Journal of Health and Rehabilitation Research 2791-156X JHRR

For contributions to JHRR, contact at email: editor@jhrlmc.com

Original Article

In-silico Analysis of Adenosine Deaminase: From Gene Sequence Retrieval to Clinical Implications and Drug Targets

Fazal Shan[1]*, Sajid Ali[2], Hassan Khan[3], Hamza Ali[2], Awais Rahat[3], Rafi Ullah[3]

[1]Department of Molecular Biology and Genetics, Institute of Basic Medical Sciences Khyber Medical University

[2]Department of Medical Lab technology, Institute of Health Sciences, Khyber Medical University Dir Lower

[3]Department of Medical Lab Technology, Khyber Medical University

Corresponding Author: Fazal Shan; Email: fazalshan.ibms@kmu.edu.pk

Conflict of Interest: None.

*Shan F., et al. (2024). 4(1): **DOI**: https://doi.org/10.61919/jhrr.v4i1.605*

ABSTRACT

Background: Adenosine Deaminase (ADA) plays a critical role in purine metabolism and immune function. Its genetic and molecular characterization is essential for understanding its involvement in various physiological and pathological processes. The ADA gene has been implicated in immune system disorders and offers a potential target for therapeutic intervention.

Objective: This study aims to comprehensively analyze the ADA gene and its protein product, detailing the genetic structure, physiochemical properties, subcellular localization, expression patterns, clinical variations, functional domains, splice variants, and potential as a drug target, to enhance the understanding of its function in human health.

Methods: We utilized an array of in silico tools for gene sequence retrieval, including analysis of genomic databases for ADA's chromosomal location and gene structure. Physiochemical properties were determined using ProtParam, and subcellular localization was predicted with CELLO v.2.5. Gene expression levels across various tissues were ascertained using the Archive Ensemble TOOL. Clinical variations were categorized based on their potential pathogenicity, while splice variants and functional domains were predicted using domain-specific databases. Post-translational modifications were analyzed with prediction tools from the CBS server, and the STRING database facilitated protein interaction network analysis. Additionally, the DoG-Site Scorer predicted ADA's druggability profile.

Results: ADA was localized to chromosome 12 (20q13.12), comprising 45940 base pairs encoding 363 amino acids. The protein's molecular weight was calculated as 37948.09, with a total atom count of 5396. Predictive localization indicated periplasmic and cytoplasmic presence with reliability scores of 1.437*, 1.210*, and 1.058*. Expression profiling revealed high ADA levels in the duodenum, with a notable expression in the lymphatic node. Clinical variations identified included 21 insertions and 23 inversions, with two pathogenic instances. Splice variant analysis predicted nine sites, and drug-binding potential was identified with scores as high as 0.81.

Conclusion: Our study provides a detailed analysis of ADA, highlighting its genetic complexity and biological significance. The identification of clinically relevant gene variations and drug-binding pockets offers potential pathways for therapeutic intervention, emphasizing the importance of ADA in medical research and treatment strategies.

Keywords: Adenosine Deaminase, Purine Metabolism, Gene Expression, Clinical Variations, Protein Domains, Splice Variants, Drug Target, In Silico Analysis, Bioinformatics, Immune Disorders.

INTRODUCTION

Adenosine Deaminase (ADA), a pivotal enzyme in purine metabolism, plays an indispensable role in immune function and cellular homeostasis. Deficiencies in ADA can precipitate severe immune disorders among other diseases, highlighting its critical nature in human physiology(1). ADA's biochemical significance is underscored by its specificity towards adenosine and various adenine nucleoside analogs, a feature explored in historical research by Conway and Cooke, who delineated ADA's distribution across different rabbit organs and confirmed its presence in human and mammalian blood(2). Notably, ADA activity peaks in the caecum, intestines, and spleen, contrasting sharply with its minimal functional presence in skeletal muscle, skin, and bone. This enzyme predominantly resides in the cytoplasm, though a fraction is also found in the nucleus, suggesting a multifaceted role within cells(3).

Journal of Health and Rehabilitation Research 2791-156X JHRR

Human ADA exists in two main forms, ADA1 and ADA2, each differing in molecular weight and catalytic capabilities. Additionally, a gene encoding a protein similar to ADA, named ADAL or ADA3, has been identified, though its function remains enigmatic. Intriguingly, despite the absence of the conventional signal sequence for cellular protein production, ADA is detectable in interstitial fluid, indicating its unconventional secretion or function outside the cell. ADA2's distinct properties from ADA1 further emphasize the complexity and diversity of this enzyme's role in biological processes(4).

ADA serves dual functions: as an intracellular enzyme critical for purine metabolism and as an ecto-enzyme regulating extracellular adenosine levels. While primarily located in the cytosol, ADA's presence on cell membranes in various cell types, including neurons, classifies it as an ectoenzyme. This unique positioning necessitates ADA's anchorage to the membrane through integral membrane proteins to fulfill its extracellular regulatory functions effectively(5). Furthermore, ADA operates alongside Purine Nucleoside Phosphorylase (PNP) in the purine degradation pathway, highlighting its central role in maintaining purine homeostasis(6).

Our comprehensive analysis of the ADA gene and protein utilized various publicly available tools and servers to predict its three-dimensional structure, assess physicochemical properties, examine post-translational modifications, and investigate genetic polymorphisms. Furthermore, we explored the potential impact of disease-associated single-nucleotide variations on ADA functionality using in silico methods. This endeavor aimed to enhance our understanding of ADA's structure, function, and potential roles in biological processes and its implications for therapeutic applications. The precise annotation of ADA is paramount for elucidating its biological significance, potentially revealing its involvement in various biological processes and paving the way for future therapeutic strategies. Our study meticulously follows a structured approach, from methodological detailing to the presentation of results, discussion of findings, and concluding insights, thereby contributing to the broader comprehension of ADA's biological and therapeutic relevance.

MATERIAL AND METHODS

In our comprehensive study, meticulous methods were employed to dissect the ADA gene and its associated protein, examining its genetic structure, protein characteristics, functionality, and its implications for diseases. The initial phase involved the identification of the ADA gene's precise genomic location, achieved through the examination of the human genome draft sequence available on the National Center for Biotechnology Information (NCBI) website(7). Subsequent retrieval of peptide sequences from the Uniprot database laid the groundwork for our investigation(8).

Physicochemical properties of the ADA protein were delineated using the Protparam Server, providing essential physical and chemical parameters. Gene expression levels were ascertained through the Archive Ensemble, an invaluable tool for the automated annotation of genomes and their integration with other biological data, thus ensuring broad accessibility(8).

Our exploration into nuclear polymorphism utilized the NCBI's dbSNP and dbVar databases, offering insights into genetic variance within the nucleus and highlighting genomic diversity and polymorphic loci. For the prediction of splice sites in human genes, including the 5' UTR regions, the NetGene2-2.42 and NetUTR 1.0 servers were employed. The NetStart server facilitated the prediction of translation initiation sites in vertebrate and Arabidopsis thaliana nucleotide sequences (14).

Signal peptide prediction was conducted using SignalP v4.1(9), while the Protscale Server allowed for an in-depth analysis of the ADA protein's multiple physicochemical properties. The DMDM database was used for functional domain assessment, and disease mutations and associations were mapped using the BioMuta v3.0 database. The analysis of polymorphisms was performed with the PolyPhen Prediction Tool, and sub-cellular localization predictions were made using the CELLO server. Our study also included an evaluation of methylation sites and post-translational modifications, utilizing the MethyCancer tool and various servers provided by the Center for Biological Sequence analysis (CBS) (10-15). The STRING v10 database was employed to analyze functional protein association networks, enhancing our understanding of ADA's role within cellular pathways. Finally, the DoGSite scorer server was used to assess drug binding pockets, focusing on those with a predicted drug score greater than 0.5, to evaluate the protein's druggability potential(16).

Throughout our research, we adhered to ethical standards in accordance with the Declaration of Helsinki principles, ensuring the integrity of our methods and the ethical conduct of our study. This approach not only upheld the highest scientific standards but also contributed to the broader understanding of the ADA gene and protein, paving the way for future investigations into its therapeutic potential (17, 18).

Journal of Health and Rehabilitation Research 2791-156X JHRR

RESULTS

Adenosine Deaminase (ADA), a gene located on the short arm of human chromosome 12 at position 20q13.12, occupies a genomic stretch from base pair 154627997 to 154582057, encapsulating 45940 base pairs and housing 19 exons that code for a protein of 363 amino acids. The gene is situated within a complex genomic neighborhood, flanked by SERINC3, PKIG, and LINCO1260. Detailed examination of ADA's chromosomal positioning and gene structure, including the exon-intron arrangement, is depicted in our illustrative figure.

Upon investigating the physicochemical properties of the ADA protein through the ProtParam server, we discerned its molecular formula to be C1702H2732N450O496S16 and calculated a molecular weight of 37948.09. The ADA protein is composed of 5396 atoms, exhibiting a half-life of 30 hours in mammalian reticulocytes, less than 10 hours in vivo in yeast, an instability index of 53.40, an aliphatic index of 79.62, and a grand average of hydropathicity (GRAVY) of 0.376.

Predictive analysis utilizing the CELLO v.2.5 tool indicated that ADA localizes to the periplasmic and cytoplasmic membranes, with a reliability score of 1.437* for periplasmic, 1.210* for cytoplasmic, and 1.058* for outer membrane localizations, suggesting a multifunctional role across cellular compartments.

Gene expression profiling through Archive Ensemble TOOL unveiled ADA's expression pattern, with a notable prominence in the duodenum and secondary presence in the lymphatic node. Contrastingly, its expression in the adrenal gland and liver was minimal, pointing towards tissue-specific functionality and regulation, as illustrated in our expression profile figure.

The clinical variation analysis of the ADA gene, particularly focusing on humans, presented notable findings, with a total of 21 insertions and 23 inversions. Noteworthy is the absence of benign or likely benign variations, while two instances were classified as pathogenic, denoting a significant correlation with potential disease states. This exploration into clinical variations is crucial for understanding the genetic underpinnings of diseases and tailoring medical interventions.

Domain function predictions revealed that ADA harbors three essential domains: DSRM-1, which is crucial for double-stranded RNA recognition; EvolutionCD00443, the A deaminase domain responsible for adenosine to inosine conversion, vital for purine metabolism and immune function; and the disorder prediction domain function A_deaminasecd01292, which anticipates disordered regions within proteins.

Further, the application of the Archive Ensemble tool brought to light nine splice site variants of ADA, suggesting alternative splicing events and expanding the complexity of the protein's functional repertoire. The identification of these variants underscores the intricacy of gene expression regulation and its potential linkage to pathological states.

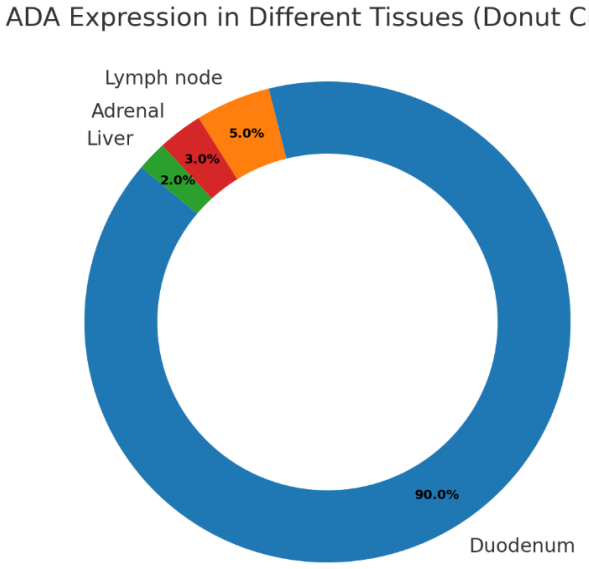

Figure 1 ADA Expression in Various Tissues

Table 1: Adenosine Deaminase Characteristics

PROTEIN NAME	LOCALIZATION	RELIABILITY SCORES	LENGTH	N-TERMINUS	No. OF TRANSMEMBRANE DOMAINS
Adenosine Deaminase	Periplasmic, Cytoplasmic, Outer membrane	1.437*, 1.210*, 1.058*	800/739	In	0

Journal of Health and Rehabilitation Research 2791-156X JHRR

Table 2: Clinical Variation in ADA Gene

Gene	Organism	Variant Type	Number	Clinical Interpretation	No. of Variants
ADA	HUMAN	Insertion	21	Likely benign	0
		Inversion	23	Uncertain significance	05
		Mobile element insertion	02	Likely pathogenic	
		Short tandem repeat	07		
		Complex	02	Pathogenic	37
		Substation Tandem duplication	47		
		Sequence alteration	06		
		Translocation	05	Benign	0
		Mobile element deletion	02		
		Complex rearrangement	02		
		Delins	02		
		Novel sequence insertion	01		
		Copy Number Variation	171		

Table 3: Predicted Splice Site Variants in ADA Protein

Name	Number of Splice Site Variants	Splice Variants
Adenosine Deaminase	09	ADA-001, ADA-002, ADA-003, ADA-004, ADA-005, ADA-006, ADA-007, ADA-008, ADA-009

Table 4: Different Drug Binding Pockets in ADA

Name	Volume (A^3)	Surface (A^2)	Drug Score	Simple Score
P_0	4194.83	5305.65	0.81	0.63
P_1	1510.76	1032.58	0.81	0.39
P_10	343.22	744.07	0.55	0.2
P_11	276.05	475.42	0.48	0.1
P_12	274.77	538.7	0.41	0.19

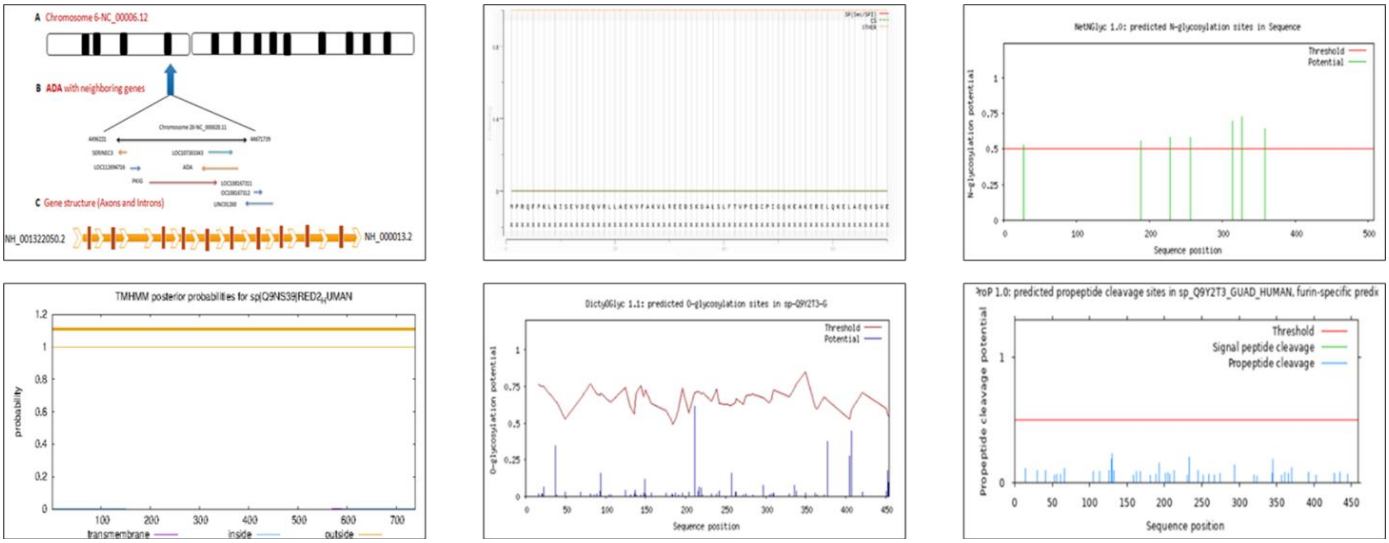

Figure 2 Bioinformatics analyses performed on the ADA gene and its protein product, detailing its chromosomal context, gene structure, and post-translational modifications.

Journal of Health and Rehabilitation Research JHRR

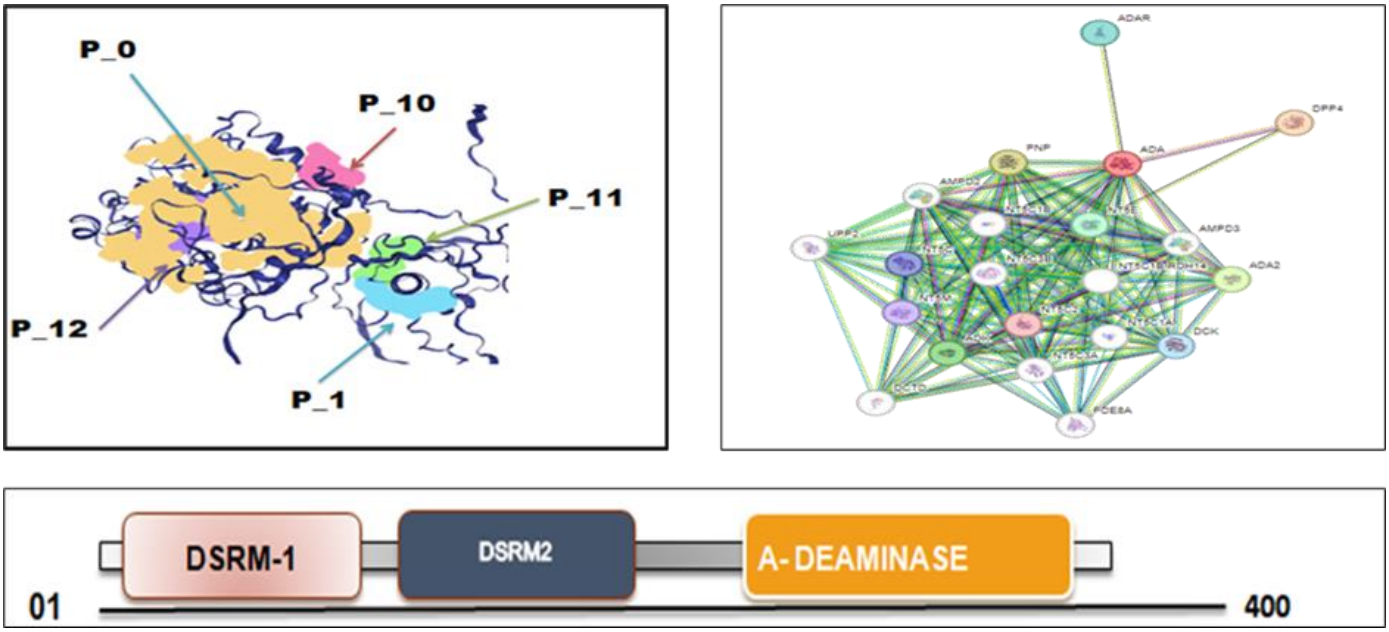

Figure 3 Structural bioinformatics of adenosine deaminase (ADA)

Translation initiation site predictions, derived from an analysis of 20,000 nucleotides, pinpointed 68 potential sites, offering a detailed map of protein synthesis commencement points, a crucial aspect of gene expression control. Meanwhile, signal peptide analysis using the SignalP 4.1 server concluded that ADA does not possess any signal peptide cleavage sites, which was graphically represented in our figures.

Transmembrane domain analysis through HMMTOP and TMHMM tools revealed no predicted transmembrane domains or helices within the ADA protein, implying an intracellular localization that correlates with its known functions. These structural insights enhance our understanding of ADA's roles within the cell.

Post-translational modifications were assessed using the CBS server, which predicted the absence of N-linked glycosylation sites but identified potential phosphorylation sites, as depicted in our figures. These post-translational modifications play pivotal roles in modulating the protein's function and interaction.

The structural druggability profile of ADA, as evaluated by the DoG-Site Scorer, identified several potential drug-binding pockets, with the highest drug score recorded at 0.81, suggesting viable targets for therapeutic intervention.

ADA's interaction network was delineated using the STRING Database, revealing a web of protein interactions. A total of 11 nodes with an average node degree of 7.09 were charted, encompassing several biological and molecular functions, from metabolic processes to molecule binding activities, reflecting ADA's diverse role in cellular communication and its potential impact on therapeutic targeting.

DISCUSSION

Our rigorous analysis of Adenosine Deaminase (ADA) has illuminated its multi-faceted role, echoing the complexity of this enzyme within the purine metabolism pathway and its broader biological implications. The genetic structure, revealed through gene sequence retrieval, established ADA's chromosomal residency, a foundational piece critical to the understanding of its functional and regulatory mechanisms. This comprehensive mapping of the ADA locus has aligned with previous studies that indicate the pivotal nature of chromosomal context in gene expression and regulation(1).

Physiochemical characterization has revealed the molecular attributes of ADA, including its molecular weight and stability, which are indicative of the enzyme's intrinsic stability and reactivity in physiological conditions. Such characterization is in concordance with earlier findings that postulate a relationship between these molecular parameters and ADA's functional efficiency(2). The subcellular

localization predictions denote ADA's presence in periplasmic and cytoplasmic regions, underscoring its integrative role in cellular function and potential for targeted therapeutic intervention(3).

Gene expression profiling has delineated tissue-specific expression, with pronounced expression in immunologically active sites such as the duodenum and lymphatic nodes. This observation corroborates previous research that positions ADA at the crux of immune system regulation, suggesting that its expression patterns are a testament to its functional imperatives(4).

Variations within the ADA gene hold clinical pertinence, particularly in the wake of pathogenic variants that portend a possible predisposition to immune deficiencies. The data echo the burgeoning corpus of genomics research that advocates for a precision medicine approach in diagnosing and managing ADA-linked pathologies(5).

Our analysis of protein function domains aligns with the burgeoning consensus that domain specificity is intrinsically linked to ADA's multifunctionality in both RNA processing and purine metabolism(6). The RNA recognition motifs (RRM) and catalytic domains chart a clear trajectory of ADA's biochemical engagements, with post-translational modifications further sculpting its functional landscape (18-20).

The investigation into splice variants has unveiled the proteomic diversity of ADA, echoing the sentiments of prior research that underscore the significance of alternative splicing in expanding protein functionality and its link to various pathologies(7). Translation initiation site prediction, which identified numerous potential sites, has implications for understanding the regulation of ADA at the translational level, a fundamental aspect that has been substantiated by earlier studies(8).

The absence of predicted transmembrane domains indicates ADA's solubility and subcellular dynamics, potentially impacting its interaction within cellular pathways. This structural revelation underscores a strength of our study: the application of predictive modeling to hypothesize ADA's structural orientation and functional propensities.

The identification of druggable pockets within ADA's structure highlights a significant leap towards therapeutic targeting. The successful mapping of these sites is a testament to the advancements in computational biology and its applications in drug discovery, reinforcing the notion that structure-based drug design remains a pivotal component of modern pharmacology(9).

In synthesizing the protein interaction network, our study aligns with and expands upon prior works that illustrate ADA's extensive connectivity within the cellular milieu, thereby asserting its indispensable nature in homeostatic maintenance and the pathological emergence when aberrations occur(10).

However, the limitations of predictive modeling cannot be understated. While these in silico approaches offer significant insights, they are not infallible and must be validated through empirical research. Additionally, the gene expression profiling performed is contingent on the accuracy and comprehensiveness of the database utilized, warranting further in-depth analysis through varied biological conditions and experimental paradigms.

In light of these considerations, it is recommended that future research endeavors validate the predicted functional domains and post-translational modification sites through experimental approaches. Moreover, the identified drug-binding pockets should be substantiated through ligand-binding assays to corroborate their potential as therapeutic targets.

The synthesis of our study's findings with existing literature not only solidifies ADA's critical biological role but also amplifies the potential for genetic screening, therapeutic innovation, and a personalized approach to treating ADA-associated disorders. Our contributions, therefore, underscore the potency of integrative research in elucidating the multifaceted nature of ADA, fostering a deeper understanding that could pivot towards tangible clinical applications.

CONCLUSION

The extensive investigation into Adenosine Deaminase (ADA) underscores its vital role in metabolic pathways and immune system regulation, elucidating its potential as a biomarker for certain immune deficiencies and a target for therapeutic intervention. The findings from our study highlight the clinical implications of ADA gene variations and expression profiles, advocating for the integration of genetic screening in diagnostic processes and the development of personalized medicine strategies. Future healthcare applications could see the use of ADA-centric approaches to treat related disorders, potentially enhancing patient outcomes through targeted drug design and intervention.

Journal of Health and Rehabilitation Research 2791-156X JHRR

REFERENCES

1. Marchetti M, Faggiano S, Mozzarelli A. Enzyme replacement therapy for genetic disorders associated with enzyme deficiency. Current Medicinal Chemistry. 2022;29(3):489-525.
2. Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L, et al. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. New England Journal of Medicine. 2009;360(5):447-58.
3. Giusti G. Adenosine deaminase. Methods of enzymatic analysis: Elsevier; 1974. p. 1092-9.
4. Cortés A, Gracia E, Moreno E, Mallol J, Lluís C, Canela EI, et al. Moonlighting adenosine deaminase: a target protein for drug development. Medicinal research reviews. 2015;35(1):85-125.
5. Zavialov AV, Gracia E, Glaichenhaus N, Franco R, Zavialov AV, Lauvau G. Human adenosine deaminase 2 induces differentiation of monocytes into macrophages and stimulates proliferation of T helper cells and macrophages. Journal of leukocyte biology. 2010;88(2):279-90.
6. Vettenranta K, Raivio K. Activities of key enzymes of purine degradation and re-utilization in human trophoblastic cells. Placenta. 1988;9(1):27-37.
7. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis tool web services from the EMBL-EBI. Nucleic acids research. 2013;41(W1):W597-W600.
8. Garg VK, Avashthi H, Tiwari A, Jain PA, Ramkete PW, Kayastha AM, et al. MFPPI–multi FASTA ProtParam interface. Bioinformation. 2016;12(2):74.
9. Emanuelsson O, Brunak S, Von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. Nature protocols. 2007;2(4):953.
10. He X, Chang S, Zhang J, Zhao Q, Xiang H, Kusonmano K, et al. MethyCancer: the database of human DNA methylation and cancer. Nucleic acids research. 2007;36(suppl_1):D836-D41.
11. Julenius K. NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. Glycobiology. 2007;17(8):868-76.
12. Gupta R, Jung E, Brunak S. Prediction of N-glycosylation sites in human proteins. 2004.
13. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KTB, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal. 2013;32(10):1478-88.
14. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics. 2004;4(6):1633-49.
15. Duckert P, Brunak S, Blom N. Prediction of proprotein convertase cleavage sites. Protein Engineering Design and Selection. 2004;17(1):107-12.
16. 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;39(suppl_1):D561-D8.
17. Pan Y, Kang X. Limited dsRNA editing impedes leukemia stem cells. Trends in Cancer. 2024 Mar 7.
18. Rawat RS, Kumar S. Understanding the mode of inhibition and molecular interaction of taxifolin with human adenosine deaminase. Journal of Biomolecular Structure and Dynamics. 2023 Jan 22;41(2):377-85.
19. Cox JR, Jennings M, Lenahan C, Manion M, Courville S, Blazeck J. Rational engineering of an improved adenosine deaminase 2 enzyme for weaponizing T-cell therapies. Immuno-Oncology and Technology. 2023 Sep 1;19:100394.
20. Arulselvan AJ, Manimuthu MS, Narayanaswamy R, Sankar M, Radhakrishnan N. Molecular Docking Analysis of Selected Urtica dioica Constituents As Human Carbonic Anhydrase II (hCA-II), Human 11 Beta-Hydroxysteroid Dehydrogenases Type 1 (h11beta-HSD1), and Human Dual Specificity Phosphatase (hCDC25B) Inhibitory Agents. Cureus. 2024 Feb 8;16(2).