Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI - 10.26479/2018.0401.04**

IN SILICO ANALYSIS OF LDLR AND PCSK9 GENES INVOLVED IN CHOLESTEROL METABOLIC PATHWAY

Nisha Gautam*[1], Satbir Kaur[2], Inderpreet Kaur[3], Parmjeet Singh[3], Kiranjeet Kaur Galhna[4], Kamaljeet Kaur[4]

Department of Human Genetics Punjabi University, Patiala, Punjab, India

ABSTRACT: Cholesterol is an important precursor for various steroid hormones which includes sex hormones (estrogens, progesterone and testosterone), vitamin D and bile acids. The two genes LDLR and PCSK9 are the protein coding genes which play major role in cholesterol metabolism. As we know that SNPs has an important role in determining the disease susceptibility. SNPs study is important in population studies and to find the variation among the populations. For this purpose, the selection of targeted SNPs is very important. Hence in the present study we performed the in silico analysis of LDLR and PCSK9 gene. A total number of ns SNPs of LDLR gene were evaluated by SIFT and I-Mutant (SNP damage prediction tools). Out of these 577 (54.38%) of nsSNPs were predicted to affect the stability of the LDLR protein. SIFT predicted 16 (1.15 %) of nsSNPs as damaging. Further the total number of nsSNPs were predicted to be deleterious by all two prediction software were 14 (1.01%). Similarly a total number of nsSNPs of PCSK9 gene were evaluated by SIFT and I-Mutant (SNP damage prediction tools). Out of these 277 (77.82%) of nsSNPs were predicted to affect the stability of PCSK9 protein. SIFT predicted 12 (3.16%) of nsSNPs as damaging. Further the total number of nsSNPs predicted to be deleterious by all two prediction software were 10 (2.63%). Besides this, basic information of these genes were also studied from GeneCards.

KEYWORDS: Single nucleotide polymorphism, Cholesterol, LDLR, PCSK9

***Corresponding Author: Ms: Nisha Gautam (Research Scholar)**

Department of Human Genetics Punjabi University, Patiala, Punjab, India

* Email Address:gautamn551@gmail.com

1.INTRODUCTION

Cholesterol is the precursor for life sustaining steroid hormones and a vital substance for cellular integrity, its insolubility in aqueous solutions creates numerous problems for its transport and excretion. It is a ubiquitous constituent of cell membranes, steroids, bile acids and signaling molecules. Cholesterols are hydrophobic and mostly insoluble in blood so they require transport within hydrophilic and spherical structures called lipoproteins which possess surface proteins (apoproteins, or apolipoproteins) that are cofactors and ligands for lipid-processing enzymes. Due to defects in choletsreol metabolizing pathway (LDLR and PCSK9) can lead to two major abnormalities which include familial hypercholesterolemia (FH) and coronary heart disease (CHD) [1]. CHD is the cause of 17.1 million deaths per year throughout the world [2]. Thus CHD is today the largest single contributor to global mortality and will continue to dominate mortality trends in the future. It has been reported that the majority of deaths (39%) in low and middle-income countries under the age of 70 years are due to CHD [3]. Premature coronary heart disease mortality is higher in UK, Indians and Asians than in European whites. Increased mortality in Indian Asians is particularly striking in men aged 30 to 39 years in whom the relative risk of death coronary heart disease is 2 and in men aged 20 to 29 whose relative risk is 3 compared with age matched European whites [4]. Prevalence of FH worldwide is 1 in 200-500 people. An estimated 1.3 million people in the U.S. have FH [5]. In India highest rates of high LDL-C in Tamilnadu (15.8%). The prevalence of the heterozygous form of FH is 1 in 500 and the prevalence of the homozygous form of FH is 1 in 1,000,000 [6]. In FH, there is reduced clearance of LDL-C from plasma because of defective activity of LDLR. FH heterozygotes have a 2- to 3-fold increase in the levels of LDL cholesterol. FH homozygotes have a 5 to 8 fold increase in LDL-C levels and develop CHD (Coronary Heart Disease) in the 2nd decade of life [7]. As we know that SNPs has an important role in determining the disease susceptibility. SNPs study is important in population studies and to find the variation among the populations. SNPs study plays an important role in personalized medicines. There are various association studies which are reported on LDLR and PCSK9 gene [8-9]. The above reported studies have shown that the these two genes (LDLR, PCSK9) are playing major role in maitaining the cholesterol homeostasis and due to mutation in these genes may cause alteration in the metabolism of cholesterol. The above described genes are the key players in cholesterol metabolism by keeping this in view we have decided to perform the in silico analysis of LDLR and PCSK9 gene.

2. MATERIALS AND METHODS

Gene databases

Different databases were used to collect the information regarding LDLR and PCSK9 genes. These include information about the cytogenetic location, pathways of LDLR and PCSK9 protein, size, molecular mass, post translational modifications, domain, orthologous of LDLR and PCSK9 genes and sub cellular locations of LDLR and PCSK9 proteins. All this information was collected from

GeneCards. GeneCards is a searchable, integrative database that provides comprehensive, user-friendly information on all annotated and predicted human genes. It automatically integrates gene-centric data from ~125 web sources, including genomic, transcriptomic, proteomic, genetic, clinical and functional information (http://www.genecards.org/). The SNPs data on LDLR and PCSK9 genes were collected from dbSNP (NCBI) (https://www.ncbi.nlm.nih.gov/snp/?term=LDLR/term=PCSK9). The Single Nucleotide Polymorphism database (dbSNP) is a free public archive for genetic variation within and across different species developed and hosted by the National Center for Biotechnology Information (NCBI) in collaboration with the National Human Genome Research Institute (NHGRI). Although the name of the database implies a collection of one class of polymorphisms only (i.e., single nucleotide polymorphisms (SNPs), it in fact contains a range of molecular variations: single nucleotide polymorphisms, short deletion and insertion polymorphisms, microsatellite markers or short tandem repeats, multinucleotide polymorphisms and heterozygous sequences. The dbSNP accepts apparently neutral polymorphisms, polymorphisms corresponding to known phenotypes and regions of no variation. It was created in September 1998 to supplement GenBank; NCBI's collection of publicly available nucleic acid and protein sequences. The information regarding distribution of single nucleotide polymorphism (SNP) on various regions of LDLR gene i.e. 5' UTR, 3' UTR, intronic and coding region was collected from this database. dbSNP is also helpful in sorting the nonsynonymous SNPs from synonymous SNPs alongwith information regarding residual change at respective SNP's position. This database also provides the total allele length and FASTA sequence of the respective SNP.

Prediction for deleterious/ damaging SNPs

Single nucleotide polymorphism (SNPs) consists of human genetic variations, occurring with an average density of ~1/1000 nucleotides of a genotype. SNPs are either neutral allelic variants or missense variants. The SNPs that are causing missense residual changes at amino acid level are likely to affect the structure and function of protein. Such effect of a residual change at the level of single nucleotide variation can be predicted by different in silico tools. These include SIFT (Sorting of Intolerance from Tolerance) and I-Mutant.

SIFT (Sorting of Intolerance from Tolerance)

SIFT predicts whether an amino acid substitution affects protein function. SIFT prediction is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences, collected through PSI-BLAST. SIFT can be applied to naturally occurring nonsynonymous polymorphisms or laboratory-induced missense mutations. It is a sequence homology based tool that sorts intolerant from tolerant amino acid substitutions and predicts whether an amino acid substitution in a protein will have a phenotypic effect. SIFT is based on the premise that protein evolution is correlated with protein function. Positions important for function should be conserved in an alignment of the protein family whereas unimportant positions

should appear diverse in an alignment.

Working of SIFT

SIFT takes a query sequence and uses multiple alignment information to predict tolerated and deleterious substitutions for every position of the query sequence. SIFT is a multistep procedure that (1) searches for similar sequences, (2) chooses closely related sequences that may share similar function to the query sequence,(3) obtains the alignment of these chosen sequences, and (4) calculates normalized probabilities for all possible substitutions from the alignment. Positions with normalized probabilities less than 0.05 are predicted to be deleterious, those greater than or equal to 0.05 are predicted to be tolerated. The SIFT prediction for substitution were given as shown in table 2.1

Table 2.1: Showing the SIFT prediction and substitution

OUTPUT	DESCRIPTION
SIFT Score	Ranges from 0 to 1. The amino acid substitution is predicted damaging if the score is <0.05 and tolerated if score is >/=0.05
Median info	Ranges from 0 to 4.32, ideally the number would be between 2.75 and 3.5. This is used to measure the diversity of the sequence used for prediction. A warning will occur if this is greater than 3.25 because this indicates that the prediction was based on closely related sequence.
Seq at position	This is the number of sequences that have an amino acid at the position of prediction.

I-MUTANT

I-Mutant 2.0 is a support vector machine (SVM) based tool for the automatic prediction of protein stability changes upon single point mutations. I-Mutant 2.0 predictions are performed either from the protein structure or, more importantly, from the protein sequence. I-Mutant 2.0 can be used both as a classifier for predicting the sign of the protein stability change upon mutation, and as a regression estimator for predicting the related Delta G (δG) values. Acting as a classifier, I-Mutant 2.0 correctly predicts (with a cross-validation procedure) 80% or 77% of the dataset, depending on the usage of structural or sequence information, respectively. When predicting DeltaG values associated with mutations, the correlation of predicted with expected/experimental values is 0.71 (with a standard error of 1.30 kcal/mol) and 0.62 (with a standard error of 1.45 kcal/mol) when structural or sequence information are respectively adopted. The web interface allows the selection of a predictive mode that depends on the availability of the protein structure and/or sequence. In the latter case, the web server requires only pasting of a protein sequence in a raw format.

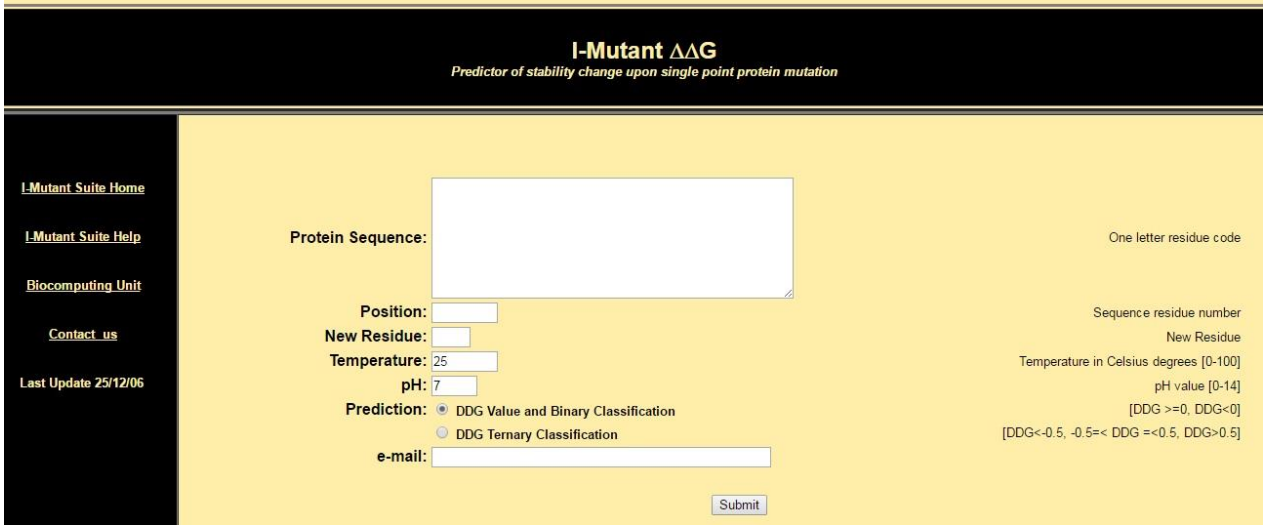

Figure 2.1: This picture is showing a snapshot of input format of I-Mutant.

3. RESULTS AND DISCUSSION

Cytogenetic location

A locus defined as the position of the gene. The cytogenetic location of LDLR and PCSK9 genes are 19p13.2 and 1p.32.3 as shown in figure 3.1(a) and 3.1(b) respectively. The molecular location of LDLR gene ranges from 11,089,36 bp to 11,133,830 bp and contain total 44,469 bases. The LDLR protein contains 860 amino acids and 95376 Dalton molecular weight of LDLR protein. PCSK9 gene located on chromosome 1 having size 25,378 bases starting with 55,039,476 bp to 55,064,853 bp. The protein coded by PCSK9 comprises of 692 amino acids having molecular mass of 74286 Da.

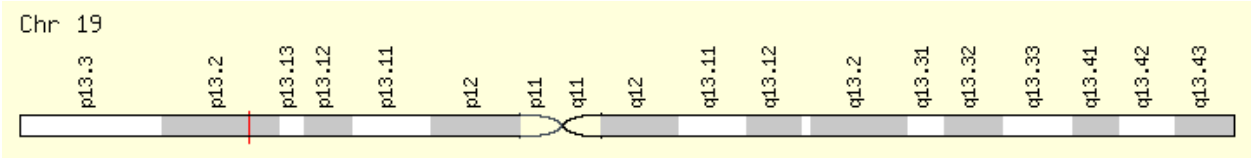

Figure 3.1 (a): The cytogenetic location of LDLR gene on chromosome 19.

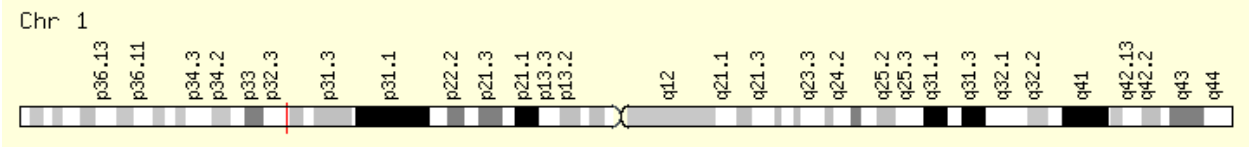

Figure 3.1 (b): Cytogenetic location of PCSK9 gene on chromosome no: 1

Post translational modification of LDLR and PCSK9 genes

The LDLR gene undergo several post- translational changes which are N- and O-glycosylated, Ubiquitinated by MYLIP leading to degradation, Glycosylation, modification sites at PhosphoSitePlus. We also have collected the information regarding PTM ofPCSK9 gene and we found that it undergo glycosylation at a cleavage by furin and PCSK5 generates a truncated inactive protein that is unable to induce LDLR degradation. Phosphorylation of PCSK9 protects the propeptide against proteolysis. Undergoes autocatalytic cleavage in the endoplasmic reticulum to release the propeptide from the N-terminus and the cleavage of the propeptide is strictly required for its maturation and activation. The cleaved propeptide however remains associated with the catalytic

domain through non-covalent interactions, preventing potential substrates from accessing its active site. As a result, it is secreted from cells as a propeptide-containing, enzymatically inactive protein.

Gene family of LDLR and PCSK9 gene

The LDLR and PCSK9 genes family includes LDLR, VLDLR, LRP5/6, LRP1, LRP2 and PCSK1, PCSK2, FURIN, PCSK4, PCSK5, PCSK6, PCSK7, MBTPS1, PCSK9 respectively, further these all play a pivotal role in cholesterol homeostasis and lipid metabolism.

Pathways of LDLR and PCSK9 genes

A pathway can trigger the assembly of new molecules. They can turn genes on and off or spur a cell to move. Major pathways of LDLR and PCSK9 genes are listed below in table 3.1 (a) and 3.1 (b) respectively.

Table 3.1(a): Showing the different pathways of LDLR gene and their role

S. No	Pathways	Role
1.	Lipoprotein metabolism	Lipid mobilization, digestion and transport, LDL-mediated lipid transport.
2.	Statin Pathway	Pharmacodynamics, , inhibition of endogenous cholesterol production.
3.	Metabolism of fat-soluble vitamins	Vitamin D and Vitamin K metabolism, retinoid metabolism and transport, absorption of fat.
4.	Folate Metabolism	Adipocyte metabolism and vitamin B12 metabolism.
5.	Clathrin-mediated endocytosis	Uptake of LDL via LDLR receptor and signal transduction.
6.	Vesicle mediated transport	LDL transport and Membrane trafficking.
7.	Insulin secretion	Clearance of circulating ApoB particles in the liver via the low density lipoprotein receptor (LDLR), LDLR-related protein 1 (LRP1).

The role of these pathways involved the activation of enzyme and catabolism of LDL, inhibition of endogenous cholesterol production by competitive inhibition of HMG-CoA reductase in statin pathway, absorption of fat in fat-soluble vitamins metabolism, adipocytes metabolism in folate metabolism, uptake of LDL in clathrin-mediated endocytosis, LDL transport in vesicle mediated transport and clearance of circulating ApoB particles in the liver via the LDLR. Such pathways show the importance of LDLR gene in maintaining the cholesterol homeostasis by regulating the metabolism of cholesterol.

Table 3.1 (b): The pathways of PCSK9 gene and their role

S.No	Pathways	Role
1.	Cholesterol metabolism	Plasma cholesterol homeostasis,
2.	Metabolism	Lipid metabolization
3.	Evolcumab metabolism	Inhibition of PCSK9 when binds to LDLR
4.	Glucose/energy metabolism	Glucose homeostasis

PCSK9 plays an important role in plasma cholesterol homeostasis pathway, lipid metabolism these pathways to maintain cholesterol level. Such pathways show the importance of PCSK9 gene in maintaining cholesterol homeostasis by regulating the metabolism of cholesterol.

Subcellular locations of LDLR and PCSK9

The cells of eukaryotic organisms are elaborately subdivided into functionally distinct membrane-bound compartments which are known as their subcellular location within the cell. The below diagrams show the subcellular locations of LDLR and PCSK9 genes and the confidence number given to each organelle is according to the differential amount of LDLR and PCSK9 genes expression present in them. The highest number depicts the largest amount of gene expression and the lowest confidence depicts the least amount of gene expression in the organelle.

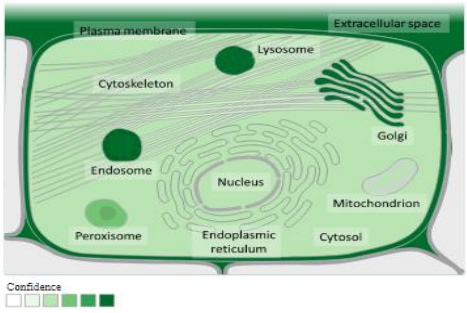

Figure 3.2 (a):The above diagram showing the differential abundance of LDLR expression in various organelles and the intensity of color depicts the relative abundance of LDLR in different organelles.

The level of expression of LDLR gene was found to be maximum in plasma membrane, extracellular matrix, lysosome, golgi apparatus and endosome as shown in the figure 3.2 (a). The maximum abundance of LDLR gene in the plasma membrane and extracellular matrix are because it code for LDLR and which is only functional when it integrate in the plasma membrane and the LDL recognize these receptors extracellularly. As shown in the above diagram it is also abundance in lysosome and endosome because these are the sites for its degradation. As we know that the endosome and lysosome is the site of LDL metabolism. It is also found abundantly in the golgi apparatus because it is the

major site of post translational modification of LDLR protein.

Table 3.2 (a): The differential expression of LDLR gene in the cellular compartments.

COMPARTMENT	CONFIDENCE
Plasma membrane	5
Extracellular	5
Lysosome	5
Golgi apparatus	5
Endosome	5
Peroxisome	3
Mitochondrian	2
Nucleus	2
Endoplasmic reticulum	2
Cytosol	2
Cytoskeleton	1

Besides this the LDLR is also found to be expressed in nucleus, endoplasmic reticulum and cytosol as given in table 3.2 (a). Nucleus is the site where LDLR gets transcribed. In the cytosol translation of LDLR takes place. The LDLR expression is also present in endoplasmic reticulum but in leaser amount because it is the minor site of post translational modification of LDLR protein.

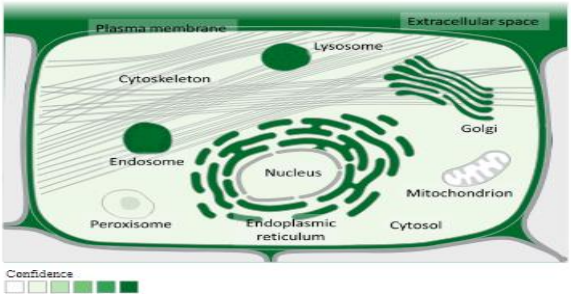

Figure 3.2 (b): The differential abundance of PCSK9 in various organelles and the intensity of color depicts the relative abundance of PCSK9 gene in different organelles.

The above diagram shows the subcellular location of PCSK9 gene and the confidence number given to each organelle is according to the differential amount of PCSK9 gene expression present in them. The highest number depicts the largest amount of PCSK9 gene expression and the lowest confidence depicts the least amount of PCSK9 gene expression in the different cell organelles. The level of expression of PCSK9 gene was found to be maximum in plasma membrane, extracellular matrix, endosome, endoplasmic reticulum and lysosome as shown in the figure 3.2 (b). The maximum abundance of PCSK9 gene in the plasma membrane and extracellular matrix are because the plasma membrane secretes this protein in an active form into the extracellular matrix. It is also found to be

abundant in endosome and lysosome because it is the site where PCSK9 bound to EGF-A domain of LDLR. The activation of PCSK9 occurs at endoplasmic reticulum by autocleavage zymogen form of PCSK9.

Table 3.2 (b): The differential expression of PCSK9 gene in the sub-cellular compartments

S.No	Compartment	Confidence
1.	Plasma membrane	5
2.	Extracellular	5
3.	Endoplasmic reticulum	5
4.	Lysosome	5
5.	Golgi apparatus	5
6.	Endosome	5
7.	Cytoskeleton	1
8.	Peroxisome	1
9.	Nucleus	1
10.	Cytosol	1

Besides this the PCSK9 is also found to be expressed in cytoskeleton, peroxisome, nucleus and cystosol as given in table 3.2(b). Nucleus is the site where PCSK9 gets transcribed. In the cytosol translation of PCSK9 takes place. Peroxisome is also the site of cholesterol metabolism so the PCSK9 is also found in peroxisome.

Orthologs and paralogs of LDLR and PCSK9 genes

Orthologous genes are homologous genes that diverged after evolution gives rise to different species which is known as speciation. Paralogous genes are homologous genes that have diverged within one species and these genes arise during gene duplication. The orthologs of LDLR gene are given in the table 3.3 (a).

Table 3.3 (a): The orthologs of LDLR gene.

Organism	Taxonomy	Gene
Mouse	Mammalia	LDLR
Chicken	Aves	___
Lizard	Reptilia	LDLR
African clawed frog.	Amphibia	LOC397757
Zebrafish	Actinopterygii	LDLRA

Paralogs of LDLR gene are LRP8, VLDLR, LRP4, LRP1B, LRP1, LRP2, EGF, SORL1, LRP5, LRP6. The orthologs of PCSK9 gene are given in the table 3.3(a)

Table 3.3 (b): Orthologs for PCSK9 gene

Sr.No	Organism	Taxonomy	Gene
1.	Chimpanzee (*Pan troglodus*)	Mammalia	PCSK9
2.	Rat (*Ratus nonvegious*)	Mammalia	PCSK9
3.	Mouse (*Mus musculus*)	Mammalia	PCSK9
4.	Dog (*Canis familiaris*)	Mammalia	PCSK9
5.	Oppossum (*Monodelphis domestica*)	Mammalia	PCSK9
6.	Platypus (*Ornythorynchus anatinus*)	Mammalia	PCSK9
7.	Chicken (*Gallus gallus*)	Aves	PCSK9
8.	Lizard (*Anolis carolinensis*)	Reptilia	PCSK9
9.	Tropical clawed frog (*Silurana tropicalis*)	Amphibia	PCSK9
10.	Zebra fish (*Danio rerio*)	Actinopterygi	PCSK9
11.	Baker yeast (*Saccharomyces cervisiae*)	Saccharomycetes	YSP3

12.	**Fission yeast (*Schizyosaccharo pombe*)**	**Schizosaccharomycetes**	**Psp3**

Paralogs for PCSK9 are not reported yet.

Disorders due to mutation in LDLR and PCSK9 genes

Every gene consists of unique nucleotide sequence and alteration of a single base will leads to mutation which ultimately cause some disorders. There are some disorders associated with LDLR and PCSK9 genes are given below in table 3.4 (a) and 3.4 (b) respectively.

Table 3.4 (a): The different disorders associated with LDLR gene.

S. No	Disorder	Definition	Gene involved	Pathways
1.	**Familial hypercholesterolemia**	**Familial hypercholesterolemia is a condition characterized by very high levels of cholesterol in the blood due to mutations in the LDLR gene.**	**LDLR**	**SREBF and miR33 in cholesterol and lipid homeostasis and folate metabolism.**
2.	**Atherosclerosis**	**Atherosclerosis is a disease in which plaque builds up inside your arteries.**	**LDLR**	**SREBF and miR33 in cholesterol and lipid homeostasis.**
3.	**Defective Apolipoprotein B-100**	**Apolipoprotein B deficiency is an autosomal dominant disorder resulting from a missense mutation.**	**LDLR**	**Lipoprotein metabolism, SREBF and miR33 in cholesterol and lipid homeostasis, Folate Metabolism.**
4.	**Hepatitis c**	**Hepatitis C, also known as chronic hepatitis c, is related to hepatitis c virus.**	**LDLR**	**Hepatitis c super pathway.**
5.	**Coronary artery disease**	**Coronary Artery Disease, also known as coronary artery anomaly, is related to hypercholesterolemia.**	**LDLR**	**Fat digestion and absorption and A-beta pathways.**
6.	**Ischemic stroke**	**Stroke, Ischemic, also known as cerebral infarction, is related to peripheral artery disease.**	**LDLR**	**Dissolution of fibrin clot and angiogenesis (CST).**

7.	**Norum disease**	**Norum Disease, also known as lecithin:cholesterol acyltransferase deficiency.**	**LDLR**	**Folate Metabolism.**
8.	**Smith-lemli-opitz syndrome**	**mith-Lemli-Opitz syndrome is an autosomal recessive frequent inborn disorder of sterol metabolism.**	**LDLR**	**Sterol Regulatory Element-Binding Proteins (SREBP) signalling and Folate Metabolism.**
9.	**Xanthomatosis**	**A lipid storage disease that is characterized by the desposition of yellowish cholesterol-rich material in tendons or other body parts.**	**LDLR**	**SREBF and miR33 in cholesterol and Clathrin-mediated endocytosis.**
10.	**Arcus senilis**	**Arcus senilis is an old age syndrome where there is a white, grey, or blue opaque ring in the corneal margin or white ring in front of the periphery of the iris**	**LDLR**	**Evolocumab Mechanism.**

Table 3.4 (b): The different disorders associated with PCSK9 gene

S.NO	Name	Definition	Gene involve	Pathways
1.	**Familial hypercholesterolemia**	**Concentration of cholesterol increased in the blood.**	**PCSK9**	**Evolocumab Mechanism**
4.	**Gastric papillary adenocarcinoma**	**Gastric Papillary Adenocarcinoma is also known as papillary adenocarcinoma of stomach, and has symptoms as symptoms including abdominal pain and dyspepsia**	**PCSK9**	**Evolocumab Mechanism**
5.	**Hypercholesterolemia, familial, autosomal recessive**	**Autosomal recessive mode of inheritance**	**PCSK9**	**Evolocumab Mechanism, Lipoprotein Metabolism**

5.	**Acrus senilis**	**Grey opaue ring in cornea**	**PCSK9**	**Evolocumab**

As shown in above tables that there are number of diseased conditions in which mutations in LDLR and PCSK9 genes are prevalent.

Drugs targeting LDLR and PCSK9 genes

There are various drugs which is used to target the LDLR and PCSK9 genes. Some drugs targeting the genes involved in cholesterol metabolic pathway so directly or indirectly influence the expression of LDLR and PCSK9 genes. Some drugs are given below along with their target, mechanism of action and their pharmacological action specific to LDLR and PCSK9 (table 3.5 (a) and 3.5 (b) respectively.

Table 3.5 (a): Showing the different drugs associated with LDLR gene.

S. No	Drug Name	Target	Mechanism of action	Pharmacological action
1.	**Porfimer**	**Low-density lipoprotein receptor**	**Cellular damage caused by porfimer is a consequence of the propagation of radical reactions.**	**Inhibitor**
2.	**Atorvastatin**	**3-hydroxy-3-methylglutaryl-coenzyme A reductase**	**Atorvastatin selectively and competitively inhibits the hepatic enzyme HMG-CoA reductase.**	**Inhibitor**
3.	**Lovastatin**	**3-hydroxy-3-methylglutaryl-coenzyme A reductase**	**Lovastatin is a prodrug that is activated in vivo via hydrolysis of the lactone ring to form the β-hydroxyacid.**	**Inhibitor**
4.	**Pravastatin**	**3-hydroxy-3-methylglutaryl-coenzyme A reductase**	**Its hydrolyzed lactone ring mimics the tetrahedral intermediate produced by the reductase allowing the agent to bind with a much greater affinity than its natural substrate.**	**Inhibitor**
5.	**Rosuvastatin**	**3-hydroxy-3-methylglutaryl-coenzyme A reductase**	**Decreased hepatic cholesterol concentrations stimulate the upregulation of hepatic low density lipoprotein (LDL) receptors**	**Inhibitor**

			which increases hepatic uptake of LDL.	
6.	**Atenolol**	**Beta-1 adrenergic receptor**	**atenolol competes with sympathomimetic neurotransmitters such as catecholamines for binding at beta(1)-adrenergic receptors in the heart and vascular smooth muscle, inhibiting sympathetic stimulation.**	**Antagonist**
7.	**Lomitapide**	**Microsomal triglyceride transfer protein large subunit**	**lomitapide inhibits microsomal triglyceride transfer protein (MTP), which prevents the formation of apolipoprotein B, and, thus, the formation of VLDL and chylomicrons as well.**	**Antagonist**
8.	**Mipomersen**	**mRNA of ApoB-100**	**Mipomersen binds to the mRNA that codes for apoB-100. This binding leads to double-stranded RNA, which is degraded by RNase H and prevents translation of the mRNA to form the apo B-100 protein.**	**Binder**
9.	**Ribavirin**	**Inosine-5'-monophosphate dehydrogenase 1**	**Ribavirin is reported to have several mechanism of actions that lead to inhibition of viral RNA and protein synthesis.**	**Inhibitor**

The above mentioned drugs play an important role in the treatment of various disorders associated with cholesterol metabolism. These drugs either decrease or increase the expression of LDLR gene in response to a particular disease.

Table 3.5 (b): Showing the different drugs associated with PCSK9 gene.

S.No	Name	Target	Mechanism of action	Pharmacological action
1.	Alirocumab	Proprotein convertase subtilisin/kexin9	preventing interaction of PCSK9 with low density lipoprotein receptors	Inhibitor
2.	Evolocumab	Proprotein convertase subtilisin/kexin9	The lumen of the endoplasmic reticulum, lomitapide inhibits (MTP), which prevents the formation of apolipoprotein B and VLDL	Inhibitor
3.	Lomitapide	Microsomal triglyceride transfer protein large subunit	The lumen of the endoplasmic reticulum, lomitapide inhibits (MTP), which prevents the formation of chylomicrons	Antagonist
4.	Estradiol	Estrogen receptor	Estradiol regulates gene transcription of messenger RNA. The mRNA produce specific proteins that express the effect of estradiol upon the target cell.	Agonist
5.	Methyltestosterone	Androgen receptor	Activation of the androgen receptor and by conversion to estradiol and activation of certain estrogen receptors.	Agonist
6.	Ritonavir	Pol polyprotein	Ritonavic inhibits the HIV viral proteinase enzyme	Inhibitor
7.	Simvastatin	Pol polyprotein	Ritonavic inhibits the HIV viral proteinase enzyme	Inhibitor
8.	Testosterone	Androgen receptor	activation of the androgen receptor (directly or as DHT), and by conversion to estradiol and activation of certain estrogen receptors	Agonist
9.	Atorvastatin Calcium	3-hydroxy methylglutaryl-coenzyme A reductase	competitive HMG-CoA reductase inhibitor	Inhibitor
10.	Antibodies	Estrogen receptor	Estradiol enters target cells	Agnoist

The above mentioned drugs play an important role in a treatment of various cholesterol metabolism

related defects. These drugs whether decrease or increase the expression of PCSK9 gene in response to a particular disease.

Distribution of SNPs over the LDLR and PCSK9 genes

It has been distinctly reported that single nucleotide polymorphism have significant role in disease susceptibility. In present study the targeted gene has following distribution of SNP's in coding and non-coding region of LDLR gene as shown in figure 3.3 (a).

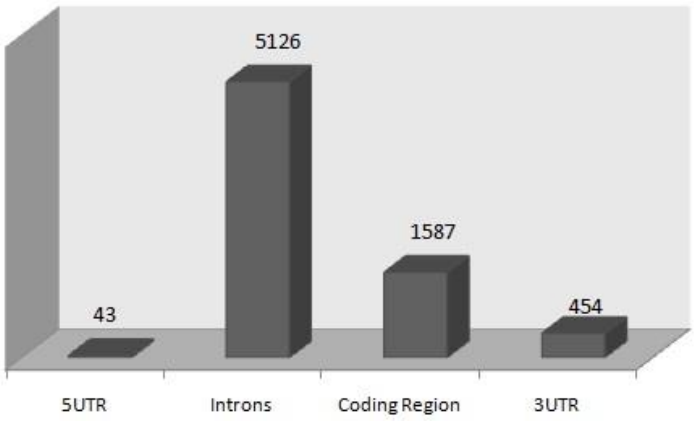

Figure 3.3 (a): SNP distribution in LDLR gene based on dbSNP

We have found that the number of SNPs present at 5'UTR , intronic region, coding region and at 3'UTR are 43, 5126, 1587, 454 of LDLR gene respectively.

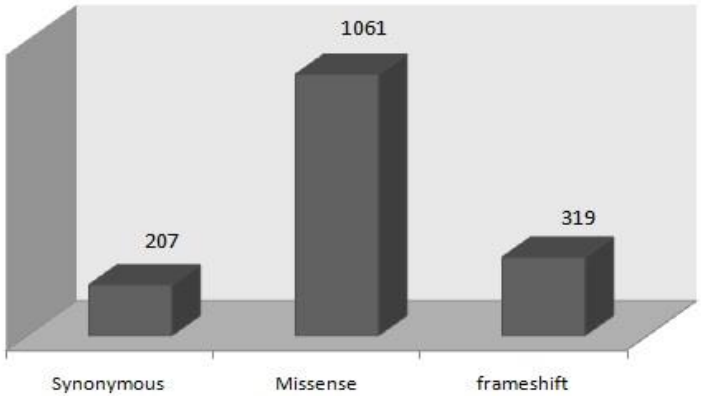

Figure 3.3 (b): SNP distribution in coding region of LDLR gene.

The targeted gene PCSK9 has following distribution of SNP's in coding and non-coding region as shown in figure 3.4 (a).

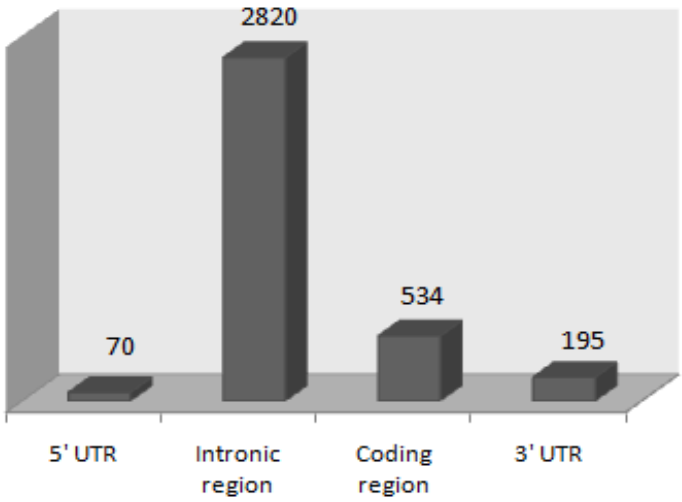

Figure 3.4 (a) : SNP distribution in PCSK9 gene based on dbSNP.

We have found that the number of SNPs present at 5'UTR, intronic region, coding region and at 3'UTR are 70, 2820, 534 and 195 of PCSK9 gene respectively.

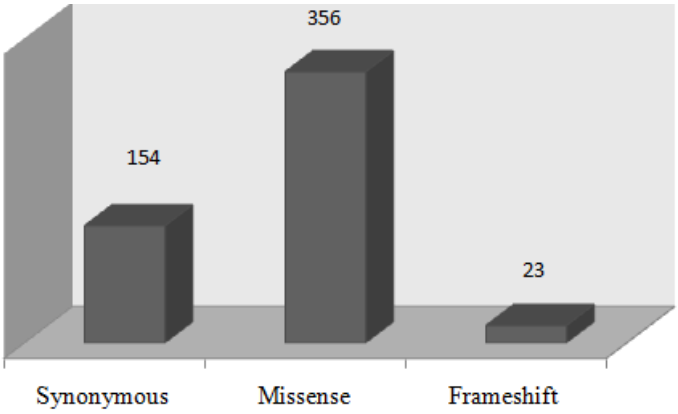

Figure 3.4 (b): SNP distribution in coding region of PCSK9 gene

Evaluation of deleterious SNP's of LDLR and PCSK9 genes by SIFT and I-Mutant

The SNP's of LDLR and PCSK9 gene were predicted to be damaging with the help of SIFT and I-Mutant were shown in table 3.4(a) and 3.4 (b) respectively.

Table 3.6 (a): List of nsSNPs of LDLR gene (Tolerated/Deleterious) that were analyzed by computational method I-Mutant 2.0 and SIFT.

S.No.	rs number	Base pair change	Type of mutation	Amino acid change	I-Mutant DDG (Kcal $mol^{-1)}$	SIFT Damaging SNP's
1.	rs121908041	G>C	Fwd.	C>S	-0.14*	Damaging
2.	rs121908025	T>G	Fwd.	W>G	-0.96*	Damaging
3.	rs121908026	C>T	Fwd.	S>C	-2.11*	Damaging
4.	rs121908039	G>A	Fwd.	C>Y	-0.87*	Damaging
5.	rs121908029	G>C	Fwd.	C>Y	-2.65*	Damaging
6.	rs121908040	G>T	Fwd.	C>F	-0.91*	Damaging
7.	rs121908038	T>A	Fwd.	L>H	-1.53*	Damaging
8.	rs13306505	C>T	Fwd.	A>V	-1.45*	Damaging
9.	rs28942085	A>G	Fwd.	Y>C	-0.70*	Damaging
10.	rs28942084	C>T	Fwd.	P>L	-0.33*	Damaging
11.	rs28942083	G>A	Fwd.	C>Y	-0.75*	Damaging
12.	rs121908036	G>C	Fwd.	D>H	-0.60*	Damaging
13.	rs28941776	G>A	Fwd.	G>D	-0.40*	Damaging
14.	rs2569548	C>A	Rev.	P>S	-0.97*	Damaging

***indicate the value of DDG in negative which means that there is decrease in stability of the protein.**

A total number of nsSNPs of LDLR gene were evaluated by SIFT and I-Mutant (SNP damage prediction tools). Out of these 577 (54.38%) of nsSNPs were predicted to affect the stability of the LDLR protein. SIFT predicted 16 (1.15 %) of nsSNPs as damaging. Further the total number of ns SNPs were predicted to be deleterious by all two prediction software were 14 (1.01%) as shown in table 3.8(a).

Table 3.6 (b): List of nsSNPs of PCSK9 gene (Tolerated/Deleterious) that were analyzed by computational method I-Mutant 2.0 and SIFT.

Sr.No	rs number	Base change	Type of mutation	Amino acid change	I - Mutant DDG(Kcal/mol)	SIFT Damaging SNPs
1.	rs72658890	A>T	Fwd	L>H	-0.57*	Damaging
2.	rs28942111	A>T	Fwd	S>R	-2.12*	Damaging
3.	rs72646508	C>T	Fwd	L>F	-0.56*	Damaging
4.	rs72646509	A>C	Fwd	P>T	-0.36*	Damaging
5.	rs72646510	A>C	Fwd	L>F	-0.10*	Damaging
6.	rs72646515	A>G	Fwd	D>N	-1.89*	Damaging
7.	rs72646517	G>A	Fwd	G>E	-1.61*	Damaging
8.	rs74646525	C>T	Fwd	P>L	-0.01*	Damaging
9.	rs7246530	C>A	Fwd	H>N	-0.57*	Damaging
10.	rs505151	A>G	Fwd	G>E	-1.31*	Damaging

*** indicate the value of DDG in negative which means that there is decrease stability of protein.**

A total number of nsSNPs of PCSK9 gene were evaluated by SIFT and I-Mutant (SNP damage prediction tools). Out of these 277 (77.82%) of nsSNPs were predicted to affect the stability of PCSK9 protein. SIFT predicted 12 (3.16%) of nsSNPs as damaging. Further the total number of nsSNPs predicted to be deleterious by all two prediction software were 10 (2.63%) as shown in table 3.8(b).

Conclusion

As we know that SNPs has an important role in determining the disease susceptibility. SNPs are important to understand because they can have profound effects on human health and they represent the vast majority of data that will explain that how the human genome controls the human disease condition such as cancer, autoimmune diseases, neurodegenerative diseases, familial hypercholesterolemia, CHD, atherosclerosis, diabetes and hypertension. The importance of SNPs comes from their ability to influence disease risk, drug efficacy and side-effects of drug. SNPs are probably the most important category of genetic changes influencing common diseases. And in terms of common diseases, 9 of the top 10 leading causes of death have a genetic component and thus most

likely one or more SNPs influence your risk. Knowledge of the position and effects of SNPs should therefore be considered in the development of better and more targeted forms of treatment. There is a great need for an effective and efficient method to filter out the pathogenic and deleterious SNPs from the available pool of variant data and to further explore the impact of those selected SNPs at the molecular level. Bioinformatics tool can be used in a cost efficient manner of prioritizing SNPs of likely functional importance, enabling an investigation of the structural basis of disease causing mutations likely to contribute of an individual's disease susceptibility. It is most important to note that the success of association studies always depends on how a research group chooses a set of SNPs to be investigated. With the help of insilico analysis, SNPs to be screened can be selected with more precisely and accurately and further the large number of samples can be used to identify the association of the selected SNP's at an acceptable level of significance. So the functional and structural impact of SNPs will not only be supportive but also facilitate the discrimination between true associations and false positives as reported recently [10].

Acknowledgements

We are thankful to the computer science department of Punjabi University, Patiala for providing us the high speed internet accessibility all the time during this work.

REFERENCES

1. Goldberg AC, Hopkins PN, Toth PP. Familial hypercholesterolemia: screening, diagnosis and management of pediatric and adult patients: clinical guidance from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. *J Clin Lipidol* 2011;5
2. Siddiqi K.H, HDL-cholesterol and cardiovascular disease: rethinking our approach. Current Opinion in Cardiology, 2015; 30 Issue 5: 536-542.
3. Thomas J Hudson , James C Engert. Genetic Variants Associated With Myocardial Infarction Risk Factors in Over 8000 Individuals From Five Ethnic Groups. *Circ Cardiovasc Genet* 2 (1), 16-25. 2009 Jan 23.
4. Chambers C.J., McGregor A, Jean-Marie J, Kooner S.J. Abnormalities of vascular endothelial function may contribute to increased coronary heart diseaserisk in UK Indian Asian. *Heart,* 1999;81:501-504.
5. Nordestgaard BG, Chapman MJ, Humphries SE, et al. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: Consensus Statement of the European Atherosclerosis Society. *Eur Heart J* 2013;34:3478-3490.
6. Garg A, and Simha, V. (2007) Update on dyslipidemia, J Clin Endocrinol Metab 92, 1581-1589.
7. Sullivan, D. (2007) Guidelines for the diagnosis and management of familial hypercholesterolaemia. *Heart Lung Circ* 16, 25-27.

8. Liu J, Parveen Abidi, Wei Chen , Fredric B. Kraemer, Hai Li. The medicinal plant goldenseal is a natural LDL-lowering agent with multiple bioactive components and new action mechanisms, *J Lipid Res*. 2006 Oct;47(10):2134-47.
9. Getz S.G. PCSK9 in South African variants of familial hypercholesterolemia . Journal of the American college of cardiology, 2014; Vol.63: 2374-35.
10. Abdul S. Bandar A. Suliman. Sayed H. A Comprehensive *In Silico* Analysis on the Structural and Functional Impact of SNPs in the Congenital Heart Defects Associated with NKX2-5 Gene—A Molecular Dynamic Simulation Approach. *Plos one*2016; 11(5): e0153999.
11. http://sift.jcvi.org/
12. http://folding.biofold.org/i-mutant/i-mutant2.0.html
13. http://www.genecards.org/Search/Keyword?queryString=ldlr
14. http://www.genecards.org/Search/Keyword?queryString=pcsk9