

Identify the Significant Hypostatic or Epistatic Effect Between the Various SNPs of LDLR gene through SNP-SNP Interaction and Haplotype Analysis - A Case Control Study

Kulwinder Kaur¹, Chandan Kumar Jha²

¹M.Sc. (Human Genetics), Department of Human Genetics, Punjabi University, Punjab, India

²Ph.D. (Human Genetics), Department of Molecular Diagnostic, Dr Lal PathLabs, Bangalore, India

Corresponding Author: Chandan Kumar Jha

DOI: <https://doi.org/10.52403/ijhsr.20220834>

ABSTRACT

The coronary artery disease (CAD) is a leading cause of the morbidity and mortality worldwide and has also become a major public health burden in India in recent decades. The *LDLR* gene mutations have been reported associated with familial CVD and are one of the main risk factors for developing coronary artery disease (CAD).

In the present study using the contingency chi-square test, statistically significant differences were observed between the CAD cases and matched healthy controls in the distribution of the genotypes of four out six SNPs investigated viz., rs688C>T (p=0.0031), rs1529729C>T (p=0.0001), rs5925G>A (p=0.006) and rs72658855C>T (p=0.0349). This suggested that these four SNPs are associated with CAD in the present material studied from Bangalore.

SNP-SNP interaction analysis revealed statistically significant hypostatic effect of rs688-rs2228671 combination, while rs688-rs72658855 combination showed statistically significant epistatic effect and the haplotype analysis demonstrated decreased (protective) risk of statistically significant association of *TATCC*, *CGTCC* and *TACCT* with CAD.

Further studies from different regions of the country are required to validate the present findings relating the association of genomic variants of *LDLR* gene with the susceptibility to CAD in people of India.

Keywords: CAD-coronary artery disease; LDLR-low-density lipoprotein receptor; SNP-SNP interaction; Haplotype analysis; OR-odds ratio; CI-confidence interval

INTRODUCTION

The coronary artery disease (CAD) is a leading cause of the morbidity and mortality worldwide and has also become a major public health burden in India in recent decades (Mukherjee, 1995, Prabhakaran *et al.*, 2016). The Registrar General of India reported that CHD led to 17% of the total deaths and 26% of adult deaths in 2001-2003, which increased to 23% of the total and 32% of the adult deaths in 2010-2013 (Gupta *et al.*, 2016).

Although many studies have reported the prevalence of CAD in India but there is little work available in literature on the role of the genetic factors in CAD. *LDLR* (low-density lipoprotein receptor) gene is one of the most accountable such factor in the development of the disease. *LDLR* is responsible for the binding and uptake of the plasma LDL particles and plays a critical role in maintaining cellular cholesterol homeostasis (Goldstein *et al.*, 1982). One thousand and seven hundred mutations in the gene have

been reported associated with familial CVD, which is one of the traditional risk factors for the disease. Recent GWAS (genome wide association studies) have highlighted that common variants at the *LDLR* locus are strongly associated with proatherogenic lipid profile and CAD status (Teslovich *et al.*, 2010 and Knowles *et al.*, 2018).

The *LDLR* gene mutations have been reported associated with familial CVD (Brown *et al.*, 1986) and are one of the main risk factors for developing coronary artery disease (CAD) (Chandan *et al.*, 2019). Limited studies are available from India in this field and it is therefore that the present investigation was undertaken to evaluate the association, if any, of the 6 SNPs of the *LDLR* gene with the coronary artery disease susceptibility in the population of Bangalore, Karnataka, South India by the help of SNPstats software (<http://bioinfo.iconcologia.net/snpstats/start.htm>) to identify the significant hypostatic or epistatic effect at the pair-wise SNP-SNP interactions between the various SNPs of *LDLR* gene and Haplotype analysis.

MATERIALS AND METHODS

This study was a hospital-based case control study. Subjects were collected from different hospitals. Informed written consent was obtained for all study subjects. The research study was approved by the Ethical Committee of Punjabi University.

Inclusion Criteria

The case control study included clinically-confirmed coronary artery disease cases. The diagnosis was confirmed using electrocardiogram (ECG), echocardiogram stress test, cardiac catheterization and angiogram heart scan. Out of 200 CAD patients, 180 were males and 20 were females.

Exclusion Criteria

The exclusion criteria involved patients with who had previously had coronary bypass surgery or percutaneous trans-luminal coronary angioplasty (PTCA) because of their treated coronary status.

Collection of Clinical History

Informed written consent was obtained from all CAD patients, as well as healthy controls. Both CAD patients as well as healthy controls were interviewed using a structured questionnaire regarding epidemiological/demographic data, past history, history of addiction, particularly smoking, family history of MI or CAD. The detailed laboratory and clinical data were collected to determine relevant clinical history.

Procedure of DNA Extraction and Genotyping

DNA was extracted from the blood samples by using the glass bead method (Chandan *et al.*, 2016). Genotyping of the studied SNPs was done using the standard laboratory protocols which involved the amplification of the DNA by either the Allele Specific PCR (AS-PCR) for SNPs rs688C>T, rs1529729C>T, rs5925G>A, rs72658855C>T and rs565675103C>A or the Amplification Refractory Mutation System PCR (ARMS-PCR) for SNP rs2228671C>T, followed by electrophoresis of the amplicons in each case. The results were read under UV light in UVP GelDoc-It Imaging system and recorded.

The AS-PCR is based on the use of sequence-specific PCR primers that allow amplification of the template DNA when the target allele is contained within the extracted DNA sample.

Analysis of snp - snp interaction between the SNPs of LDLR gene

The SNPstats software (<http://bioinfo.iconcologia.net/snpstats/start.htm>) was used to identify the significant hypostatic or epistatic effect at the pair-wise SNP-SNP interactions between the various SNPs of *LDLR* gene. To illustrate this analysis, a logistic regression model like odds ratio (OR) and 95% confidence intervals (CI) was fitted to the present genotype data to test for each model of hypostatic or epistasis effect. The p-value was calculated by the Bonferroni-correction

with a statistically significant threshold of $p \leq 0.05$.

Haplotype analysis

The haplotype analysis was carried out using the online programme SNPSTATs web tool for SNP analysis provided by Institut Català d'Oncologia and accessed through

<http://bioinfo.iconcologia.net/snpstats/start.htm>. When analyzing multiple SNPs of the same gene, as is the case in the present study, the programme will perform linkage disequilibrium statistics, haplotype frequency estimations and analysis of haplotype association with disease outcome.

Linkage disequilibrium (LD) analysis

The linkage disequilibrium (LD) is a situation when two alleles at two different loci are inherited together more frequently than that expected under the independent assortment. In other words, linkage disequilibrium describes the correlation between two alleles located on the same chromosome. During recombination sections of different copies of chromosome pairs are interchanged. The LD results indicate the alleles closely located to those causal of disease to be identified as associated with the disease. Linkage disequilibrium has been calculated using the online web-based SNP analysis SNPstats accessed through the link: http://bioinfo.iconcologia.net/en/SNPStats_web.

RESULT AND DISCUSSION

CAD is a multifactorial disease affected by both genes and the environment (Shirodkar and Marsden, 2011). Traditional risk factors, such as lipid-rich diet, advanced age, smoking, hypertension, diabetes mellitus and dyslipidemia have been reported to be associated with an increased risk of CAD (Musunuru and Kathiresan 2010).

In the present study using the contingency chi-square test, statistically significant differences were observed between the

CAD cases and matched healthy controls in the distribution of the genotypes of four out six SNPs investigated viz., rs688C>T ($p=0.0031$), rs1529729C>T ($p=0.0001$), rs5925G>A ($p=0.006$) and rs72658855C>T ($p=0.0349$). This suggested that these four SNPs are associated with CAD in the present material studied from Bangalore.

Using the odds ratio (OR) test, the increased risk of developing CAD in Bangalore city (Karnataka state) population was found to be associated with LDLR SNPs viz., rs688C>T (OR = 3.0, 95% CI = 1.43-6.2, $p=0.0037$), rs1529729C>T (OR = 2.36, 95% CI=1.3-4.29, $p=0.0047$), rs2228671C>T (OR=2.8, 95% CI=1.07-7.34, $p=0.034$) and rs72658855C>T (OR=1.66, 95% CI=1.07-2.58, $p=0.0215$). On the other hand, rs5925G>A (OR = 0.477, 95% CI=0.28-0.78, $p=0.003$) was found associated with a decreased susceptibility to the CAD in this study.

In the present study, the covariates correlation of genotype frequency of SNP with CAD showed non-significant value, except in the SNP rs688C>T, the family history of CAD showed significant value ($p=0.04$) and in rs2228671C>T, the LDL-cholesterol concentration showed significant value ($p=0.0047$).

SNP-SNP Interaction

Interaction between two or more SNPs (or genetic loci in general) is commonly referred to as epistasis (Cordell, 2002). Different epistatic effects have been observed in a variety of species, which are useful to define different genetic inheritance models of interaction between two SNPs.

Analysis of all possible 10 pair-wise SNP-SNP interactions, only rs688-rs2228671 and rs688-rs72658855 showed statistically significant association with CAD in population of Bangalore city (Karnataka state). In the homozygous rs2228671 CC genetic background, homozygous rs688 TT genotype conferred a reduced risk of CAD (OR= 0.24) compared to the wild-type rs688 CC genotype. And the heterozygous rs688 CT genotype combined with heterozygous

rs2228671 CT led to a potentially protective effect (OR = 0.23) on CAD compared to the wild-type rs688 CC genotype. And in the homozygous rs72658855 CC genetic background, the heterozygous rs688 CT genotypes conferred a higher risk of CAD (OR= 3.67) compared to the wild-type rs688 CC genotype.

The results presented in Table 3 show that SNP pair rs688-rs2228671 demonstrated statistically significant hypostatic effect

with OR=0.24 (95% CI= 0.11-0.52, p=0.03). On the other hand, SNP pair rs688-rs72658855 showed statistically significant epistatic effect with OR=3.67 (95% CI= 1.22-11.02, p = 0.0037) (Table 4). The remaining eight SNP-SNP interactions showed no hypostatic or epistatic effect as test of interaction was found to be statistically non-significant in each case (Tables 1-10).

Table 1. SNP-SNP interaction between LDLR gene SNPs rs688 and rs5925 in the present study.

rs688	rs5925								
	GG			GA			AA		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	10	5	1.00	15	25	3.33 (0.96-11.63)	3	6	4.0 (0.69-23.09)
CT	34	22	1.29 (0.39-4.30)	83	113	2.72 (0.90-8.26)	13	11	1.69 (0.44-6.47)
TT	10	3	0.60 (0.11-3.21)	26	14	1.08 (0.31-3.78)	6	1	0.33 (0.03-3.58)

Interaction p value: 0.73

OR = Odds Ratio, CI = Confidence interval.

Table 2. SNP-SNP interaction between LDLR gene SNPs rs688 and rs1529729 in the present study.

rs688	rs1529729								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	1	6	1.00	21	29	0.23 (0.03-2.06)	6	1	0.03 (0.00-0.55)
CT	12	31	0.43 (0.05-3.96)	102	110	0.18 (0.02-1.52)	16	5	0.05 (0.01-0.54)
TT	5	5	0.17 (0.01-1.94)	31	13	0.07 (0.01-0.64)	6	0	0.00

Interaction p value : 0.74

OR = Odds Ratio, CI = Confidence interval.

Table 3. SNP-SNP interaction between LDLR gene SNPs rs688 and rs2228671 in the present study.

rs688	rs2228671								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	22	36	1.00	4	0	0.00	2	0	0.00
CT	121	142	0.72 (0.40-1.29)	8	3	0.23 (0.05-0.96)	1	1	0.61 (0.04-10.27)
TT	41	16	0.24 (0.11-0.52)	1	1	0.61 (0.04-10.27)	0	1	---

Interaction p value : 0.03*

OR = Odds Ratio, CI = Confidence interval, *statistically significant

Table 4. SNP-SNP interaction between LDLR gene SNPs rs688 and rs72658855 in the present study.

rs688	rs72658855								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	11	6	1.00	15	28	3.42 (1.06-11.09)	2	2	1.83 (0.20-16.51)
CT	26	52	3.67 (1.22-11.02)	102	89	1.60 (0.57-4.50)	2	5	4.58 (0.67-31.20)
TT	11	11	1.83 (0.50-6.72)	31	7	0.41 (0.11-1.50)	0	0	---

Interaction p value : 0.0037*

OR = Odds Ratio, CI = Confidence interval, *statistically significant

Table 5. SNP-SNP interaction between LDLR gene SNPs rs5925 and rs1529729 in the present study.

rs5925	rs1529729								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
GG	6	7	1.00	37	21	0.49 (0.14-1.64)	11	2	0.16 (0.02-1.00)
GA	11	32	2.49 (0.69-9.04)	99	119	1.03 (0.34-3.17)	14	1	0.06 (0.01-0.61)
AA	1	3	2.57 (0.21-31.71)	18	12	0.57 (0.15-2.12)	3	3	0.86 (0.12-5.94)

Interaction p value : 0.14

OR = Odds Ratio, CI = Confidence interval.

Table 6. SNP-SNP interaction between LDLR gene SNPs rs5925 and rs2228671 in the present study.

rs5925	rs2228671								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
GG	52	30	1.00	2	0	0.00	0	0	---
GA	113	146	2.24 (1.34-3.74)	8	4	0.87 (0.24-3.12)	3	2	1.16 (0.18-7.31)
AA	19	18	1.64 (0.75-3.60)	3	0	0.00	0	0	---

Interaction p value : 0.33

OR = Odds Ratio, CI = Confidence interval.

Table 7. SNP-SNP interaction between LDLR gene SNPs rs5925 and rs72658855 in the present study.

rs5925	rs72658855								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
GG	11	7	1.00	42	22	0.82 (0.28-2.42)	1	1	1.57 (0.08-29.41)
GA	32	54	2.65 (0.93-7.53)	89	92	1.62 (0.60-4.38)	3	6	3.14 (0.59-16.85)
AA	5	8	2.51 (0.58-10.88)	17	10	0.92 (0.27-3.16)	0	0	---

Interaction p value : 0.84

OR = Odds Ratio, CI = Confidence interval.

Table 8. SNP-SNP interaction between LDLR gene SNPs rs1529729 and rs2228671 in the present study.

rs1529729	rs2228671								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	17	40	1.00	1	0	0.00	0	2	---
CT	143	148	0.44 (0.24-0.81)	8	4	0.21 (0.06-0.80)	3	0	0.00
TT	24	6	0.11 (0.04-0.31)	4	0	0.00	0	0	---

Interaction p value : 0.07

OR = Odds Ratio, CI = Confidence interval.

Table 9. SNP-SNP interaction between LDLR gene SNPs rs1529729 and rs72658855 in the present study.

rs1529729	rs72658855								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	5	12	1.00	12	27	0.94 (0.27-3.26)	1	3	1.25 (0.10-15.11)
CT	34	54	0.66 (0.21-2.04)	117	94	0.33 (0.11-0.98)	3	4	0.56 (0.09-3.44)
TT	9	3	0.14 (0.03-0.74)	19	3	0.07 (0.01-0.33)	0	0	---

Interaction p value : 0.85

OR = Odds Ratio, CI = Confidence interval.

Table 10. SNP-SNP interaction between LDLR gene SNPs rs2228671 and rs72658855 in the present study.

rs2228671	rs72658855								
	CC			CT			TT		
	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)	Cases	Controls	OR (95% CI)
CC	43	67	1.00	137	120	0.56 (0.36-0.89)	4	7	1.12 (0.31-4.07)
CT	4	1	0.16 (0.02-1.48)	9	3	0.21 (0.05-0.83)	0	0	---
TT	1	1	0.64 (0.04-10.53)	2	1	0.32 (0.03-3.65)	0	0	---

Interaction p value : 0.79

OR = Odds Ratio, CI = Confidence interval.

Haplotype analysis

In the present study the haplotype analysis was performed by using SNPstats software (<http://bioinfo.iconcologia.net/snpstats/start.htm>). For SNPs rs688C>T, rs5925G>A, rs1529729C>T, rs2228671C>T and rs72658855C>T, 15 major haplotypes viz., CGCCT, TATCC, TGTCC, CGCCC, TGCCC, CATCC, TACCC, CGTCC, CACCC, TACCT, TGCCT, TGTCT, CATCT, CACCT, and CGTCT were observed (Table 11). The pair-wise Linkage Disequilibrium (LD) matrix showed low LD with CAD in haplotype analysis as shown in Figure 1. The test results showed that 3 haplotypes viz., TATCC, CGTCC and TACCT, out of

15 considered, demonstrated statistically significant association with CAD.

Table 11. Haplotype association of LDLR gene SNPs with CAD.

Haplotype	Frequency	OR (95% CI)	p
CGCCT	0.2148	1.00	-
TATCC	0.2034	0.39 (0.16 - 0.99)	0.048*
TGTCC	0.0777	0.43 (0.14 - 1.35)	0.15
CGCCC	0.0736	1.75 (0.63 - 4.91)	0.29
TGCCC	0.0593	0.42 (0.14 - 1.26)	0.12
CATCC	0.0588	1.35 (0.30 - 6.18)	0.7
TACCC	0.0533	2.70 (0.37 - 19.74)	0.33
CGTCC	0.0509	0.17 (0.05 - 0.57)	0.0042*
CACCC	0.0349	1.35 (0.34 - 5.44)	0.67
TACCT	0.0296	0.14 (0.03 - 0.74)	0.021*
Haplotype	Frequency	OR (95% CI)	p
TGCCT	0.0292	0.28 (0.05 - 1.65)	0.16
TGTCT	0.0251	0.00 (-Inf - Inf)	1
CATCT	0.0222	0.44 (0.08 - 2.57)	0.36
CACCT	0.0193	0.53 (0.06 - 5.02)	0.58
CGTCT	0.008	0.00 (-Inf - Inf)	1

*Statistically significant.

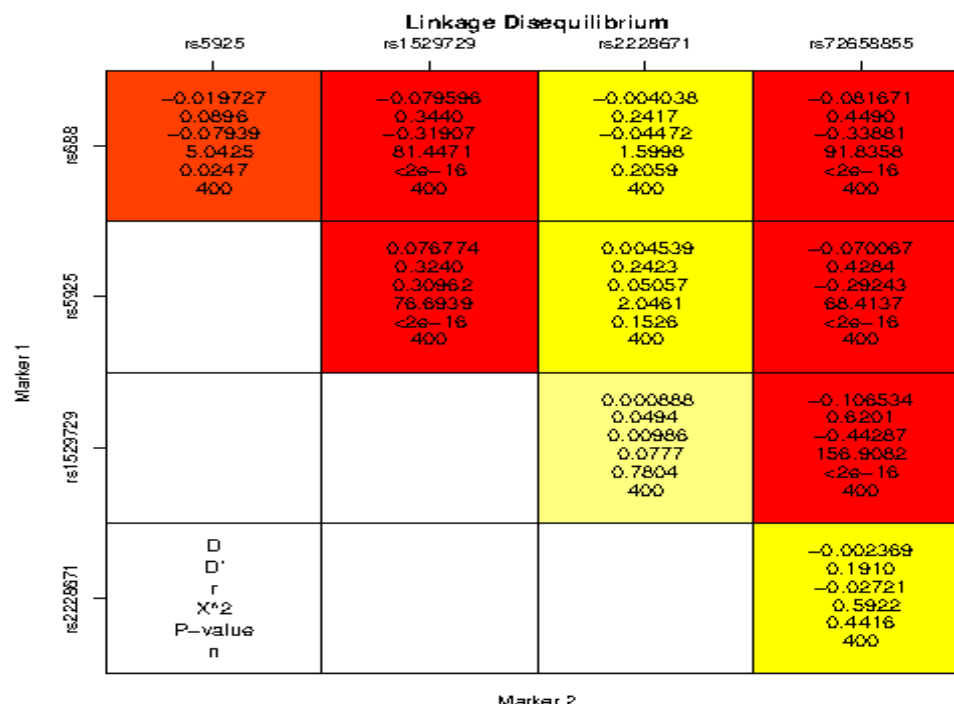


Figure 1. Linkage disequilibrium matrix of rs688, rs5925, rs1529729, rs2228671 and rs72658855 SNPs of LDLR gene.

It is believed that haplotype analysis can provide more information than single-locus analysis (Rafii, et al., 2002). The analysis of possible 15 haplotypes for the *LDLR* gene from different combinations of the five studied variable SNPs, 3 haplotypes viz., *TATCC*, *CGTCC* and *TACCT* were found linked with significant potentially protective effect on CAD risk (OR=0.39, p=0.048; OR=0.17, p=0.0042 and OR=0.14, p=0.021 respectively). These haplotypes possess protective effects which are consistently over represented in CAD patients relative to the healthy controls, suggesting a leading role of selected polymorphisms in CAD risk determination. The present haplotypes results were corroborated by the pair-wise linkage disequilibrium (LD) analysis using five variable SNPs of *LDLR* gene which showed low LD values with CAD (Figure 1).

CONCLUSION

The statistically significant OR values were found for SNPs rs688C>T (OR=3.0, 95% CI = 1.43-6.2, p = 0.0037) and rs1529729C>T (OR=2.36, 95% CI = 1.3-4.29, p= 0.0047) under the Co-dominant model of inheritance, while SNPs rs5925G>A (OR= 0.477, 95% CI = 0.28-0.78, p=0.003), rs2228671C>T (OR=2.8, 95% CI =1.07-7.34, p = 0.034) and rs72658855C>T (OR=1.66, 95% CI=1.07-2.58, p=0.0215) showed statistically significant AIC values under the Dominant model of inheritance.

SNP-SNP interaction analysis revealed statistically significant hypostatic effect of rs688-rs2228671 combination, while rs688-rs72658855 combination showed statistically significant epistatic effect.

The haplotype analysis demonstrated decreased (protective) risk of statistically significant association of *TATCC*, *CGTCC* and *TACCT* with CAD.

The results from this study are specific to the population of Bangalore (Karnataka), South India and may not be generalized for people of India. However, to validate the present findings relating the association of

genomic variants of *LDLR* gene with the susceptibility to CAD in people of India, further studies from different regions of the country are required.

Acknowledgement: None

Conflict of Interest: None

Source of Funding: None

Ethical Approval: Approved

REFERENCES

1. Brown, M.S.; Goldstein, J.L. 1986. A receptor-mediated pathway for cholesterol homeostasis. *Science*, 232, 34–47.
2. Cordell, H. J. 2002. Epistasis: what it means, what it doesn't mean and statistical methods to detect it in humans. *Human Molecular Genetics*, 11 (20): 2463-2468.
3. Goldstein, J.L.; Anderson, R.G.; Brown, M.S. 1982. Receptor-Mediated Endocytosis and the Cellular Uptake of Low Density Lipoprotein. *Ciba Found. Symp.* 92, 77–95.
4. Gupta, R., Mohan, I., & Narula, J. 2016. Trends in coronary heart disease epidemiology in India. *Annals of Global Health*, 82 (2): 307-315.
5. Jha CK, Chahal SMS, Khullar N, Banu S and Mir R. 2016. High-quality genomic DNA extraction from long term stored (LTS) whole blood samples using glass bead method. *Int J Health Sci Res.* 6(5):288-292.
6. Jha, C.K.; Mir, R.; Khullar, N.; Banu, S.; Chahal, S.M.S. 2018. LDLR rs688 TT genotype and T allele are associated with increased susceptibility to coronary artery disease—A case-control study. *J. Cardiovasc. Dev. Dis.* 5, 31.
7. Knowles, J.W.; Ashley, E.A. 2018. Cardiovascular disease: The rise of the genetic risk score. *PLoS Med.* 15, e1002546
8. Mukherjee, A. K. 1995. India's health-today and tomorrow. *Journal of the Indian Medical Association*, 93(8): 312-315.
9. Musunuru, K. and Kathiresan, S. 2010. Genetics of coronary artery disease. *Annual Review of Genomics and Human Genetics*, 11: 91-108.
10. Prabhakaran, D., Jeemon, P. and Roy, A. 2016. Cardiovascular diseases in India: current epidemiology and future directions. *Circulation*, 133 (16): 1605-1620.

11. Rafii, S., O'regan, P., Xinarianos, G., Azmy, I., Stephenson, T., Reed, M., Thacker, J. and Cox, A. 2002. A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. *Human Molecular Genetics*, 11 (12): 1433-1438.
12. Shirodkar, A. V. and Marsden, P. A. 2011. Epigenetics in cardiovascular disease. *Current Opinion in Cardiology*, 26 (3): 209-215.
13. Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J. and Johansen, C. T. 2010. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, 466 (7307): 707.

Web references

1. Haplotype analysis:
<http://bioinfo.iconcologia.net/snpstats/start.htm>.
2. Linkage Disequilibrium:
http://bioinfo.iconcologia.net/en/SNPStats_web.

How to cite this article: Kulwinder Kaur, Chandan Kumar Jha. Identify the significant hypostatic or epistatic effect between the various SNPs of LDLR gene through SNP-SNP interaction and haplotype analysis - a case control study. *Int J Health Sci Res.* 2022; 12(8):255-262.
DOI: <https://doi.org/10.52403/ijhsr.20220834>
