

Identification of PCSK9 Gene Mutation in Bangladeshi Population for Prognosis of Cardiovascular Disease

Md. Mahbubur Rahman

Department of Biochemistry, Rajshahi Medical College, Rajshahi-6100,
Bangladesh and Institute of Biological Sciences, University of Rajshahi,
Bangladesh

Md. Tanjil Islam Shovon

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Md. Mahmudul Hasan Maruf

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Ratna Khatun

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Md. Jahangir Alam

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Md. Ragib Shariar

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Md. Zahidus Sayeed

Department of Cardiology,
Rajshahi Medical College Hospital, Rajshahi-6100, Bangladesh

Zennat Ferdousi

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Kazi Md. Faisal Hoque

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Khandaker Md. Khalid-Bin-Ferdaus

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

Md. Abu Reza

Molecular Biology and Protein Science Laboratory,
Department of Genetic Engineering & Biotechnology,
University of Rajshahi, Rajshahi-6205, Bangladesh

ABSTRACT

Cardiovascular Diseases (CVDs) is one of the most common causes of mortality and morbidity in world as well as in Bangladesh. Dyslipidemia or abnormal lipid level leads to atherosclerosis which is one of the major risk factors for developing CVD. The ultimate consequences of CVD include myocardial infarction, heart failure and stroke. Low-density lipoprotein cholesterol levels in the blood are significantly correlated with atherosclerosis progression or continuously building up plaques in inner wall of cardiac artery. The abnormality of HDL and LDL level is greatly influenced by genetic factors. PCSK9 is an endogenous inhibitor of the low-density lipoprotein receptor and plays a significant role in the metabolism of cholesterol. The current study aims to identify the PCSK9 gene mutations as well as to carry out a gross assessment of secondary causes to reveal the actual magnitude of risk for future generation of Bangladeshi population. Blood samples were collected from 100 subjects for lipid profiling and DNA was extracted from the blood of the patients with abnormal lipid profile. Socio-demographic analyses were carried out to assess the magnitude of the risk factors associated with CVD. For mutation analysis of the PCSK9 gene, genomic DNA was extracted from the blood. Exon 1 and 2 were selected for segmental amplification and subsequent genetic analysis. Mutations in the amplified sequences were identified by multiple sequence alignment. Bioinformatics analysis revealed seven-point mutations in exon-1- and three-point mutations in exon-2 of PCSK9 gene of some of the CVD patients. This study will be helpful to evaluate the possibility of using PCSK9 as genetic markers for prognosis of dyslipidemia and CVD.

Keywords: CVD, Dyslipidemia, Lipid profiling, PCSK9 gene, Mutations.

INTRODUCTION

Cardiovascular Diseases (CVDs) are considered as an exceedingly influential cause of mortality and morbidity in Bangladesh [1]. CVD is most often manifested either in the form of stroke or coronary artery disease (CAD) such as angina and myocardial infarction. These two are the major cause of death in Bangladesh contributing 16.27% and 14.31% of total death respectively [2]. The main pathological feature of CVD is atherosclerosis which is secondary to accumulation of Low Density Lipoprotein (LDL) in the wall of artery resulting plaque formation

along with narrowing and hardening of the artery that can reduce or completely stop blood circulation [3]. Low level of High Density Lipoprotein (HDL) which scavenges and removes LDL from the body is strongly associated with atherosclerosis which also caused by elevated level of LDL [4], [5]. The excess burden of CAD among South Asians appears to be primarily due to dyslipidemia that is characterized by high levels of LDL, triglycerides (TG), and total cholesterol and low levels of HDL and apoA1.

The abnormality of HDL and LDL level is greatly influenced by genetic factors ranges from 40–60% [6]. Though secondary causes such as diet, gender, age, obesity, smoking, alcohol, exercise, drugs or metabolic disorders have been also reported in numerous studies. About 20 genes have been identified as monogenic causes of Dyslipidemia. However, mutation of the gene namely *PCSK9* has been identified as one of the most responsible factors for autosomal-dominant hypercholesterolemia as well as Tangier disease[7], [8]. The liver secretes PCSK9, which functions as an endogenous inhibitor of the low-density lipoprotein receptor and plays a significant role in the metabolism of cholesterol [9]. With the progress in molecular biology, polymorphisms and genetic changes in the genes potentially associated with plasma HDL and LDL metabolism is being investigated all over the world. Many countries have conducted population-based study to find out genetic risk factors and to overcome CVD related complications and thus they have complete database for the CVD related genetic mutations for their population, which they utilize for prognosis and risk factor calculation of individuals [10]–[12]. However, although a good number of population and clinical research have been conducted on CVD patients in Bangladesh, but no study has been carried out on the genetic factors which are also associated with CVD. Here, we report the first genetic study showing mutation in PCSK9 gene in the patients with dyslipidemia leading to CVD. Therefore, the current study aims to identify the genetic factors as well as to carry out a gross assessment of secondary causes to reveal the actual magnitude of risk of CVD for future generation of Bangladeshi population.

MATERIALS AND METHODS

Sample and Data Collection

A total 100 subjects were selected including 75 CVD patients and 25 healthy controls with no prior family history of CVDs for this study. A questioner was prepared and used for collecting socio-demographic data from the participants. Each participant signed a consent form to take part in the study. Blood samples were drawn from participants. Approximately 2.5 ml blood was collected from each individual for Genomic DNA extraction and serum isolation. The experiments were carried following the protocol approved by the Institutional Ethics Committee for Experimentations on animal, human, microbes and living natural sources, Institute of Biological Sciences, University of Rajshahi, Bangladesh (Memo No.: 110/320/IAMEBBC/IBSC).

Lipid Profiling

The blood samples for serum preparation were kept in small glass tube and incubated for half an hour at 25°C for clotting. Subsequently, they were centrifuged at 4,000 rpm for 12 minutes. The serum was transferred into fresh tubes. Lipid profiling was conducted using the kits of Human (Wiesbaden, Hessen, Germany) in Microlab-300 clinical chemistry analyzer (ELITech Group, Puteaux, France).

Genomic DNA Extraction and Polymerase Chain Reaction (PCR)

Patient and control DNA was isolated from the blood by following manufacturer's guideline of the TIANamp blood DNA kit, China (Tiangen Biotech Company Ltd., Beijing, China). DNA was visualized by run on 1% agarose gel electrophoresis. PCSK9 gene reference sequence was retrieved from the NCBI database (Accession number: NG_009061.1) and two pairs of primers, one pair each for exon 1 and exon 2 were designed (table 1). GoTaq® G2 Hot Start Green Master Mix (Promega, USA) was used for PCR amplification. GoTaq® G2 Hot Start Green Master Mix possesses 5' → 3' exonuclease activity and hence reduces the chance of PCR errors or misincorporation of wrong bases. After PCR amplification, PCR product was cleaned up using Wizard® SV Gel and PCR Clean-Up System kit (Promega, USA) according to the manufacturer's protocol.

Table 1. Primer properties of PCSK9 gene

Primer	Sequence	Base pairs	Tm	GC content
PCSK9 Ex1 Fwd	5'-GATGGGCTGGTGGCGTGATCTGCG-3'	27	68.2°C	68.2
PCSK9 Ex1 Rev	5'-GCGGTGGAAGGTGGCTGTGGTTC-3'	20	64.8°C	65.2
PCSK9 Ex2 Fwd	5'-CTCTTTTGTAAATTTGTAAGTAGGGGTGAG-3'	32	57.4°C	57.4
PCSK9 Ex2 Rev	5'-GCTCCAGCAGGTCGCCACTCATCTTC-3'	26	64.9°C	61.5

DNA Sequencing

Sanger dideoxy DNA sequencing method was used for DNA sequencing. In the current study, the cycle sequencing PCR was conducted using a BigDye Chain Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, USA) according to manufacturer's instructions. Purification of the product was performed using a BigDye X Terminator® Purification Kit. Finally, sequencing was performed on the ABI PRISM 3100 automated DNA sequencer (Applied Biosystems, USA).

Mutation Screening

Identity of the obtained sequence was first verified using NCBI Nucleotide BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), BLASTN. Multiple sequence alignment was performed using DNAMAN software version 4.1 (Lynnon Biosoft, CA, USA) to identify the mutation present in the obtained sequences, comparing with the NCBI reference sequence. Mutations were identified and translated into corresponding amino acid sequence by online Swiss Bioinformatics Resources Portal-ExPasy translation tool (<https://web.expasy.org/translate/>). The change of specific amino acid was recorded.

Bioinformatics Analysis

Prediction of Damaging or Disease-Causing Mutations:

In this study various bioinformatics tools were utilized to determine the disease-causing mutations among the detected mutations in exon 1 and 2 of PCSK9 gene such as, PolyPhen-2: Polymorphism Phenotyping v2, PANTHER: Protein Analysis Through Evolutionary Relationships, PhD-SNP: Predictor of human Deleterious Single Nucleotide Polymorphisms and SNPs&GO. Polyphen-2 is a widely used bioinformatics tools to predict pathogenic mutations. Each substitution's potential impact on a protein's structural and functional characteristics is predicted by Polyphen-2. The examination of structural and evolutionary conservation yields

this prediction. Three categories of predictions are typically made by polyphen-2: "Probably Damaging," "Possibly Damaging," and "Benign."

The score range for Polyphen-2 ranges from 0 to 1 [13]. Whether a specific substitution is pathogenic or not is predicted by PANTHER. It is measured by the evolutionary conservation process. Here, the protein's functional outcome is determined by how long the protein has been conserved [14]. On the basis of the reliability index (RI) score, PhD-SNP forecasts results. 'Neutral' and 'disease' results are the two possible outcomes [15]. SNPs&GO uses a system called a support vector machine (SVM) to predict results and determine if a certain protein substitution site is linked to a disease or not. With a probability score of >0.5 and an accuracy of close to 82%, it is likely that the underlying protein structure has been altered [16].

Prediction of Protein Stability Change:

To determine the changes in protein stability, MUpro computational tool was utilized. The MUpro (<http://mupro.proteomics.ics.uci.edu/>) server is commonly used to assess the stability of protein. This web server was built using Support Vector Machines (SVM) and neural networks, two machine learning techniques. With these techniques, the effect of single-site amino acid changes on protein stability is estimated, and the results are reported as an increase or decrease based on a positive or negative score [17].

Comparative 3D Protein Modeling:

We used the Robetta server (<http://robetta.bakerlab.org/>) to build the 3D structure of the native and mutated structure of exon 1 and 2 of *PCSK9* gene. When the known structure is accessible, the Robetta server employs a comparison modeling strategy. Robetta, on the other hand, employs the de novo strategy in the absence of the established structures [18]. For validation, Robetta generated model 1 for every mutated protein was selected. Validation was carried out by using the SAVES v6.0 server (<https://saves.mbi.ucla.edu/>). SAVES v6.0 server use five different validation tools. In this study we utilized ERRAT [19] and PROCHECK [20] to validate the modelled protein structure. After that, structural analysis between the native protein and mutant protein models was performed by ChimeraX software [21].

RESULTS

Analysis of Serum Lipid Profile and Demographic Parameters

As expected, serum lipid profile was positively correlated to CVD among the current study subjects compared to that of healthy controls. The result of lipid profile for the selected CVD subjects for molecular study has been pointed out. The mean of serum cholesterol (mg/dl, mean±SD) was 191.96±76.96 in CVD patients and 164.94±21.88 in control; triglycerides level was found to be 205.79±89.74 in patients and 150.58±47.52 in control; serum Low Density Lipoprotein Cholesterol 120.32±65.06 in patients and 95.68±27.34 in controls. High Density Lipoprotein Cholesterol was 35.20±24.94 in patients and 47.36±15.78 in control. Other relevant socio-demographic risk factors associated with the prevalence of CVD calculated in the current study is shown in figure 1. Our results suggests that family history, hypertension, diabetes, smoking, statins and chronic kidney disorders (CKD) contributes 25%, 36%, 28%, 20%, 11% and 2% respectively for manifestation of CVD among the study subjects.

Prevalence of Risk Factors in CVD

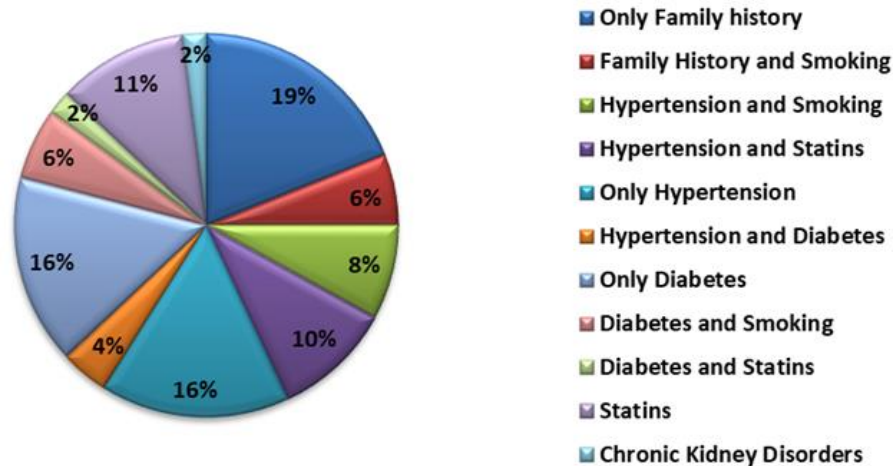


Figure 1. Socio-demographic risk factors associated with the prevalence of CVD, percentage of Family History, Hypertension, Diabetes, Smoking, Statins and Chronic Kidney Disorders.

Extraction of Genomic DNA from Blood and Quality Assessment

The blood samples were collected from the patients and the controls. The blood genomic DNA was extracted from each sample. The yielded DNA products were visualized by agarose gel electrophoresis (1% w/v) to assess the quality of DNA (Figure 2). All the extracted DNA showed bright high molecular weight bands (above 10kb) without much of smearing which indicates the quality of DNA was good. Optical density of extracted DNA was also checked at 260 and 280 nanometer wavelengths. Concentration of extracted was found to be between 120µg/ml to 346 µg/ml) and 260/280 ratio of all the extracted DNA ranged between 1.78 to 1.97, which indicates the purity of DNA was good enough to carry out downstream experimentations.

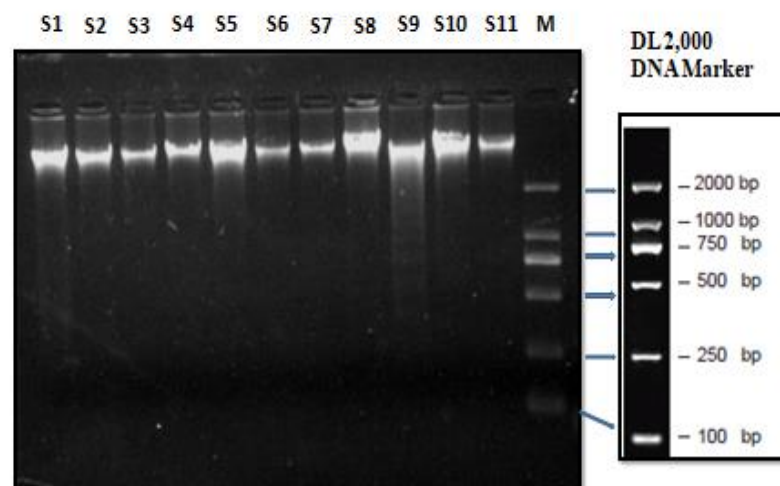


Figure 2. Visualization of extracted genomic DNA on 1% agarose gel electrophoresis. M represents DL 2,000 DNA marker ranging from 100bp to 2000bp. S1 to S11 denotes Genomic DNA from representative sample 1 to sample 11.

PCR amplification of *PCSK9* Gene

For amplification of target regions of *PCSK9* gene, PCR was carried out. To ensure the assessment of accurate size of PCR product, amplified product was visualized on 1% agarose gel under UV trans-illuminator. PCR amplifications using primer *PCSK9* Ex1 Fwd and *PCSK9* Ex1 rev yielded ~550 bp PCR product (Figure 3a) whereas primer *PCSK9* Ex2 Fwd and *PCSK9* Ex2 rev generated product with ~275 bp size (Figure 3b).

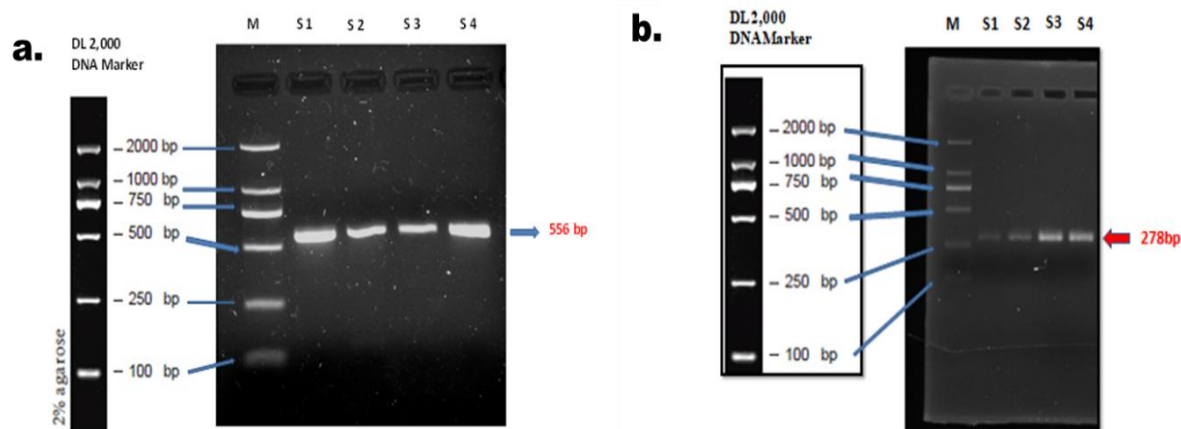


Figure 3. PCR amplification of *PCSK9* gene segments. a) Amplification of exon 1 and b) Amplification of exon 2. *PCSK9* EX-1. M represents DL 2,000 DNA marker ranging from 100bp to 2,000 bp. S1, S2, S3 and S4 denotes representative sample 1, 2, 3, 4 respectively. 3 μ l of PCR products were run on the gel at 80V for 45 minutes.

BLAST Search, Mutation Screening

Amplified PCR products of each sample, both CVD patients and controls were subjected to PCR purification and sequenced using Sanger sequencing. All the PCR products showed 99-100% identity with human *PCSK9* gene fragments. Subsequently, obtained sequences were subjected to multiple sequence alignment was performed separately for exon 1 and exon 2 fragments using DNAMAN to identified mutations. Interestingly we found seven mutations (A118G, T300C, GC447-448CG, GCC519-521CTG) in exon 1 of *PCSK9* gene of CVD patients with abnormal lipid profile (Figure 4).

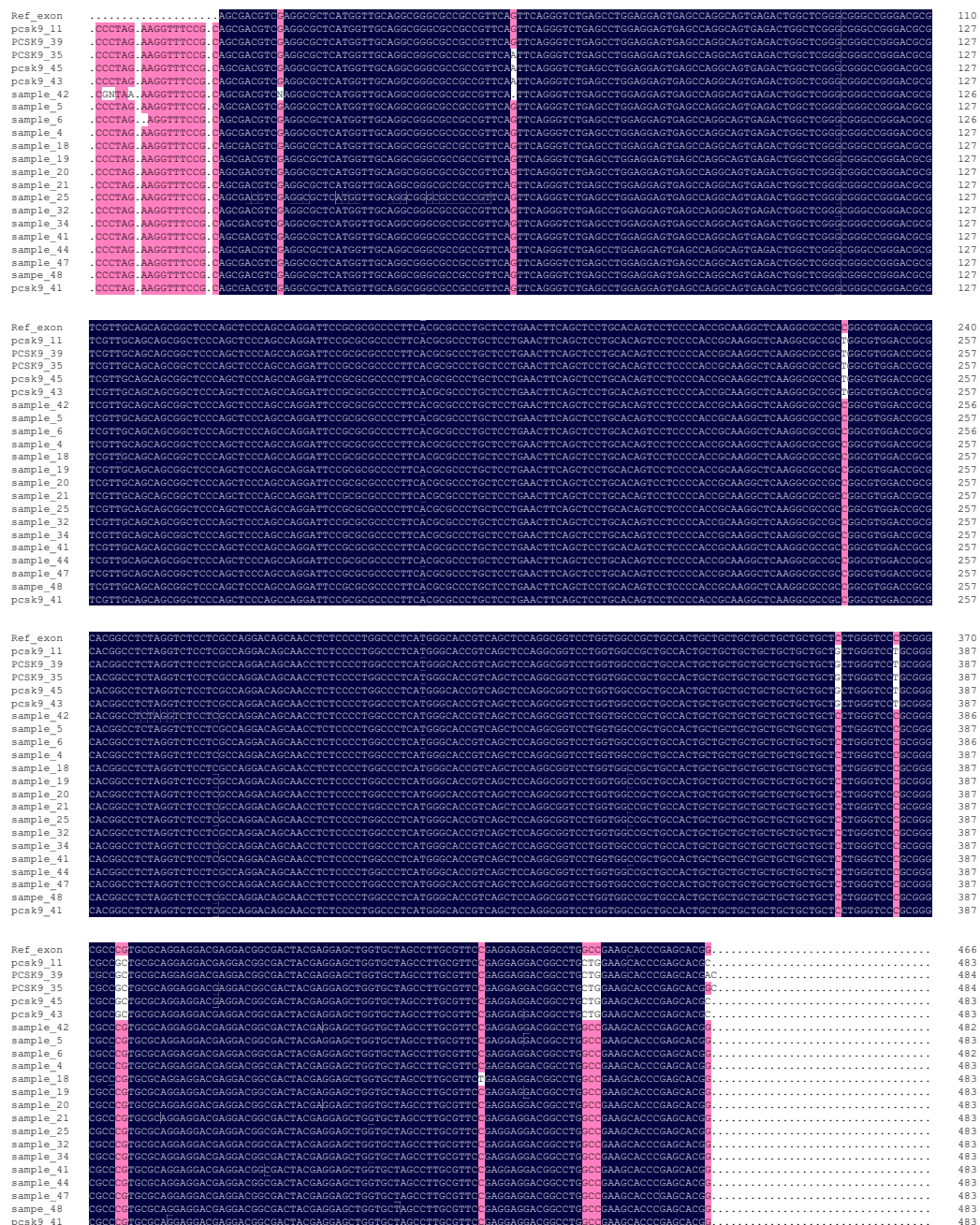
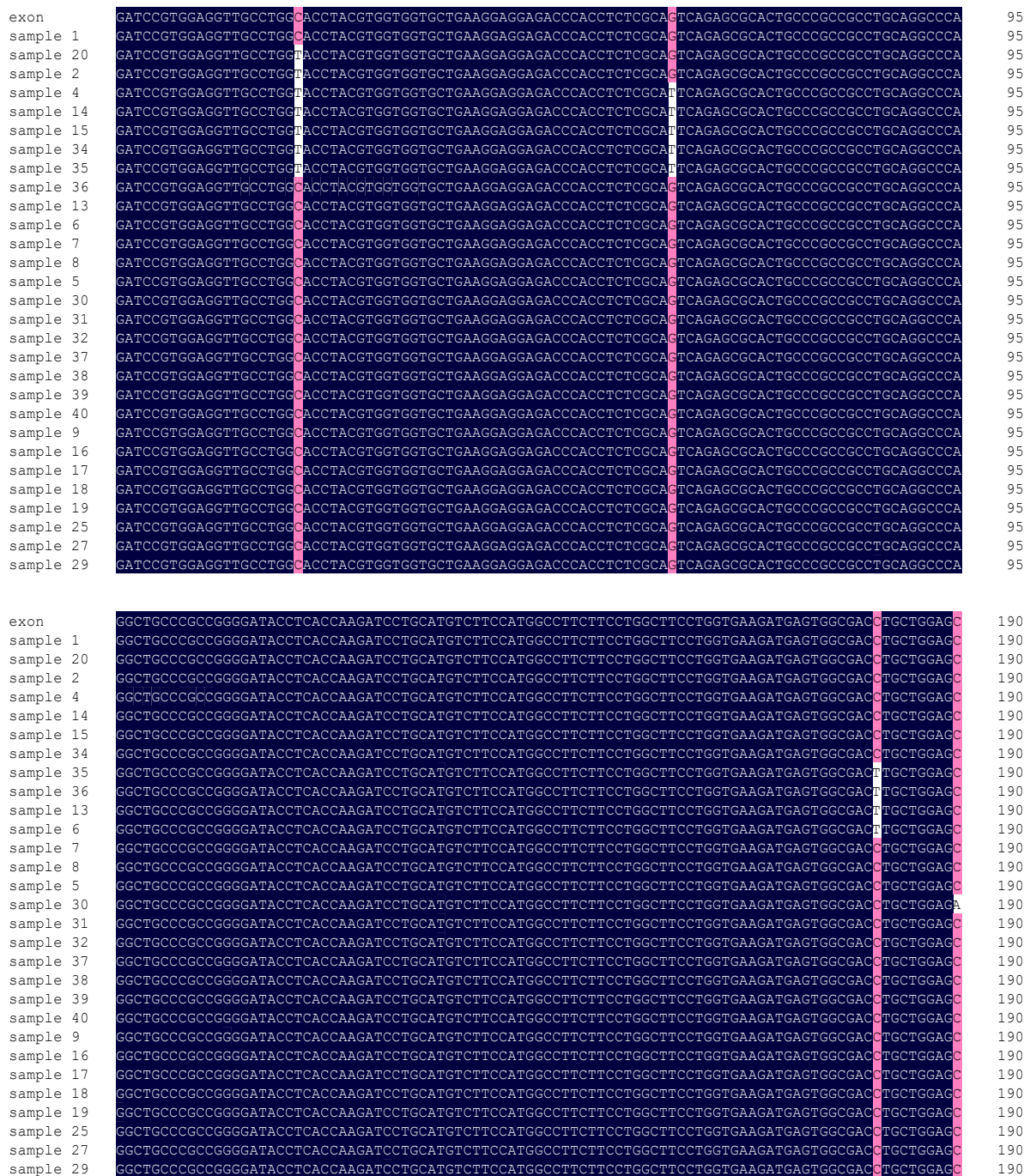


Figure 4. Multiple Sequence Alignment of exon 1 of PCSK9 gene with the reference exon sequence from NCBI (NG_009061.1). Mutations are seen in position: A118G, T300C, GC447-448CG, GCC519-521CTG. Pink color shows the variable region and black color shows the conserved regions.

On the other hand, Exon 2 showed (Figure 5) three-point mutations (C590T, G632T, C748T). As these mutations are in the exons of the gene, these brings change in the amino acid sequence in the PCSK9 protein resulting in abnormal structure and altered or reduced functions causing dyslipidemia and ultimately manifest CVD.



Bioinformatics Analysis

Prediction of Damaging or Disease-Causing Mutations:

In our study, for exon 1 all the mutation gives Benign for both HumDiv and HumVar prediction by Polyphen-2 Server. In case of exon 2, Q20H substitution gives probably damaging and

L63Q gives possibly damaging results in both HumDiv and HumVar analysis in Polyphen 2 computational tool (Table 2).

Table 2. Results of Polyphen-2 computational tool

Exon 1				
Mutation	HumDiv		HumVar	
R29A	BENIGN	0.002	BENIGN	0.001
A53L	BENIGN	0.054	BENIGN	0.006
Exon 2				
Mutation	HumDiv		HumVar	
Q20H	PROBABLY DAMAGING	0.995	PROBABLY DAMAGING	0.922
L63Q	POSSIBLY DAMAGING	0.829	POSSIBLY DAMAGING	0.541

Other three computational tools namely PhD-SNP, PANTHER and SNPs&Go give neutral results for Both mutation R29A and A53L in exon 1 of PCSK9. In case of exon 2, these three tools give the prediction as disease for both mutations (Table 3).

Table 3. Results of PhD-SNP, PANTHER and SNPs&Go computational tools

PhD-SNP			
Exon 1			
Mutation	Prediction	RI	Probability
R29A	Neutral	3	0.349
A53L	Neutral	3	0.344
Exon 2			
Mutation	Prediction	RI	Probability
Q20H	Disease	7	0.836
L63Q	Disease	5	0.746
PANTHER			
Exon 1			
Mutation	Prediction	RI	Probability
R29A	Neutral	7	0.147
A53L	Neutral	6	0.189
Exon 2			
Mutation	Prediction	RI	Probability
Q20H	Disease	7	0.847
L63Q	Disease	10	0.993
SNPs&Go			
Exon 1			
Mutation	Prediction	RI	Probability
R29A	Neutral	8	0.079
A53L	Neutral	9	0.053
Exon 2			
Mutation	Prediction	RI	Probability
Q20H	Disease	7	0.862
L63Q	Disease	7	0.839

Prediction of Protein Stability Change:

For predicting the protein stability changes, MUpro computational tool was utilized. R29A substitution in exon 1 and both mutation such as L63Q and Q20H in exon 2 decrease the stability of the proteins. On the other hand, only A53L mutation in exon 1 increase the stability of the protein (Table 4).

Table 4. MUpro results which indicates the stability of the proteins

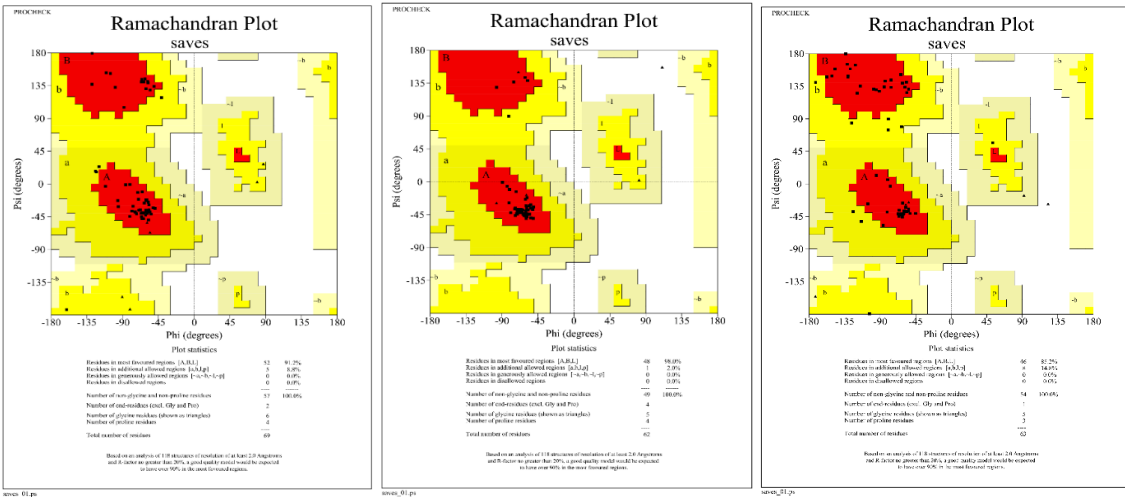
Exon 1		
Substitution	Prediction SVM	Score
R29A	DECREASE stability	-0.41416769
A53L	INCREASE stability	0.21662801
Exon 2		
Substitution	Prediction SVM	Score
L63Q	DECREASE stability	-2.1408564
Q20H	DECREASE stability	-0.74680409

Comparative 3D Protein Modeling:

Native and Mutant protein structures of Exon 1 and 2 of PCSK9 were modelled by using Robetta web server. Later, the structures were validated by ERRAT and PROCHECK validation tools. In ERRAT validation, all the protein structures show overall quality factor more than 80%. Validation results are shown in Table 5.

Table 5. Results of ERRAT validation tool of the modelled protein structure

Protein Models	Overall Quality Factor
PCSK9 Exon1 Ref	98.3607
PCSK9 Exon1 R29A A53L	100
PCSK9 Exon2 Ref	88.2353
PCSK9 Exon2 Q20H	92.8571
PCSK9 Exon2 L63Q	92
PCSK9 Exon2 Q20H L63Q	85.7143



(a)

(b)

(c)

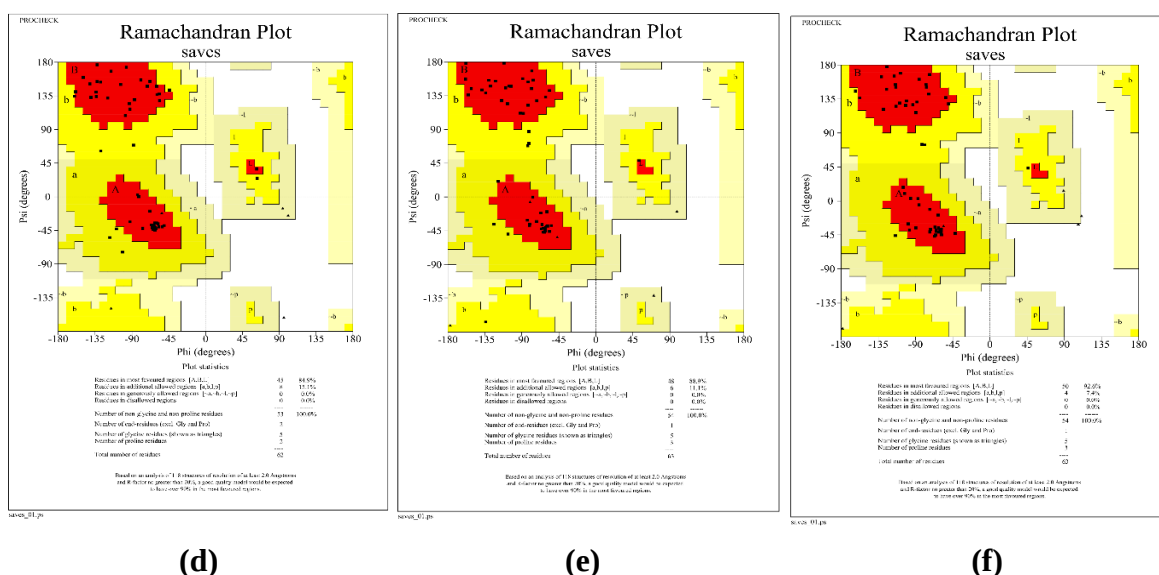


Figure 6. Ramachandran plot of native and mutated PCSK9 protein structure (a) PCSK9 Exon 1 Native Protein, (b) PCSK9 Exon 1 Mutated Protein with R29A and A53L Mutations, (c) PCSK9 Exon 2 Native Protein, (d) PCSK9 Exon 2 Mutated Protein with Q20H Mutation, (e) PCSK9 Exon 2 Mutated Protein with L63Q Mutation, (f) PCSK9 Exon 2 Mutated Protein with Q20H and L63Q Mutations.

In PROCHECK, all the models give more than 80% residues in most favored regions (Figure 6).

To compare the mutated proteins with the wild one, each mutated protein structure was superimposed on the native one (Figure 7 and 8). For each superimposition, Root Mean Square Deviation (RMSD) data was generated. RMSD values of the mutated models by comparing with the native one is given in Table 6.

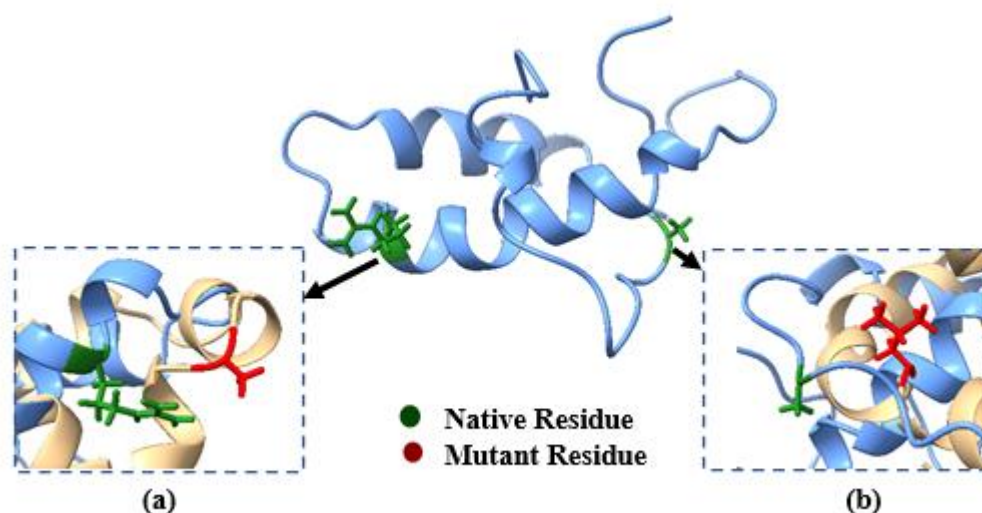


Figure 7. Comparison between native *PCSK9* gene Exon 1 and its mutated forms. (a) R29A: native Arginine substituted by Alanine at position 29, (b) A53L: native Alanine substituted by Leucine at position 53.

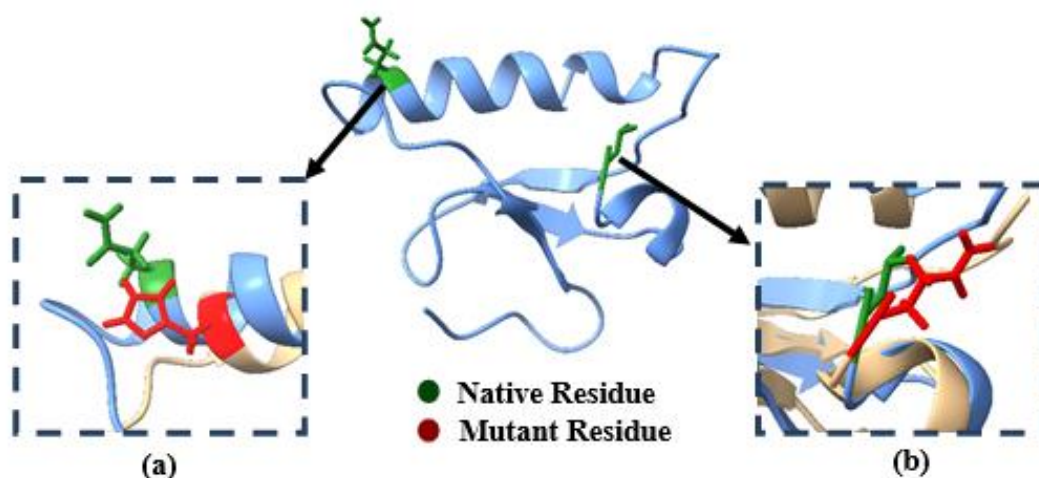


Figure 8. Comparison between native PCSK9 Exon 2 and its mutated forms, (a) Q20H: native Glutamine substituted by Histidine at position 20, (b) L63Q: native Leucine substituted by Glutamine at position 63.

Table 6. RMSD values from the superimposition of mutant protein structure with native protein

Exon No	Sample No	Substitutions	RMSD
Exon 1	11, 35, 39, 43, 45	R29A, A53L	1.35
Exon 2	4	Q20H	1.208
	2, 20	L63Q	1.276
	14, 15, 34, 35	Q20H, L63Q	1.351

DISCUSSION

CVD is a major health concern around the world including Bangladesh. The excess burden of CAD among South Asians appears to be primarily due to dyslipidemia that is characterized by high levels of LDL, triglycerides (TG), and total cholesterol and low levels of HDL and apoA1 [22]. Excessive use of fats, deep frying, reuse of cooking oil, and overcooking leading to destruction of folates may all contribute to dyslipidemia in this population [2]. However, genetic factors may also contribute in developing dyslipidemia leading to CVD which we studied in the current article for the first time in Bangladesh. As a matter of fact, CVD is a complex disease and life style, food habit, genetic trait along with environmental components and different physical and physiological conditions of an individual contributes for development of this diseases. Some forms of CVD exhibit a simple pattern of inheritance suggestive of a single causal gene that confers a large effect on phenotype. In 1985, Lehrman and colleagues directly sequenced the low-density lipoprotein receptor (LDLR) gene in a patient with homozygous familial hypercholesterolemia and uncovered 5 kilo base deletion that eliminated several exons, representing the first demonstration of a mutation for Mendelian CVD [23].

Since then, population-based study for identifying mutations in respective gene for determining the genetic risk for a country or ethnicity is of immense interest of research [24]. The present study recruited a total of 100 subjects including 75 CVD patients with abnormal HDL and LDL

level and 25 healthy individuals with no formal family history of CVD. Abnormal lipid profile was confirmed according to a set of guidelines provided by The American Heart Association, NIH and NCEP for fasting LDL and HDL levels. CVD patients with dyslipidemia were considered for PCSK9 gene mutation study. Statistical analysis of the biochemical parameters of the current study subjects showed that total cholesterol, LDL and TG level are comparatively higher in CVD patients than the healthy control groups as anticipated. However, in some patients with severe CVD, the level of HDL was found to be significantly lower. These findings are consistent with the other study studies related to lipid profile abnormality conducted on Bangladeshi population [25].

Hypertension, a major component of lipitension, damages the endothelium through altered shear stress and oxidative stress, resulting in increased endothelial cell synthesis of collagen and fibronectin, reduced nitric oxide-dependent vascular relaxation, and increased permeability to lipoproteins. Hypertension is also associated with an up-regulation of lipid oxidation enzymes [26]. LDL, especially oxidized LDL, is a major cause of endothelial dysfunction. Thus, lipitension contributes to atherosclerosis and the resultant vascular risk, through its effects on the level of the endothelium. In our study, the association of hypertension with CVD was 36% cases. The results reflect the fact that that hypertension might complicate and increase the fatality risk among those patients with CVD which is in accordance with results of previous studies related to Bangladeshi population.

Diabetes has been better studied in Bangladesh compared to the other immediate CVD risk factors. The pooled prevalence from these studies is significantly lower than the recent nationwide survey in adults (10%). Multiple factors, such as method of blood sugar measurement for diabetes and diagnostic criteria used for T2DM, are likely to explain this variation. We found that the prevalence of T2DM ranges from 5%–7% in rural areas whereas, 10%–17% in urban areas. The rural prevalence is slightly higher than that in South India and Pakistan (less than 5%), but it is considerably higher than China (3%) [27]. The present study showed the 28% prevalence of Diabetes in CVD patients which is compatible to the other study [28].

Smoking is considered to be a significant risk factor for the development of atherosclerosis. The atherosclerotic effects of cigarette smoke are due to a substantial degree to thrombosis-related events [23]. Nicotine is considered to be responsible for an increase in blood viscosity and platelet aggregation, since it prevents the production of prostacyclin that would limit platelet aggregation. The increase in platelet aggregation, the effect of nicotine on blood coagulation time, and the increase in blood viscosity, in combination with the increase in levels of LDL and VLDL, the reduction in HDL, and inflammatory processes, promote the creation of atherogenic plaque and the development of atherosclerosis [29]. The prevalence of smoking in Bangladesh is well established. While only 5%–7% females (both in rural and urban areas) are smokers, the proportion among males is high (37%–63%). The proportions do not differ markedly between urban and rural areas (29%–49% vs 26%–37%) [30] and our study showed the prevalence of smoking is 28% in CVD patients which is compatible to previous study but the studies on the strength of the association of tobacco consumption, particularly smokeless tobacco, with CVDs, however, are not yet available.

However, although there are several reports of unique mutations in the respective genes related to CVD, there were no such study reported in Bangladesh before this one. Along with the socio-demographic parameter analysis for CVD patient, dyslipidemia is another major risk factor which is largely contributed by the mutation in the proprotein convertase subtilisin/kexin 9 (PCSK9) gene. PCSK9 gene yields a 72-kDa protein molecule which contains 692 amino acids. Main function of PCSK9 protein is to bind to particular cell surface receptors to lead them toward intracellular degradation compartments, causing degradation of cell surface low-density lipoprotein receptors (LDLRs). Mutations in the gene alter PCSK9 protein's ability to function normally. In the current study, we have PCR amplified exon-1 and exon-2 of PCSK9 gene and sequenced them in order to identify mutations that might be related to CVD in Bangladeshi population. In the current study, we could detect seven mutations in exon-1 and three mutations in exon-2 of PCSK9 of some CVD patients. This is the first report of identifying mutation in PCSK9 gene in Bangladesh which might be used in future (through studying larger population size studies) as genetic marker or genetic risk factor for Bangladeshi population.

In computational analysis of the detected mutations in the exon 1 and exon 2 of PCSK9 gene reveals that, R29A and A53L alteration in exon 1 shows benign in Polyphen 2 tool, whether for exon 2, one alteration L63Q shows Possibly Damaging and another mutation in exon 2 Q20H shows Probably Damaging in Polyphen 2 tool (Table 2). This indicate that mutation in the exon 2 has the capability to cause the disease.

Other computational analysis by using PhD-SNP, PANTHER and SNPs&Go computational analysis also supports the results of Polyphen 2 tool. The results of these three tools are shown in Table 3.

In terms of protein stability analysis, MUpro webserver shows some interesting data (Table 4). In MUpro analysis R29A alteration shows the decrease of the protein stability where another mutation in exon 1 A53L shows increase stability. In exon 2 both mutations Q20H and L63Q shows decrease stability of the mutated proteins.

After that for structural comparison modelling of the native and the mutated protein structure was carried out by Robetta. All the modelled structure was validated using ERRAT and PROCHECK servers. All the models show more than 85% overall quality factor in ERRAT validation which indicated these models' quality is good (Table 5). In PROCHECK validation a Ramachandran Plot was given which shows the percentage of residues in most favoured regions. If more than 80-90% residues are in most favoured regions, the quality of the structure can be considered as good quality models [20]. In our study, all the modelled structures show more than 80% residues are in most favoured regions (Figure 6), which indicate the good quality of our modelled structures. Native structure of the exon 1 shows 91.2% residues are in most favoured regions and the mutated protein which contains both R29A and A53L mutations shows 98.0% residues are in most favoured regions. In exon 2, the native protein shows 85.2%, mutated protein that contain Q20H mutation shows 84.9%, protein with L63Q shows 88.9% and mutated protein which contains the both Q20H and L63Q alterations shows 92.6% residues in most favoured regions.

After modelling of all the native and mutated structure for both exon 1 and exon 2, each mutated protein structure was superimposed on native protein structure (Figure 7 and 8). By superimposing two models the root-mean-square deviation (RMSD) can be obtained which indicates the deviation of the two compared models. A more significant deviation is indicated by a greater RMSD value. At the times of Exon 1, mutated structure with R29A and A53L mutations give higher RMSD value than others mutated structure. In the exon 2 the mutated structure with double mutations Q20H and L63Q also give more RMSD value than the single mutated structure. RMSD values of the mutated proteins comparing with the native protein structure are given in table 6.

CONCLUSION

Mutation in *PCSK9* gene is strongly associated with lipid profile alteration and atherosclerosis. We conducted mutation screening for 75 patients and 25 control through molecular analysis such as DNA extraction, PCR amplification and at last step we conduct sequencing of the *PCSK9* gene specifically for exon 1 and 2. In this current study, two mutation R29A and A53L was found in exon 1 of PCSK9 protein. These two mutations in exon 1 shows benign in most of the computational analysis, though R29A mutation can decrease the stability of protein structure. In exon 2, mutation found in two positions such as, Q20H and L63Q.

Both mutations in exon 2 shows damaging characteristics in all computational analysis and both have the capability of decreasing the stability of the protein structure. These Q20H and L63Q mutation also shows structural changes when comparing with the native protein structure. The findings of this study can aid in develop prognosis tools for early detection of the cardiovascular disease specifically in Bangladeshi population which can reduce the risk of the CVD by aware people for future CVD.

Conflict of interest

The authors declare that there are no conflicts of interest.

Author contribution

Md. Abu Reza conceptualized and supervised the whole project. Md. Mahbubur Rahman, Ratna Khatun and Md. Jahangir Alam, Md. Mahmudul Hasan Maruf performed most of the experiments. Md. Tanjil Islam Shovon performed the bioinformatics analysis. Md. Zahidus Sayeed and Md. Ragib Shariar identified the CVD patients, collected sociodemographic data. Zennat Ferdousi, Kazi Md. Faisal Hoque and Khandaker Md. Khalid-Bin-Ferdaus analyzed the sociodemographic data. The first draft of the manuscript was written by Md. Mahbubur Rahman, Md. Mahmudul Hasan Maruf and Md. Tanjil Islam Shovon. All authors read and approved the manuscript.

Funding

The authors cordially express their humble gratitude towards Professor Golam Shabbir Sattar, the honorable Vice Chancellor of the University of Rajshahi, for partial financial support through allocation of a special research grant (A-701/6/109 (research), 2022) and by patronizing academic research. Current research project was also supported by the Grant for Advanced Research in Education (GARE) of Ministry of Education (Project ID: LS20201440 and

LS2019879), Ministry of Science and Technology grant SI no: 615; Project ID: SRG 225615 and HPNSP research grants of Bangladesh Medical Research Council (BMRC).

References

- [1] R. R. Huxley *et al.*, "Age-and sex-specific burden of cardiovascular disease attributable to 5 major and modifiable risk factors in 10 asian countries of the western pacific region," *Circ. J.*, vol. 79, no. 8, pp. 1662–1674, 2015, doi: 10.1253/circj.CJ-15-0661.
- [2] K. Fatema, N. A. Zwar, A. H. Milton, L. Ali, and B. Rahman, "Prevalence of risk factors for cardiovascular diseases in Bangladesh: A systematic review and meta-analysis," *PLoS One*, vol. 11, no. 8, pp. 1–14, 2016, doi: 10.1371/journal.pone.0160180.
- [3] M. Rafieian-Kopaei, M. Setorki, M. Douadi, A. Baradaran, and H. Nasri, "Atherosclerosis: Process, indicators, risk factors and new hopes," *Int. J. Prev. Med.*, vol. 5, no. 8, pp. 927–946, 2014.
- [4] D. Pahlajani, "Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes," *Yearb. Cardiol. 2016*, vol. 373, no. 3, pp. 273–273, 2016, doi: 10.5005/jp/books/12834_83.
- [5] R. Garg, "Association of atherosclerosis with dyslipidemia and co-morbid conditions : A descriptive study," vol. 6, no. 1, 2015, doi: 10.4103/0976-9668.149117.
- [6] A. Qasim and D. Rader, "Human genetics of variation in high-density lipoprotein cholesterol," *Curr. Atheroscler. Rep.*, vol. 8, no. 3, pp. 198–205, 2006, doi: 10.1007/s11883-006-0074-0.
- [7] H. Cohen and R. Shamir, "Lipid disorders in children and adolescents," in *Pediatric endocrinology*, CRC Press, 2006, pp. 305–316.
- [8] S. R. Daniels and S. C. Couch, "Lipid Disorders in Children and Adolescents," in *Sperling Pediatric Endocrinology: Expert Consult - Online and Print*, CRC Press, 2020, pp. 1004–1021.
- [9] J. D. Horton, J. C. Cohen, and H. H. Hobbs, "PCSK9: A convertase that coordinates LDL catabolism," *J. Lipid Res.*, vol. 50, no. SUPPL., pp. S172–S177, 2009, doi: 10.1194/jlr.R800091-JLR200.
- [10] D. Allard *et al.*, "Novel mutations of the PCSK9 gene cause variable phenotype of autosomal dominant hypercholesterolemia," *Hum. Mutat.*, vol. 26, no. 5, p. 497, 2005, doi: 10.1002/humu.9383.
- [11] Y. Cho, S. Y. Shin, S. Won, C. L. Relton, G. D. Smith, and M. J. Shin, "Alcohol intake and cardiovascular risk factors: A Mendelian randomisation study," *Sci. Rep.*, vol. 5, no. 1, p. 18422, 2015, doi: 10.1038/srep18422.
- [12] J. Miao *et al.*, "Role of insulin in the regulation of proprotein convertase subtilisin/kexin type 9," *Arterioscler. Thromb. Vasc. Biol.*, vol. 35, no. 7, pp. 1589–1596, 2015, doi: 10.1161/ATVBAHA.115.305688.
- [13] I. Adzhubei, D. M. Jordan, and S. R. Sunyaev, "Predicting functional effect of human missense mutations using PolyPhen-2," *Curr. Protoc. Hum. Genet.*, vol. 76, no. SUPPL.76, pp. 7–20, 2013, doi: 10.1002/0471142905.hg0720s76.
- [14] P. D. Thomas *et al.*, "PANTHER: A library of protein families and subfamilies indexed by function," *Genome Res.*, vol. 13, no. 9, pp. 2129–2141, 2003, doi: 10.1101/gr.772403.
- [15] E. Capriotti, R. Calabrese, and R. Casadio, "Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information," *Bioinformatics*, vol. 22, no. 22, pp. 2729–2734, 2006, doi: 10.1093/bioinformatics/btl423.

-
- [16] R. Calabrese, E. Capriotti, P. Fariselli, P. L. Martelli, and R. Casadio, "Functional annotations improve the predictive score of human disease-related mutations in proteins," *Hum. Mutat.*, vol. 30, no. 8, pp. 1237–1244, 2009, doi: 10.1002/humu.21047.
- [17] J. Cheng, A. Randall, and P. Baldi, "Prediction of protein stability changes for single-site mutations using support vector machines," *Proteins Struct. Funct. Genet.*, vol. 62, no. 4, pp. 1125–1132, 2006, doi: 10.1002/prot.20810.
- [18] D. E. Kim, D. Chivian, and D. Baker, "Protein structure prediction and analysis using the Robetta server," *Nucleic Acids Res.*, vol. 32, no. WEB SERVER ISS., pp. W526–W531, 2004, doi: 10.1093/nar/gkh468.
- [19] C. Colovos and T. O. Yeates, "ERRAT: an empirical atom-based method for validating protein structures," *Protein Sci.*, vol. 2, no. 9, pp. 1511–1519, 1993.
- [20] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J. Appl. Crystallogr.*, vol. 26, no. 2, pp. 283–291, 1993, doi: 10.1107/s0021889892009944.
- [21] E. F. Pettersen *et al.*, "UCSF ChimeraX: Structure visualization for researchers, educators, and developers," *Protein Sci.*, vol. 30, no. 1, pp. 70–82, 2021, doi: 10.1002/pro.3943.
- [22] E. A. Enas, V. Mohan, M. Deepa, S. Farooq, S. Pazhoor, and H. Chennikkara, "The metabolic syndrome and dyslipidemia among Asian Indians: a population with high rates of diabetes and premature coronary artery disease," *J. Cardiometab. Syndr.*, vol. 2, no. 4, pp. 267–275, 2007, doi: 10.1111/j.1559-4564.2007.07392.x.
- [23] K. Ball, "The Basis for Action," pp. 822–826, 1974.
- [24] W. Lieb and R. S. Vasan, "Scientific Contributions of Population-Based Studies to Cardiovascular Epidemiology in the GWAS Era," *Front. Cardiovasc. Med.*, vol. 5, no. June, pp. 1–6, 2018, doi: 10.3389/fcvm.2018.00057.
- [25] S. K. Das *et al.*, "Serum Lipid Profile Among Urban and Rural Bangladeshi Population Obesity & Weight Loss Therapy," vol. 2, no. 8, pp. 8–11, 2012, doi: 10.4172/2165-7904.1000148.
- [26] M. Kaplan and M. Aviram, "Oxidized Low Density Lipoprotein : Atherogenic and Proinflammatory Characteristics during Macrophage Foam Cell Formation . An Inhibitory Role for Nutritional Antioxidants and Serum Paraoxonase," vol. 37, no. 8, pp. 777–787, 1999.
- [27] M. Wasay, I. A. Khatrri, and S. Kaul, "Stroke in South Asian countries," *Nat. Publ. Gr.*, vol. 10, no. 3, pp. 135–143, 2014, doi: 10.1038/nrneurol.2014.13.
- [28] T. J. Chahil and H. N. Ginsberg, "Diabetic Dyslipidemia," *Endocrinol. Metab. Clin. North Am.*, vol. 35, no. 3, pp. 491–510, 2006, doi: 10.1016/j.ecl.2006.06.002.
- [29] R. Ross and J. Glomset, "Response to Injury and Atherogenesis," pp. 675–684, 1976.
- [30] A. Ashraf *et al.*, "Health and Demographic Surveillance Systems in Asia," no. May, 2009, doi: 10.3402/gha.v2i0.1997.
-