Journal of Population Therapeutics & Clinical Pharmacology

RESEARCH ARTICLE
DOI: 10.53555/jptcp.v31i1.4277

IN-SILICO CHARACTERIZATION OF NON-SYNONYMOUS SINGLE NUCLEOTIDE POLYMORPHISMS IN HUMAN MDM2: IMPLICATIONS FOR CANCER SUSCEPTIBILITY

Awais Khan[1], Azhar Hussain Shah[2], Syed Basit Ali Shah[3], Aatka Jamil[4], Khuzin Dinislam[5], Attiq Ullah[6*]

[1]Department of Biotechnology and Genetic Engineering, Hazara University, Mansehra 21300, Pakistan / Department of Innovative Technology of Biomedical Engineering & Medical Devices, Ming Chi University of Technology, Taiwan
[2]Department of Biotechnology and Genetic Engineering, Hazara University, Mansehra 21300, Pakistan
[3,6*]College of Life Sciences, Northeast Forestry University, No 26, Hexing Road, Harbin 150040 China
[4]National Agriculture Research Center Islamabad 44000, Pakistan
[5]Department of Chemistry, Bashkir State Medical University, Street Lenia, 3, Ufa, Republic of Bashkortostan 450008, Russia

***Correspondence:** Attiq Ullah
*College of Life Sciences, Northeast Forestry University, No 26, Hexing Road, Harbin 150040 China, Email: attiqak464@gmail.com

Abstract

The oncogene MDM2 (Murine Double Minute 2) was first found in DNA and bound to the paired acentric chromosome. 0SNPs (single nucleotide polymorphisms) are important for determining the genetic basis of several complicated human disorders. Non-synonymous single nucleotide polymorphisms (nsSNPs) are mutation in only a one nucleotide in sequences of any gene that has changed the specific synthesized protein in its function, structure and morphogenesis. Finding the affected SNPs in the genetic make-up of diseases remains a difficult task. The present *Insilico* study explored and identified the genetic variation that affects and modifies the expression of MDM2 genes. SIFT found that 36 SNPs in the MDM2 gene are harmful while polyphen2 analyzed 12 SNPs in MDM2 gene. An aggregate result was obtained by examining six tools with diverse views, and seven nsSNPs were shown to be the most likely to have a detrimental impact. I-Mutant and Project HOPE approaches were utilized to predict the mutant proteins' severe structural and functional instability, while the InterPro was utilised the predict SNP in crucial functional domains. Only 9 SNPs rs764034976, rs753663917, rs761546875, rs755429424, rs764918809, rs763077439, rs765555199, rs1475420873 and rs550783815 were discovered that influence the structure, function, and stability of the MDM2 protein synthesized by MDM2 gene. To our knowledge, there has never been any research on the SNPs of the MDM2 gene. This is the first study on the MDM2 gene using computation methods.

Keywords: MDM2, Oncogene, Cancer biology, Therapeutic targets, Nucleotide polymorphisms, Non-Synonymous Single Nucleotide Polymorphisms.

Introduction

The Murine Double Minute Clone 2 (MDM2) oncogene, discovered three decades ago, represents a crucial player in cancer biology, capable of triggering tumorigenesis (Momand et al., 2000). Initially identified as a gene linked to the tumor suppressor p53, located at 12q15, MDM2's significance lies in its intricate role as an E3 ubiquitin ligase. This function allows MDM2 to regulate the activation of p53, a well-documented negative regulator of cellular growth (Momand et al., 2000)

Maintaining low levels of p53 in normal cells is critical due to its potent growth inhibitory actions. The absence of MDM2 renders p53 lethal, highlighting the delicate balance required for cellular homeostasis. MDM2 employs various mechanisms to inhibit p53 activation, including promoting p53 degradation by transporting it from the nucleus to the cytoplasm and facilitating ubiquitination for proteasomal destruction. Additionally, MDM2 inhibits the transactivation domain of p53, preventing the action of intact p53. The multifaceted role of p53 involves the activation of genes responsible for cell cycle disruption and apoptosis (Momand et al., 2000).

MDM2's carcinogenic effects extend beyond p53 regulation, encompassing the ubiquitination of transcription factors like ATF, E2F, retinoblastoma protein, Histones H2A, and H2B. This ubiquitination influences crucial cellular processes such as cell proliferation, DNA repair, transcription, ribosomal biosynthesis, and even serine metabolism when MDM2 is chromatin-bound (Cissé et al., 2020).

In human malignancies, a prevalent genetic change involves the overexpression of MDM2, often surpassing 50-fold amplification. Intriguingly, tumors with MDM2 amplification can grow without the need for p53 mutations. The prevalence of MDM2 gene amplification in human cancer biology fluctuates between 3.5% and 4.4%. The recognition of MDM2's significance has elevated it to a high priority in cancer treatment, leading to the development of compounds targeting the MDM2-p53 connection (Dembla et al., 2018; Kato et al., 2018).

Therapeutic advancements, exemplified by molecules like Nutlins, have emerged as pharmacological inhibitors disrupting the MDM2-p53 interaction. Despite challenges such as limited clinical effectiveness and dose-limiting thrombocytopenia, ongoing research explores new compounds inhibiting MDM2 or its interactions with p53 (Aydin et al., 2020; Cissé et al., 2020; Cornillie et al., 2020).

Beyond its role in oncology, MDM2 is increasingly recognized for its involvement in non-neoplastic disorders. Understanding genetic diseases is facilitated by exploring nucleotide polymorphisms, particularly single nucleotide polymorphisms (SNPs), which account for the majority of biological variations (Nie et al., 2019). Among the three MDM genes (MDM1, MDM2, and MDM3), MDM2, found on chromosome 12q14.3-q15, emerges as a pivotal gene exhibiting overexpression and significant roles in gene transformation.

OBJECTIVES

- To identify and scanning of functional SNPs in Human MDM2 gene by Insilco approaches.
- To predict and map the potential impact of (nsSNPS) on the structure and function of Human MDM2 protein.

MATERIAL AND METHODS

2.1 Datasets of SNPs:

MDM2 SNPs were found in the database dbSNPs. To scan and identify functional SNPs of the MDM2 gene, present study used ENSEMBLE, SNP500 cancer, Gene cards, and Uniprot databases. The dbSNPs database was created with the help of the National Center for Biotechnology

Information (NCBI) and the National Human Genome Research Institute (NHGRI). This is a fantastic resource for genomic variability data, and it's maintained on a frequent basis as new additions are submitted (Sherry et al., 2001). Data from dbSNP is also updated and shared through Ensemble's variation resource. Ensemble's data mining capabilities are a crucial aspect for even further data processing (Y.-A. Chen et al., 2019). GenCard and Uniprot are the two databases that provide similar information and similar data. MDM2 nucleotides and protein sequence were also obtained from the ENSEMBL database (A. D. Yates et al., 2020).

2.2 Deleterious nsSNPs Predication:

SIFT (Sorting intolerance from tolerance) was used to determine whether non-synonymous SNPs had a tolerant or detrimental influence on protein function (Ng & Henikoff, 2003). By comparing the homologous sequence and physical characteristics of proteins, SIFT predicts that impact of nsSNPs on structure of protein. The program's algorithms employed several data bases such as SWISSPROT and Tremble to check for homogeneous sequence (Boeckmann et al., 2003). SIFTs were used to classify SNPs as intolerant or tolerant based on homologous alignment. As essential amino acids, they are under a lot of selective pressures and are extremely conservative in the protein family. Variations in such resultant amino acids are regarded as harmful. The threshold value, less than 0.05 is termed as intolerant, however when it became greater than 0.05, it indicates the probability of specific amino acids as a, tolerant. If the value increases, SNPs will have a lower functional influence on protein function (Ng & Henikoff, 2002). A higher tolerance index indicates that a mutation's effect on protein structure and function will be less pronounced.

2.3 Structure Homology Based Method to find the Consequences of Coding nsSNPs:

For categorization, this tool employed a nave Bayesian method with a score range of 0-1The mutations are classified as benign, perhaps detrimental, or probably destructive based on the score. Scores close to 1 are regarded as likely harmful, having a strong influence on the structure of protein (Acevedo et al., 2016).

2.4 Functional characterization of nsSNPS

In current studies, the nsSNPS analyzers such as PROVEAN and P-MUT were used to analyze the characterization of functional nsSNPS, SNP&GO (Colombo et al., 2014) and PhD-SNP (Capriotti & Fariselli, 2017). For the categorization and explanation of non-synonymous SNPs' effects on protein. PHD SNP analyses and classifies genes into 2 groups: harmful or harmless (Capriotti et al., 2006). MutPred is a high-efficiency variant of the SIFT algorithm (Pejaver et al., 2020). MutPred was used to determine the structural and functional differences between wild and mutant proteins. The G value dictates whether a mutation affects a protein's function. The G-value ranges from 0 to 1, with the greater the G-value, the greater the impact (Kamaraj & Purohit, 2016). SNP&GO also employed the SVM algorithm. PMUT employed neutral networking to analyses several types of sequence information and identify mutations based on the information acquired. The PMUT was used to process altered protein sequences in FASTA format, and the result might be diseased or neutral depending on the sequence's likelihood score.

2.5 nsSNPs prediction on molecular structure of protein:

SNPeffect was used to predict the presence of nsSNPS in the coding area of the MDM2 protein. A conservation score was calculated and predicted for the landscape of protein homeostasis(Powers et al., 2009). SNPeffect employs a variety of methods, including TANGO to predict mutant aggregation propensity, WALTZ to predict amyloid propensity, and LIMBO to estimate chaperone binding propensity. The TANGO score of wild and mutant amino acids was determined using distinct TANGO scores (Dtango), which is used to determine aggregating susceptibility (Moreira et al., 2013). The tendency increases if the score is larger than 50, decreases if it is less than -50, and has no impact if it is between -50 and 50.

Non-synonymous mutations have an impact on protein stability (INPS):
The structure and function of protein is depending upon its stability so, in current studies the I-MUTANT 3.0 program was preferred to calculate the possible instabilities in MDM2 protein. SVM is also used in this algorithm. By comparing the I-MUTANT and MUPRO results, the validity of the I-MUTANT result was established. These systems utilised the very same method to determine whether a mutation would improve or reduce stability (Chitranshi et al., 2017) The SRide server was used to predict the residues which were stabilizing the normal and mutant proteins.

Identification of functional SNPs in Conserved Region
To anticipate evolutionarily conserved amino acids, the ConSurf online tool was employed. The important structural and functional amino acid residues are conserved so, ConSurf online server has been used to forecast conserved amino acids. It is based on a Bayesian method, and conservation is given as a score ranging from 1 to 4, 5 to 6, and 7 to 9 for conserved. The color scheme and conservation of the supplied protein's result were displayed with the use of the color score (Mayrose et al., 2004) .

Prediction of 3D structures by using computational tools
The three-dimensional structures of proteins provide the better understanding of function specially its binding affinity with substrates even with or without mutations etc. so, there are some experimental approaches, such as X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR), to explore the 3D structures of protein but these procedures are both costly and time-consuming. The discovered proteins structures has been storing in the Protein Data Bank (PDB) (Engel et al., 2003)

Refinement of protein model
Local stereochemistry must be improved in tertiary structures acquired experimentally or computationally. The bind angel, torsion angel, and ring side chains should all be within statistically predicted ranges, internal conflicts should be avoided, and the distribution of hydrophobic and polar residues must be as predictable (Celestino, 2015), so the LocPREFMD program was used to revise the predicted structure of MDM2. By improving the structure, the global and local structural quality of the protein model may be enhanced (Heo et al., 2013) .

Verification of 3D structure by using different tools:
The Ramachandran Plot, verify 3D, and ERRAT tools were all utilized to validate the 3D structure of MDM2 proteins. The Verify3D tool was used to determine which parts of the protein had correct conformation or to identify misfolded regions since it assesses the structure at the molecular level (von Grotthuss et al., 2003). psy (φ,Ψ) angel showed by Ramachandran plot of proteins. The Ramachandran plot results show if the main chain torsion angel is stereo chemically viable (Holland et al., 1997).

The ERRAT method is used to verify proteins and assess the structural quality of proteins (Colovos & Yeates, 1993).

Site-directed mutagenesis with Insilco and RMSD calculation
In the software TRITON interfaced with MODELLER, a 3D structure of mutant protein was produced based on homologous modeling (Šali & Blundell, 1993). PROCHECK estimated the stereochemical quality of each model (Laskowski et al., 1996) ERRAT checked the environmental profile. The NOMAD and CHIMERA databases were used for minimization of mutant structures, also to calculate the RMSD of native and mutant structures (S. Chen et al., 2014).

Simulation of the effect of mutations on the MDM2 binding pocket

By comparing MDM2 protein to its CHIMERA template, binding residues were investigated and COACH has been used to verify the outcome. Simulation in motion The CABS-flex server was used to examine the MDM2 protein and its mutations (S. Chen et al., 2014).

RESULTS

Single nucleotide polymorphisms (SNPs) Dataset Retrieval:

MDM2 SNPS data was retrieved from a variety of databases, including the NCBI dbSNP database, Gene Cards, SNP500 cancer, Ensemble genome browser, and Uniprot. From the MDM2 gene variation database, only nsSNPS were chosen. A total of 36 SNPS were chosen from MDM2, with 12 non-synonymous (nsSNPS) being determined to be harmful. The detail of nsSNP is shown in Figure 3.1 and table 3.1

SNP distribution in the MDM2 gene region of humans

Analysis of Non-synonymous SNPs

For the analysis of nsSNPS, send to Sorting Intolerant from Tolerant to forecast the tolerance index (TI) of altered proteins (SIFT). By matching homologous sequences and evaluating conserved amino acids, SIFT determines the TI (0-1). The TI value was used to assess the functional impact of a given amino acid alteration; the greater the TI, the smaller the functional impact.

Table 1: SIFT study of nsSNPs in the human MDM2 protein's coding region

S.NO	SNPs ID	Mutation	TI	S.NO	SNPs ID	Mutation	TI
1	rs1246173207	E31Q	0.01	19	rs867680779	P322L	0
2	rs759244097	K76M	0	20	rs764918809	C328R	0
3	rs746439458	E31G	0.03	21	rs1357831520	L331F	0
4	rs764034976	G89E	0	22	rs1451286579	L331R	0
5	rs753663917	F92C	0	23	rs773519801	R332C	0
6	rs761546875	G93R	0	24	rs763077439	R332H	0
7	**rs1555187172**	**R189C**	**0**	**25**	**rs759644422**	**S397R**	**0.01**
8	rs1442753433	S192F	0	26	rs767686947	D399H	0.03
9	**rs774688124**	**S202N**	**0.03**	**27**	**rs1293117772**	**S401F**	**0.01**
10	**rs1162032400**	**L274F**	**0.01**	**28**	**rs1173684536**	**S408G**	**0.01**
11	rs1385029841	L274P	0.04	29	rs754521846	S412I	0.03
12	rs755429424	S304P	0	30	rs201788800	Q414L	0
13	rs748459725	D307G	0	31	rs201686188	Q414H	0
14	rs1161921915	C314Y	0	32	rs765555199	C444Y	0
15	rs577942747	P319T	0	33	rs1475420873	G459D	0
16	rs1353702185	P319R	0	34	rs1044633873	N478H	0
17	rs78419579	L321F	0	35	rs550783815	C481S	0
18	rs1221150148	P322S	0	36	rs777525974	V483L	0.03

Variants with a tolerance index of less than 0.05 are regarded harmful, whilst others are classified tolerant.

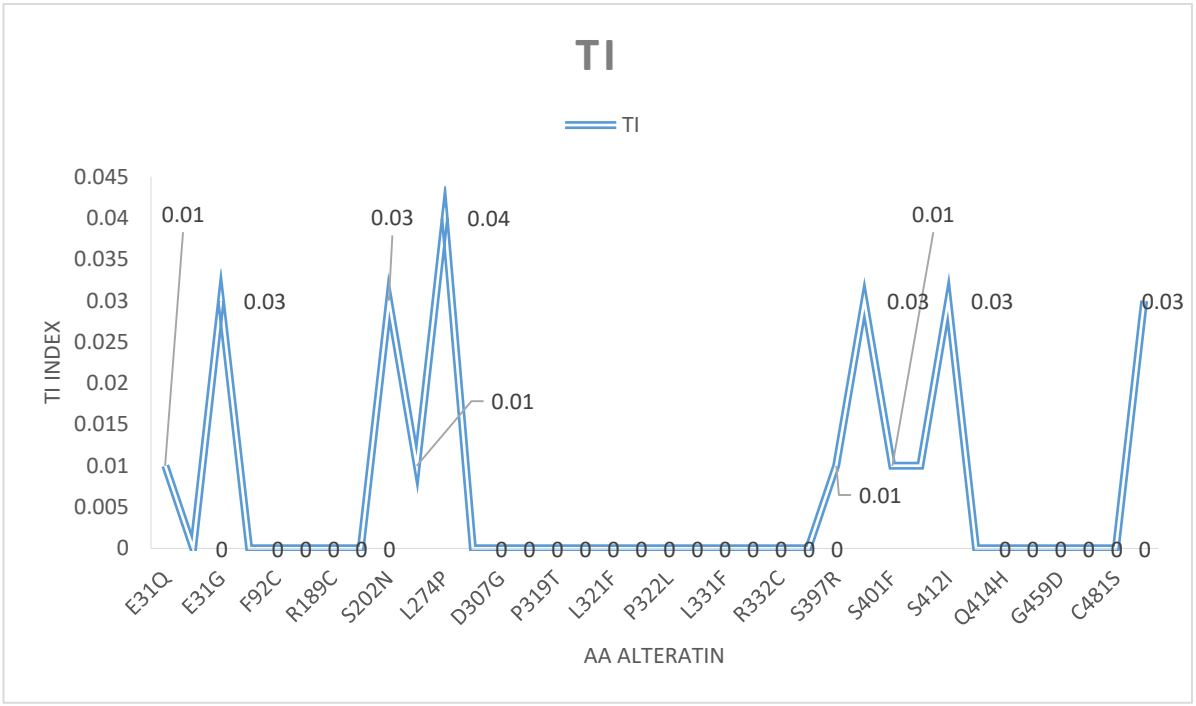

Figure 1: The SIFT technique was used to predict potential amino acid changes in human MDM2

3.2.2 Functional Modification of coding nsSNPS Prediction

The Poly Phen2 server was used to predict the structural and functional effects of nsSNPS on MDM2. The PolyPhen2 server analyzed the supplied protein's homologous sequence and 3D structure to known structures. The likelihood score was calculated based on similarity and the score reflected the impact of mutation on the protein. The server received 36 nsSNPS, all of which were judged to be probably harmful, but none of which were found to be potentially harmful.

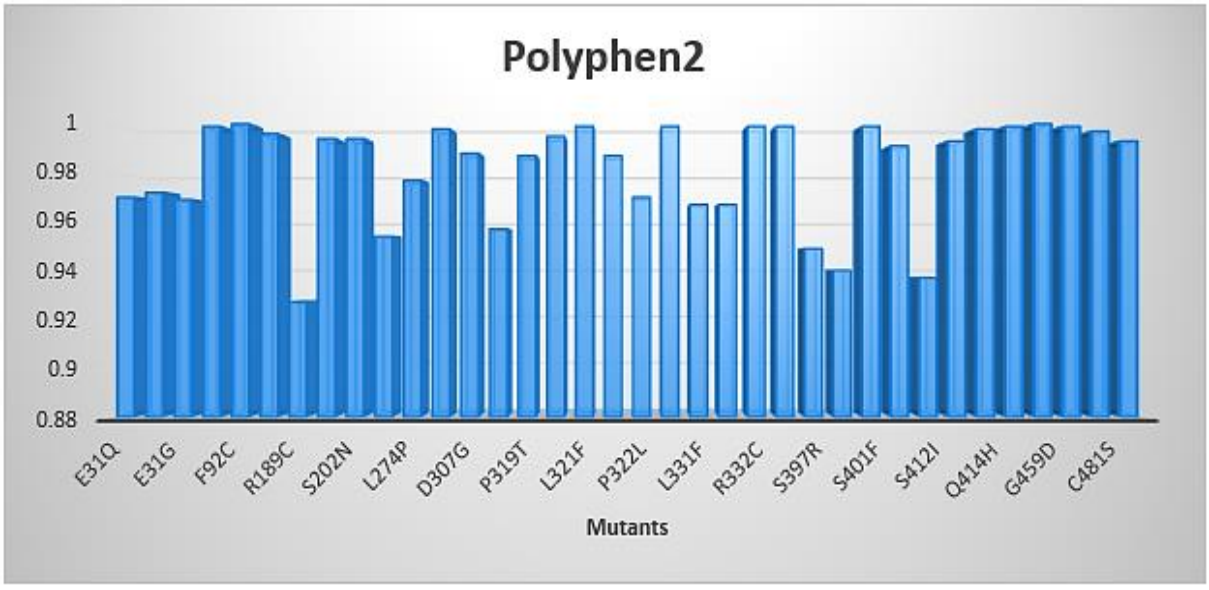

Figure 2: PolyPhen2 alterations analysis in the human MDM2 protein.

The PolyPhen2 algorithm was used to predict potential amino acid alterations in the human MDM2 protein.

The SNPS then transmitted to analyzer of nsSNPS, PhD SNP, P-MUT, PROVEAN, MutPred, and SNPS&GO to ensure the validity of SIFT and PolyPhen2 findings. SIFT and PolyPhen2 both judged SNPS to be harmful, thus they were chosen for further investigation.

Table 2: Polyphen2 server identified the following SNPs to be substantially harmful

SNPs ID	Mutation	Polyphen2	Effect	SNPs ID	Mutation	Polyphen2	Effect
rs1246173207	E31Q	0.97	Probably Damaging	rs867680779	P322L	0.97	Probably damaging
rs759244097	K76M	0.972	Probably damaging	rs764918809	C328R	0.999	Probably damaging
rs746439458	E31G	0.969	Probably damaging	rs1357831520	L331F	0.967	Probably damaging
rs764034976	G89E	0.999	Probably damaging	rs1451286579	L331R	0.967	Probably damaging
rs753663917	F92C	1	Probably damaging	rs773519801	R332C	0.999	Probably damaging
rs761546875	G93R	0.996	Probably damaging	rs763077439	R332H	0.999	Probably damaging
rs1555187172	R189C	0.927	Probably damaging	rs759644422	S397R	0.949	Probably damaging
rs1442753433	S192F	0.994	Probably damaging	rs767686947	D399H	0.94	Probably damaging
rs774688124	S202N	0.994	Probably damaging	rs1293117772	S401F	0.999	Probably damaging
rs1162032400	L274F	0.954	Probably damaging	rs1173684536	S408G	0.991	Probably damaging
rs1385029841	L274P	0.977	Probably damaging	rs754521846	S412I	0.937	Probably damaging
rs755429424	S304P	0.998	Probably damaging	rs201788800	Q414L	0.993	Probably damaging
rs748459725	D307G	0.988	Probably damaging	rs201686188	Q414H	0.998	Probably damaging
rs1161921915	C314Y	0.957	Probably damaging	rs765555199	C444Y	0.999	Probably damaging
rs577942747	P319T	0.987	Probably damaging	rs1475420873	G459D	1	Probably damaging
rs1353702185	P319R	0.995	Probably damaging	rs1044633873	N478H	0.999	Probably damaging
rs78419579	L321F	0.999	Probably damaging	rs550783815	C481S	0.997	Probably damaging
rs1221150148	P322S	0.987	Probably damaging	rs777525974	V483L	0.993	Probably damaging

Table 3: Identification of deleterious mutation by different software

variant ID	allele change	A.A change	SNP&GO	PHD SNP	PROVEAN	PANTHER
rs1246173207	G/C	E31Q	Neutral 9	Neutral 6	Neutral -1.148	Neutral 2
rs746439458	A/G	E31G	Neutral 7	Neutral 5	D -2.825	Disease 1
rs759244097	A/T	K76M	Neutral 4	Neutral 1	D -3.317	Disease 6
rs764034976	G/A	G89E	Disease 7	Disease 8	D -6.932	Disease 7
rs753663917	T/G	F92C	Disease 0	Disease 3	D -3.961	Disease 7
rs761546875	G/C	G93R	Disease 3	Disease 5	D -5.932	Disease 6
rs1555187172	C/T	R189C	Neutral 6	Neutral 4	D -2.757	Disease6
rs1442753433	C/T	S192F	Neutral 8	Neutral 6	D -2.519	Disease4
rs774688124	G/A	S202N	Neutral 5	Neutral 3	Neutral -0.888	Disease 4
rs1162032400	C/T	L274F	Neutral 9	Neutral 8	Neutral -1.402	Neutral 1
rs1385029841	T/C	L274P	Neutral 5	Neutral 2	Neutral -1.661	Disease 1
rs755429424	T/C	S304P	Disease 0	Disease 2	D -3.166	Disease 5
rs748459725	A/G	D307G	Disease 1	Disease 4	D -5.561	Disease 5
rs1161921915	G/A	C314Y	Disease 5	Disease 6	D -8.855	Disease 9
rs577942747	C/A/T	P319T	Neutral 7	Neutral 5	Neutral -2.353	Disease 2
rs1353702185	C/G	P319R	Neutral 9	Neutral 6	D -3.664	Neutral 5
rs78419579	C/G/T	L321F	Neutral 8	Neutral 2	Neutral -2.265	Neutral
rs1221150148	C/T	P322S	Neutral 7	Neutral 2	D -2.962	Disease 0
rs867680779	C/T	P322L	Neutral 7	Neutral 3	D -4.539	Neutral 3
rs764918809	T/C	C328R	Disease 7	Disease 6	D -9.428	Disease 9
rs1357831520	C/T	L331F	Neutral 2	Disease 0	D -2.972	Disease 4
rs1451286579	T/G	L331R	Disease 5	Disease 6	D -4.537	Disease 8
rs773519801	C/T	R332C	Neutral 2	Neutral 7	D -6.047	Disease 9
rs763077439	G/A	R332H	Disease 5	Disease 6	D -3.798	Disease 8
rs759644422	T/G	S397R	Neutral 9	Neutral 8	Neutral -2.204	Disease 2
rs767686947	G/C	D399H	Neutral 9	Neutral 9	D -3.055	Disease 1
rs1293117772	C/T	S401F	Neutral 9	Neutral 8	D -3.520	Disease 3
rs1173684536	A/G	S408G	Neutral 9	Neutral 7	Neutral -2.184	Disease 2
rs754521846	G/T	S412I	Neutral 9	Neutral 5	D -3.523	Neutral 0
rs201788800	A/T	Q414L	Neutral 5	Neutral 6	D -3.104	Disease 6
rs201686188	A/C/G	Q414H	Neutral 4	Neutral 5	Neutral -2.278	Disease 4
rs765555199	G/A	C444Y	Disease 6	Disease 7	D -9.915	Disease 9
rs1475420873	G/A	G459D	Disease 6	Disease 8	D -6.097	Disease8
rs1044633873	A/C	N478H	Neutral 10	Neutral 8	Neutral -2.432	Neutral 4
rs550783815	G/C	C481S	Disease 3	Disease 3	D-8.744	Disease 7
rs777525974	G/T	V483L	Neutral 6	Neutral 3	Neutral -1.492	Neutral 1

*** Deleterious (D)**

SNP-effect

Table 1: SNPs influence nsSNPS analyses in the human MDM2 protein's coding area

Variant id	D tango score	Aggregation tendency	D waltz	Amyloid propensity	dLIMBO	Chaperone binding tendency
rs764034976	-0.01	Not affect	0.04	Not affect	0	Not affect
rs753663917	0	0	0	Not affect	0	Not affect
rs761546875	0.90	Not affect	-0.15	Not affect	-0.02	Not affect
rs755429424	0	Not affect	0	Not affect	0	Not affect
rs748459725	23.02	Not affect	0.06	Not affect	0	Not affect
rs1161921915	0	Not affect	0.02	Not affect	0	Not affect
rs764918809	0.90	Not affect	-0.15	Not affect	0	Not affect
rs1451286579	0.90	Not affect	-0.16	Not affect	0	Not affect
rs763077439	1.93	Not affect	0.23	Not affect	0	Not affect
rs765555199	2.88	Not affect	0.34	Not affect	-117.46	Decrease
rs1475420873	-0.03	Not affect	0	Not affect	0	Not affect
rs550783815	-0.05	Not affect	0	Not affect	0	Not affect

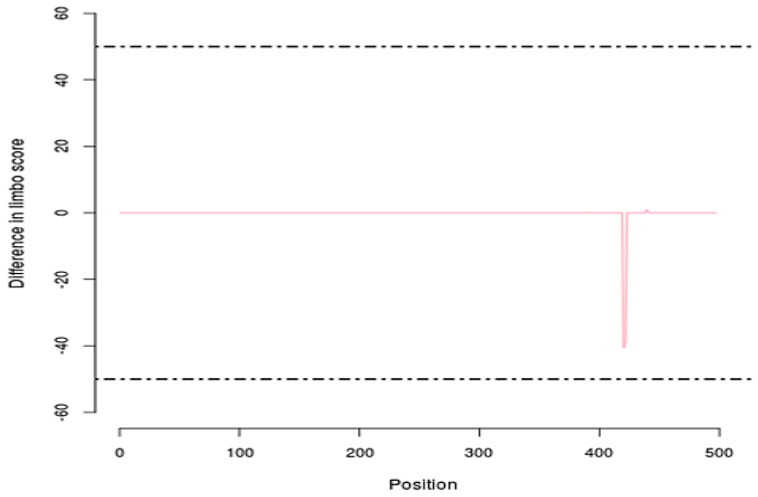

Figure 3: Difference in limbo score analysis of MDM2 protein

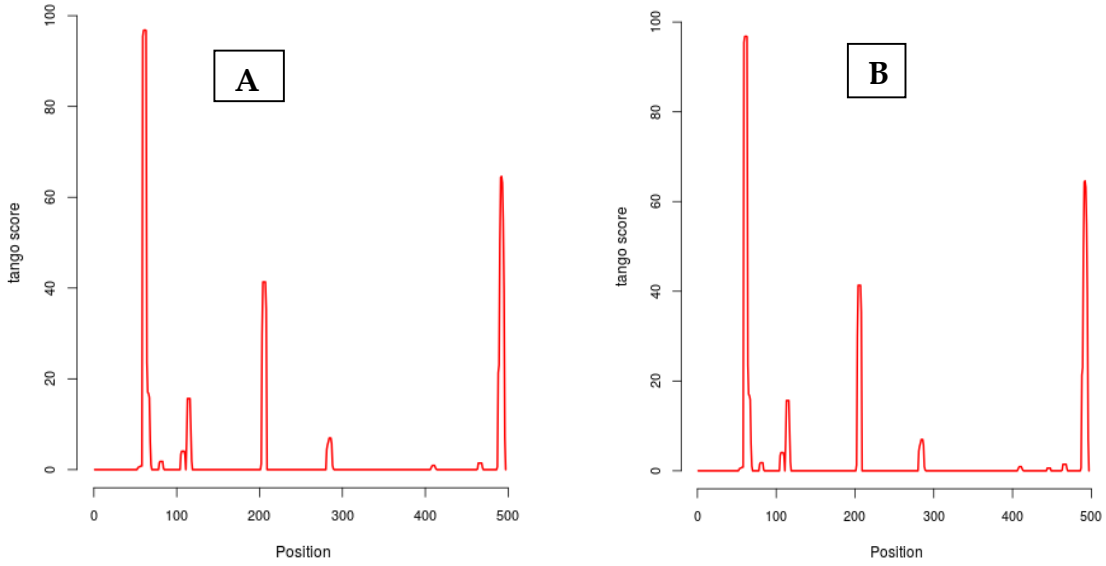

Figure 1: TANG0 aggregation differs between WT and mutant C444Y

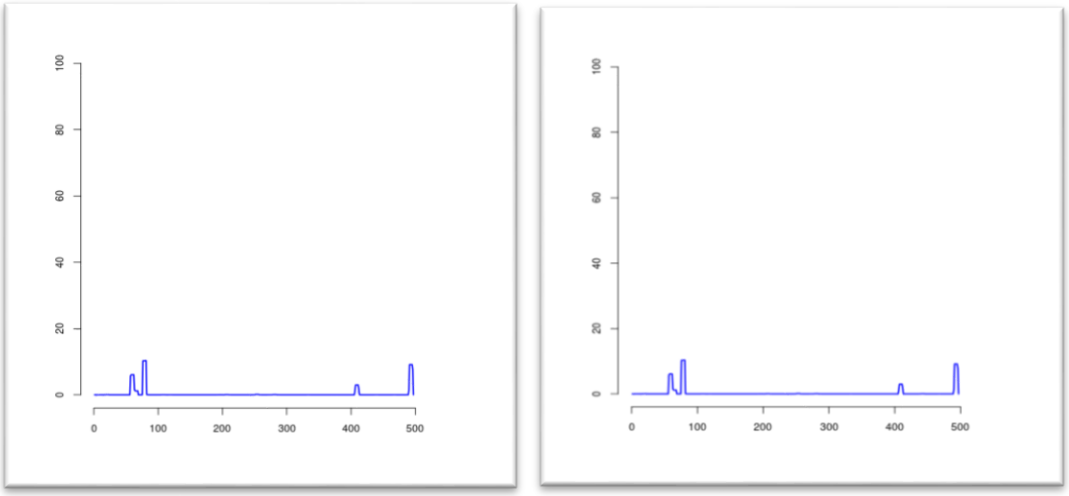

Figure 2: WALTZ amyloid propensity differences between WT and mutant (C444Y)

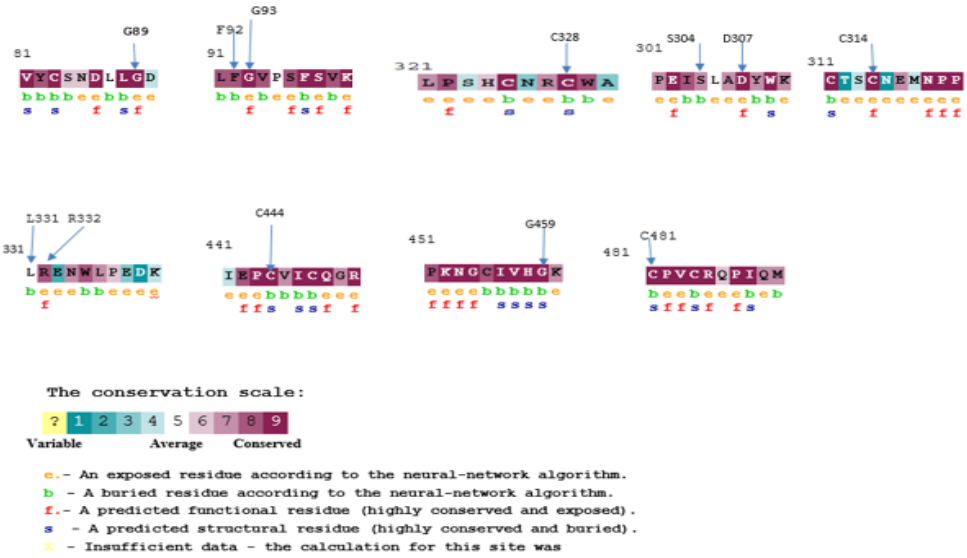

Figure 3: The ConSurf result is display in a colour scheme, with the conservation scale indicating the conservation status of each MDM2 amino acid.

Table 5: Amino acid conservation profile in MDM2 that coincides in location with high-risk ns SNPS

SnpID	Residue positions	Residue	Color	Buried or exposed	Function/structure
rs764034976	89	G	9	E	F
rs753663917	92	F	8	B	I
rs761546875	93	G	9	E	F
rs755429424	304	S	7	B	I
rs748459725	307	D	9	E	F
rs1161921915	314	C	9	E	F
rs764918809	328	C	9	B	S
rs1451286579	331	L	5	B	I
rs763077439	332	R	9	E	F
rs765555199	444	C	9	B	S
rs1475420873	459	G	9	B	S
rs550783815	481	C	9	B	S

****CS: conservation score (1-4=variable, 5 = average and 6-9 = conserved); (f): predicted functional site; (s): predicted structure site**.

Effects of mutations on protein stability

Table 2: For a subset of nsSNPs, MUTANT and Mutpro predictions were used.

position	WILD	Mutant	Ddg VALUE	I-MUTANT	MUPRO
89	G	E	-0.85	DECREASE	decrease
92	F	C	-0.54	DECREASE	Decrease
93	G	R	-0.31	DECREASE	Decrease
304	S	P	-1.00	DECREASE	Decrease
307	D	G	-0.66	INCREASE	Decrease
314	C	Y	0.34	INCREASE	Decrease
328	C	R	-1.24	DECREASE	Decrease
331	L	R	0.08	INCREASE	Decrease
332	R	H	-0.96	DECREASE	Decrease
444	C	Y	-1.12	DECREASE	Decrease
459	G	D	-0.40	DECREASE	Decrease
481	C	S	-1.72	DECREASE	decrease

****DDG (change in free energy variation upon mutation): DG (delta G; New protein) –DG (delta G; Wild type) in Kcal/mol,DDG<0: Decrease Stability, DDG>0: Increase stability**

Evaluation of the tertiary structure

The tertiary structure of MDM2 protein was predicted using SWISS MODEL program. However, no trustworthy template was identified, even after all templates were verified and determined to be smaller than the minimal value after the structure was predicted using the Ab initio approach using trRosetta.

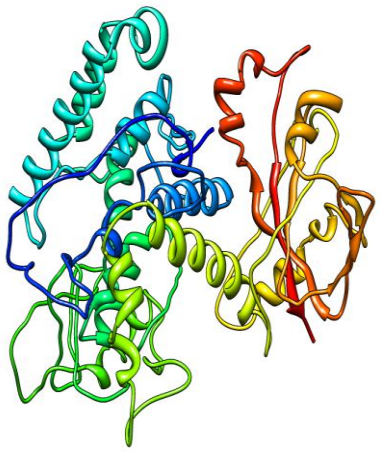

Figure 4: predicted structure of MDM2 human onco protein

Verification of model

After modelling, locPREFMD refines the 3D structure. The revised models were then tested on a variety of servers based on various character traits. Errate compared the model to a highly refined structure and anticipates the degree of error function based on residue location using a statistical technique. Verify 3D predicted the amino acid compatibility in a protein sequence as well as its 3D structure. The structure is deemed correct tertiary structure if the score of 80 percent amino acids is >-0.2. Otherwise, it is not considered accurate. In our MDM2 predicted model and altered structure, about 84.14 percent of residues had a 3D-1D score of >- 0.2.

The protein conformation was visually depicted using a Ramachandran plot In the Ramachandran plot, certain regions were designated based on the (,) angles group mirror of the three types of secondary structure. Different colour schemes in different portions of the plan indicated the most beneficial, authorised, generously allowed, and restricted sites. 90 percent of residues in core areas

are required for optimum structure. The majority of MDM2 residues are found in the core area, according to our findings. For phi-psi values, this orientation is appropriate. 94.545 percent of residues were found in the preferred zone, whereas 4.242 percent were found in the authorized region. A total of 1.212 percent was found in the outlier zone, with no residue found in the forbidden area.

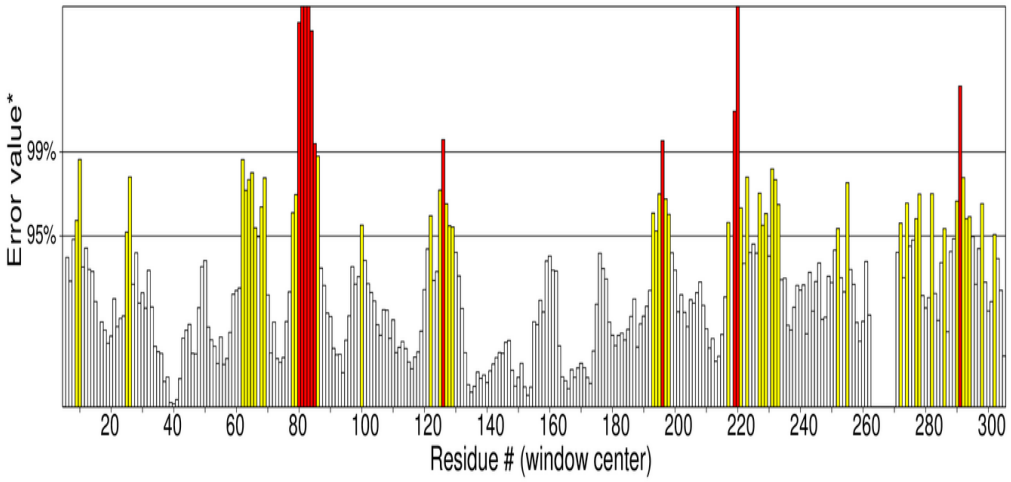

Figure 5: An error evaluation plot of MDM2 human predicted protein (Regions of the structure that can be rejected at a 95% confidence level are highlighted in silver; areas that can be rejected at a 99 percent confidence level are represented in yellow; and regions that may be rejected at a 1% confidence level are shown in red).

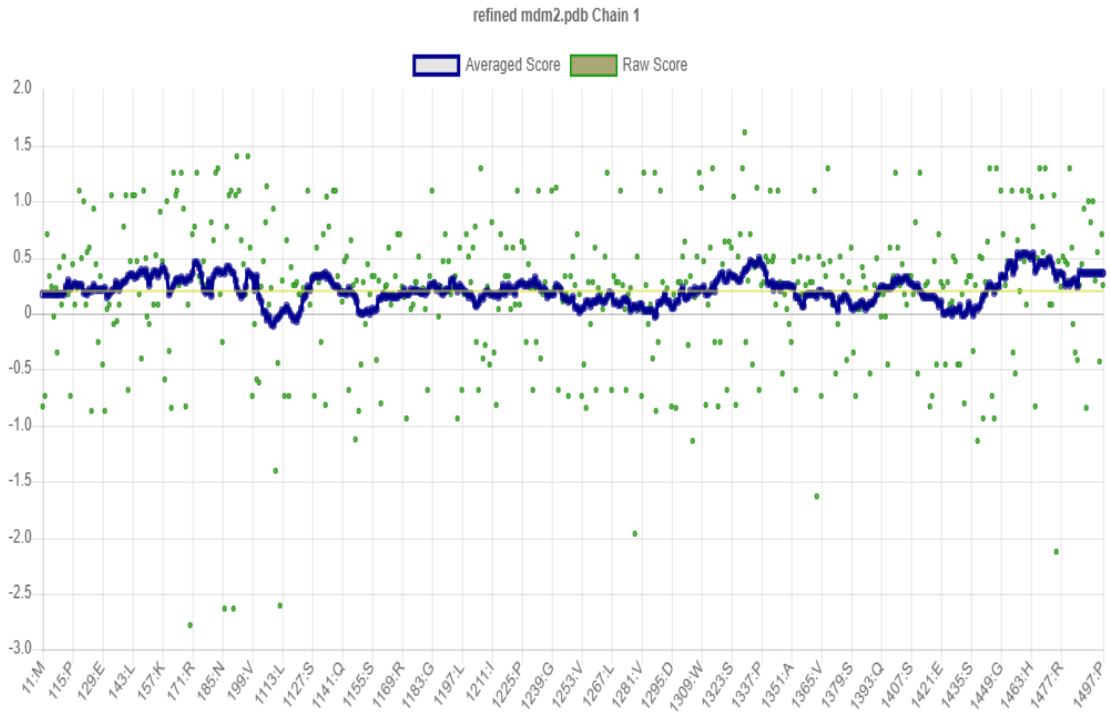

Figure 6: The 3D-ID score is visually displayed in the verify3D

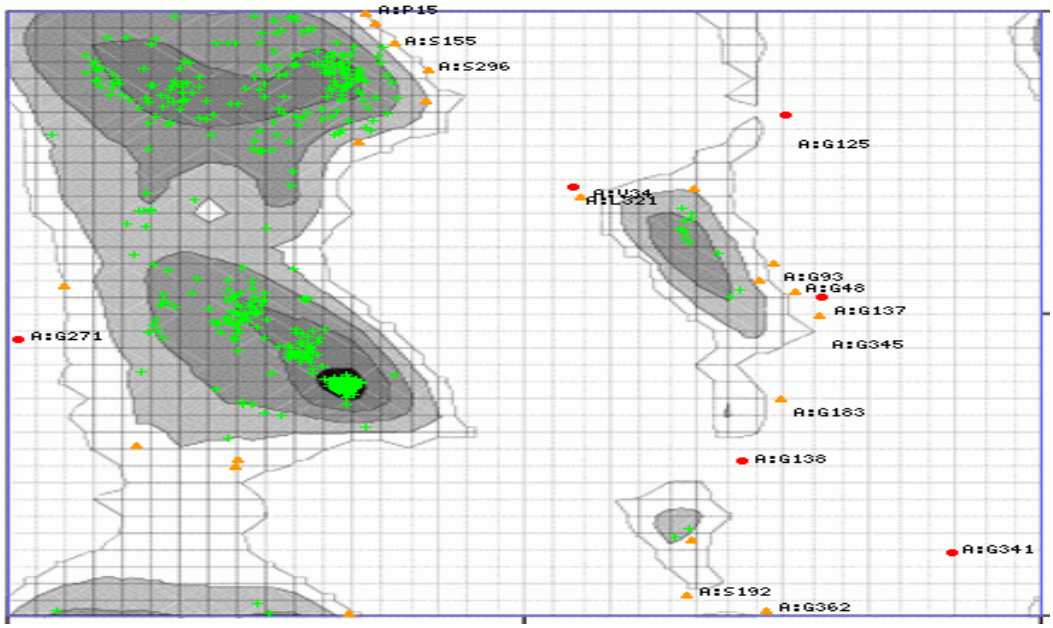

Figure 7: The favoured, allowed, and "generously allowed" areas as specified by ProCheck are shown by the silver yellow, light silver, and beige silver regions in the Ramachandran plot.
In-silico site directed mutagenesis and RMSD calculations

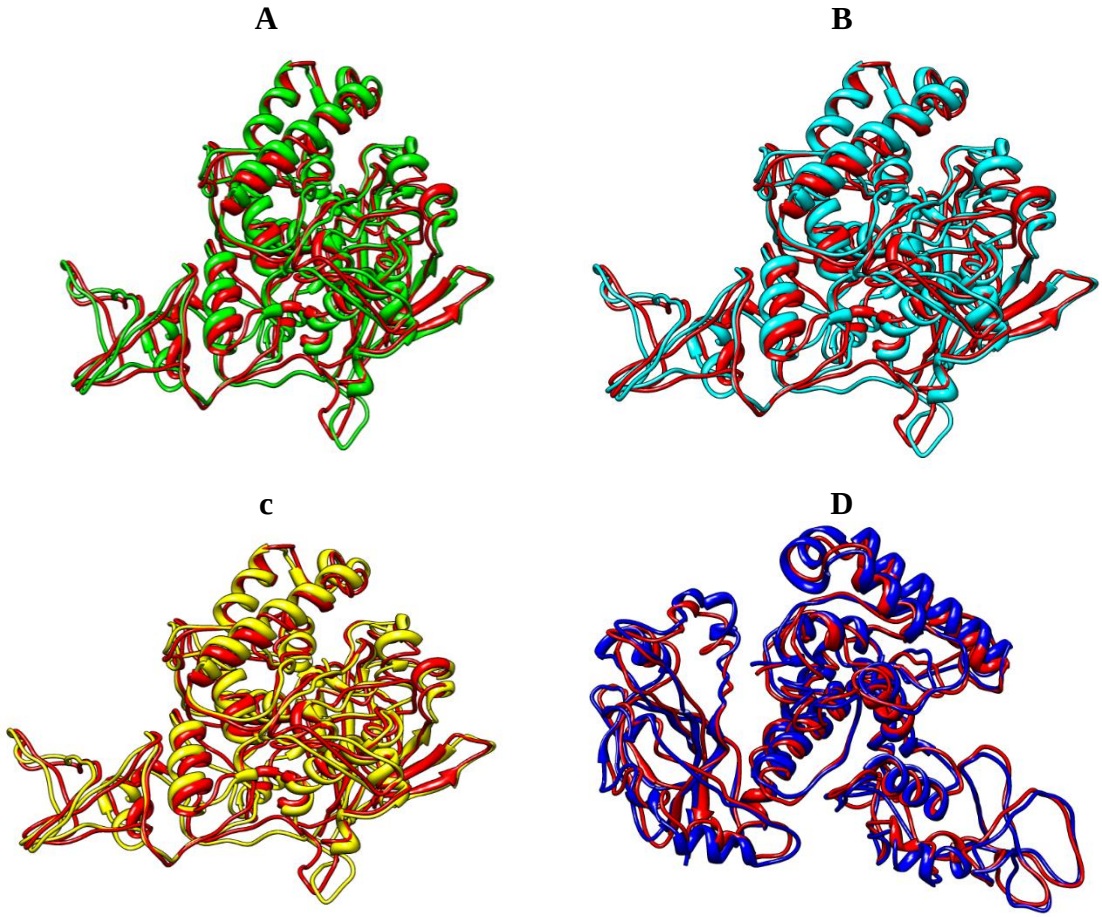

Figure 8: (A) Wild-type and mutant protein structures superimposed (G89E). MDM2 normal protein structure (red) superimposed on mutant protein structure (green) (B) Wild-type and mutant protein structures superimposed (F92S). MDM2 (red) natural protein structure superimposed with mutant Cyaan chrom. (C) Wild-type and mutant protein structures superimposed (G93R). MDM2

natural protein structure (red) superimposed on mutant yellow hue. (D) Wild-type and mutant (G459D) protein structures superimposed. MDM2 normal protein structure (red) superimposed on mutant protein structure (blue).

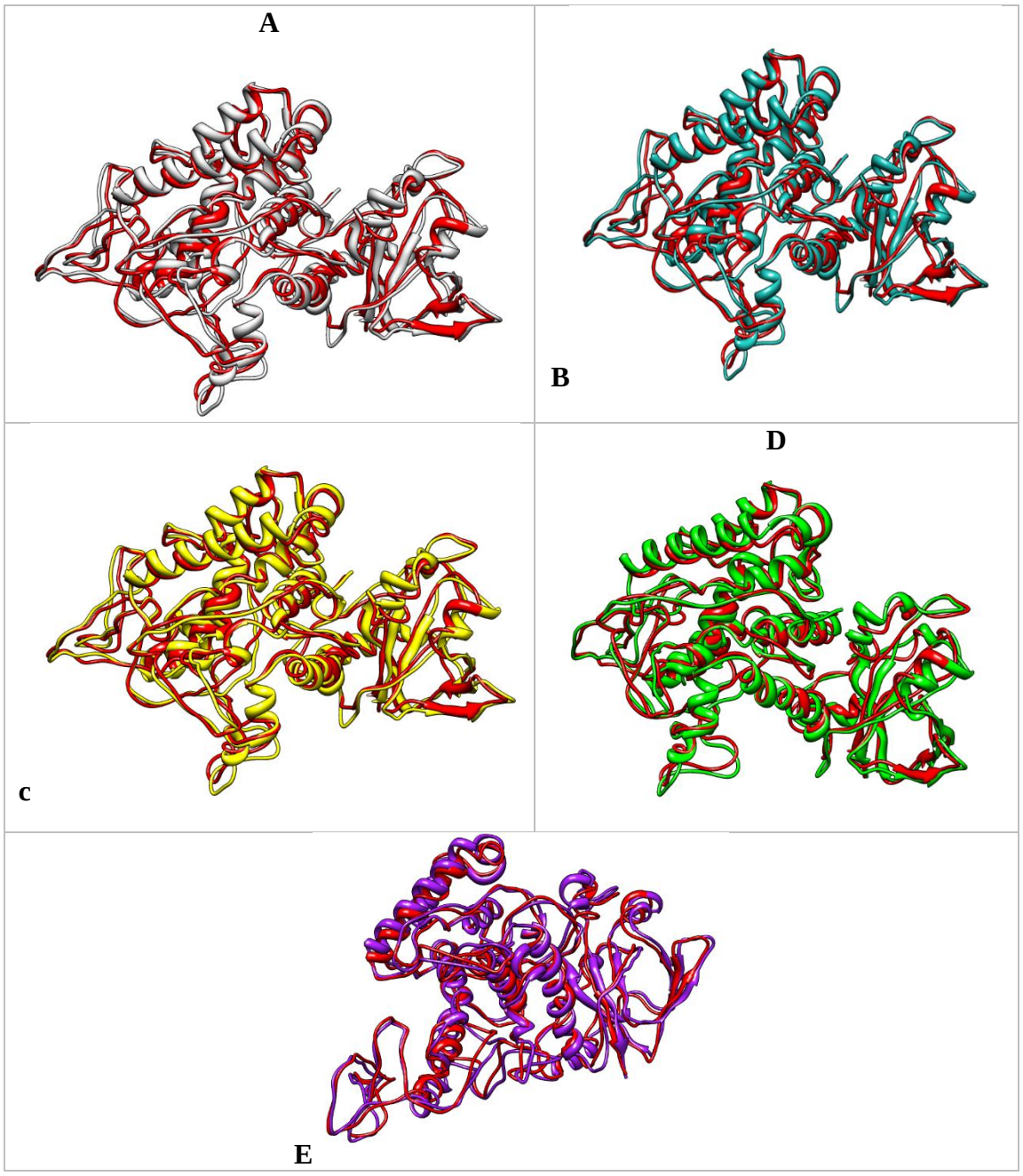

Figure 9: (A) Structures of wild-type and mutant proteins superimposed (R332H). MDM2 natural protein structure (red) superimposed on mutant structure (light grey). (B) Wild-type and mutant protein structures superimposed (C444Y). MDM2 natural protein structure (red) superimposed with mutant light marine green hue. (C) Wild-type and mutant protein structures superimposed (C481S). MDM2 natural protein structure (red) superimposed on mutant yellow hue**.** (D) Wild-type and mutant protein structures superimposed (S304P). MDM2 normal protein structure (red) superimposed on mutant protein structure (green). (E)Wild-type and mutant protein structures superimposed (C328R). MDM2 natural protein structure (red) superimposed over mutant purple hue.

Table 3: RMSD (A) values for wild type and mutant protein models

WILD /Mutant	Main chain RMSD Å	C-Alpha RMSD Å
G89E	55.920	55.953
F92C	41.775	41.781
G93R	47.998	48.006
S304P	32.023	32.019
C328R	71.420	71.412
L331R	53.311	53.339
C444Y	85.072	85.066
G459D	55.040	55.044
C481S	38.049	38.068

MD simulations

Over a 100ns time period, RMSD describes the backbone analysis of C atom dynamics of the total 10 samples one verses 10 models. (Mdm2normal) (G89E), (F92C), (G93R), (S304P), (C328R), (R332H), (C444Y), (G459D), (C481S) have been reported in our study to have a static with minor fluctuation with time interval. The (black) color code presenting normal model system starting with initial trajectories of RMSD 2Å at 1ns followed by abrupt changes in RMSD values i.e. 6.5Å at 27 ns showed that protein has high loop region. Cα was noticed for a time periods of 100ns against complex (Mdm2 normal) (G89E), (F92C), (G93R), (S304P), (C328R), (R332H), (C444Y), (G459D), (C481S).

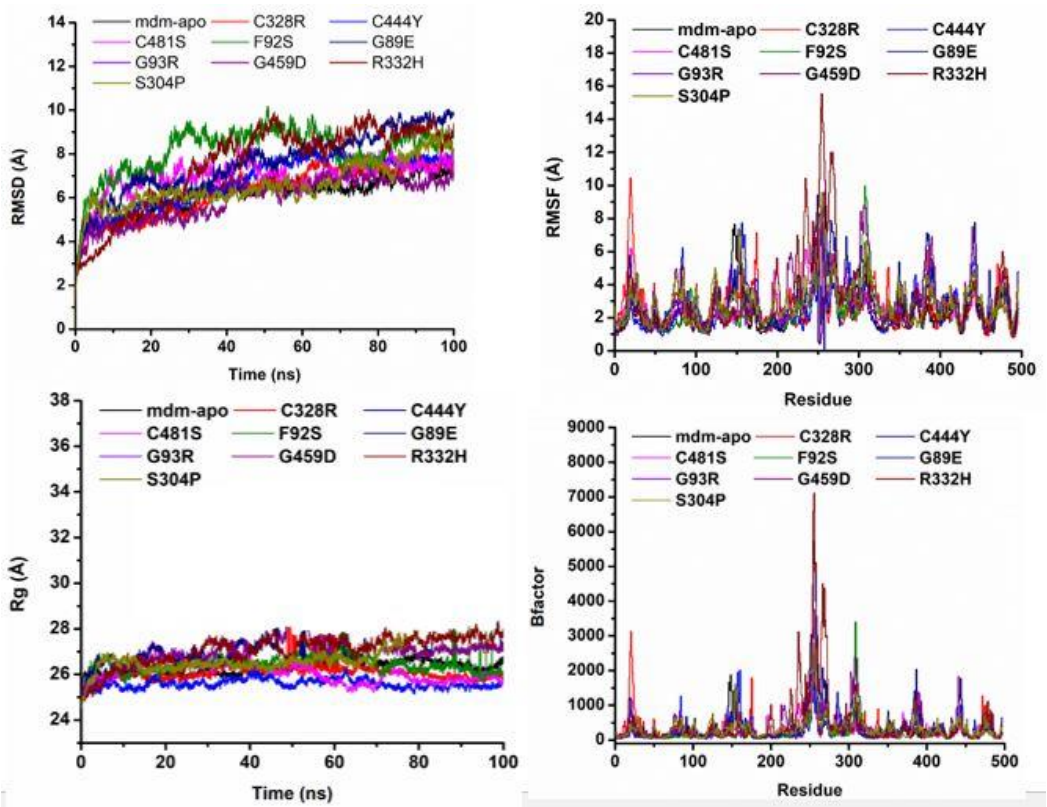

Figure 16: Depicting RMSD, RMSF, and Radius of gyration and beta-factor of all the 10 samples system. (In correspondence with the RMSD graph, noticeable structural changes occurred at a time of 17ns, 28ns at residue leu73 with high mean RMSF value 4Å for sample (S304P), 3.5Å for complex (C481S) and 3.8Å for (C444Y). The same trend has been followed at Ile100 with RMSF score 2.3Å, 2Å, 2.1Å against (G93R, D307G, C314Y, L331R and R332H) respectively. Other residues with high RMSF values include Lys150, Tyr180, Ile270 and Lys335 at different time intervals in all samples).

Hydrogen bond analysis

	B

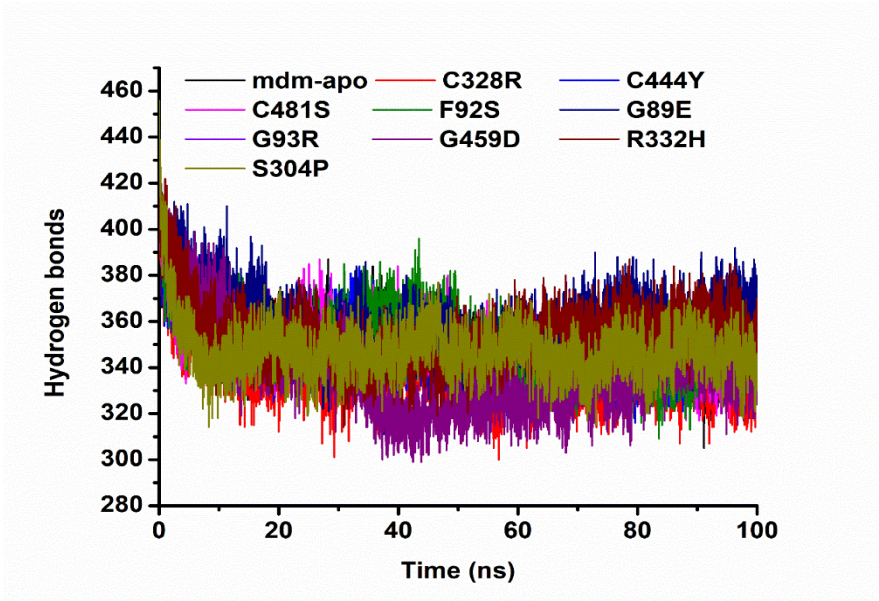

Figure 17: Presenting the number of hydrogen bonds of all 10 samples with time interval of 100ns. (The patterns of hydrogen bonds for all three complexes were exhibited in each frame within three frames to examine the degree of intermolecular interaction over the simulation duration. All 10 samples have showed the highest number of hydrogen bonds with target protein, indicating strong interactions. Figure 0.25 depicts the number of hydrogen bonds in all three complexes).

SASA analysis

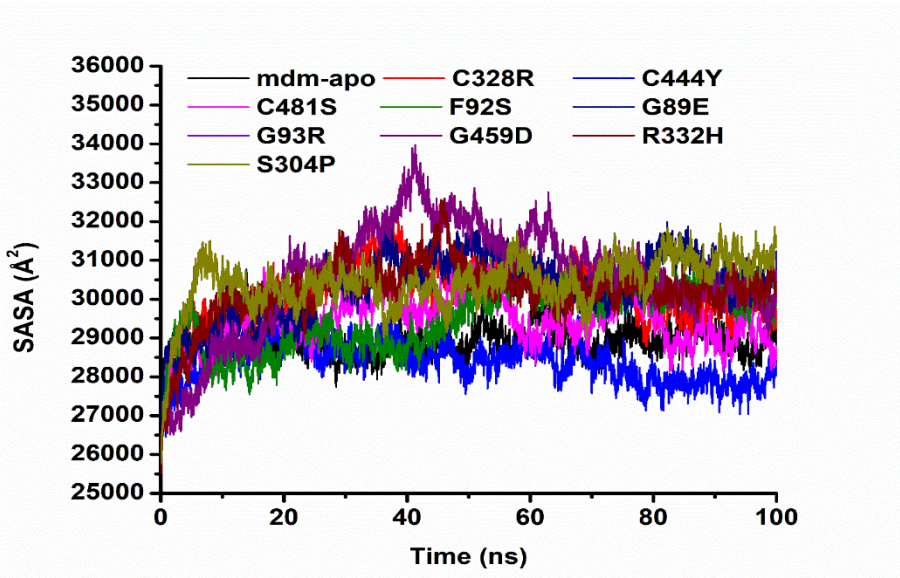

Figure 18: Graph presenting the coulomb, vdW and GB binding energies of all the 10 samples system. (The binding energies and components of samples (1234567890) systems investigated were compared and contrasted using prime MM-GBSA calculations. The ΔG binding of all systems is quite favorable i.e., -33.9, -38.8 and -49.2 respectively which isn't show that much variation. All the systems exhibit more favorable Coulombic interactions, resulting in lower binding energies. As a result, their preferred Coulombic interactions adequately compensate for their electrostatic solvation energy of Generalized Born. This is due to the Coulombic interactions, which are unfavorable. All the 10 samples show a clear pattern which is considerably the same to each other. Hence, this discovery attracted our curiosity, prompting us to investigate the conformational characteristics of

the systems. Indeed, the results of dynamics Simulations might give a logical explanation for the in vitro findings).

DISCUSSION

MDM2 is a protein found in the human body. It's an E3 ubiquitin binder that directs the p53 peptide for destruction by the proteasome. As a result, MDM2's most significant job is to regulate p53 activity; it is the best-documented specific inhibitor of p53. Because p53 possesses significant growth inhibitory actions, it is critical that it is kept at low levels in normal cells. In fact, in the lack of MDM2, p53 is fatal. Non-synonymous SNPS cause amino acid variations in a protein chain, which affects the protein's structure and function. The majority of genetic illnesses are caused by nsSNPs. Operationally critical areas are preserved under high selection pressure, and amino acid degeneracy is used to combat protein sequence variation. It's indeed difficult to differentiate between neutral and harmful SNPS, as well as to identify the amino acid that plays a significant role in illness development (C. M. Yates & Sternberg, 2013). Using different algorithms and information available in data sources, insilico analysis can assist discriminate between neutral and harmful SNPS. On the basis of structure and evolutionary information, a mutated peptide is evaluated, and pretty precise results are obtained (Papić *et al.*, 2020). Non-synonymous SNPS in the coding area can produce amino acid changes, which can alter protein function and increase disease susceptibility. Individual vulnerability to disease can be determined by distinguishing harmful nsSNPs from tolerant nsSNPs. It is not essential for all variants to have a significant functional impact; some may be tolerated well. nsSNPs associated to illnesses or other traits, on the other hand, frequently have molecular relevance (Ebben & Johnson, 2005). They have the potential to alter enzyme function, destabilise protein structures, and cause protein interactions to be disrupted. The identification and characterization of functionally relevant nsSNPs from those that are not is currently a key problem in molecular and population genetics (Jayaraman *et al.*, 2013).

SNPS in amidotrasferase may play a role in cancer by influencing purine synthesis. Various SNPS of MDM2 are accessible in databases, but it is necessary to identify those SNPS that cause protein structural disruption and, as a result, sickness. Using computational approaches, the SNP databases have been evaluated in this work to uncover SNPS that could potentially be harmful to MDM2. A search for nsSNPs in MDM2 yielded 138 hits.

lDs of mutants were uploaded to SIFT and PolyPhen2 services to investigate the functional implications of nsSNPs. SIFT identified 36 nsSNPs as non-tolerable, whereas PolyPhen2 identified 9 nsSNPs as potentially or certainly harmful. The discrepancies in results might be due to the server's usage of a different algorithm. As a consequence, the SNPS ID was presented to a variety of analysing tools, including SNPS&GO, P-Mut, Phd SNP, and PROVEAN, to ensure the authenticity of the results. In nearly all tools, at least 9 SNPS with rslD of rs764034976, rs753663917, rs761546875, rs755429424, rs764918809, rs763077439, rs765555199, rs1475420873 and rs550783815 were recognised as disastrous. There hasn't been a single study that has found a link between the harmful nsSNPs and any disease. As a result, to complement this result, the verification of these nsSNPs in any illness is required. Our findings also demonstrated that using a combination of different methods to prioritise putative functional nsSNPs can be quite effective. PolyPhen 2 and SIFT was claimed to have superior performance in finding functional nsSNPs than other Insilco techniques (Thusberg & Vihinen, 2009). As a result, even if a clear link between a genomic locus and a specific phenotype has been established, it's often not clear that the recognised DNA version has a causal relationship only with disease and that the section illustrates is not due to chain unbalance with the exact functional variant (Rickert *et al.*, 2003). The above-mentioned server can be utilized to discriminate between casual and non-casual associations between an nsSNP and the phenotypic of interest in this scenario.

The amyloid and facilitator binding SNPeffects were used to test the influence of SNPS on protein aggregation tendency. The SNPeffect results reveal that rs750794128 increases aggregation tendency with a dTANGO of 98.17, whereas rs772368910 decreases 'TANGO with a score of -95.83 but increases amyloid propensity with a dWALTZ score of 238.27. When nsSNPs were entered to the SNPeffect server to see whether they seemed to have any influence on the biochemical phenotype of MDM2, the preponderance of the variants were not anticipated to affect the protein's aggregation tendency, amyloid propensity, or chaperone interaction tendency. If these variations are capable of causing some harmful changes to MDM2, they are most likely not related to the protein's aggregating tendency, amyloidosis propensity, or chaperone binding tendency, as predicted by SNPeffect. The links among nsSNPs and their placement in protein complexes have been examined by several groups (Fernald *et al.*, 2011; Yue *et al.*, 2006).

Because the tertiary structure was not found in the PDB, 1gph was used to simulate the 3D structure of the protein and its mutation. 1. A (PDB id) as a template with a 45.39 percent resemblance. The mutation and wild type models were subjected to minimizing the objective function and RMSD calculations. Mutations, deletions, and insertions cause the mutant structure to differ from the native structure (Han *et al.*, 2006) and the deviation between the two structures is evaluated by their RMSD values which could affect stability and functional activity (Bhavaniramya *et al.*, 2019).

To investigate the differences in dynamics and durability of native and fatal mutations, a molecular dynamics (MD) simulation has been used to investigate the explicit solvent behavior of native and mutant structures. For the MDM2 gene variant, the RMSD value of the native structure and the mutant modelled structure were compared. A higher RMSD value suggests that the mutants' structural variation is greater. (C444Y) had the highest RMSD 85.072, the lowest was 32.023 for (S304P), while the remaining nine were in the range of in between the highest and lowest. Almost all the mutation deviates the greatest from the wild type structure, according to the RMSD value.

CONCLUSION

This study is very helpful to identify the genotypic and phenotypic association linked with nsSNPs of MDM2 gene. Protein stability was reduced by the nine deleterious SNPs identify in this study. Through computational approach the present study identifies the SNPs main targets of mutations. SNPs mostly occur at domain regions. SNPs regulate the MDM2. Further experiments needed on large scale on population base studies. Effect of SNPs on structure and functions of proteins need to investigate. Mutations in MDM2 of different diseased models could be comparing to notice to MDM2 role in disease.

REFERENCES

1. Acevedo, A., S. Alférez, A. Merino, L. Puigví, and J. Rodellar, 2016. Automatic recognition system of nucleated peripheral blood cell images. Paper presented at the Int. J. Lab. Hematol.**,** 38, 1-1.
2. Alogheli, H., G. Olanders, W. Schaal, P. Brandt, and A. Karlén, 2017. Docking of macrocycles: comparing rigid and flexible docking in glide. J. chem. linformation modeling, 57(2): 190-202.
3. Aydin, G., M. N. Paksoy, M. D. Orhan, T. Avsar, M. Yurtsever, and S. Durdagi, 2020. Proposing novel MDM2 inhibitors: Combined physics-driven high-throughput virtual screening and in vitro studies. Chem. Biol. Drug Des. 96(1): 684-700.
4. Bachmann, H., D. Molenaar, F. Branco dos Santos, and B. Teusink, 2017. Experimental evolution and the adjustment of metabolic strategies in lactic acid bacteria. FEMS Immunol Med Microbiol, 41(Supp_1): S201-S219.
5. Baeissa, H., G. Benstead-Hume, C. J. Richardson, and F. M. Pearl, 2017. Identification and analysis of mutational hotspots in oncogenes and tumour suppressors. Oncotarget, 8(13): 21290.

6. Bali, V., and Z. Bebok, 2015. Decoding mechanisms by which silent codon changes influence protein biogenesis and function. Int J Biochem Cell Biol, 64, 58-74.
7. Bhavaniramya, S., R.Vanajothi, S. Vishnupriya, M. S. Al-Aboody, R. Vijayakumar, and D. Baskaran, 2019. Computational characterization of deleterious snps in Toll-like receptor gene that potentially cause mastitis in dairy cattle. Biocatal Agric Biotechnol 19, 101151.
8. Bill, K. L. J., J.Garnett, I. Meaux, X. Ma, C. J. Creighton, S. Bolshakov, . . . B. C. Prudner, 2016 . SAR405838: a novel and potent inhibitor of the MDM2: p53 axis for the treatment of dedifferentiated liposarcoma. Clin cancer Res, 22(5): 1150-1160.
9. Biswas, S., E. Killick, A. G. Jochemsen, and J. Lunec, 2014. The clinical development of p53-reactivating drugs in sarcomas–charting future therapeutic approaches and understanding the clinical molecular toxicology of Nutlins. Expert Opin Investig Drugs, 23(5): 629-645.
10. Boeckmann, B., A. Bairoch, R. Apweiler, M.-C. Blatter, A. Estreicher, E. Gasteiger, . . . I. Phan, 2003 . The SWISS-PROT protein knowledgebase and its supplement trembl in 2003. Nucleic Acids Res, 31(1): 365-370.
11. Bond, G. L., W. Hu, and A. Levine, 2005. A single nucleotide polymorphism in the MDM2 gene: from a molecular and cellular explanation to clinical effect. Cancer Res, 65(13): 5481-5484.
12. Botstein, D., and N. Risch, 2003. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat genet , 33(3): 228-237.
13. Brister, J. R., D. Ako-Adjei, Y. Bao, and O. Blinkova, 2015. NCBI viral genomes resource. Nucleic acids Res, 43(D1): D571-D577.
14. Bromberg, Y., and B. Rost, 2007. SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic acids Res, 35(11): 3823-3835.
15. Budowle, B., and A. Van Daal, 2008. Forensically relevant SNP classes. Biotech
16. 44(5): 603-610.
17. Capriotti, E., R. Calabrese, and R. Casadio, 2006. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinform, 22(22): 2729-2734.
18. Capriotti, E., and P. Fariselli, 2017. Phd-snpg: a webserver and lightweight tool for scoring single nucleotide variants. Nucleic acids Res, 45(W1): W247-W252.
19. Cardon, L. R., and L. J. Palmer, 2003. Population stratification and spurious allelic association. The Lancet, 361(9357): 598-604.
20. Celestino, P. D. S. F. 2015. Oncogenic Mechanisms of Activation and Resistance of the type III Receptor Tyrosine Kinase family. École normale supérieure de Cachan-ENS Cachan,
21. Chakravarti, A. 1999. Population genetics—making sense out of sequence. Nat Genet, 21(1): 56-60.
22. Chang, F., S. Syrjänen, A. Tervahauta, and K. Syrjänen, 1993. Tumourigenesis associated with the p53 tumour suppressor gene. Br J Cancer , 68(4): 653-661.
23. Chen, H., G. Reed, J. Guardia, S. Lakhan, O. Couture, E. Hays, and N. Chandar, 2013. Vitamin D directly regulates Mdm2 gene expression in osteoblasts. Biochem Biophys Res Commun 430(1): 370-374.
24. Chen, J., S. M. Ng, C. Chang, Z. Zhang, J.-C. Bourdon, D. P. Lane, and J. Peng, 2009. P53 isoform Δ113p53 is a p53 target gene that antagonizes p53 apoptotic activity via bclxl activation in zebrafish. Genes Dev 23(3): 278-290.
25. Chen, S., A.-E. Rotaru, F. Liu, J. Philips, T. L Woodard, K. P. Nevin, and D. R. Lovley, 2014. Carbon cloth stimulates direct interspecies electron transfer in syntrophic co-cultures. Bioresour Technol 173, 82-86.
26. Chen, Y.-A., E. Yogo, N. Kurihara, T. Ohno, C. Higuchi, M. Rokushima, and K. Mizuguchi, 2019. Assessing drug target suitability using targetmine. F1000Res, 8.

27. Cheng, C.-H. C., and X. Zhuang, 2020. Molecular Origins and Mechanisms of Fish Antifreeze Evolution. Antifreeze Proteins Volume 1, 275-313.
28. Chitranshi, N., Y. Dheer, Wall, R. Vander V. Gupta, M. Abbasi, S. L. Graham, and V. Gupta, 2017. Computational analysis unravels novel destructive single nucleotide polymorphisms in the non-synonymous region of human caveolin gene. Gene Rep 6, 142-157.
29. Cissé, M. Y., S. Pyrdziak, N. Firmin, L. Gayte, M. Heuillet, F. Bellvert, . . . G. Arena, 2020. Targeting MDM2-dependent serine metabolism as a therapeutic strategy for liposarcoma. Sci Transl Med , 12(547).
30. Colombo, E., L. Spaccini, G. Negri, L. Volpi, C. Gervasini, D. Lazarevic, . . . L.Larizza, 2014. Viable phenotype associated with novel mutations of integrin 3, encoding a laminins receptor involved in structural and functional organization of lung, kidney and skin epithelia. Paper presented at the "Il Sequenziamento di Nuova Generazione in Genetica Umana e Medica"(Evento SIGU).
31. Colovos, C., and T. O. Yeates, 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein sci, 2(9): 1511-1519.
32. Consortium, U. 2015. Uniprot: a hub for protein information. Nucl Acids Res, 43(D1): D204-D212.
33. Coop, G., J. K. Pickrell, J. Novembre, S. Li, J. Kudaravalli, D. Absher, . . . J. K Pritchard, 2009. The role of geography in human adaptation. PLoS Genet 5(6): e1000500.
34. Cornillie, J., A. Wozniak, H. Li, Y. Gebreyohannes, J. Wellens, D. Hompes, . . . P. Schöffski, 2020. Anti-tumor activity of the MDM2-TP53 inhibitor BI-907828 in dedifferentiated liposarcoma patient-derived xenograft models harboring MDM2 amplification. Clin Transl Oncol, 22(4): 546-554.
35. Craig, A. J., J. Von Felden, T. Garcia-Lezana, S. Sarcognato, and A. Villanueva, 2020. Tumour evolution in hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol, 17(3): 139-152.
36. Danielson, M. L., and M. A. Lill, 2012 . Predicting flexible loop regions that interact with ligands: the challenge of accurate scoring. Proteins Struct.Funct, and Bioinform, 80(1): 246-260.
37. Dembla, V., N. Somaiah, P. Barata, K., Fu, S. Hess, F. Janku,. . . V. Subbiah, 2018. Prevalence of MDM2 amplification and coalterations in 523 advanced cancer patients in the MD Anderson phase 1 clinic. Oncotarget, 9(69): 33232.
38. Doss, C. G. P., R. Rajasekaran, C. Sudandiradoss, K. Ramanathan, R. Purohit, and R. Sethumadhavan, 2008. A novel computational and structural analysis of nssnps in CFTR gene. Genom Med, 2(1): 23-32.
39. Ebben, J. J., and A. C. Johnson, 2005. Efficiency, flexibility, or both? Evidence linking strategy to performance in small firms. Strateg Manag J, 26(13): 1249-1259.