Exploring the Functional Impact of KRAS Gene Mutations in Lung Cancer through Computational Analysis

Kizhakedathil Moni Philip Jacob[1], Chowdary Chandu[2], Koyye Rahul[2] and Vijaykumar Sudarshana Deepa[2*]
1. Department of Biotechnology, Bannari Amman Institute of Technology, Sathyamangalam, Tamil Nadu- 638401, INDIA
2. Department of Biotechnology, National Institute of Technology, Tadepalligudem, Andhra Pradesh- 534101, INDIA
*sudarshanadeepa@nitandhra.ac.in

Abstract

Lung cancer accounts for 12.3% of global cancer cases with variability across regions, races and genders. Lung cancer involves oncogene activation, such as mutated KRAS and tumor suppressor gene dysregulation like p53. KRAS mutations are commonly seen in pancreatic, colorectal and lung cancers. These mutations disrupt normal cell signaling and promote tumor growth. In this study, 63 KRAS mutations from 28,964 COSMIC database samples were analyzed.

The mutants were screened for their pathogenicity and nine highly deleterious mutations were identified. The stability of the mutants was checked with that of the wild type kras. Physiochemical properties were predicted using Protparam tools revealing that the properties of the mutants showed no major deviations from the wild type. The three-dimensional structures of the wild type and mutants were modelled using Robetta and the structures were validated using MolProbity server. The KRAS wild type and mutant structures were docked with Guanosine-5'-triphosphate (GTP) using Autodock Vina. The results indicated that the mutant D57N exhibited more affinity than the wild type and other mutants under study.

Keywords: Lung cancer, KRAS, mutation, docking.

Introduction

Lung cancer is the most prevalent of all the cancer cases[15]. Its incidence varies notably by geographic location, race and gender. Research indicates that women may be at an increased risk for lung cancer due to exposure to carcinogenic substances in tobacco smoke. Lifelong smokers face a lung cancer risk that is 20 to 30 times higher than that of non-smokers. While smoking rates in the United States are decreasing, China and Eastern Europe are experiencing a smoking epidemic which could lead to millions of new lung cancer diagnoses in this century[15,22,23]. Preventing lung cancer is feasible, as quitting smoking can substantially lower the risk.

However, the benefits of cessation are not immediate; there is usually a delay of about seven years before risk reduction becomes noticeable[22]. The development of lung cancer is associated with the activation of various oncogenes including c-Myc, mutated KRAS (which appears in 15%-20% of non-small cell lung cancers, especially adenocarcinomas, but is absent in small cell lung cancers) and the overexpression of EGFR, cyclin D1 and BCL2. These oncogenes are essential for cell growth, division and apoptosis, all of which contribute to the onset and advancement of the disease[9]. Most lung cancers express both telomerase RNA (hTR) and its catalytic subunit (hTERT), providing a pathway for cellular immortality.

Tumor suppressor genes (TSGs) frequently involved in lung cancer include p53 (mutated in about 90% of small cell lung cancers and 50% of non-small cell lung cancers), Retinoblastoma (Rb) (affected in roughly 90% of small cell lung cancers and about 20% of non-small cell lung cancers) and p16 (observed in over 50% of non-small cell lung cancers but less than 1% of small cell lung cancers). These TSGs play critical roles in controlling cell growth and division and their dysfunction significantly contributes to the development and progression of lung cancer[20].

KRAS is a key protein involved in cell signaling, crucial for regulating cell proliferation and other essential cellular functions. It acts as an important hub in the cell's communication network, receiving signals from upstream sources and transmitting activating signals to various downstream pathways including the mitogen-activated protein kinase (MAPK) pathway[8]. KRAS exists in two primary states: an inactive form bound to guanosine diphosphate (GDP) and an active form bound to guanosine triphosphate (GTP). This switching between states is vital for its role in cellular signaling[28].

When KRAS is active and bound to GTP, it can interact with and activate several effector proteins, such as RAF kinases, PI3K and RalGDS. This activation occurs when a guanosine exchange factor (GEF) displaces GDP in the nucleotide binding site, allowing GTP to bind, aided by the higher concentration of GTP relative to GDP in the cell[2]. Deactivation of active KRAS occurs through the hydrolysis of GTP to GDP. Although KRAS has a low intrinsic GTPase activity, this process is significantly enhanced by GTPase activating proteins (GAPs), which accelerate the hydrolysis reaction and facilitate the transition from the active to inactive state[2,12].

In pancreatic, colorectal and lung cancers, mutations in the KRAS gene, particularly missense mutations that result in the substitution of a single amino acid, are common. These

mutations have profound effects on KRAS functionality and are linked to cancer development[10]. In this study, we have explored the effects of the KRAS gene mutations on lung cancer computationally, predicting its properties and stability by comparing it with the wild type KRAS.

Material and Methods

Collection and Screening of KRAS mutants: Variants in the KRAS gene associated with lung cancer were sourced from the COSMIC database[25]. The FASTA sequence of the wild type KRAS was obtained from the UniProt database with the identifier P01116.

Pathogenicity analysis of the mutants: To study the effect of missense mutations on the pathogenicity of KRAS mutants, Meta-SNP server was used. Meta-SNP integrates four methods: PANTHER[26], PhD-SNP[4], SIFT[21] and SNAP[13] which help in filtering disease causing mutations from neutral mutations. The FASTA sequences of the wild type and mutant KRAS proteins were given as inputs and pathogenicity was determined for all 63 mutants.

Stability analysis of the mutants: In order to predict the stability of missense mutations, iStable, I-Mutant2.0 and MUpro servers were used[5-7].

Physicochemical Properties of the mutants: ProtParam tool[11] (www.expasy.org/tools/protparam.html) was used to predict the various physicochemical parameters of wild type and mutant proteins. Amino acid sequences were given as inputs. Properties such as theoretical isoelectric point, extinction coefficient, instability index, aliphatic index (AI) and GRAVY (Grand Average of Hydrophobicity) scores were predicted.

KRAS Structure Prediction: PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) was used to evaluate the protein secondary structures of wild type and mutant proteins[3]. The three dimensional structure of the wild type and mutant KRAS was predicted using Robetta server[24] using *Ab-intio* modelling. The generated models were validated using Molprobity[29]. The best models were used for molecular docking.

Molecular Docking and Binding analysis: Molecular docking was performed between guanosine triphosphate (GTP) and wild/mutant KRAS protein using AutoDock Vina[27]. Post-dock binding analysis of the protein-ligand complex was performed using protein ligand interaction profiler server (https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index)[1].

Results and Discussion

Over the past two decades, high-throughput sequencing has driven a major increase in Single Nucleotide Polymorphism (SNP) data, leading to extensive mutation databases with numerous variants still unclassified. Previous studies highlight the significance of non-synonymous SNPs in both simple and complex diseases, emphasizing the role of in silico prediction tools to understand how these mutations affect protein structure and function. Protein-protein interactions (PPIs) are essential for cellular functions like signaling and regulation. Mutations at PPI interfaces, such as single amino acid substitutions, can disrupt protein complex formation by altering shape, size, secondary structure, hydrophobicity and causing conformational changes. Additionally, the mutation of a smaller amino acid to a larger one can create inter-molecular clashes, while a smaller mutation may result in gaps, disrupting protein folding and functionality[17].

Efficiently categorizing SNPs, particularly through combined data from *in silico* tools (e.g. structural analysis and molecular dynamics simulations), has enhanced predictions of genetic variant impact, improving our understanding of molecular pathology in rare disorders. These computational methods are crucial for advancing personalized medicine, pharmacogenomics and genetic disease prognosis[16]. In this study, we focus on characterizing functional effects of curated KRAS mutations using *in silico* pathogenicity and stability predictions, structure modeling and docking analysis.

Collection and Screening of KRAS mutants: The COSMIC database, which houses causative mutations for various cancer types, was used to gather mutations associated with lung cancer in the KRAS protein. A total of 190,793 KRAS mutations were identified from the COSMIC database, in which 28,964 were linked to the onset of lung cancer. Among the 28,964 mutation samples, 63 distinct mutations were identified and further analyzed for its pathogenicity and stability.

Pathogenicity of the KRAS mutants: The pathogenicity of the 63 mutants was assessed using Meta-SNP server. The server integrates various tools like PANTHER, PhD-SNP, SIFT and SNAP. In PANTHER, PhD-SNP and SNAP, the output is given in values ranging between 0 and 1. If the predicted values of mutations are greater than 0.5, it is designated as disease-causing mutants and if the values are less than 0.5, they are designated as neutral. In case of SIFT tool, the values less than 0.05 are predicted as disease-causing mutations and greater than 0.05 are considered to be neutral. Among the 63 mutants, 8 mutants: A18D, D57N, G10V, G12C, G13C, G13D, G13V and T58I were predicted to be disease causing by PANTHER, PhD-SNP, SIFT, SNAP and Meta-SNP (Table 1). These mutants were analyzed further for their stability and properties.

Stability Prediction of KRAS mutants: The stability of a protein is the overall balance of forces that dictate whether it will assume its functional folded structure or an abnormal conformation that may impair its function[26]. The eight mutants were assessed for its stability upon mutation (Table 2). The iStable server[16] predicts protein stability changes due to mutations using thermodynamic parameters. A single

amino acid mutation can significantly affect protein structure resulting in changes in folding free energy and stability. I-Mutant2.0 was employed to assess the change in folding free energy (ddG)[17]. Stability of the mutants was also assessed using MUpro18. MUpro's confidence score ranges from −1 to 1, reflecting the likelihood of stability changes, with higher scores indicating greater confidence[18].

From the table 2, It could be observed that the mutations – G10V, G13C and G13V decreased the stability of the proteins. With respect to A18D and D57N, MUpro and iStable predicted that the mutation caused a decrease in the stability, whereas iMutant predicted an increase in stability. For the mutants G12C and G13D, MUpro and iStable predicted that the mutation caused an increase in the stability, whereas iMutant predicted an decrease in stability. For G13D, iMutant and iStable predicted that the mutation caused a decrease in the stability, whereas MUpro predicted an increase in stability.

Physicochemical Properties of the mutants: The physiochemical properties of the eight mutants were predicted using Protparam tool (Table 3). All mutants did not show much difference in molecular weight (~21kDa). With respect to the theoretical pI, the mutants - G10V, G12C, G13C, G13V and T58I had pI similar to that of the wild type KRAS (6.33).

Table 1
Pathogenicity values of various KRAS mutants

Amino Acid	Coding Sequence	PANTHER		PhD-SNP		SIFT		SNAP		Meta-SNP	
A18D	c.53C>A	Disease	0.683	Disease	0.9	Disease	0	Disease	0.685	Disease	0.777
D57N	c.169G>A	Disease	0.509	Disease	0.814	Disease	0	Disease	0.875	Disease	0.799
G10V	c.29G>T	Disease	0.665	Disease	0.934	Disease	0	Disease	0.845	Disease	0.902
G12C	c.34G>T	Disease	0.644	Disease	0.885	Disease	0.02	Disease	0.65	Disease	0.762
G13C	c.37G>T	Disease	0.863	Disease	0.927	Disease	0.01	Disease	0.72	Disease	0.807
G13D	c.38G>A	Disease	0.742	Disease	0.923	Disease	0	Disease	0.795	Disease	0.808
G13V	c.38G>T	Disease	0.78	Disease	0.925	Disease	0.01	Disease	0.785	Disease	0.831
T58I	c.173C>T	Disease	0.765	Disease	0.793	Disease	0.01	Disease	0.765	Disease	0.771

Table 2
Stability prediction of KRAS mutants

Amino Acid	Coding Sequence	i-Mutant2.0 SEQ	DDG	MUpro	Conf. Score	iStable	Conf. Score
A18D	c.53C>A	Increase	-0.02	Decrease	-0.33645484	Decrease	0.527204
D57N	c.169G>A	Increase	0.27	Decrease	-1	Decrease	0.580317
G10V	c.29G>T	Decrease	-0.44	Decrease	-0.64207936	Decrease	0.822543
G12C	c.34G>T	Decrease	-1.59	Increase	0.35992912	Increase	0.668212
G13C	c.37G>T	Decrease	-1.56	Decrease	-0.85930258	Decrease	0.695908
G13D	c.38G>A	Decrease	-1.01	Increase	0.1377395	Increase	0.698219
G13V	c.38G>T	Decrease	-0.57	Decrease	-0.68586877	Decrease	0.686209
T58I	c.173C>T	Increase	0.22	Decrease	-0.75989536	Increase	0.707965

Table 3
Physicochemical properties of KRAS mutants

Mutants	Molecular weight (daltons)	Theoretical pI	Ext. coefficient, assuming all pairs of Cys residues form cystines	Ext. coefficient, assuming all Cys residues are reduced	Instability index	Aliphatic index	GRAVY
Wild type	21655.83	6.33	0.631	0.619	38.56	85.03	-0.432
A18D	21699.84	5.98	0.629	0.618	38.56	84.5	-0.46
D57N	21654.84	6.83	0.631	0.619	38.9	85.03	-0.432
G10V	21697.91	6.33	0.63	0.618	39.46	86.56	-0.407
G12C	21701.92	6.33	0.635	0.618	40.23	85.03	-0.416
G13C	21701.92	6.33	0.635	0.618	37.51	85.03	-0.416
G13D	21713.87	5.98	0.629	0.618	37.91	85.03	-0.448
G13V	21697.91	6.33	0.63	0.618	37.91	86.56	-0.407
T58I	21667.88	6.33	0.63	0.619	39.35	87.09	-0.404

The eight mutants and wild type KRAS exhibited an Extinction coefficient of ~ 0.63, assuming all pairs of Cys residues form cystines and ~ 0.618 for Extinction coefficient, assuming all Cys residues are reduced. The estimated half-life is 30 h in mammalian reticulocytes, *in vitro*; >20 h in yeast, *in vivo* and >10 h in *Escherichia coli*, *in vivo* in wild type KRAS and mutants. The instability index was < 40 indicating its stability. The aliphatic index was between 84 and 87 indicating these mutants are thermostable. GRAVY scores of the mutants ranging between -0.46 and -0.404 indicates its hydrophilicity and indicates they are often soluble in aqueous environments like cytoplasm or extracellular fluids.

KRAS Structure Prediction and Validation: PSIPRED tool was used to predict the secondary structures of the wild and mutant KRAS. It was revealed that the secondary structures were primarily composed of coils, helices and strands (Fig. 1).

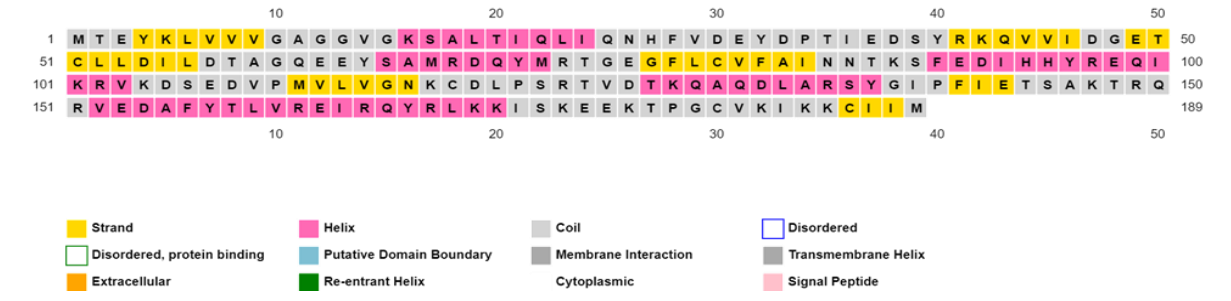
The wild type and KRAS mutants were modeled using the Robetta server, which utilizes a deep learning-based approach to modeling. Three dimensional models help in understanding the interactions of the protein with that of the ligand. Five models were generated for each input. The generated models were validated through MolProbity server and the best model which demonstrated high-quality Ramachandran and rotamer distributions were chosen for further analysis. The Ramachandran plot, based on phi and psi dihedral angles, helps verify the protein backbone conformation by evaluating each residue and categorizing them into allowed, favored and outlier regions[14]. The best wild type and KRAS mutant constructs displayed 98.5% of all residues in the 98% favored regions and 99.5-100% of all residues in the allowed regions (>99.8%) in the Ramachandran plot.

Molecular Docking and Binding analysis: The KRAS Wild type and the mutant proteins were docked with Guanosine triphosphate (GTP) ligand using Autodock Vina. Active site docking was performed, where the grid box was set around the active site aminoacids -Gly10, Lys11, Lys16, Met122, Gly12, Asp13, Phe61, Val62, Gln63, Asn66, Tyr75, Gly116, Gly117, Lys118, Cys119, Asp59, Ser117, Thr153, Asn154, Asp154 and Gln155[19]. The grid parameters of the wild type and mutant KRAS are listed in table 4.

The dockings were performed in triplicates and the average binding energy in kcal/mol are tabulated in table 5. The wild type has a binding energy of -7.5 ± 0.46 kcal/mol. Most mutations result in slight variations in binding energy, with values ranging from -6.07 kcal/mol for T58I, which shows the weakest binding, to -7.93 kcal/mol for D57N, which has the strongest binding. Other mutations, such as A18D, G12C, G13C and G13D, exhibit intermediate binding energies, generally slightly lower than the wild type. Mutations like G10V and G13V also show reduced binding compared to the wild type, indicating a range of effects on receptor-GTP binding affinity across the different variants.

PLIP was used to evaluate the various interactions between the protein and GTP (Table 5). Post-dock analysis was conducted to examine the various non-covalent interactions, including hydrogen bonds, hydrophobic interactions and electrostatic interactions, between the ligand and the protein. Hydrogen bonds, in particular, play a crucial role in determining protein stability within docking complexes[18]. The wild type receptor has a binding energy of -7.5 ± 0.46 kcal/mol, with key hydrogen bonds at Ala11, Gly13, Val14, Gly15, Lys16, Tyr32, Asn116, Ala146 and Lys147, hydrophobic interactions at Ala18 and Phe28 and a salt bridge at Lys16.

(a)



(b)

Fig. 1: Secondary Structure prediction of (a) wild type KRAS and (b) mutant D57N

Table 4
Grid parameters of wild type and mutant KRAS used for docking studies

Receptor	Size_x	Size_y	Size_z	Centre_x	Centre_y	Centre_z
Wild type	70	70	70	25.812	1.903	-8.751
A18D	50	56	88	8.064	4.452	-8.233
D57N	70	70	70	28.244	0	7.086
G10V	60	60	60	25.066	8.235	-3.289
G12C	60	60	60	22.673	-6.612	-12.087
G13C	60	60	60	11.86	24.197	21.043
G13D	60	60	60	21.484	-12.828	-8.526
G13V	60	60	60	11.755	20.332	17.335
T58I	50	72	90	-13.421	18.158	-11.576

Table 5
Binding energy and Protein Ligand Interaction of KRAS wild type and mutants with GTP

Receptor	Binding Energy with GTP (kcal/mol)	Hydrogen Bond	Hydrophobic interactions	Salt Bridges
Wild type	-7.5±0.46	Ala11, Gly13, Val14, Gly15, Lys16, Tyr32, Asn116, Ala146, Lys147	Ala18, Phe28	Lys16
A18D	-7.13±0.46	Glu37, Asp38, Asp57, Gln61, Tyr64, Thr74, Arg164, Lys170	-	Arg41, Arg167
D57N	-7.93±0.23	Gly13, Gly15, Lys16, Ser17, Val29, Asn57, Lys117, Ser145, Lys147	Lys117, Lys147	Lys16
G10V	-6.83±0.46	Gly13, Val14, Gly15, Lys16, Ser17, Tyr32, Gly60, Glu62, Glu63	-	Glu62
G12C	-7.1±0.17	Gly13, Gly15, Lys16, Ser17, Tyr32, Thr35, Lys147	Ala59	Lys147
G13C	-7.23±0.38	Cys13, Gly15, Ala18, Asp30, Tyr32, Thr35, Asn116, Lys117, Asp119, Ser145, Ala146, Lys147	Ala18	Asp119
G13D	-7.6±0.17	Asp13, Glu31, Tyr32, Asn85, Asn116, Lys117, Ser145, Ala146, Lys147	Phe28	Lys117, Asp119
G13V	-6.93±0.38	Val13, Asp30, Tyr32, Asn85, Asn116, Lys117, Ser145, Ala146, Lys147	Ala18, Phe28	Lys117, Asp119
T58I	-6.07±0.12	Asn26, Lys42, Gln43, Lys185	-	Lys42, Lys185

The D57N mutant has a slightly stronger binding energy of -7.93 ± 0.23 kcal/mol, with hydrogen bonds at Gly13, Gly15, Lys16, Ser17, Val29, Asn57, Lys117, Ser145 and Lys147, hydrophobic interactions at Lys117 and Lys147 and a salt bridge at Lys16. In the D57N mutant, several new hydrogen bonds and hydrophobic interactions were introduced, particularly involving Lys117 and Ser145, suggesting altered interactions compared to the wild type.

Conclusion

In this study, we investigated KRAS mutations associated with lung cancer, analyzing a total of 63 mutations sourced from the COSMIC database. These mutations were screened for pathogenicity and we identified 8 mutations with the highest pathogenic potential. Subsequently, we assessed the stability of these 8 mutants, finding that all were stable. Despite the mutations, the physicochemical properties of the 8 mutants remained nearly identical to the wild type. We further examined the impact of these mutations by predicting secondary structure changes and modeling the protein.

To evaluate the effects of mutations on GTP binding, docking studies were performed, revealing that the D57N mutant exhibited stronger binding energy compared to the

wild type, suggesting increased activity. The remaining mutants showed similar binding energy to the wild type. Additionally, the protein-ligand interactions in the mutant complexes were comparable to those observed in the wild type, indicating similar interaction patterns.

Acknowledgement
The authors would like to thank NIT, Andhra Pradesh for providing the necessary facilities to carry out the work.

References
1. Adasme M.F., Linnemann K.L., Bolz S.N., Kaiser F., Salentin S., Haupt V.J. and Schroeder M., PLIP 2021: Expanding the scope of the protein–ligand interaction profiler to DNA and RNA, *Nucleic Acids Res*., **49(W1)**, W530-W534 **(2021)**

2. Bos J.L., Rehmann H. and Wittinghofer A., GEFs and GAPs: critical elements in the control of small G proteins, *Cell*, **129(5)**, 865-877 **(2007)**

3. Buchan D.W., Moffat L., Lau A., Kandathil S.M. and Jones D.T., Deep learning for the PSIPRED Protein Analysis Workbench, *Nucleic Acids Res*., **52(1)**, W287–W293 **(2024)**

4. Capriotti E., Calabrese R. and Casadio R., Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information, *Bioinformatics*, **22(22)**, 2729-2734 **(2006)**

5. Capriotti E., Fariselli P. and Casadio R., I-Mutant 2.0: predicting stability changes upon mutation from the protein sequence or structure, *Nucleic Acids Res*., **33(2)**, W306-W310 **(2005)**

6. Chen C.W., Lin M.H., Liao C.C., Chang H.P. and Chu Y.W., iStable 2.0: Predicting protein thermal stability changes by integrating various characteristic modules, *Comput. Struct. Biotechnol. J.*, **18**, 622-630 **(2020)**

7. Cheng J., Randall A. and Baldi P., Prediction of protein stability changes for single-site mutations using support vector machines, *Proteins*, **62(4)**, 1125-1132 **(2006)**

8. Cox A.D., Der C.J. and Philips M.R., Targeting RAS membrane association: back to the future for anti-RAS drug discovery?, *Clin. Cancer. Res.*, **21(8)**, 1819-1827 **(2015)**

9. Fong K.M., Sekido Y., Gazdar A.F. and Minna J.D., Lung cancer, 9: Molecular biology of lung cancer: clinical implications, *Thorax*, **58(10)**, 892-900 **(2003)**

10. Forbes S.A. et al, COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer, *Nucleic Acids Res*., **39(1)**, D945-D950 **(2010)**

11. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D. and Bairoch A., Protein Identification and Analysis Tools on the Expasy Server, In Walker John M., ed., The Proteomics Protocols Handbook, Humana Press, 571-607 **(2005)**

12. Hunter J.C., Manandhar A., Carrasco M.A., Gurbani D., Gondi S. and Westover K.D., Biochemical and structural analysis of common cancer-associated KRAS mutations, *Mol. Cancer. Res.*, **13(9)**, 1325-1335 **(2015)**

13. Johnson A.D., Handsaker R.E., Pulit S.L., Nizzari M.M., O'Donnell C.J. and De Bakker P.I., SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap, *Bioinformatics*, **24(24)**, 2938-2939 **(2008)**

14. Kizhakedathil M.P.J., Saravanakumar T., Anbarasu S., Kallesh S. and Thiruvalluvan V., Designing of Chimeric Vaccine against Canine Distemper Virus Targeting Hemaglutanin Protein, *Biointerface Res. Appl. Chem*., **13**, 347 **(2022)**

15. Kratzer T.B., Bandi P., Freedman N.D., Smith R.A., Travis W.D., Jemal A. and Siegel R.L., Lung cancer statistics, *Cancer*, **130(8)**, 1330-1348 **(2024)**

16. Kumar D.T., Emerald L.J., Doss C.G.P., Sneha P., Siva R., Jebaraj W.C.E. and Zayed H., Computational approach to unravel the impact of missense mutations of proteins (D2HGDH and IDH2) causing D-2-hydroxyglutaric aciduria 2, Metab. *Brain Dis.*, **33(5)**, 1699-1710 **(2018)**

17. Kumar S.U. et al, Comprehensive in silico screening and molecular dynamics studies of missense mutations in Sjogren-Larsson syndrome associated with the ALDH3A2 gene, *Adv. Protein Chem. Struct. Biol.*, **120**, 349-377 **(2020)**

18. Loganathan Y., Jain M., Thiyagarajan S., Shanmuganathan S., Mariappan S.K., Kizhakedathil M.P.J. and Saravanakumar T., An Insilico evaluation of phytocompounds from *Albizia am*ara and *Phyla nodiflora* as cyclooxygenase-2 enzyme inhibitors, *DARU J. Pharm. Sci*., **29(2)**, 311-320 **(2021)**

19. Maurer T. et al, Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity, Proc. Natl. Acad. Sci., **109(14)**, 5299-5304 **(2012)**

20. Minna J.D., Kurie J.M. and Jacks T., A big step in the study of small cell lung cancer, *Cancer Cell*, **4(3)**, 163-166 **(2003)**

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

22. Peto J., Cancer epidemiology in the last century and the next decade, *Nature*, **411(6835)**, 390-395 **(2001)**

23. Peto R., Darby S., Deo H., Silcocks P., Whitley E. and Doll R., Smoking, smoking cessation and lung cancer in the UK since 1950: combination of national statistics with two case-control studies, *BMJ*, **321(7257)**, 323-329 **(2000)**

24. Raman S. et al, Structure prediction for CASP8 with all-atom refinement using Rosetta, *Proteins*, **77(S9)**, 89-99 **(2009)**

25. Sondka Z. et al, COSMIC: a curated database of somatic variants and clinical data for cancer, *Nucleic Acids Res*., **52(D1)**, D1210-D1217 **(2024)**

26. Thomas P.D., Ebert D., Muruganujan A., Mushayahama T., Albou L.P. and Mi H., PANTHER: Making genome-scale phylogenetics accessible to all, *Protein Sci.*, **31(1)**, 8-22 **(2022)**

27. Trott O. and Olson A.J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *J. Comput. Chem*., **31(2)**, 455-461 **(2010)**

28. Vetter I.R. and Wittinghofer A., The guanine nucleotide-binding switch in three dimensions, *Science*, **294(5545)**, 1299-1304 **(2001)**

29. Williams C.J. et al, MolProbity: more and better reference data for improved all-atom structure validation, *Protein Sci.*, **27(1)**, 293-315 **(2018)**.

(Received 17th November 2024, accepted 21st January 2025)