, 95%CI 2.13-5.89, p=0.000) except the Dominant model (OR=0.77, 95%CI 0.39-1.52, p=0.6). Structure and functional analysis revealed that the SNP in the proline rich domain is responsible for interaction with HRMT1L2 and WWOX. In conclusion, it was observed that the Arg coding G allele is highly associated with increased risk of prostate adenocarcinoma in the Pakistani population (p=0.000).

Keywords: p53 - prostate - polymorphism - rs1042522 - SNP

Introduction

Cancer is a global public health concern due to its fatality, disease burden and potential for increased incidence (Acikgoz and Ergor, 2013; Sahin, 2013, Tan and Polat, 2013). Evolution of cells towards cancerous one involves complex multi step genetic and epigenetic aberrations, conferring selective advantages. Regardless of the substantial research efforts made over the past decades, molecular perceptive of cancer is still a major challenge (Rivlin et al., 2011). Despite the vast diversity among the genes responsible for tumorigenesis, TP53, encoding a hub protein standout as a key player in stress responses, preservation of genomic stability and tumor suppression (Reinhartd and Schumacher, 2012; Xu et al., 2012). It is also a prime regulator of multiple signaling pathways (Levine and Oren, 2009) and has a rigid correlation with almost all kinds of human malignancies.

Prostate adenocarcinoma is one of the most commonly diagnosed cancers and a leading cause of cancer mortality in men (Siegel et al., 2011; Tafrihi et al., 2014). PCa risk factors Identification is vital for the development of potential remedies and to further our perceptive against it. Due to the complexity of origin and causes of PCa, it is difficult to pinpoint the root causes but a variety of risk factors including advanced age, culture, environmental variations and genetic alterations have been identified as major contributors in its development (Zhang et al., 2012). The absolute range of PCa alterations are still not characterized (Berger et al., 2011) and thus only limited data is available in relation to the functional SNPs association in their respective pathways (Hu et al., 2013; Karimpur-Zahmatkesh et al., 2013; Zhang et al., 2014).

rs1042522 is a common variant of TP53 at codon 72 (CGC to CCC) in exon 4, alter the p53 activities and has found differentially distributed worldwide. The SNP results in proline to arginine substitution in the proline rich region, vital for p53-mediated apoptosis (Ricks-Santi et al., 2012). Pro72Arg polymorphism of p53 have already been exploited in different international studies with
Mohammad Haroon Khan et al


reference to different ethnic groups but unfortunately, no literature is available on the subject of its allelic frequencies and its association with PCa in Pakistani population. Critical review of the relevant literature on rs1042522 polymorphism from diverse ethnic groups provoked us to investigate the status of rs1042522 polymorphism in PCa patients and its association with increased risk in Pakistani population.

Materials and Methods

Approval of the study

This study was approved by the Ethics Committee and Institutional Review Committee of the Institute of Biomedical and Genetic Engineering (IB&GE), Islamabad, Pakistan. All the samples were collected along with the signed informed consents from the donors participating in this study.

Sampling

Peripheral blood samples included 146 PCa cases and 107 controls of the same ethnic and age groups were collected from distinct random individuals of Pakistan with informed signed consent. Clinical characteristics of the subjects including Gleason grade, PSA level and diagnosis of prostatitis were obtained from their medical records. The PCa patients were between 45 to 90 years of age with mean age of 71.01±12.05, which was significantly greater than the controls (age ranges from 38-85 with mean age of 66.08±12.26) (Table 1). It was also observed that all the controls had PSA levels smaller 4.0 ng/ml and normal digital rectal exams. Individuals with benign prostatic hyperplasia were not considered in this study. Genomic DNA was extracted using the standard pheno-chloroform method.

Genotyping

Primers were designed for TP53 exon 4 by using the online resource of Primer3 web version 3.0.0 (http://primer3.wi.mit.edu). PCR was carried out to amplify the DNA sequence including p53 codon 72 by using (Forward) 5’TGCTCTTTTTACCCATCTC3’ and (Reverse) 5’ACTTCAATGCCTGGCCGTAT3’ primers. The detection of rs1042522 polymorphism was carried through RFLP to discriminate the C from G allele. The purified amplified fragments were digested with restriction enzyme BstU1 (Fermentas, Vilnius, Lithuania) at 57°C for 16hrs. The digested fragments were electrophoresed on 4% agarose gel and then visualized under UV light after ethidium bromide staining. The expected fragment sizes were, the Pro uncut allele had a band of 353bp, the Arg allele which is recognized by BstU1 was digested into two fragments of 212 and 141bp while the heterozygote had all the three bands. For quality control, 25% samples were repeated and >95% concordance was recorded. After PCR-RFLP, the samples having Pro/Arg or Arg/Arg bands were sequenced through ABI genetic analyzer for further confirmation and validation of the observed results.

Statistical methods

The statistical analysis for the case-control study was carried out with the help of online available tools. Differences in subjects were compared through Fisher’s exact test. Hardy-Weinberg equilibrium (HWE) with chi-square was calculated by comparing the expected genotype frequencies to the observed frequencies. Unconditional logistic regression was used for the calculation of odd ratios (OR) with 95% confidence intervals to estimate the effect of the SNP presence on disease risk. Two-tailed p-values ≤0.05 was hypothesized to be statistically significant.

Structure and function analysis

Nucleotide alteration can significantly affect the structure and function of a gene product. Associating these alterations with phenotypic characters is one of the important areas of research (Waheed et al., 2012). Sequence of Human p53 protein with accession number P04637 was retrieved from the UniProt (www.uniprot.org) database for detailed structure based assessment. Proteins naturally exist in complex folded structures rather than as linear polypeptides. These high order 3D conformations define their biological functions. For a detail insight, three-dimensional structure of human p53 was predicted through Modeller 9.11 (Eswar et al., 2006), refined through ModRefiner (Xu and Zhang, 2011) for quality enhancement and was validated for quality assurance through PROCHECK (Laskowski et al., 1993) and APOLLO (Wang et al., 2011).

The confirmed polymorphism, Pro72Arg was substituted in the native sequence of p53 using MUTATE_MODEL, available as a Python script at http://salilab.org/modeller/wiki/Mutate_model, to get the altered protein for investigating structural and functional deviations. The mutant model was compared against the native in 3D through PDBeFOLD (http://www.ebi.ac.uk/msd-srv/ssm) for structural similarities. Amino acid substitution can cause physiochemical differences, which in turn can affect the protein interactions, therefore physiochemical properties were predicted through ProtParam (http://au.expasy.org/tools/protparam.html). SIFT (Ng and Henikoff, 2003), PolyPhen2 (Adzhubei et al., 2010), Mutation Assessor (Reva et al., 2011) and PROVEAN (Choi et al., 2012) were used for the analysis of functional impact of altered residue due to polymorphism.

Meta analysis

It was observed in the critical review about the topic that, some of the published articles reported non-significant association of TP53 P72R polymorphism and increased risk of PCa. We therefore carried out a systematic review and meta-analysis of the previous available results of TP53 P72R polymorphism in association with PCa from different countries to explore the notorious available data. This detailed study will help to further our knowledge about the diseases development and the biological phenomenon enhancing the risk of PCa which in turn will be used as cancer diagnostics and therapeutics.

The data searched was carried out through different databases like Pubmed, Science direct, Human genetic mutation database, Chinese biomedical literature and Google scholar from 1999 to 2013. The keywords

used to search relevant studies were TP53 P72R polymorphism OR TP53 Pro72Arg polymorphism OR TP53 single nucleotide polymorphism OR TP53 codon 72 polymorphism OR rs1042522 and prostate cancer OR PCa OR prostate adenocarcinoma. Full text papers were retrieved wherever possible and the selection criteria for research articles were then narrowed down by keeping some of the important parameters constant and publications incorporating the full desired information were included in the Meta analysis. All the publications were manually examined for inclusion according to the following parameters, i) only case-control studies were selected; ii) having full information of Author/s; ii) Year of publication with country and population under study; iii) Total number of cases in the study; iv) cases genotypes; v) Total number of controls and; vi) Genotypes of the controls. A flow diagram was constructed for the collected data included in the Meta-analysis according to the PRISMA statement (Figure 4). The data was carefully taken as whole exact results from each eligible study. The genotypes of cases and controls were analyzed and the Meta analysis was performed for the total included data. Both the fixed and random effect model along with 95% CI was used to measure the genetic association and also to calculate the pooled effect estimates.

Although we have set very precise and strict inclusion criteria, there still exists heterogeneity due to a number of potential factors like the study design etc. To assess this inter study heterogeneity in a more precise fashion, I-squared, Cochran’s Q and Chi2-p statistic was used to quantitatively evaluate the proportion of the total variation due to heterogeneity. The Forest plot was used for the graphical representations of the meta-analysis results while funnel plot for the publication biasness. Forest plots and funnel plots were constructed on the basis of different genotypic models.

Results

A total of 212 samples including 120 PCa cases and 92 controls were analyzed through PCR-RFLP. The amplified fragments of TP53 containing the exon-4, when digested with BstU1 restriction enzyme followed by electrophoresis, containing the homozygous pro allele produced a single band of 353bp, homozygous arg allele produced two fragments of 212 and 141 bp while the heterozygous samples produced three bands of 353, 212 and 141 bp. A representative pattern of PCR-RFLP and sequencing is depicted in Figure 1 while Table 1 is representing the clinical features and Table 2 is presenting the cases and controls.

It has been clear from the research that biological function of a protein is usually correlated with its sequence and so any change in the sequence will impose changes in its biological function by means of changing the high order conformations. Human p53 structural variation due to Pro72Arg polymorphism present in our samples was therefore investigated in the present study and it was found that the substitution is in the proline rich domain, responsible for Interaction with HRMT1L2 and WWOX.

The biochemical differences, nature and location of

<table>
<thead>
<tr>
<th>Table 1. Clinical Features of the PCa Patients and Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factors</strong></td>
</tr>
<tr>
<td><strong>No. %age</strong></td>
</tr>
<tr>
<td>Age in years</td>
</tr>
<tr>
<td>BMI (kg/m2)</td>
</tr>
<tr>
<td>PSA level (ng/ml)</td>
</tr>
<tr>
<td>Smoking</td>
</tr>
<tr>
<td>Hard Drinks</td>
</tr>
<tr>
<td>Disease stage Localized</td>
</tr>
<tr>
<td>Disease stage Locally advance</td>
</tr>
<tr>
<td>Disease stage Bone metastasis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Genotypes and Allelic Frequencies of TP53 Pro72Arg Polymorphism in Prostate Adenocarcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genotype</strong></td>
</tr>
<tr>
<td><strong>No. %age</strong></td>
</tr>
<tr>
<td>C-allele</td>
</tr>
<tr>
<td>G-allele</td>
</tr>
<tr>
<td>CC</td>
</tr>
<tr>
<td>CG</td>
</tr>
<tr>
<td>GG</td>
</tr>
</tbody>
</table>

Figure 2. Representation of A) Structures of Proline and Arginine residues; B) Partial diagram representing secondary structure comparison of native and polymorphic p53 proteins. The red window shows the amino acid substitution but has no affect on the structure; C) Partial diagram representing alignment of secondary structure elements of native and polymorphic p53 proteins. The red window shows that the polymorphism has no affect on the structure elements.
Table 3. Association of TP53 Pro72Arg Polymorphism and Risk of Prostate Adenocarcinoma

<table>
<thead>
<tr>
<th>Genetic Model</th>
<th>OR</th>
<th>LCI</th>
<th>UCI</th>
<th>Z-stat</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allelic contrast (C vs G)</td>
<td>2.90</td>
<td>1.20</td>
<td>4.23</td>
<td>5.55</td>
<td>0.000</td>
</tr>
<tr>
<td>CC vs CG</td>
<td>5.91</td>
<td>2.63</td>
<td>13.28</td>
<td>4.30</td>
<td>0.000</td>
</tr>
<tr>
<td>CC vs CG</td>
<td>0.47</td>
<td>0.22</td>
<td>0.99</td>
<td>-1.99</td>
<td>0.460</td>
</tr>
<tr>
<td>CG vs GG</td>
<td>12.63</td>
<td>6.46</td>
<td>24.69</td>
<td>7.41</td>
<td>0.000</td>
</tr>
<tr>
<td>Dominant (CC+CG vs GG)</td>
<td>10.18</td>
<td>5.45</td>
<td>19.45</td>
<td>7.27</td>
<td>0.000</td>
</tr>
<tr>
<td>Recessive (CC vs GG+CG)</td>
<td>1.29</td>
<td>0.66</td>
<td>2.54</td>
<td>0.74</td>
<td>0.460</td>
</tr>
<tr>
<td>Log additive GG/CC/CG</td>
<td>1.24</td>
<td>0.64</td>
<td>2.41</td>
<td>0.63</td>
<td>0.530</td>
</tr>
</tbody>
</table>

*OR=odd ratio; LCI=95% lower confidence interval; UCI=95% upper confidence interval

---

Table 4. Details of Studies Included in the Meta-Analysis

<table>
<thead>
<tr>
<th>Author/s</th>
<th>Population</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Henner et al., 2001</td>
<td>Caucasian</td>
<td>109</td>
<td>66  41</td>
</tr>
<tr>
<td>Suzuki et al., 2003</td>
<td>Japanese</td>
<td>114</td>
<td>46  48</td>
</tr>
<tr>
<td>Huang et al., 2004</td>
<td>Taiwanese</td>
<td>200</td>
<td>92  42</td>
</tr>
<tr>
<td>Wu et al., 2004</td>
<td>Taiwanese</td>
<td>92</td>
<td>20  11</td>
</tr>
<tr>
<td>Leiros et al., 2005</td>
<td>Caucasians</td>
<td>39</td>
<td>2   17</td>
</tr>
<tr>
<td>Quinones et al., 2006</td>
<td>Chile</td>
<td>60</td>
<td>14  22</td>
</tr>
<tr>
<td>Hirata et al., 2007</td>
<td>Japanese</td>
<td>167</td>
<td>22  89</td>
</tr>
<tr>
<td>Hirata et al., 2009</td>
<td>Japanese</td>
<td>140</td>
<td>20  75</td>
</tr>
<tr>
<td>Xu et al., 2010</td>
<td>Chinese</td>
<td>209</td>
<td>41  129</td>
</tr>
<tr>
<td>Ricks-Santi et al., 2010</td>
<td>African</td>
<td>245</td>
<td>73  24</td>
</tr>
<tr>
<td>Doosti &amp; Dehkordi, 2011</td>
<td>Iranian</td>
<td>187</td>
<td>15  98</td>
</tr>
<tr>
<td>Rogler et al., 2011</td>
<td>Caucasian</td>
<td>118</td>
<td>9   44</td>
</tr>
<tr>
<td>Our Study, 2013</td>
<td>Pakistani</td>
<td>146</td>
<td>27  101</td>
</tr>
</tbody>
</table>

*Caucasian; **Southern Chinese; ***African descent

---

Table 5. Log Odd Ratios, Standard Error, Variance and p-value of the Studies Included in the Meta-analysis Under Allelic Contrast and Homozygous Genotypic Models

<table>
<thead>
<tr>
<th>Author</th>
<th>Allelic Contrast</th>
<th>CC vs GG</th>
<th>CC vs GC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Henner et al., 2001</td>
<td>0.15</td>
<td>1.07</td>
<td>0.78</td>
</tr>
<tr>
<td>Suzuki et al., 2003</td>
<td>0.17</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>Huang et al., 2004</td>
<td>0.49</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Wu et al., 2004</td>
<td>0.06</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>Leiros et al., 2005</td>
<td>-0.31</td>
<td>-0.67</td>
<td>-0.67</td>
</tr>
<tr>
<td>Quinones et al., 2006</td>
<td>0.56</td>
<td>1.06</td>
<td>1.05</td>
</tr>
<tr>
<td>Hirata et al., 2007</td>
<td>0.01</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Hirata et al., 2009</td>
<td>0.09</td>
<td>0.37</td>
<td>0.37</td>
</tr>
<tr>
<td>Xu et al., 2010</td>
<td>0.17</td>
<td>-0.31</td>
<td>-0.31</td>
</tr>
<tr>
<td>Ricks-Santi et al., 2010</td>
<td>-0.26</td>
<td>-0.48</td>
<td>-0.48</td>
</tr>
<tr>
<td>Doosti &amp; Dehkordi, 2011</td>
<td>0.01</td>
<td>-0.37</td>
<td>-0.37</td>
</tr>
<tr>
<td>Rogler et al., 2011</td>
<td>0.17</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>Our Study, 2013</td>
<td>0.17</td>
<td>0.78</td>
<td>0.78</td>
</tr>
</tbody>
</table>

*LOR=Log odd ratio; SE=Standard error; Var=Variance; p=p-value

---

Table 6. Log Odd Ratios, Standard Error, Variance and p-value of the Studies Included in the Meta-analysis Under Heterozygous, Dominant and Recessive Genotypic Models

<table>
<thead>
<tr>
<th>Author</th>
<th>GC vs GG</th>
<th>Dominant model</th>
<th>Recessive model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Henner et al., 2001</td>
<td>2.09</td>
<td>1.81</td>
<td>-0.13</td>
</tr>
<tr>
<td>Suzuki et al., 2003</td>
<td>-0.37</td>
<td>-0.13</td>
<td>0.26</td>
</tr>
<tr>
<td>Huang et al., 2004</td>
<td>0.52</td>
<td>0.66</td>
<td>0.02</td>
</tr>
<tr>
<td>Wu et al., 2004</td>
<td>1.5</td>
<td>1.34</td>
<td>-0.12</td>
</tr>
<tr>
<td>Leiros et al., 2005</td>
<td>-0.16</td>
<td>0.13</td>
<td>0.36</td>
</tr>
<tr>
<td>Quinones et al., 2006</td>
<td>0.36</td>
<td>0.56</td>
<td>0.57</td>
</tr>
<tr>
<td>Hirata et al., 2007</td>
<td>0.19</td>
<td>0.13</td>
<td>0.23</td>
</tr>
<tr>
<td>Hirata et al., 2009</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>Xu et al., 2010</td>
<td>-0.01</td>
<td>-0.21</td>
<td>0.02</td>
</tr>
<tr>
<td>Ricks-Santi et al., 2010</td>
<td>-0.07</td>
<td>-0.23</td>
<td>0.42</td>
</tr>
<tr>
<td>Doosti &amp; Dehkordi, 2011</td>
<td>-0.52</td>
<td>-0.57</td>
<td>0.42</td>
</tr>
<tr>
<td>Rogler et al., 2011</td>
<td>-0.12</td>
<td>0.22</td>
<td>0.57</td>
</tr>
<tr>
<td>Our Study, 2013</td>
<td>0.25</td>
<td>0.23</td>
<td>0.32</td>
</tr>
</tbody>
</table>

*LOR=Log odd ratio; SE=Standard error; Var=Variance; p=p-value

---

Figure 3. Graphic Representation of the Changes Enforced by A72P Polymorphism in the Native 3D Conformation

Figure 4. Flow Diagram for the Collected Data Included in the Meta-Analysis According to the PRISMA Statement
amino acid substitution can affect the protein in various ways and is therefore important to determine whether it can alter the protein function. Isoelectric focusing point (pI) and charge is important for the solubility and interaction of a protein (Nandi et al., 2005; Khaldi and Shields, 2011). Theoretical pI of the p53 proteins were predicted by using Bioinformatics tools and were observed 6.34 and 6.47 for the native and polymorphic respectively.

A total of 250 relevant journal papers associated with TP53 Pro72Arg polymorphism and prostate cancer risk were retrieved in due time allocated for data retrieval. Only 12 studies were fulfilling our required criterion as mentioned above and thus data was taken out of these in addition to our’s study for the Meta analysis. The studies included in the analysis consisted of a total of 1686 PCa cases and 1888 controls. The included studies were carried out in different populations in different countries (Table 4).

When all the included studies were pooled in the meta-analysis, there was a significant association between the increased PCa risk and TP53 Pro72Arg polymorphism in any of the genetic model [Allelic contrast (C vs G): OR=1.10, 95% CI 1.0-1.2, p=0.06, Q=66.43, I-squared=81.94% and p=0.00; CC vs GG: OR=1.29, 95% CI 1.05-1.59, p=0.02, Q=55.55, I-squared=78.38% and p=0.00; CC vs GC: OR=0.81, 95% CI 0.68-0.97, p=0.02, Q=29.87, I-squared=59.82% and p=0.00; GC vs GG: OR=1.27, 95% CI 1.08-1.48, p=0.00, Q=83.06, I-squared=85.55% and p=0.00; Dominant model: OR=1.26, 95% CI 1.08-1.46, p=0.49, Q=83.73, I-squared=85.67% and p=0.00; Recessive model: OR=0.94, 95% CI 0.80-1.11, p=0.49, Q=34.11, I-squared=64.82% and p=0.00] (Figure 5). Publications bias was estimated through both Begg’s and Egger’s test along with funnel plots (Figure 6).

Discussion

Prostate adenocarcinoma is globally a leading cause of cancer related mortality (Ricks-Santi et al., 2012), but unfortunately the molecular mechanisms underlying its development and progression still remains poorly understood like other cancers (Boyd et al., 2012). To date, the mechanisms for increased risk of PCa in Pakistani population have not yet elucidated. We genotyped the Pro72Arg SNP (rs1042522) of tumor suppressor gene TP53 in a PCa case-control study of men at the Institute of...
Biomedical and Genetic Engineering, Islamabad, Pakistan
to explore its association with PCa genetic susceptibility.

In the study samples, controls were tended to be
younger than cases with age range from 40-70 for
controls and 40-85 for cases respectively. Similarly PSA
level was also comparatively lower in the controls. The distribution of three genotypes namely, pro/pro, pro/arg
and arg/arg were observed 18.49%, 69.18% and 12.33%
in PCa patients and 14.95%, 26.17% and 58.88% in
controls respectively. Highly significant differences were
observed in the distribution of genotypes between patients
and controls (df=2, p=0.00). The allele frequencies of
patients and controls were fitted in the Hardy-Weinberg
Equilibrium with frequencies of 0.78 (Controls) and
0.27 (patients). Logistic regression was used to evaluate
the association between Pro72Arg polymorphism and
prostate adenocarcinoma. Highly significant association
was revealed by all the genetic models [allelic contrast
(C vs G): OR=2.90, 95% CI: 1.24-2.3; p=0.000; CC
vs GG: OR=5.91, 95% CI: 2.63-13.28; p=0.000; CC vs
GC: OR=0.47, 95% CI: 0.22-0.99, p=0.046; GC vs GG:
OR=12.63, 95% CI: 6.46-24.69, p=0.000; Dominant
model (OR=10.18, 95% CI: 5.45-19.04, p=0.000) except
the Recessive model: (OR=1.29, 95% CI: 0.66-2.54,
p=0.000 and log additive model: OR=1.24, 95% CI: 0.64-
2.41, p=0.53). Based on the results, it is apparent that
the genotypes are in significant association with increased
risk due to the increasing number of polymorphic alleles.

Knowledge of protein structure provides an insight
into its interactions (Aydin et al., 2011), which define
the protein’s biological role and functions (Cheng et
al., 2005). Residues substitutions by any mean can
affect the protein high order structure which determines
protein functions. Human p53 structural variation due
to Pro72Arg polymorphism present in our samples that
showed statistically significant association was therefore
investigated in the present study. Pro72Arg substitution is
in the proline rich domain, responsible for Interaction
with HRMT1L2 and WWOX. The substitution is replacing
a non-polar residue with a positively charged residue which
is bigger and less hydrophobic than the wild type which
might lead to bumps and changes in hydrophobicity and
lead to loss of hydrophobic interactions, either in the core
or surface of the protein (Figure 2).

3D structures are more conserved than sequence
(Capriott et al., 2010), thus native and polymorphic
structures were aligned in 3D which actually produces
a measure to assess the level of similarity of the aligned
structures and RMSD value of 0.0040 was observed
(Figure-3). The biochemical differences, nature and
location of amino acid substitution can affect the protein
in various ways and is therefore important to determine
whether it can alter the protein function. Isoelectric
focusing point (pI) and charge is important for the
solubility and interaction of a protein (Nandi et al., 2005;
Khaldi and Shields, 2011). It is clear from the research
that, pl of a protein can vary due to insertions, deletions,
substitutions and the ecology of the organism (Kiraga
et al., 2007). Theoretical pl of the p53 proteins were
predicted by using Bioinformatics tools and were observed
6.34 and 6.47 for the native and polymorphic respectively.

Theoretical pl of the polymorphic p53 protein was higher
than the native by 0.14, which is equivalent to 1-2 net
positive charge increase per molecule. Differences were
also observed in the aliphatic index and GRAVY. The
substitution Pro72Arg was predicted to be TOLERATED
with a score of 0.43 by SIFT, regarded as benign by
PolyPhen2 with scores of 0.000 (sensitivity: 1.00 and
specificity: 0.00), neutral by mutation assessor and
PROVEAN with a PROVEAN score of-0.230.

The results showed that, the Arg coding G allele
was extremely significantly associated with the disease
prevalence in our sampled population. To the best of our
knowledge, this is a pioneer study to test the association
of this very common polymorphism of TP53 and PCa
risk in Pakistani population. Regression analysis was
specifically performed on the data by using different
models to critically pinpoint the alleles association with
PCa risk. Extremely significant associations were revealed
by all the studied models between the SNP and PCa risk.

The TP53 Pro72Arg SNP has found to be related with
changes in the efficiency and function of TP53 product
(Murphy, 2006; Shi et al., 2009; Whibley et al., 2009).
The pro coding C allele is responsible for an enhanced
transcriptional transactivation which thus induces elevated
cell-cycle arrest in G1 (Thut et al., 1995) while the G
which is responsible for arg residue promotes induction of
apoptosis (DuMont et al., 2003). It has also been observed
that, the Arg residue, which is due to G allele, denatures
at high temperatures and is comparatively less stable
thermodynamically (Khoo et al., 2009) and has been
hypothesized to be under selective pressure (Kiraga et al.,
2007). The thermodynamically stable C allele responsible
coding Pro residue promotes the induction of TP53 for
repair against oxidative damage in populations living in
hot climates (Khoo et al., 2009). It has been proposed that
C allele could provide selective advantage in populations
living in comparatively colder areas around the globe by
reducing implantation failure, as in the case of Pakistani
population, providing better ability for the induction of
cell-cycle arrest (Bensad et al., 2006). Depending upon
the ethnicity of the population studied, G allele responsible
coding Arg residue have been found to be extremely
associated with increased risk of prostate adenocarcinoma.

A large number of case-control studies have already
been conducted to assess the association TP53 Pro72Arg
polymorphisms and increased risk of different cancers
including PCa in humans. The results of all the reported
studies are inconsistent, a comprehensive meta-analysis
was therefore included in this study to provide further
insights and to further explore this debated area. Majority
of the previous publications reported the polymorphism
non-significantly associated with PCa risk, while others
found it significantly associated. The meta-analysis
will help to evaluate the potential association of the
polymorphism and increased PCa risk and to achieve
more reliable conclusion. The results of meta-analysis
showed statistically significant heterogeneity among the
included studies which may be due to different potential
factors. Careful investigation of the results in this study
reflects that population under study is the most important
contributing factor for this heterogeneity which is due to
the differential genotypes distribution of TP53 Pro72Arg polymorphisms among different populations. In conclusion, we examined the TP53 Pro72Arg polymorphism and the association of PCa risk in a case-control study of Pakistani population. It was observed that the TP53 CGC to CCC polymorphism in exon-4 is associated with the increased risk of PCa which was also confirmed through the meta-analysis of the published data. It is therefore recommended that TP53 Pro72Arg polymorphisms may be a potential biomarker for PCa. It can thus be concluded that depending upon the environment, nature favors the selection of one allele over the other during the course of evolution. On the basis of this systematic study, it is hoped that differences in allelic frequencies in different populations associated with a large number of health problems can provide a valuable foundation for unifying the mechanisms of complex diseases like PCa and others.

Acknowledgements

We would like to thank the lab assistants of sampling lab from IRNUM hospital, Islamabad, Pakistan for their cooperation in blood sampling. We also extend our thanks to Ms. Raisa Bano for her technical assistance. We are also thankful to IBandGE for providing the space and financial support.

References


Mohammad Haroon Khan et al


