International Journal of Advanced Biochemistry Research

ISSN Print: 2617-4693
ISSN Online: 2617-4707
IJABR 2024; SP-8(1): 839-844
www.biochemjournal.com
Received: 08-10-2023
Accepted: 12-12-2023

Gaurav Patel
Animal Genetics and Breeding, National Dairy Research Institute, Karnal, Haryana, India

Ayushi Singh
Animal Genetics and Breeding, Indian Veterinary Research Institute, Izzatnagar Bareilly, Uttar Pradesh, India

Ymberzal Koul
Animal Genetics and Breeding Division, National Dairy Research Institute, Karnal, Haryana, India

Utsav Surati
Animal Genetics and Breeding, National Dairy Research Institute, Karnal, Haryana, India

Sachin Patidar
Department of Animal Husbandry, Govt. of Rajasthan, Rajasthan, India

Corresponding Author:
Gaurav Patel
Animal Genetics and Breeding, National Dairy Research Institute, Karnal, Haryana, India

In-depth analysis of deleterious non-synonymous SNPs in the bovine ANPEP gene and their structural consequences

Gaurav Patel, Ayushi Singh, Ymberzal Koul, Utsav Surati and Sachin Patidar

DOI: https://doi.org/10.33545/26174693.2024.v8.i1Sk.456

Abstract
This study delves into the genetic variations within the Aminopeptidase N (ANPEP) gene in cattle, focusing on missense variants and their potential impact on protein structure, stability, and function. A dataset containing SNPs was retrieved from Ensembl-Biomart and Uniprot databases. Deleterious non-synonymous SNPs were identified using SIFT, PANTHER, POLYPHEN-2, PhDSNP, PredictSNP, and SNAP2. The study also predicts the structural and functional consequences of these SNPs using I-Mutant 2.0, MUPRO, mCSM, and HOPE. Furthermore, protein-protein interaction analysis was conducted using STRING. Results reveal a set of common deleterious variants with potential implications for protein stability. The study enhances our understanding of genetic variations in the ANPEP gene in cattle, offering insights into potential functional consequences.

Keywords: Cattle, ANPEP gene, nsSNPs

Introduction
A single nucleotide polymorphism (SNP) is a source variance in a genome. A single base mutation in DNA is called an SNP. SNPs account for 90% of all genetic variations in the genome and are the most prevalent and basic type of polymorphism. For SNPs, there are numerous publicly accessible databases, including dbSNP, GWAS Central, and SwissVar. Only the missense variants, also known as non-synonymous SNPs (nsSNPs), are particularly significant because they alter the translated amino acid residue sequence. Given their widespread association with many disorders, nsSNPs most likely contribute significantly to the functional diversity of coding proteins in human populations. By decreasing protein solubility or by disrupting protein structure, nsSNPs may have an impact on how proteins function and they may affect gene regulation by transcription and translation [1, 2].
Aminopeptidase N (ANPEP) is a vital component of the purine phosphoribosyl transferases (PRT) and assumes a crucial role in the purine salvage pathway, facilitating the conversion of preformed purine and phosphoribosyl pyrophosphate (PRPP) substrates into nucleotide monophosphates [3]. In cattle, the ANPEP gene plays a significant role in various physiological processes, and its proper functioning is essential for normal cellular activities. Research on the ANPEP gene in cattle is essential for understanding its role in bovine physiology. Genetic studies on this gene may provide insights into its impact on metabolic pathways and overall cellular function. Furthermore, investigating the inheritance patterns and potential mutations of the ANPEP gene in cattle could contribute to our understanding of genetic disorders or variations that may affect bovine health.

Materials and Methods
Data set
The data of the gene ANPEP was retrieved from Ensembl-Biomart Databases (source: dbSNP; http://www.ensembl.org/ biomart/martview/) and Uniprot (https:// www.uniprot.org/uniprot/). We retrieved the information of SNPs (SNP ID, location, Gene stable ID, residue alteration, etc.).

Identification of deleterious nsSNPs

The substituted amino acids that alter protein function and phenotypic changes was predicated by SIFT (Sorting Intolerant from Tolerant) score from Ensembl database. SIFT predicted the deleteriousness of the SNP in the form of a tolerance index (TI) score ranging from 0.0 to 1.0, with a TI score of 0.05 or less as intolerant or deleterious [4, 5].

Further, to verify the identified deleterious SNPs from SIFT; PANTHER, POLYPHEN-2, PhDSNP, PredictSNP, and SNAP2 were used. Results of PhDSNP were obtained by the consensus classifier of PredictSNP.

Prediction of structural and functional effect on ANPEP gene

For prediction of protein stability change three different web servers I-Mutant 2.0, MUPRO, and mCSM were used. These are Support Vector Machine (SVM) - based web server for the automatic prediction of protein stability changes upon single-site mutations [6, 7]. The input is a FASTA sequence of protein along with the residues change was provided. I-Mutant 2.0 predicts free energy change and RI value (reliability index). If the DDG value is negative, then the mutated protein will have less stability and vice versa for high stability.

HOPE version 1.1.1 (https://www3.cmbi.umcn.nl/hope/) was used to recognise the structural effect of nonsynonymous change in the ANPEP protein sequence. It also provided 3D structure visualization of altered protein and superimposition of wild and mutant after providing altered protein sequence as input [8].

Prediction of Protein- Protein interaction

For protein-protein interaction "Search Tool for the Retrieval of Interacting Proteins" (STRING; http://string-db.org/) was used [9].

Results and Discussion

A total of 174 missense variant and 68 synonymous variants were filtered out in the bovine ANPEP gene during variant calling. The highest number of SNPs were present in intronic region (2598), followed by the downstream region (270), upstream region (230).

Identification of deleterious nsSNPs in cattle ANPEP gene

Based on SIFT score of variants obtained from the ensemble, 86 variants are found to be deleterious out of 174. As ANPEP gene has two protein transcripts, common variants were removed. After removing common variants (same variant id with similar substitution) from both transcripts, we got 54 unique deleterious variants based on SIFT score. Which were further analyzed using PANTHER, POLYPHEN-2, PhDSNP, PredictSNP, and SNAP2. Twenty nsSNPs named A411D, D185N, D430V, D430H, F179L, F218S, I864S, K476M, P79S, S273F, S576P, T461P, T509P, T810P, T847P, V23G, Y191S, Y195D, Y853C, and Y890C were found common deleterious (Supplementary file; Figure 1).

Several similar studies done previously. In the bovine SLC11A2 gene, deleterious SNPs were predicted using SIFT, PolyPhen and Panther [10].

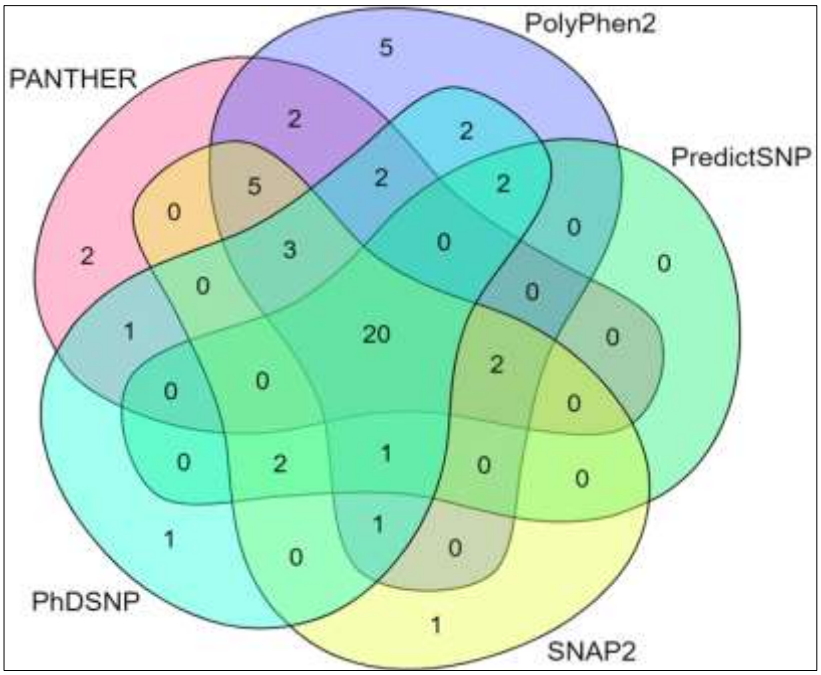

Fig 1: Common deleterious nsSNPs

Prediction of change in protein stability due to mutation for the common mutation

For these three programs were used, I-Mutant 2.0, muPro, and mCSM. In muPro all variants decrease the protein stability except S273F. Through I-Mutatnt prediction, most variants decrease the protein stability except S273F, Y853C, and Y890C. In the mCSM program A411D, F218S, I864S, P79S, Y191S, and Y195S variants were highly destabilizing the protein structures and D430V was stabilizing, and rest were destabilizing structures (Table 1).

When considering a protein's structural and functional properties, stability is crucial. Any alteration in the stability of proteins may result in abnormal protein aggregation, misfolding, or destruction [11].

Table 1: Prediction of Protein stability

Deleterious variant	muPro	DDG	I-Mutant2.0	DDG	mCSM prediction	mCSM
A411D	Decrease	-0.3431432	Decrease	-1.13	Highly Destabilizing	-2.937
D185N	Decrease	-1.2290843	Decrease	-0.1	Destabilizing	-0.55
D430H	Decrease	-1.2098438	Decrease	-1.17	Destabilizing	-0.28
D430V	Decrease	-0.531635	Decrease	-0.4	Stabilizing	0.431
F179L	Decrease	-0.9521074	Decrease	-1.17	Destabilizing	-1.741
F218S	Decrease	-1.7815225	Decrease	-2.27	Highly Destabilizing	-3.08
I864S	Decrease	-2.4348956	Decrease	-2.41	Highly Destabilizing	-3.235
K476M	Decrease	-0.3666694	Decrease	-0.27	Destabilizing	-1.158
P79S	Decrease	-1.1829543	Decrease	-1.94	Highly Destabilizing	-2.57
S273F	Increase	0.1516802	Increase	0.68	Destabilizing	-0.84
S576P	Decrease	-0.6185008	Decrease	-1.11	Destabilizing	-0.156
T461P	Decrease	-1.2319006	Decrease	-1.85	Destabilizing	-0.385
T509P	Decrease	-0.7971856	Decrease	-1.2	Destabilizing	-0.32
T810P	Decrease	-1.3522509	Decrease	-1.12	Destabilizing	-0.62
T847P	Decrease	-1.3522509	Decrease	-1.12	Destabilizing	-0.281
V23G	Decrease	-2.2746834	Decrease	-4.33		
Y191S	Decrease	-1.1389335	Decrease	-2.34	Highly Destabilizing	-3.682
Y195D	Decrease	-1.6506274	Decrease	-1.46	Highly Destabilizing	-3.052
Y853C	Decrease	-1.0610254	Increase	0.34	Destabilizing	-0.737
Y890C	Decrease	-1.0610254	Increase	0.34	Destabilizing	-1.191

Prediction of functional impact

Based on the result of I-Mutant 2.0, muPRO, and Mcsm; six variants A411D, F218S, I864S, P79S, Y191S, and Y195D were highly decrease protein stability. In-depth functional impact analysis of six variants for insights into potential biological implications were done using HOPE server (Table 2).

Table 2: Functional Impact of deleterious nsSNPs

Mutation	Comparative Size of Mutant	Hydrophobicity (of Mutant)	Key Points	Prediction
A411D	Bigger	Less hydrophobic	• Located in the metalloprotease region • Wild was neutral, mutant residue charge is negative • Highly conserved position	Difference between charge and size can disturb the interaction with metal-ion "zinc" Based on this conservation information this mutation is probably damaging to the protein.
F218S	Smaller	Less hydrophobic	• Mutant residue located in the metalloprotease region of the protein mutant residue is located near a highly conserved position.	Based on conservation scores this mutation is probably damaging to the protein.
I864S	Bigger	-	• The mutant residue is located near a highly conserved position. • The mutated residue is located in a domain that is important for the binding of other molecules	Based on conservation scores this mutation is probably damaging to the protein.
P79S	smaller	Less hydrophobic	• Mutation located in the Metalloprotease region	Based on conservation information mutation is probably damaging to protein.
Y191S	smaller		• Located in the Metalloprotease region	In some rare cases mutation might occur without damaging the protein.
Y195D	smaller	Less hydrophobic	• Located in the Metalloprotease region • wild-type residue charge was NEUTRAL, the mutant residue charge is NEGATIVE	Based on this conservation information this mutation is probably damaging to the protein.

Protein-Protein interaction analysis

In the protein-protein interaction analysis by STRING, a total of ten proteins were found to be associated with functional ANPEP protein with a high confidence score (< 0.95) (Figure 2). Proteins like GCLM, GGT6, LAP3, FOLH1, SFN, GGT7 were found to be associated with ANPEP protein.

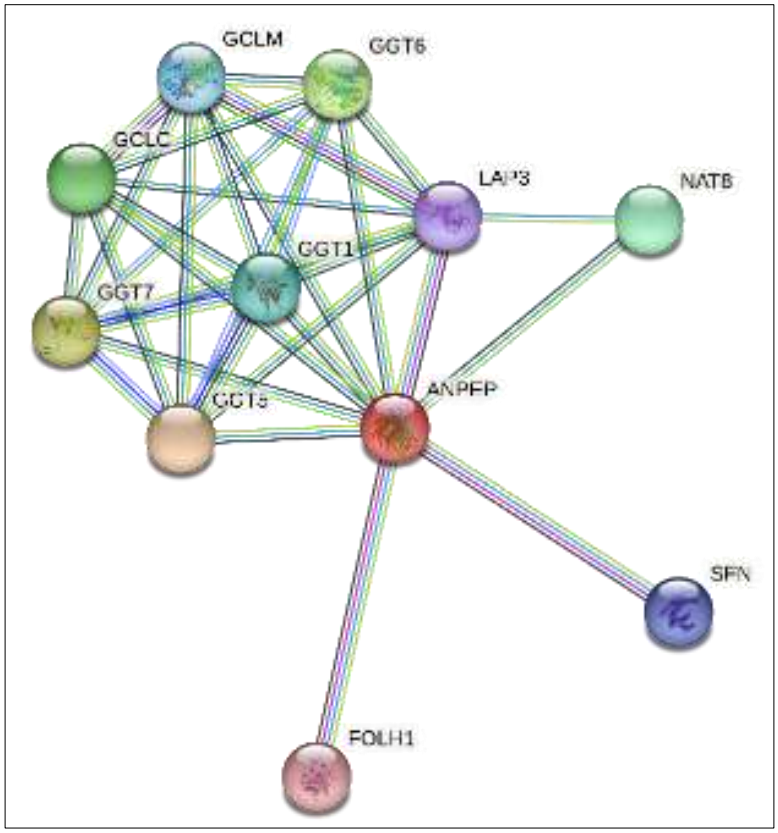

Fig 2: Protein interaction network of ANPEP

Conclusion

This comprehensive analysis sheds light on the impact of missense variants in the bovine ANPEP gene, providing valuable information on potential disruptions to protein stability and function. The identified deleterious variants may have implications for bovine physiology, warranting further investigation into their role in metabolic pathways and cellular functions. This research contributes to the broader understanding of genetic variations in key genes, offering insights into potential implications for bovine health and providing a foundation for future studies in this field.

Acknowledgements

The authors are thankful to the Directors of ICAR-National Bureau of Animal Genetic Resources and National Dairy Research Institute for providing the necessary facilities.

Conflict of interest

Authors have no conflict of interest in this study.

References

1. Barroso I, Gurnell M, Crowley VE, Agostini M, Schwabe JW, Soos MA, *et al*. Dominant negative mutations in human PPARγ associated with severe insulin resistance, diabetes mellitus and hypertension. Nature. 1999 Dec;402(6764):880-883.
2. Chasman D, Adams RM. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: Structure-based assessment of amino acid variation. Journal of molecular biology. 2001 Mar 23;307(2):683-706.
3. Sørensen KD, Abildgaard MO, Haldrup C, Ulhøi BP, Kristensen H, Strand S, *et al*. Prognostic significance of aberrantly silenced ANPEP expression in prostate cancer. British journal of cancer. 2013 Feb;108(2):420-428.
4. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome research. 2001 May 1;11(5):863-874.
5. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic acids research. 2003 Jul 1;31(13):3812-3814.
6. Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic acids research. 2005 Jul 1;33(suppl 2):W306-10.
7. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins: Structure, Function, and Bioinformatics. 2006 Mar 1;62(4):1125-1132.
8. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC bioinformatics. 2010 Dec;11(1):1-10.
9. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, *et al*. The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic acids research. 2010 Nov2;39(suppl 1):D561-568.
10. Patel SM, Koringa PG, Reddy BB, Nathani NM, Joshi CG. In silico analysis of consequences of non-synonymous SNPs of Slc11a2 gene in Indian bovines. Genomics Data. 2015 Sep 1;5:72-79.
11. Kong L, Rupp B. Protein stability: a crystallographer's perspective. Acta Crystallographica. Section F, Structural Biology Communications, 2016 Jan 26, 72(Pt 2).

Supplementary

Variant ID	AA co-ord	Panther preservation time	Panther prediction	Polyphen-2	Polyphen-2 score	Predict SNP	Predict SNP accuracy	SNAP2	SNAP2 Score	PhD-SNP prediction	PhD-SNP expected accuracy
rs715733499	V20M	176	Probably benign	Possibly damaging	0.71	Neutral	0.73834499	Neutral	-1	Neutral	0.660879
rs134133286	V23G	456	Probably damaging	Probably damaging	0.978	Deleterious	0.7556615	Effect	47	Deleterious	0.73260309
rs438435099	P79S	1038	Probably damaging	Probably damaging	1	Deleterious	0.7556615	Effect	69	Deleterious	0.85822785
rs445214795	D141A		Not scored: Invalid substitution	Benign	0.091	Neutral	0.75203963	Neutral	-25	Deleterious	0.60798122
rs473115546	F179L	1038	Probably Damaging	Probably damaging	1	Deleterious	0.71871275	Effect	40	Deleterious	0.73260309
rs800590783	D185N	1038	Probably Damaging	Possibly damaging	0.838	Deleterious	0.50595948	Effect	66	Deleterious	0.73260309
rs467858592	Y191S	1038	Probably Damaging	Probably damaging	1	Deleterious	0.7556615	Effect	87	Deleterious	0.87523992
rs434115333	Y195D	1038	Probably Damaging	Probably damaging	1	Deleterious	0.60697259	Effect	85	Deleterious	0.85822785
rs478348406	N199I	30	Probably Benign	Possibly damaging	0.859	Deleterious	0.71871275	Neutral	-27	Deleterious	0.67620995
rs482101480	F218S	1038	Probably Damaging	Probably damaging	1	Deleterious	0.86908365	Effect	53	Deleterious	0.67620995
rs521298929	S273F	1038	Probably Damaging	Probably damaging	1	Deleterious	0.86908365	Effect	80	Deleterious	0.87523992
rs451279161	A411D	1038	Probably Damaging	Probably damaging	1	Deleterious	0.86908365	Effect	71	Deleterious	0.88474971
rs466935231	K429E	456	Probably Damaging	Probably damaging	0.983	Neutral	0.73834499	Effect	36	Neutral	0.68183996
rs445188646	K429N	456	Probably Damaging	Probably damaging	0.976	Neutral	0.73834499	Effect	45	Neutral	0.58230958
rs462828853	D430V	842	Probably Damaging	Probably damaging	0.998	Deleterious	0.71871275	Effect	77	Deleterious	0.73260309
rs478170687	D430H	842	Probably Damaging	Probably damaging	1	Deleterious	0.71871275	Effect	75	Deleterious	0.60798122
rs462828853	D430A	842	Probably Damaging	Probably damaging	1	Neutral	0.63151762	Effect	27	Neutral	0.55202703
rs450798252	V433L	456	Probably Damaging	Benign	0.21	Neutral	0.73834499	Neutral	-40	Neutral	0.55202703
rs461930410	T448P		Not Scored: Invalid Substitution	Benign	0.028	Deleterious	0.86908365	Effect	55	Deleterious	0.67620995
rs461930410	T448A		Not Scored: Invalid Substitution	Benign	0.076	Neutral	0.63151762	Neutral	-19	Neutral	0.58230958
rs440177651	T461P	324	Possibly Damaging	Probably damaging	0.999	Deleterious	0.86908365	Effect	37	Deleterious	0.81731169
rs473086748	K476M	1038	Probably Damaging	Probably damaging	1	Deleterious	0.71871275	Effect	74	Deleterious	0.60798122
rs876275665	T509P	361	Possibly Damaging	Probably damaging	0.999	Deleterious	0.7556615	Effect	59	Deleterious	0.60798122
rs472902583	I559S		Not Scored: Invalid Substitution	Probably damaging	0.999	Neutral	0.6025641	Effect	41	Deleterious	0.58885542
rs465380808	D567A	842	Probably Damaging	Probably damaging	0.999	Neutral	0.62529138	Effect	53	Deleterious	0.7733853
rs467615355	S576P	456	Probably Damaging	Probably damaging	0.999	Deleterious	0.86908365	Effect	28	Deleterious	0.88474971
rs461884564	W582R	1038	Probably Damaging	Probably damaging	1	Neutral	0.75291375	Effect	58	Neutral	0.55202703
rs457659198	Y598C	176	Probably Benign	Possibly damaging	0.614	Deleterious	0.54946365	Neutral	-13	Deleterious	0.7733853
rs462682986	D617V		Not Scored: Invalid Substitution	Benign	0.097	Deleterious	0.86908365	Effect	12	Deleterious	0.7733853
rs440225374	I641N	176	Probably Benign	Probably damaging	1	Deleterious	0.86908365	Effect	6	Deleterious	0.67620995
rs452140124	T682P	1037	Probably Damaging	Probably damaging	1	Neutral	0.82622462	Effect	34	Neutral	0.7828765
rs467826408	T740I		Not Scored: Invalid Substitution	Benign	0.176	Neutral	0.73834499	Effect	4	Neutral	0.660879
rs519626872	N776H		Not Scored: Invalid Substitution	Possibly damaging	0.488	Neutral	0.82622462	Neutral	-28	Neutral	0.68183996
rs448419296	T810P	324	Possibly Damaging	Probably damaging	0.998	Deleterious	0.60548272	Effect	73	Deleterious	0.60798122
rs447485103	D840A	910	Probably Damaging	Probably damaging	0.99	Neutral	0.62529138	Effect	33	Deleterious	0.73260309

rs471782704	Y853C	1038	Probably Damaging	Probably damaging	0.999	Deleterious	0.71871275	Effect	57	Deleterious	0.58885542
rs444423342	I864S	456	Probably Damaging	Probably damaging	1	Deleterious	0.71871275	Effect	59	Deleterious	0.81731169
rs455617729	E876D	456	Probably Damaging	Benign	0.152	Neutral	0.82622462	Neutral	-24	Neutral	0.68183996
rs433817941	Q879K	456	Probably Damaging	Possibly damaging	0.692	Neutral	0.73834499	Neutral	-79	Neutral	0.44670846
rs468706779	K884M	361	Possibly Damaging	Probably damaging	0.999	Neutral	0.73834499	Neutral	-43	Neutral	0.68183996
rs454751786	K907N	30	Probably Benign	Possibly damaging	0.885	Neutral	0.73834499	Neutral	-40	Neutral	0.55202703
rs211627711	N922H		Not scored: Invalid substitution	Benign	0.014	Neutral	0.73834499	Neutral	-71	Neutral	0.83210379
rs1116627945	D926H		Not scored: Invalid substitution	Possibly damaging	0.865	Neutral	0.6025641	Neutral	-54	Neutral	0.68183996
rs41970584	S30F	361	Possibly damaging	Probably damaging	1	Deleterious	0.86908365	Effect	8	Neutral	0.660879
rs519626872	N813H		Not scored: Invalid substitution	Possibly damaging	0.641	Neutral	0.73834499	Neutral	-37	Deleterious	0.7733853
rs448419296	T847P	324	Possibly Damaging	Probably damaging	0.998	Deleterious	0.60548272	Effect	58	Deleterious	0.73260309
rs447485103	D877A	910	Probably Damaging	Probably damaging	0.994	Neutral	0.62529138	Effect	37	Deleterious	0.81731169
rs471782704	Y890C	1038	Probably Damaging	Probably damaging	0.999	Deleterious	0.71871275	Effect	54	Deleterious	0.7733853
rs444423342	I901S	456	Probably Damaging	Probably damaging	1	Deleterious	0.71871275	Effect	44	Neutral	0.55202703
rs455617729	E913D	456	Probably Damaging	Benign	0.128	Neutral	0.82622462	Neutral	-30	Deleterious	0.7733853
rs433817941	Q916K	456	Probably Damaging	Possibly damaging	0.664	Neutral	0.65307311	Neutral	-82	Deleterious	0.60798122
rs468706779	K921M	361	Possibly Damaging	Probably damaging	0.999	Neutral	0.73834499	Neutral	-55	Deleterious	0.73260309
rs454751786	K944N	30	Probably Benign	Possibly damaging	0.748	Neutral	0.73834499	Neutral	-41	Deleterious	0.58885542
rs1116627945	D963H		Not scored: Invalid substitution	Probably damaging	0.967	Neutral	0.6025641	Neutral	-71	Neutral	0.68183996