ISSN: 0976-3031

Available Online at http://www.recentscientific.com

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research
Vol. 8, Issue, 6, pp. 17977-17984, June, 2017

International Journal of Recent Scientific Research

DOI: 10.24327/IJRSR

Research Article

IN SILICO ANALYSIS OF HUMAN *GSDMA* AND *GSDMD* GENES FOR FUNCTIONAL AND STRUCTURAL IMPACT OF NON-SYNONYMOUS SNPs

Praveen P. Balgir* and Suman Rani

Department of Biotechnology, Punjabi University, Patiala, India

ARTICLE INFO

Article History:

Received 15th March, 2017
Received in revised form 25th April, 2017
Accepted 28th May, 2017
Published online 28th June, 2017

Key Words:

GSDMA, *GSDMD*, *in silico*, I-Mutant, MuPro, MutPred, nsSNPs, PolyPhen2, PROVEAN, SIFT, SNP& GO

ABSTRACT

Non-Synonymous Single Nucleotide Polymorphism (nsSNPs) are the main cause of defects in Genotypes and are critical for the prediction of genetic basis of various diseases. The major issue in analysis of variation at genetic level is to differentiate between mutation that can influence gene function from those that are neutral. The present study employed multiple *in silico* tools to predict nsSNPs of *GSDMA* and *GSDMD* genes with functional implications, before proceeding to study them at population level. Seven different tools such as SIFT, PolyPhen2, PROVEAN, SNP & GO, MuPro, I-Mutant and MutPred were used for *insilico* analysis. 10 nsSNP (missense) of *GSDMA* and 11 nsSNPs of *GSDMD* were analyzed using above softwares. By combining results of different methods, 2 nsSNPs of *GSDMA* namely rs191833662 (T2I) and rs115509258 (G200D) and 1 ns SNP of *GSDMD* the rs62000416 (L186M) were predicted to be deleterious or disease related by all softwares. The study is the first approach for *insilico* analysis of polymorphism in *GSDMA* and *GSDMD* genes that will be useful for further population and functional analysis.

INTRODUCTION

Point mutation or single nucleotide changes in genome constitute the heritable changes in the genome of individuals. Such variations with allelic frequency ≥1% are said to be polymorphic amongst human populations (Dabhi and Mistry, 2014) and termed Single Nucleotide Polymorphisms (SNPs). Out of all types of SNPs, the non-synonymous SNPs (nsSNP) also known as missense mutation are important as they are result in variation in coding regions of the protein, resulting in change in amino acid sequence, leading to diversity of encoded human proteins. These may affect gene regulation, alter transcriptional factor binding and also affect protein function. In human Genome about 2% of nsSNPs are reported to be the underlying cause for differential expression of traits, variation in drug responses and common diseases such as diabetes, hypertension, asthma, cancer, Sickle cell anemia and many more (Hassan *et al*., 2016).

Database of SNP (dbSNP) is one of the most extensive database that serves the public as freely available stock of genetic variations. 1000 Genome Project also showed that most of the human genetic level variations are expressed as SNPs.

To identify the SNPs responsible for phenotypic changes, in silico analysis using various bioinformatic tools and databases allows for selecting SNPs according to their structural and functional importance. This can help in finding the SNPs associated with diseases and development of new disease biomarkers that can be further used for drug discovery (Barroso *et al*., 1999; Chasman and Adams, 2001; Lander, 1996; Smith *et al*., 1994).

For present study, two human genes belonging to Gasdermin super family namely GasderminA (*GSDMA*) and Gasdermin D (*GSDMD*) have been analyzed for presence of functional SNPs using various bioinformatic tools.

Gasdermin family genes are expressed in various epithelial tissues from skin to gastrointestinal tract (Tamura *et al*., 2007). *GSDMA* is expressed in skin and stomach (Saeki *et al* 2000) whereas *GSDMD* is expressed in stomach and esophagus. Differentiated expression pattern of Gasdermin family genes in human cancer cell lines reveal some of their cellular functions, however their function is still poorly understood. Gene expression of *GSDMA* and *GSDMD* was found to be suppressed in gastric cancer tissue specimens (Saeki *et al*., 2009, Komiyama *et al*., 2010).

MATERIALS AND METHODS

Polymorphic data for identification of functional nsSNP of human *GSDMA* and *GSDMD* and information such as protein accession number, SNP ID were obtained from NCBI dbSNP

Corresponding author:* **Praveen P. Balgir
Department of Biotechnology, Punjabi University, Patiala, India

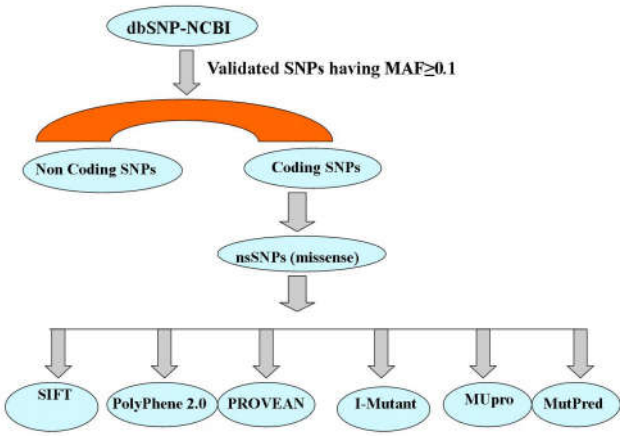

Fig 1 Schematic representation of computational approach employed for *in silico* analysis of Single nucleotide polymorphisms (SNPs) of *GSDMA* and *GSDMD* genes.

(http://www.ncbi.nlm.nih.gov/snp/), Swiss-Prot databases (http://expasy.org/), UniProt database (http://www.uni prot. org),1000 genome (http://www. 1000genome.org/) and the ensemble genome browser (http://www.ensembl.org/index.html).

Non-Synonymous functional SNP analysis

Structural and functional context of nsSNP were predicted using various *in silico* tools such as SIFT, PolyPhen2, I-Mutant 3.0, PROVEAN, SNP&GO, MutPred, MuPro. nsSNPs predicted to be deleterious by these *in silico* algorithms were categorized as high risk nsSNPs.

Prediction of functional nsSNP using SIFT (http://sift.jcvi.org/)

SIFT (Sorting Intolerant from Tolerant) was used to analyze the effect of nsSNP on protein function (Ng and Henikoff, 2003). SIFT depicts the deleterious or non-tolerated SNPs on the basis of conserved amino acid residues in protein. SIFT prediction gives tolerance index (TI) score ranging from 0.0 to 0.1 which gives the normalized probability whether change in amino acid due to SNP is tolerated or not. A nsSNP with TI score of ≤0.05 are considered structurally damaging and tolerated if score is >0.05.

Prediction of structural and functional impact of nsSNPs using PolyPhen2 (http://genetics.bwh.harvard.edu/pp2)

PolyPhen2 (Polymorphism Phenotyping2) gives amino acid substitution impact on structure and function of protein while using physical and comparative approaches. PolyPhen2 calculates the position-specific independent count (PSIC) score for every variant, giving difference between different variants. The outcomes of the PolyPhen2 is in the form of probably damaging, possibly damaging or benign with score ranging from 0 to 1(Ramensky *et al*., 2002).

Prediction of functional nsSNP using PROVEAN (http://provean.jcvi.org/index.php)

This algorithm was used to predict the effect of amino acid substitution on protein's biological function. This tool allows the separation between deleterious and neutral amino acids based on threshold. The score <-2.5 depicts that change is deleterious and score having threshold >-2.5 depicts that change is neutral (Manickam *et al*., 2014).

Prediction of disease causing nsSNP using SNP&GO (http://snps-and-go.biocomp.unibo.it/snps-and-go/)

SNP & GO help to predict that a single amino acid substitution causes disease condition in humans with scoring efficiency 82%. It is based on Support Vector Machine (SVM). The reliability index (RI) with value >5 depicts the disease relatedness of the mutation (Calabrese *et al*., 2009).

Prediction of functional nsSNP using I-Mutant (Version 3.0) (http://gpcr2.biocomp.unibo.it/cgp/predictor/I-Mutant3.0/I-Mutant3.0.cgi)

I-Mutant is a suite of SVM based method that predicts the change in protein stability upon single amino acid variation. It gives the free energy change value (DDG) by calculating the unfolding Gibbs free energy value for the wild type protein and subtracting it from that of mutant protein (DDG= G Mutant-G Wild type) (Capriotti *et al*., 2006).

MutPred (http://mutpred.mutdb.org/)

MutPred is a machine learning based software and adventitious server that predicts the impact on specific features of protein structure and function, helps in experimental studies of phenotype changing variants (Li *et al*., 2009). MutPred is also used to predict whether a variant is disease associated or neutral. It is based on SIFT, gain/loss of 14 different structural and functional properties of protein. It gives the general score (g) and top 5 property scores (p).

MUpro (http://www.igb.uci.edu/servers/servers.html)

MUpro uses another SVM algorithm that predicts the protein stability changes arising due to single amino acid substitutions (Cheng *et al*., 2006). This prediction is used for better understanding of protein structure.

RESULTS

NCBI dbSNP is the most substantial database for SNPs. It contains all type of SNPs in genome whether they are validated or not. As per this NCBI dbSNP, *GSDMA* gene have 1373 SNPs, of which 119 are validated by 1000 genome and having MAF≥0.10. Those that are not validated and having MAF<0.10 were not considered for study. Out of these 119 SNPs, 94 fall in intronic regions, 9 at 3'UTR, 8 at 5'UTR, 8 stop codon gains and 10 are non-synonymous missense SNPs. These 10 nsSNP (missense) were selected for further analysis and others were excluded. Similarly for *GSDMD* polymorphism total of 326 SNPs were found in coding region listed in dbSNP and out of these 11 nsSNPs (missense) were selected for further analysis. A list of 10 nsSNPs of *GSDMA* and 11 nsSNPs of *GSDMD* was submitted to SIFT.

Table 1 Functionally significant nsSNPs of *GSDMA* predicted using SIFT and PolyPhen2

SNP	Amino Acid Position	Allele	Amino Acid Change	SIFT Score	Prediction	PolyPhen2 Score	Prediction
rs191833662	2	C/T	T/I	0.01	Deleterious	0.649	Possibly Damaging
rs3894194	18	G/A	R/Q	0.14	Tolerated	0.051	Benign
rs140044904	98	C/T	T/M	0.09	Tolerated	0.56	Possibly Damaging
rs7212944	130	G/A	E/K	0.64	Tolerated	0.016	Benign
rs553099772	182	C/T	P/S	0.12	Tolerated	0.04	Benign
rs115509258	200	G/A	G/D	0.04	Deleterious	0.712	Possibly Damaging
rs200722398	253	G/A	V/I	0.28	Tolerated	0.021	Benign
rs369568940	302	G/A	A/T	1	Tolerated	0.001	Benign
rs56030650	314	C/A/G	T/N	0.1	Tolerated	0.529	Possibly Damaging
rs559726482	331	G/A	A/T	0.01	Deleterious	0.224	Benign

SIFT Score ≥ 0.05 Tolerated, if ≤ 0.05 Deleterious
PolyPhen 2 Score ≥ 0.5 Possibly Damaging, if ≤o.5 Benign

Table 2 PROVEAN and SNP & GO prediction of functionally significant nsSNPs of *GSDMA*.

SNP	Amino Acid Position	Allele	Amino Acid Change	SNP & GO	Prediction	PROVEAN	Prediction
rs19183362	2	C/T	T/I	9	Neutral	-2.969	Deleterious
rs3894194	18	G/A	R/Q	9	Neutral	-0.990	Neutral
rs140044904	98	C/T	T/M	9	Neutral	-0.632	Neutral
rs7212944	130	G/A	E/K	10	Neutral	-0.262	Neutral
rs553099772	182	C/T	P/S	10	Neutral	0.306	Neutral
rs115509258	200	G/A	G/D	7	Neutral	-6.174	Deleterious
rs200722398	253	G/A	V/I	10	Neutral	-0.107	Neutral
rs369568940	302	G/A	A/T	10	Neutral	-0.406	Neutral
rs56030650	314	C/A/G	T/N	8	Neutral	-1.564	Neutral
rs559726482	331	G/A	A/T	8	Neutral	-0.755	Neutral

PROVEAN Score > -2.5 Deleterious, <-2.5 Neutral
SNP&GO Reliability index (RI)= 0 TO 10
If 1- Disease, if >1- Neutral

Table 3 MuPro prediction of functionally significant ns SNPs of *GSDMA*

SNP	Amino Acid Position	Allele	Amino Acid Change	MuPro		
				Confidence Score (Neural Network)	Confidence Score (SVM)	Effect on protein stability
rs191833662	2	C/T	T/I	-0.55783263051882	-0.39533449	Decreases
rs3894194	18	G/A	R/Q	-0.828008075988088	-0.97430434	Decreases
rs140044904	98	C/T	T/M	-0.709087151050457	-0.61873006	Decreases
rs7212944	130	G/A	E/K	-0.970588845677499	-1	Decreases
rs553099772	182	C/T	P/S	-0.997665856467833	-1	Decreases
rs115509258	200	G/A	G/D	-0.694525361597673	-0.56322404	Decreases
rs200722398	253	G/A	V/I	-0.71951067185924	-0.42449228	Decreases
rs369568940	302	G/A	A/T	-0.813799149322277	-1	Decreases
rs56030650	314	C/A/G	T/N	-0.768613874744082	-0.092893227	Decreases
rs559726482	331	G/A	A/T	-0.83594037290925	-1	Decreases

MuPro Score <0 Protein stability decreases, >0 Protein stability increases

Table 4 I- mutant prediction of functionally significant ns SNPs of *GSDMA*

SNP	Amino Acid Position	WT	MT	PH	Temp	DDG value prediction Kcal/mol	Prediction effect
rs191833662	2	T	I	7.0	25	-0.01	Protein Stability decreases
rs3894194	18	R	Q	7.0	25	-0.10	Protein Stability decreases
rs140044904	98	T	M	7.0	25	-0.44	Protein Stability decreases
rs7212944	130	E	K	7.0	25	-0.66	Protein Stability decreases
rs553099772	182	P	S	7.0	25	-1.05	Protein Stability decreases
rs115509258	200	G	D	7.0	25	-1.35	Protein Stability decreases
rs200722398	253	V	I	7.0	25	-0.93	Protein Stability decreases
rs369568940	302	A	T	7.0	25	-0.76	Protein Stability decreases
rs56030650	314	T	N	7.0	25	-0.52	Protein Stability decreases
rs559726482	331	A	T	7.0	25	-1.27	Protein Stability decreases

WT: wild type amino acid. MT: mutant type amino acid. DDG: delta DG (units of free energy) (DDG < 0: decreased stability, DDG > 0: increased stability).

Out of the 10 nsSNPs of *GSDMA,* 3 were predicted to be deleterious by SIFT and of *GSDMD,* 1 nsSNP was predicted to be deleterious and the results are listed in Table 1 and Table 6 respectively. Same list of nsSNPs of both the genes were submitted to PolyPhen2 and the results predict that 4 nsSNPs of *GSDMA* and 2 nsSNPs of *GSDMD* were predicted to be

Table 5 MutPred prediction of functionally significant nsSNPs of *GSDMA*

SNP	Amino Acid Position	MutPred		
		Score (g)	Prediction	Molecular Mechanism Disrupted (P)
rs191833662	2	0.357	Neutral	Loss of disorder (P = 0.0349) Loss of MoRF binding (P =0.1896) Loss of helix (P= 0.2271) Gain of catalytic residue at E5 (P= 0.2699) Loss ofglycosylation atT2 (P = 0.2836)
rs3894194	18	0.383	Neutral	Loss of solvent accessibility (P= 0.0044) Loss of methylation at R18 (P = 0.0661) Loss of catalytic residue at R18 (P = 0.0745) Loss of sheet (P= 0.0817)
rs140044904	98	0.328	Neutral	Loss of stability (P = 0.1058) Loss of phosphorylation at T98 (P = 0.0418) Loss of glycosylation at T98 (P = 0.0657) Loss of mubiquitination at K102 (P = 0.093)
rs7212944	130	0.183	Neutral	Loss of disorder (P = 0.1908) Gain of sheet (P = 0.1945) Gain of methylation at E130 (P = 0.0245) Gain of MoRF binding (P = 0.0281) Loss of ubiquitination at K132 (P = 0.0311) Gain of glycosylation at E130 (P = 0.1621)
rs553099772	182	0.364	Neutral	Gain of phosphorylation at T127 (P = 0.1847) Gain of catalytic residue at P182 (P = 0.0159) Gain of sheet (P = 0.0827) Gain of disorder (P = 0.2767) Loss of loop (P = 0.2897)
rs115509258	200	**0.872**	**High Confidence**	Loss of glycosylation at P178 (P = 0.4561) Loss of sheet (P= 0.0126) Gain of helix (P= 0.0496) Loss of MoRF binding (P =0.0497) Loss of methylation at K199 (P = 0.1138)
rs200722398	253	0.465	Neutral	Loss of loop (P = 0.1242) Loss of helix (P = 0.0033) Gain of loop (P = 0.0079) Loss of disorder (P = 0.163) Gain of catalytic residue at V253 (P = 0.187) Gain of relative solvent accessibility (P=0.2363)
rs369568940	302	0.442	Neutral	Loss of stability (P = 0.0959) Gain of disorder (P = 0.246) Loss of helix (P = 0.3949) Gain of solvent accessibility (P= 0.4946) Gain of phosphorylation at A302 (P = 0.5727)
rs56030650	314	0.240	Neutral	Gain of catalytic residue at T314 (P = 0.0491) Loss of glycosylation at T314 (P = 0.0575) Loss of methylation at K309 (P = 0.0843) Gain of sheet (P= 0.1208) Gain of disorder (P = 0.1648)
rs559726482	331	0.727	Harmful	Gain of methylation at K326 (P =0.0616) Loss of ubiquitination at K326 (P =0.1166) Loss of stability (P = 0.3989) Gain of glycosylation at A331 (P = 0.521) Loss of helix (P = 0.5596)

g > 0.5 deleterious
P score < 0.05 high confidence

Table 6 SIFT and PolyPhen2.0 prediction of functionally significant nsSNPs of *GSDMD*.

SNP	Amino Acid Position	Allele	Amino Acid Change	SIFT Score	Prediction	PolyPhen 2 Score	Prediction
rs375764221	7	A/C/G	R/Q	0.58	Tolerated	0.036	Benign
rs553636785	109	C/T	A/V	0.18	Tolerated	0.0342	Benign
rs62000416	186	A/C	L/M	0.01	Deleterious	0.991	Possibly Damaging
rs149736517	205	C/T	T/M	0.22	Tolerated	0.69	Possibly Damaging
rs143242888	231	C/T	L/F	0.39	Tolerated	0.046	Benign
rs138749323	249	A/G	R/H	0.25	Tolerated	0.372	Benign
rs74645610	396	A/G	Q/R	0.82	Tolerated	0.004	Benign
rs144173624	403	A/G	E/K	0.26	Tolerated	0.061	Benign
rs200540390	434	A/G	E/K	0.18	Tolerated	0.067	Benign
rs143727728	451	A/G	E/K	0.43	Tolerated	0.253	Benign
rs138643473	453	A/G	T/A	1	Tolerated	0.012	Benign

SIFT Score ≥ 0.05 Tolerated, if ≤ 0.05 Deleterious
PolyPhen 2 Score ≥ 0.5 Possibly Damaging, ≤ 0.05 Benign

possibly damaging, the results are listed in Table 1 and Table 6 respectively. I-Mutant 3.0 predicted that all the mutations submitted effect the protein stability (Table 4 and Table 9 respectively). All the nsSNPs were submitted to PROVEAN. Out of which 2 nsSNPs of *GSDMA* and 1 nsSNP of *GSDMD* were predicted to be deleterious and rest all were neutral listed in Table 2 and Table 7 respectively. SNP&GO predicted that effect of all nsSNPs in query were neutral, as shown in Table 2 and Table 7 respectively.

respectively).The efficacy and accuracy of prediction of these *in silico* algorithms for SNPs, such as they are deleterious or not; is increased by combining results of different methods. So, 2 nsSNPs of *GSDMA* namely rs191833662 (T2I) and rs115509258 (G200D) and 1 ns SNP of *GSDMD* namely rs62000416 (L186M) are commonly predicted to be deleterious or disease related by all softwares.

Table 7 PROVEAN and SNP & GO prediction of functionally significant nsSNPs of *GSDMD*.

SNP	Amino Acid Position	Allele	Amino Acid Change	SNP & GO	Prediction	PROVEAN	Prediction
rs375764221	7	A/C/G	R/Q	9	Neutral	-0.868	Neutral
rs553636785	109	C/T	A/V	9	Neutral	-2.007	Neutral
rs62000416	186	A/C	L/M	9	Neutral	-1.568	Neutral
rs149736517	205	C/T	T/M	10	Neutral	-1.969	Neutral
rs143242888	231	C/T	L/F	9	Neutral	-1.939	Neutral
rs138749323	249	A/G	R/H	9	Neutral	-0.490	Neutral
rs74645610	396	A/G	Q/R	10	Neutral	0.547	Neutral
rs144173624	403	A/G	E/K	10	Neutral	-1.777	Neutral
rs200540390	434	A/G	E/K	9	Neutral	-2.641	Deleterious
rs143727728	451	A/G	E/K	10	Neutral	-0.079	Neutral
rs138643473	453	A/G	T/A	10	Neutral	-0.029	Neutral

PROVEAN Score > -2.5 Deleterious, <-2.5 Neutral
SNP&GO Reliability index (RI)= 0 TO 10
If 1- Disease, if >1- Neutral

Table 8 MuPro prediction of functionally significant nsSNPs of *GSDMD*

SNP	Amino Acid Position	Allele	Amino Acid Change	MuPro		
				Confidence Score (Neural Network)	Confidence Score (SVM)	Effect on protein stability
rs375764221	7	A/C/G	R/Q	-0.999935348452093	-1	Decreases
rs553636785	109	C/T	A/V	-0.693761773268883	0.44060811	Increases
rs62000416	186	A/C	L/M	-0.828018978529109	-0.84435898	Decreases
rs149736517	205	C/T	T/M	-0.559995192186926	-0.67988575	Decreases
rs143242888	231	C/T	L/F	-0.884757015308498	-1	Decreases
rs138749323	249	A/G	R/H	-0.906384690613997	-0.36762533	Decreases
rs74645610	396	A/G	Q/R	0.6382094551384849	0.14842856	Increases
rs144173624	403	A/G	E/K	-0.999214387507013	-1	Decreases
rs200540390	434	A/G	E/K	-0.891687340636073	-0.95781075	Decreases
rs143727728	451	A/G	E/K	-0.912279845831437	-1	Decreases
rs138643473	453	A/G	T/A	-0.997066456851101	-0.53623911	Decreases

MuPro Score <0 Protein stability decreases, >0 Protein stability increases

Table 9 I- mutant prediction of functionally significant nsSNPs of *GSDMD*.

SNP	Amino Acid Position	WT	MT	PH	Temp	DDG value prediction Kcal/mol	Prediction effect
rs375764221	7	R	Q	7.0	25	-1.50	Protein Stability decreases
rs553636785	109	A	V	7.0	25	-1.42	Protein Stability decreases
rs62000416	186	L	M	7.0	25	0.01	Protein Stability increases
rs149736517	205	T	M	7.0	25	-0.70	Protein Stability decreases
rs143242888	231	L	F	7.0	25	-0.31	Protein Stability decreases
rs138749323	249	R	H	7.0	25	-0.05	Protein Stability decreases
rs74645610	396	Q	R	7.0	25	-0.57	Protein Stability decreases
rs144173624	403	E	K	7.0	25	-0.07	Protein Stability decreases
rs200540390	434	E	K	7.0	25	-0.17	Protein Stability decreases
rs143727728	451	E	K	7.0	25	-0.04	Protein Stability decreases
rs138643473	453	T	A	7.0	25	-0.16	Protein Stability decreases

WT: wild type amino acid. MT: mutant type amino acid. DDG: delta DG (units of free energy) (DDG < 0: decreased stability, DDG > 0: increased stability).

In MuPro the protein stability decreases for all nsSNPs of *GSDMA* and in *GSDMD* polymorphism due to 2 nsSNPs, protein stability increases and for rest of all stability decreases (Table no 3 and 8 respectively). From MutPred it can be predicted that the 2 nsSNPs of *GSDMA* and 10 nsSNPs of *GSDMD* show probability of being harmful (Table No.5 and 10

DISCUSSION

A single amino acid substitution may results in phenotypic change. A set of powerful softwares/tools/algorithm have been used to predict the phenotypic effect of nsSNP on structure and function of target proteins.

Table 10 MutPred prediction of functionally significant nsSNPs of *GSDMD*

SNP	Amino Acid Position	MutPred		
		Score	Prediction	Molecular Mechanism Disrupted (P)
rs375764221	7	0.510	Harmful	Loss of MoRF binding (P =0.003) Loss of helix (P= 0.3949) Gain of disorder(P = 0.2425) Loss of phosphorylationat S3 (P =0.3697) Loss of stability(P = 0.5083)
rs553636785	109	0.696	Harmful	Loss of disorder(P = 0.0674) Gain of sheet (P= 0.0827) Loss of glycosylation at S113 (P =0.1706) Loss of phosphorylation at S111 (P =0.1797) Loss of loop (P= 0.2237)
rs62000416	186	0.687	Harmful	Loss of glycosylation at S185 (P =0.0618) Gain of MoRF binding (P =0.0665) Loss of sheet (P= 0.0817) Gain of disorder(P = 0.152)
rs149736517	205	0.600	Harmful	Gain of methylation at R183 (P =0.1892) Loss of methylation at K204 (P = 0.0523) Loss of ubiquitination at K203 (P =0.0605) Loss of disorder (P = 0.1209) Loss of sheet (P= 0.1907)
rs143242888	231	0.744	High Confidence	Gain of MoRF binding (P =0.2511) Gain of methylation at K235 (P =0.0494) Gain of loop (P= 0.0851) Loss of sheet (P= 0.1158) Loss of ubiquitination at K235 (P =0.1264) Loss of stability (P = 0.1268
rs138749323	249	0.602	Harmful	Gain of ubiquitination at K248 (P =0.0483) Loss of phosphorylation at S252 (P =0.0752) Loss of methylation at K248 (P = 0.0778) Gain of sheet (P = 0.1208) Gain of catalytic residue at R249 (P =0.1457)
rs74645610	396	0.526	Harmful	Gain of phosphorylation at S395 (P =0.1114) Loss of disorder (P = 0.176) Loss of helix (P= 0.2662) Gain of loop (P= 0.2754) Gain of solvent accessibility (P= 0.3194)
rs144173624	403	0.539	Harmful	Gain of ubiquitination at E403 (P =0.0121) Gain of methylation at E403 (P =0.0424) Gain of helix (P= 0.062) Gain of catalytic residue at E403(P = 0.0687) Loss of loop (P= 0.0986)
rs200540390	434	0.783	High Confidence	Gain of ubiquitination at E434 (P =0.0114) Gain of methylation at E434 (P =0.0119) Gain of glycosylation at E434 (P =0.0354) Loss of disorder (P = 0.1429) Gain of solvent accessibility (P= 0.1505)
rs143727728	451	0.482	Neutral	Gain of ubiquitination at E451 (P =0.0198) Loss of sheet (P= 0.0357) Loss of solvent accessibility (P= 0.0807) Gain of catalytic residue at E451 (P = 0.1429) Gain of glycosylation at E451 (P =0.1683)
rs138643473	453	0.503	Harmful	Loss of phosphorylation at T453 (P =0.0662) Loss of sheet (P= 0.1398) Gain of loop (P= 0.2754) Loss of disorder (P = 0.3265) Loss of solvent accessibility (P= 0.4274)

g > 0.5 deleterious
P score < 0.05 high confidnce

The information can be helpful for ascertaining genotype-phenotype relatedness and further relation to disease biology. A number of studies on polymorphism using *in silico* analysis have helped in prediction of functional nsSNPs associated with genes such as ADRB1 and ADRB2 (Balgir *et al*., 2016). Our results indicate that use of different *in silico* tools help in selection of functional nsSNPs. In *GSDMA* gene 2 nsSNPs namely rs191833662 (T2I) and rs115509258 (G200D) and 1 ns SNP of *GSDMD* that is rs62000416 (L186M) were commonly predicted to be deleterious or disease related by all softwares.

In *GSDMA* one nsSNP rs191833662, that is mutation at amino acid position 2, a Threonine is substituted by Isoleucine. Isolecucine is more hydrophobic than threonine. Isoleucine is unable to form hydrogen bond as threonine forms between theirside-chain and main chain backbone (Yu *et al*., 1984). So this mutation can effect the structure and ultimately function of the protein. The other nsSNP of *GSDMA* is rs115509258, a glycine is replaced by aspartic acid at position 200 of the protein. Glycine is hydrophobic amino acid while aspartic acid is polarized negatively charged amino acid, mainly forming salt

bridges in protein structure. Glycine is often found at the surface of proteins, within loop- or coil regions. It provides high flexibility to the polypeptide chain at these locations. So this replacement may alter the protein structure and consequently function.

In the nsSNP of *GSDMD* rs62000416, leucine is replaced by methionine at position 186 of the protein. Leucine and Methionine both are hydrophobic amino acids and normally buried inside the protein core. However Methionine is involved in formation of disulfide bonds and hence its presence can influence the protein structure if the particular methionine is involved in such bonding. Whereas, some studies shows that this replacement is un- likely to destroy the protein structure (Gassner *et al*., 1996). This may not influence the structure and ultimately function of the protein.

CONCLUSION

In this era of research and technology advancement, Next Generation Sequencing (NGS) generates high throughput data regarding SNPs, however analyzing the biological function of nsSNPs experimentally is money and time consuming. So prediction of structural and functional impact of nsSNPs on protein applying various *in silico* tools can pinpoint relevant SNPs for investigation. Therefore in our study we tried a bioinformatic approach by applying different tools for detection and evaluation of nsSNPs of *GSDMA* and *GSDMD* genes. Out of all the nsSNPs analyzed with various softwares discussed above, a total of 3 nsSNPs were predicted to be deleterious such as 2 of *GSDMA* and 1 of *GSDMD* by all the tools. These polymorphisms could directly or indirectly effect the intra and intermolecular interactions amongst amino acids that can lead to disease risk. Further in-vitro analysis is needed to check the effect of these predicted nsSNPs. The study is the first organized approach for *in silico* analysis of functional SNPs of *GSDMA* and *GSDMD* genes.

References

Balgir, P.P., Goel, R.K., Kaur, J., Sharma, M., Dhiman, S.R. (2016): Insilico Prediction of Functional and Structural Impact of Novel Nonsynonymous SNPS in Human Adrenergic Beta-Receptors. *Int. J. Recent. Sci. Res*., 7(5): 11347-11353.

Barroso, I., Gurnell, M., Crowley, V.E., Agostini, M., Schwabe, J.W. (1999): Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature.*, 402: 880-883.

Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P.L., Casadio, R. (2009): Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum. Mutat*., 30(8):1237-1244.

Capriotti, E., Calabrese, R., Casadio, R. (2006): Predicting the insurgence of human genetic diseases associated to single point protein mutations with Support Vector Machines and evolutionary information. *Bioinfo*., 22: 2729-2734.

Chasman, D., Adams, R.M. (2001): Predicting the functional consequences of nonsynonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. *J. Mol. Biol*., 307: 683-706.

Cheng, J., Randall, A., Baldi, P. (2006): Prediction of Protein Stability Changes for Single Site Mutations Using Support Vector Machines. Proteins Struct. *Funct. Bioinfo*., 62: 1125-1132.

Dabhi, B. and Mistry, K.N. (2014): In silico analysis of single nucleotide polymorphism (SNP) in human TNF-α gene. Meta. Gene., 2: 586-595.

Gassner, N. C., Baase, W.A., Matthews, B.W. (1996): A test of the "jigsaw puzzle" model for protein folding by multiple methionine substitutions within the core of T4 lysozyme. *Proc. Natl. Acad. Sci*. U.S.A.. 93:12155–12158.

Hassan, M.M., Omer, S.E., Khalf-allah, R.M.,Mustafa, R.Y., Ali, I.S., Mohamed, S.B. (2016): Bioinformatics Approach for Prediction ofFunctional Coding/Noncoding Simple Polymorphisms(SNPs/Indels) in Human BRAF *Gene. Adv.Bio.,* Article ID 2632917, 15 pages.

Komiyama, H., Aoki, A., Tanaka, S., Maekawa, H., Kato, Y., Wada, R., Maekawa, T., Tamura, M., Shiroishi, T. (2010): Alu derived cis-element regulates tumorigenesis-dependent gastric expression of GASDERMIN B (GSDMB). *Genes Genet Syst*, 85: 75-83.

Lander, E.S. (1996): The new genomics: global views of biology. *Science*, 274: 536-539.

Li, B., Krishnan, V.G., Mort, M.E., Xin, F., Kamati, K.K., Cooper, D.N., Mooney, S.D., Radivojac, P. (2009): Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinfo*., 25(21):2744-2750.

Manickam, M., Ravanan, P., Singh, P., Talwar, P. (2014): In silico identification of genetic variants in glucocerebrosidase (GBA) gene involved in Gaucher's disease using multiple software tools. Front Genet., 5: 148.

Ng, P.C., and Henikoff, S. (2003): SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 31(13): 3812-3814.

Ramensky V, Bork P, Sunyaev S. (2002): Human non-synonymous SNPs: server and survey. *Nucleic Acids R*es. 1; 30(17):3894-900.

Saeki, N., Kuwahara, Y., Sasaki, H., Satoh, H., Shiroishi, T. (2000): Gasdermin (Gsdm) localizing to mouse Chromosome 11 is predominantly expressed in upper gastrointestinal tract but significantly suppressed in human gastric cancer cells. *Mamm Genome*., 11(9):718-24.

Saeki, N., Usui, T., Aoyagi, K., Kim, D.H., Sato, M., Mabuchi, T., Yanagiha,K., Ogawa, K., Sakamoto, H., Yoshida, T., Sasaki, H. (2009): Distinctive expression and function of four GSDM family genes (GSDMA-D) in normal and malignant upper gastrointestinal epithelium. *Cancer*., 48: 261-271.

Smith, E.P., Boyd, J., Frank, G.R., Takahashi, H., Cohen, R.M. (1994): Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man. N. Engl. J. Med., 331: 1056–1061.

Tamura, M., Tanaka, S., Fujii, T., Aoki, A., Komiyama, H., Ezawa, K., Sumiyama, K., Sagai, T., Shiroishi, T.(2007): Members of a novel gene family, Gsdm, are expressed exclusively in the epithelium of the skin and gastrointestinal tract in a highly tissue-specific manner. *Genomics*., 89: 618-629.

Yu, M.H. and King, J. (1984): Single amino acid substitutions influencing the folding pathway of the phage P22 tail spike endorhamnosidase (protein folding/amino acids/DNA

sequences/temperature-sensitive mutants). *Biochem*., 81, pp, 6584-6588.

Web links

http://www.ncbi.nlm.nih.gov/snp/
http://expasy.org/
http://www.uniprot.org
http://www.1000genome.org/
http://www.ensembl.org/index.html
http://sift.jcvi.org/
http://genetics.bwh.harvard.edu/pp2
http://provean.jcvi.org/index.php
http://snps-and-go.biocomp.unibo.it/snps-and-go/
http://gpcr2.biocomp.unibo.it/cgp/predictor/I-Mutant3.0/I-Mutant3.0.cgi
http://mutpred.mutdb.org/
http://www.igb.uci.edu/servers/servers.html

How to cite this article:

Praveen P. Balgir and Suman Rani.2017, In Silico Analysis of Human Gsdma and Gsdmd Genes for Functional and Structural Impact of Non-Synonymous SNPs. *Int J Recent Sci Res.* 8(6), pp. 17977-17984.

****