Original Article

Screening of Type 2 Diabetes Patients of Khyber Pakhtunkhwa for CDKAL1 Variant (rs10946398)

Maqsood Ali¹, Haji Bahadar*, Haseenullah Shah¹, Sajjad Afridi¹, Sajid Ali¹, Kiran Ijaz², Fazli Khuda³, Hamayun Khan⁴, Mohsin Raziq⁵ and Zakiullah⁶

¹Department of Pharmacy, University of Peshawar, Peshawar, Pakistan
²Institute of Pharmaceutical Sciences, Khyber Medical University, Peshawar, Pakistan
³Department of Biotechnology, Abdul Wali Khan University, Mardan, Pakistan

INTRODUCTION

Diabetes Mellitus (DM) is a metabolic disorder distinguished by chronically high blood glucose levels. Decreased insulin secretion and insufficient response of cells to insulin called insulin resistance, or both of the two factors can cause DM [1, 2]. According to the International Diabetes Federation (IDF) 8.3% of the world's adult population, will have diabetes by 2030, up from the current estimated 371 million cases [3]. Only 50% of persons with diabetes are aware of it, and among those who do, 90% have type 2 diabetes, whereas only 10% have type 1. According to the National Diabetes Fact Sheet for 2011, 8.3% of Native Americans, regardless of age, have diabetes, with 11.3% of those affected being adults aged 20 and older. Around 215,000 people under the age of 20 had diabetes in 2010, primarily either Type 1 DM (T1DM) or Type 2 DM (T2DM), with those 165 and older accounting for 25% of the total. 7 million Americans are estimated to have diabetes, 27% of whom have not yet been given a diagnosis. In Europe, the prevalence of diabetes is 6 to 8%, affecting 55 million people, yet only 21.2 million (or 39% of this population) have received a diagnosis [3-5]. About 38 million people die from diabetes each year worldwide, making it one of the top four causes of mortality. Eighty percent of all NCD morbidity occurs in developing and underdeveloped nations. In 2025, the prevalence of Diabetes will rise to 5.4%, a sign of both the disease's expanding prevalence and the mounting cost and...
major threat to the community [3, 6]. As per the International Diabetes Federation report, in Pakistan, in 2021, the diabetes prevalence rate in adult was 26.7%, positioning it among the highest globally. This alarming situation highlights the urgent need for targeted healthcare strategies. The situation demands immediate attention towards genetic predispositions and lifestyle factors contributing to the rising trend, emphasizing the importance of both prevention and management in the Pakistan. Recent epidemiologic studies have reported a high relationship between genetic variables and its association with T2DM, even though environment has significant effects in pathophysiology of the disease. Researchers have been attempting to identify genes that raise T2DM risk since the 1980s, but minimal progress was made until the recent development of GWAS (Genome-Wide Association Studies) due to the complicated pathophysiology of T2DM. The discovery of T2DM susceptibility loci and the improvement of our understanding of T2DM pathophysiology have both been considerably aided by GWAS. GWAS (Genome-Wide Association Studies) increased the count of T2DM (Type 2 Diabetes Mellitus) risk genes from the previously identified two, using earlier techniques, to over 60. CDKAL1, FTO, TCF7L2, PPARG, IGF2BP2, NOTCH2, SLC30A8, WFS1, JAZF1, and KCNJ11 are a few examples of the genes with substantial associations with Diabetes [7]. The 37 kb long CDKAL1 gene, which is found at 6p22.3, encodes 579 amino acids. The cyclin-dependent kinase 5 (CDK5) regulatory subunit-associated protein 1-like 1 (CDKAL1) gene’s intron 5 has a collection of single-nucleotide polymorphisms which is linked to type 2 diabetes risk [8]. The CDKAL1 gene encodes a tRNA decorating enzyme called methyl transfer enzyme, which produces the 37th base of tRNA Lys’s 2-methylthio-N6-threonylcarbamoyladenosine (UUU). The 37 kb long CDKAL1 gene, which is found at 6p22.3, encodes 579 amino acids. The CDKAL1 gene codes for a tRNA-modifying enzyme known as a methyltransferase, responsible for synthesizing the 37th base of tRNA Lys’s 2-methylthio-N6-threonylcarbamoyladenosine (UUU) [9]. It has been proposed that Cyclin-dependent kinase 5 (CDK5) operates more efficiently when the expression of CDKAL1 is reduced. Consequently, the interaction of CDKAL1 with the CDK5 activator p35 might regulate insulin secretion induced by CDK5 [10, 11]. In adipose tissue, CDKAL1 is required for normal mitochondrial morphology and operation. These findings imply that CDKAL1, a type 2 diabetes susceptibility gene, performs novel roles in controlling mitochondrial activity, and hence impaired ATP generation may lead to decrease first phase insulin secretion and ultimately to T2DM [9]. The following variants show association with Diabetes Mellitus rs10946398, rs7758992, rs9465871, rs4712523, rs4712524. Among them the variants of CDKAL1 rs10946398 is highly susceptible and possess strong association with Diabetes Mellitus[10, 11]. However, till date, no such study has been reported in the population of Khyber Pakhtunkhwa, Pakistan. The goal of this study was to determine the association of CDKAL1 (rs10946398) with T2DM in the Khyber Pakhtunkhwa population of Pakistan. Therefore, this case control study has been designed to assess the presence of CDKAL1 (rs10946398) in local T2DM patients of Khyber Pakhtunkhwa, Pakistan.

**METHODS**

**Screening of Subjects and Blood Sample Collection**

Using the assumptions of a 5% threshold of significance, 90% power, and a 4 odds ratio from Table 10 of the WHO’s sample size calculation manual, the sample size for the case-control study was determined. The Department of Pharmacy at the University of Peshawar’s ethics committee granted clearance (411/EC/LIFE/UOP-2021), dated October 28, 2021, to proceed. 200 patients and healthy persons were selected from the tertiary care facilities in Peshawar, Khyber Pakhtunkhwa. Patients and controls both agreed to participate in the study by signing a carefully crafted consent form based on variety of aspects connected to T2DM which includes family history, economic status, age, and way of life, vital signs, diagnosis, complications, and past medical history. The following list of criteria was used to select study participants for inclusion and exclusion [12]. Guidelines for Inclusion criteria were T2DM patients with family history, regardless of gender and T2DM individual with age ≥ 30 years. Guidelines for Exclusion criteria were Diabetic patients who are insulin dependent, T2DM sufferers without a familial history, Diabetic expectant mothers (gestational diabetes). Patients with T2DM under 30 years of age and diabetics who also have a chronic illness, such as HIV with chronic disease like HIV. Three milliliters (3ml) of whole blood samples were collected, and placed in properly labeled EDTA tubes for preservation at -20°C.

**PCR Protocol for Detecting Polymorphism**

**DNA Extraction**

For the isolation of DNA from whole blood samples, we utilized the WizPrepTM gDNA Mini Kit, following the manufacturer’s protocol. Each participant provided 3ml of whole blood, collected in EDTA tubes. The extracted DNA was quantified for purity and concentration using a spectrophotometer, then stored at -20°C until further analysis. This approach ensures the integrity of the DNA for subsequent polymerase chain reaction (PCR) processes.

**Primer Design**

The CDKAL1 gene region containing the rs10946398 polymorphism was amplified using specific primers.
designed for high specificity and efficiency. The forward primer sequence was 5’-CTGCTTGCTGTGGGGAAGA-3’, and the reverse primer sequence was 5’-CTCAATGCTGTTCATAGGCAC-3’. These primers were validated through in silico analysis using the NCBI primer BLAST tool, ensuring their specificity to the target sequence and minimizing off-target effects.

**PCR Amplification**

The PCR amplification was conducted in a 20µl reaction mixture, consisting of 10 ng of genomic DNA, 10 pmol of each primer, 200 µM of each dNTP, 1x PCR buffer, and 1.5 U of Taq DNA polymerase. The amplification protocol initiated with an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 30 seconds, and extension at 72°C for 1 minute. A final extension phase at 72°C for 5 minutes concluded the PCR. This protocol was optimized to ensure the amplification of the CDKAL1 gene region encompassing rs10946398.

**Gel Electrophoresis**

The PCR products were verified through agarose gel electrophoresis to confirm the successful amplification of the target DNA segment. A 2% agarose gel was prepared, into which 5µl of the PCR product mixed with loading dye was loaded per well. The gel ran at 100V for 30 minutes in 1x TAE buffer. A 100 bp DNA ladder was used as a molecular weight marker to estimate the size of the PCR products. The expected size of the amplified CDKAL1 fragment was approximately 157 bp. This method allows for the visual confirmation of PCR product integrity and specificity, critical for accurate genotyping in the subsequent analysis.

**Restriction Digestion for Genotype Identification of CDKAL1 Gene (rs10946398)**

**Enzyme Selection**

The Acil restriction enzyme was specifically chosen for its ability to recognize and cleave the polymorphic site within the CDKAL1 gene that encompasses the rs10946398 SNP. This enzyme facilitates the distinction between the various alleles based on their digestion patterns, allowing for accurate genotype identification. The selection was based on the unique recognition site present in the allele variants, making Acil highly suitable for differentiating between the genotypes associated with type 2 Diabetes Mellitus susceptibility.

**Digestion Process**

For the restriction digestion process, 10µl of the PCR product was mixed with 1µl (10 units) of Acil restriction enzyme and 2µl of the appropriate 10x reaction buffer. The volume was brought up to 20µl with nuclease-free water. This mixture was then incubated at 37°C for 16 hours to ensure complete digestion of the PCR products. The precise conditions for the digestion reaction were optimized to achieve efficient and specific cleavage of the DNA, thereby allowing for the subsequent genotype determination.

**Visualization**

Following the restriction digestion, the products were analyzed using agarose gel electrophoresis to visualize the genotype-specific fragment sizes. A 3% agarose gel, prepared in 1x TAE buffer and stained with ethidium bromide, was used for this purpose. The digested samples were loaded into the wells of the gel alongside a 50bp DNA ladder, enabling the identification of fragment sizes corresponding to the different CDKAL1 genotypes. The expected fragment sizes for the rs10946398 SNP after digestion with Acil are as follows:

- For the GG genotype (non-risk allele), a single band of 157bp, indicating no cleavage by the enzyme.
- For the GC genotype (heterozygous risk allele), two bands of 121bp and 36bp, showing partial digestion.
- For the CC genotype (homozygous risk allele), two bands of 121bp and 36bp, similar to the GC genotype but indicating a different allele composition.

The visualization of these distinct patterns on the agarose gel allows for the straightforward identification of the individual’s genotype at the rs10946398 locus of the CDKAL1 gene. This genotype information is crucial for understanding the genetic predisposition to Type 2 Diabetes Mellitus within the studied population.

**Data Evaluation**

The connection of the CDKAL1 (rs10946398) gene in T2DM patients and controls was examined statistically using SPSS version 26. The odds ratio (OR) was calculated in SPSS version 26.0 using regression analysis, with a P-value of 0.05 or less being regarded as significant.

**RESULTS**

100 cases of diagnosed patients of type 2 Diabetes Mellitus were considered in this study which comprised of 65% males and 35% females. Apart from the 100 cases, 100 samples were taken consists of 77% males and 23% females as control group. Average age for our T2DM participants was kept 51 ± 10 years. Table 1 indicates the demographic features of our study participants.

**Table 1: Percentage of Onset by Age and Gender among Research Participants**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Type</th>
<th>Case</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants Age</td>
<td></td>
<td></td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>31-40 yrs</td>
<td>6</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41-50 yrs</td>
<td>41</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-60 yrs</td>
<td>34</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>61-70 yrs</td>
<td>19</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71-80 yrs</td>
<td>0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2 shows the clinical features of the study participants in terms of different comorbidities.

Table 2: Comorbidities among Participants

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Comorbidities</th>
<th>Case (%)</th>
<th>Control (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Renal Diseases</td>
<td>05</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>High cholesterol</td>
<td>06</td>
<td>03</td>
</tr>
<tr>
<td>3</td>
<td>Blood Pressure</td>
<td>34</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Retinal Damage</td>
<td>39</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Heart Diseases</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Hepatitis B Virus (HBV)</td>
<td>00</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Hepatitis C Virus (HCV)</td>
<td>01</td>
<td>0</td>
</tr>
</tbody>
</table>

**CDKAL1 Gene Amplification**

The CDKAL1 gene was amplified using the standard PCR technique known as PCR-SSP. After the procedure was completed under the optimal circumstances to acquire precise primer binding, the Gel-doc system was used to view the PCR data. The CDKAL1 gene’s 157 bp amplicon size required the use of a 100 bp hyper ladder to view the PCR results. The technique is fully discussed in section 2.6. The electropherograms for the case and control groups of the CDKAL1 gene are shown in figures 1.

**CDKAL1 Gene SNPs after Applying Restriction Enzyme**

The restriction enzyme ACIL, a 50 bp ladder, and a 1% agarose gel were used to identify the polymorphism in the CDKAL1 gene. The PCR result yielded 157, 121, 36 base pair fragments when it was digested at the sites guanine-guanine, guanine-cytosine and cytosin-cytosin respectively. Electropherograms for the case and control groups of CDKAL1 gene SNPs are shown in figures 2 and 3.

**Risk Carrying Alleles of CDKAL1**

The results of the binary logistic regression analysis conducted on the control case column and Wild Carriage CDKAL1 column in SPSS have yielded some interesting findings. In the case of Wild Carriage CDKAL1 GG, the Sig value was found to be 0.001, with an odd ratio of 0.395 and lower and upper limits of 0.846 and 1.231, respectively. However, after adjusting for confounding factors, the Sig value increased to 0.018, and the odd ratio decreased to 0.482, with the upper and lower limits becoming 0.417 and 0.774. As for Allele Carriage CDKAL1 CC, the Sig value was found to be 0.007, with an odd ratio of 2.434 and minimum and maximum limits of 0.611 and 3.649, respectively. The odd ratio climbed to 2.837 and the Sig value to 0.045 after controlling for confounding variables; the top and lower limits of 0.646 and 1.231, respectively. The findings provide valuable insights into the relationship between the Wild Carriage CDKAL1 and Allele Carriage CDKAL1 genes and the occurrence of the disease being studied. As shown in table 4, carriage of alleles is frequent.

**Table 4: Frequency of Risk Carrying Alleles of CDKAL1 Gene**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Control N(%)</th>
<th>Case N(%)</th>
<th>Unadjusted OR (95% CI)</th>
<th>p-value</th>
<th>Adjusted OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GG</td>
<td>55 (55)</td>
<td>23 (29)</td>
<td>Ref</td>
<td>Ref</td>
<td>Ref</td>
<td>Ref</td>
</tr>
<tr>
<td>GC</td>
<td>36 (36)</td>
<td>56 (66)</td>
<td>4.833 (1.979-10.509)</td>
<td>0.027</td>
<td>4.993 (1.167-21.354)</td>
<td>0.30</td>
</tr>
<tr>
<td>CC</td>
<td>9 (9)</td>
<td>15 (15)</td>
<td>2.788 (1.545-5.033)</td>
<td>0.001</td>
<td>2.700 (1.482-4.920)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**Figure 1:** Electropherogram of the PCR product (Case) using a 50 bp ladder for the CDKAL1 gene.

**Figure 2:** PCR product electropherogram for CDKAL1 gene SNPs using a 50 bp ladder (Case)
DISCUSSION

Among diverse populations, the prevalence of T2DM varies significantly. Data from the International Diabetes Federation indicates that South Asians are more frequently to have T2DM than Europeans [14]. Therefore, it is crucial to carry out genetic research on South Asians in order to comprehend the genetics of T2DM in these populations. We examined how the SNP (rs10946398) in the CDKAL1 gene influences T2D prevalence in patients of Pakistan's KP region in the aforementioned case-control research. In the current study, we discovered that the distribution of CDKAL1 genotypes among T2DM subject's vs healthy controls differed significantly. Our findings revealed a significant, (p<0) connection between T2DM in KP patients and the CDKAL1 (rs10946398) gene polymorphism. Our results were found to be consistent with those of earlier research carried out in Han Chinese and Korean populations, which supported the connection between the CDKAL1 variation (rs10946398) and T2DM [15-17]. Similarly, according to another published research paper indicates a favorable connection between the CDKAL1 variations (rs10946398) with type 2 diabetes in Tehran-based Iranians [18]. CDKAL1 rs109463986 have been previously associated with T2DM in European-derived populations [19]. Similar to this, meta-analysis uncovered a strong relationship between CDKAL1 for rs10946398 and T2DM in populations from Europe and America. Also, we found in this meta-analysis that T2DM in populations from both Europe and America was proportional to the CDKAL1 gene variant rs10946398 [19, 20]. Furthermore, a meta-analysis research found that the rs10946398 variant of the CDKAL1 gene was strongly related with a higher risk of developing T2DM in populations from western (Europe and the United States) and Asian countries [20]. In another published study, we discovered a strong correlation between Taiwanese people's CDKAL1 rs10946398 genotype with T2DM. The connection was much stronger in CC than CA carriers. The CC genotype was strongly linked to T2DM according to stratified studies [21]. In accordance with our findings, a study conducted in the German KORA population found a connection between the CDKAL1 polymorphism and T2DM predisposition. According to a study carried out in Uttara khand, India, T2DM was more common among people with CDKAL1 heterozygous GC and mutant CC [22, 23]. Our results, which conflicted with those of earlier research that had revealed no association between CDKAL1 polymorphism and T2DM across populations in the United Arab Emirates, suggest that rs10946398 variations may not be specifically linked to T2DM susceptibility in the general population [24]. The results of the binary logistic regression analysis on the Wild Carriage CDKAL1 and Allele_Carrige_CDKAL1 genes have important implications for the understanding of the genetic basis of diabetes. The significant association of the Wild Carriage CDKAL1 GG genotype with diabetes suggests that individuals carrying this genotype may be at a higher risk of developing the disease but statistically it is non-significant. The decrease in the association after adjusting for confounding factors implies that other factors such as lifestyle, diet, and environmental factors may also play a role in the development of diabetes in individuals carrying this genotype. Similarly, the significant association of the Allele Carriage CDKAL1 CC genotype with diabetes indicates that this gene may be a strong predictor of the disease. This finding is consistent with previous studies that have shown that the CDKAL1 gene is associated with type 2 diabetes [25, 26]. The presence of the CC genotype may influence the production of insulin, leading to insulin resistance and ultimately the development of diabetes. The findings of this study provide evidence for the association of the Allele Carriage CDKAL1 CC allele with diabetes. The study highlights the importance of genetic factors in the development of diabetes, as well as the need to consider other lifestyle and environmental factors. Further research is needed to confirm these findings and to explore the underlying mechanisms by which these genes contribute to the development of diabetes.

CONCLUSIONS

The results of our study revealed a strong association of CDKAL1 (rs10946398) variant with T2DM susceptibility in the local population of Khyber Pakhtunkhwa and can be further studied as a potential genetic marker for the same.

Acknowledgements

We sincerely thank the participants, the University of Peshawar, and our fellow researchers.

Authors Contribution

Conceptualization: Z
Methodology: MA, HS, SA1
Formal analysis: KI
Writing-review and editing: HB, SA2, FK, HK, MR

All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Source of Funding

PMU1-22/HEREF/2014-15/Vol-I/V/3408 the Higher Education Research Endowment Fund was used to get funding for this study from the Khyber Pakhtunkhwa government.
REFERENCES


