Association of Mitochondrial HVS-I Region Variants with Type 2 Diabetes in Pakistani Diabetic Subjects

Syed Ali Shah Bukhari1, Muhammad Shahid2, Saad Farooq1, Muhammad Irfan1, Muhammad Shakeel1 and Ishtiaq Ahmad Khan3

1Jamil-ur-Rahman Centre for Genome Research ICCBS, University of Karachi, Karachi, Pakistan
2Department of Medicine, National Institute of Diabetes and Endocrinology, Dow University of Health Sciences, Karachi, Pakistan

ABSTRACT
Objective: To analyse mitochondrial hypervariable segment I (HVS-I) variations in Pakistani type 2 diabetic subjects.
Study Design: Case-control study.
Place and Duration of the Study: National Institute of Diabetes and Endocrinology, Dow University of Health Sciences, Karachi, Pakistan, between January 2019 to January 2021.
Methodology: DNA from whole blood was isolated, and mitochondrial HVS-I region (16024-16370) of 92 individuals, including 47 controls and 45 diabetics, was amplified, sequenced, and analysed.
Results: Ninety-two variable sites in the sequenced region were identified and individuals were classified into 56 different haplotypes according to phylotree 17.0 classifications, where major haplotype M5 was nearly 2-fold higher in diabetes. Fischer’s exact test revealed variant 16189T>C significantly associated with diabetes (Odds ratio = 12.9, 95% CI = 0.6917 – 2400248) as compared to controls. The authors further analysed 1000 Genomes Project data of Pakistani Control subjects (i.e. PJL, n=96) and found that besides 16189T>C (Odds ratio = 5.875, 95% CI = 1.093 – 31.57, p<0.0339), 16264C>T (Odds ratio = 16, 95% CI = 0.8026 – 314.7, p<0.0310) also showed significant association with diabetic subjects. Comparing diabetic subject data with global control population data of the 1000 Genomes Project, significant associations of eight variants in the studied region were found.
Conclusion: Based on the results of this case-control study, it can be concluded that specific variations in the mitochondrial hypervariable segment I (HVS-I) region are significantly associated with type 2 diabetes in the Pakistani population. The major haplotype M5 was found to be higher in diabetic subjects and variants 16189T>C and 16264C>T were significantly associated with diabetes. These findings suggest that mitochondrial DNA variations may play a role in the development of type 2 diabetes in the Pakistani population.

Key Words: Diabetes Mellitus, HVS-I region, Diabetic subjects, Mitochondrial genomics, Pakistani population.

How to cite this article: Bukhari SAS, Shahid M, Farooq S, Irfan M, Shakeel M, Khan IA. Association of Mitochondrial HVS-I Region Variants with Type 2 Diabetes in Pakistani Diabetic Subjects. J Coll Physicians Surg Pak 2023; 33(07):980-985.

INTRODUCTION

Diabetes mellitus (DM) is a multifactorial disorder affecting millions of people worldwide. Asia, the epicentre of diabetes, contains more than 60% of the global diabetic population. The prevalence of DM is rising exponentially in Pakistan. According to the National Diabetes Survey of Pakistan (NSDP), the prevalence of DM in Pakistan has surpassed all the previous estimations and currently, it is about 26%, which means 1 in every four persons has diabetes. Moreover, the cost of treatment of DM is extremely high in Pakistan posing a huge economic burden. Not only urban areas are affected but its prevalence is also increasing in rural areas. This alarming situation merits taking necessary steps at research and clinical grounds to reduce the burden of this rapidly emerging health disaster. Even after post-genome era, genome-wide association studies are largely lacking in Pakistan and the genomic data from the native Pakistani population is still in infancy. In recent years, the search for genetic determinants of DM has changed dramatically and they are known to have key role in the prognosis, diagnosis, and understanding pathogenesis of different diseases. Due to their vital role in aerobic metabolism, mitochondria are inherently intriguing when considering the pathophysiology of DM.

Mitochondrial dysfunction is directly linked to DM and its complication. Moreover, oxidative processes in mitochondria may generate extensive reactive oxygen species (ROS) that affect the mitochondrial oxidative environment leading to more chances of oxidation-based DNA damage. It is estimated that each mitochondrion has more than 10 copies of mitochondrial (mt) DNA. Mitochondrial DNA mutation rate is about 10-fold higher than nuclear DNA, because of increased exposure to ROS. Furthermore, mitochondria lack protective histones and an efficient DNA repair system, which results in limited defence
mechanisms against endogenous or exogenous damaging agents such as oxidative stress and leads to a high mutation rate. Different mitochondrial mutations are associated with DM and its complications in various populations. Recently, mitochondrial haplotypes are characterised and classified into various sub-classifications to assist in the medicine responses. Most of the variations in the mitochondrial genome are present in the region termed as hypervariable segment (HVS). Different populations investigate the genetic differences in this region to pinpoint diabetic haplotypes that are susceptible or resistant to the disease. Though data is available from different parts of the world related to the involvement of mitochondrial mutations in DM. Limited studies are done so far in Pakistan to identify the mutations in a mitochondrial genome and to the best of the authors' knowledge no study presents such mutations in diabetic patients. The present study aimed to identify these mutations to improve the understanding of the pathogenesis, diagnosis, and prognosis of DM and its complications in the Pakistani population. It will help to understand the pathogenesis of this disease at the molecular level in the Pakistani population. Additionally, it provides a ground for future studies on a large scale to validate and further establish these findings.

**METHODOLOGY**

This case-control study was conducted after ethical approval (IRB-1135/DUHS/Approval/2018). Blood samples from 45 diabetic patients were recruited from the National Institute of Diabetes and Endocrinology, Dow University of Health Sciences and blood samples from volunteer healthy controls after taking their informed consent, from January 2019 to January 2021. Samples were collected from 45 subjects (Db) and 47 normal controls (NC) after informed consent. Samples were collected in EDTA anticoagulated tubes and used for DNA isolation and sequencing.

Subjects of either gender aged 17 to 77 years were included in the study. For controls, subjects without a history of any disease including genetic disorders were included in the study.

Cases were diabetic subjects, person with known history of diabetes in accordance with T2DM WHO 1999 diagnostic criteria were included in the study. The HbA1c cut point for disease diagnosis was set as >6.5% at the time of diagnosis. Pregnant females, patients with psychotic disorders, and subjects having a history of any disease including genetic disorders were excluded from the study.

DNA was isolated using phenol-chloroform isoamy alcohol based extraction method. Quantity and purity of isolated DNA were estimated by NanoDrop 2000 spectrophotometer (ThermoFischer, MA, USA). Isolated DNA was amplified using GoTaq PCR master mix (cat # M7122 (Promega, WI, USA)) following the manufacturer’s instructions. PCR primers designed by Azzawi et al. were used in the study (Forward: 5’ CAGTCTTGTAAACCGGAGATG 3’; Reverse: TGATTTCACGGAGATGTTG) to amplify the mitochondrial HVS-1 region which yielded 508bp product. PCR products were purified using the AMPure XP magnetic beads as per the manufacturer’s instructions (cat#A63800, Beckman Coulter, CA). Approximately 10-20 ng of purified DNA was used for Sanger sequencing. Briefly, purified PCR products were mixed with BigDye™ Terminator Sequencing reaction buffer and BigDye™ Terminator v3.1 Ready Reaction Mix (cat# 4337455), and forward primers were used for the sequencing. The samples were sequenced on the genetic analyzer system ABI 3500. Sequence data were evaluated for quality and 346bp sequence ranging from 16024 to 16370 was used for alignment with fully annotated mitochondrial genome sequence available at NCBI (NC_012920). Data were further processed for mutations and variation analysis using MEGA-X software. Publicly available data of different populations of mitochondrial DNA was retrieved from 1000 Genomes Project data. These populations include SAS, EAS, EUR, AFR, and AMR. The frequencies of variants presented in the diabetic populations were compared with this data.

All statistical analyses were carried out by using Graph Pad Prism v 5.0. Fisher’s exact test was performed to study the associations. A p-value of less than 0.05 was considered statistically significant. Odds ratio (OR) was given to determine the odds of having certain mitochondrial mutations in diabetic subjects than in controls. The confidence interval (CI) was calculated to address if this finding was significant so upper 95% CI is chosen for this study.

**RESULTS**

DNA from the whole blood was isolated and its quality and purity was checked using NanoDrop. About 100 ng of DNA was then used to amplify the HVS-1 region fragment which yielded 508bp of amplified product. The amplified product was then subjected to Sanger sequencing and HVS-1 from 16024 to 16370 was used to study variation. Ninety-two subjects including 47 controls comprising of 33 males (70%) and 14 females (30%), and 45 diabetes including 26 males (58%) and 19 females (42%) were included in the study.

In 346 bases analysed, the authors identified 92 variable sites. Among 92 variable sites, 53 variations with more than a 2-fold difference in DM were found in comparison to the control (Figure 1). Moreover, 32 variations were exclusively presented in DM (Figure 2), and 25 variable sites were presented only in control. However, 35 variations were shared between controls and diabetic subjects. Among these studied variations, 16189>T>C was found to be significantly associated with diabetic subjects as compared to control, OR = 12.9, 95% CI (0.7 to 240.6; p=0.0248).

Comparison with PJL data (n = 96) from 1000 Genomes project data showed that only 2 samples have 16189 T>C variation compared to 5 out of 45 diabetic subjects. Fischer’s Exact test showed significant association OR = 5.8, 95% CI (1.1 to 31.6; p=0.034). Moreover, this significance was only limited to the PJL population of SAS and all the other populations of SAS (i.e. BEB, GIH, ITU and STU) did not show any association.
Table I: Diabetic mtDNA variants associated with South Asian (SAS) Control population: Comparison of Diabetes data and 1000 genome SAS data showing significant association of mitochondrial variation 16189T>C and 16264C>T. (Associations study was carried out using fisher’s exact test and P<0.05 was considered statistically significant; n.s.: not significant, Con: Controls, Db: Diabetics, AFR: Africans, AMR: Americans, EUR: Europeans, SAS: South Asian). The square bracket shows frequency and percentage of a mutation by number of total mitochondrial genomes from that particular region.

<table>
<thead>
<tr>
<th>Mitochondrial DNA Variation</th>
<th>Db Frequency (n=45)</th>
<th>Db vs. Con (n=47) OR(95% CI; p-value)</th>
<th>Db vs. PJL(n=96) OR(95% CI; p-value)</th>
<th>Db vs. BEB (n=86) OR(95% CI; p-value)</th>
<th>Db vs. GIH (n=103) OR(95% CI; p-value)</th>
<th>Db vs. EUR (n=503) OR(95% CI; p-value)</th>
<th>Db vs. EAS(n=504) OR(95% CI; p-value)</th>
<th>Db vs. SAS (n=489) OR(95% CI; p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16048G&gt;A</td>
<td>2 (4.4%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=1; 1%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=1; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
<tr>
<td>16163A&gt;G</td>
<td>2 (4.4%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
<tr>
<td>16189T&gt;C</td>
<td>5 (11%)</td>
<td>n=12.9 (0.7 to 240.6; p=0.025)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=7; 8.1%] n.s.</td>
<td>[n=7; 7%] n.s.</td>
<td>[n=7; 6.9%] n.s.</td>
<td>[n=7; 7%] n.s.</td>
<td>[n=7; 7%] n.s.</td>
</tr>
<tr>
<td>16207A&gt;G</td>
<td>2 (4.4%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
<tr>
<td>16214C&gt;T</td>
<td>2 (4.4%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
<tr>
<td>16218C&gt;T</td>
<td>2 (4.4%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
<tr>
<td>16243T&gt;C</td>
<td>2 (4.4%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
<tr>
<td>16246C&gt;T</td>
<td>3 (6.6%)</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
<td>[n=0; 0%] n.s.</td>
</tr>
</tbody>
</table>

Further analysis showed that SNP 16264 C>T was also significantly associated with DM [OR = 15.2, 95% CI (0.8 to 301.8; p=0.034)] with no case in PJL population while 3 cases in diabetic subjects (Table I).

To have the broader picture, 8 variations were selected i.e. 16048G>A, 16163A>G, 16189T>C, 16207A>G, 16214C>T, 16218C>T, 16243T>C, 16264C>T exclusively presented in diabetic subjects with frequency 2 or more and compared it with different global populations i.e. East Asian (EAS), South Asian (SAS), African (AFR), American (AMR) and European populations (EUR). Except 16189T>C, each variation showed significant associations with at least 3 of 5 global population of 1000 genome project. The statistical details and variations with respect to each population are presented in Table II. It is interesting to note that population SAS of 1000 Genomes Project in which Pakistan is grouped showed significant association to all these variations except 16189T>C.

Haplogroups of the studied subject were classified according to the phylo tree classification 17.0 using online resource of mitomap. All the subjects were grouped into 56 unique minor haplotypes. The most common haplotypes
were H, U, M5, R and W which accounted for more than 45% of the population. Subjects including haplogroups with a frequency less than 5% were merged with their respective major haplogroup (Figure 3). The haplogroup M5 and R were present more than 2-folds (16% vs. 7% and 9% vs. 4%) in DM than in control.

DISCUSSION

In Pakistan, there has been lack of evidence of mitochondrial genetic variations associated with diabetes. Here, the authors sequenced and found the variations present in the mitochondrial HVS-1 region of Pakistani diabetic subjects (Figure 1). One of the major factors for the development of DM and its complication are genetic variations. Some mutations are reported to have strong associations with diabetes and increased risk of diabetes, however other seems protective. The mutation rate in HVS region of mitochondria is higher than in other parts of mitochondrial DNA. These mutations can interfere with promoters and will affect the binding affinity of promoters modulating the transcription and replication processes. The mitochondrial variation 16189 T>C lies in the control region of mitochondrial replication and may affect the replicative processes of the mitochondria. It is associated with increased risk of DM as evident from the meta-analysis of the Asian studies. However, no correlation is found in European populations. The altered mitochondrial metabolism can be the key deterministic factor of the pathophysiology of diseases that involve oxidative stress like DM. Though subjects were classified into various haplotypes but no significant association related to haplotypes was found between controls and diabetic subjects except for Haplogroup M5 and R which were nearly 2-fold greater in diabetics as compared to the control but did not show any significant association as presented in many other studies. This suggests that in general, no ethnic group in studied population is less or more prone to the DM, but considering less sample number, this conclusion could not be strong enough and further validations with large sample number is necessary. Furthermore, in order to have a true picture of haplotypes, high-resolution haplotypes from full mitochondrial genome sequencing can give a better idea about haplotype associations.

Moreover, it was found that the highly prevalent mutation in diabetics i.e. 16189 T>C was more prevalent in the region of Africa (Table II) followed by different other regional groups suggesting the population-based differences. Interestingly according to the 1000 Genome Project data frequency of 16189T>C in the population in the Pakistani region was the least and EAS population in which Pakistan is placed has the least frequency as compared to all the other global groups (Table II). There have been a number of studies published that identifies the population specific variations responsible for diabetes. Mitochondrial mutation 16189 T>C has been an important discriminant of diabetes in some populations but not in others. The T16189C variant has been reported in cardiomyopathy and DM and has been suggested to affect mitochondrial DNA replication.
Mitochondrial HVS-I variants in Pakistani diabetic subjects

16189T>C, 16207A>G, 16214C>T, 16243T>C, 16264C>T are associated with diabetes, however, study with the increased number of subjects is required to further validate and authenticate the data. These variations in the control region have been involved in the compromised replicative process of mitochondria and in the pathogenesis of various diseases involving oxidative stress like cancers and metabolic disorders.22-24

These results show that the Pakistani diabetic subjects have characteristic variable sites and some of these are distinctly different to any healthy population which can be the key variations to be referred for future studies. Furthermore, the study only reflects the data from a very limited population dataset. Therefore, a more detailed study with strict age and gender-matched study samples is required.

CONCLUSION

The study identifies variations present in HVS-1 region of Pakistani Diabetic subjects. We Mutations 16189 T>C associated with Pakistani diabetic subjects were identified. Moreover, using 1000 Genomes Project data for comparison it was found that mutation 16264C>T and nearly six other variations are also significantly associated with diabetes comparing it with global data. Research with increased number of sample size and whole mitochondrial genome sequencing can be done in future studies.

FUNDING:
The study was funded by PCMD (Panjwani Centre for Molecular Medicine and Drug Research) through an internal grant for the project 1401-2018.

ETHICAL APPROVAL:
This case-control study was conducted after ethical approval from Dow University of Health Sciences (IRB-1135/DUHS/Approval/2018/).

PATIENTS’ CONSENT:
Informed consent was obtained from the patients to publish the data.

COMPETING INTEREST:
The authors declared no competing interests.

AUTHORS’ CONTRIBUTION:
SASB, MS, SF, MI: The acquisition, analysis and interpretation of data for the work drafting and revising the manuscript critically for important intellectual content.
MS, IAK: The acquisition and analysis of data for the work. All the authors have approved the final version of the manuscript to be published.

REFERENCES


