Association of a single-nucleotide polymorphism in C12orf43 region with the risk of coronary artery disease

Najma Qammar1, Maryam Zain1, Raheela Jabeen2, Farah Deeba1, Nadia Iqbal1, Hafiz Muhammad Rashad Javeed2, Fatema Suliman Alatawi1, Mohsen Suliman Alatawi1, Sanaa Almowallad1, Amnna A. Alharbi3, Hüseyin Şahin5

1 Department of Biochemistry and Biotechnology, The Women University, Multan-60000 Punjab, Pakistan
2 COMSATS University Islamabad, Vehari Campus, Vehari, Vehari-61100, Punjab, Pakistan
3 Department of Biochemistry, Faculty of Science, University of Tabuk, Tabuk, Saudi Arabia
4 Department of Pediatrics, College of Medicine, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
5 Tekirdağ Namık Kemal University faculty of medicine, Tekirdağ, Turkey

Abstract

The genetics of organisms play a vital role in the development of coronary artery disease (CAD), with its heritability estimated at approximately 50-60%. For this purpose, we examined the relationship between CAD risk and C12orf43/rs2258287 polymorphisms in the Pakistani population. In this study based on the genetic approach to dyslipidemia, a total of 200 subjects were included from the southern Punjab. The biochemical analysis of parameters (total cholesterol, triglycerides, blood glucose, high-density lipoprotein, and low-density lipoprotein) was carried out along with molecular analysis using an ARMS-PCR-based assay for single-nucleotide polymorphism (SNP) C12orf43/rs2258287 to identify the genotype. Genotypes showed a substantial correlation with both family history and metabolic markers. The cholesterol, low-density lipoprotein cholesterol (LDL-C), triglycerides and blood glucose levels were higher while the high-density lipoprotein cholesterol (HDL-C) level was lower significantly (p<0.05) in cases than in controls. Age, pulse rate, diabetes, physical activity, smoking, family history, and dietary habits were also significantly associated (p<0.05) with CAD individuals. The SNP C12orf43/rs2258287 also showed an association with CAD in the population of southern Punjab. Based upon this study, it could be concluded that CAD is characterized by an unfavorable lipid profile in association with SNP C12orf43/rs2258287.

1. Introduction

Coronary artery disease (CAD) is the predominant manifestation of cardiovascular disease. It is caused by the development of fatty deposits within the blood arteries that provide blood to the heart [1]. CAD encompasses a broad spectrum of conditions, ranging from stable angina and asymptomatic atherosclerosis to acute coronary syndrome. CAD is a very complicated disorder as both environmental and genetic factors have significant impact on it [2]. Although it is inherited partially, a lot of work has been done to understand its genetic basis and other complicated cardiovascular diseases linked to it [3]. Past studies concluded it to be a heritable disorder in humans and have initiated the unraveling of the genetic underpinnings of this disease. The evidences, i.e. twin studies, specific genetic factors, familial studies and genome-wide association studies, indicated that CAD could be related to a genetic basis in addition to environmental factors [4,5]. Clinical reports beyond the 1950s have concluded that the RF for CAD is heritable [6] which is between 40-58% [7]. Furthermore, a more significant predictor was a family history of CAD in the parents or siblings [8,9]. Substantial advancements in the field of genetic underpinning of CAD helped the researcher to understand the genomes of different living things [10]. Genome-wide association studies (GWAS) have sorted the approx. 60 genetic variants that had a strong association with CAD risk. CAD is a quintessentially complicated phenotype resulting from very stochastic, complex, and non-linear interactive effects among genomics, proteomics, genetic and some environmental factors. GWAS has identified over 100 single nucleotide polymorphisms (SNPs) linked with CAD. The SNP C12orf43/rs2258287 has an association with CAD risk and is involved in blood coagulation, lipid metabolism and atherosclerosis as well (Erdmann et al. 2009). It is newly discovered loci for CAD, so very few studies to date are reported. SNP (rs2258287)

* Corresponding author.
E-mail address: najmaqammar@gmail.com (Hüseyin Şahin); aymanelsabagh@gmail.com (N. Qammar).
Doi: http://dx.doi.org/10.14715/cmb/2024.70.2.4

24
is located downstream to HNF1A, within C12orf43 gene (alias FLJ12448, hypothetical protein LOC64897), in tail-to-tail manner, with in-between 500bp only. C12orf43 is also known as custos and has 7 exons. The gene is situated on the complementary strand of DNA and exhibits widespread expression in the kidney (RPKM 7.1), small intestine (RPKM 5.6), duodenum, brain, and 23 more organs. RT-PCR investigations have demonstrated that C12orf43 is expressed widely in the heart, aorta, and associated tissues of both humans and mice, all of which play a role in the development of atherosclerosis [11]. Additionally, the risk allele rs2258287 has been mapped to a higher plasma level of the low-density lipoprotein cholesterol [12]. The main objective of the study was the comparison of anthropometric parameters and biochemical analysis in the case (having CAD) and control groups.

2. Materials and Methods

2.1. Study design

The study included 200 subjects, i.e., 100 cardiac (case) and 100 non-cardiac (control), recruited from Heart Care and Medical Center, Mian Channun, Southern Punjab, Pakistan. Subjects from both the cardiac and non-cardiac categories were selected from the same institution. The majority of the participants were individuals with a history of heart conditions. The subjects included in the study were of all ages and of both genders (male and female) as well. During the study, anthropometric data i.e., age, weight, height, blood pressure (BP), pulse rate, body mass index (BMI), gender, eating habits, smoking and family history and biochemical parameters i.e., high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), cholesterol, triglycerides and glucose (S 16) were recorded at the hospital and some at the Lab (Zeenat Laboratories Mian Channun, Southern Punjab, Pakistan). People who had hepatitis C or B, were seropositive for those viruses, kidney or liver disease, cancer, HIV, or any other ongoing acute infectious disease before biochemical analysis could not be included in the study.. Incursion criteria of the cardiac patients were electrocardiogram (ECG) and past heart disease history which were present in the patient diagnostic file. A consent or ethical form was filled out in and signed before taking the blood samples from the subjects/patients (S 17) and their personal information was promised to be kept secret. The study approval was taken by the institutional ethics committee of The Women University Multan, Southern Punjab, Pakistan, and all the procedures were adopted in compliance with the Helsinki Declaration.

2.2. Blood sample collection and DNA extraction

As part of the study, a 5 ml blood sample was collected from both the control and case subjects who readily volunteered for the study. The blood was extracted using a syringe and then stored in vacutainers containing gel and clot activator, as well as sterile EDTA.. The blood samples were labeled properly and stored at 4 °C. Half-blood of the control and case subjects was used for DNA extraction (v≈ 3 ml), which was stored in sterile EDTA vacutainers at 4°C. Half-blood (v≈ 2 ml) was used for biochemical assays i.e., HDL-C, LDL-C, cholesterol, triglycerides, and glucose. The Red Cell Lysis Buffer (RCLB) Method and the Salt Precipitation Method (REF) were both used to extract genomic DNA. All the samples of genomic DNA were stored at −20°C.

2.3. Biochemical analysis

Various biochemical parameters including cholesterol, HDL-C, LDL-C, glucose, and triglyceride levels were analyzed using serum collected from blood samples stored in gel and clot activator vacutainers. Cholesterol, HDL-C, LDL-C, and triglycerides in serum was measured through spectrophotometric analysis on the biochemical analyzer using reagent kits (Equation no 1, 2, 3 and 4; cholesterol/HDL-C/LDL-C/triglycerides liquicolor [22] by Human, Germany, Ref 10028/10084/10094/10724, respectively). Moreover, the glucose was determined from serum through a photometric system on the biochemical analyzer using reagent kits (Equation no 5; Merck, ref 5.17604.0001). Samples were also calibrated using biochemical photometric system (Merck, Microlab 300).

2.4. Calculations

The following equation was used to determine various biochemical parameters.

\[
C = \frac{\text{Absorbance (sample)}}{\text{Absorbance (standard)}} \times \text{[mg/dl]} \quad \text{Eq. 1, 4} \quad [13]
\]

\[
C_{\text{sample}} = \frac{C_{\text{calibrator}} \times \text{Absorbance (sample)}}{\text{Absorbance (calibration)}} \times \text{[mg/dl]} \quad \text{Eq. 2,3} \quad [13,14]
\]

Glucose Concentration = Standard Concentration \times \Delta\text{Abs} \times \text{mg/dl} \quad [15]

2.5. Genotyping

Expansion Genotypes were ascertained by the Fractory Mutation System-Polymerase Chain Reaction (ARMS-PCR). Three primers, as two forward primers (F2A 5’-CGTCATAGGAGGCTTTGATAACG-3’, F2B 5’-CGTCATAGGAGGCTTTGGATAACT-3’) and one common reverse primer (R 5’-ACTGCTTTGGCAACAAAGAGCT-3’) were used for amplification of SNP C12orf43/rs2258287 through ARMS-PCR. The prepared ARMS-PCR reaction mixture contained 5 µl/2X of the Fermentas #1081 PCR master mix, 1 µl DNA, 3.4 µl ampule water, 0.6 µl/3% DMSO (Dimethylesulfo-oxide), 1.5 µl/µM reverse primer (R1), and 1.5 µl/µM forward primer (F1A) for a total 10 µl volume. These primers allowed the detection of 578 bp fragments for genotype G/T of SNP C12orf43/rs2258287. Reaction conditions for PCR of the reaction mixture were as follows: 2 minutes at 95°C, 35 cycles for 1 minute at 95°C, 1 minute at 61.2°C, 1 minute at 72°C, and a final extension for 5 minutes at 72°C. Reaction products were then analyzed on 1% agarose gel by using stained ethidium bromide and visualized through a UV transilluminator.

2.6. Statistical analysis

All the data were analyzed using the software Statistical Package for the Social Sciences (SPSS, IBM Statistics, version 25, USA). Clinical characterization of subjects was given as mean ± SD. The P-value was also calculated to check significant differences among characteristics of controls and cases by an Independent-Sample T-test. Using statistical software and an online calculator, allelic and genotyping frequencies were calculated and then checked for Hardy-Weinberg Equilibrium implementation. The Odds Ratio (OR) calculator (www.medcalc.org) was utilized to examine the correlation between polymorphism and trait. The chi-square test was used for the qualitative compari-
son of the phenotypic characteristics e.g., to analyze the association of BMI, diabetes, family history of CAD and diabetes, smoking, and dietary habits. The p-value < 0.05 was considered significant.

3. Results

Table 1 illustrates the characteristics of 200 samples, of which 100 were cases and 100 were controls. The subjects in this study were both male and female. Both female and male subjects were included in the current study. In some cases, the male was 55% and the female was 45%, while in control, the male was 45% and the female was 55%. A significant difference was observed in the mean age (years) of cases (54.60±12.38), and control (43.67±13.03), as the mean age of the case subject was higher than that of control subjects. Male individuals had a mean age of 50±13.942 and female subjects had a mean age of 48.27±13.708. The mean BMI (kg/m²), weight (kg), and height (cm) of the controls and patients did not differ significantly.

The CAD subjects’ proportion with diabetes was significantly higher in cases i.e., 48 (24%) than in controls 14 (7%). Similarly, there was also a marked difference in the mean blood glucose level in cases (143.74±62.38) and controls (110.24±55.86). The normal range of blood glucose in the fasting is 72~108 mg/dl. Compared to the control group (185.91±24.49), the mean cholesterol level was higher in the case group (215.01±28.39). The normal range of cholesterol level is about <200 mg/dl. The level of triglycerides between 200~239 mg/dl is taken as borderline but if it exceeds this limit then subjects are prone to CVD.

A marked difference (p<0.05) in HDL-C was observed, as the mean HDL-C level was lower in cases (36.18±5.51) while higher in controls (51.74±16.65). The normal range of HDL-C was 35~55 mg/dl (Table 1; S1-4).

In both groups, there was a substantial difference in the LDL-C level. 149.40±47.78 was the mean LDL-C level in cases compared to 105.36±30.99 in controls. The ideal range for LDL-C levels was less than 100 or between 100 and 129 mg/dl; 130 to 159 mg/dl is regarded as noticeably high; if it is more than 160 mg/dl, individuals are at risk for coronary artery disease (CAD). The study groups exhibited a significant difference in triglyceride levels, with the mean triglyceride level in cases (215.55±39.62) being greater than that in controls (139.38±39.32). The ideal triglyceride level is within the range of less than 150 milligrams per deciliter (mg/dl). A level exceeding 150 mg/dl is regarded as a risk factor for atherosclerosis, however if it rises to 200 mg/dl, the risk of coronary artery disease (CAD) increases significantly. The prevalence of smoking among individuals with coronary artery disease (CAD) was significantly higher (20.5%) compared to the control group (5%) (Table 1). The success rate of genotyping of 200 DNA samples for SNP C12orf43/rs225828 was 58%-67%. The genotype distribution of SNP C12orf43/rs2258287 in controls and cases is shown in Table 2.

The terms SNP (single-nucleotide polymorphism), χ² (Chi-square value), OR odds ratio (recessive genetic model), p significance level, and 95% CI stands for 95% confidence interval. It was found that in cases, the genotype frequencies of the C12orf43/rs2258287 polymorphism G/T were 4.48% (GG), 37.31% (GT), and 58.20% (TT). In controls, they were 18.37% (GG), 55.10% (GT), and 26.53% (TT). When C12orf43/rs2258287 homozygote genotype (GG) as a reference group was used, there was

Table 1. Comparison of biochemical and anthropometric parameters in case and control groups.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases</th>
<th>Controls</th>
<th>p-value</th>
<th>S/NS (p&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample no (n)</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n)</td>
<td>55</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (n)</td>
<td>45</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age(years)</td>
<td>54.60±12.38</td>
<td>43.67±13.03</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>73.35±8.24</td>
<td>70.65±14.94</td>
<td>0.253</td>
<td>NS</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>162.39±10.67</td>
<td>161.16±11.80</td>
<td>0.438</td>
<td>NS</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.94±7.42</td>
<td>27.35±6.00</td>
<td>0.538</td>
<td>NS</td>
</tr>
<tr>
<td>Pulse rate (beats/min)</td>
<td>83.96±13.10</td>
<td>79.64±13.25</td>
<td>0.021</td>
<td>S</td>
</tr>
<tr>
<td>Diabetic subjects</td>
<td>48(24%)</td>
<td>14(7%)</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>Glucose (mg/dl)</td>
<td>143.74±62.38</td>
<td>110.24±55.86</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>Cholesterol (mg/dl)</td>
<td>215.01±28.39</td>
<td>185.91±24.49</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>HDL-C (mg/dl)</td>
<td>36.18±5.51</td>
<td>51.74±16.65</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>LDL-C (mg/dl)</td>
<td>149.40±47.78</td>
<td>105.36±30.99</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>Triglycerides (mg/dl)</td>
<td>215.55±39.62</td>
<td>139.38±39.32</td>
<td>0.000</td>
<td>S</td>
</tr>
<tr>
<td>Smokers (n)</td>
<td>41</td>
<td>10</td>
<td>0.000</td>
<td>S</td>
</tr>
</tbody>
</table>

Abbreviation: S; significant, NS; non-significant.

Table 2. SNP loci genotypic and allelic frequency distribution and association with CAD.

<table>
<thead>
<tr>
<th>SNP C12orf43/rs2258287</th>
<th>GG</th>
<th>GT</th>
<th>TT</th>
<th>χ²</th>
<th>OR</th>
<th>p</th>
<th>95%CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>3</td>
<td>25</td>
<td>39</td>
<td>0.23</td>
<td>0.77</td>
<td>0.687</td>
<td>0.162</td>
</tr>
<tr>
<td></td>
<td>(4.48%)</td>
<td>(37.31%)</td>
<td>(58.20%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>9</td>
<td>27</td>
<td>13</td>
<td>0.46</td>
<td>0.54</td>
<td>0.444</td>
<td>0.587</td>
</tr>
<tr>
<td></td>
<td>(18.37%)</td>
<td>(55.10%)</td>
<td>(26.53%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR 0.0009 1.74-8.57
an association of GT genotype with increased coronary artery disease risk (GT vs. GG, OR 2.78, CI 0.67–11.44, p = 0.16). The TT genotype also showed an association with the increased risk of CAD (TT vs. GG, OR 9.00, CI 2.11-38.35, p=0.003). When C12orf43/rs225828 combined GG/GT genotypes as a reference group were used in the recessive genetic model, the TT genotype showed an association with increased CAD risk (TT vs GG/GT, OR 3.86, CI 1.74-8.57 p=0.0009). The C12orf43/rs225828 polymorphism showed no association with CAD risk in the dominant genetic model (GG vs. GT/TT, OR 0.21, CI 0.05-0.82, p=0.02).

4. Discussion
Both sexes were involved in CAD cases in the current investigation. Worldwide, CAD is a leading cause of death for both men and women. Women often experience CAD roughly ten years after men [16]. But in both men and women, the incidence of CAD rises beyond the age of 35. Men have a 49% lifetime chance of developing CAD after the age of 40, compared to a 32% risk for women [17]. It has been found that some CAD RFs affect the disease onset separately in women and men. Higher level of RFs e.g., smoking, diabetes, physical activity, and hypertension at an earlier age in men enhances the chances of CAD[16]. Cardiovascular disorders account for a significant portion of diabetes-related mortality and morbidity, and they also cost the country $37.3 billion annually in cardiovascular-related expenses [18]. The rate of heart disease is 2.5 times greater in males and 2.4 times greater in females in diabetic patients as compared to those without diabetes [19]. The current study also presents information regarding causative role of RFs in CAD. Common conditions that coexist with type-2 diabetes such as dyslipidemia and hypertension are clear CAD risk factors and diabetes confers the independent risk [20]. In CAD patients, lipid abnormalities are mainly characterized by increased levels of serum triglycerides, LDL-C, and decreased levels of HDL-C [21]. The increased concentration of triglycerides results in increased VLDL-C production and also diminished triglycerides clearance [22]. The results of the present study indicated clearly dyslipidemia in the patients of CAD (cases) as compared to controls. The mean of total cholesterol, LDL-C, and triglycerides were significantly higher while the mean of HDL-C level was remarkably lower in cases than controls. This abnormal lipid profile in patients with CAD confirms the severity of the disease. The protective impact of HDL is unaffected by other RFs or lipids and is just as potent as that of LDL-C. There is a 50% difference in risk for every 10 mg/dl variation in HDL-C levels [18]. Moreover, it is observed that the risk related to cholesterol is significantly influenced by the coexistent RF level, although other RFs of any level showed an independent association of CAD with the value of serum total cholesterol. Thus, any cholesterol level, low or high, has a different significance that depends on the presence of other RFs [23]. Hypertension is also strongly related to CAD, as a significant association was observed in cases and controls in the present study. Hypertension is a major independent RF for CAD development, renal failure, and stroke [24].

Another significant RF of CAD is family history [25]. It is observed that a family history of CAD plays a major role in the onset of CAD. Patients with the age <50 years and having a family history of premature CAD showed increased CAD mortality risk [26]. Similarly, there was a strong correlation found between smoking and the risk of CAD. When compared to not smoking during one’s lifetime, the risk of death and morbidity from CAD is approximately doubled by cigarette smoking, and this risk increases with the quantity and duration of smoking [27]. In the present study, physical activity showed a greater effect among cases than control as individuals with physical activity were at lower risk of CAD. According to studies, it is observed that individuals with physical activity experienced less or half the incidence of CHD or sudden death than sedentary individuals [28]. For the general population (children, the elderly, and adults), recommendations for physical exercise have become an important element of preventive policies [29]. One of the key risk factors for the emergence of CAD in the general population is poor eating habits [30]. Cardiovascular patients are most likely to smoke and eat poorly [31], and may be clustered with other unfavorable behaviors like low activity and excessive drinking levels, increasing the cumulative effect of the multiple risks [32].

In this study, we also looked at the link between SNP C12orf43/rs2258287 G/T and CAD risk in the Pakistani population. We found that SNP C12orf43/rs2258287 G/T may raise the risk of CAD in the Pakistani population. C12orf43/rs2258287 in 11q22.3 locus is a member of the SNPs group, which is found to have a strong association with CAD in various researches [33]. The intron 7 of the Hepatocyte nuclear factor-1a (HNF1A) is included in the region of the chromosome possessing this SNP. HNF1A encodes the transcription factor that is only present in the liver. Thus, HNF1A variants can possibly lead to maturity, the onset of diabetes in the young, influencing the C-reactive protein’s plasma concentration, an influential threat pointer in the support of CAD. Besides, the peril allele of the HNF1A locus side (rs2258287) has been reported to increase the LDL-C level of plasma [11]. A remarkable association was observed between SNP C12orf43/rs2258287 with CAD in the present study sample. T allele frequency was higher in our population thus studied. In cases (CAD subjects), the frequency of the T allele was 77%. In controls (healthy subjects), T allele frequency was also higher (54%) compared with G allele (46%). Given this increased frequency, it seemed that RF had a more significant role in the population’s development of coronary artery disease.

5. Conclusion
The positive association between variant and CAD risk could be detected successfully in the current study, which can be beneficial to look in more detail into the genetic causes of CAD in a unique ethnic group in Pakistan. Based on this study, it could be concluded that CAD is characterized by an unfavorable lipid profile in association with SNP C12orf43/rs2258287.

Acknowledgements
I would like to express my deep gratitude to Dr. Ayman El Sabagh, for their guidance and useful critiques of this research work.

Competing interest
The authors declared that they have no competing conflict
of interest.

Consent for publications
The author read and approved the final manuscript for publication.

Ethics approval and consent to participate
No human or animals were used in the present research.

Informed consent
The authors declare not used any patients in this research.

Availability of data and material
The data that support the findings of this study are available from the corresponding author upon reasonable request

Authors' contributions
Najma Qammar, Maryam Zain, Raheela Jabeen, Farah Dceba, Nadia Iqbal, Hafiz Muhammad Rashad Javeed, designed, conducted the research, analysis the data as well as prepared the manuscript. Fatema Suliman Alatawi, Mohsen Suliman Alatawi, Sanaa Almowallad, Amnna A. Alharbi, Hüseyin Şahin, also contributed during writing the manuscript and advised scientific suggestion as well as revised/edited the manuscript.

Funding
None.

References