

# Identification and Analysis of Deleterious SNPs of Human GSR Gene: Their Effects on Structure and Functions of Associated Protein and Other Diseases

**Bharti Vyas**

Jamia Hamdard

**Ratul Bhowmik**

Jamia Hamdard

**Mymoona Akhter** (✉ [makhtar@jamiahamdard.ac.in](mailto:makhtar@jamiahamdard.ac.in))

Jamia Hamdard

**Farhan Jalees Ahmad**

Jamia Hamdard

---

## Research Article

**Keywords:** Glutathione-disulfide reductase (GSR), SNPs, human GSR gene, proteins, diseases, Glutathione reductase (GR)

**Posted Date:** November 11th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1043830/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Scientific Reports on March 31st, 2022. See the published version at <https://doi.org/10.1038/s41598-022-09295-6>.

# Abstract

Hereditary glutathione reductase deficiency, caused by mutations of the GSR gene, is an autosomal recessive disorder characterized by decreased glutathione disulfide (GSSG) reduction activity and increased thermal instability. This study implemented computational analysis to screen the most likely mutation that might be associated with hereditary glutathione reductase deficiency and other diseases. Using ten online computational tools, the study revealed 4 nsSNPs among the 17 nsSNPs identified as most deleterious and disease associated. Structural analyses and evolutionary confirmation study of native and mutant GSR proteins using the HOPE project and ConSrf. HOPE revealed more flexibility in the native GSR structure than in the mutant structure. The mutation in GSR might be responsible for changes in the structural conformation and function of the GSR protein and might also play a significant role in inducing hereditary glutathione reductase deficiency. LD and haplotype studies of the gene revealed that the identified variations rs2978663 and rs8190955 may be responsible for obstructive heart defects (OHDs) and hereditary anemia, respectively. These interethnic differences in the frequencies of SNPs and haplotypes might help explain the unpredictability that has been reported in association studies and can contribute to predicting the pharmacokinetics and pharmacodynamics of drugs that make use of GSR.

## Introduction

Glutathione-disulfide reductase (GSR) protein is also known as Glutathione reductase (GR) enzyme encoded by the GSR gene in humans, which is located on chromosome 8p21 and consists of 13 exons. Glutathione-disulfide reductase is a 522-amino acid protein that initiates the synthesis of mitochondrial and cytosolic GR. It is a member of the class-1 pyridine nucleotide-disulfide oxidoreductase family. Glutathione reductase is a homodimeric flavoprotein, a central enzyme of cellular antioxidant defense, and reduces oxidized glutathione disulfide (GSSG) to the sulfhydryl form GSH, which works as a cellular antioxidant<sup>1</sup>.

Glutathione reductase catalyzing the reaction ( $GSSG + NADPH + H^+ \rightarrow 2 (GSH) + NADP^+$ ) is an important enzyme in this cellular system. Because it maintains the ratio of GSH/GSSG, it is involved in many cellular functions, including the activation of dormant cells<sup>2,1</sup>. GSH plays a key role in two ways: maintaining function and preventing oxidative stress in red blood cells<sup>3</sup> and clearing electrophilic xenobiotics.

Hereditary glutathione reductase deficiency generally impairs cellular energy balance and increases the level of oxidative stress in red blood cells and is responsible for hereditary hemolytic anemia<sup>4,2</sup>.

Single gene mutations/defects have often been found to be involved in more than one disease, and through this study, the linkage and association of the GSR gene with various defects/disorders, such as obstructive heart defects (OHDs) and hereditary anaemia in humans has been investigated. This study employed a linkage disequilibrium (LD) and haplotype-based approach to examine GSR genes for linkage and association. This study defined a high-resolution haplotype structure for the GSR gene and examined the association of the individual SNPs and haplotypes with the GSR gene in Han Chinese (CHB) individuals from the International Hapmap Project.

Therefore, to study and reveal the pathogenic variants of the GSR gene and to predict the effect on the functionality of the GSR protein, the current study was performed. Further examination of the association of the

individual SNPs and haplotypes with GSR by subjecting the gene sequences along with nonsynonymous mutations to the various insilico methods was carried out.

## Result

### Mutation spread of GSR gene

The SNPs present in the GSR gene were retrieved using the ClinVar database, and 18 various mutations were found in the gene. Among the coding regions of mutations identified, 17 were related to missense mutations, and 1 was related to nonsense mutations. An overview of the complete methodological approaches is summarized in a schematic diagram (Fig. 1).

### Identification of deleterious missense mutation

Six pathogenicity prediction web servers were used to predict the deleterious effect of nsSNPs. Six variants were observed to be damaged by using the SIFT tool and were further subjected to cross check by using five different tools (WHESS.db submodule PolyPhen-2 server, Mutation testers, SNPs&Go, SNAP2 and Provean tools). Out of a total of 17 nsSNPs identified, only four were predicted to be the most deleterious nsSNPs in all computational algorithms. The results are presented in Table 1.

Table 1

The list of deleterious nsSNPs predicted by SIFT, SNAP2, POLYPHEN, SNP&GO, Mutation tester, Provean and SNAP2 tools.

		Amino acid substitution (nsSNP ID)			
Tool	Parameter	V289A	R233C	A199T	R153C
	nsSNPs	rs151187899	rs145851500	rs141805635	rs8190955
SIFT Prediction	Score	0	0	0	0
	Prediction	DAMAGING	DAMAGING	DAMAGING	DAMAGING
POLYPHEN Prediction	Score	0.882	1	1	0.932
	Prediction	PROBABLY DAMAGING	PROBABLY DAMAGING	PROBABLY DAMAGING	PROBABLY DAMAGING
SNP&GO	Score	0.837	0.801	0.873	0.524
	Prediction	Disease	Disease	Disease	Disease
Mutation tester	Score	1	0.999	0.999	0.999
	prediction	Disease causing	Disease causing	Disease causing	Disease causing
SNAP2	Score	76	54	31	27
	prediction	Effect	Effect	Effect	Effect
PROVEAN	Score	-6.982	-3.733	-5.830	-9.026
	prediction	Deleterious	Deleterious	Deleterious	Deleterious

## Identification of nsSNPs on the domains of proteins

InterPro, a domain identification tool, predicts the domains and active sites of a protein through the functional analysis of protein families. It predicted four functional domains of GSR, which are pyridine nucleotide-disulfide oxidoreductase domains and FAD/NAD(P)-binding domains (65-390), and demonstrated that all 4 nsSNPs identified are positioned on these domains (Figure S1).

## Prediction of stability of the mutant protein

Protein stability was analysed by using the CUPSAT, I-Mutant-2 and DUET servers. The results revealed that four variants destabilized the GSR residue, namely, (V289A) rs151187899, (R233C) rs145851500, (A199T) rs141805635, and (R153C) rs8190955. The results are presented in table 2

Table 2

List of damaged nsSNPs and affected amino acids their Prediction, DDG Values after mutation by using I-Mutant, CUPSAT and DUET

Gene Name	nsSNP ID	Amino Acid	CUPSAT		DUET		I-Mutant2
			Prediction	DDG Value (Kcal/mol)	Prediction	DDG Value (Kcal/mol)	DDG Value (Kcal/mol)
GSR	rs151187899	V289A	Destabilizing	-5.4	Destabilizing	-0.68	-0.68
	rs145851500	R233C	Destabilizing	-0.13	Destabilizing	-1.352	-1.52
	rs141805635	A199T	Destabilizing	-2.65	Destabilizing	-1.621	-1.21
	rs8190955	R153C	Destabilizing	-0.11	Destabilizing	-2.869	-1.869

## Structural effect of point mutation on human GSR protein:

The Project HOPE server revealed that the mutant residues of R153C and R233C are more hydrophobic than the wild-type residues, and these variations in size and hydrophobicity disrupt the H-bond interactions with the adjacent molecules due to loss of hydrophobic interactions in the core of the protein. The A199T mutant residue is larger, but the loss of hydrophobic interactions might disturb protein function by this mutation. The flexibility and rigidity of a protein structure are essential for exhibiting a specific function. In addition, V289A mutants exposed different properties and hence may significantly affect the function of the FAD/NAD-binding domain (Table S2).

## Evolutionary conservation analysis.

The evolutionary conservancy of amino acid residues of the native GSR was examined by the ConSurf web server. It identified structural and functional residues of the 4 high-risk nsSNPs of the GSR protein using evolutionary conservation and solvent accessibility. We observed that R153 and V289 are exposed and

functional, whereas residues A199 and R233 are buried and structural. All 2 of these residues are highly conserved (Table S1).

## Study of LD and Haplotype

LD (linkage disequilibrium) and haplotype were used to analyse the various genetic parameters of the GSR gene, and the genotype data of CHB (Han Chinese) were retrieved from the International Hapmap Project. Haplotype block reveals the combination of alleles at neighboring loci on the chromosome that may be transmitted together, and LD provides information on the measurement of the involvement of alleles (genetic marker) in a nonrandom mode.

Information about the predisposition of various diseases due to genetic variants has been evaluated. These parameters work as vital biomarkers for functional associations with a variety of diseases. The LD and haplotype study revealed an important block in the GSR gene, with three important SNPs having nonrandom associations, as represented in Fig. 2. Three critical SNPs out of five SNPs identified were rs3757918, rs8190955 and rs2978663 with minor allele frequencies T:C, C:T and A:G, respectively, with  $r^2 \geq 0.8$ , showing a high correlation between the loci (Table S3).

Furthermore, one of the haplotype blocks generated of the GSR gene involving 2 SNPs revealed that these variations may be responsible for obstructive heart defects ((rs2978663) OHDs) and hereditary anaemia (rs8190955) with different population frequencies. The TA haplotype was prominently found with a frequency of 0.533 in the studied population in comparison to haplotypes CG and CA.

## Discussion

Glutathione-disulfide reductase is a flavoprotein involved in the glutathione redox cycle maintaining proper function and preventing oxidative stress in RBCs. The literature reveals that deregulation of glutathione-disulfide reductase protein leads to activation of dormant cells and deregulation of the cell cycle. Changes in the GSR protein during the glutathione antioxidant defence system play a vital role in executing its function, but any nsSNPs in the GSR gene lead to aberrant conformations, which in turn leads to glutathione reductase deficiency. Therefore, it becomes necessary to identify the effects of deleterious nsSNPs of GSR and their association with various diseases.

This study aimed to determine the most deleterious nsSNPs and their effects on the function of GSR proteins, characterize the haplotype structure of these genes and investigate markers associated with disease. Among the 17 molecular consequences found in the ClinVar database, using six insilico SNP prediction tools, SIFT, Mutation Tester, Polyphen-2, SNP&GO, PROVEAN, SNAP2, and PredictSNP 4, significantly deleterious nsSNPs were finally screened out. The selection was made on the basis of prediction scores generated by these tools. The study tried to find the disease relevance effect for these mutations, as several of these assorted 4 mutations were highly pathogenic. These 4 nsSNPs revealed different domains of the protein, where two nsSNPs were positioned in the pyridine nucleotide-disulfide oxidoreductase domain that interacts with the GSR association domain<sup>21</sup>. One nsSNP was located in the FAD/NAD-binding domain, which serves as the FAD/NAD-binding domain involved in oxidative metabolism of a variety of hydrocarbons (rubredoxin reductase, putidaredoxin reductase, terpredoxin reductase, ferredoxin-NAD<sup>+</sup> reductase components of benzene 1,2-dioxygenase, toluene

1,2-dioxygenase, chlorobenzene dioxygenase, biphenyl dioxygenase), NADH oxidase and NADH peroxidase<sup>1, 2, 3</sup>. The fourth nsSNP was present in the mitochondrial apoptosis-inducing factor, the C-terminal domain, which is crucial for cell apoptosis<sup>22</sup>.

The stability of the protein structure is crucial for the proper function of the protein. Alternations in the stability of proteins may cause misfolding and degradation of proteins. Therefore, to study the structural and functional activity of proteins, protein stability studies were carried out. By using the CUPSAT and DUET servers, four deleterious nsSNPs that may affect the stability of the GSR protein were identified. Furthermore, evolutionary conservancy in the GSR protein sequence is vital to determine whether a mutation has any negative effect on the host. Using the ConSurf server, we observed that highly deleterious nsSNPs with high conservation scores were located in highly conserved regions, therefore increasing the risk of hereditary anaemia by altering the GSR protein sequence.

The Project Hope server revealed that these four highly risky nsSNPs negatively affect the structure of the GSR protein, among which 2 nsSNPs were structural and 2 nsSNPs were functional residues according to ConSurf.

Linkage studies identified the 8p21 region as a susceptibility locus for obstructive heart defects (OHDs), hereditary anaemia. The LD pattern and haplotype structure for GSR in Han Chinese was characterized, and it revealed that there is LD across the GSR locus with little recombination. The rs3757918, rs8190955 and rs2978663 markers are loci enclosing a small part of the gene. Linkage disequilibrium has been reported between the common polymorphism found on GSR at positions 30619688 and 30627495. The analysis revealed that of the five nsSNPs identified, only three nsSNPs occurred and were linked in Han Chinese individuals. The results also indicated that only MAF (minor allele frequency) values of 0.467, 0.011 and 0.422 showed relatively strong linkage disequilibrium. The genotype of rs2978663 with the GSR gene increased the risk of occurrence of obstructive heart defects (OHDs). Right-sided and left-sided obstructive heart defects (OHDs) are subtypes of congenital heart defects in which the heart valves, arteries, or veins are abnormally narrow or blocked. Previous studies have suggested that the development of OHDs involves a complex interplay between genetic variants and maternal factors. Using data from 569 OHD case families and 1,644 control families enrolled in the National Birth Defects Prevention Study (NBDPS) between 1997 and 2008, we conducted an analysis to investigate the genetic effects of 877 single nucleotide polymorphisms (SNPs) in 60 candidate genes associated with the risk of OHDs and their interactions with maternal use of folic acid supplements and prepregnancy obesity. Applying log-linear models based on the hybrid design, we identified a SNP in the methylenetetrahydrofolate reductase (MTHFR) gene (C677T polymorphism) with a main genetic effect on the occurrence of OHDs. In addition, multiple SNPs in betaine-homocysteine methyltransferase (BHMT and BHMT2) were also identified to be associated with the occurrence of OHDs through significant main infant genetic effects and interaction effects with maternal use of folic acid supplements. We also identified multiple SNPs in glutamate-cysteine ligase, catalytic subunit (GCLC) and DNA (cytosine-5-)-methyltransferase 3 beta (DNMT3B) that were associated with an elevated risk of OHDs among obese women. Our findings suggested that the risk of OHDs was closely related to a combined effect of variations in genes in the folate, homocysteine, or glutathione/transsulfuration pathways, maternal use of folic acid supplements and prepregnancy. Obstructive heart defects associated with candidate genes, maternal obesity, and folic acid supplementation resolution in prepregnancy obese patients and maternal genotypes of SNPs in the GSR gene were associated with an increased risk of OHDs<sup>23</sup>. The genetic variant of GSR (rs8190955) was also found to be significantly associated

with anaemia. This study demonstrates a potential connection between anaemia and oxidative stress, which could accelerate the production of ROS in addition to reducing the ability of the antioxidant defence system caused by SNPs of enzymes. LD and haplotype data should be useful in drug development and in understanding the genetic associations of GSR with adverse drug effects.

## Method

Hereditary red blood cell enzymopathies of GSR gene-related information were collected from the database Online Mendelian Inheritance In Man (OMIM)<sup>5</sup> and other reported literature. The dataset of chromosome number and position of the GSR gene in the human genome was collected from the ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>) only for missense variants<sup>5</sup>. Missense variants were chosen for further analysis because of their higher impact on the structure and function of proteins.

### Identification of deleterious missense variants

Prediction of the deleterious effect of missense variants was performed by using seven different tools: Sorting Intolerant from Tolerant (SIFT)<sup>6</sup> (<http://sift.jcvi.org>), Polymorphism Phenotypingv2 (PolyPhen-2)<sup>7</sup> (<http://genetics.bwh.harvard.edu/pph2/>), SNP-GO<sup>8</sup> (<https://snps-and-go.biocomp.unibo.it/snps-and-go/>), SNAP2<sup>9</sup>, Provean<sup>10</sup>, predictor of human deleterious single nucleotide Polymorphisms (PhD-SNP) and MUTATION TASTER<sup>11</sup> (<http://www.mutationtaster.org/>) tools. The SIFT tool is a sequence homology-based tool; if the score is equal to or less than 0.05, the nsSNPs are considered deleterious nsSNPs. The WHES.db module of the PolyPhen-2 server is a sequence and structure evolutionary conservation based on classifying the damaging effect of amino acid substitution; if the score lies between 0.801–1.00, then the nsSNPs are considered probably damaging. The PROVEAN server provides a pairwise sequence alignment (PSA) score and identifies nonsynonymous variants. Single nucleotide polymorphisms & Gene Ontology (SNPs&GO) and predictors of human deleterious single nucleotide polymorphisms (PhD-SNPs) are both support vector machine (SVM)-based tools used to predict evolutionary information, protein sequences and functions if the given mutation can be classified as disease related. Furthermore, mutation tester servers were used to evaluate the DNA sequence variants for disease-causing potential. Mutation tester scores that ranged from <0.5 were considered disease-causing.

### Identification of nsSNPs on the domains of GSR

The InterPro<sup>12</sup> (<https://www.ebi.ac.uk/interpro/>) tool was used to identify the location of point mutations on the domains of glutathione-disulfide reductase protein, which can recognize motifs, active sites and domains of a protein.

### Predicting mutation effect on protein stability

Protein stability analysis of mutant proteins was performed using three different tools: CUPSAT, the DUET server and I-Mutant2<sup>13</sup>. CUPSAT was used to assess the effect of protein stability due to point mutation<sup>14</sup>. This tool uses structural environment-specific atom potential and torsion angle potentials to predict  $\Delta\Delta G$  and the difference in the free energy between native and mutant proteins. Protein stability was also studied by using the DUET server<sup>15</sup> and the I-Mutant 2.0 tool, which supports the vector machine SVM-based tool that predicts the

change in protein stability upon single point mutation. The results obtained are in the form of stability of protein Gibbs free energy in the form of DDG values.

### **Analysing of GSR protein evolutionary conservation.**

To understand the evolutionary conservation of the amino acids in the protein sequence, ConSurf<sup>16</sup> (<https://consurf.tau.ac.il>) was used to analyse the phylogenetic relationships between homologous sequences. Considered those nsSNPs of GSR that were found to be conserved for further analyses.

### **Prediction of the structural effect of nsSNPs on the human GSR protein.**

To identify the effect of the nsSNPs on the structure of the protein, HOPE61 (<https://www3.cmbi.umcn.nl/hope>) was used. HOPE is a web server that identifies the structural effects of point mutations in a protein sequence<sup>17</sup>. P00390 (UniProt-Accession Code of GSR) and the 4 SNPs were used individually as the input.

### **LD and Haplotype Block analysis**

LD plays a key role in mapping complex disease or trait-associated genes. Haplotype block provides information on patterns of genetic variation that are associated with health and disease, and it can be used to examine stretches of DNA near the SNP cluster to identify the gene or genes responsible for causing the disease. Linkage disequilibrium (LD) is used for the study of population genetics for the nonrandom association of alleles at different loci [18,19]. The Haploview tool[20] from the MIT/Harvard Broad Institute was used to study the genotype data for quantitative genetic parameters such as LD, and haplotype block data of Han Chinese (CHB) were retrieved from the International Hapmap Project. The data were visualized and analysed for any linkages and generation of LD and haplotype blocks.

## **Conclusion**

GSR is a central enzyme in cellular antioxidant defence. The study of the functional SNPs of GSR provided significant insight into the deleterious effects of the nsSNPs identified in protein stability and cell functions. Two of the nsSNPs (rs2978663 and rs8190955) identified were also found to be associated with obstructive heart defects (OHDs) and hereditary anemia. It can be concluded that in this study LD and haplotype study has unrevealed the relation of GSR with hereditary anemia and OHD and that understating the oxidative stress pathways at molecular levels may be helpful in developing new interventions for the disease.

## **Declarations**

### **Acknowledgement:**

The authors are thankful to the Department of Biotechnology, Ministry of Science & Technology, Government of India for the Bioinformatics infrastructure facility at Jamia Hamdard under the BTISNet program.

### **Funding:**

One of the authors, Ms. Bharti Vyas, acknowledges the fellowship obtained from the Indian Council of Medical Research (ICMR), New Delhi.

## Conflict of interest:

None of the authors have no conflicts of interest to declare

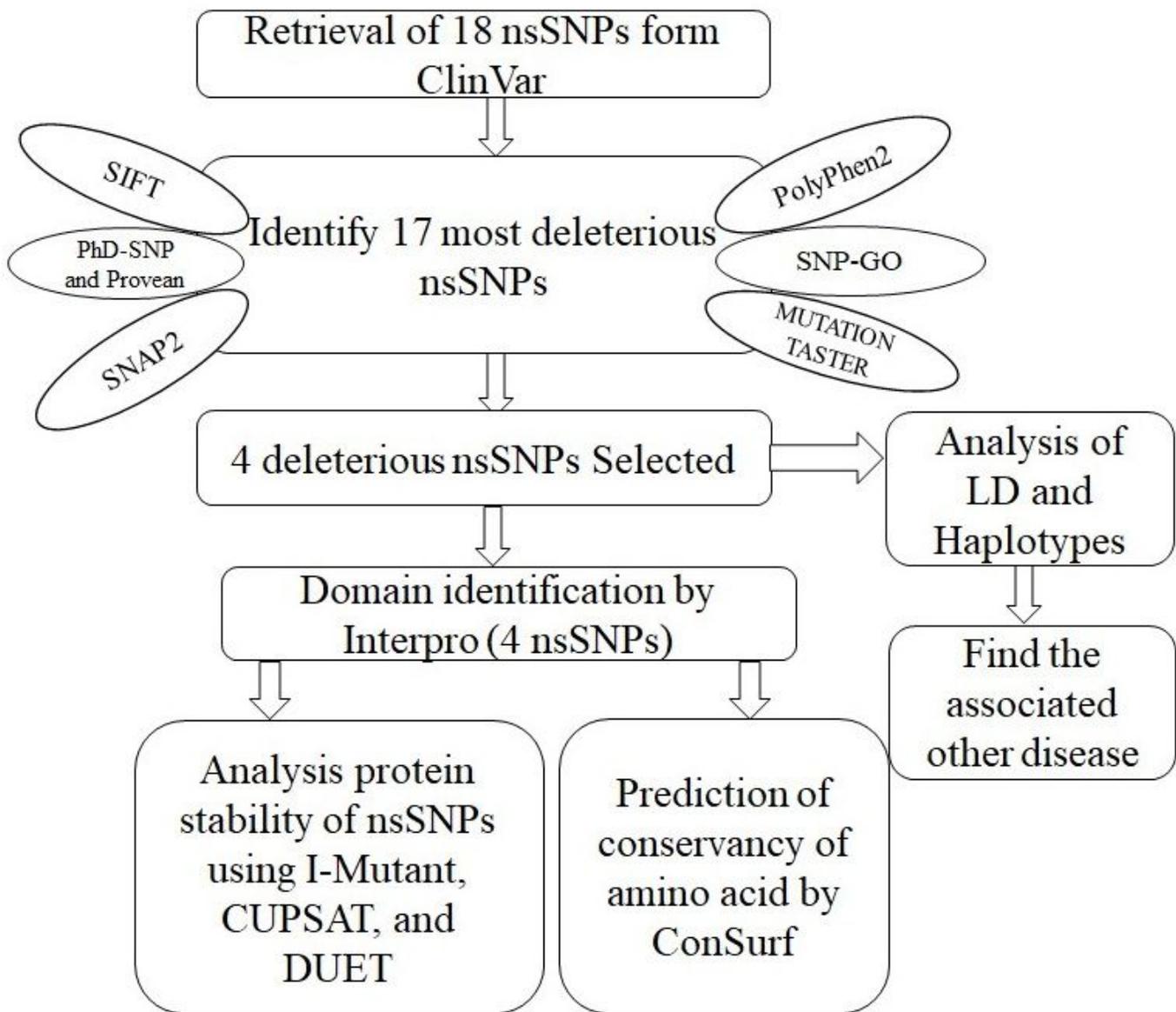
## References

1. Pai EF, Schulz GE. The catalytic mechanism of glutathione reductase as derived from x-ray diffraction analyses of reaction intermediates. *J Biol Chem.* 1983;258(3):1752-1757. doi:10.1016/S0021-9258(18)33050-3
2. Kamerbeek NM, van Zwieten R, de Boer M, et al. Molecular basis of glutathione reductase deficiency in human blood cells. *Blood.* 2007;109(8):3560-3566. doi:10.1182/blood-2006-08-042531
3. Deponte M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim Biophys Acta.* 2013;1830(5):3217-3266. doi:10.1016/j.bbagen.2012.09.018
4. Roos D, Weening R, Voetman A, et al. Protection of phagocytic leukocytes by endogenous glutathione: studies in a family with glutathione reductase deficiency. *Blood.* 1979;53(5):851-866. doi:10.1182/blood.V53.5.851.851
5. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018;46(D1):D1062-D1067. doi:10.1093/nar/gkx1153
6. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 2012;40(Web Server issue):W452-457. doi:10.1093/nar/gks539
7. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;Chapter 7:Unit7.20. doi:10.1002/0471142905.hg0720s76
8. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics.* 2013;14 Suppl 3:S6. doi:10.1186/1471-2164-14-S3-S6
9. Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics.* 2015;16(S8):S1. doi:10.1186/1471-2164-16-S8-S1
10. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinforma Oxf Engl.* 2015;31(16):2745-2747. doi:10.1093/bioinformatics/btv195
11. Steinhaus R, Proft S, Schuelke M, Cooper DN, Schwarz JM, Seelow D. MutationTaster2021. *Nucleic Acids Res.* 2021;49(W1):W446-W451. doi:10.1093/nar/gkab266
12. Apweiler R, Attwood TK, Bairoch A, et al. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* 2001;29(1):37-40. doi:10.1093/nar/29.1.37
13. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 2005;33(Web Server issue):W306-310. doi:10.1093/nar/gki375
14. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. *Nucleic Acids Res.* 2006;34(Web Server):W239-W242. doi:10.1093/nar/gkl190
15. Pires DEV, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* 2014;42(Web Server issue):W314-319.

doi:10.1093/nar/gku411

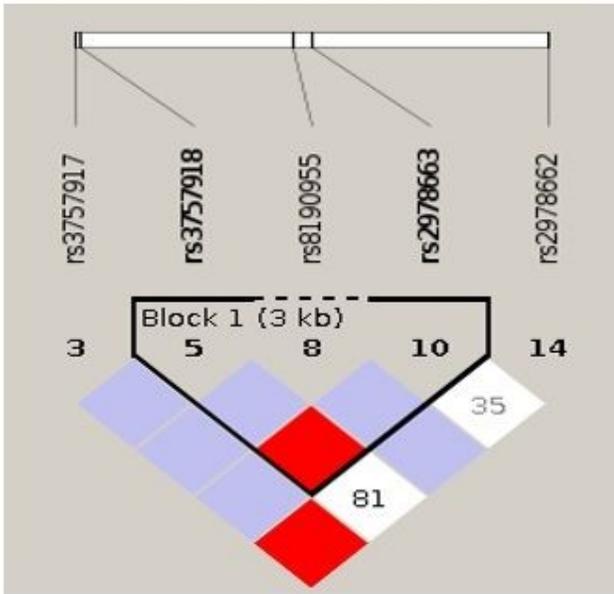
16. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* 2016;44(W1):W344-W350. doi:10.1093/nar/gkw408
17. Venselaar H, Te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics.* 2010;11:548. doi:10.1186/1471-2105-11-548
18. Blomhoff A, Olsson M, Johansson S, et al. Linkage disequilibrium and haplotype blocks in the MHC vary in an HLA haplotype specific manner assessed mainly by DRB1\*03 and DRB1\*04 haplotypes. *Genes Immun.* 2006;7(2):130-140. doi:10.1038/sj.gene.6364272
19. Long JR, Zhao LJ, Liu PY, et al. Patterns of linkage disequilibrium and haplotype distribution in disease candidate genes. *BMC Genet.* 2004;5:11. doi:10.1186/1471-2156-5-11
20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21(2):263-265. doi:10.1093/bioinformatics/bth457
21. Wang T, Xie X, Liu H, et al. Pyridine nucleotide-disulphide oxidoreductase domain 2 (PYROXD2): Role in mitochondrial function. *Mitochondrion.* 2019;47:114-124. doi:10.1016/j.mito.2019.05.007
22. Daugas E, Nochy D, Ravagnan L, et al. Apoptosis-inducing factor (AIF): a ubiquitous mitochondrial oxidoreductase involved in apoptosis. *FEBS Lett.* 2000;476(3):118-123. doi:10.1016/s0014-5793(00)01731-2
23. Xinyu Tang, Mario A. Cleves, Todd G. Nick, Ming Li, Stewart L. MacLeod, Stephen W. Erickson, Jingyun Li, Gary M. Shaw, Bridget S. Mosley, Charlotte A. Hobbs, and the National Birth Defects Prevention Study. Obstructive Heart Defects Associated with Candidate Genes, Maternal Obesity, and Folic Acid Supplementation.

## Figures



**Figure 1**

Schematic diagram summarizing the study.



1.



2.

Figure 2

1. LD 2. Haplotype block

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [paper3supplimentaryfiletable1.docx](#)