3rd International and 12th Iranian Conference on Bioinformatics

Mazandaran,Iran | February 28-29, 2024

University of Science and Technology of Mazandaran

Behshahr, Iran

Main Topics :

* AI & Machine Learning in Biology
* Big Data in Biology
* Structural Bioinformatics
* Biological Sequences Analysis
* Modeling in Computational Biology
* Computational Drug Design & Discovery
* Systems Biology

SilicoGene Intelligence Omics Analysis

CIVILICA

Organizers

Computational Biology Research Center

Iranian Bioinformatics Society

دانشگاه علم و فناوری مازندران

وزارت علوم، تحقیقات و فناوری

دانشگاه علوم پزشکی و خدمات بهداشتی درمانی مازندران

Indexed by

ISC Islamic World Science Citation Center

CIVILICA

WELLCOME MESSAGES

Dr. Ali Ghanbari Sorkhi

University of Science and Technology of Mazandaran

Greetings to all the participants of the 3rd International and 12th National Conference on Bioinformatics. As the scientific secretary of this conference, I am delighted and honored to invite you to this grand conference, which will be held at the University of Science and Technology of Mazandaran, located in the beautiful city of Behshahr.

This conference aims to bring together the bioinformatics community from Iran and beyond, to showcase the latest developments and innovations in this exciting and interdisciplinary field. The conference program consists of a pre-conference workshop on Artificial Intelligence in Bioinformatics on February 27, followed by seven keynote speeches by distinguished speakers from different countries on February 28 and 29. In addition, there will be numerous oral and poster presentations by researchers and students, covering a wide range of topics and applications in bioinformatics.

This conference is the result of the collaborative efforts of many individuals and organizations, including the University of Science and Technology of Mazandaran, the Iranian Bioinformatics Society, and others. I would like to express my sincere appreciation to all of them for their support and contribution. I hope that this conference will provide a fruitful and stimulating opportunity for networking, learning, and sharing among the bioinformatics community, and will enhance the recognition and impact of bioinformatics in Iran and the world.

Dr. Ali Ghanbari Sorkhi
Scientific Chair of the 3rd International and 12th National Iranian Conference on Bioinformatics

WELLCOME MESSAGES

Dr. Jamshid Pirgazi

University of Science and Technology of Mazandaran

As a member of the bioinformatics community in Iran, I am filled with immense satisfaction and gratitude that we, by God's power and blessing, have successfully laid the groundwork for the 3rd International and 12th National Conference on Bioinformatics in Iran. This prestigious event is set to take place at the University of Science and Technology of Mazandaran, placed in the historic city of Behshahr.

We are thrilled to welcome you to a diverse program that includes a workshop entitled Artificial Intelligence in Bioinformatic on the day preceding the conference, February 27. This will be followed by seven keynote speeches delivered by esteemed speakers from various corners of the globe. The second and third days, February 28 and 29, will feature a significant number of oral presentations and poster sessions, providing a platform for sharing cutting-edge research and innovative ideas in the field of bioinformatics.

This monumental event has been made possible through the collective efforts of numerous individuals from the University of Science and Technology of Mazandaran, the Iranian Bioinformatics Society, and various other organizations. I extend my deepest gratitude to each one of them for their invaluable contribution. It is my sincere hope that this international conference will serve as a catalyst for the exchange of knowledge and ideas, fostering collaborations, and enhancing the visibility of bioinformatics, its practitioners, and researchers at both the national and international levels.

Dr. Jamshid Pirgazi
Executive Secretary of the 3rd International and 12th National Iranian Conference on Bioinformatics

TABLE OF CONTENT

Plan Table

Detailed schedule

28th febraury

ICB12 Schedule

28th February 2024

Iranian Conference on Bioinformatics

Time	Speaker	Title
7:00-8:00		Breakfast + admission
8:00-8:50	Dr. Rabbani, Dr. Sadeghi, Dr. Eslahchi	Conference opening ceremony
9:00-9:45	 Hesam Montazeri	Statistical Methods for Cancer Genomics: From Cancer Dependency Maps to Driver Discovery
9:50-10:10	Babak Bahri Alabadi	Unlocking Drug Combinations: A Novel Graph-Based Approach for Predicting Anti-Cancer Synergy
10:10-10:30	Reyhaneh Mortazavi	A Siamese neural network for immunotherapy response prediction
10:30-11:00	Break and Poster presentations	
11:00-11:30	 Seyed Amir Malekpour	Boolean Logics in Gene Regulatory Inference from Single Cell RNA-seq data
11:35-11:55	Mohammad Taheri-Ledari	A method to visualize pseudo-potential landscape of high-dimensional Boolean networks
12:00-12:20	Ali Karimi	DeepSiPPI: Enhancing Protein-Protein Interaction Prediction through Siamese Neural Networks and Sequence-to-Image Transformation
12:20-14:00	Lunch and Poster presentations	
14:00-14:45	 Syed Helder	Machine learning approaches for understanding inter- and intra-patient heterogeneity in breast cancer using bulk and single cell profiling
14:50-15:10	Shokoofeh Ghiam	A Deep Neural Network Method to Reveal Genetic Links Among COVID-19, Alzheimer's Disease, and Multiple Sclerosis Based on Single-Cell RNA-seq
15:15-15:35	Parishad Mokhber	Enhancing Single-Cell RNA Sequencing Analysis through Federated Learning: A Privacy-Preserving Approach to Mitigating Batch Effects
15:40-16:00	Sajedeh Bahonar	A Novel Mechanistic Simulation Model for Single-Cell DNA Sequencing
16:00-16:30	Break and Poster presentations	
16:30-16:50 (online)	 Maryam Hassanlou	Deciphering the c-MYC miRNA Regulatory Network through Machine Learning
16:55-17:15	Hamed Emami	Unraveling Protein Dynamics in the Presence of Cellulose Nanocrystals: An Essential Dynamics Analysis
17:30-18:15	Emad Tajkhorshid	Visualizing Functional lipid–Protein Interactions with Computational Microscopy
18:15-21:00	Networking and Dinner	

29th febraury

ICB12

Schedule
29th February 2024

Iranian Conference on Bioinformatics

Time	Speaker	Title
8:00-8:45		• Breakfast + admission
• 9:00-9:45 • 9:45-10:05 • 10:10-10:30 (online)	 Mehdi Sadeghi • Naser Elmi • Zeynab Arman	• Cellular decision making, noise, differentiation and spatial arrangement • Identification of specific gene modules related to the Meyloid Blast Crisis (MBC) phase of CML Using MAGI Algorithm • A Hybrid Deep Learning-based Model for Off-target Prediction in CRISPR System
10:30-11:00 Break and Poster presentations		
• 11:00-11:45 • 11:45-12:05 • 12:10-12:30	 Laleh Haghverdi • Mina Karimpour • Fereshteh Fallah	• Inference of temporal dynamics from high-throughput single-cell omics data • Single-Cell Perturbation Response Prediction using Conditional Autoencoders • Deciphering Tumor Microenvironment Dynamics and Cellular Communication in Breast Cancer Progression through Single-Cell RNA Sequencing
12:30-14:00 Lunch and Poster presentations		
• 14:00-14:45 • 14:50-15:10 • 15:15-15:35 • 15:40-16:00 (online)	 Jamshid Pirgazi • Seyed Hossein Khoshraftar • Seyed Alireza Khanghahi • Milad Lagzian	• Drug target interaction prediction based on machine learning methods • Identification and Analysis of Crucial Genes Underlying Gastric Cancer via Bioinformatics Techniques • Exploring Traditional Machine Learning and Deep Learning for Predicting Intracytoplasmic Sperm Injection (ICSI) Success Rates • Logical engineering of cellulase enzyme isolated from Acidothermus cellulolyticus bacteria for increasing thermal stability
16:00-16:30 Break and Poster presentations		
• 16:30-17:15	 Martin Nowak	• Evolutionary Dynamics
• 17:15-18:15		• Conference closing ceremony
18:15-21:00 Networking and Dinner		

Keynote Speaker

Martin Nowak

Affiliations

- Distinguished Professor of Mathematics and Biology at Harvard, previously at Oxford and Princeton

Main Research Areas

- Evolution, evolutionary dynamics, evolutionary game theory, adaptive dynamics

Awards and Honors

- Director of Harvard's program for evolutionary dynamics from 2003 until 2020
- Published over 700 academic papers with 140,000+ citations and four critically acclaimed books
- Author of "Evolutionary Dynamics: Exploring the Equations of Life," widely praised in the scientific community

Keynote Speaker

Emad Tajkhorshid

Affiliations

- J. W. Hastings Endowed Chair of Biochemistry
- Professor of Chemistry, Bioengineering, Biophysics, and Neuroscience

Main Research Areas

- Drug Discovery, Infectious Diseases, Ion Channels, Membrane Biology, Molecular Pharmacology, Neuroscience Protein Biochemistry and Protein Structure, Signal Transduction, Virology

Awards and Honors

- Director, NIH Resource Center for Macromolecular Modeling and Visualization at Beckman Institute for Advanced Science and Technology and University of Illinois at Urbana Champaign
- Authored more than 300 research articles with over 40,000 citations in top-tier journals such as Nature, Science, Cell, and PNAS
- Served on the Editorial Boards of multiple major journals, including Biophysical Journal, Journal of Biological Chemistry, PLoS Computational Biology, and Biochemical and Biophysical Research Communication

Keynote Speaker

Mehdi Sadeghi

Affiliations

- National institute of genetic engineering and biotechnolog

Main Research Areas

- Bioinformatics
- Theoretical biology

Keynote Speaker

Jamshid Pirgazi

Affiliations

- Professor of Artificial Intelligence University of Science and Technology of Mazandaran

Main Research Areas

- Bioinformatics, Deep Learning, Machine Learning, Transformers, and Pattern RecognitionProtein Biochemistry and Protein Structure, Signal Transduction, Virology

Awards and Honors

- heads the Machine Learning Research Lab at the University of Science and Technology of Mazandaran
- has done extensive research in the field of DTI drug-target prediction analysis

Keynote Speaker

Syed Haider

Affiliations

- Institute of Cancer Research, London

Main Research Areas

- Bioinformatics
- Machine Learning
- Computational Biology
- Cancer Genomics

Awards and Honors

- has been leading the Breast Cancer Research Bioinformatics Group at The Institute of Cancer Research in London since November 2016
- His team utilizes bioinformatics and machine learning to pinpoint therapeutic vulnerabilities and biomarkers in breast cancer
- Previously, he played pivotal roles at the University of Oxford and the Ontario Institute for Cancer Research, making significant contributions to the field of oncology
- Dr. Haider's expertise lies in developing computational approaches for discovering target genes and biomarkers related to breast cancer risk

Keynote Speaker

Laleh Haghverdi

Affiliations

- Berlin Institute for Medical Systems Biology

Awards and Honors

- Awarded the Helmholtz Doctoral Prize and Erwin Schrödinger Prize for her outstanding research in Geometric Diffusions for Reconstruction of Cell Differentiation Dynamics
- Dr. Laleh Haghverdi and her team develop efficient computational methods for analysis of large single-cell omics data sets, resolution of complex lineage trees and temporal processes, and data

Keynote Speaker

Hesam Montazeri

Affiliations

- Assistant Professor at the Department of Bioinformatics, IBB, University of Tehran

Main Research Areas

- Probabilistic Graphical Models
- Statistical Machine Learning
- CancerGenomics
- bioinformatics
- EvolutionaryModels
- Phylogenetics

SYSTEMS BIOLOGY

Systems Biology is a research field focused on comprehending entire biological systems, such as protein complexes, metabolic pathways, and gene regulatory networks. It employs computational and mathematical analyses to understand how cells, tissues, and organisms function as interconnected systems. The systemic perspective acknowledges the interdependence of biological components, where the behavior of individual parts influences the entire system.

Systems Biology is a scientific approach that studies living organisms as integrated and interconnected systems. Instead of focusing on individual components, it examines how genes, proteins, cells, and organs collaborate to create the complex functions of life. It employs computational models and advanced technologies to analyze and understand the dynamic interactions within biological systems. This holistic perspective allows researchers to unravel the underlying principles governing the behavior of organisms.

SYSTEMS BIOLOGY

Evolutionary Stages of Systems Biology:

Molecular Biology to Systems Molecular Biology: Transformation from understanding gene structure to exploring molecular pathways and networks.
Systems-Mathematical Biology: Incorporating general systems theory and nonlinear dynamics to study living organisms.
Convergence of Molecular and Mathematical Systems Biology: Integration leading to advancements in systems-based medicine, biotechnology, and drug development.
Technological Advances:

The advent of high-throughput approaches has revolutionized biological data analysis. High-powered computing technologies enable the observation and understanding of complex biological systems, supporting the transition from single-molecule studies to systems-level investigations.

Major Research Topics:

Gene Regulatory Networks
Modeling Metabolic Interactions
Protective Mechanisms Induced by Antibiotics
Cell Signaling Pathways

- Najarian, K., Najarian, S., Gharibzadeh, S., & Eichelberger, C. N. (2009). "Systems Biology and Bioinformatics: A Computational Approach." CRC Press.
- Palsson, B. (2015). "Systems Biology." Cambridge University Press.
- Klipp, E., Herwig, R., Kowald, A., Wierling, C., & Lehrach, H. (2005). "Systems Biology in Practice: Concepts, Implementation, and Application." John Wiley & Sons

BIG DATA IN BIOLOGY

Big Data in biology refers to the massive and complex datasets generated in biological research, encompassing genomic, transcriptomic, proteomic, and other molecular data. The volume, velocity, and variety of this data require advanced computational and analytical approaches for meaningful interpretation.

Key Aspects:

Genomic Sequencing: The exponential growth in DNA sequencing technologies has led to vast genomic datasets, enabling comprehensive insights into genetic variations and relationships.

Transcriptomics and Expression Data: High-throughput techniques capture gene expression levels across various conditions, providing a holistic view of cellular activity.

Proteomics and Metabolomics: Large-scale studies of proteins and metabolites contribute to understanding complex biological processes and pathways.

BIG DATA IN BIOLOGY

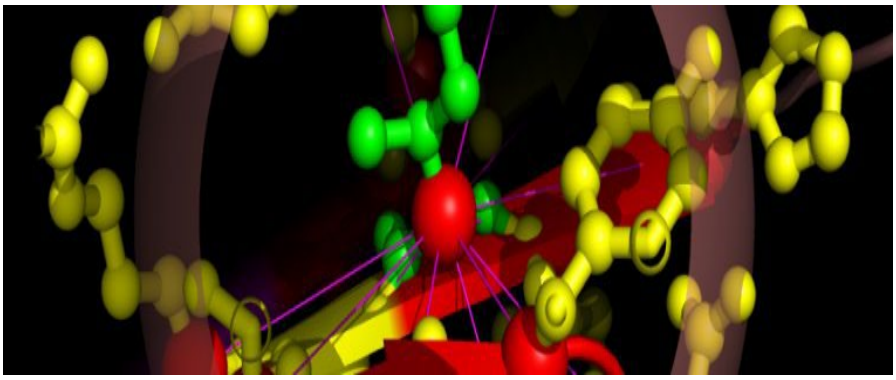
Applications:

- Chen, H., & Mehmood, R. (2017). "Big Data Integration in Bioinformatics." Springer.

Disease Research: Big Data facilitates the identification of biomarkers, understanding disease mechanisms, and personalized medicine.
Drug Discovery: Analyzing vast datasets accelerates the discovery of potential drug targets and aids in drug development.
Biological Networks: Mapping intricate biological networks helps unravel interactions between genes, proteins, and pathways.
Challenges:

Data Integration: Harmonizing diverse datasets for meaningful analysis poses a significant challenge.
Computational Resources: Analyzing Big Data requires substantial computing power and storage capacities.

STRUCTURAL BIOINFORMATICS



In recent decades, remarkable advancements in the fields of biology and computer science have given rise to a dynamic and critically important discipline known as “Structural Bioinformatics.” Structural Bioinformatics was the first major effort to show the application of the principles and basic knowledge of the larger field of bioinformatics to questions focusing on macromolecular structure, such as the prediction of protein structure and how proteins carry out cellular functions, and how the application of bioinformatics to these life science issues can improve healthcare by accelerating drug discovery and development.

Structural bioinformatics focuses on the study and analysis of three-dimensional structures of biological macromolecules such as proteins, nucleic acids, and complexes. It combines principles from biology, computer science and mathematics to reveal the relationship between the structure, function and evolution of these molecules. The history of Structural Bioinformatics as a research field traces back several decades. Initially, the primary focus was on experimental methods for solving biomolecular structures

STRUCTURAL BIOINFORMATICS

With the advent of computational methods and breakthroughs in computer science, the possibility of more accurate and rapid modeling of biomolecular structures became available. Subsequent prominent advances in advanced algorithms, specialized software, and three-dimensional databases of biomolecular structures have transformed Structural Bioinformatics into a pivotal realm in biological research.

One of the key goals of structural bioinformatics is to understand how the structure of a macromolecule relates to its function. By studying the arrangement of atoms and residues in a molecule, researchers can gain insights into its biological activity, such as enzyme catalysis, protein-protein interactions, and molecular recognition processes.

Additionally, structural bioinformatics plays a crucial role in drug discovery and design. By analyzing the structure of target proteins and their binding sites, researchers can identify potential drug candidates and develop computational models to predict their binding affinity and efficacy.

Structural bioinformatics' main goal is to create models and methods for studying and analyzing structural data. As well as solving various biological problems, these methods will help improve our understanding of biological systems. In summary, the purpose of these methods can be categorized into one or more of the following topics: developing databases to collect and store data, visualization and comparison of structures, protein classification, structural analysis, structure prediction, and simulation.

- Gu, J., & Bourne, P. E. (2009). Structural Bioinformatics (2nd ed.). Wiley-Blackwell.
- BYJU'S. (n.d.). Bioinformatics. BYJU'S. https://byjus.com/biology/bioinformatics/
- Samish, I., Bourne, P. E., & Najmanovich, R. J. (2015). Achievements and challenges in structural bioinformatics and computational biophysics. Bioinformatics, 31(1), 146–150. https://doi.org/10.1093/bioinformatics/btu769

BIOLOGICAL SEQUENCE ANALYSIS

Biology sequence analysis, at the intersection of biology and computational sciences, represents a cornerstone in the exploration of genetic information encapsulated in DNA, RNA, and proteins. Emerging in the late 1960s and early 1970s, this field gained prominence with groundbreaking work on tRNA sequence analysis. The transformative moment came with the advent of the Basic Local Alignment Search Tool (BLAST) in 1990, which revolutionized the study of biological sequences and paved the way for subsequent breakthroughs. The field's full potential was showcased during monumental projects like the Human Genome Project, where computational methods were indispensable for decoding the intricacies of the human genome.

Biological sequence analysis relies on a diverse toolkit of algorithms and tools sourced from bioinformatics databases and computational models. Foundational tools such as BLAST and FASTA facilitate sequence comparison and alignment, while advanced methodologies like phylogenetic tree construction, genetic algorithms, Markov chains, and dynamic programming (e.g., Smith-Waterman, Needleman-Wunsch) enable researchers to explore evolutionary relationships.

BIOLOGICAL SEQUENCE ANALYSIS

Scoring matrices like PAM and BLOSUM, Bayesian alignment algorithms, and progressive alignment models (e.g., ClustalW, Gibbs Sampler) enrich the analytical repertoire, empowering researchers to glean meaningful insights from biological data.

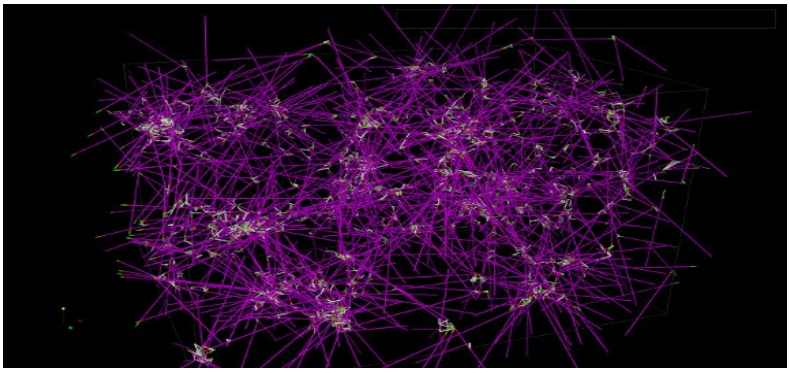
The impacts of biological sequence analysis are far-reaching, influencing realms such as medicine, evolutionary biology, and biotechnology. Computational approaches have facilitated the development of targeted therapies for genetic disorders, deepened our understanding of species evolution, and propelled advancements in personalized medicine. As the field continues to evolve, the integration of cutting-edge technologies like artificial intelligence and machine learning promises to unlock new dimensions in understanding biological sequences. These innovative approaches hold the potential to drive transformative discoveries across scientific disciplines, opening avenues for novel insights and applications in the intricate tapestry of life sciences.

There have been a number of works in this field, including:

1. RNA structure analysis and prediction
2. Comparison of sequences
3. Global and local sequence alignment
4. Improving database searching by sequence
5. Multiple-sequence alignment
6. Pattern and profile methods of identifying distant homologs
7. Genomic analysis
8. Protein structure prediction

- *Briefings in Bioinformatics,* Volume 20, Issue 4, July 2019, Pages 1280–1294, "BioSeq-Analysis: a platform for DNA, RNA and protein sequence analysis based on machine learning approaches " https://doi.org/10.1093/bib/bbx165
- R., Sean. E, Krogh A., G. M. (1998) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. illustrate. Edited by R. Durbin. Cambridge University Press, 1998. Available

MODELING IN COMPUTATIONAL BIOLOGY



Computational biology is an interdisciplinary field that applies techniques from computer science and mathematics to the understanding and modeling of biological systems. Modeling plays a crucial role in computational biology, allowing researchers to simulate and analyze complex biological processes

Bioinformatics emerged in the 1970s, driven by the need to analyze biological data using computational methods.The Human Genome Project, starting in 1990, exemplified computational biology's impact, mapped 85% of the human genome by 2003, with continued efforts leading to a "complete genome" by 2021.In addition to helping sequence the human genome, computational biology has helped create accurate models of the human brain, map the 3D structure of genomes, and model biological systems

MODELING IN COMPUTATIONAL BIOLOGY

Computational biology finds diverse applications in various biological domains:

Anatomy: Computational anatomy focuses on modeling and simulating biological structures at the visible or gross anatomical scale. It leverages 3D measurements, such as magnetic resonance imaging, to extract anatomical coordinate systems.

Data and Modeling: Mathematical biology uses theoretical models to understand biological systems, drawing on discrete mathematics, topology, Bayesian statistics, and algebra. Bioinformatics, a key subfield, involves storing, retrieving, and analyzing biological data, often focusing on genetics.

Systems Biology: This field explores interactions within biological systems, employing computational techniques from biological modeling and graph theory to study complex interactions at cellular levels.

Evolutionary Biology: Computational biology aids evolutionary biology by using DNA data for phylogenetic reconstruction, fitting population genetics models, and predicting evolutionary outcomes.

Genomics: Computational genomics studies genomes, such as the Human Genome Project, aiming to sequence entire genomes for personalized medicine. Sequence homology and alignment are essential tools for comparing genetic data.

Overall, computational biology enhances our understanding of biological systems across scales, from anatomy and evolution to genomics and neuroscience.

In conclusion, computational modeling is a powerful tool in the field of biology that enables researchers to study complex biological systems and processes. By combining mathematical and computational techniques with biological knowledge

- Hogeweg, Paulien (7 March 2011). "The Roots of Bioinformatics in Theoretical Biology". PLOS Computational Biology. 3. 7 (3): e1002021. Bibcode:2011PLSCB…7E2021H. doi:10.1371/journal.pcbi.1002021. PMC 3068925. PMID 21483479.
- Biomedical Information Science and Technology Initiative. (2000, July 17). NIH working definition of bioinformatics and computational biology
- Ninh Laboratory of Computational Biology. (2013, February 18). The Sub-fields of Computational Biology. Retrieved April 18, 2022

COMPUTATIONAL DRUG DESIGN AND DISCOVERY

Pharmaceutical compounds, commonly known as medications, are recognized chemical entities capable of inducing biological alterations within living organisms, ranging from the activation to the inhibition of proteins associated with various diseases. Their purpose extends to the treatment, diagnosis, enhancement, and prevention of medical conditions. In instances where existing medicinal solutions prove inadequate, the initiation of de novo drug discovery projects becomes imperative. These endeavors embark on the quest for entirely new therapeutic agents, aiming to uncover innovative compounds that can address unmet medical needs and broaden the scope of available treatments.

In the realm of drug discovery, computational methods have emerged as powerful tools, revolutionizing traditional, time-consuming processes. This interdisciplinary approach integrates principles from chemistry, biology, informatics, and computer science to accelerate the identification and design of novel pharmaceutical compounds. Computational methods employ algorithms and simulations to predict and analyze the interactions between drugs and biological targets, significantly expediting the identification of potential candidates.

COMPUTATIONAL DRUG DESIGN AND DISCOVERY

- Computational Omics Lab, Centre of Bioinformatics, University of Allahabad, Prayagraj, Indi.(2020). Computational Approaches for Drug Target Identification. https://link.springer.com/chapter/10.1007/978-981-15-6815-2_8

Drug design, an inventive process rooted in the understanding of biological targets, involves crafting new medications. This intricate process revolves around designing molecules that align in shape and charge with the biomolecular target, aiming for a precise and effective therapeutic outcome.

The integration of computer systems has emerged as a catalyst for transforming drug design and discovery, mitigating costs, and enhancing efficiency. This approach, known as Computer-Aided Drug Design (CADD), leverages the proliferation of databases, the increased computational power of systems, and the evolution of computational methods across scientific disciplines. CADD encompasses diverse theoretical and computational approaches that have become integral to modern drug design and discovery.

Some of main subjects in this topic are available as follows:

- Computational drug design
- Bioactivity prediction of compounds
- Adverse drug reaction predictions
- Drug-target binding affinity prediction
- Drug repurposing
- Drug-target association prediction
- QSAR
- Molecular docking studies

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN BIOLOGY

In recent years, the convergence of artificial intelligence (AI) and machine learning (ML) with the field of biology has opened up unprecedented avenues for advancing biological research and understanding complex biological systems.

.In the early stages, AI applications were primarily focused on automating routine laboratory tasks, but the landscape rapidly evolved with the increasing availability of biological data. Machine learning algorithms began to play a pivotal role in deciphering complex biological patterns, aiding in genomics, proteomics, and drug discovery. Over time, AI has proven instrumental in predicting protein structures, identifying biomarkers, and understanding intricate biological processes. The marriage of AI and biology has not only accelerated the pace of scientific discovery but has also paved the way for personalized medicine, offering tailored treatments based on an individual's genetic makeup. As technology advances, the synergy between AI and biology continues to deepen, promising groundbreaking insights and innovations in the understanding and manipulation of living systems.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN BIOLOGY

- Bhardwaj, A., Kishore, S., & Pandey, D. K. (2022)." Artificial Intelligence in Biological Sciences". Life, 12(9), 1430. https://doi.org/10.3390/life12091430
- Hassoun, S., Jefferson, F., Shi, X., Stucky, B., Wang, J., Rosa, E. Jr., … (2021). Artificial Intelligence for Biology. Journal of Bioinformatics, 15(3), 123-145. https://doi.org/10.1093/icb/icab188

Applications of artificial intelligence and machine learning in biology are diverse and encompass various areas. Below are some of these applications:

1.Protein Structure Prediction:

• Using machine learning algorithms to predict the three-dimensional structure of proteins.

• Identifying common patterns in protein structures using advanced models.

2.Disease Prediction:

• Early detection of various diseases based on biological data using artificial intelligence.

• Predicting the likelihood of genetic diseases through the analysis of genetic sequences with machine learning algorithms.

3.Personalized Medicine:

• Tailoring personalized treatments based on individual genetic and biological information using machine learning.

• mproving accuracy in diagnosis and intervention with diseases using machine learning algorithms.

4.Drug Development:

• Designing drug molecules using machine learning algorithms based on biological needs.

WORKSHOPS

Dr. Changiz Eslahchi
Professor of algorithm in bioinformatics, Department of Computer Science, Shahid Beheshti University

Dr. Kaveh Kavoosi
Professor in bioinformatics, Department of bioinformatic, Tehran University

Dr. Hezam Montazeri
Assistant Professor in bioinformatics, Department of bioinformatic, Tehran University

Dr. Alireza Fotuhi Siah Pirani
Assistant Professor in bioinformatics, Department of bioinformatic, Tehran University

Dr. Ali Ghanbari Sorkhi
Assistant professor in bioinformatic, Department of Computer Engineering, University of Science and Technology of Mazandaran

Dr. Sobhan Ahmadian
M.Sc. Student of bioinformatics, Amirkabir University of Technology

Theory:

- **Transformers**
- **Autoencoders**
- **Siamese Networks**
- **Multimodal Models**
- **Deep Neural Networks**
- **Recommender Systems**
- **Graph Neural Networks**
- **Large Language Models**
- **Convolutional Networks**
- **Generative Adversarial Networks**

Practical:

- **Deep Learning with Pytorch**

COMMITTEESS LIST

Dr. Ahmad Reza Rabbani
Conference Chairman

Dr. Ali Ghanbari Sorkhi
Scientific Chair

Dr. Changiz Eslahchi
Scientific Chair

Dr. Jamshid Pirgazi
Executive Secretary

SCIENTIFIC COMMITTEESS

Mehdi Sadeghi

Scientific Committee

Fatemeh Zare-Mirakabad

Scientific Committee

Bahram Goliaei

Scientific Committee

COMMITTEE LIST

Hamid Pezeshk

Scientific Committee

Kaveh Kavousi

Scientific Committee

Sajjad Gharaghani

Scientific Committee

COMMITTEE LIST

Mohammad Hossein Karmi-Jafari

Scientific Committee

Najmeh Salehi

Scientific Committee

Changiz Eslahchi

Scientific Committee

COMMITTEE LIST

Jamshid Pirgazi

Scientific Committee

Hesam Montazeri

Scientific Committee

Ali Ghanbari Sorkhi

Scientific Committee

COMMITTEE LIST

Alireza Fotuhi Siahpirani

Scientific Committee

Mohammad Ganj Tabesh

Scientific Committee

Seyed-Amir Mirashi

Scientific Committee

COMMITTEE LIST

Seyed Shahriar Arab

Scientific Committee

Yazdan Asgari

Scientific Committee

Seyed Abolfazl Montahari

Scientific Committee

COMMITTEE LIST

Ali Sharifi-Zarchi

Scientific Committee

Mohieddin Jafari

Scientific Committee

Mehdi Mirzaie

Scientific Committee

COMMITTEE LIST

Abbasali Emamjomeh

Scientific Committee

Ghasem Hoseini Salekdeh

Scientific Committee

Esmaeil Shahsavand

Scientific Committee

COMMITTEE LIST

Zahra Razaghi-Moghadam

Scientific Committee

Zeynab Mousavian

Scientific Committee

Nazanin Hoseinkhan

Scientific Committee

COMMITTEE LIST

Pegah Khosravi

Scientific Committee

Reza Hassanzadeh

Scientific Committee

Rosa Aghdam

Scientific Committee

COMMITTEE LIST

Zarrin Minuchehr

Scientific Committee

Seyed Amir Malekpour

Scientific Committee

Najmeh Khorasani

Scientific Committee

COMMITTEE LIST

Saeideh Khodaee

Scientific Committee

Zahra Salehi

Scientific Committee

Reza Javanmard Alitappeh

Scientific Committee

COMMITTEE LIST

Mohammad Mehdi Pourhashem Kallehbasti

Scientific Committee

Malihe Danesh

Scientific Committee

Mohammad M. AlyanNezhadi

Scientific Committee

COMMITTEE LIST

Milad Niaz Azari

Scientific Committee

Mostafa Gholami

Scientific Committee

Majid Iranpour Mobarakeh

Scientific Committee

COMMITTEE LIST

Mousa Nazari

Scientific Committee

Mohammad Hossein Olyaei

Scientific Committee

Ehsan Ataei

Scientific Committee

COMMITTEE LIST

Mehdi Yaghoubi

Scientific Committee

Sajad Ahmadian

Scientific Committee

Vahid Ghasemi

Scientific Committee

COMMITTEE LIST

Ali Kermani

Scientific Committee

Omid Abbaszadeh

Scientific Committee

Hassan Mohabatkar

Scientific Committee

COMMITTEE LIST

Mohammad Reza Bakhtiarizadeh

Scientific Committee

Yousof Gheisari

Scientific Committee

Mohammad Hossein Rohban

Scientific Committee

COMMITTEE LIST

Ali Najafi

Scientific Committee

Mahya Mehrmohamadi

Scientific Committee

EXECUTIVE COMMITTEE

Malihe Danesh

Publication Chair

Reza Javanmard Alitappeh

International Relations Chair

Mohammad Mehdi Pourhashem Kallehbasti

Industrial Relations Chair

COMMITTEE LIST

Ali Kermani

Financial Chair

Mohammad M. AlyanNezhadi

IT Support

Mousa Nazari

Workshop Chair

Majid Mokhlesi

Financial Chair

STUDENT EXECUTIVE COMMITTEE LIST

Abolfazl Mohammadi Shakbae

Mohhadeseh Keyhanian

Mehrshad Najafi

Parnian Modanlo

Kiana Babazadeh

Jalal Alipour

Amir Rezaali

Mohammadreza Shakouri

Hesam Pourrostami

COMMITTEE LIST

Alireza Khorramfard

Ali Karimi

Nastaran Zarafshan

Halime Davaji

Seyede Zahra Ahmadi

Amir Hesam
Khoda karami

Mehdi Hoseinzadeh

Akram Sabokro

Niloufar Sedighzad

COMMITTEE LIST

Seyed Valiolah Ekraseyedi

Ehsan Mollaei

Maryam Khajevand

Zahra Habibi

Zahra Farrokhi

Hosein Hoseinifar

Hosein Ranayi

Hosein Samadi

Fereshteh Noroozi

COMMITTEE LIST

Farideh Bahari

Mahsa Sa'adat

Milad Besharatifard

Maryam Rastegari

Sajedeh Bahonar

Fatemeh Reisi

Abstracts

ICB12-1402-1013

Identification of Novel miRNAs Encoded by MSI1 Gene Using Bioinformatic and Experimental Methods

L. Lachinani[a], M. Forouzanfar[a], K. Dormiani[a*], B. M. Soltani[b], K. Dolatshahi[c], S. M. Hakimian[d,e], S. Dokanehiifard[f], M. H. Nasr Esfahani[a*]

a Department of Animal Biotechnology, Cell Science Research Center, Royan Institute for Biotechnology, ACECR, Isfahan, Iran.

b Department of Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran.

c Department of Medicine, Najafabad Branch, Islamic Azad University, Najafabad, Iran.

d Ordibehesht Breast Clinic, Isfahan, Iran.

e Poursina Hakim Digestive Diseases Research Center, Isfahan University of Medical Sciences, Isfahan, Iran.

f Department of Human Genetics, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Biomedical Research Building, Miami, FL, USA.

*Corresponding authors' email addresses: k_dormiani@royaninstitute.org; mh.nasr-esfahani@royaninstitute.org

Abstract: miRNAs play critical roles in many cellular processes and developmental pathways. It is documented that aberrant expression of miRNAs is associated with several disorders [1]. RNA-binding protein Musashi1 (MSI1) shows an increased expression level in several cancers and has been introduced as a prognostic marker in some malignancies [2]. It is expected that if any miRNA is encoded by this gene, it might have a role in cancer development or could be considered as a prognostic biomarker. Accordingly, in this study, we aimed to find novel miRNA(s) inside the intronic regions of the MSI1 gene. For discovery of novel miRNAs, several approaches including experimental and computation-driven methods are employed [3]. In this study we used the combination of both experimental and computation-driven methods to discover novel miRNAs inside MSI1 gene. We employed several bioinformatics software to predict novel miRNA precursors, evaluate the ability of producing mature miRNAs by predicted stem-loop structures and predict target genes for the potential mature miRNAs. Among several miRNA precursors predicted by bioinformatics studies, we selected two precursors within intron 4 of MSI1 gene, named MSM2 and MSM3. For experimental analysis, corresponding precursor miRNAs were transfected into HEK293T cells and exogenous expression of the mature miRNAs were detected by RT-qPCR and sequencing analysis. Two mature miRNAs, MSM3-3p and MSM3-5p were generated by MSM3 precursor and one, MSM2-5p, was derived from MSM2. Moreover, endogenous expression of MSM2-5p and MSM3-3p was detected in MCF-7 and SH-SY5Y cell lines. Additionally, the interaction between the MSM3-3p and 3′UTR region of PDE11A was confirmed by using luciferase assay. Overall, our data demonstrated that MSI1 gene encodes two novel miRNAs.

Keywords: Novel miRNAs; MSI1 gene; Cancer; Bioinformatics

References

[1] J. Lu, G. Getz, "MicroRNA Expression Profiles Classify Human Cancers", Nature, 435 (2005), 834–838.

[2] A. E. Kudinov, J. Karanicolas, "Musashi RNA-Binding Proteins as Cancer Drivers and Novel Therapeutic Targets", Clin. Cancer Res. 23 (2017), 2143–2153.

[3] C. P. Gomes, "A Review of Computational Tools in MicroRNA Discovery", Front. Genet. 4 (2013), 81.

ICB12-1402-1015

Molecular docking simulation of Anticancer properties of Cisplatin as an ovarian cancer inhibitor

Atefi, A[a*]. Zare Bidoki, M[b]. Dehghan, P[c]
a Research and Development Department, Ista NanoPharmed Aria Pharmaceutical Company, Yazd, Iran.
b M.Sc. of Genetics, Faculty of Biological Sciences, Yazd University, Yazd, Iran.
c Research and Development Department, Ista NanoPharmed Aria Pharmaceutical Company, Yazd, Iran.
*Corresponding author's email address: aref.atefi@yahoo.com

Abstract: Cisplatin is a platinum coordination complex (cis-[Pt(NH3)2Cl2]), classified as an alkylating agent, and is a well-known chemotherapeutic drug [1]. It was used for the treatment of numerous human cancers, such as ovarian cancer [2]. FAK, a protein tyrosine kinase, plays a crucial role in regulating cellular functions such as adhesion, motility, proliferation, and survival in different cell types, including its activation and overexpression in advanced cancers, making it a promising target for cancer treatment [3]. This study aimed to delineate the binding affinity of Cisplatin as FAK in ovarian cancer inhibitors using AutoDock4 software. The crystal structure data of Cisplatin N, N-dimethylformamide solvate was obtained from the RCSB site (PDB: 1DDP). The present study used AutoDock 4, Discovery Studio visualizer, MGLTools, and PyMOL software for molecular docking and visualization. The crystallographic structure of FAK (PDB: 1MP8) was used from the RCSB protein data bank. The binding pocket has a volume of 687.721 A3 and a surface area of 586.372 A2, as calculated by the CASTp database. The binding free energy of the studied compound in the best conformation ΔGbind (kcal/mol) was -5.78. The amino acids involved in the connection site with the studied compound have three residues, including GLN686, GLU682, and GLU683, which form three hydrogen bonds in the interaction with the receptor. The findings of this study demonstrated that the bonding of Cisplatin to the FAK protein resulted in a reduction in protein activity. Consequently, Cisplatin exhibits potent potential as a chemotherapeutic agent for the treatment of ovarian cancer.

Keywords: Molecular docking; Cisplatin; Ovarian Cancer; FAK

References

[1] Ghosh, Sumit. "Cisplatin: The First Metal Based Anticancer Drug." Bioorganic chemistry 88 (2019): 102925.

[2]Zoń, Aleksandra, and Ilona Bednarek. "Cisplatin in Ovarian Cancer Treatment—Known Limitations in Therapy Force New Solutions." International Journal of Molecular Sciences 24, no. 8 (2023): 7585

[3] Yoon, Hyunho, Joshua P Dehart, James M Murphy, and Ssang-Taek Steve Lim. "Understanding the Roles of Fak in Cancer: Inhibitors, Genetic Models, and New Insights." Journal of Histochemistry & Cytochemistry 63, no. 2 (2015): 114-28.

ICB12-1402-1019

Enhancing Heart Failure Prediction Accuracy through Effective Preprocessing and Principal Component Analysis

A.Dibaji[a*], S.Sulaimany[a]
a Social & Biological Network Laboratory (SBNA), Department of Computer Engineering, University of Kurdistan, Sanandaj, Iran
*Corresponding author's email address: Abolfazl.Dibaji@uok.ac.ir

Abstract: Accurate prediction of heart failure is crucial for early intervention and preventative care. This study aims to improve prediction accuracy using a Heart Failure Prediction dataset of 299 samples with 12 distinct features and a target variable. We addressed data imbalance using the NearMiss algorithm and normalized the data to ensure uniformity. Subsequently, Principal Component Analysis (PCA) was used to distill the dataset to 7 principal features, which, when aggregate with the original features, formed a restructured dataset. Several machine learning models were evaluated, and the random forest algorithm emerged as the most accurate, achieving an 83.5% prediction success rate. This outcome not only represents a significant improvement over previous studies [1] but also highlights the importance of meticulous preprocessing and feature optimization in predictive modeling.

Keywords: Heart Failure Prediction; PCA; Machine Learning; Preprocessing; Random Forest

References

[1] N. A. Abujabal and A. B. Nassif, "Enhanced Heart Failure Prediction Using Feature Selection-based Machine Learning Models," in *2023 Advances in Science and Engineering Technology International Conferences (ASET)*, IEEE, Feb. 2023, pp. 1–6. doi: 10.1109/ASET56582.2023.10180853.

ICB12-1402-1020

Molecular docking simulation of Antibacterial properties of Teixobactin as Bacillus anthracis inhibitor

Atefi, A[a*]. Zare Bidoki, M[b]. Dehghan, P[c]
a Research and Development Department, Ista NanoPharmed Aria Pharmaceutical Company, Yazd, Iran.
b M.Sc. of Genetics, Faculty of Biological Sciences, Yazd University, Yazd, Iran.
c Research and Development Department, Ista NanoPharmed Aria Pharmaceutical Company, Yazd, Iran.
*Corresponding author's email address: aref.atefi@yahoo.com

Abstract: Anthrax is an acute zoonotic disease caused by Bacillus anthracis. This bacterium is at the top of the list of dangerous agents for use in biological weapons and bioterrorism. A wide range of antibiotics are used to treat Anthrax, but usually, treatment may not be effective [1]. EA1 is an abundant, highly antigenic, surface-layer protein of Bacillus anthracis vegetative cells [2]. Teixobactin is a peptide-like secondary metabolite of some species of bacteria that kills some gram-positive bacteria. It is an inhibitor of cell wall synthesis [3]. This study aimed to delineate the binding affinity of Teixobactin as EA1 with Bacillus anthracis inhibitors. The present study used AutoDock Vina, Discovery Studio visualizer, and MGLTools software for molecular docking and visualization. The crystal structure data of Acetyl-DELTA1-5-Arg10-teixobactin was obtained from the Crystallography Open Database (COD: 7119658). The crystallographic structure of EA1 (PDB: 8OPR) was used from the RCSB protein data bank. The binding pocket was calculated by the SCFBio databank. The binding free energy of the studied compound in the best conformation ΔGbind (kcal/mol) was -6.7. The amino acids involved in the connection site with the studied compound have three residues, including Asn419, Tyr420, and Ser422, which form hydrogen bonds in the interaction with the receptor. The findings of this study demonstrated that the bonding of Teixobactin to the EA1 protein resulted in a reduction in protein activity. Consequently, Teixobactin exhibits potent potential as an antibacterial agent for the treatment of Anthrax.

Keywords: Molecular docking; Teixobactin; Bacillus anthracis; EA1

References

[1] Aref Atefi, Mohammad Hassan Shahhossieni, Hossein Dehghan, Parisa Dehghan, Jamshid Ayatollahi, Robab Shikhpour. "An Overview of Anthrax. " 2018, Sokanvaran Pub, Tehran, Iran.

[2] Williams DD, Turnbough CL Jr. "Surface layer protein EA1 is not a component of Bacillus anthracis spores but is a persistent contaminant in spore preparations. " J Bacteriol. 2004 Jan;186(2):566-9 .

[3] Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, et al. "A new antibiotic kills pathogens without detectable resistance. " Nature 2015. 517 (7535): 455–459.

ICB12-1402-1021

Distribution of Estrogen Receptor 1 (ESR1) gene (rs2234693) polymorphism among breast cancer patients from Mazandaran province, Iran; a case-control study

Pardis Mirzaeyan[1], Mohammad Shokrzadeh[2*], Ali Salehzadeh[3], Farzam Ajamian[4]
1 Department of Biology, Rasht Branch, Islamic Azad University, Rasht, Iran;
(pardis.mirzaeyan@yahoo.com)
2 Pharmaceutical Research Center, Department of Toxicology and Pharmacology, Faculty of Pharmacy, Mazandaran University of Medical Sciences, Sari, Iran;
(mslamuk@yahoo.com)
3 Department of Biology, Rasht Branch, Islamic Azad University, Rasht, Iran;
(salehzadeh@iaurasht.ac.ir)
4 Department of Biology, Faculty of science, University of Guilan, Rasht, Iran;
(ajamian@guilan.ac.ir)
*Corresponding author's email address: mslamuk@yahoo.com

Abstract: Breast cancer accounts for about one-third of woman's cancers today and it has been increasing in Iranian women, recently. Changes in the estrogen signaling pathway, including changes in the estrogen receptor α (ESR1) gene, play an important role in the development of breast cancer. The current study was conducted to investigate the distribution of the rs2234693 ESR1 polymorphism in breast cancer patients. Peripheral blood samples were taken from 60 breast cancer patients and 60 healthy individuals, referring to the Imam Khomeini hospital, Sari, Iran, to determine the (rs2234693 T>C) single nucleotide polymorphism (SNP). Genomic DNA was extracted using the salting-out method and genotyping was performed by PCR-RFLP assay. Statistical analysis was performed using Medcalc version 15 software. frequency of TC, TT, and CC genotypes among healthy controls were 65%, 15%, and 20%, respectively. The patient group had a frequency of 38.5%, 38.5%, and 23% for TC, TT, and CC genotypes, respectively. Our results showed a significant difference in the prevalence of ESR1 gene genotypes between patients and control groups ($\chi 2$= 10.408, p= 0.005). This study showed that the ESR1 polymorphism may be involved with the increased risk of breast cancer development and the rs2234693 T>C SNP could be regarded as potent marker for breast cancer screening approaches.

Keywords: gene polymorphism; Breast Cancer; ESR1; RFLP-PCR.

ICB12-1402-1026

FN1, COMP, and CXCL5, serve as diagnostic biomarkers for pancreatic cancer, as revealed through analysis of microarray data

Parisa Alirezae[a], Seyed Hossein Khoshraftar[a], Saba Hadi[a], Anahita soleimani[a], Mohammad Mostafa Pourseif[b*]
a Department of Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
Department of Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: Introduction Pancreatic cancer (PC) is a malignancy with few warning signs before the disease reaches its final stages. Currently, the early diagnosis of PC is very difficult because most patients have non-specific symptoms, which leads to delaying the correct diagnosis. In this study, an attempt was made to discover biomarkers based on differentially expressed genes (DEGs) that can act as a driving factor in tumorigenesis.
Methods At the beginning of the study, datasets related to Microarray data extracted by using the GEO database. In the following, quality control was checked with PCA. Subsequently, the LIMMA package was utilized to identified DEGs. Furthermore, the TCGA RNA-seq results were employed to validate our findings. Then to identify biological processes and pathways in which genes with differential expression are involved. GeneCards and Reactome databases were used. Lastly, the relationship between the DEGs and survival time of the disease was assessed using the Kaplan-Meier plotter.
Results: In the studied data set, a number of genes have significant differential expression (adj.P.val <0.05, Log2FC>3). Among them, three genes FN1, COMP, CXCL5, had increased expression. Then, using the information available in GeneCards and Reactome databases, these genes had regulatory control over vital pathways, including PI3K-Akt-mTOR, cellular immune pathways. And the differential rate has been confirmed by different graphs.
Conclusion: This study led to the identification of three genes with increased expression in pancreatic cancer and potentially capable of early detection of this disease.
Keywords: Pancreatic cancer; biomarker; microarray

References

[1] Sung, H.F., J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F, Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. 2021.

[2] Vogell, A. and M.L. Evans, Cancer Screening in Women. Obstetrics and Gynecology Clinics of North America, 2019. 46(3): p. 485-499

ICB12-1402-1027

Pathogenicity prediction of a missense mutation (p.Pro250Arg) in the FGFR3 gene associated with achondroplasia

Pegah Mirjalili[1], Mehri Khatami[1*]
1 Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: m.khatami@yazd.ac.ir

Abstract: Achondroplasia (ACH) is an autosomal dominant disease caused by missense mutations in the FGFR3 (fibroblast growth factor receptor 3) gene, it is the most common cause of short stature in humans. FGFR3 is expressed in chondrocytes and mature osteoblasts and regulates bone growth. These mutations, which cause the gain of FGFR3 function, affect many tissues, especially the cartilaginous growth plate in the growing skeleton, leading to a variety of manifestations and complications. The structure of FGFR3 is Ig-like C2-type 3 and its gene is positioned on 4 p 16.3. We aimed to perform a bioinformatics analysis to find of pathogenic impact of an SNP in FGFR3 at position 250 causes wild-type proline change to mutant arginine amino acid (p.Pro250Arg). Prolines are known to be very rigid amino acid and therefore induce a special backbone conformation which might be required at this position. This missense mutation can disturb this special conformation. This assessment has been analyzed using several bioinformatics tools such as Polyphen-2, SIFT, HOPE, Netsurf-2.0, and UniProt that revealed this variation causes a change in the structure of protein and ultimately disorder in protein function. FGFR3 gene data were obtained from the National Center for Biotechnology Information (NCBI) website. Polyphen-2 analysis indicated that this mutation is predicted to be PROBABLY DAMAGING with a score of 0.999. SIFT for this substitution predicted that to AFFECT PROTEIN FUNCTION with a score of 0.00. In addition, HOPE analysis showed this variant's MetaRNN score is 0.9890754. It can range from 0.0 to 1.0. The higher, the more likely it is to be pathogenic and the wild-type residue is highly conserved, but a few other residue types have been observed at this position too. In conclusion, the mutated residue is located in a domain that is important for the binding of other molecules and in contact with residues in a domain that is important for the activity of the protein. The mutation might affect this interaction and thereby disturb signal transfer from the binding domain to the activity domain.

Keywords: Achondroplasia (ACH); FGFR3 gene; Missense mutation

References

[1]Chen S, Dong H, Yong L, Zhang Y, Pan L.Heterozygous variant in FGFR3 underlying severe phenotypes in the second trimester: a case report.BMC Medical Genomics.2023;16:80:1-5

[2]Hogler W, M.Ward.L . New developments in the management of achondroplasia: Wien Med Wochenschr.2020; 170:104-111

ICB12-1402-1033

Identification and Analysis of Crucial Genes Underlying Gastric Cancer via Bioinformatics Techniques

Seyed Hossein Khoshraftara[a], Parisa Alirezaiea[a], Saba Hadia[a], Anahita Soleimania[a], Sara Kamrani[b], Akbar Amirfiroozi[a], Mohammad Mostafa Pourseif[c*]
a Department of Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
b Department of Biology, Faculty of Sciences, University of Zanjan, Zanjan, IRAN
c Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: Introduction: Gastric cancer (GC) has become a significant concern globally, particularly in East Asian nations, in recent times. According to the GLOBOCAN database, GC is the fourth most common cancer and the third leading cause of cancer-related deaths worldwide. This study aimed to explore biomarkers based on differentially expressed genes (DEGs) that may act as a driver factor in tumorigenesis.

Two sets (400 samples) of data were obtained from the GEO database, and each of these sets underwent quality analysis through the application of PCA. Subsequently, the LIMMA package was utilized to identify genes that exhibited differential expression. To determine the common DEGs between them, a Venn diagram was employed. Furthermore, the TCGA RNA-seq results were employed to validate our findings. Lastly, the relationship between the DEGs and the survival time of the disease was assessed using the Kaplan-Meier plotter.

In two datasets, 14 genes were found to be significantly dysregulated (adj.P.val < 0.05). Among them, three genes ATP4A, ESRRG, and ADIPOQ show a significant decrease in expression in both datasets. Based on the information provided by the GeneCards and Reactome pathway databases, these genes exert regulatory control over various crucial pathways, including Ion channel transport, Nuclear Receptor transcription pathway, and AMPK Signaling Pathway. According to Kaplan-Meier plots DEGs are correlated with poor prognosis in GC.

Our investigation revealed that the three genes ATP4A, ESRRG, and ADIPOQ play significant roles in GC and they are correlated with poor prognosis in GC.

Keywords: Gastric Cancer; Biomarker; Microarray; Bioinformatics Analysis

ICB12-1402-1036

Prediction of Pathogenicity of rs1252692732 p.Gly151Val missense mutation (rs1252692732) in MTTP Gene Associated with Abetalipoproteinemia by Bioinformatics Tools

Fateme Zahra Kazempour[1*], Mehri Khatami[1]
1 Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: kzpfateme@gmail.com

Abstract: Abetalipoproteinemia typically presents in infancy with failure to thrive, diarrhea, vomiting, and malabsorption of fat. Hematologic manifestations may include acanthocytosis (irregularly spiculated erythrocytes), anemia, reticulocytosis, and hemolysis with resultant hyperbilirubinemia. .[1] Abetalipoproteinemia is caused by a homozygous autosomal recessive mutation in the MTTP gene. More than 33 mutations that cause the disease have been identified. This gene codes a microsomal triglyceride protein (MTP) that mediates intracellular chylomicron or VLDL assembly and transport in the intestinal mucosa and hepatocytes.[2] Untreated individuals may develop spinocerebellar degeneration and retinitis pigmentosa. [3] This study was conducted to find the pathogenicity impact of an SNP in MTTP (rs1252692732). The MAF of this substitution by ClinVar was predicted to be 0.00001. A substitution at position p.Gly151Val was predicted to affect protein function with a score of 0.00 by SIFT, and this mutation is likely to be harmful with a score of 1.000 by Polyphen-2. The MetaRNN score of this type by Hope is 0.9744465. The mutant residue is larger than the wild-type residue and is more hydrophobic than the wild-type residue. Consequently, the mutation is located in a domain annotated in UniProt as deleterious. Then, this missense mutation can destroy performance and can be pathogenic.

Keywords: Abetalipoproteinemia; MTTP gene; VLDL; Bioinformatics

References

[1] Burnett JR, Hooper AJ, Hegele RA. Abetalipoproteinemia. 2018 Oct 25 [updated 2022 May 19]. MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Amemiya A, editors. Seattle (WA): University of Washington, Seattle; 1993–2023. PMID: 30358967.
[2] Junaid SZS, Patel K. Abetalipoproteinemia. 2023 Jul 31. Publishing; 2023 Jan–. PMID: 30020727..
[3] Welty FK. Familial hypobetalipoproteinemia and abetalipoproteinemia. InCholesterol 2022 Jan 1 (pp. 465-480). Academic Press.

ICB12-1402-1040

A Bioinformatics study: predicting the pathogenicity of two SNPs in the CFTR gene which is associated with Cystic Fibrosis (CF) disease

Farideh Askari[1*], Mehri Khatami[1], Reyhane Chamani[1]
1 Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: faride.ask1381@gmail.com

Abstract: Cystic fibrosis (CF) impacts various physiological systems, including respiratory, exocrine pancreas, digestive, hepatobiliary, and sweat glands, with the CFTR gene regulating ion and water secretion and absorption in epithelial tissues. The most frequently encountered mutation associated with CF is DeltaF508, which results in compromised folding and trafficking of the encoded protein. Our study aimed to identify pathogenic single nucleotide polymorphisms (SNPs) within the CFTR gene using bioinformatics servers. Two SNPs, rs397508718 (G149R) and rs397508729 (T161N and T161D) were studied using SIFT, PolyPhen-2, I-Mutant 2.0, and Hope servers. Results showed that rs397508718 affected protein function with a score of 0.02 by the SIFT and was predicted to be possibly damaging with a score of 1.000 by PolyPhen-2 programs. I-Mutant analysis showed an increase in stability for the mutant state. In addition, the Hope server indicated that the wild-type (G149) residue is the most flexible amino acid. This flexibility might be necessary for the protein's function. The rs397508729 affected protein functions with a score of 0.00 by SIFT and was predicted to be damaging with a score of 1.000 by PolyPhen-2 programs. I-Mutant analysis showed also an increase in stability. Hope also indicated that the size difference between wild-type (T161) and mutant residues (N161 and D161) causes the new residue to not be in the correct position to make the same hydrogen bond. In conclusion, the study suggests that the G149R, T161N, and T161D variants of CFTR could affect protein function, potentially be pathogenic, and damage the protein.

Keywords: CFTR gene; SIFT; PolyPhen-2; I-Mutant 2.0; Hope

References

[1] Savant A, Lyman B, Bojanowski C, Upadia J. Cystic Fibrosis. 2001 Mar 26 [updated 2023 Mar 9]. In: Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Amemiya A, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2023. PMID: 20301428 .

[2] Jia S, Taylor-Cousar JL. Cystic Fibrosis Modulator Therapies. Annu Rev Med. 2023 Jan 27;74:413-426. doi: 10.1146/annurev-med-042921-021447. Epub 2022 Aug 16. PMID: 35973718 .

[3] McDonald EF, Meiler J, Plate L. CFTR Folding: From Structure and Proteostasis to Cystic Fibrosis Personalized Medicine. ACS Chem Biol. 2023 Oct 20;18(10):2128-2143. doi: 10.1021/acschembio.3c00310. Epub 2023 Sep 20. PMID: 37730207; PMCID: PMC10595991

ICB12-1402-1041

Identification of hub genes as novel biomarkers in Ovarian Cancer using High-Throughput Transcriptome data analysis

Saba Hadi[a], Anahita Soleimani[a], Seyed Hossein Khoshraftar[a], Parisa Alirezaie[a], Mohammad Mostafa Pourseif[b*], Mahmoud Shekari Khaniani[b*]
a Department of Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
Department of Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding authors' email addresses: mmpourseif@gmail.com, mahmoud.khaniani@gmail.com

Abstract: Introduction: Ovarian cancer (OC) is a malignancy of the female reproductive system and is the third most common cancer among all cancers in women. Due to the fact patients are diagnosed in advanced stages, it is the fifth cause of cancer death among women [1, 2]. The central objective of this study is to discover novel biomarkers associated with OC to predict prognosis and develop improved treatment strategies.

The GSE66957 dataset was obtained from GEO database, consisting of 57 cancer and 12 normal samples. Principal Component Analysis (PCA) was executed to assess the consistency of the samples. Finally, the statistical analysis was carried out using the LIMMA package. Differentially expressed genes (DEGs) were identified using the criteria of $|log2FC|> 5$ and adj P-value < 0.05. The Kaplan-Meier plotter was employed to assess the correlation between gene expression and survival outcomes. Microarray analysis was conducted to assess the expression profile of 69 samples. A total of 37 DEGs were identified, of which 33 were up-regulated and 4 were down-regulated. Among them, two genes CFB, and PPL have been discovered in previous studies to be related to cancers such as Pancreatic and lung. These two genes show a significant increase in expression in this dataset. Based on the information provided by the GeneCards and Reactome pathway databases, these genes exert regulatory control over various crucial pathways, including complement activation pathways, and innate and adaptive immune systems. According to Kaplan-Meier plots DEGs are correlated with poor prognosis in OC.
Our investigation revealed that the CFB and PPL genes play significant roles in OC and they are correlated with poor prognosis in OC.

Keywords: ovarian cancer; biomarker; microarray

References

[1] Sung, H.F., J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F, Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. 2021.

[2] Vogell, A. and M.L. Evans, Cancer Screening in Women. Obstetrics and Gynecology Clinics of North America, 2019. 46(3): p. 485-499.

ICB12-1402-1044

Computational Modeling and Machine Learning Evaluation of Factors Impacting Regional Nasal Dosage Delivery for Micro-Sized Peptide-Based Therapeutics

Fatemeh Ebrahimi Tarki[a], Mahkame Sharbatdar[b*], Mahboobeh Zarrabi[a*], Ghamartaj Khanbabaei[c], Ahya Abdi Ali[d]

a Department of Biotechnology, Faculty of biological sciences, Alzahra university, Tehran, Iran
b Department of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran
c Department of Pediatric Pulmonology, Mofid Children's Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
d Department of Microbiology, Faculty of Biological Science, Alzahra University, Tehran, Iran
*Corresponding authors' email addresses: m.sharbatdar@kntu.ac.ir; mzarrabi@alzahra.ac.ir

Abstract: Nasal peptide delivery offers advantages over systemic administration [1], [2], but lacks a comprehensive understanding of crucial parameters for regional dosage deposition in the upper respiratory system. This study evaluated Peptide-based therapeutics (PTPs) deposition in a healthy adult male (37 years old). Nasal geometries were constructed from CT-scan images, treating each side of the nostril as a distinct nasal passage model. Computational Flow Dynamics (CFD) simulated PTP trajectories using the Lagrangian tracking approach, considering forces like drag, Saffman's lift, thermophoretic, and Brownian forces [3]. The steady-laminar and steady-kω-SST models incorporated three flow rates (8.7, 15, and 30 L/min) and two flow regimes (laminar and turbulent). Deposition analysis in four nasal regions (vestibule, nasal valve, anterior turbinate, and nasopharynx) was conducted for varying PTP diameter (1–100 µm), spray cone angle (32º, 79º), and injection speed (2, 19.2 m/s). The anterior turbinate emerged as a favorable site for local and systemic nasal drug delivery [4]. In this study, low injection speed and spray cone angle play pivotal roles in maximizing anterior turbinate deposition. Utilizing 24 distinct inhalation and PTP delivery scenarios generated through numerical simulations, machine learning models underwent training with five-fold cross-validation to predict the delivered dose, eliminating the need for future partial differential equation solvers. The random forest and gradient boosting models yielded R2 scores of 0.91 and 0.90. The deposition location in the nasal cavity, the diameter of PTP, and injection velocity emerged as the most crucial factors influencing the delivered dose

Keywords: Intranasal delivery; Therapeutic peptides; CFD; Machine Learning

References

[1] W. Al Bakri, M. D. Donovan, M. Cueto, Y. Wu, C. Orekie, and Z. Yang, "Overview of intranasally delivered peptides: key considerations for pharmaceutical development," Expert Opinion on Drug Delivery, vol. 15, no. 10, pp. 991–1005, Oct. 2018, doi: 10.1080/17425247.2018.1517742.

[2] S. Grassin-Delyle et al., "Intranasal drug delivery: An efficient and non-invasive route for systemic administration," Pharmacology & Therapeutics, vol. 134, no. 3, pp. 366–379, Jun. 2012, doi: 10.1016/j.pharmthera.2012.03.003.

[3] J. S. Kimbell et al., "Characterization of Deposition from Nasal Spray Devices Using A Computational Fluid Dynamics Model of The Human Nasal Passages," Journal of Aerosol Medicine, vol. 20, no. 1, pp. 59–74, Mar. 2007, doi: 10.1089/jam.2006.0531.

[4] M. Kiaee, H. Wachtel, M. L. Noga, A. R. Martin, and W. H. Finlay, "Regional deposition of nasal sprays in adults: A wide ranging computational study," Numer Methods Biomed Eng, vol. 34, no. 5, p. e2968, May 2018, doi: 10.1002/cnm.2968

ICB12-1402-1045

Revolutionizing NSCLC Treatment: Exploring the Potential of In Silico-designed Multiepitope Vaccines

Ata Moghimi[1], Nasrin Banihoseinaian[1], Elham Baghbani[1], Saeid Nour Olyaei[1], Behzad Baradaran[1*]

1 Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

*Corresponding author's email address: behzad_im@yahoo.com

Abstract: Non-Small Cell Lung Cancer (NSCLC) presents a formidable challenge in oncology, necessitating innovative strategies for improved therapeutic outcomes [1]. This study explores cutting-edge immunotherapy, focusing on the potential of in silico-designed multiepitope vaccines to reshape NSCLC treatment. Four NSCLC-associated antigensVimentin, PRAME, TXNDC5, and AKAP4 were retrieved from the NCBI database. B cell prediction involved ABCpred, while T cell CD8+ predictions used NetCTL 1.2 and IEDB servers, and CD4+ predictions employed NetMHCIIpan-4.0. Subsequently, comprehensive analyses, including validation, allergenicity, toxicity, and physicochemical assessments, were performed via web servers. Employing AAY, GPGPG, and KFER as linkers, and HBHA and CpG-Oligonucleotides (ODNs) as adjuvants linked by EAAAK, the final construct underwent disulfide engineering, molecular docking, immune simulation, and codon adaptation for effective vaccine production. Identifying 8 linear B-cell epitopes and 16 and 10 epitopes for CD4+ and CD8+ cells, respectively, the predicted epitopes were linked to facilitate proper protein folding. No allergenicity was observed, and the vaccine exhibited proper antigenicity (0.726722). Molecular docking studies suggested a high affinity for human receptors. Codon optimization and in silico cloning into the E. coli K12 strain indicated potential for expression, with a CAI of 1.0 and GC parameters of 52.82. Notably, most current NSCLC vaccines lack efficacy, especially in targeting all three crucial oncoproteins. Our novel vaccine shows promise and uniqueness, emphasizing the need for continued research to ensure comprehensive evaluation in human subjects. The vaccine, utilizing antigens found in CAR T-cell treatments, is designed for combination use, potentially amplifying inflammatory factors and enhancing CAR T-cell effectiveness [2].

Keywords: NSCLC, Multiepitope vaccine, in silico design, Immunotherapy

References

1. Horvath, L., et al., Overcoming immunotherapy resistance in non-small cell lung cancer (NSCLC)-novel approaches and future outlook. Molecular Cancer, 2020. 19: p. 1-15. 2. Ansah, E.O., A. Baah, and E.B. Agyenim, Vaccine Boosting CAR-T Cell Therapy: Current and Future Strategies. Advances in Cell and Gene Therapy, 2023. 2023.

ICB12-1402-1048

Pathogenicity prediction of a single nucleotide polymorphism (rs104894107: p.G130V) in human FXN gene

Marzieh Charostad[a*], Mehri Khatami[a]
a Department of Biology, Yazd University, Yazd, Iran.
*Corresponding author's email address: m.charostad79@gmail.com

Abstract: Friedreich's Ataxia (FRDA or FA) is an autosomal recessive disorder that causes difficulty in walking, impaired speech, and loss of arm and leg coordination. This condition is caused by a mutation in the FXN gene on chromosome 9 which produces a protein called Frataxin. In FRDA, cells produce less Frataxin which leads to degeneration of nerve tissue in the spinal cord and ataxia. In this study, we have investigated the pathogenicity of one type of SNP, rs104894107 (p.G130V), in this disease. rs104894107 affected the protein function with a score of 0.00 by the SIFT and was predicted to be possibly damaging with a score of 1.000 by the PolyPhen-2 programs. Also, it was checked by I-Mutant, showing a large decrease in its stability with a DDG value of approximately -1.69. Moreover, the HOPE server indicated that the interaction between domains could be disturbed by the mutation, which might affect protein function and the signal transduction between the domains. the wild-type residue (p.G130) is the most flexible of all residues. Mutation of this glycine can abolish this function. While the wild-type residue (p.G130) was buried in the core of the protein, the mutant residue (p.V130) is bigger and probably will not fit. The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles. In conclusion, this study suggested that the SNP variant of FXN would have an effect on the protein function and could be pathogenic and damage protein.

Keywords: Friedreich's Ataxia (FRDA or FA); FXN gene; SNP; Missense mutation

References

[1] Krasilnikova MM, Humphries CL, Shinsky EM. Friedreich's ataxia: new insights. Emerg Top Life Sci. 2023 Dec 14;7(3):313-323. doi: 10.1042/ETLS20230017. PMID: 37698160.

[2] Hafiz S, De Jesus O. Ataxia. 2023 Aug 23. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–. PMID: 32965955.

[3] Williams CT, De Jesus O. Friedreich Ataxia. 2023 Aug 23. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–. PMID: 33085346.

ICB12-1402-1050

lactate dehydrogenase A inhibition via peptides to regulate the Warburg effect: cell signaling analysis

Farzaneh Jafary[1, 2], Ali Moradi[2*], Mohamad Reza Ganjalikhany[3*], Ameneh Khodarahmi[2]
1 Isfahan Endocrine and Metabolism Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
2 Department of Biochemistry, Faculty of Medicine, Shahid Sadoughi University of Medical Sciences, Yazd, Iran.
3 Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran.
*Corresponding authors' email addresses: ganjalikhany@sci.ui.ac.ir , moradi20018@gmail.com

Abstract: Lactate dehydrogenase A (LDHA) plays a critical role in the Warburg effect that is detected in numerous types of solid tumors. Hence, LDHA has been proposed as a potential target for suppression in the field of cancer treatments. Here, we used in-silico study to design inhibitory peptide for LDHA. After that, we checked the effect of peptide on cell signaling in Balb/c mice. The expression levels of three proteins, namely cdc20, cdk2, and Rb1, were analyzed in tumor tissue samples obtained from both the control and treatment groups. The findings indicate a decrease in the expression levels of cdc20 and cdk2 in the treatment groups. However, the expression of Rb1 was shown to be significantly higher in the treatment groups when compared to both the control and sham groups. The inactivation of the RB1 is a common occurrence in cancer, either through direct mutations or indirectly through the dysregulation of RB1 regulators [1]. On the other hand, activation of CDK2 is associated with the progression of malignancy and the invasive behavior of cancer cells [2]. Other study was shown that inhibiting mitotic exit by the downregulation of CDC20 proved to be a more effective approach in eliminating cancer cells [3]. Therefore, increases of Rb1 and decreases of cdc20 and cdk2 expression indicated the effect of peptide in cancer treatment. Finally, discovering of these peptides may be a new way for dealing with cancer, and we will decide to optimize these peptides in future work.

Keywords: Lactate dehydrogenase A; Warburg effect; peptide inhibitors; therapeutic targets; cancer

References:

[1] Vélez-Cruz R, Johnson DG, "The Retinoblastoma (RB) Tumor Suppressor: Pushing Back against Genome Instability on Multiple Fronts". Int J Mol Sci., 16 (2017) 1776.

[2]Liu FY, Wang L-P, Wang Q, et al. "miR-302b regulates cell cycles by targeting CDK2 via ERK signaling pathway in gastric cancer". Cancer Med., 5(2016) 2302-2313.

[3] Kidokoro T, Tanikawa C, Furukawa Y, Katagiri T, Nakamura Y, Matsuda K. "CDC20, a potential cancer therapeutic target, is negatively regulated by p53". Oncogene., 27(2008) 1562–1571.

ICB12-1402-1052

Using the recommended protein marker gpd to investigate the phylogeny relationships among 23 species of bioindicator fungal genus *Lactarius*

Mohammad Mahdavi[a*], Mohammadali Tajik Ghanbari[b]
a Department of plant protection, Sari agricultural sciences and natural resources university, Sari, Iran.
b Department of plant protection, Sari agricultural sciences and natural resources university, Sari, Iran.
*Corresponding author's email address: mahdavi8262@yahoo.com

Abstract: This research has been carried out in the form of examining the similarity of the gpd protein amino acid sequence among the isolates of 23 species of the bioindicator fungal genus *Lactarius*. In this study, 48 amino acid sequences of gpd protein related to 23 species of *Lactarius* fungi were extracted from UniProt database and clustered by version 11 of MEGA software. gpd is a highly conserved enzyme in the glycolytic pathway and catalyzes the conversion of glyceraldehyde-3-phosphate to glycerate-1,3-biphosphate. *Lactarius* is a genus of ectomycorrhizal fungi that has several edible species. Species of this genus, commonly known as milk caps, are characterized by the milky liquid (latex) they exude when cut or injured. These fungi act as environmental indicators and their presence in the environment is used to evaluate the quality of the environment. In this research, the mentioned sequences were compared among 48 isolates of this genus and showed similarities and differences, and the isolates of each species stood together and the species were clearly separated. In the end, it was observed that the gpd protein marker can distinguish the species of this genus well and is a suitable candidate for phylogenetic analyzes and taxonomic studies.

Keywords: gpd; Lactarius; Mycoindicator; phylogenetic analysis; phylogeny.

ICB12-1402-1053

Molecular Modeling of Monomeric Spike SARS-CoV-2 - Ferritin Nanoparticle Vaccine

Farnaz Garmeh Motlagh[1] and Maryam Azimzadeh Irani[1]*
[1]Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran.
*Corresponding author's email address: m_azimzadeh@sbu.ac.ir

Abstract: The emergence of COVID-19, caused by SARS-CoV-2, has triggered a worldwide health crisis. One of the critical elements in SARS-CoV-2 infectivity is its Spike protein, which binds to the ACE2 receptor, triggering a robust immune response [1]. Therefore, developing an efficient vaccine is crucial. Nanoparticle vaccines have demonstrated superior effectiveness [2,3].

In this in silico study, ferritin - monomeric Spike protein as immune system stimulator to produce the SARS-CoV-2 nanoparticle vaccine was modeled for the first time. The MD simulations of the complex were performed using GROMACS 2021.4 software. The stability of the monomeric Spike SARS-CoV-2-ferritin complex was examined through analysis of MD simulations using the Bio3D package in R [4].

The RMSD analysis indicated an average RMSD value of 12.1 Å. This relatively high RMSD value can be attributed to the substantial size of the complex. RMSF analysis revealed an average RMSF value of 3.7 Å. Generally, these results demonstrated noticeable stability within the modeled complex. Individual structural analysis of the RBD and ferritin indicated significant structural stability within these components. The secondary structure analysis further supported these findings, revealing a lack of profound changes in the Spike and ferritin secondary structures. Furthermore, the visualisation of dynamics revealed that the RBD maintained an upward active conformation that is crucial for its interaction with the ACE2 receptor, without interfering with other domains within the complex.

The study revealed that the designed nanoparticle vaccine is structurally stable and potentially able to provide immunity against the SARS-CoV-2 virus.

Keywords: Ferritin; SARS-CoV-2 Spike; Nanoparticle vaccine; Molecular dynamics simulations

References

[1] Li, F., Li, W., Farzan, M., and Harrison, S. C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309, 1864–1868. https://doi.org/10.1126/science.1116480

[2] Oyewumi, M. O., Kumar, A., & Cui, Z. (2010). Nano-microparticles as immune adjuvants: correlating particle sizes and the resultant immune responses. Expert Review of Vaccines, 9(9), 1095–1107. https://doi.org/10.1586/erv.10.89

[3] Wuertz, K.M., Barkei, E.K., Chen, W.H., Martinez, E.J., Lakhal-Naouar, I., Jagodzinski, L.L., Modjarrad, K. (2021). A SARS-CoV-2 spike ferritin nanoparticle vaccine protects hamsters against Alpha and Beta virus variant challenge. Npj Vaccines, 6(1), 129. https://doi.org/10.1038/s41541-021-00392-7

[4] Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006). Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics 22, 2695–2696. doi:10.1093/bioinformatics/btl461

ICB12-1402-1055

Semantic Segmentation Based on Artificial Intelligence: An Effective Paradigm for Automated Estimation of Fetal Head Circumference in Ultrasound Images

Seyed Vahab Shojaedini [a]*, Mohammad Momenian[b]
a Biomedical engineering Department, Iranian Research Organization for Science & Technology, Tehran, Iran
b Department of Computer Engineering, Faculty of Engineering, Islamic Azad University, E-Campus, Tehran, Iran
*Corresponding author's email address: shojadini@irost.ir

Abstract: For several decades the obstetricians have measured fetus head circumference (HC) manually from ultrasound images captured from womb in order to estimate its gestational age, size, and weight. Unfortunately, some factors may hamper the performance of this technique including its subjective nature, different diagnoses by different specialists and its time-consuming procedure. Computer Aided Diagnosis has been introduced as an effective paradigm to address the aforementioned problems, but the low contrast of the ultrasonic images, its low signal-to-noise ratio, acoustic shadows, and speckle noise are the limiting parameters of these methods. Thus, in recent years, the use of artificial intelligence has become an inevitable choice to solve this problem due to its ability to construct comprehensive models for the fetus and the background.

In this article, a new scheme is presented in artificial intelligence framework in order to estimate fetal HC, which is based on semantic segmentation by using deep learning. In the proposed scheme, the region related to the fetus is separated from the background by using a UNET deep neural network which has been trained by utilizing ultrasound images captured from the womb, in which the fetal head area has been labeled by a specialist. The UNET may promote the performance of HC estimation by applying the sliding window technique which needs fewer images to increase the model performance. Furthermore, such an architecture may promote localization tasks as well as creating distinguished class label for each pixel thanks to create a local patch for each pixel.

In order to evaluate the effectiveness of the semantic segmentation scheme it has been implemented as a software package which may test on real HC18 dataset contains a total of 999 two-dimensional labeled ultrasound images of the standard plane that can be used to measure the HC. The testbed was prepared on python3 and TensorFlow paradigms on Intel ® Core i7-10700 computer with Ubuntu 20.04 operation system, 32 GB RAM, and an NVIDIA 2080 Ti. We used Google Colab, a GPU framework made available by Google, to run the program.

The results obtained from the above evaluations demonstrated the effectiveness semantic segmentation paradigm in estimation fetal head parameters in such way that the Absolute Difference measure (i.e., ADF) between the ground truth around the fetal head and the estimation of the proposed method has been in range of [2.35 - 2.75] millimeters. In the same way, Dice's parameter for the proposed method has been obtained in range of [96.75 – 97.7] percent and Jaccard's parameter has been in range of [94-95] percent. These parameters indicate the acceptable similarity of the head circumference boundary extracted from the proposed method to the actual fetal head circumference boundary. The low difference between actual and estimated HC of fetus (e.g., the ADF) as well as the high similarity between real and estimated borders (e.g., Dice and Jaccard measures) showed that the use of deep learning based semantic segmentation can be developed and utilized as an option with the potential of fetal head circumference estimation in practical applications.

Keywords: Fetal Head Circumference; Artificial Intelligence; Semantic Segmentation; Deep Learning; UNET.

ICB12-1402-1056

Transfer Learning Based on DenseNet-121 Model:

A Deep Learning Approach to Promote Detecting Breast Cancer in Thermogram

Mohamad Firouzmand [a], Seyed Vahab Shojaedini [a]*, Mehdi Abedini[b], Mahsa Monajemi[c]
a Biomedical engineering Department, Iranian Research Organization for Science & Technology, Tehran, Iran
b Department of Computer Engineering, Faculty of Engineering, Islamic Azad University, E-Campus, Tehran, Iran
c Department of biomedical Engineering, Faculty of Engineering, Islamic Azad University, Qazvin branch, Qazvin, Iran
*Corresponding author's email address: shojadini@irost.ir

Abstract: In the last decade, the use of thermography has become an attractive research topic in the field of medical imaging for diagnosing breast cancer, thanks to its harmlessness, cheapness, and early detection potential. However, the variety of the thermal patterns related to the cancerous lesion and their separation from the normal body temperature profile has made the interpretation of thermograms more complicated than the images obtained from other diagnostic modalities. This issue has caused much attention to be paid to the use of artificial intelligence-based methods, in particular, deep neural networks as a tool to detect the presence of cancerous lesions in breast thermograms. Since deep neural networks are considered as a member of the family of big data methods, their optimal performance in this field is highly dependent on the availability of a large collection of breast thermograms, which such a vast database is not yet available due to the young age of breast thermography. In this article, the use of transfer learning concept is proposed and examined as a solution to improve the performance of deep neural networks in breast cancer diagnosis by thermograms.
In our approach, the DenseNet-121 model which has been pre-trained by ImageNet database, containing than 14 million images and 1000 classes, is reused as the starting point for a model for distinguishing healthy and cancerous breast tissues via a fine-tuning scheme. This allows us to address the challenge of the large amount of computing and storage resources required to develop an effective breast cancer deep learning based detector. The testbed for evaluation of the proposed scheme was provided by utilizing Tensorflow which is an open-source set of Python machine learning module, with a NVIDIA 2080 TI GPU with 32 GB RAM. Furthermore, Google Colab, a GPU framework made available by Google were used in order to run the computer code. The tests were performed on DMR-IR dataset which includes 760 thermal images for the sick and 762 thermal images for the healthy class and have been divided into three categories of training, validation and test subsets, with ratios of 60%, 15% and 25% among the total images respectively. The results of multiple experiments demonstrate the acceptable performance of the proposed technique in terms of obtaining 84% accuracy, 75% sensitivity, and 93% specificity in distinguishing healthy and cancerous breasts. According to demonstrated results,
the use of transfer learning paradigm based on DenseNet-121 may be considered as an option with acceptable potential in constructing deep neural models in order to interpret breast cancer thermograms.

Keywords: Breast cancer; Thermography; Deep learning; Transfer learning; DenseNet-121.

ICB12-1402-1057

Semantic Segmentation Based on Artificial Intelligence: An Effective Paradigm for Automated Estimation of Fetal Head Circumference in Ultrasound Images

Mohammad Jahangiri[1*], Leila Zarandi-Miandoab[1]
1 Department of Biology, Faculty of Sciences, Azarbaijan Shahid Madani University, Tabriz, Iran
* Corresponding author's email address: m.jahangiri7739@gmail.com

Abstract: Marine microalgae have been used as a source of nutritional and medicinal supplements for many years due to the presence of various compounds and nutrients. Bioactive compounds with algal origin have many capabilities such as antihypertensive, immune modulator, antioxidant, anticancer and antimicrobial. Proteins make up a high percentage of compounds in microalgae. Bioactive peptides are compounds that are formed after enzymatic hydrolysis of proteins and have different properties and functions. This study investigated in-silico the properties of bioactive peptides in the phycoerythrin protein of Porphyridium purpureum microalgae. Phycoerythrin protein, which is one of the effective proteins in the photosynthesis of microalgae, has two subunits, alpha and beta. The bioinformatics investigation of potential peptides in these subunits in 4 selected fields with activities related to human health, determined that the frequency of occurrence of these peptides in the alpha subunit is higher than beta. Most of these peptides have DPP-4 inhibitory properties and then ACE inhibitory properties, which are effective in improving type2 diabetes and high blood pressure, respectively. Also, by simulating the effect of digestive enzymes (pepsin, trypsin, and chymotrypsin) on phycoerythrin protein subunits, 15(non-toxic) peptide fragments with different biological activities were created, and the abundance, activity potential, and digestive absorption of these peptides were also investigated. The results of this study can be a basis for further experimental and laboratory investigations to make medicine and food supplements from this protein and its peptides.

Keywords: Porphyridium purpureum microalgae; phycoerythrin; Bioactive peptides; Bioinformatics

ICB12-1402-1059

Bioinformatics approach for shedding light on controversial information of IGF1/2 expression pattern in hepatocellular carcinoma

A.H.Farrokhzad[a] , M.Vosough[a*]
a Department of Regenerative Medicine, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, Academic Center for Education, Culture and Research (ACECR), Tehran, Iran
*Corresponding author's email address: masvos@Royaninstitute.org

Abstract: Hepatocellular carcinoma (HCC) is one of the most common and lethal malignant tumors worldwide [1]. The present study aimed to identify the expression patterns of IGF1/2 and pathways associated with HCC by using bioinformatics methods. The raw microarray data of GSE84402, including 14 pairs of HCC tissues and the corresponding non-cancerous tissues, were obtained from the Gene Expression Omnibus (GEO) database. These data were contributed by Wang et al. [2]. After defining two groups and analyzing with GEO2R, a group of genes was achieved; in order to find the hub genes between them, a group of genes was screened out according to adjusted P-value <0.05 and — logFC—≥2; second, we constructed the protein-protein interaction (PPI) network from the STRING database. Finally, among some genes, IGF1/2 was chosen according to its good betweenness centrality. To define the expression patterns of IGF1/2 in HCC tissue, we used the UALCAN and GEPIA databases and used the KEGG database to construct a network of HCC; The signaling pathway of IGF1/2 was obtained from REACTOME database. In conclusion, based on bioinformatics analyses, we observed down-regulation of IGF1/2 in cancer samples compared to control, although the in vivo and in vitro results were not similar. We found that down-regulation of IGF1/2, increases cancerous phenotype through increasing the activity of Wnt/β- catenin, mTOR and ERBB2 signaling pathway and resulting in enhanced survival, proliferation and differentiation of cells; So targeting these molecular pathways can be a promising therapeutic approach in future.

Keywords: Insulin-like growth factor system; Hepatocellular carcinoma; System biology; Wnt/β-catenin; ERBB2

References

[1] A.Alqahtani ,Z.Khan "Hepatocellular Carcinoma: Molecular Mechanisms and Targeted Therapies" Medicina 2019

[2] H.Wang,X.Hou "STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4". Mol. Cancer2017

ICB12-1402-1060

From Melanoma to Skin Aging: The Impact of DCT and TYRP1 in Melanogenesis

M. Kaboli[a], A. Amiri-yekta[b*]
a.Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, Tehran, Iran
b. Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, Tehran, Iran
*Corresponding author's email address: amir.amiriyekta@gmail.com

Abstract: Melanoma is responsible for over 90% of skin cancer deaths. Skin color and cumulative UV exposure are both known risk factors for skin cancer as well as skin aging, which is not only a physiologic phenomenon but also a health risk, resulting in increased incidence of infection and skin cancers [1]. This study aims to pinpoint essential genes that contribute significantly to Melanoma development and skin aging. The raw microarray data of GSE35388, including eight samples of Expression data from normal melanocytes and melanoma cells, were obtained from the Gene Expression Omnibus (GEO) database. These data were based on the GPL570 platform and were contributed by Xiao D et al. [2]. First, two groups were defined for analysis using GEO2R, and a list of genes was obtained; after that, we filtered out a group of genes based on their adjusted P-value < 0.05 and — logFC >-2; second, we constructed the protein-protein interaction network from the STRING database to identify hub genes. Finally, two genes were chosen according to their excellent betweenness centrality and Tau specificity score (from the HUMAN PROTEIN ATLAS database): DCT and TYRP1. which, Based on the REACTOME database, are in the same signaling pathway leading to melanin biosynthesis and have been linked to various skin aging phenotypes. In conclusion, skin aging and skin cancer might be associated; they could coexist in one genetically predisposed individual but might even influence each other [3]. Therefore, targeting these genes can be a potential therapeutic strategy for skin phenotypes.

Keywords: Melanogenesis; Skin Pigmentation; Melanoma; Skin Cancer; Skin Aging

References

[1] H. Lee, Y. Hong, "Structural and functional changes and possible molecular mechanisms in aged skin " Int. J. Mol. Sci. 2021

[2] D. Xiao, J. Ohlendorf, "Identifying mRNA, microRNA and protein profiles of melanoma exosomes " 2012.

[3] L. C. Jacobs, "Genetic Determinants of Skin Color, Aging, and Cancer " 2015.

ICB12-1402-1061

Expression data from human breast tumors and their paired normal tissues protein-protein interaction

Hanieh Abdollahzadeh[1*],Yaghub Pazhang[1]
1.Department of Biology, Faculty of Sciences, Urmia University, Urmia, Iran.
*Corresponding author's email address: st_h.abdollahzadeh@urmia.ac.ir

Abstract: Breast cancer is globally the second most prevalent form of cancer which is affecting approximately 2.09 million women.[1] Numerous previous studies have revealed the peroxisome proliferator-activated receptor (PPAR) signaling pathway and the risk of cancer. However, these studies have yielded inconsistent results due to unknown role of the PPAR signaling pathway in cancer promotion. Hence, this study aims to investigate the PPAR signaling pathway in breast cancer. The RNA-Seq expression of a from a breast cancer was obtained from the Gene Expression Omnibus (GEO) dataset. In the training dataset (GSE15852), we chose and analyzed it by Cytoscape and Gephi software (fig1). These analyses were done using twenty-one genes that exhibited differential expression of the PPAR signaling pathway by enrichment analysis and some have a role in cell cycle process such as TGFB2. Therefore, the PRAR signaling pathway may be related to breast cancer progression.[2] The result was, collectively most of the genes are related to the PPAR signaling pathway (Table 1). Some genes are in the cell cycle such as TGFB2.

Keywords: System biology; protein-protein interaction; breast cancer; tumor; gene expression

References

[1] Y. E. Hadisaputri *et al.*, "Antiproliferation activity and apoptotic mechanism of soursop (Annona muricata l.) leaves extract and fractions on mcf7 breast cancer cells," *Breast Cancer Targets Ther.*, vol. 13, pp. 447–457, 2021, doi:10.2147/BCTT.S317682.
[2] H. Wang *et al.*, "The PPAR signaling pathway as a potential biomarker for the diagnosis of breast cancer," *Int. J. Clin. Exp.Med.*, vol. 12, no. 6, pp. 7327–7336, 2019, [Online]. Available: https://www.embase.com/search/results?subaction=viewrecord&id=L2002255314&from=export

ICB12-1402-1067

An in-silico study on some synthesized triazole derivatives as tyrosinase inhibitors

Marzieh Kazazi[a], Atiyeh Mahdavi[a*]
a Department of Biological Sciences, Institute for Advanced Studies in Basic Sciences (IASBS), 444 Prof. Sobouti Blvd., Gava Zang, Zanjan 45137-66731, Iran
*Corresponding author's email address: a.mahdavi@iasbs.ac.ir

Abstract: Tyrosinase is a crucial enzyme in melanogenesis pathway. However, overexpression or overactivity of this enzyme result in some disorders such as melanoma, age spots, melasma, pregnant spots, etc. In these days, a lot of research is being done to find safe and potent tyrosinase inhibitors [1]. In order to predict the type and position of interactions between ligands and enzyme, we examined the molecular docking studies of three novel synthesized triazole derivatives (L1, L2 and L3) on tyrosinase activity [2]. The protein data bank (www.rcsb.org) provided the mushroom tyrosinase pdb file (PDB ID: 2Y9X) and the ligands were designed and optimized using Gaussian 09W software. Then, introduce tyrosinase and ligands close to each other by means of Visual Molecular Dynamics (VMD) software. Then, direct and rigid dockings were run by AutoDock 4.2.6 software [3]. AutoDock 4.2.6, Discovery Studio, LigPlot and PyMOL software were used to analyze the docking results. The outcomes demonstrated that there were a variety of interactions between the ligands and the enzyme, including hydrophobic, electrostatic and hydrogen bond interactions. Additionally, the compounds interact with the important enzyme residues. The binding energies of all the compounds are roughly the same, with L3 having the lowest binding energy. The overall results indicate, the L3 anticipation as more potent inhibitor than the other ones and might be more considered for further studies in food, pharmaceutical and agricultural industries.

Keywords: Mushroom tyrosinase; Inhibitors; Triazole compounds; Molecular docking

References

[1] Mohammadsadeghi, N., Mahdavi, A., Saadati, F., & Mohammadi, F. (2023). In silico and in vitro studies of novel derivatives of tyrosol and raspberry ketone as the mushroom tyrosinase inhibitors. Food Chemistry, 424, 136413.

[2] Kashyap, A., & Silakari, O. (2018). Triazoles: Multidimensional 5-membered nucleus for designing multitargeting agents. In Key Heterocycle Cores for Designing Multitargeting Molecules (pp. 323-342). Elsevier.

[3] Rezaei, M., Mohammadi, H. T., Mahdavi, A., Shourian, M., & Ghafouri, H. (2018). Evaluation of thiazolidinone derivatives as a new class of mushroom tyrosinase inhibitors. International journal of biological macromolecules, 108, 205-213.

ICB12-1402-1068

Identification of biomarkers for breast cancer in different reproductive ages using computational methods

Sana Zarghami[1]
1 University of science and culture, Royan institute , Tehran, Iran
*Corresponding author's email address: sanazarghami4444@gmail.com

Abstract: Identification of biomarkers for breast cancer in different reproductive ages using computational methods Breast cancer is the most common non-skin cancer in women worldwide [1], characterized by molecular and morphological heterogeneity [2].The study aimed to identify potential hub genes in the development network of breast cancer in premenopausal (reproductive) and postmenopausal age groups. Using data from the Gene Expression Omnibus (GEO) Dataset study GSE 102484, comprising 683 samples, patients with breast cancer were divided into the two age groups, and analysis was conducted using GEO2R. Genes exhibiting a log fold change greater than " -1.2 " with an adjusted p-value below " 0.05" were plotted in a network, revealing the progesterone receptor (PGR) as the hub gene with a high degree in the Cytoscape network. Elevated expression of PGR was found in the premenopausal age group, indicating a direct correlation between its high expression and increased female sex hormones (esterogen and progesterone) in this age group. The findings suggest potential differential diagnostic and therapeutic approaches between various age groups, with the examination of PGR as a biomarker in diagnosing and treating breast cancer in the premenopausal age group proving effective. The study highlights the importance of considering age-specific factors in breast cancer research and treatment.

Keywords: Breast cancer; age groups; hub genes; PGR(progesterone receptor); biomarker; differential diagnosis; therapeutic approach**.**

References:

[1] Wang X, He X, Guo H, Tong Y. Variants in the 8q24 region associated with risk of breast cancer: Systematic research synopsis and meta-analysis. Medicine (Baltimore). 2020;99(8):e19217.
[2] Budh DP, Sapra A. Breast Cancer Screening. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2023, StatPearls Publishing LLC.; 2023.

ICB12-1402-1069

c.5244 del G: A Novel Mutation in the ABCA4 Gene is Associated with Retinitis Pigmentosa

Elham Norouz Dolatabadi[a,b], Asiyeh Jebelli[c], Shahin Rashidi[d], Ebrahim Sakhinia[e], Effat Alizadeh[b*]
a Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
b Department of Medical Biotechnology, Tabriz University of Medical Sciences, Tabriz, Iran
c Department of Biological Sciences, Faculty of Basic Sciences, Higher Education Institute of Rab-Rashid, Tabriz, Iran
d Department of Molecular Genetics, University of Tabriz, Tabriz, Iran
e Department of Laboratory and Regenerative Medicine, University of Manchester, United Kingdom
*Corresponding author's email address: e.alizadeh.2010@gmail.com

Abstract:

Background: Retinitis pigmentosa (RP) is an inherited retinal disease that affects approximately 1 in 3,500 individuals worldwide. It is characterized by progressive degeneration of photoreceptor cells, leading to visual impairment and eventual blindness. RP exhibits significant genetic heterogeneity, with mutations identified in over 70 genes associated with the disease. The ABCA4 gene has been extensively studied due to its involvement in various retinal dystrophies, including Stargardt disease and cone-rod dystrophy.[1]

Methods: In this study, we obtained blood samples from a RP patient with characteristic symptoms such as night blindness and peripheral vision loss. Genomic DNA was extracted using standard protocols, followed by whole exome sequencing (WES) technique. Data analysis of WES was performed based on a specific gene panel including known RP-associated genes. The resulting sequence data was analyzed using bioinformatics tools to identify potential disease-causing variants, which was confirmed by Sanger sequencing. The secondary structure of normal and mutant mRNA was analyzed by the RNA fold database.

Results: WES analysis revealed a homozygous variant reported as a likely pathogen in Franklin and VarSome; However, there is no evidence related to it in Clinvar. This novel mutation, c.5244 del G in exon 37 of the ABCA4 gene, is a frameshift variation leads to a premature stop codon downstream, resulting in a short mRNA and a truncated protein product. The secondary structure of mRNA was changed after the mentioned mutation, which could affect its downstream functions .

Discussion and conclusion: The ABCA4 gene encodes an ATP-binding cassette transporter involved in transporting retinoids across photoreceptor outer segments' disc membranes [2]. Mutations affecting this gene have been associated with impaired clearance of toxic retinoid derivatives, leading to photoreceptor cell death [3]. In conclusion, the identification of this novel mutation (c.5244 del G) and comparison of the secondary structure of the normal and mutant mRNA expands our understanding of the genetic landscape underlying RP and highlights the importance of comprehensive genetic testing for accurate diagnosis and prognosis prediction.

Keywords: WES; ABCA4 gene; Retinitis pigmentosa; mRNA secondary structure

References

[1] Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. The Lancet. 2006;368(9549):1795-809 .

[2] Molday RS. ATP-binding cassette transporter ABCA4: molecular properties and role in vision and macular degeneration. Journal of bioenergetics and biomembranes. 2007;39:507-17.

ICB12-1402-1070

Molecular profiling of downregulated genes in T-DM1 resistance: insights from OE-19 esophageal cancer cells

Fateme Yazdani[1], Negar Mottaghi-Dastjerdi[1*] , Mohammad Soltany-Rezaee-Rad[2] , Hamed Montazeri[1]
1 Department of Pharmacognosy and Pharmaceutical Biotechnology, School of Pharmacy, Iran University of Medical Sciences, Tehran, Iran
2 Behestan Innovation Factory, Tehran, Iran
*Corresponding author's email address: *Mottaghi.n@iums.ac.ir

Abstract:

Introduction: Esophageal cancer is a challenging disease because it is aggressive and has limited treatment options. T-DM1, a medication targeting HER2-positive cancer cells, has shown promise in improving patient outcomes. However, drug resistance is still a significant obstacle [1, 2]. To develop effective therapeutic strategies, it is crucial to understand the molecular mechanisms that cause T-DM1 resistance.

Methods: In this study, we aimed to identify the essential regulatory genes and pathways involved in the progression of TDM1 resistance in OE-19 EC cells. We extracted expression datasets from GEO omnibus, analyzed gene interactions, analyzed the reconstructed protein-protein interaction network, and performed enrichment analysis of the hub genes

Results: We identified six hub genes (ALDH1A1, SLPI, CEBPA, RAC2, PIWIL1, and PRODH) as the key downregulated genes (Figure 1) that are mostly involved in the synthesis of GDP-mannose, Fructose Catabolism, Fructose Metabolism, and ROS And RNS Production In Phagocytes. The heterocycle catabolic process, gamma-aminobutyric acid metabolic process, and intrinsic apoptotic signaling pathway in response to oxidative stress were shown to be significantly enriched in GO analysis. The most prominent cellular components were the chromatoid body, NADPH oxidase complex, and mitochondrial membrane. Retinal dehydrogenase, aldehyde dehydrogenase (NAD+), and oxidoreductase activity were significantly enriched in molecular functions.

Conclusion: This study thoroughly examines the molecular landscape behind TDM-1 resistance in esophageal cancer cell lines. The discovered hub genes and pathways offer prospective treatment targets and shed light on the complicated resistance mechanisms at work. Further experimental validation of these findings is required for future clinical application.

Keywords: Biological networks; Differentially expressed genes; Esophageal cancer; Protein-protein interactions; Therapeutic target

References:

[1] He, S., et al., Advances and challenges in the treatment of esophageal cancer. Acta Pharm Sin B, 2021. 11(11): p. 3379-3392.

[2] Yang, J., et al., Understanding Esophageal Cancer: The Challenges and Opportunities for the Next Decade. Front Oncol, 2020. 10: p. 1727

ICB12-1402-1072

Re-purposing Clove Oil in Ovarian Cancer Treatment: A Multi-Dimensional Approach Involving Bulk RNA-Seq and scRNA-SeqDescription

Bahar Bahrami [1] Masoud Tohidfar*,[2]
1 Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
2 Dept. of Cell and Molecular Biology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
*Corresponding author's email address: m_tohidfar@sbu.ac.ir

Abstract: Ovarian cancer (OC) is currently the fifth leading cause of cancer-related deaths among women, with approximately 140,000 fatalities globally per year. To improve the prognosis of OC patients, novel therapeutic approaches are essential. Cancer Stem Cells (CSCs) are integral to ovarian cancer's entire development process, including initiation, metastatic progression, therapeutic resistance, and disease recurrence. Thus, targeted therapy, particularly against Ovarian Cancer Stem Cells (OCSCs), is expected to be more effective and less toxic, potentially improving patient survival and reducing tumor relapse.

Our initial analysis utilized GSE13237 dataset, which includes data from 11 pairs of primary and metastatic OV tumors. We employed an RNA-seq pipeline along with the DESeq2 package, setting a threshold of p.value <0.05 and |logFC| > 1 to identify highly variable genes. This yielded 2134 significant genes. Subsequently, using the clusterProfile package and David online enrichment tool for KEGG, we pinpointed clove oil—a mild analgesic used for toothache and suggested for Diabetic Cardiomyopathy therapy—as a potential agent targeting the COL1A1 gene. This gene plays a significant role in metastasis and stem cells regulatory pathways.

To ensure accuracy, we further analyzed pooled scRNA-seq datasets GSE184880 and GSE158937, comprising 8 samples (5 healthy and 3 high-grade serous metastatic OC). The results significantly enriched the COL1A1 gene in the cancer cell cluster, providing clear evidence of its association with OCSCs.

By understanding the genetic underpinnings and potential therapeutic targets like the COL1A1 gene, we can pave the way for more effective and personalized treatments for ovarian cancer, offering hope for improved outcomes.

Keywords: Ovarian cancer,cancer stem cells,Clove oil,COL1A1(Collagen type I alpha 1)

ICB12-1402-1077

Prediction of pathogenicity of rs1800751 (Pro1812Thr) nucleotide change in BRCA1 gene related to breast cancer by bioinformatics tools

Melika Kamboozi[1], Mehri Khatami[1], Fatemeh Farrokhzadeh[1]
1 Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: Melikakamboozia@gmail.com

Abstract: Breast cancer is one of the most common types of cancer in women, and the BRCA1 gene is associated with an increased risk of developing this disease. Understanding the function of the BRCA1 gene and its role in the development of breast cancer has important implications for personalized medicine and targeted therapies for individuals at high risk. Ongoing research aims to uncover how BRCA1 gene mutations contribute to cancer development and explore potential treatment options. In the present study, a missense mutation in this gene causes the proline amino acid at position 1812 to change to Threonine. Its pathogenicity assessment has been analyzed using several bioinformatics tools such as polyphen-2, I-Mutant 2.0, and SIFT.

PolyPhen-2 is an online tool designed to predict the impact of amino acid substitutions on protein structure and function. It does so by analyzing multiple sequence alignments of 3D protein structures. The Expasy website can predict the effects of missense mutations on protein structure, function, and stability. The SIFT tool is also used to predict whether an amino acid substitution affects protein function, based on the protein sequence.

Polyphen-2 analysis indicated that this mutation is predicted to be probably damaging with a score of 1.000. Substitution at position 1812 from P to T is predicted to affect protein function with a score of 0.00 by the SIFT. According to the results from Expasy, this mutation caused a change in the hydrophobicity index from -1.600 (proline) to -0.700 (threonine).

According to this consideration, rs1800751 might have pathogenic effects on the BRCA1 gene. This theory should be proved with experimental studies because the score of pathogenicity of this SNP is very high.

Keywords: BRCA1 gene; Breast cancer; rs1800751; Missense mutation

ICB12-1402-1078

Solubility, mathematical modeling and thermodynamic properties of budesonide solubility in 1-Propanol+Water from T = (293.2 to 313.2) K

Mahmoud Omidi[a,b], Seyyed Mohammad Hasan Hashemi[c], Fereshteh Talebi[d], Esmaeil Mohammadian[a,d*]
a Molecular Medicine Research Center, Hormozgan Health Institute, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
b Department of Pharmacology and Toxicology, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
c Department of Pharmaceutics, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
d Department of Medicinal Chemistry, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
*Corresponding author's email address: e.mohamadyan1390@gmail.com

Abstract: Budesonide (BDS) is a corticosteroid that is used for treating asthma, rhinitis, chronic obstructive pulmonary disease (COPD), autoimmune hepatitis, and inflammatory bowel disease [1]. BDS with aqueous solubility of 16 µg/mL belongs to class II of biopharmaceutical classification system. By increasing the aqueous solubility of these drugs, enhanced bioavailability, dose reduction, and enhanced efficiency can be achieved. Cosolvency as a simplest technique for the improvement aqueous solubility of poorly soluble drugs are widely used in pharmaceutical industry. A binary mixture of 1-propanol and water as cosolvent system was used to determine the solubility of BDS. Solubility of BDS in mass fractions of 1-propanol/water was measured using shake-flask method. After 48 h (Based on the dissolution rate, the amount of absorption reached its maximum value after 48 hours), the absorbance of the diluted solutions was obtained with using a UV–vis spectrophotometer and the concentrations were calculated using a previously created calibration curve (R2= 0.999). In the next step, the solubility of BDS in binary mixtures of 1-Propanol+Water at 293.2, 298.2, 303.2, 308.2 and 313.2 K was investigated. Also, BDS solubility in binary mixture compositions at various temperatures was calculated using the trained versions of the van't Hoff equation, Jouyban-Acree model and a combination of the van't Hoff equation and Jouyban-Acree model [2]. The accuracy of these models was evaluated using the mean relative deviations (MRDs) as a criterion. The MRDs for the fitting of the solubility data of BDS in binary mixtures of 1-propanol + water with Jouyban-Acree and Jouyban-Acree-van't Hoff models were found to be 12.30% and 9.23%, respectively. Finally, according to van't Hoff's and Gibb's equations, some apparent thermodynamic quantities were calculated in all the mixed solvents, such as the dissolution enthalpy, the dissolution entropy, and the Gibbs free energy change
Keywords: Solubility; Budesonid; 1-Propanol; Jouyban-Acree model; Mathematical model

References

[1]Johansson S-Å, Andersson K-E, Brattsand R, Hedner P "Topical and systemic glucocorticoid potencies of budesonide and beclomethasone dipropionate in man" European journal of clinical pharmacology", 1982.

[2] Jouyban A, "Review of the cosolvency models for predicting solubility of drugs in water-cosolvent mixtures", Journal of Pharmacy & Pharmaceutical Sciences, 2008.

ICB12-1402-1079

In silico discovery of small-molecule K-Ras inhibitor

M. Shirzad [a], S. Kia[a]*, S. Shabani Mashcool [b]
a Department of Engineering Sciences, Faculty of Advanced Technologies, University of Mohaghegh Ardabili, Namin, Iran.
b Laboratory of Bioinformatics and Drug Design, Institute of Biochemistry and Biophysics, University of Tehran.Tehran, Iran.
*Corresponding author's email address: s.kia@uma.ac.ir

Abstract: Mutational activation of the Ras oncogene products (H-Ras, K-Ras, and N-Ras) is frequently observed in human cancers, making them promising anticancer drug targets [1]. Mutation of the proto-oncogene K-Ras is one of the most common molecular mechanisms in non-small cell lung cancer. Many drugs for treating lung cancer have been developed, however, due to clinical observed K-Ras mutations, corresponding chemotherapy and targeted therapy for such mutation are not efficient enough. In this study, on the basis of the crystal structure of 4DSU, GDP-bound K-RasG12D mutant [2], 25 structural analogues of vismodegib, a hedgehog pathway inhibitor [3] were rationally designed. The designing of these compounds was based on the structure of 4DSU protein, and the related groups were replaced by bioisosteres to improve the affinity and selectivity.

Performing molecular docking:

Using PyRx, Autodock, Discovery Studio, UCSF Chimera and PyMol, make the desired changes (such as adding charge, converting protein and ligands to Pdbqt, adding hydrogen, removing extra molecules, etc.) on the protein and ligands. The desired docking on the protein and ligands was done, and analyze the obtained results using the Ligplot, Protein Ligand Interaction profiler, and PDBSum programs. Docking analysis showed that the VIS17 can effectively bind (ΔG_{bind}= -9.4 Kcal/mol) to the key amino acids of the enzyme active site and form a hydrogen bond with the LYS 89A.The structure and Functional groups of VIS17 was shown below.

In conclusion, our studies in finding novel potent compound (VIS17) with confirmed mechanism showed great potential for further optimisation and other medicinal chemistry relevant studies.

Keywords: K-Ras; 4DSU; vismodegib; docking; anticancer drug

References

[1] AT Shaw, MM Winslow, M Magendantz. "Selective killing of K-Ras mutant cancer cells by small molecule inducers of oxidative stress". Proc. Natl Acad. Sci. USA 108(21), (2011) 8773–8778.

[2] T. Maurer, LS. Garrenton, Oh A et al. "Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity". Proc. Natl Acad. Sci. USA 109(14), (2012) 5299–5304.

[3] S.Y. Quah, M.S. Tan, K.L. Ho, N.A. Manan et al. "In silico and saturation transfer difference NMR approaches to unravel the binding mode of an andrographolide derivative to K-Ras oncoprotein". Future medicinal chemistry, 12(18), (2020) 1611-1631.

A whole genome sequencing approach for unraveling the genetic basis of antibiotic resistance in Iranian Helicobacter pylori isolates

MohammadReza Najafi disfani, Parastoo Saniee*
Department of Microbiology and Microbial Biotechnology, Faculty of Life Science and technology, Shahid Beheshti University, Tehran, Iran
*Corresponding author email: p_saniee@sbu.ac.ir

Introduction: Antimicrobial resistance (AMR) among Helicobacter pylori strains is the major cause of eradication failure [1]. Based on studies conducted in Iran, the level of resistance has reached an alarming level and most Iranian H. pylori isolates reported as multi-drug resistant [2]. However the underlying cause of this high resistance is not clear. In this study, whole genome sequencing was used to analyze 31 randomly selected Iranian H. pylori strains to trace genomebased resistance.

31 H. pylori strains were isolated from Iranian patients with gastric disorders and subjected to DNA extraction. Whole genome sequencing was performed by the Illumina platform. After assembly and mapping of raw data, all isolates were analyzed by ResFinder, Comprehensive Antimicrobial Resistance Database (CARD) and ABRicate tools which are widely used for analyzing whole genome sequencing data to identify genes that confer resistance to various antibiotics [2].

ResFinder found some unidentified point mutations along with AMR to tetracycline, doxycycline and minocycline in only 1/31 isolates. ABRicate analysis showed AMR to tetracycline in only 2/31 isolates, one of them was the same as founded isolate by ResFinder. Although CARD showed some strict-hit AMR with a low percentage identity of matching region, no perfect-hit AMR was found in examined isolates.

Despite a high prevalence of antibiotic resistance in Iran which detected by laboratory susceptibility testing, bioinformatics tools revealed limited genetic resistance in only 2/30 of recruited isolates. This discrepancy between phenotypic and whole genome-based results in Iranian H. pylori strains indicates the need for further investigations.

Keywords: Helicobacter pylori; WGS; Antibiotic resistance

References:

[1] F. Ghotaslou R, Leylabadlo HE, Asl YM. Prevalence of antibiotic resistance in Helicobacter pylori: a recent literature review. World J Methodol. 2015; 5(3): 164-174.

[2] Jafari-Sales, Abolfazl a ; Jafari, Behboud b et al. Helicobacter pylori: a systematic review of drug resistance in Iran. Reviews and Research in Medical Microbiology 35(1):p 36-41, January 2024.

[3] Fernández-Caso, B.; Miqueleiz, A.; Alarcón, T. Whole Genome Sequencing for Studying Helicobacter pylori Antimicrobial Resistance. Antibiotics 2023, 12, 1135

ICB12-1402-1081

Stabilization of Aspergillus flavus urate oxidase enzyme by site-directed mutagenesis

Maryam Mousarezaei, Shahriar Shakeri*, Masoud Torkzadeh-Mahani, Mojtaba Mortazavi
Department of Biotechnology, Institute of Science High Technology & Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran
*Corresponding author e-mail : Shahriar.shakeri@yahoo.com

Abstract: Uricase enzyme, also known as urate oxidase, is a key and important enzyme in the catabolism of purines and it produces allantoin by catalyzing the reaction process of uric acid decomposition. In medicine, uricase is utilized as a diagnostic tool to measure uric acid levels in blood and other biological fluids. Elevated uric acid levels in the blood can lead to conditions such as gout, diabetes, cardiovascular disease, renal nephropathy, and tumor lysis syndrome. Hence, urate oxidase can serve as a medicinal enzyme to reduce uric acid levels.However, a major challenge with pharmaceutical enzymes like urate oxidase is their low thermal stability under physiological conditions, resulting in reduced enzyme activity above 25 degrees Celsius. Site-directed mutagenesis allows precise modification of gene and protein function, potentially leading to the creation of a protein with improved properties. In a particular study, the enzyme's structure was analyzed using bioinformatics tools, and amino acid 164 was selected for mutation due to its location outside the active site and its potential for mutation and kinetic and thermodynamic optimization. Using the site-directed mutagenesis method, it was converted to arginine. The introduction of arginine is expected to create new opportunities for hydrogen and ionic bonds as well as electrostatic bonds in the protein, making it a suitable substitute for the desired amino acid. This modification is predicted to enhance the stability of the tetramer structure, which is the active form of the enzyme, due to the creation of new electrostatic, ionic, and hydrogen bonds.

Keywords: urate oxidase;site-directed mutagenesis; Stabilization; uric acid; Structural bioinformatics

References:

[1] N. Motamedi, M. Barani, A. Lohrasbi-Nejad, M. Mortazavi, A. Riahi-Medvar, R. S. Varma, et al, “Enhancement of thermostability of Aspergillus flavus urate oxidase by immobilization on the Ni-Based Magnetic Metal-Organic framework,” Nanomaterials (Basel) 2021 Vol. 11 Issue 7.
[2] Z. Taherimehr, M. Zaboli and M. Torkzadeh-Mahani, “New insight into the molecular mechanism of the trehalose effect on urate oxidase stability,” Journal of Biomolecular Structure and Dynamics 2022 Vol. 40 Issue 4 Pages 1461-1471.
[3] S. Mirzaeinia, M. Pazhang, M. Imani, N. Chaparzadeh and A. R. Amani-Ghadim , “Improving the stability of uricase from Aspergillus flavus by osmolytes: Use of response surface methodology for optimization of the enzyme stability”, Process biochemistry 2020 Vol. 94 Pages 86-98.
[4] L. Rezaeian Marjani, M. Imani and H. Zarei Jaliani, “Enhancement of Pharmaceutical Urate Oxidase Thermostability by Rational Design of De Novo Disulfide Bridge,” Iran J Biotechnol 2020 Vol. 18 Issue 3 Pages e2662

ICB12-1402-1082

Evaluation of neural network for computing accuracy of genomic breeding value

Reza Behmaram*
Department of Animal Science, University of Mohaghegh Ardabili, Ardabil, Iran
*Corresponding author's email address: rezabehmaram881978@gmail.com

Abstract: Today, parametric and nonparametric statistical methods are used for genomic selection of traits with additive and epistatic genetic architectures. This study was conducted to investigate the effect of non-parametric method of neural network compared with of genomic best linear unbiased prediction (GBLUP) for calculating genomic breeding value using simulated data. Therefore, a genome containing 4 chromosomes, each 100 CM long, was created. Then, for supplying variation, genomic data were simulated with 1200 evenly distributed single nucleotide markers (SNP) and 120 randomly distributed quantitative trait loci (QTL) on each chromosome. In addition, additive allelic effects of QTL were determined with gamma distribution. Finally, shape, scale, and heritability level were considered 0.3, 1.72 and 0.25 respectively. Results showed that by considering only additive allelic effects, using nonparametric neural network approach had lower accuracy of genomic breeding value compared with genomic best linear unbiased prediction (0.72 ± 0.06 for neural network versus 0.76 ± 0.04 for GBLUP). While, including other effects more than additive effects resulted in higher accuracies of genomic breeding value in neural network method rather than genomic best linear unbiased prediction (0.71 ± 0.03 for neural network versus 0.69 ± 0.05 for GBLUP). Although in both approaches, prediction accuracy decreased significantly in next generations after reference population, but this decrease was higher for neural network. In conclusion, non-parametric neural network method can be used as well as genomic best linear unbiased prediction in conditions such as considering more effects than only additive allelic effects in genomic selection.

Keywords: accuracy; breeding value; genomic selection; neural network; simulation.

References

[1] Howard. R., Carriquiry, AL. and Beavis, WD., 2014. Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. G3: Genes| Genomes| Genetics: g3. 114.010298.

[2] Esfandyari, H. and Sørensen, A. C., 2017. "Xbreed: An R Package for Genomic Simulation of Purebreds and Crossbreds," in 68th Annual Meeting of the European Federation of Animal Science, Tallinn, Estonia, 28 Aug - 1 Sep 2017.

[3] Sahebalam, H., Gholizadeh, M., Hafezian, H. and Farhadi, A., 2019. Comparison of parametric, semiparametric and nonparametric methods in genomic evaluation. J. Genet. 98, 102.

[4] VanRaden P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91, 4414–4423.

ICB12-1402-1084

A Hybrid Deep Learning-based Model for Off-target Prediction in CRISPR System

Zeynab Arman[1], Leila Safari[1], Kian Sahafi[1], Alireza Khanteymoori[2]
1 Department of Computer Engineering, University of Zanjan, Zanjan, Iran
2 Department of Neurozentrum, Universitätsklinikum Freiburg, Freiburg, Germany
*Corresponding author's email address: zeynabarman1402@gmail.com

Abstract: The CRISPR system, as a gene editing method, has revolutionized the field of biology, provided that the exact target sites (on-targets), for gene editing are determined accurately, to avoid unintended side effects that could potentially harm cellular function. To address this issue, computational methods have been developed to accurately predict off-target locations. In this research, a hybrid deep learning model incorporating two neural networks, BiLSTM and CNN, has been proposed for identifying off-target sites in the CRISPR system. Due to the length and complexity of DNA sequences, a specialized encoding method is suggested for feeding information into the model. Utilizing k-mer sequence embeddings of various sizes using DNAtoVec, and calculating sequence mismatches at both nucleotide and K-mer levels, this model is capable of identifying specific patterns and features in sequences. Furthermore, the use of data augmentation and under sampling techniques has provided a balanced dataset to address the issue of data imbalance in this research. The evaluation results indicate that the proposed model surpasses the baseline models, achieving accuracy and F1-Measure metrics for predicting off-target sites that exceed 0.98.

Keywords: CRISPR; off-target positions; Increased Data; Deep Learning.

References:

[1] Listgarten, J., et al., Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. Nature biomedical engineering, 2018. 2(1): p. 38-47.

[2] Abadi, S., et al., A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. PLoS computational biology, 2017. 13(10): p. e1005807.

[3] Peng, H., et al., Recognition of CRISPR/Cas9 off-target sites through ensemble learning of uneven mismatch distributions. Bioinformatics, 2018. 34(17): p. i757-i765.

ICB12-1402-1088

Computational Designing of Chimeric Antigen Receptor Extracellular Regions: Homology modelling, Docking and Molecular Dynamic Simulations

Ramezani[a], V. Jajarmi[a], M. Soleimani[b], J. Ranjbari[a], W M, Van Weerden [c], M.Hassani[a]*
K.Bagherzadeh[d e]*
a Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Science, Tehran, Iran
b Department of Tissue Engineering and Applied Cell Science, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Science, Tehran, Iran
c Department of Urology, Erasmus Medical Center, Rotterdam, Netherlands
d Eye Research Center, The Five Senses Health Institute, Rassoul Akram Hospital, Iran University of Medical Sciences, Tehran, Iran
e Stem Cell and Regenerative Medicine Research Center, Iran University of Medical Sciences, Tehran, Iran.
*Corresponding authors'email addresses: Mahmoud.hassani@yahoo.com *Bagherzadeh.k@iums.ac.ir

Abstract: Chimeric antigen receptors (CAR) are recombinant receptors that have both the ability to bind to specific antigens and the ability to activate immune cells [1]. CAR T Cell technology provides a new way for the treatment of tumors including prostate cancer [2]. CARs generally consist of three parts: the extracellular, transmembrane and the intracellular region. The extracellular region usually includes the antigen recognition domain (commonly single-chain variable fragment of an antibody and nanobody (Nb) and spacer domain [2,3]. The advance in protein engineering and computational biology approaches lead to development of engineered nanobody-based CAR constructs. Herein, three nanobodies with different spacer length (short, intermediate and long) were designed to target prostate-specific membrane antigen (PSMA). The designed nanobodies structures were homology modeled followed by 250 ns of MDs to assess the stability and dynamic properties of the modeled structures. The minimized structures were then docked over PSMA by the application of ClusPro web server. Accordingly, the highest docking score and lowest Kd values of (-701.4 kj/mol) and (8.7e-08) were obtained for the one carrying the short spacer length CAR, respectively.

Keywords: Chimeric antigen receptor; Prostate-specific membrane antigen; Homology modelling; Molecular dynamic simulations; Molecular docking.

References:
[1] M. Sadelain, R. Brentjens and I. Rivière, "The basic principles of chimeric antigen receptor design," Cancer discovery., 3 (2013) 388-398.
[2] H. Yu, J. Pan, Z. Guo, C. Yang and L. Mao, "CART cell therapy for prostate cancer: status and promise," OncoTargets and therapy., 12 (2019) 391.
[3] M. Hassani, F. Hajari Taheri, Z. Sharifzadeh, A. Arashkia, J. Hadjati, W. M. van Weerden, et al., "Construction of a chimeric antigen receptor bearing a nanobody against prostate a specific membrane antigen in prostate cancer," Journal of cellular biochemistry., 120 (2019)10787-10795

ICB12-1402-1089

Binary Gravitational Search Algorithm Feature Selection method for DTI

Nasim Nozarpourshami[a*], Mahsa HassanpourAzizi[b], Jamshid Pirgazi[c]
a Ayatollah Boroujerdi University, Boroujerd, Iran
b University of Mazandaran, Babolsar, Iran
c Faculty of University of Science and Technology of Mazandaran, Behshahr, Iran
*Corresponding author's email address: Nasimpourshami@gmail.com

Abstract: Anticipating the interplay between pharmaceuticals and proteins stands as an indispensable stride within the realm of drug advancement, leading to the discovery of innovative methods. Experimental approaches relying on clinical treatments to identify these relationships are time-consuming, expensive, labor-intensive, and complex [1,2,3]. New computational methods, particularly those employing machine learning algorithms, offer a cost and time-efficient alternative to traditional experimental approaches[4]. This paper presents an innovative computational framework designed to forecast drug-target interactions, where various features are extracted from a dataset consisting of drugs and their corresponding SMILES, and proteins with their protein sequences. To address the abundance of features, BGSA (Binary Gravitational Search Algorithm) is applied. The carefully chosen features have the potential to significantly enhance the accuracy of the classification model, and minimizing the learning process's computational cost[5]. To improve the performance of BGSA, mutation operators are used to introduce random changes in the search process to avoid getting stuck in local optima. Such actions facilitate the algorithm's exploration of uncharted space and improve the quality of the solutions [6]. The Super Vector Machine (SVM) classification is utilized to predict the efficiency of the method using the ultimately chosen features. The SVM classifier's accuracy on the designated golden standard datasets (Enzyme, ion channels, G-protein-coupled receptors, nuclear receptors) is provided below before feature selection: 95.94%, 95.08%, 87.99%, and 70.83% and after feature selection are 98.63%, 98.47%, 96.06%, and 97.22% respectively. According to the experimental results, the proposed method showcases a commendable level of accuracy in predicting drug-target interactions (DTI) and harmonizes effectively with methodologies advocated in existing publications.

Keywords: Drug-Target Interactions; Feature Selection; Binary Gravitational Search; SVM

References

[1] Ye, Q., Hsieh, CY., Yang, Z. et al. "A unified drug–target interaction prediction framework based on knowledge graph and recommendation system." Nat Commun 12, 6775 (2021) .

[2] DiMasi, J. A., Hansen, R. W. & Grabowski, H. G. "The price of innovation: new estimates of drug development costs." J. Health Econ. 22, 151–185 (2003)

[3] Paul, S. M. et al. "How to improve R&D productivity: the pharmaceutical industry's grand challenge." Nat. Rev. Drug Discov. 9, 203–214 (2010)

[4] B. T. Yaseen, "Drug Target Interaction Prediction Using Convolutional Neural Network (CNN)," 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Istanbul, Turkiye, 2023, pp. 1-5.

[5] M. HassanpourAzizi, N. Nozarpourshami and J. Pirgazi, "Chaotic Modified Binary Grey Wolf Optimizer Feature Selection method for Drug-Target interaction prediction Based on drug-protein Features," 2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Behshahr, Iran, Islamic Republic of, 2022.

ICB12-1402-1090

Omics Technologies and Data mining approaches in Canine Osteosarcoma: Insights from High-Throughput Genetics and Bioinformatic Analysis

Elham Yekzaman[a], Seyed Amir Hossein Hosseini[a*], Mohamad Zamani Ahmad Mahmudi[b]
a Shahid Bahonar University of Kerman, Kerman, Iran
b Department of Clinical Science, Shahid Bahonar University of Kerman, Kerman, Iran
*Corresponding author's email address: a.hosseini.bio@gmail.com

Abstract: Osteosarcoma is the most common primary bone malignancy which originates from primitive bone-forming mesenchymal cells. In canines, it constitutes 85% of the tumors that occur in the bone. By the emersion of omics technologies and molecular phenotyping analysis (e.g., genomics, transcriptomics, proteomics, and metabolomics), our recent achievements in high-throughput genetics have attracted more attention than ever before. The resulting information from these methods can be utilized to enhance our understanding of diseases through the assessment of various correlations. The data can be fitted into a logical framework that shows the affected pathways and different factors that contribute to the development of a disease. Precision medicine, tumor classification and phenotyping small animals as models for enhancing the understanding of human biology and diseases are some examples of these technologies. In this study, we used 4 datasets and analyzed their microarray data with different bioinformatic software, such as R, geWorkbench, GSEA and PennCNV. We assessed the obtained results from our analyzes, the up and down regulation of gene expressions and clustering and different gene sets that were harvested from our analyzes. We also detected CNV's in one of the data sets with 570 cases. Our findings showed relations between osteosarcoma and gene sets related to lipid cycles, such as Adipogenesis, Acid metabolism, and Oxidative phosphorylation, which were acquired from MSigDB. The change in lipid profile is undeniable, and it should be considered in future studies of cancer treatment and other related research.

Keywords: Osteosarcoma; Omics technologies; Gene expression

References

[1] G. Selvarajah, J. Kirpensteijn, "Prognostic and predictive biomarkers of canine osteosarcoma," Veterinary journal (London, England: 1997) vol. 185,1 (2010): 28-35.

[2] Y. Hasin, M. Seldin, A. Lusis, "Multi-omics approaches to disease," Genome biology vol. 18,1 83. 5 May. 2017

ICB12-1402-1091

Analyzing Monomeric and Dimeric ORF8 Interactions with Antiviral Small Molecules

Elham Abdollahi[1] and Maryam Azimzadeh Irani[1]*
[1] Faculty of Life Science and Biotechnology, Shahid Beheshti University, Tehran, Iran.
*Corresponding author's email: m_azimzadeh@sbu.ac.ir

Abstract: Open reading frame (ORF8) is a SARS-CoV-2 accessory protein with monomeric and dimeric structures [1][2][3]. ORF8 exerts multiple different functions that interfere with host immune responses, including downregulation of MHC class I molecules, inducing ER stress and triggering inflammatory response through IL-17 pathway [1]. In terms of connections with human protein networks, ORF8, along with the membrane(M) protein and Nsp7, was statistically identified as a significant protein of SARS-CoV-2 [1][4]. Therefore, studying how ORF8 structures interact with antiviral drugs can help treat COVID-19 and identify new targets for vaccines and drugs**.** ORF8's 3D structures were obtained from Protein Data bank with PDB ID 7F5F [2] for monomeric structure and PDB ID 7JTL [3] for dimeric structure. Remdesivir CID 121304016, Artemisinin CID 68827 and Ivermectin B1a CID 6321424 were selected as antiviral small molecules and their structures were obtained from PubChem database [5]. Blind molecular docking between each chosen drug with ORF8 monomeric and dimeric structures was performed using Autodock4 [6]. The ORF8 amino acids that form hydrogen bonds with Remdesivir are ASP46, CYS44, GLU42 (monomer) and GLN111, ALA3 (dimer); with Artemisinin is ARG84, LYS27 (monomer) and ARG34 (dimer) and with Ivermectin B1a are LYS51, HIS23, PHE24, CYS66, PHE 69(monomer) and GLN13, ARG145 (dimer).

The results indicate that small molecules binding sites are not overlapping. Furthermore, based on the ORF8 amino acids involved in hydrogen bonds with the drugs, the binding epitopes of monomeric and dimeric ORF8 differ. This information can help develop drugs that target specific binding sites and epitopes.

Keywords: SARS-CoV-2; ORF8; Molecular Docking; Antiviral small molecules

References

[1] Vinjamuri S, Li L, Bouvier M. SARS-CoV-2 ORF8: one protein, seemingly one structure, and many functions. Front Immunol. (2022)

[2] Chen X, Zhou Z, Huang C, Zhou Z, Kang S, Huang Z, Jiang G, Hong Z, Chen Q, Yang M, He S, Liu S, Chen J, Li K, Li X, Liao J, Chen J, Chen S. Crystal structures of bat and human coronavirus ORF8 protein Ig-like domain provide insights into the diversity of immune responses. Front Immunol. (2021)

[3] Flower TG, Buffalo CZ, Hooy RM, Allaire M, Ren X, Hurley JH. Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. Proc Natl Acad Sci U S A. (2021)

[4] Assadizadeh M, Azimzadeh Irani M. Oligomer formation of SARS-CoV-2 ORF8 through 73YIDI76 motifs regulates immune response and non-infusion antiviral interactions. Front Mol Biosci. (2023)

[5] https://pubchem.ncbi.nlm.nih.gov

[6] Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Computational Chemistry* 2009, 16: 2785-91

ICB12-1402-1093

Identification of specific gene modules related to the Meyloid Blast Crisis (MBC) phase of CML Using MAGI Algorithm

N. Elmi[a] , GE Kazemi-Sefat[b] ,K. Kavousi[a] , S.Talebi[b] , K. Mousavizadeh[c*]
a Institute of Biochemistry and Biophysics (IBB), Department of Bioinformatics, Laboratory of Complex Biological Systems and Bioinformatics (CBB), University of Tehran, Tehran, Iran
b Department of Medical Genetics, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
c Department of Pharmacology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: Mousavizadeh.k@iums.ac.ir

Abstract: Chronic myeloid leukemia (CML) is a type of cancer that is classified into three phases: Chronic, Accelerated, and Blast crisis, while the myeloid blast crisis (MBC) phase is resistant to treatment [1]. Despite evidence in the BCR-ABL fusion gene as the most significant genetic abnormality and a key for the initiation and supporting the continuation of CML, there are fascinating proofs that CML harbors many other genetic alterations in oncogenes as well as tumor suppressor genes that contribute to the progression of this malignancy. Some of these changes are drivers that map to signaling pathways and control cell growth, cell cycle progression, and apoptosis. While the others act as cancerpredisposing variants, most of them are germline. The affected genes may work together in connected networks as modules and play a role in the pathogenesis of MBC-CML. In this study, we tried to find potential disease modules, as a set of genes that enrich the novel variants in cases with MBC-CML compared to controls which were Imatinib-responsive CML. To this end, we utilized the MAGI algorithm [2], which merges different omics data including gene co-expression network, ProteinProtein interaction, and genetic variant data in case and control groups. MAGI has been used for neurodevelopmental diseases such as autism and Intellectual disability. But for the first time, we propose using this algorithm for cancer diseases. To this end, we used HPRD data as PPI network input, and we constructed a gene co-expression network for CML patients based on the GSE42519 micro-array dataset [3]. We also used European Genome-Phenome Archive (EGA) data with accession EGAS00001003071 as our case and control samples [4]. We selected 18 samples with the MBC-CML case group and 44 samples with the CML control group. We analyzed the fastq files based on GRCh38 genome assembly using BWAMEM as the aligner, and Mutect2 as the variant caller. We used the Mutect2 algorithm from the GATK4 toolkit to call somatic and germline variants. Then Several tools were used to annotate and interpret variants including CGI, Varsome, InterVar, CancerVar, and Franklin. The annotated VCF files were filtered to find Tier 1 and Tier 2 variants according to AMP classification and Tier 3 variants which were pathogenic (P) based on ACMG classification. Finally, we created a table of important gene variants for case and control groups. Then the MAGI algorithm was run to find the modules including novel genes based on PPI and co-expression network. The Module with the best score was selected as the best module and enriched with the EnrichR web tool. The Enriched terms were highly associated with CML and cancer and as a result, the novel genes can be potentially associated genes with CML resistance to the treatment.

Keywords: Myeloid blast crisis; CML; Network; Gene Module; MAGI algorithm

ICB12-1402-1094

Efficient classification of single cell ATAC sequence data using machine learning

H. Haririmonfared[a], N. Elmi[b], K. Kavousi[b*], B. Majidi[a*]
a Department of Computer Engineering, Khatam University, Tehran, Iran.
b Institute of Biochemistry and Biophysics (IBB), Department of Bioinformatics, Laboratory of Complex Biological Systems and Bioinformatics (CBB), University of Tehran, Tehran, Iran
*Corresponding authors' emails: b.majidi@khatam.ac.ir ,
kkavousi@ut.ac.ir

Abstract: Based on recent advances in next-generation sequencing technology and microfluidics, we are now able to study the structure of DNA of each cell type. Transposase-accessible chromatin sequencing (ATAC-seq) method [1], which uses Tn5 transposase to determine chromatic accessibility across the genome, provides more detailed information about chromatin packaging and also the molecular mechanism behind gene expression regulation in different tissues and phenotypes. On the other hand, Single-cell omics technologies gifted us the opportunity to analyze thousands of cells simultaneously and distinguish between different cell types. One of the challenges in the single cell analysis including ATAC-seq data is cell type annotation. To do this step one approach is using classification algorithms [2]. In this paper, we have investigated the performance of 6 well-established machine learning methods for the classification of single scATAC-seq data. The main performance criteria in this research were computational complexity and computational resource consumption. Hence, a new classification method based on Extremely Randomized Trees (ERT) [3] was proposed which had faster performance while keeping the classification accuracy constant. For evaluation, the performance of four public scATAC-seq datasets which were from different studies from various organisms and tissues was used to evaluate the performance of the methods. The results showed that these methods performed well in some specific cell types in a particular scATAC-seq dataset. In both intra-dataset and inter-dataset tests, while support vector machine (SVM) and nearest mean classifier (NMC) showed overall better performance than other methods in all 4 datasets; ERT method was able to perform the classification operation in significantly less time with almost the same accuracy.

Keywords: Single cell ATAC-seq; Machine Learning; Extremely Randomized Trees;Classification

References

[1]J. D. Buenrostro, B. Wu, H. Y. Chang, and W. J. Greenleaf, "ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide," (in eng), *Curr Protoc Mol Biol,* vol. 109, pp. 21.29.1-21.29.9, Jan 5 2015, doi: 10.1002/0471142727.mb2129s109.

[2]H. Guo, Z. Yang, T. Jiang, S. Liu, Y. Wang, and Z. Cui, "Evaluation of classification in single cell atac-seq data with machine learning methods," *BMC Bioinformatics,* vol. 23, no. 5, p. 249, 2022/09/21 2022, doi: 10.1186/s12859-022-04774-z.

[3]P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning,* vol. 63, no. 1, pp. 3-42, 2006/04/01 2006, doi: 10.1007/s10994-006-6226-1.

ICB12-1402-1101

Unraveling Therapeutic Prospects for Endometriosis: Modulating AKT1 with an Established Breast Cancer Drug, Capivasertib

Mahdokht Mohammad Esmaili [1], Soodeh Mahdian [2], Maryam Shahhoseini [3]*
1 Department of Cell and Molecular Biology Science, Academic Center for Education Culture and Research, Tehran, Iran.
2 Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, Academic Center for Education Culture and Research, Tehran, Iran.
3 Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, Academic Center for Education Culture and Research, Tehran, Iran.
*Corresponding author's email address: m.shahhoseini@royaninstitute.org

Abstract: Endometriosis, a prevalent and debilitating gynecological disorder affecting millions of women worldwide, is characterized by the presence of endometrial-like tissue outside the uterine cavity [1]. This study explores novel therapeutic strategies for endometriosis by targeting RAC-alpha serine/threonine-protein kinase (AKT1) with drug Capivasertib. Leveraging the Gene Expression Omnibus (GEO) dataset, significant alternation in gene expression patterns associated with endometriosis were identified. Network analysis using STRING and Cytoscape highlighted AKT1 as a central player in endometriosis-associated pathways. In silico exploration of DrugBank revealed Capivasertib, an FDA-approved inhibitor of pathway leading to AKT1 upregulation in breast cancer [2]. Despite no prior application in endometriosis, the potential of Capivasertib to modulate AKT1 expression introduces an innovative therapeutic hypothesis. We propose the strategic use of Capivasertib as a transformative approach for endometriosis, building upon the established safety profile in breast cancer.

Additionally, our in-silico investigation suggests that the upregulation of AKT1 may serve as a potential biomarker for endometriosis. Differential expression levels of AKT1 could offer diagnostic and prognostic insights, enhancing clinical assessments .

This study assumes significance in addressing the unmet medical needs of individuals with endometriosis. With a dual focus on precision therapeutics and biomarker identification, the potential implications extend to improved patient outcomes, personalized treatment strategies, and a deeper understanding of molecular mechanisms driving endometriosis.

Keywords: Endometriosis; AKT1; Capivasertib; Therapeutic Prospects; Biomarker;

References

[1] K. Zondervan, C. M. Becker, "Endometriosis," N Engl J Med, 13 (2020) 1244-1256.

[2] N C. Turner, M. Oliveira, "Capivasertib in hormone receptor-positive advanced breast cancer," N Engl J Med, 22 (2023) 2058-2070.

ICB12-1402-1102

he role of imatinib on the differentiation genes expression in K562 cell line: A bioinformatic study

Seyed Mahdi Mousavi[1*], Yaghub Pazhang[1]
1: Department of Biology, Faculty of Sciences, Urmia University, Urmia, Iran.
*Corresponding author's email address: st_ma.mousavi@urmia.ac.ir

Abstract: Bcr-Abl protein is a tyrosine kinase enzyme that is involved in the growth and carcinogenesis of K562 cells in chronic myeloid leukemia cancer [1, 2].Therefore, the inhibitors of this enzyme such as imatinib can reduce the growth of K562 cells[2] . In this study, the aim of this work was to investigate the role of imatinib in the expression of the differentiation genes of K562 cells into red blood cells [3]. In this study, the GEO site the GEO2R tool were used to find and analyze the dataset. After surveying the site, and checking the different datasets, the GSE19567 dataset was chosen to analyze the imatinib effect on K562 cell line gens expression profile[4]. Analysis of the GSE19567 dataset displayed over-expression of ALAS2, HBB, HBD, GYPA, FECH, TAL1 genes in imatinib treated K562 cells by 22.3, 11.7, 10.7, 4.2, 3.01, and 2.4 times, respectively. The results showed that the expression of genes involved in the metabolism of heme and hemoglobin and also Glycophorin gene has been increased in treated cells. According the role of the genes in the differentiation of the K562 cells, it can be suggested that imatinib induces differentiation in the cells. It may be related to inhibition the Bcr-Abl protein which causes K562 cell growth.

Keywords: Bcr-Abl tyrosine kinase; GEO, Imatinib; K562.

References

[1]Desogus, A., et al., *Bcr-Abl tyrosine kinase inhibitors: a patent review.* Expert Opinion on Therapeutic Patents, (2015). **25**(4): p. 397-412.
[2]Tanaka, A., et al., *Tyrosine kinase inhibitor imatinib augments tumor immunity by depleting effector regulatory T cells.* Journal of Experimental Medicine,(2020). **217**(2).
[3]Shin, D., et al., *Multiplexed single-cell RNA-seq via transient barcoding for simultaneous expression profiling of various drug perturbations.* Science Advances,(2019). **5**(5): p. eaav2249.
[4]Larson, R.A., et al., *Nilotinib vs imatinib in patients with newly diagnosed Philadelphia chromosome-positive chronic myeloid leukemia in chronic phase: ENESTnd 3-year follow-up.* Leukemia, (2012). **26**(10): p. 2197-2203.

ICB12-1402-1103

Computational vaccine designing against Nipah virus G and F proteins using structural modeling and dynamic simulations

Ali Baradaran [a,b], Mohammad Mostafa Pourseif [b,c]*, Behzad Jafari [d], Sarah Gholami [e]
a Faculty of Pharmacy, Urmia University of Medical Sciences, Urmia, Iran
b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
c Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
d Department of Medicinal Chemistry, Faculty of Pharmacy, Urmia University of Medicinal Sciences, Urmia, Iran
e Department of Medical Biotechnology, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: The Nipah virus (NiV), a member of the Henipavirus genus, sharing similarities with the Hendra virus [1]. The emergence of NiV as a deadly zoonotic pathogen and sporadic human-to-human transmission as well as the absence of specific treatments underscore the urgent need for comprehensive vaccine design and development strategies to mitigate the potential threat of a future pandemic [2]. Urgency in rational vaccine development is underscored, necessitating the application of bioinformatics advancements to meet the pressing need for timely and cost-effective approaches [3]. This study focuses on computationally designing a candidate vaccine with potential broad-spectrum protection against the NiV.This study revolves around the thorough immunoinformatics analysis of the Nipah virus's G and F surface antigens. Through the utilization of machine learning (ML) algorithms for B-cell epitope (BCE) identification, regions with significant potential were systematically pinpointed. Furthermore, prediction of HLA class I and II-associated peptide binders (including B40:01, B44:03/02, A68:02/01, A02:03/01, DRB101:01, DRB107:01) was conducted using the IEDB NXG standalone tool, and analyzed with R language programming. Molecular docking and dynamic simulations (200 ns) using ZDOCK and GROMACS software were employed to study the conformational changes and stability of the designed vaccine in interaction with human immune cells, specifically anti-NiV Fab antibodies and HLA molecules associated with resistance to NiV.The results from the ML-based analysis and dynamics simulations showed that the immunodominant regions of the designed vaccine exhibit high-affinity and stable interactions with anti-NiV Fab antibodies and the selected MHC-I/II molecules. These identified regions serve as the basis for in silico vaccine design against the Nipah virus, ensuring comprehensive immunoprotection. Our computational vaccine design strategy considers the dynamic nature of viral proteins, harnessing the power of in silico methods to expedite the vaccine development process. In silico epitope mapping and dynamic simulations revealed that the candidate vaccine can form a more stable interaction with the specific antibodies and HLA molecules and will be useful in experimental developing an anti-NiV candidate vaccine.

Keywords: Computational vaccinology; Dynamic simulations; Epitope; Nipah virus; Structural modeling.

ICB12-1402-1104

Evaluation of 5-fluorouracil related genes in breast cancer involved in the microtubule structure

Fatemeh Mirzamahmoudi *[a], Dr.Yaghub Pazhang [b], Dr.Hamid Maadi
a Department of Biology, Urmia University, Urmia, Iran
b Department of Cell and Molecular Biology & Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: st_f.mirzamahmoudi@urmia.ac.ir

Abstract: Krüppel-like factors (KLFs) are implicated in several cellular functions. Numerous biological processes, including cell stemness, proliferation, apoptosis, autophagy, and migration, have been linked to it. [1]. KIF20A is an essential protein for cell division and causes chromosome migration during mitosis. Therefore, it is necessary for the survival and morphogenesis of cells [2]. The formation of microtubules is achieved by polymerizing the heterodimers of α- and β-tubulins. There are at least nine isoforms of α-tubulin in humans, including tubulin α8 (TUBA8), and nine isoforms of β-tubulin in humans including tubulin β2A class IIA (TUBB2A), β2B class IIB (TUBB2B) [3]. One of the most popular chemotherapy agents for the treatment of solid tumors is 5-fluorouracil (5-FU). The current study aimed to determine the effect of 5-FU on tubulin gene expression [4]. By utilizing NCBI's GEO dataset and GEO2R, the genes that are differently expressed under different experimental settings within the gene expression patterns caused by 5-FU in breast cancer cell lines were identified, then the GSE2584 dataset was chosen. The results showed that in the treated cells, the expression of the TUBA8, TUBB2, KLF5, and KIF20A genes was reduced in comparison to control cells. Based on the results it can be suggested that the 5-FU may modulate the microtubule formation by downregulating the tubulin gene expression.
Keywords: KLF5; KIF20A; TUBB2; TUBA8; 5-fluorouracil drug; MCF7 cells.

References

[1] Luo, Y. and C. Chen, The roles and regulation of the KLF5 transcription factor in cancers. Cancer science, 2021. 112(6): p. 2097-2117.

[2] Jin, Z., et al., Expression, regulating mechanism and therapeutic target of KIF20A in multiple cancer. Heliyon, 2023.

[3] Nami, B. and Z. Wang, Genetics and expression profile of the tubulin gene superfamily in breast cancer subtypes and its relation to taxane resistance. Cancers, 2018. 10(8): p. 274.

[4] Hernández-Vargas, H., et al., Transcriptional profiling of MCF7 breast cancer cells in response to 5-Fluorouracil: Relationship with cell cycle changes and apoptosis, and identification of novel targets of p53. International journal of cancer, 2006. 119(5): p. 1164-1175.

ICB12-1402-1107

Mathematical modeling for azole drugs solubility in the binary solvent systems at various temperatures

Mahmoud Omidi[a,b], Seyyed Mohammad Hasan Hashemi[c], A.Maryam Babaei[d], C.F.Esmaeil Mohammadian[a,d*]
a. Molecular Medicine Research Center, Hormozgan Health Institute, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
b. Department of Pharmacology and Toxicology, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
c. Department of Pharmaceutics, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
d. Department of Medicinal Chemistry, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
*Corresponding author's email address: e.mohamadyan1390@gmail.com

Abstract: Azoles-based drugs in two distinct classes including imidazoles (include clotrimazole, econazole, ketoconazole, miconazole, and tioconazole) and triazoles (fluconazole, itraconazole, Posaconazole, and voriconazole) can be used to inhibit the growth of wide range fungal infections including athlete's foot, onychomycosis, ringworm, and vaginal candidiasis. Due to their hydrophobic structure, many azole drugs have low aqueous solubility. Various cosolvency systems have been investigated to improve the aqueous and non-aqueous solubility of these drugs. Using mathematical models as an alternative approach to estimating drug solubility is an easy and fast method. In this study, to predict the solubility of some azole-based drugs in binary systems of solvents at various temperatures, several mathematical models were developed based on the Jouyban-Acree model. To obtain more accurate mathematical models, the Jouyban-Acree model is combined with the physicochemical parameters of the Abraham solvation parameters and the Hansen solubility parameters. In order to assess the applicability of the proposed methods, we calculated the average mean relative deviation (MRDs %) values for the solubility data of azole-based drugs in some cosolvency systems. The result of the present study has shown that the combination models based on Abraham solvation parameters have reduced the MRDs % by about 3 times compared to Hansen solubility parameters. This study provides essential information in the pharmaceutical industry that can be invaluable in selecting the appropriate solvents for extracting and purifying these drugs.

Keywords: Solubility; Azole drug; Abraham solvation parameters; Hansen solubility parameters; Mathematical model.

References

[1] Mohammadian, E., et al., Simulation of mesalazine solubility in the binary solvents at various temperatures. Journal of Molecular Liquids, 2022. 357: p. 119160.

[2] Karimzadeh, Z., et al., Employing Abraham and Hansen parameters for solubility prediction of ketoconazole in binary solvents at various temperatures. Journal of Solution Chemistry, 2021: p. 1-20.

ICB12-1402-1109

Analysis of Co-Expression Network of BMPR1B Gene with Other Genes Involved in litter Size of Sheep

Zahra Roudbari[a]
a Department of Animal Science, Faculty of Agriculture, University of Jiroft, Jiroft, Iran
*Corresponding author's email address: roudbari.zahra@ujiroft.ac.ir

Abstract: The BMPR1B gene, known as FecB, is critical in sheep reproductive biology and carries a mutation at the 746bp position, which significantly increases the rate of ovulation and litter size. This gene is deeply connected to reproductive processes, as the mutation influences both regulate external estrous characteristics and follicle stimulating hormone in the estrous cycle [1]. Research has shown that this mutation in BMPR1B can control the activity of certain genes within reproductive tissues like the hypothalamus, pituitary gland, and ovaries throughout the estrous cycle [2]. The co-expression network, an essential tool in biological system analysis, sheds light on genes that change expression together [3]. In this study, a co-expression network was constructed using the Cytoscape software for BMPR1B, linking it with other relevant genes. This led to the identification of 10 genes (BMP7, BMPR2, BMP15, BMP2, BMP4, BMP6, FSHR, GDF5, GDF9, SMAD9) that interact with BMPR1B and were organized into two co-expression modules via K-means clustering. Further investigation into the roles of these genes, using the String database for interaction analysis, indicated their significant involvement in ovarian steroidogenesis and TGF-beta signaling pathways. These insights have greatly expanded our understanding of the genetic factors that influence ovulation and litter size in sheep. The genes studied may also serve as valuable biomarkers for selecting sheep with superior reproductive traits in breeding programs.

Keywords: Litter Size; Regulated Network; BMPR1B gene; Ovine Reproduction.

References

[1] R. Kumar, R.C.Sharma, Gopal, R. Gowane, A. Kumar.Genetic variability in 5' untranslated region of BMPR1B gene from prolific Avishaan strain of sheep. Indian Journal of Small Ruminants., 28(2022):16-19.
[2] X. Wang, X. Guo, X. He, R. Di, X. Zhang, J. Zhang, M. X. Chu. Integrated Proteotranscriptomics of the Hypothalamus Reveals Altered Regulation Associated with the FecB Mutation in the BMPR1B Gene That Affects Prolificacy in Small Tail Han Sheep. Biology, 12(2022):72-72.
[3] R. Xulvi-Brunet and H. Li. Co-expression networks. Bioinformatics, 26(2010):205-214.

ICB12-1402-1110

Association of IL18 and CSF1R overexpression leads to immunosuppressive phenotypes in microglial cells in GBM microenvironment: A transcriptome-wide meta-analysis

Elnaz Amanzadeh Jajin*,a
a Functional Neurosurgery Research Centre, Shahid Beheshti University of Medical Sciences
*Corresponding author's email address: elnazamanzadehjajin@gmail.com

Abstract: GBM is the most aggressive brain tumor in which glial cells undergo several changes leading to tumor progression. The role of astrocytes in GBM progression is well established, but its role in connection with other glial cells and induction of invasion still remains unclear. Hereby, we used single RNA sequencing data analysis in order to perform an integrative analysis of glioblastoma tissue samples. Cell type identification was performed using SingleR package which executes cell type identification based on machine learning algorithms. Among about 10000 cells data analyzed, we obtained three main clusters of cells including astrocyte, microglia, and oligodendrocytes. Interestingly, each cell type included several clusters. We hypothesized the presence of various cell phenotypes for each cell type among the identified clusters. Accordingly, clustering was performed to identify the more similar clusters for identified cell types. This process led to find two phenotypes for astrocyte including active astrocyte and astrocyte/stem cell, three phenotypes for microglia including CD163/CD14 microglia, CD33+ or steady state microglia, and active microglia, immature oligodendrocyte, and neurons. Studying the gene networks and related signaling pathways using Cytoscape and GSEA analysis led to find that astrocytes show proliferative and growth factor producing activities, while active astrocyte phenotype shows mitochondrial dysfunction and apoptotic changes. Microglial subpopulations show a wide range of diversity in expression levels of marker genes including interleukins. Accordingly, it seems that each phenotype of microglia show increase in differential expression levels of interleukins and growth factors, so that active microglia show higher expression levels of VEGF, while CD163/CD14 microglia shows higher expression levels of TGF-β and HBEGF, and CD33+ microglia show higher expression levels of PDGFB in comparison to other types. Similar results found for astrocytes, so that stem cell/astrocytes show higher expression levels of EGFR and PDGFR than active astrocytes, while active astrocytes show higher expression levels of VEGF. In addition, our findings revealed that astrocytes overexpression of CSF1 as a tumorigenic factor. However, overexpression of CSF1R in microenvironment were found in microglia subpopulations. In addition, studying the common genes between three phenotypes of microglia showed that there are 12 common genes between CD163/CD14, and active microglia microglia, 75 common genes between CD163/CD14 and CD33+, and 112 common genes between CD33+ and active microglia. GSEA analysis showed that PI3K, ECM-receptor interaction, and Focal adhesion signaling pathways are among the most frequent singling pathways in microglia and astrocytes withing the microenvironment glioblastoma. According to the results of pathway enrichment analysis, immunosuppressive microglia show overexpression of cytokines and inflammatory responses. Our results for the first time showed association of overexpression between CSF1R and IL18 as tumorigenic and inflammatory cytokine leading to transition of microglia to immunosuppressive phenotype. Considering that the simultaneous overexpression of IL18 and CSF was found in other cancers as prognostic factor. On the other hand, it was found that using IL-18
and CSF as adjuvants for tumor vaccines can induce immune response in colon and spleen cancer. In line with these findings, we found overexpression of CSF1R in various clusters of microglial cells which was associated with overexpression of IL18 in all those clusters. Considering that no studies has reported the association between CSF1R and IL18 in GBM or other types of brain tumors, it is the first time that this association is studied.

Keywords: glioblastoma multiforme, reactive astrocyte, GBM microenvironment, IL18, CSF1R, microglia

ICB12-1402-1111

Integration of genomic data and machine learning to identify signature genes in pyroptotic cell death in neural cells of AD patients

E. Amanzadeh Jajin
Functional neurosurgery Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Ira
*Corresponding author's email address: elnazamanzadehjajin@gmail.com

Abstract: Alzheimer's Disease is the most prevalent neurodegenerative diseases in which neural cells die in various ways. Pyroptosis is a cell death type which induces inflammation in neural cells in AD via formation of inflammosome, but its regulatory mechanisms in this disease are not clear yet. Neural cell loss leads to AD progression and no exact treatment method has been found to suppress this process. In this study, we aimed to explore the regulatory mechanisms underlying pyroptosis in AD brain tissue. For this purpose, multiple microarray datasets of AD samples were integrated and analyzed to find differentially expressed genes in comparison with healthy control samples. Then, ML algorithms were used to find the signature genes via LASSO algorithm. In addition, the signature genes were used to differentiate AD and control samples. GSEA analysis was used to find key signaling pathways. WGCNA analysis was used to find the key genes associated with this cell death. Our results revealed that four genes including TLR1, IL-32, TAC1, and S100A4 are signature genes for pyroptosis in AD. These genes are expressed differentially in AD samples compared to the control groups and can be used to differentiate AD and control samples. GSEA analysis showed that inflammatory response, IL6-JAK-STAT Signaling, TNF-alpha signaling via NF-κB, and cytokine-cytokine receptor signaling pathways are closely related to the pyroptosis. WGCNA analysis yielded one cluster of mostly associated genes among which cytokines and TLRs were main regulatory nodes in the results of protein-protein interaction analysis. We developed a robust and novel signature for pyroptosis identification in AD based on the machine learning algorithms and bioinformatic tools. Notably, the found genes have not been studied in AD as therapeutic or prognostic factors, while we found them to be effective signature genes with potential role in regulation of pyroptosis in AD.
Keywords: Alzheimer's Disease; pyroptosis; machine learning; signature gene.

ICB12-1402-1113

Celastrol strongly binds to AMPK activating site: A docking study

Parisa Zallou[1*], Yaghub Pazhang[1]
1 Department of Biology, Faculty of Sciences, Urmia University, Urmia, Iran
*Corresponding author's email address: st_p.zallou@urmia.ac.ir

Abstract: The AMP-activated protein kinase (AMPK) is a central regulator of cellular metabolism and energy homeostasis in mammalian tissues. According to key role of AMPK in controlling energy homeostasis, its activation can be related with metabolic disease, including type 2 diabetes and cancer treatment. Therefore, AMPK activators can be potential anti-cancer agents "[1]". The aim of this study is evaluation of Berberine, Caffeic acid, Celastrol, Curcumin (Cum), Luteolin, Metformin"[2]", Salicylat, Staurosporine (Stu)"[3]", R34 and 6VT on AMPK by molecular docking. Three-dimensional (3D) structure of the compounds "[4]" and AMPK (two states, 6C9F and 5KQ5) were obtained from the PubChem database and Protein Data Bank (PDB) respectively. Finally, the molecular docking was studied using Autodock 4.1 software. Analyses of molecular docking exhibited that STU ligand is present in both active sites of 6C9F and 5KQ5 proteins with highest binding free energy with -11.68 kcal/mol in 6C9F and -13.34 kcal/mol in 5KQ5 compared to the other compounds. After the STU, Celastrol displayed the highest binding energy with -10.44 kcal/mol and Metformin showed the lowest binding energy with -3.03 kcal/mol. According to the obtained results, it was suggested that among the compounds (except for STU) Celastrol may have had activating effect on AMPK. Therefore, AMPK can be a potential target to celastrol in cancer treatment.

Keywords: AMPK; Cancer treatment; Celastrol; Molecular docking.

References

[1] Carling, D. and M.C.S. Centre, "*AMPK signalling in health and disease*", Current opinion in cell biology ,2017. **45**: p. 31-37.
[2]Cioce, M. and C. Pulito, "*Metformin: Metabolic rewiring faces tumor heterogeneity.* Cells", 2020. **9**(11): p. 2439.
[3]Tanramluk, D. and A. Schreyer, "*On the origins of enzyme inhibitor selectivity and promiscuity: a case study of protein kinase binding to staurosporine.* Chemical biology & drug design", 2009. **74**(1): p. 16-24.
[4]Gullett, N.P. and A.R. Amin. " *Cancer prevention with natural compounds*", 2010. Elsevier

ICB12-1402-1115

Understanding the Activity of BMPR1B within the BMP/SMAD Pathway in Sheep

Zahra Roudbari
Department of Animal Science, Faculty of Agriculture, University of Jiroft, Jiroft, Iran
*Corresponding author's email address: roudbari.zahra@ujiroft.ac.ir

Abstract: BMPR1B gene is a significant component in the BMP/SMAD signaling pathway, holding importance in numerous biological functions, especially reproductive processes. In sheep, BMPR1B has been identified as a key gene affecting litter size, influencing prolificacy and fertility [1]. This gene's activity within the BMP/SMAD pathway plays a critical role in regulating essential processes like ovarian follicular development, ovulation, and embryonic growth [2-3]. To investigate this, a study was conducted using network construct to explore the impact of BMPR1B activity and the BMP/SMAD pathway. The research aimed to comprehend the genetic factors influencing litter size in sheep and their implications for breeding and production. The result revealed that BMP ligands bind to BMPR1B, leading to receptor dimerization and activation. Upon ligand binding, BMPR1B forms a heterotetrameric complex with BMPR2, leading to phosphorylation and activation. This activated BMPR1B then phosphorylates receptor-regulated SMADs (R-SMADs), predominantly SMAD1, SMAD5, and SMAD8.These phosphorylated R-SMADs create complexes with SMAD4 (co-SMAD), which subsequently translocate into the nucleus, acting as transcription factors. They regulate gene transcription by binding to specific DNA sequences and interacting with other transcriptional regulators. The activity of BMPR1B and the BMP/SMAD signaling is under stringent regulation by extracellular antagonists, intracellular inhibitors, and feedback mechanisms. Understanding the impact of BMPR1B activity on the BMP/SMAD pathway offers valuable insights into the molecular mechanisms governing fertility traits, specifically litter size, in sheep. The findings also illuminate the intricate interactions within the BMP/SMAD pathway and their role in regulating litter size, providing crucial information for the sheep farming industry.

Keywords: BMPR1B gene; BMP/SMAD pathway; Pathway Analysis; Ovine Reproduction.

References

[1] D. Marius, Adriana, M. Marius, G. Aipătioaie, A. Ionascu, et al., Single-Nucleotide Polymorphisms Identified within Exon 2 of Fertility-Associated Bone Morphogenetic Protein (BMP15) Gene in Three Romanian Sheep Breeds. Agriculture.13(2023):996-996.

[2] Y. Gong, X. Wang, X. He, P. Yu, M. X. Chu, R. Di. Progress on the effect of FecB mutation on BMPR1B activity and BMP/SMAD pathway in sheep. Yi Chuan= Hereditas. 45 (2023):295-305 .

[3] L. Xie, X. Miao, O. Luo, H. Zhao, X.W. Qin. Impact of FecB Mutation on Ovarian DNA Methylome in Small-Tail Han Sheep. Genes, 14(2023):203-203.

ICB12-1402-1117

Deciphering transcriptomic signatures in myotonic dystrophy: a bioinformatics case-control study

F. Farzi [a], A. Fatemeh farzi [a], C. Mohammad Mostafa Pourseif [b]*
a Department of Medicinal Biotechnology, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: Myotonic dystrophy (DM), characterized by its intricate multi-system involvement and diverse subtypes, poses significant challenges in therapeutic strategies and interdisciplinary research. Among the pivotal genes associated with myotonia dystrophy, the MEF2 gene assumes a central role [1]. This study employs bioinformatics analysis to explore a microarray dataset (GSE13608; Affymetrix Human Genome U133 Plus 2.0 Array) sourced from the NCBI's GEO database [2]. Focusing on skeletal muscle biopsies from DM1 (10 samples) and DM2 patients (20 samples), the analysis compares these profiles with those of normal individuals (n = 6). Transcriptome profiling serves as a pivotal tool in deciphering the molecular intricacies underlying neuromuscular disorders, offering insights into the regulatory mechanisms that underlie the complexities of myotonic dystrophy.

The analysis pipeline involved loading the data using GEOquery, quality control, log2 transformation, and generating descriptive plots (e.g., box-and-whisker plots, expression value distribution plots, and mean-variance trends) using R programming [3]. Notably, our focus centered on the dysregulation of MEF2 and MEF2-related genes. Out of 36 samples, seven genes displayed significant (adj.P.val < 0.01 and |log2 FC| > 1) dysregulation including, COL4A3, nebulin, MTUS1, COPE, TTN, AMPD1, and GRAPL, proposing as significant biomarkers. In the context of myotonic dystrophy, our analysis reveals distinct expression patterns, illustrated through box-and-whisker plots showcasing variability in gene expression across the case and control conditions. The GeneCards, Reactome and KEGG pathway databases showed that these genes exert regulatory control over various crucial biological processes and pathways, including Tibial muscular dystrophy, Metabolic pathways, and ECM-receptor interaction, extracellular matrix organization, collagen formation, signal transduction, striated muscle contraction pathway, and collagen biosynthesis. Our study not only contributes valuable insights into the molecular signatures associated with myotonic dystrophy but also suggests potential therapeutic interventions targeting these dysregulated pathways.

Keywords: Bioinformatics analysis; Microarray; Myotonic dystrophy; Transcriptome profiling.

References

[1] Piasecka, A., et al., MEF2C shapes the microtranscriptome during differentiation of skeletal muscles. Sci Rep, 2021. 11(1): p. 3476.

[2] Bachinski, L.L., et al., Altered MEF2 isoforms in myotonic dystrophy and other neuromuscular disorders. Muscle Nerve, 2010. 42(6): p. 856-63.

[3] Rahbar Saadat, Y., et al., Modulatory Role of Vaginal-Isolated Lactococcus lactis on the Expression of miR-21, miR-200b, and TLR-4 in CAOV-4 Cells and In Silico Revalidation. Probiotics Antimicrob Proteins, 2020. 12(3): p. 1083-1096.

ICB12-1402-1119

The impact of adding a functional group to the antimalarial drug Artemisinin on its penetration into biological membranes using molecular dynamics simulation

Samaneh Boroomand [a], Delara Mohammad-Aghaie [a*], Fatemeh Mohammadpour [b*]
a Department of Chemistry, Shiraz University of Technology, 71555-313, Shiraz, Iran
b Department of Physics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran
*Corresponding authors's email addresses: d_aghaie@sutech.ac.ir (D.M.A)
*Corresponding author's email address: f.mohammadpour@cfu.ac.ir

Abstract: Artemisinin is one of the most important natural products derived from plants, which has attracted the attention of international communities to the extent that it has been dubbed the "Hot" drug [1]. This compound is extracted from the sweet wormwood plant (Artemisia annua) and has a chemical formula of C15H22O5 [2]. In addition to its effectiveness in treating malaria, this compound also has anti-inflammatory and Coronavirus effects [3]. However, due to its disadvantages such as low solubility in water, which is directly related to decreased permeability in membranes, its use is limited. Therefore, due to the importance of this compound in a wide range of diseases, researchers have been seeking to synthesize and develop other derivatives of Artemisinin. One of the most important synthesized derivatives of this compound is Artemisone, which has an amino functional group at carbon C10 compared to Artemisinin. This study aims to investigate the effect of the functional group of Artemisone on the permeability of these compounds in biological membranes using molecular dynamics simulations. Based on the results obtained from simulations such as MSD, density, and Area per lipid, it was observed that the addition of an amino functional group to the Artemisinin molecule improves the penetration of the molecule into the studied biological membrane compared to the Artemisinin molecule.

Keywords: olecular dynamic simulation; Artemisinin; Artemisone; Biological membrane.

References

[1] Li. Z. Zhao, T. Li, J., Yu, Q., Feng, Y. Xie, Y. and Sun, P. Nanomedicine Based on Natural Products: Improving Clinical Application Potential. J. Nanomater. (2022).

[2] Lal. M, Chandraker, S.K. and Shukla, R. Antimicrobial properties of selected plants used in traditional Chinese medicine. In Functional and Preservative Properties of 119-143, (2020) Academic Press.

[3] Orege, J.I. Adeyemi, S.B. Tiamiyu, B.B. Akinyemi, T.O. Ibrahim, Y.A. and Orege, O.B. Artemisia and Artemisia-based products for COVID-19 management: current state and future perspective. Advances in Traditional Medicine, 23(1), (2023).85-96.

ICB12-1402-1120

Comparison of the toxicity of Doxorubicin in water and natural deep eutectic solvents based on choline chloride by molecular dynamics simulation

Samaneh Boroomand [a], Fatemeh Mohammadpour [b]*, Delara Mohammad-Aghaie [a]*
a Department of Chemistry, Shiraz University of Technology, 71555-313, Shiraz, Iran
b Department of Physics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran
*Corresponding author's email addresses: f.mohammadpour@cfu.ac.ir; d_aghaie@sutech.ac.ir (D.M.A)

Abstract: Doxorubicin (DOX) is a widely used anticancer drug for treating various types of cancer like bladder, stomach, breast, thyroid, ovaries, etc. However, it is associated with significant side effects, particularly dose-dependent cardiotoxicity [1]. In recent times, natural deep eutectic solvents (NDES) have emerged as a viable alternative to minimize toxicity and enhance drug delivery efficiency and pharmaceutical activity of drugs [2]. Therefore, the objective of this study is to investigate the impact of choline chloride-based natural deep eutectic solvents on the solubility of DOX drugs using molecular dynamics simulation. The simulation results, including energy, hydrogen bond interactions, and aggregation of the drug, indicate that these solvents can potentially reduce the toxicity of DOX drug compared to water solvent.

Keywords: Molecular dynamic simulation; Doxorubicin; Toxicity; Natural deep eutectic solvents.

References

[1] Kopsida, M., Clavero, A.L., Khaled, J., Balgoma, D., Luna-Marco, C., Chowdhury, A., Nyman, S.S., Rorsman, F., Barbier, C.E., Bergsten, P. and Lennernäs, H. Inhibiting the endoplasmic reticulum stress response enhances the effect of doxorubicin by altering the lipid metabolism of liver cancer cells. Mol. Metab., (2024).101846.

[2] Pradeepkumar, P., Rajan, M., Almoallim, H.S. and Alharbi, S.A. Targeted Delivery of Doxorubicin in HeLa Cells Using Self-Assembled Polymeric Nanocarriers Guided by Deep Eutectic Solvents. Chemistry Select, (2021), 7232-7241.

ICB12-1402-1122

Drug design for the treatment of autoimmune diseases by Increasing the binding affinity of IL-21R Protein to IL-21.

Sara Yazifeshenas [a], Faeze Hajihasan [b]
a Department of Tehran university/Faculty of Biological Sciences Engineering
b Department of Tehran university/Faculty of Biological Sciences Engineering
*Corresponding author's email addresses: sarah.vazifeshenas75@gmail.com; faezehibb2002@gmail.com

Abstract: Increased production of IL21 is a characteristic of some autoimmune diseases, including systemic lupus erythematosus and rheumatoid arthritis, which contributes to the production of autoantibody as well as the pathological features of autoimmune disease. Therefore, in recent researches, antibodies have been designed with a single specificity (only to bind to interleukin 21) to inhibit this interleukin. Since the use of monoclonal antibodies as large-sized proteins has a high production cost and side effects that limit their use, smaller peptides can be designed and selected that contain only the binding region of the receptor to the ligand and groups. The sides of this region have the direction of complete binding of the receptor to the ligand. To reduce the cost of production, the economic bacteria E. coli can be used to express the designed peptide. Therefore, in order to limit the binding of interleukin 21 with its natural receptor on the cell surface, by creating a mutation in the extracellular part of the interleukin 21 receptor (IL-21R), a small peptide is designed that has an increased tendency to bind to interleukin 21. The binding of the designed peptide to interleukin 21 will prevent its binding to the IL-21R receptor on the cell surface and subsequent signals.

Keywords: Interleukin 21; autoimmune disease; IL-21R receptor; E. coli bacteria

ICB12-1402-1123

Predicting the pathogenicity of a missense mutation (Glu88Lys) in the *NCAPG* gene involved in childhood hepatocellular carcinoma using bioinformatics tools

Negar Karami, Mehri Khatami*, Mohammad Hosein Darvand Araghi
Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: m.khatami@yazd.ac.ir

Abstract: Pediatric hepatocellular carcinoma (HCC) is a rare and malignant liver tumor in children with a poor prognosis. This development is mainly observed in children over 10 years old, whether they have a cirrhotic or non-cirrhotic background. NCAPG, a protein associated with chromosomal condensation during mitosis, ensures the correct separation of sister chromatids during chromosome condensation and fusion. It also plays a role in condensing and stabilizing chromosomes during both meiosis and mitosis. NCAPG has been demonstrated to be highly expressed in various cancers, and its related molecular mechanism impacts tumor cell proliferation, invasion, metastasis, and apoptosis such as hepatocellular carcinoma, gastric cancer, colorectal cancer, ovarian cancer, lung adenocarcinoma, breast cancer and prostate cancer. Moreover, the sensitivity of tumor cells can be decreased by NCAPG to reduce the reaction to the original chemotherapy, leading to drug-resistant tumor cells. To sum up, NCAPG is capable of being a new diagnosis and treatment target for several cancers and is also a very promising prognostic marker. Our objective is to assess the pathogenesis of a missense mutation in NCAPG. As a result of this mutation, the glutamic acid at position 88 turns into lysine amino acid.
Several bioinformatics tools, including polyphen-2, I-Mutant 2.0, and SIFT; have been utilized to analyze this assessment. Polyphen-2 analysis confirmed the possible harm caused by this mutation, with a score of 0.917 in the HumDiv model. This missense mutation has been proven by I-Mutant to decrease the stability of NCAPG. The SIFT scored 0.00 for the protein function affected by the mutation at position 88 NCAPG.
According to this analysis, the substitution of glutamic acid with lysine in position 88 may result in pathological effects for the NCAPG protein. The score of pathogenicity of this SNP is very high, so it's important to prove this theory through experimental studies.

Keywords: NCAPG gene; Pediatric Hepatocellular Carcinoma (HCC); Prediction analysis; Missense mutation; Condensin I complex.

References

[1]X. Cai, J. Gao, C. Shi, W. Guo, D. Guo, S. Zhang "The role of NCAPG in various of tumors," Biomedicine & Pharmacotherapy. 2022. Volume 155 (11): 113635.
[2]Y. Shi, C. Ge, D. Fang, "NCAPG facilitates colorectal cancer cell proliferation, migration, invasion and epithelial–mesenchymal transition by activating the Wnt/β-catenin signaling pathway," Cancer Cell International. 2022 Mar 15;22(1):119.
[3] H. Sun, H. Zhang, Y. Yan, Y. Li, G. Che, C. Zhou, C. Nicot, H. Ma, "NCAPG promotes the oncogenesis and progression of non-small cell lung cancer cells through upregulating LGALS1 expression, " Molecular cancer. 2022 Feb 18; 21(1):55.
[4] Z. Yuan, L. Ge, P. Su, Y. Gu, W. Chen, X. Cao, S. Wang, X. Lv, T. Getachew, JM. Mwacharo, A. Haile, W. Sun, "NCAPG Regulates Myogenesis in Sheep, and SNPs Located in Its Putative Promoter Region Are Associated with Growth and Development Traits, " Animals. 2023 Oct 11; 13(20):3173.

ICB12-1402-1124

Bioinformatics Analysis of Microarray Data to Identify Hub Genes as Novel Biomarkers for Colorectal Cancer

Anahita Soleimani[a], Saba Hadi[a], Seyed Hossein Khoshraftar[a], Parisa Alirezaie[a], Mohammad Mostafa Pourseif[b]*, Akbar Amirfiroozi[a], Mahmoud Shekari Khaniani[a]*

a Department of Medical Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran

b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran

*Corresponding author's email address: mmpourseif@gmail.com; mahmoud.khaniani@gmail.com

Abstract: Colorectal cancer (CRC) is the third most common cancer and the second cause of cancer-related death worldwide [1,2]. As early detection is associated with lower mortality, novel biomarkers are urgently needed for timely diagnosis. This study aimed to identify hub genes as CRC biomarkers by analyzing differentially expressed genes (DEGs).

Methods: The GSE113513 dataset was obtained from the GEO database, consisting of 14 pairs of CRC primary lesions and non-cancerous surrounding tissue. Principal Component Analysis (PCA) assessed the sample consistency. LIMMA package was used for statistical analysis. DEGs were identified using $|log2FC|> 3$ and adj P-value < 0.05. Gene expression and survival outcomes correlation were determined by the Kaplan-Meier plotter.

Results: The expression profiles of 28 samples were evaluated using microarray analysis. There were 110 DEGs found in all, with 93 being down-regulated and 17 being up-regulated. Among these DEGs, two genes, MMP7 and CDH3, have been previously linked to various cancers such as breast, lung, pancreatic, etc. In this dataset, these genes exhibited a significant increase in their expression. As per the information available in GeneCards, WikiPathways, and Reactome pathway databases, these genes regulate crucial pathways, including the Wnt signaling pathway, pluripotency, cell-cell communication, and ILK signaling. The DEGs identified in this study were found to be correlated with poor prognosis in colorectal cancer based on the Kaplan-Meier plots.

Conclusion: The findings of our investigation indicate that the genes MMP7 and CDH3 have significant contributions to colorectal cancer (CRC) and are connected to a poor prognosis in CRC.

Keywords: Colorectal Cancer; Biomarker; Microarray.

References

[1] Sung, H.F., J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F, Global Cancer Statistics 2020: GLOBOCAN

[2] Islami F, Goding Sauer A, Miller KD, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. CA Cancer J Clin. 2018;68(1): 31-54. doi:10.3322/caac.21440.

ICB12-1402-1125

Identification of potential inhibitors of kidney glutaminase using molecular docking simulations

Mohammad Hosein Darvand Araghi[1], Reyhane Chamani[*], Negar Karami
1 Department of biology, Yazd University, Yazd, Iran
*Corresponding author's email address: creyhane@yahoo.com

Abstract: Kidney-type mitochondrial glutaminase (GLS) is a phosphate-activated amidohydrolase primarily active in the brain and kidney. It plays a crucial role in energy generation, neurotransmitter production, and kidney acid-base balance by breaking down glutamine. Alternative splicing generates multiple transcript variants. Targeting this protein may hold promise for suppressing and treating cancer, given its involvement in metabolic diseases. [1]. As a result, we have sought to explore the potential of inhibiting this protein as a new avenue for controlling metabolic diseases. Previous studies have shown that the molecule (PubChem ID:71577426) was able to inhibit GLS [2]. Using this molecule as a template, a molecular library from SWISS-SIMILARITY database were created and a set of 57 molecules were identified. Subsequently, PyRex software were used in an in-silico environment to conduct docking simulations of GLS with the molecules in the library [3]. The resulting data revealed that among this library, three molecules (CIDs: 3675176, 3679991, 3680003) interacted with GLS with a minimum ΔG less than -12.1. In addition interacting residues were identified using Discovery studio software as follows: molecule (3680003) with residues chain C: GLU170, ASN182, HIS188, LEU179, LYS178, PHE180, LEU181, PHE176,chain D: LEU172, ARG173, chain B:LYS184, PHE186, TYR258, , molecule (3675176) with residues chain B: LEU187,HIS194, LEU185, LYS184, TYR258, ASN188, GLU189, chain C: ARG175, LYS178, TYR252, PHE180, LEU181, ASN182, GLU183, LEU179, PHE176, chain D: ARG173, and molecule (3679991) with residues chain C: HIS188, LEU181, PHE176, LEU179, TYR252, ARG175, ASN182, GLU183, chain B: TYR258, ASN188, LEU185, LYS184, chain D: ARG173. In conclusion, it seems that these molecules will be able to inhibit GLS and potentially control metabolic diseases, especially cancer. For the future studies, it is recommended that the toxicity and efficacy of these potential inhibitors be evaluated *in vitro* and *in vivo*.

Keywords: Glutaminase; Pyrex; GLS inhibitor; docking.

References

[1] L. Lin, H. Chang, T. Lin, J. Lin, SH. Hsia, K. Hsu, T. Huang" Structure-based virtual screening discovers novel kidney-type glutaminase inhibitors" Biomedicine & Pharmacotherapy, Volume 154,2022, 113585,
[2] M.M. Robinson, S.J. McBryant, T. Tsukamoto, C. Rojas, D.V. Ferraris, S. K. Hamilton, J.C. Hansen, N.P. Curthoys, Novel mechanism of inhibition of rat kidney-type glutaminase by bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl) ethyl sulfide (BPTES), Biochem J. 406 (3) (2007) 407–414.
[3] O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, Journal of Computational Chemistry 31 (2010)455-461

ICB12-1402-1126

In-silico Comparison of Interaction Some Bee Venom Peptides with the Mouse mMHC-MOG:37-46 complex as a Potential Treatment of MS Disease

K.Mahnam *[a], S.F.Razavi [b], F.Pouryousef [c], M.Shakhsi-Niai [d]
a,d Department of Biology, Faculty of Basic Sciences, Shahrekord University, Shahrekord, Iran
b & c Department of Biology, Faculty of Basic Sciences, Azarbaijan Shahid Madani University, Tabriz, Iran

*Corresponding author's email address: mahnam.karim@sci.sku.ac.ir

Abstract: Multiple Sclerosis disease is an autoimmune-mediated disorder, which affects the central nervous system (CNS) and destroys the myelin sheath, causing a variety of neurological dysfunction [1].

Myelin oligodendrocyte glycoprotein (MOG) is an encephalitogenic protein that can induce a demyelinating immune response in several experimental models of inflammatory demyelinating diseases [2]. The previous results revealed that MOG35-55 is an immunodominant epitope for T-cell responses in MS disease [3].

Bee venom peptides have become more prominent in recent years due to their medicinal properties and therapeutic aspects [4,5]. This research investigates the interaction of bee venom peptides with the MHC-MOG:37-46 complex in mouse (allele H-2 kb) to prevent the interaction of TCR with this complex which can lead to the treatment of MS disease in the mouse. At first MOG:37-46 fragment was docked to the MHC binding cleft and then 100 ns MDS was done on the complex. Then, the three-dimensional structure of eleven bee venom peptides was obtained from the pep-fold server, and after 100 ns MDS of them, 20 structures were obtained from the last 20 ns of MDS, and these structures were docked to the MOG:37-46-MHC complex. Then the best structure that was obtained from docking for each peptide was subjected to 100 ns MDS and MM/PBSA binding free energy of peptides was calculated. The binding free energy results and other parameters showed that peptide melittin (GIGAVLKVLTTGLPALISWIKRKRQQ) had better interaction with this complex and can be considered as a potential drug in MS disease in mouse.

Keywords: Bee venom peptides; MS disease; Molecular dynamics simulation (MDS), MOG:37-46.

References

[1] P.M. Stuart, "Major histocompaΘbility complex (MHC): mouse." Els., 14(2015) 3105–3117.

[2] O. Corbali, T. Chitnis, "Pathophysiology of myelin oligodendrocyte glycoprotein anΘbody disease". FronΘers in Neurology.,

14(2023), 1137998.

[3] W. Ambrosius, S. Michalak, W. Kozubski, A. Kalinowska, A."Myelin oligodendrocyte glycoprotein anΘbody-associated disease:

current insights into the disease pathophysiology,diagnosis and management". InternaΘonal journal of molecular sciences, 22(2020).

[4] R. Wehbe, Rim, et al. "Bee venom: Overview of main compounds and bioacΘviΘes for therapeuΘc interests." Molecules 24(2019).

ICB12-1402-1128

Unraveling Protein Dynamics in the Presence of Cellulose Nanocrystals: An Essential Dynamics Analysis

Hamed Emami [a], Yasaman Mahmoodi, Masoud Shahriaria , Faramarz Mehrnejad*
a. Department of Life Sciences Engineering, Faculty of New Sciences & Technologies, University of Tehran, 14395-1561, Tehran, Iran.
*Corresponding author's email address: Mehrnejad@ut.ac.ir

Abstract: The dynamic behavior of proteins is crucial for their function, stability, and interactions. Among the various methods for analyzing protein dynamics, essential dynamics (ED) stands out by identifying the most relevant motions, characterized by their large amplitude, that contribute significantly to the overall dynamics of the system. By applying principal component analysis (PCA) to the covariance matrix of atomic positions, ED can effectively extract these dominant modes of fluctuation, correlation, and conformational change [1]. Herein, molecular dynamics (MD) simulations were applied to investigate the essential dynamics of insulin in the presence of cellulose nanocrystals (CNC). A 1000-ns (1-µs) MD simulation was conducted using the GROMACS package [2]. To reduce the dimensionality of peptide dynamics, PCA was applied to the generated trajectories. The covariance matrix was constructed from atomic coordinates, and an eigenvalue decomposition was performed using the Gmx covar and Gmx anaeig tools to obtain orthogonal collective modes. The results were interpreted by projecting displacement vectors onto PCA modes. The first 25 PCs accounted for, respectively, 55.1%, 34.5%, and 29.73% of the total movements observed during the 1-µs simulation. Our ED analysis revealed that the presence of CNC significantly influenced the peptide's dynamics. The collective modes identified through PCA revealed alterations in their amplitude and direction, suggesting that CNC-induced interactions modulated the peptide's conformational changes and fluctuation patterns. These findings provide valuable insights into the intricate interplay between insulin and CNC at the molecular level

Keywords: Essential Dynamic; Principal Component Analysis; Covariance Matrix; Peptide.

References

[1] T. Meyer et al., "Essential Dynamics: A Tool for Efficient Trajectory Compression and Management," J Chem Theory Comput, vol. 2, no. 2, pp. 251–258, 2006, doi: 10.1021/ct050285b.

[2] P. Bauer, B. Hess, and E. Lindahl, "GROMACS 2022 Manual." Zenodo, Feb. 2022. doi: 10.5281/zenodo.6103568.

[3] C. C. David and D. J. Jacobs, "Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins," in Protein Dynamics: Methods and Protocols, D. R. Livesay, Ed., Totowa, NJ: Humana Press, 2014, pp. 193–226. doi: 10.1007/978-1- 62703-658-0_11.

ICB12-1402-1129

Designing mutated protein with increased binding affinity to Il-22 in

Fatemeh Askariyeh*
a. Department of Student at Tehran University
*Corresponding author's email address: askariyehfatemeh@gmail.com

Abstract: Interleukins, a class of cytokines, serve as crucial signaling proteins integral to the orchestration of immune responses. Specifically, interleukin-22 (IL-22), a member of the IL-10 cytokine family, holds a pivotal position in immune regulation. Its multifaceted role encompasses both the promotion of tissue repair and participation in inflammatory process. In the context of autoimmune pathologies such as psoriasis and rheumatoid arthritis, IL-22 assumes a pathogenic significance, contributing substantively to the perpetuation of chronic inflammatory cascades.

The present investigation is centered on the deliberate engineering of a mutated protein, denoted as mIL-22R1, characterized by an augmented binding affinity to IL-22 in order to inhibiting IL-22. This strategic modification aims to curtail IL-22 activity, thereby presenting a targeted and prospective therapeutic paradigm for mitigating autoimmune disorders intricately linked to IL-22 dysregulation.

Keywords: Mutated, autoimmune diseases, interleukin, IL-22, Il-22R1

ICB12-1402-1133

Prediction of Pathogenicity of a Single Nucleotide Polymorphism (rs2133785693) in *human CDKN1C* gene Associated with Wilms Tumor

Fatemehalsadat Motamedfar [a]*, Mehri Khatami[a]
a Department of Biology, Yazd University, Yazd, Iran.
*Corresponding author's email address: fmotamedfar@gmail.com

Abstract: Wilms tumor (WT), or nephroblastoma, is the most common pediatric kidney cancer. Although healthy children develop WT, it is more common in children with cancer predisposition disorders like Denys-Drash syndrome, WAGR syndrome, or Beckwith-Wiedemann syndrome (BWS). Patients with BWS are at an approximately 800-fold increased relative risk of developing WT. Beckwith-Wiedemann Syndrome is a rare genetic disorder that causes macroglossia, hemihyperplasia, omphalocele, neonatal hypoglycemia, macrosomia, embryonal tumors (such as Wilms tumor), visceromegaly, adrenocortical cytomegaly, kidney abnormalities, and ear creases / posterior helical ear pits. Pathogenic variants of heterozygous maternal Cyclin-dependent kinase inhibitor 1C (*CDKN1C*) are one of the five molecular mechanisms found in individuals with BWS. This study aims to assess the pathogenicity of a missense mutation in *CDKN1C* (rs2133785693). This mutation changes the tryptophan at position 79 to cysteine (W79C). This assessment has been analyzed using several bioinformatics tools including polyphen-2, I-Mutant 2.0, and SIFT. Results showed that rs2133785693 affected the protein function with a score of 0.0 by the SIFT. Polyphen-2 web-based tool analysis indicated that this mutation was likely deleterious and is predicted to be probably damaging with PSIC index score of 1.00. Additionally, I-Mutant suggested that this missense mutation will decrease the stability of the CDKN1C protein. In conclusion, W79C might have pathogenic effects on the CDKN1C protein. Further experimental research is required to validate this hypothesis.

Keywords: CDKN1C gene; Beckwith-Wiedemann syndrome(BWS); Wilms tumor(WT); rs2133785693; SNP.

References

[1] Nirgude, Snehal, et al. "Cancer predisposition signaling in Beckwith-Wiedemann Syndrome drives Wilms tumor development." British Journal of Cancer (2023): 1-13.

[2] Quarello, Paola, et al. "Implications of an underlying Beckwith–Wiedemann Syndrome for Wilms Tumor treatment strategies." Cancers 15.4 (2023): 1292.

[3] Ali, Ashiq, et al. "Therapeutic potential of natural antisense transcripts and various mechanisms involved for clinical applications and disease prevention." RNA biology 21.1 (2024): 1-18.

ICB12-1402-1134

In Silico Design of a Novel Antibody-Drug Conjugate Targeting Mesothelin in Pancreatic Cancer

F. Soltan Mohammadi[1], M. Ghalam-Zan [2,*]
1 Department of Science, Shahed university, Tehran, Iran
2 Department of Science, Tehran university, Tehran, Iran
*Corresponding author's email address: mahdieh.ghalamzan@ut.ac.ir

Abstract: Antibody-drug conjugates (ADCs) represent a promising new class of therapeutics that combine the specificity of immunotherapy with the potency of chemotherapy, offering a targeted approach to cancer treatment, including pancreatic cancer[1], [2]. This research focuses on the design of an ADC that targets Mesothelin (MSLN), a surface antigen overexpressed in pancreatic tumor cells. In this study, we aimed to design a mesothelin directed antibody conjugate composed of a humanized single-chain variable fragment (scFv) and GrB using precise in silico techniques.
The challenge lies in the phenomenon of 'shedding', where proteolytic enzymes cleave various amino acid sequences, preventing the ADC from binding to the mesothelin on the cancer cell surface[3]. To overcome this, we utilized the findings of Liu, X et al., to retrieve the sequence of a humanized single-chain variable fragment (scFv), known as 15B6 scFv, which recognizes the residual mesothelin post-shedding[4].
We then conjugated 15B6 to granzyme B (GrB) to design our ADC, employing precise in silico methodologies[5]. Four distinct linker peptides were used for the conjugation of 15B6 to GrB, and the 3D structures of these antibody conjugates were predicted using I-TASSER.
The conjugate whose linker peptide least affected the structural conformation of 15B6 and GrB was chosen. We also compared the solubility and melting temperature of the selected conjugate with those of 15B6 and GrB, and evaluated its physicochemical properties and flexibility.
Finally, we compared the binding capacity and the dissociation constant (Kd) of the selected conjugate to MSLN with those of 15B6, and identified the residues contributing to antigen binding using LigPlot+ software.
However in vivo and in vitro investigation is required to determine the stability, feasibility and effectiveness of such construct, in silico techniques, such as those employed in this study, could be utilized for the early development of immune based therapeutics.
Keywords: Antibody-Drug Conjugate; Mesothelin; Pancreatic Cancer; In Silico Design; Granzyme B

References

[1]improved delivery of a potent payload by counteracting MUC16/CA125 inhibitory effects.," PLoS One, vol. 18, no. 5, p. e0285161, 2023, doi: 10.1371/journal.pone.0285161.

[2] D. Aggarwal et al., "Antibody-drug conjugates: the paradigm shifts in the targeted cancer therapy," Frontiers in Immunology , vol. 14. 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fimmu.2023.1203073

ICB12-1402-1136

Morbillivirus Cellular Receptor Interaction: Introduce Potential Important Regions Involved in Novel Cellular Receptor

Mkhosravi [a]* F. mohammadzadeh[b]
a Department of Pathobiology, Faculty of Veterinary Medicine, Amol University of Special of Modern Technologies, Amol, Iran
b BSc Student in Veterinary Laboratory Sciences, Department of Pathobiology, Faculty of Veterinary Medicine, Amol University of Special of Modern Technologies, Amol, Iran
*Corresponding author's email address: mojtaba.kh.dvm@gmail.com

Abstract: Morbilliviruses are single-stranded RNA viruses that belong to the Paramyxoviridae family, including several deadly human and animal infecting viruses. The first step of virus-cell entry has been promoting through the interaction between hemagglutinin protein (H) and cellular receptors via its head domain. Up to now, signaling lymphocyte activation molecule (SLAM) and Nectin-4 are recognized as cellular receptors. Morbilliviruses could develop CNS disorders and no receptor has been introduced for nerves cells "[1]". The prior findings revealed that relatively same region of the head domain is associated with cellular receptor interaction. This study is designed to find potential region of head domain involving to novel cellular receptor interaction. For this reason, alignment analysis at amino acid level was conducted among several reference strains of morbillivirus. Our result showed that five nearly conserved regions designed as A,B,C,E and D. The E region (512 to 543 aa) is associated with the both SLAM and Nectin-4 receptors "[2]". Furthermore, 3D structure measles H protein and human SLAM revealed that regions A and D are situated upon the protein's surface. Additionally, these two regions contain several highly conserved residues and also offer a relatively suitable surface for interacting with the novel receptor. By considering with the pattern exhibited by the H protein in relation to previously recognized receptors, it can be concluded that H protein may employ a comparable pattern or alternatively share distinct region to bind the new receptor. Further biochemical studies at the in vitro level are highly recommended to substantiate this matter.

Keywords: Morbillivirus; H portion; slam; nectin-4; novel receptor.

References

[1] T. Hashiguchi, T.Ose, M . Kubota, N. Maita, , J, Kamishikiryo, K. Maenaka, , Y.Yanagi, "Structure of the measles virus hemagglutinin bound to its cellular receptor SLAM" Nat. Struct. Mol. Biol.,18 (2011) 135–141.
[2] F.Xu, Sh.Tanaka, ,H. Watanabe , Y. Shimane,M.Iwasawa, K.Ohishi ,T. Maruyama , " Computational Analysis of the Interaction Energiesbetween Amino Acid Residues of the Measles Virus Hemagglutinin and Its Receptors" Viruses., 10(2018) 1-18.

ICB12-1402-1137

Multi-Omics Integration for Pancreatic Adenocarcinomas Subtyping

F. Dorri-Najafabadi[a], M. Emadi-Baygi[a]*, M.lotfi-Shahreza[b]
a Department of Genetics, Faculty of Basic Sciences, Shahrekord University, Shahrekord, Iran
b Department of Computer Engineering, Shahreza Campus, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: emadi-m@sci.sku.ac.ir

Abstract: Cancer is a leading cause of death worldwide. Precision oncology aims to identify new molecule-based cancer subtypes from large-scale cancer multi-omics data(1), allowing for more accurate and personalized treatments. Multi-omics studies analyze high-dimensional datasets at various levels to reveal the complexity of cells and their environment(2). Integrating multi-omics data has become increasingly important, with machine learning playing a key role in comparing and identifying patterns in biological data(3). Our study utilized multi-omics data, including transcriptomics RNA-sequencing, DNA methylation, and gene mutations, to identify three molecular subtypes and assess sample similarity within the subtypes. We applied various pre-processing steps, including annotation, quality control, filtering, normalization, and feature selection. Then, we executed ten classical clustering algorithms to recognize patients with different molecular features using the "MOVICS" package in R. We filtered out low express genes, noncoding genes, and removed probes with detection P value > 0.01, all non-CpG probes, all SNP-related probes, all multi-hit probes, and probes located on sex chromosomes. Finally, we identified three molecular subtypes and quantified the sample similarity within the subtypes using the silhouette score. Our study highlights the importance of multi-omics integration and pre-processing steps in understanding molecular subtypes. The use of the "MOVICS" package in R provides an accessible and powerful toolset for researchers to analyze multi-omics data. Integrating multi-omics and clinical data can help identify robust and clinically actionable cancer subtypes. We hope that our findings will contribute to the development of more effective cancer therapies and personalized medicine.

Keywords: Multi-Omics; Cancer Subtyping; Pancreatic Adenocarcinoma; MOVICS.

References

[1] D. Wang, J. Gu, "Integrative clustering methods of multi-omics data for molecule-based cancer classifications," Quantitative Biology. Quant Biol , 4 (2016) 58-67.
[2] YJ. Heo, C. Hwa, GH. Lee, JM. Park, JY. An, "Integrative Multi-Omics Approaches in Cancer Research: From Biological Networks to Clinical Subtypes," Molecules and Cells. Mol Cells., 44(2021) 433–443.
[3] S. Graw, K. Chappell, CL. Washam, A. Gies, J. Bird, MS 2nd. Robeson, SD. Byrum, "Multi-omics data integration considerations and study design for biological systems and disease," Molecular Omics. Mol Omics., 17 (2021) 170-185.

ICB12-1402-1140

Predicting Functional miRNA Targets in C-Fos with High Accuracy via Machine Learning

Ali Malmir[1], Maryam Hassanlou[2*], Mahsa Mohammad Amoli[3], Majid Sadeghizadeh[1*]
1 Department of Molecular Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran
2 Farzanegan campus, semnan University, Semnan, Iran
3 Endocrinology and Metabolism Research Institute (EMRI), Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran
*Corresponding authors' email addresses: M.hassanlou@semnan.ac.ir; sadeghma@modares.ac.ir

Abstract: Introduction: The proto-oncogene C-Fos plays a crucial role in cell proliferation and differentiation[1], yet its intricate control network, particularly involving microRNAs (miRNAs), remains largely uncharted. This study delves into the possibility of C-Fos harboring encoded miRNA regulatory information, employing machine learning to decipher its miRNA regulatory landscape.

Materials and Methods: We compiled diverse datasets encompassing C-Fos expression, miRNA expression, and predicted miRNA target sites within the C-Fos gene. Utilizing feature engineering techniques, we extracted sequence and positional features surrounding potential target sites. Subsequently, we implemented various machine learning algorithms, including random forests and support vector machines, to classify true miRNA target sites based on these features. We evaluated model performance using cross-validation and compared results across algorithms.

Results: Our models achieved high accuracy in identifying true miRNA target sites within the C-Fos gene, exceeding 80% in most cases. By analyzing the learned feature weights, we uncovered key sequence and positional motifs crucial for miRNA recognition and binding. Interestingly, we identified clusters of enriched motifs within the C-Fos gene, suggesting the presence of miRNA regulatory hotspots. Furthermore, we mapped these hotspots to specific functional domains within C-Fos, revealing a potential link between miRNA regulation and C-Fos function.

Conclusions: Our study demonstrates the effectiveness of machine learning in decoding miRNA regulatory networks encoded within genes. By applying this approach to C-Fos, we unraveled a complex miRNA regulatory landscape with hotspots likely linked to specific C-Fos functions. This understanding paves the way for further investigations into miRNA-mediated control of C-Fos in diverse cellular processes, including cancer development[2]. Future research will focus on experimentally validating predicted target sites and elucidating the functional consequences of this newly discovered miRNA regulatory network in various biological contexts.

Keywords: C-Fos; microRNA; miRNA regulatory network; machine learning; target site prediction; cell proliferation.

References

[1] *osteosarcoma cells accompanied by altered expression of Wnt2 and Fzd9.* PLoS One, 2017. **12**(6): p. e0180558.

[2]Arivazhagan, L., G. Venkatraman, and S.K. Rayala, *Increased expression of MicroRNA 551a by c-Fos reduces focal adhesion kinase levels and blocks tumorigenesis.* Molecular and Cellular Biology, 2019.

ICB12-1402-1141

Deciphering the c-MYC miRNA Regulatory Network through Machine Learning

Ali Malmir[1], Maryam Hassanlou[2*], Mahsa Mohammad Amoli[3], Majid Sadeghizadeh[1*]
1 Department of Molecular Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran
2 Farzanegan campus, semnan University, Semnan, Iran
3 Endocrinology and Metabolism Research Institute (EMRI), Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran
*Corresponding author's email addresses: M.hassanlou@semnan.ac.ir; sadeghma@modares.ac.ir

Abstract: Introduction: The oncogene c-MYC plays a pivotal role in cancer development and progression[1], yet its intricate control network, particularly involving microRNAs (miRNAs), remains elusive[2]. This study delves into c-MYC's potential to encode miRNA regulatory information using machine learning, aiming to decipher its miRNA regulatory landscape.

Materials and Methods: We compiled diverse datasets encompassing c-MYC expression, miRNA expression, and predicted miRNA target sites. Using feature engineering techniques, we extracted sequence and positional features surrounding potential miRNA target sites within the c-MYC gene. Subsequently, we employed different machine learning algorithms, including random forests and support vector machines, to classify true miRNA target sites based on these features. We evaluated model performance using cross-validation and compared results across algorithms.

Results: Our models achieved high accuracy in identifying true miRNA target sites within the c-MYC gene, exceeding 85% in most cases. By analyzing the learned feature weights, we uncovered key sequence and positional motifs crucial for miRNA recognition and binding. Interestingly, we identified clusters of enriched motifs within the c-MYC gene, suggesting the presence of miRNA regulatory hotspots. Furthermore, we mapped these hotspots to specific functional domains within c-MYC, revealing a potential link between miRNA regulation and c-MYC function.

Conclusions: Our study demonstrates the feasibility of using machine learning to decode miRNA regulatory networks encoded within genes. By applying this approach to c-MYC, we unveiled a complex miRNA regulatory landscape with hotspots likely linked to specific c-MYC functions. This understanding paves the way for further investigations into miRNA-mediated control of c-MYC in cancer. Future research will focus on validating predicted target sites experimentally and elucidating the functional consequences of this newly discovered miRNA regulatory network in cancer biology

Keywords: c-MYC; microRNA; miRNA regulatory network; machine learning; target site prediction; cancer.

References

[1] Hoffman, B., et al., *The proto-oncogene c-myc in hematopoietic development and leukemogenesis.* Oncogene, 2002. **21**(21): p. 3414-3421.
[2] Pauley, K.M. and E.K. Chan, *MicroRNAs and their emerging roles in immunology.* Annals of the New York Academy of Sciences, 2008. **1143**(1): p. 226-239.

ICB12-1402-1142

Drug repositioning and the candidates' mechanism of action for vitiligo

F. Masumeh Jalalvand[a], Rasoul Salehi[b,c], Reza Nedaeinia[c] Leila Abkhooie[a], Sharifeh Khosravi[b] A&C. Sedigheh Momenzadeh[d]*

a Department of Medical Biotechnology, Faculty of Medicine, Lorestan University of Medical Sciences, Khorramabad, Iran

b Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

c Pediatric Inherited Diseases Research Center, Research Institute for Primordial Prevention of Non-Communicable Disease, Isfahan University of Medical Sciences, Isfahan, Iran

d Department of Medical Biotechnology, School of Allied Medicine, Bushehr University of Medical Sciences, Bushehr, Iran

*Corresponding author's email address: fmbiotec@gmail.com

Abstract: Aim: Vitiligo is a complex disease with hypochromic or achromic spots [1]. Topical treatment drugs are not 100% effective and are also associated with side effects [1–3]. Method: To introduce new drug candidates for vitiligo, we have made a systematic search in the GEO Ncbi. Array data were selected and filtered by parameters of lesional and non-lesional samples. Meta-analysis was performed using the ImaGEO online tool [4]. Differentially expressed genes were used as signatures for screening against drugs in the connectivity map database. Finally, the results were filtered based on the highest Z score and MOA classification [5,6]. The highest reverse Z score (z score<-80) was used for target prediction using Swiss target prediction. The drug-target network was constructed, and the target nodes with the highest degree were identified. Findings: Twenty-eight genes were significant after a meta-analysis of the two datasets of GSE75819 and GSE65127 with a total of 38 samples. 1002 compounds were identified with reverse profiles. The compound with the highest z score includes CDK, DNA synthesis, HDAC, IGF-1, JAK, JNK, MEK, mTOR inhibitors, and Glucocorticoid receptor agonists. The highest degree nodes in the constructed target-compound network were enriched in the most significant pathways in cancer, PI3K-Akt signaling pathway, Ras signaling pathway, and EGFR Tyrosine Kinase Inhibitor Resistance.
Conclusion: There are a couple of candidates that have not previously been proposed for vitiligo treatment and as therapy candidates worth examining. Based on the drug target network, some cancer pathways are also involved in vitiligo, which indicates the shared pathways of cancer and vitiligo.
Keywords: Vitiligo; Drug repositioning; Connectivity map.

References

[1]Marchioro HZ, Silva de Castro CC, Fava VM, et al. Update on the pathogenesis of vitiligo. An Bras Dermatol. 2022;97(4):478–90.
[2]Lee JH, Kwon HS, Jung HM, et al. Treatment Outcomes of Topical Calcineurin Inhibitor Therapy for Patients with Vitiligo: A Systematic Review and Meta-analysis. JAMA Dermatology. 2019;155(8):929–38.
[3]Wang X, Mccoy J, Lotti T, et al. Topical cream delivers NB-UVB from sunlight for the treatment of vitiligo. Expert Opin Pharmacother. 2014;15(18):2623–7.
[4]Toro-Domínguez D, Martorell-Marugán J, López-Domínguez R, et al. ImaGEO: integrative gene expression meta-analysis from GEO database. Bioinformatics. 2019;35(5):880–2.
[5]Liu TP, Hsieh YY, Chou CJ, et al. Systematic polypharmacology and drug repurposing via an integrated L1000-based Connectivity Map database mining. R Soc Open Sci. 2018;5(11).
[6]Konuma T, Ogawa K, Okada Y. Integration of genetically regulated gene expression and pharmacological library provides therapeutic drug candidates. Hum Mol Genet. 2021;30(3–4):294–304.

ICB12-1402-1143

Investigation of the internal co-expression relationship evolution in chronic rhinosinusitis with nasal polyps

Nasibeh Khayer [a,*], Maryam Jalessi [a,b]
a Skull Base Research Center, The Five Senses Health Institute, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
b ENT and Head and Neck Research Center and Department, Hazrat Rasoul Hospital, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: khayer.n@iums.ac.ir

Abstract: Chronic rhinosinusitis (CRS) is a widespread disorder affecting nearly a third of adults. It is characterized by persistent inflammation of the nasal and sinus lining cells. CRS exhibits a heterogeneous nature, with two major subtypes distinguished by distinct pathophysiological mechanisms: CRS with nasal polyps (CRSwNP) and CRS without nasal polyps (CRSsNP) [1]. Notably, up to 60% of individuals with CRSwNP also experience concurrent lower airway diseases, such as asthma, which typically onset in adulthood [2]. The main therapeutic options for CRSwNP are currently limited to intranasal glucocorticoid administration, sinus surgery, or a combination of both [3]. Despite the significant morbidity and negative impact on quality of life, our understanding of the underlying molecular mechanisms and specific biomarkers associated with CRSwNP remains incomplete [4]. Therefore, identifying causal genes and dysregulated pathways paves the way for novel therapeutic interventions. The objective of the present study is a comprehensive analysis of CRSwNP-related transcriptome data, aiming to decipher sophisticated molecular relations at an elevated level of complexity compared to the classical pairwise interaction approach. For such purpose, we used a novel computational method, i.e., liquid association (LA), as a powerful tool to capture dynamic co-expression relationships [5]. This approach captured the internal evolution of the co-expression relation between a pair of genes (X, Y) under a change in the expression profile of a third gene (Z), named the switch gene. The results suggested that the *S100a9*, *Nfe2l2*, *Ppl* and *Tgfbr3* genes can be considered potential therapeutic targets in the CRSwNP.

Keywords: Chronic rhinosinusitis; Liquid association tool; Therapeutic targets.

References

[1]Hopkins, C., *Chronic rhinosinusitis with nasal polyps.* New England Journal of Medicine, 2019. **381**(1): p. 55-63.
[2]Ragab, A., P. Clement, and W. Vincken, *Objective assessment of lower airway involvement in chronic rhinosinusitis.* American journal of rhinology, 2004. **18**(1): p. 15-21.
[3]Bachert, C., L. Zhang, and P. Gevaert, *Current and future treatment options for adult chronic rhinosinusitis: focus on nasal polyposis.* Journal of Allergy and Clinical Immunology, 2015. **136**(6): p. 1431-40.
[4]Hastan, D., et al., *Chronic rhinosinusitis in Europe–an underestimated disease. A GA2LEN study.* Allergy, 2011. **66**(9): p. 1216-23.
[5]Ho, Y.Y., et al., *Modeling liquid association.* Biometrics, 2011. **67**(1): p. 133-41.

ICB12-1402-1144

Unlocking Drug Combinations: A Novel Graph-Based Approach for Predicting Anti-Cancer Synergy.

Babak Bahri Aliabadi[a*], Fatemeh Yassaee Meybodi[b], Changiz Eslahchi[a,b,]
a Data and computer sciences, Shahid Beheshti university, Tehran, Iran.
b Institute for Research in Fundamental Sciences(IPM), Tehran, Iran.
*Corresponding author's email address: b_bahrialiabadi@sbu.ac.ir

Abstract: In the quest for effective cancer therapies, exploring combinatorial drug regimens is crucial to harness synergistic interactions and overcome drug resistance. However, identifying synergistic drug pairs faces challenges due to the vast combinatorial space and experimental limitations. This study introduces ClusterSyn, a novel machine learning-powered framework for classifying anti-cancer drug synergy scores. Utilizing drug synergy scores on cancer cell lines, ClusterSyn employs a two-step approach involving drug clustering and synergy score prediction using a fully connected deep neural network. For each cell line in the training dataset, a drug graph is constructed, with nodes representing drugs and edge weights denoting synergy scores between drug pairs. Drugs are clustered using the Markov clustering (MCL) algorithm, aligning similar synergy profiles. Subsequently, vectors representing the similarity of drug pairs to each cluster are input into the deep neural network for synergy score prediction (synergy or antagonism). Comparative analysis with clustering and regression-based methods, including DeepSynergy and DeepDDS, demonstrates the superior performance of ClusterSyn on diverse datasets such as Oniel and Almanac. These results underscore the remarkable potential of ClusterSyn as a versatile tool for predicting anti-cancer drug synergy scores.

Keywords: Drug synergy; Clustering; Prediction; Machine learning.

ICB12-1402-1147

Functional investigation of MYL9 gene in signaling pathways and gene networks in chicken

M. Mohammadnezhad, M. Gholizadeh *
Department of Animal Science, Sari Agricultural Sciences and Natural Resources University, Sari, Iran
*Corresponding author's email address: m.gholizadeh@sanru.ac.ir

Abstract: MYL9 gene is a light chain myosin gene that plays an important role in processes such as contraction, cell motility, and participating in the construction of myofibril proteins of muscle cells in chicken. This study was carried out with the aim of functional investigation of MYL9 gene in signaling pathways and gene networks in chicken. Using data from the COXPRESdb database, the list of 50 high co-expressed genes and gene networks related to MYL9 were extracted. Gene ontology analysis for co-expressed genes was performed using David's bioinformatics tool. Corrected P-value values (Benjamin correction) less than 0.05 were considered as significant values for gene ontology results. The genes TAGLN and CSRP1 showed the highest levels of co-expression with z-scores of 11.8 and 10.0, respectively. The co-expression map (UMAP) showed that the MYL9 gene, along with 20 highly co-expressed genes, forms a Semi-centralized structure in the lower right corner of the map. There were 14 genes directly associated with the MYL9 gene. These genes included DMD, PALLD, PDLIM7, RBPMS, FHL2, PDLIM1, RBPMS2, MYOCD, CSRP1, ACTG2, SMTN, HSPB1, TPM2 and TAGLN. Performing ontology analysis using DAVID software showed that the molecular function of MYL9 gene and 50 genes in the list high co-expressed genes were significantly (P-value < 0.05) related with biological pathways including actin binding, actin strand binding, alpha-actinin binding, myosin binding II, extracellular matrix structural component, muscle structural components, muscle alpha-actinin binding and binding Collagen. It was also found that these genes were also significantly related to the KEGG enrichment pathway of motor proteins.

Keywords: MYL9; COXPRESdb; co-expression; KEGG; signaling pathways.

References

[1] S.N. Albooshoke, M. Tahmoorepur, "Identification of RNA isoforms expressed in skeletal muscle of native and commercial chickens by sequencing method" Iranian Journal of Animal Science., 2 (2018) 179-191.

ICB12-1402-1148

Comprehensive discovery of markers effective in disease severity in the genome of Iranian patients with covid-19

Fariba Esmaeili *, Dariush Salimi
*Corresponding author's email address: esfariba416@gmail.com

Abstract: This article examines the use of artificial intelligence algorithm statistical techniques for automatic identification of biomarkers with the aim of prognosticating the severity of infection in patients with COVID-19. The operation can be applied to other viruses as well. In this research, a group of people (110 people) were selected for the study and they were classified into two groups of patients with high and low severity, and the evolution-based harmony search algorithm was used to extract patterns from the virus genome and also, a multi-objective approach was used to investigate the relationship between these patterns and the severity of the disease. In the next part of the study, the targets of unique miRNAs related to the genome of samples of two groups are extracted and considered as a marker to predict the severity of infection. Also, taking into account the effect of different amino acids on increasing the potential of their interactions, a method for grouping and weighting was proposed, which has improved the search capability of the algorithm. The results of this model were in many cases consistent with previously identified markers determined through experiments. The obtained results show that there is a significant difference in the mutation rate in the candidate regions in patients with high and low severity. Also, the relationship between mutations and the age of the patients was investigated and a significant correlation was observed between the frequency of mutations in the designated areas and the age of the patients.

Keywords: artificial intelligence algorithm; mutations; biomarkers of molecular diagnosis.

ICB12-1402-1149

TWIST1 and Stomach Adenocarcinoma: Deciphering the Molecular Landscape of Metastasis-Related Genes

S. Ebrahimi-Jouzdani[a], H. Rahimi-Tamandegani[b], M. Emadi-Baygi[a]*
a Department of Genetics, Faculty of Basic Sciences, Shahrekord University, Shahrekord, Iran
b Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran
*Corresponding author's email address: emadi-m@sci.sku.ac.ir

Abstract: The TWIST1 gene encodes a basic helix-loop-helix (bHLH) transcription factor crucial for gene expression control[1]. It forms homodimers and heterodimers when binding to DNA E-box sequences and is involved in epithelial-mesenchymal transition (EMT), a key process in metastasis[2, 3]. High TWIST1 expression indicates poor prognosis in metastatic cancer[4]. This study investigated TWIST1 expression levels in stomach adenocarcinoma and their association with invasion and metastasis-related genes. Genes associated with cancer hallmarks were obtained from several databases, and stomach adenocarcinoma transcriptome data was obtained from The Cancer Genome Atlas (TCGA). High-expression and low-expression groups were created based on TWIST1's average expression in normal samples. Differentially expressed genes (DEGs) were analyzed (using Limma, an R package); Furthermore, univariate and multivariate Cox regression analyses were performed. Weighted gene co-expression network analysis (WGCNA) was conducted, and gene enrichment analysis (using R packages named gProfiler2 and enrichR) revealed pathways related to invasion and metastasis. 271 out of 8716 DEGs were up-regulated ($logFC \geq 1$, adjusted p-value ≤ 0.01), and four significant genes with p-value < 0.05 were identified from multivariate Cox analysis. WGCNA revealed "greenyellow" module with a 0.65 correlation rate with TWIST1 trait. The findings suggest a potential connection between TWIST1 and other cancer hallmarks, highlighting its potential as a therapeutic target and prognostic indicator for cancer patients. This study emphasizes the importance of multi-omics data analysis in understanding cancer invasion and metastasis mechanisms.

Keywords: TWIST1; Cancer Hallmarks; Transcriptome Profiling; Differential Expression Analysis; Survival Analysis

References

[1] Qin, Q., et al., Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. Cell Research, 2012. 22(1): p. 90-106.
[2] Liu, T., et al., The EMT transcription factor, Twist1, as a novel therapeutic target for pulmonary sarcomatoid carcinomas. Int J Oncol, 2020. 56(3): p. 750-760.
[3] Razzaque, M.S. and A. Atfi, Regulatory Role of the Transcription Factor Twist1 in Cancer-Associated Muscle Cachexia. Frontiers in Physiology, 2020. 11.
[4] Zhao, Z., et al., Multiple biological functions of Twist1 in various cancers. Oncotarget, 2017. 8(12): p. 20380-20393.

ICB12-1402-1150

In silico analysis of identification dysregulated genes in doxorubicin treatment human cervical cancer cell (HeLa)

Noorolhoda Khalighi[a]. Somayeh Reiisi[b]*
a Department of pathobiology, Faculty of veterinary Medicine, Shahrekord University, Shahrekord, Iran
b Department of Genetics, Faculty of Basic Sciences, Shahrekord University, Shahrekord, Iran
*Corresponding author's email address: s.reiisi@sku.ac.ir; s.reiisi@yahoo.com

Abstract: Cervical cancer, the second most diagnosed gynecological cancer globally after breast cancer, is seeing increasing incidence and mortality rates despite effective screening and vaccination programs [1]. Doxorubicin (DOX), an anthracycline drug used in clinical chemotherapy for cervical carcinoma [2], acts by intercalating into DNA, inhibiting topoisomerase II, disrupting mitochondria, generating free radicals [3], and inducing apoptosis through various mechanisms [4]. DOX triggers early activation of p53, followed by caspase-3 activation and DNA fragmentation in tumor cells [5]. The objective of this investigation is to analyze the gene expression profile of HeLa cells treated by doxorubicin, aimed at delineating novel functional mechanisms of the drug.

Methods: The raw data set GSE125249 was downloaded from the Gene Expression Omnibus. Differentially expressed genes (DEGs) between doxorubicin-treated HeLa cells and untreated cells were identified using R packages such as GEOquery, limma, BiocGenerics, affy, and oligo. The DEGs were then used to construct a protein-protein interaction (PPI) network via the STRING database. Subsequently, Cytoscape software was employed to recognize hub genes. The DEGs also underwent gene ontology (GO) analysis and KEGG pathway analysis.
Results: The study detected 998 DEGs (|log2FC (fold change) |>1), comprising 335 upregulated and 663 downregulated DEGs between doxorubicin-treated and untreated cervical cancer cells (HeLa). Dysregulated genes were enriched in pathways such as NF-kappa B signaling, TNF signaling, and p53 signaling in KEGG, and in biological processes including Positive Regulation of p38MAPK cascade and regulation of mitotic cytokinesis. CytoHubba plugin in Cytoscape identified 10 hub genes: CENPA, PLK1, CDCA8, KIF2C, CDK1, AURK, KIF23, KIF20A, TTK, and ASPM. These hub genes were notably enriched in cell cycle, p53 signaling, FoxO signaling, and cellular senescence pathways.
Conclusion: Transcriptome datasets from doxorubicin-treated cancer cells offer insights into cancer-specific biological functions, pathways, and novel mechanisms for predicting clinical outcomes in cervical cancer.

Keywords: Cervical cancer; doxorubicin; Transcriptome; In silico.

References

[1] SD A, Pasumarthi D, Pasha A, Doneti R, Botlagunta M, Pawar SC. Identification of differentially expressed genes in cervical cancer patients by comparative transcriptome analysis. BioMed Research International. 2021 Mar 20;2021.

[2] Liu X, Yang X, Chen F, Chen D. Combined application of doxorubicin and naringin enhances the antitumor efficiency and attenuates the toxicity of doxorubicin in HeLa cervical cancer cells. International Journal of Clinical and Experimental Pathology. 2017;10(7):7303.

[3] Carvalho C, Santos RX, Cardoso S, Correia S, Oliveira PJ, Santos MS, Moreira PI. Doxorubicin: the good, the bad and the ugly effect. Current medicinal chemistry. 2009 Sep 1;16(25):3267-85.

[4] Wang X, Wang H, Mou X, Xu Y, Han W, Huang A, Li Y, Jiang H, Yang X, Hu Z. Lysophosphatidic acid protects cervical cancer HeLa cells from apoptosis induced by doxorubicin hydrochloride. Oncology Letters. 2022 Aug 1;24(2):1-8.

[5] Wang S, Konorev EA, Kotamraju S, Joseph J, Kalivendi S, Kalyanaraman B. Doxorubicin induces apoptosis in normal and tumor cells via distinctly different mechanisms: intermediacy of H2O2-and p53-dependent pathways. Journal of Biological Chemistry. 2004 Jun 11;279(24):25535-43

ICB12-1402-1151

Targeting the Tumor Stroma: Identifying Crucial Genes in Cancer-Associated Fibroblasts for High-Grade Serous Ovarian Cancer Therapeutics

M. Safari-Kharkeshi[a], N. Cheraghi[a], N. Salehi[b], H. Montazeri[c], Gh.Hossein[a]*
a Department of Animal Biology, Developmental Biology group, School of Biology, University College of Science, University of Tehran, Tehran, Iran
b School of Biological Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
c Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
*Corresponding author's email address: ghossein@ut.ac.ir

Abstract: Ovarian cancer ranks as the eighth leading cause of cancer-related deaths worldwide [1]. Histologically, about 90% of ovarian tumors originate from epithelial cells, with over 70% classified as high-grade serous epithelial ovarian cancer (HGSOC) and a less than 35% five-year survival rate [2]. Recent reports emphasize the critical role of the tumor microenvironment in cancer progression, with cancer-associated fibroblasts (CAFs) being a pivotal cell type [3]. Therefore, pinpointing key genes expressed in CAFs for potential therapeutic targeting is of utmost importance. In this study, a meta-analysis of gene expression associated with CAFs was conducted using microarray data from the GEO datasets, GSE126132 and GSE40595 were extracted, normalized, and subjected to PCA analysis using R programming. Differentially expressed genes (DEGs) with significant expression differences in CAFs isolated from HGSOC tumor tissues compared to normal ovarian fibroblasts (NOFs) were identified in human samples. To unravel interactions between DEGs, String software was utilized, and hub genes were identified using Centiscape and Cytohubba. Additionally, a weighted gene co-expression network analysis (WGCNA) was performed to identify hub genes from co-expressed genes. This comprehensive approach resulted in the identification of 28 hub genes from String analysis and 30 top hub genes using the WGCNA package. Our results underscore the significance of genes associated with the mTORC1 signaling pathway, MYC Target v1, and Complement pathways in the function of CAFs promoting HGSOC. Specifically, genes CCT6A, PSMD12, and COPS5 in the mTORC1 signaling pathway, SYNCRIP, COPS5, and PWP1 in the MYC Target v1 signaling pathway, and USP14, USP16 in the Complement signaling pathway in CAFs were linked to metabolic changes promoting the survival of tumor cells. The identified key genes may present potential targets in the treatment of high-grade serous ovarian cancer, warranting experimental validation.

Keywords: Cancer-associated fibroblasts; High grade serous ovarian cancer; Hub genes; WGCNA analysis; Metabolic pathways.

References

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA: Cancer J. Clin., 71 (2021) 209-249.
[2] S. Shah, A. Cheung, M. Kutka, M. Sheriff, S. Boussios, "Epithelial ovarian cancer: providing evidence of predisposition genes," I. J. E. R. P. H., 19 (2022) 8113.
[3] M. Sarkar, T. Nguyen, E. Gundre, O. Ogunlusi, M. El-Sobky, B. Giri, TR. Sarkar, "Cancer-associated fibroblasts: The chief architect in the tumor microenvironment," Front. Cell Dev. Biol., 11 (2023) 1089068.

ICB12-1402-1157

Identification of crucial genes and pathways associated with Hepatocellular carcinoma based on bioinformatics analysis

Hamed Pourhossein [a], Saba Hadi [b], Anahita Soleimani [b], Seyed Hossein Khoshraftar [b], Parisa Alirezaie [b], Mohammad Mostafa Pourseif [c]*
a Department of Laboratory Sciences, Faculty of Paramedicine, Tabriz University of Medical Sciences
b Department of Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
c Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: Among the various forms of liver cancer diagnosed in 2020, hepatocellular carcinoma (HCC) took the lead, constituting approximately 75-85% out of the nearly 906,000 reported cases. Additionally, HCC ranked as the third highest cause of cancer-related deaths globally in 2020, with an estimated 5-year survival rate of 20%. The primary aim of this study is to ascertain novel biomarkers associated with HCC, enabling the prediction of prognosis and the advancement of treatment strategies [1,2].
Methods: The GSE101685 dataset was obtained from GEO database, consisting of 24 cancer and 8 normal samples. To assess the consistency of the samples, Principal Component Analysis (PCA) was carried out. Subsequently, the statistical analysis was performed utilizing the LIMMA package. Differential expression of genes (DEGs) was determined based on the criteria of $|log2FC|> 3$ and adj P-value < 0.05. To examine the relationship between gene expression and survival outcomes, the Kaplan-Meier plotter was utilized.
Results: Microarray analysis was conducted to assess the expression profile of 69 samples. A total of 126 DEGs were identified, of which 43 were up-regulated and 83 were down-regulated. Among them, two genes SLCO1B3, and CLEC4M have been discovered in previous studies to be related to cancers such as Prostate and lung. These two genes show a significant decrease in expression in this dataset. Through the utilization of the GeneCards and Reactome pathway databases, it has been determined that these genes possess the ability to regulate significant pathways, including Bile acid and bile salt metabolism and Complement cascade. According to Kaplan-Meier plots DEGs are correlated with poor prognosis in HCC.
Conclusion: Our examination reveals that the SLCO1B3 and CLEC4M genes are vital in HCC and strongly associated with poor prognosis.
Keywords: hepatocellular carcinoma; biomarker; microarray.

References

[1] Sung, H., et al., Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin, 2021. 71(3): p. 209-249.
[2] Brar, G., et al., Hepatocellular Carcinoma Survival by Etiology: A SEER-Medicare Database Analysis. Hepatology Communications, 2020. **4**(10): p. 1541-1551.

ICB12-1402-1158

Deciphering Tumor Microenvironment Dynamics and Cellular Communication in Breast Cancer Progression through Single-Cell RNA Sequencing

Fereshteh Fallah Atanaki[a], Kaveh Kavousi[a] *
a Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran 14176-14411, Iran
*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Tumor microenvironment as well as inter- and intracellular communication within it, play a critical role in tumor progression and drug resistance. This study intends to detect and describe these complicated relationships during the progression of breast cancer using single-cell RNA sequencing. Focusing on breast cancer, a significant and widely prevalent malignancy, our microscopic examination aims to shed light on the molecular nuances underlying the transition from normalcy to BRCA1 precancerous states and ultimately to triple-negative breast cancer (TNBC).
By leveraging scRNA-seq data, we identify key genes that orchestrate substantial shifts between different stages of breast cancer. The complex interactions and processes in which these discovered genes engage in it are then clarified using network analysis. The importance of these key genes is further validated through machine learning models, providing a comprehensive understanding of their role in driving the transition from one stage to another. By shedding light on the molecular underpinnings of breast cancer at a microscopic level, our study aims to pave the way for the development of more effective treatment strategies, thereby offering hope for improved outcomes for patients battling this challenging disease.

Keywords: Tumor microenvironment; Cancer progression; Triple-negative breast cancer; Multilayer network; Machine learning.

References

[1] Karaayvaz M, Cristea S, Gillespie SM, Patel AP, Mylvaganam R, Luo CC, Specht MC, Bernstein BE, Michor F, Ellisen LW. "Unraveling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq." Nat Commun. 2018 Sep 4;9(1):3588. doi: 10.1038/s41467-018-06052-0. PMID: 30181541; PMCID: PMC6123496.

[2] Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, Bryant VL, Penington JS, Di Stefano L, Tubau Ribera N, Wilcox S, Mann GB; kConFab; Papenfuss AT, Lindeman GJ, Smyth GK, Visvader JE. "A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast." EMBO J. 2021 Jun 1; 40(11):e107333. doi: 10.15252/embj.2020107333. Epub 2021 May 5. PMID: 33950524; PMCID: PMC8167363.

[3] Ji Zhang, Meige Guan, Qianliang Wang, Jiajun Zhang, Tianshou Zhou, Xiaoqiang Sun, "Single-cell transcriptome-based multilayer network biomarker for predicting prognosis and therapeutic response of gliomas", Briefings in Bioinformatics, Volume 21, Issue 3, May 2020, Pages 1080–1097, https://doi.org/10.1093/bib/bbz040

ICB12-1402-1159

Advancing Feature Selection: A Hybrid Approach for High-Dimensional and Incomplete Data

B. Ebrahimi [a], N. Bagherpour [a]*
a Department of Engineering Sciences, University of Tehran, Tehran, Iran
*Corresponding author's email address: negin.bagherpour@ut.ac.ir

Abstract: A recently encountered challenge in data science and more specifically in machine learning is growing amount of data. Exclusion of superfluous data and thereby focusing on essential variables, known as feature selection, proves vital for model performance optimization. This study undertakes a thorough investigation into pivotal feature selection strategies. The benefits and limitations of each method is clearly stated. An advanced methodology is also presented for tackling incomplete datasets, alongside introducing an innovative hybrid model that unites the Partial Mutual Information Criterion (PMIC), state-of-the-art null value completion strategies and neural network synergies to improve feature selection processes. Finally, the suggested strategy is implemented in Python and numerical test results are reported on a few randomly generated data sets and three well-known datasets breast_cancer, iris and diabetes. We evaluate both the similarity (of selected features to known important features) and the accuracy of imputed data. The reported results confirm the efficiency of this hybrid algorithm in feature selection for large data sets suffering from incomplete data.

Keywords: Feature selection; Non-negative Latent Factor; Maximal Information Coefficient; Dimensionality; Incomplete Data.

References

[1] J. Hu, "An improved feature selection method for classification on incomplete data: Non-negative latent factor-incorporated duplicate MIC," Expert Systems with Applications, pp. 37-51, 2023.

[2] D. N. Reshef, "Detecting novel associations in large data sets. Science," Science, pp. 1518-1524, 2011.

[3] K. Li, "Deep feature screening: Feature selection for ultra high-dimensional data via deep neural networks," Neurocomputing, pp. 126-186, 2023.

ICB12-1402-1165

Ensemble based variational autoencoders for detecting protein complexes in protein-protein interaction networks

Malihe Danesh*, Parisa Rostami
Department of Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran
*Corresponding author's email address: m.danesh@mazust.ac.ir

Abstract: Protein-protein interaction (PPI) networks are composed of multiple protein complexes which play the essential roles in many biological functions and identifying different forms of a disease. Each protein complex is a group of some proteins interacting with each other. Nowadays, due to the limitations of experimental methods, computational approaches are used to identify the complexes. In this regard, measurement errors lead to the noisy and uncertain interactions, which makes it difficult to obtain reliable clusters. To face the challenge, a new method based on Ensemble Variational Autoencoders named EVA is proposed in this study, that benefits from deep embedding and consensus clustering together to deal with the uncertainty. Using variational autoencoder, it is possible to filter the noise by creating meaningful representations of the proteins and extracting important features of co-complex ones. In addition, the ensemble learning approach integrate multiple deep models to seek better embeddings of the proteins and lead to the more qualitative clustering of PPI networks.
In this regard, a similarity matrix is generated first using second-order proximity of pairwise proteins. Then, several varioational autoencoders are trained to embed the data points into the low dimensional feature space. Next, the resulting representations of each network are extracted and clustered independently. Finally, the base clusterings are combined to obtain a robust reliable complexes of the proteins. The proposed method was evaluated by four real datasets of PPI networks in different density and dimensions including Krogan-core, Krogan-extended, Collins and Gavin. According to the results of F-score and MCC (Matthews's correlation coefficient) evaluation metrics, the proposed method achieved significant efficiency compared to the recent clustering methods of protein interaction networks.

Keywords: Protein complex; variational autoencoder; ensemble learning; protein-protein interaction.

References

[1] S. Dilmaghani, MR. Brust, CHC. Ribeiro andetc, "From communities to protein complexes: A local community detection algorithm on PPI networks", PLoS One., 17(1): e0260484, 2022.

[2] M. Danesh, M. Dorrigiv, and F. Yaghmaee, "Ensemble-based clustering of large probabilistic graphs using neighborhood and distance metric learning", J. Supercomput., 77:4107-4134, 2021.

[3] K. Han K, F. Gui, X. Xiao et al, "Efficient and effective algorithms for clustering uncertain graphs", PVLDB., 12(6):667–680, 2019.

[4] M. Danesh, M. Dorrigiv, and F. Yaghmaee, "A survey of clustering large probabilistic graphs: Techniques, evaluations, and applications", Expert Syst., 40(6): e13248, 2023.

ICB12-1402-1166

Transcriptomics analysis of lung samples of COPD patients identifies shared pathways involved in the disease among independent studies

Arman Mokaram Doust Delkhah[a], Masoud Arabfard[b]*
a Department of Cell and Molecular Biology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
b Chemical Injuries Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: arabfard@bmsu.ac.ir

Abstract: Chronic obstructive pulmonary disease (COPD) manifests as a heterogeneous disease distinguished by enduring respiratory symptoms attributable to structural anomalies affecting the airways and alveoli. These anomalies give rise to persistent airflow obstruction, often exhibiting a progressive nature(1). Individuals who engage in cigarette smoking, the most prominent risk factor for COPD, generally experience pulmonary inflammation. However, those who develop COPD demonstrate a heightened or aberrant response to inhaling toxic agents(2, 3). Differentially expressed genes play a role in various pathogenesis pathways, including those responsible for inflammation(4). In the present study, we employed a transcriptomics merge-analysis approach to identify the Differentially Expressed Genes (DEGs) and their associated pathways. Two microarray datasets, GSE38974 and GSE27597, published in 2012 and 2011, respectively, were used in the differential expression analysis. The merged data encompassed a total of 17 normal smokers and 71 COPD smoker patients. The results of the Differentially Expressed Genes (DEG) analysis using criteria of $|logFC| > 1$ and Adjusted p-value < 0.05 revealed 35 upregulated and 37 downregulated genes between normal smokers and COPD smoker patients. Furthermore, the differentially expressed genes were involved in several pathways contributing to COPD pathogenesis such as TNF, IL-17, Chemokine, C-type lectin receptor, NF-kappa β, and Toll-like receptor signaling pathways (Adjusted p-value < 0.001).

Keywords: COPD; Lung Tissue; Transcriptome Analysis.

References

[1]Venkatesan P. GOLD COPD report: 2024 update. The Lancet Respiratory Medicine. 2024;12(1):15-6.
[2]MacNee W. Pathology, pathogenesis, and pathophysiology. Bmj. 2006;332(7551):1202-4.
[3]Song Q, Chen P, Liu X-M. The role of cigarette smoke-induced pulmonary vascular endothelial cell apoptosis in COPD. Respiratory Research. 2021;22(1):39.
[4]Huang X, Li Y, Guo X, Zhu Z, Kong X, Yu F, Wang Q. Identification of differentially expressed genes and signaling pathways in chronic obstructive pulmonary disease via bioinformatic analysis. FEBS Open Bio. 2019;9(11):1880-99.

ICB12-1402-1167

Transcriptomic Data Mining to Identify Circular RNAs as Biomarkers in the Development of Grape Clusters

S. Ranjbar [a], A. Emamjomeh[*b]
a Department of Plant Breeding and Biotechnology (PBB), Faculty of Agriculture, University of Zabol, Zabol, Iran
b Laboratory of Computational Biotechnology and Bioinformatics (CBB), Department of Bioinformatics, Faculty of Basic Sciences, University of Zabol, Zabol, Iran
*Corresponding author's email address: aliimamjomeh@uoz.ac.ir

Abstract: Circular RNAs (circRNAs) represent a novel class of RNA molecules with the potential to influence critical biological processes. These molecules exhibit tissue-specific expression patterns that are depend on the cell type and developmental stage. Sistan wine grapes, characterized by seed lessness and small berry size, were the focus of our investigation. In this study, we aimed to identify the index circRNAs that are active in cluster formation and investigated the effects of gibberellin treatment on their expression levels. Eight detection tools were used to predict the expressed circRNAs. Ten attribute weighting algorithms were used to identify the most important circRNAs from total reliable identified circRNAs as features. Before performing the algorithms, data were normalized and changed into 0-1 amounts. Three ranges greater than 0.5, 0.75 and 0.95 were defined as selection thresholds. In total, Chi Squared and Rule models had the most selected attributes (weigh>0.9) and PCA, SV, Relief and Deviation had the lowest one (weight<0.5). The number of 17 circRNAs was selected considering the threshold of 0.95 and at least six weighting algorithms as biomarkers. These circRNAs can be considered as biomarkers in the three stages of grape cluster formation treated with gibberellic acid or non-treated plants. Among these, 7 were related to treatment and others were controls. The only circRNA identified by eight weighting methods is 11:16250294-16250532 from the group of control plants, which had higher expression in the first and third stage of cluster formation. It corresponds to gene ID VIT_11s0103g00420 and is associated with the gene ontology (GO) term "regulatory ncRNA-mediated gene silencing" in the biological process (BP) category.

Keywords: circRNAs; data mining; gibberellin treatment; grape; weighting algorithms.

References

[1] G. Kalwan, SS. Gill, P. Pariyadarshni, *et al,* "Approaches for identification and analysis of plant Circular RNAs and their role in stress responses," J. Environ Exp Bot., 205 (2022) 105099.
[2] Y. Shiri, M. Solouki, E. Ebrahimie, A. Emamjomeh, *et al,* "Unraveling the transcriptional complexity of compactness in sistan grape cluster," J. Plant Sci., 270 (2018) 198-208.
[3] Y. Shiri, M. Solouki, E. Ebrahimie, A. Emamjomeh, *et al,* "Gibberellin causes wide transcriptional modifications in the early stage of *grape* cluster development," J. Genomics., 112 (2020) 820-830.
[4] S. Ranjbar, A. Emamjomeh, M. Ebrahimi, A. Ghorbani, *et al*, "Comparative analysis of bioinformatic tools to predict and quantify active circular RNAs during grape cluster development," J. *Biol Plantarum.,* 68 (2024) 1-11.
[5] E. Ebrahimie, F. Ebrahimi, M. Ebrahimi, S. Tomlinso, *et al*, "A large-scale study of indicators of sub-clinical mastitis in dairy cattle by attribute weighting analysis of milk composition features: highlighting the predictive power of lactose and electrical conductivity," J. J Dairy Res., 85.2 (2018) 193-200.
[6] E. Ebrahimie, F. Zamansani, IO. Alanazi, EM. Sabi, *et al*, "Advances in understanding the specificity function of transporters by machine learning," J. Comput Biol Med., 138 (2021) 104893.

ICB12-1402-1169

High-throughput transcriptome analysis for identification of *Helicobacter pylori* vaccine target antigens

Amir Bajmanlou Balaei [a,b], Amir Abbas Barzegari [a], Fatemeh Khadivi Derakhshan [c], Ailar Nakhlband [d], Negar Hossein Nezhad Daneshvar [b], Mohammad Mostafa Pourseif [b,e*]

a Department of Biology, Faculty of Basic Science, University of Maragheh, Maragheh, East Azerbaijan Province, Iran

b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran

c Department of Biology, Faculty of Basic Science, Urmia Branch, Islamic Azad University, Urmia, West Azerbaijan Province, Iran

d Research Center of Psychiatry and Behavioral Sciences, Aging Research Institute, Tabriz University of Medical Sciences, Tabriz, Iran

e Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran

*Corresponding author's email address: mmpourseif@gmail.com

Abstract: *Helicobacter pylori* (Hp), a Gram-negative bacterium, establishes enduring colonization in the human stomach by inducing immunoregulatory responses [1]. Due to its involvement in chronic gastritis, peptic ulcers, and gastric cancer, researchers aim to prevent these conditions through novel, safe, and highly effective vaccines [2]. Microarray and RNA-seq techniques play crucial roles in aiding vaccinologists to identify novel vaccine-targeted antigens contributing significantly to the pathogenesis of various pathogens. In this study, we analyze microarray (GSE60427 and GSE123623) and RNA-seq (GSE164216) datasets from the NCBI GEO database. After quality control using R program Bioconductor packages, principal component analysis assesses sample uniformity within each dataset. The Limma package conducts statistical analysis, and a Venn diagram identifies common differentially expressed virulence genes (DEVGs). Kaplan-Meier plotter assesses the correlation between gene expression and survival outcomes.

High-throughput transcriptome analysis of 48 samples reveals 24 significantly upregulated and 23 downregulated genes (adj.P.val < 0.01 and |log2 FC| > 2). Top less-annotated genes (HP1167 and HP1440) were identified in the upregulated list, and (HP1588 and HP0415) in the downregulated list, with predicted key roles in bacterial pathogenicity, proposing them as notable vaccine antigens. Genes like dppA, gpsA, pdxJ, rnhB, tagD (downregulated), and flgK, flaA, cag7, fliD, flag (upregulated) were discovered as known antigens in bacterial invasion pathways. GeneCards and KEGG pathway databases indicate their significant molecular function in pathogenicity-related pathways, suggesting them as vaccine targets in subsequent in silico and experimental studies.

This study not only provides valuable insights into the molecular signatures associated with Hp's pathogenicity but also highlights their potential for developing prophylactic interventions in clinical settings.

Keywords: Antigen; Helicobacter pylori; Microarray; RNA-seq; Transcriptome; Vaccine.

References

[1] Tubau-Juni, N., et al., Identification of new regulatory genes through expression pattern analysis of a global RNA-seq dataset from a Helicobacter pylori co-culture system. Sci Rep, 2020. 10(1): p. 11506.

[2] Sutton, P. and J.M. Boag, Status of vaccine research and development for Helicobacter pylori. Vaccine, 2019. 37(50): p. 7295-7299.

ICB12-1402-1171

Logical engineering of cellulase enzyme isolated from Acidothermus cellulolyticus bacteria for increasing thermal stability

Fatemeh Mohammadian
a Department of Biology, Faculty of Science, University of Sistan and Baluchestan, Zahedan, Iran *
m.lagzian@science.usb.ac.

Abstract: Cellulases of microbial origin have shown their potential application in industry today.However, the search for thermostable natural cellulases remains challenging. Therefore, alternative approaches like rational engineering offer promising prospects for their production. In this study, we isolated and identified a highly cellulase-producing strain from the microbial flora of Taftan volcanic region. Molecular methods, including 16srRNA gene sequencing, were employed to characterize the bacterial strain and the cellulase enzyme gene, which were the main focus of our research . Our findings highlighted the Acidothermus cellulolyticus strain as an excellent cellulase producer.Subsequently, we performed homology modeling of the cellulase enzyme using AlphaFold server patterns and assessed the model's quality for the next steps. The constructed 3D model exhibited high structural quality, rendering it suitable for subsequent stages such as in-silico mutagenesis. Notably, it possessed four unique mutations (G360I, G464F, G4R, N479C) that would stabilize the enzyme. These mutations, as validated by FoldX5, amplified the enzyme's stability, achieving a score of -17.54 kcal/mol. We employed Schrödinger software to carry out molecular docking and MM-GBSA scoring for the wild-type enzyme and the mutants variants. The wild-type enzyme yielded an average docking and MM-GBSA score of -34.819 kcal/mol, serving as a reference point for evaluating the mutants. The research then proceeded with Alanine scanning using Discovery Studio 2018 to pinpoint the most stable enzyme-stabilizing mutations, leading to the creation of single, double, and triple mutants while preserving key residues. We identified G360I, G4R as the most potent mutant variant of the cellulase enzyme.To assess the behavior of the mutant cellulase under thermal stress, we plan to conduct molecular dynamics simulations that compare the mutant and native structures. In conclusion, we discovered a unique cellulase from an extreme environment and enhanced its stability through a streamlined computer-assisted protein engineering process. This modified enzyme could be beneficial for industrial applications that require cellulase activity at high temperatures. [1-3]
Keywords: the cellulase enzyme; thermal stability; laboratory experiments; computational methods

References

[1] Gilmore, S.P., et al., Designing chimeric enzymes inspired by fungal cellulosomes. Synthetic and Systems Biotechnology, 2020. 5(1): p. 23-32.
[2] Tang Z, Jin W, Tang Y, et al. Research on homology modeling, molecular docking of the cellulase and highly expression of the key enzyme (Bgl) in Pichia pastoris. Int J Biol Macromol. 2018; 115:1079–1087.
[3] Chalak A, Villares A, Moreau C, et al. Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. Biotechnol Biofuels. 2019 Sep;12 (1):206.

ICB12-1402-1172

Deciphering the Regulatory Symphony: MicroRNA Orchestration of the RAS/MAPK Signaling Pathway and its Interplay across Cellular Networks in Colorectal Cancer Tumorigenesis – A Systems Biology Perspective

Alireza Pasdar [a], Elnaz Farzadifar [b]
a Department of Medical Genetics and Molecular Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
b Biology Department, Faculty of Sciences, Science and Arts University, Yazd, Iran
farzadifarelnaz@gmail.com

Abstract: Background: Colorectal cancer (CRC) stands as a global health challenge, prompting an exploration of its molecular intricacies. The RAS/MAP kinase pathway, pivotal in cancer development, intertwines with microRNAs (miRNAs), shaping CRC's dynamic cellular landscape. This study utilizes a systems biology approach, employing advanced bioinformatics tools to unravel miRNA-mediated regulatory networks in CRC. Investigating miRNA dysregulation, identifying hub genes, and exploring interplay between signaling pathways aim to provide nuanced insights into CRC's molecular landscape. Protein-protein network analysis uncovers the interconnected web of miRNA-targeted genes within the RAS/MAPK signaling pathway. Utilizing STRING and Cytoscape, we pinpoint 13 hub proteins orchestrating network dynamics. GO and KEGG enrichment analyses reveal intricate regulatory mechanisms. Cluster analysis unveils 817 clusters, emphasizing CRC significance. Hub gene promoter motif analysis delves into transcriptional regulatory elements governing CRC pathogenesis.
Dysregulated miRNA expression in CRC is explored, emphasizing miRNAs' dual roles as oncogenes or tumor suppressors. A comprehensive list of miRNAs targeting the RAS/MAPK pathway in CRC is presented. Protein-protein interaction networks highlight 13 hub proteins, predominantly linked to transcriptional regulation. Enriched pathways, such as the MAPK cascade, underscore their pivotal roles. Cluster analysis reveals CRC's significance, emphasizing involvement in glucose metabolism. Promoter motif analysis uncovers significant motifs targeted by dysregulated miRNAs. Transcription factor motifs reveal biological roles, with implications in signal transduction and cell proliferation. Enrichment in nucleus-based terms and disruption of cell polarity underscore potential implications in cancer development.
This integrative study unravels the regulatory intricacies of miRNAs in the RAS/MAPK signaling pathway, shedding light on CRC's molecular landscape. Identification of hub genes, enriched pathways, and regulatory motifs provides valuable insights for diagnostics and therapeutics, offering promising targets for simultaneous inhibition in cancer cell growth.

Keywords: Colorectal cancer, Network analysis, subnetwork analysis, Promoter motif analysis, RAS/MAPK signaling pathway

ICB12-1402-1173

The promotor methylation level of ATG12 correlates with tumor progression and autophagy in Bladder cancer; A bioinformatic study

Mir Shayan Jahangirzadeh[a], Mehdi Asghari Vostakolaei[b], Asadollah Asadi[c], Masoumeh Valipour [a, *]
a Department of Biology, Faculty of Basic Science, Azarbaijan Shahid Madani University, Tabriz, Iran.
b Cancer Immunology and Immunotherapy Research Center, Ardabil University of Medical Sciences, Ardabil, Iran.
c Department of Biology, Faculty of Science, University of Mohaghegh Ardabili, Ardabil, Iran.
*Corresponding author's email address: m.valipour@azaruni.ac.ir

Abstract: Bladder cancer (BC) is among the top ten most common cancer types in the world [1]. The heterogeneity of BC calls for substantial research with more in-depth molecular characterization, with the hope of identifying new diagnostic and treatment options [2]. DNA methylation is an epigenetic mechanism involving the transfer of a methyl group. DNA methylation regulates gene expression by recruiting different proteins or by inhibiting the binding of transcription factors to DNA. Methylation can alter gene expression and regulate different signaling pathways such as autophagy [3]. Autophagy has a conserved mechanism whose main role is to recycle cellular components and maintain homeostasis through the removal of undesirable proteins [4]. ATG12 is one of the most important genes involved in autophagy pathways [5]. **Aim:** This study aimed to explore the function of the ATG12 gene in the autophagy pathway involved in BC based on bioinformatic databases. **Methods:** Clinical information and promotor methylation level of ATG12 in patients with BC was obtained from The Cancer Genome Atlas (TCGA) based on the following selection criteria including basic clinical information of sample types (normal and primary tumor), gender, age, stage, and nodal metastasis. All requirements were conducted on a large sample size (>400). Datasets were then compared to obtain the significant p-value. **Results:** Our results indicate that the promotor methylation level of ATG12 is decreased in BC compared with normal tissue. This decrease was observed in different BC stages (1-4), different age groups (21-40, 41-60, 61-80, and 81-100 ys), gender, and nodal metastasis status of BC. There were no significant differences in gender and nodal metastasis (N0, N1, N2, and N3) groups. **Conclusion:** Our study revealed that ATG12 can be considered an important factor in BC progress with diagnostic and prognostic values.

Keywords: ATG12; Bladder cancer; Methylation; Autophagy; The Cancer Genome Atlas (TCGA).

References

[1]Saginala, K., et al., Epidemiology of bladder cancer. Medical sciences, 2020. 8(1): p. 15.

[2]Minoli, M., et al., Evolution of urothelial bladder cancer in the context of molecular classifications. International journal of molecular sciences, 2020. 21(16): p. 5670.

[3]Mattei, A.L., N. Bailly, and A. Meissner, DNA methylation: a historical perspective. Trends in Genetics, 2022. 38(7): p. 676-707.

[4]Klionsky, D.J., et al., Autophagy in major human diseases. The EMBO journal, 2021. 40(19): p. e108863.

[5]Rubinstein, A.D., et al., The autophagy protein Atg12 associates with antiapoptotic Bcl-2 family members to promote mitochondrial apoptosis. Molecular cell, 2011. 44(5): p. 698-709.

ICB12-1402-1174

Assessment of the isolated compounds from Cumin (Cuminum cyminum) as inhibitor of VEGFR-2 using molecular docking methodology

MohammadMoein Bashi Mehrizi, Seyed Mohammad Mahdi Mortazavian*, Hossein Hadavand Mirzaei
Master's student of the Department of Agronomy and Plant Breeding Sciences, College of Aburaihan, University of Tehran, Tehran, Iran.
Associate Professor of Plant Molecular Genetics and Genetic Engineering, Department of Agronomy and Plant Breeding Sciences, College of Aburaihan, University of Tehran, Tehran, Iran.
Department of Molecular Physiology, Agricultural Biotechnology Research Institute of Iran, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran.
*Corresponding author's email address: mortazavian@ut.ac.ir

Abstract: Cancer is the second leading cause of morbidity and mortality after cardiovascular disease all over the world. At the moment, more than 60% of the drugs that are used in cancer treatment are derived from natural sources. The Apiaceae family contains over 3700 plant species, many of which have medicinal properties and are used in traditional medicine. The vascular endothelial growth factor receptor (VEGFR) has a key role in angiogenesis. This process accelerates the conversion of tumors from benign to malignant abnormally in cancer patients. Therefore, the identification of angiogenesis inhibitors can help us in the treatment of cancer disease. Herein, we assessed the inhibitory potential of the 91 natural compounds isolated from the cumin plant (Cuminum cyminum) against the VEGFR-2 protein by in-silico strategy. First of all, the physicochemical properties of the compounds were calculated using the Swiss ADME a freely accessible web server as screening methodology. According to the obtained results, 49 compounds were selected to evaluate their energy binding in the active site of VEGFR-2 protein using AutoDocktools software. The results show that the average binding energy was -5.31 kcal/mol. The highest energy binding was indicated by germacrene-D with -7.51 kcal/mol, and the lowest energy binding with -3.43 kcal/mol was obtained from the interaction of Z-3-hexanol with target protein.

Keywords: Cuminum cyminum; Molecular docking; VEGFR-2; Antiangiogenesis.

References

[1] Nadeem M, Riaz A (2012) Cumin (*Cuminum cyminum*) as a potential source of antioxidants. Pakistan J Food Sci 22:101–107

[2] Peng FW, Liu DK, Zhang QW, et al (2017) VEGFR-2 inhibitors and the therapeutic applications thereof: a patent review (2012-2016). Expert Opin Ther Pat 27:987–1004. https://doi.org/10.1080/13543776.2017.1344215

[3] Mirzaei HH, Hosseini SM (2023) In-Silico Assessments of Fruticulin-A and Demethylfruticulin-A Isolated from Salvia Species Against Important Anticancer Targets. Res J Pharmacogn 10:31–41. https://doi.org/10.22127/RJP.2023.385883.2062

ICB12-1402-1176

Medical Image Segmentation using EfficientNet-based U-Net Architecture

Parisa Rostami, Mostafa Nikoseresht, Malihe Danesh*
Department of Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran
*Corresponding author's email address: m.danesh@mazust.ac.ir

Abstract: Medical image segmentation plays a pivotal role in computer-aided diagnosis and treatment planning. This study focuses on improving the accuracy of medical image segmentation, a crucial step in diagnosing and treating diseases. We explore the use of a powerful neural network, EfficientNet, combined with the U-Net architecture. We compare the performance of our model with a conventional U-Net variant utilizing the ResNet34 backbone, aiming to assess the efficacy of EfficientNet in this domain.

Our experiments involve extensive training on a curated dataset, encompassing diverse anatomical regions, including large bowel, small bowel, and stomach. We evaluate the segmentation performance through metrics such as Intersection over Union (IoU) and Dice Coefficient. Additionally, we discuss the importance of proper data preprocessing, including input normalization and augmentation, in achieving robust segmentation results.

The presented EfficientNet-based U-Net architecture holds significant potential for real-world deployment in medical image analysis, offering improved segmentation accuracy and computational efficiency. This work contributes to the ongoing exploration of deep learning architectures in medical imaging, paving the way for advancements in clinical diagnosis and treatment planning.

Keywords: Medical imaging; Segmentation; U-Net Architecture; EfficientNet.

References

[1] Jimin Cao, Xu Liang, Xueli Chen, Yonghua Zhan, "Medical Image Segmentation based on U-Net: A Review," Journal of Imaging Science and Technology, vol. 64, 2020, DOI: 10.2352/J.ImagingSci.Technol.2020.64.2.020508.

[2] N. Siddique, S. Paheding, CP. Elkin, V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," IEEE Access, vol. 9, 2021, pp. 82031–82057, DOI: 10.1109/access.2021.3086020.

[3] N. Siddique, S. Paheding, AA. Reyes Angulo, MZ. Alom, V. K. Devabhaktuni, "Fractal, recurrent, and dense U-Net architectures with EfficientNet encoder for medical image segmentation," Journal of Medical Imaging, vol. 9, no. 6, 064004, 2022.

ICB12-1402-1177

In-silico analysis of a missense mutation (Leu309Pro) in the *PYGM* gene associated with Glycogen storage disease type V

Leila Keshavarz, Dr. Mehri Khatami*
Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: m.khatami@yazd.ac.ir

Abstract: Glycogen storage disease type V (also called McArdle's disease and GSDV) is a disorder of muscle metabolism. Muscle phosphorylase enzyme deficiency causes this disease. One of the effects of this disease is the inability to break down glycogen "fuel" reserves. It should be noted that McArdle's disease leads to reversible acute kidney failure, severe muscle damage, pain and fatigue with vigorous exercise. In most cases, the age of onset of this disease occurs in the first decade of life, but it can be different. In some cases, diagnosis of the disease due to neglect of myalgia and fatigue in old age leads to involvement of proximal muscles and continuous weakness. The diagnosis of GSDV in an individual is established by the identification of biallelic PYGM (encoding glycogen phosphorylase, muscle form). In this article, the pathogenicity of the missense mutation of the leucine amino acid at position 309 to proline in the PYGM gene has been investigated in the PolyPhen-2 and HOPE bases.

PolyPhen-2 is a tool that predicts the potential impact of amino acid substitutions on human protein function. Regarding the Leu309Pro amino acid mutation, according to the polyPhen-2 results to both HumDiv and HumVar, this mutation is indicated to be probably damaging with a score of 1000. Moreover HOPE is an online web service where the user can submit sequences and mutations. Based on HOPE's survey of genomic variants, the MetaRNN score of this species is 0.9898901. It can be from 0.0 to 1.0. The higher it is, the more likely it is to be pathogenic, so the mutation is likely to damage the protein.

In conclusion, it is likely that this missense mutation in the PYGM gene is pathogenic and possibly damaging.

Keywords: McArdle disease; GSDV; PYGM.

References

[1] A Martín M, Lucia A, Arenas J, et al. "Glycogen Storage Disease Type V". University of Washington; Updated 2019.

[2] Quinline R , Martinuzzi A , Schoser B. "Pharmacological and nutritional treatment for McArdle disease (Glycogen Storage Disease type V)". Cochrane Library; 2014.

ICB12-1402-1178

The Expression Level of SQSTM1 Related to the Progression and Autophagy of Gastric Cancer Based on TCGA

Ali Nosrat[1,2], Vahid Asghariazar[3], Hamed Haghi-Aminjan[1], Kazem Nejati-Koshki[1,*]
1 Pharmaceutical Sciences Research Center, Ardabil University of Medical Sciences, Ardabil, Iran
2 Deputy of Research and Technology, Ardabil University of Medical Sciences, Ardabil, Iran.
3 Cancer Immunology and Immunotherapy Research Center, Ardabil University of Medical Sciences, Ardabil, Iran.
*Corresponding author's email address: kazemnejati1364@gmail.com

Abstract: Gastric cancer(GC) is one of the main causes of morbidity and mortality worldwide [1]. Autophagy, an intracellular homeostatic pathway, plays an important role in the physiopathological processes of human diseases such as cancer [2]. The expression of p62/SQSTM1 (P62) was considered to evaluate the clinical significance of autophagy in GC. P62 is an autophagy-related receptor protein that is vital in regulating multiple signaling pathways and tumorigenesis [3]. Aim: This examination aimed to investigate the function of the SQSTM1 gene in the autophagy pathway involved in GC based on bioinformatic databases. Methods: Clinical data and expression level of SQSTM1 in patients with GC were obtained from The Cancer Genome Atlas (TCGA) based on the following selection criteria including patient's gender, race, sample types (normal and primary tumor), stage, tumor grade, and histological subtype. All requirements were conducted on a large sample size (>400). Datasets were then compared to obtain the significant p-value. Results: Our results demonstrate that the expression level of SQSTM1 is increased in GC compared with normal tissue. This increase was discovered in different GC stages (1-4), different races (Asian, Caucasian), genders, tumor grades (1-3), and five different histological subtypes. There were no significant differences in the African-American race and 2 histological subtypes (IntAdenoMucinous and IntAdenoPapillary). Conclusion**:** Our study revealed that SQSTM1 can be considered a substantial factor in GC development with diagnostic and prognostic values.

Keywords: SQSTM1; P62; Autophagy; The Cancer Genome Atlas (TCGA); Gastric Cancer.

References

[1]Xin-Bo, X., et al., Expression and prognostic significance of the DNA damage response pathway and autophagy markers in gastric cancer. Neoplasma, 2021. 68(6).

[2]Qian, H.-r. and Y. Yang, Functional role of autophagy in gastric cancer. Oncotarget, 2016. 7(14): p. 17641.

[3]Xu, Y., et al., SQSTM1/p62 as a prognostic marker promotes the progression of gastric cancer through epithelial-mesenchymal transition. 2023.

ICB12-1402-1179

In silico analysis of single nucleotide polymorphisms (SNPs) in human *HGD* gene which is associated with Alkaptonuria

Farideh Askari*, Mohammad Mehdi Heidari[1]
[1] Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: faride.ask1381@gmail.com

Abstract: Alkaptonuria, resulting from homogentisate 1,2-dioxygenase deficiency, presents with dark urine, ochronosis, and arthritis. Manifestations include pigment deposition and complications such as valve calcification, renal stones, and hypothyroidism. Early recognition is crucial for effective management. The *HGD* gene encodes homogentisate 1,2-dioxygenase and in this study sought to discover disease-causing single nucleotide polymorphisms (SNPs) in the *HGD* gene using bioinformatics servers. One SNP rs28941783 (G161R) was studied using SIFT, PolyPhen-2, I-Mutant 2.0, and Hope servers. The result showed that rs28941783 affected protein function with a score of 0.00 by the SIFT and was predicted to be possibly damaging with a score of 1.000 by PolyPhen-2. Moreover, a large decrease in the stability of protein was predicted by I-Mutant (DDG = -1.59). Hope server indicated that mutation of wild-type residue, glycine, can abolish the protein's function because glycine is the most flexible of all residues and this flexibility might be necessary for the protein's structure and function. The mutated residue is located in a domain and in contact with another domain that is important for the activity of the protein. The interaction between these domains could be disturbed by the mutation, which might affect the function of the protein. This interaction may be important for the correct function of the protein. Ultimately, the study suggests that the G161R variant of *HGD* gene could affect protein function.

Keywords: HGD gene; SIFT; PolyPhen-2; I-Mutant; Hope.

References

[1] Abdelkhalek ZS, Mahmoud IG, Omair H, Abdulhay M, Elmonem MA. Homogentisate 1,2-dioxygenase (HGD) gene variants in young Egyptian patients with alkaptonuria. Sci Rep. 2023 Sep 1;13(1):14374. doi: 10.1038/s41598-023-41200-7. PMID: 37658095; PMCID: PMC10474279.

[2] Sharabi AF, Goudar RB. Alkaptonuria. 2023 Aug 8. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan–. PMID: 32809406.

[3] Zaib S, Rana N, Hussain N, Ogaly HA, Dera AA, Khan I. Identification of Potential Inhibitors for the Treatment of Alkaptonuria Using an Integrated In Silico Computational Strategy. Molecules. 2023 Mar 14;28(6):2623. doi: 10.3390/molecules28062623. PMID: 36985595; PMCID: PMC10058836

ICB12-1402-1180

Bioinformatic Discovery of Novel FGFR-1 Inhibitors from Natural Products: A Promising Approach for Lung Cancer Treatment

Nayereh.Shariat Gonabady[a], Zeinab. Bagheri[b*]
a Department of Cell and Molecular Biology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
b Department of Cell and Molecular Biology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
*Corresponding author's email address: email: ze_bagheri@sbu.ac.ir

Abstract: Lung cancer is a prevailing and substantially type of cancer that is attributed to a significant number of cancer-related fatalities due to its highly metastatic nature. While current therapies suggest some benefits, they often face limitations such as toxicity and drug resistance, highlighting the urgent need for novel and more effective treatment options. Natural products, with their rich phytochemical profiles and reduced toxicity, have emerged as promising candidates for therapeutic development. This study employed a bioinformatic approach to identify bioactive phytochemicals with potential efficacy in targeting fibroblast growth factor receptor 1 (FGFR-1), a key regulator of the extracellular signal-regulated kinase (ERK) pathway in cancer progression. A comprehensive analysis of over 8,000 phytocompounds from the ZINC15 database was conducted against FGFR-1 as the target protein, obtained from the Protein Data Bank (PDB). Docking simulations were performed using Schrodinger Maestro 11.5 to evaluate the glide emodel and docking score of these phytochemicals with FGFR-1. The top-performing compounds were further assessed using Pyrx virtual screening to assess their binding affinity and calculate root mean square deviation (RMSD). To assess their pharmacokinetic properties, the compounds were evaluated using the Admetlab2 database. Among the screened phytochemicals, ZINC5075238 with Glide emodel= -68.52, docking score = -6.933 binding affinity= -8 and RMSD=0 exhibited significantly higher binding interaction to FGFR-1 compared to the commercially available FGFR-1 inhibitor pemigatinib, while also possessing favorable pharmacokinetic properties. This finding presents a promising lead for the development of novel natural product-derived FGFR-1 inhibitors. Further studies are warranted to validate the efficacy and safety of ZINC5075238 in preclinical models and clinical trials. In conclusion, this study demonstrates the potential of utilizing a bioinformatic approach to identify novel bioactive phytochemicals with promising therapeutic applications in lung cancer. The discovery of ZINC5075238 as a FGFR-1 inhibitor emphasizes the value of natural products as therapeutic agents.

Keywords: FGFR-1; lung cancer; Drug design; Anti-cancer agents; Phytochemicals

ICB12-1402-1181

Analysis and Prediction Hormonal Effects of Metformin on PCOS in Mice using Machine Learning algorithms

Mehri Mirhoseini, dept. of Nursing (School of Nursing and Midwifery Amol) Mazandaran University of Medical Sciences Amol, Iran
Mohammad Taghi Ghorbanian, dept. of Biology Damghan University Damghan, Iran
C. Corresponding author Karim Mostafapour*, A. Presenter of the Papera, C. *
*Corresponding author's email address: s_mostafapour@webmail.guilan.ac.ir

Abstract: polycystic ovarian syndrome (PCOS) forecasting is a useful tool for increasing woman's awareness of their reproductive health. This illness is the most common endocrine disease in women of reproductive age [1]. In this study, machine learning algorithms are used to predict the PCOS. The six important classification models are applied to forecast PCOS model in mice. Considering the positive effects of metformin as a blood sugar-lowering and sex hormone-regulating drug that can improve the physiological and histological activity of the ovary [2]. the present study aims to investigate and predict the therapeutic effect of metformin on important hormonal parameters and changes The amount of blood sugar is to improve polycystic ovary syndrome with the help of machine learning [3]. Models were compared and evaluated using statistical measures such as Accuracy, Balanced Accuracy, AUC-ROC Curve, and F1 Score. Based on these evaluations, the best model was selected. kNN is thought to be the best model for forecasting PCOS. The results demonstrate that In comparison to the patient group, all the studied parameters in PCOS improved following metformin treatment and the Kneighborsclassifier's Accuracy, Balanced Accuracy, AUC-ROC Curve, and F1 Score are 1.

Keywords: polycystic ovarian syndrome; hyperandrogenism; Hyperinsulinemia; machine learning; kNN.

References

[1] Hu Q, Jin J, Zhou H, Yu D, Qian W, Zhong Y, et al. Crocetin attenuates DHT-induced polycystic ovary syndrome in mice via revising kisspeptin.

[2] Zhang Y, Hu M, Meng F, Sun X, Xu H, Zhang J, et al. Metformin ameliorates uterine defects in a rat model of polycystic ovary syndrome. EBioMedicine. 2017;18:157-70.

[3] Hetal Bhavsar, Amit Ganatra, A survey Variations of Support Vector Machine classification Technique. International Journal of Advanced Computer Research. 2012.

ICB12-1402-1182

In silico identification of G-quadruplex structure in the STAT-6 gene sequence

Saeedeh Ghazaey Zidanloo [1*] Nazanin Mohajeri[1]
[1] Department of Cell and Molecular Biology, Kosar University of Bojnord, Bojnord, Iran.
Saeedeh Ghazaey Zidanloo, Department of Cell and Molecular Biology, Kosar University of Bojnord, Bojnord,
*Corresponding author's email address: Ghazaeys@kub.ac.ir

Abstract: Signal Transducers and Activators of Transcription (STAT) are a family of transcription factors that regulate a variety of cellular processes such as proliferation, differentiation, and survival. The STAT6 gene is located on chromosome 12q13.3 with 23 exons and 15955 nucleotides. It is expressed in almost all organs of the body and may act as a T helper type 2 (Th2) inducing transcriptional activator [1,2]. Therefore, STAT6 is involved in the pathophysiology of numerous allergic diseases. Nevertheless, it is similarly complicated in several tumor developments, especially lymphomas and solitary fibrous tumors, and contributes to the regulation of the tumor microenvironment [2]. Therefore, regulating its gene expression is important for disease treatment. One of the ways to control the expression of this gene is to induce G-quadruplex structures, which act as a barrier to the gene expression apparatus [3]. Interestingly, there is a database that lists quadruplex-forming G-rich sequences (QGRS) (http://tubic.tju.edu.cn/greglist/) [4]. The sequence of the STAT6 gene was obtained from GenBank at NCBI. We used a web-based server, QGRS Mapper, which predicts quadruplex-forming G-rich sequences (QGRS) in nucleic acid sequences (http://bioinformatics.ramapo.edu/QGRS/) [5]. STAT6 was a large gene and the GC content of its sequence was very high, based on the data of the online software, we found 131 sequences prone to G-quadruplex formation, among which 13 regions had a score above 35, which indicates that the probability of the structure formation in these regions is very high. According to the recent evidence for the in vivo location and role of DNA G-quadruplexes in several cellular pathways including DNA replication and gene expression, effective treatment strategies can be defined for drug design and targeted treatment of the diseases.

Keywords: STAT6; Regulation of gene expression; G-quadruplex structure; In Silico analysis.

References

[1]Zhu M, Li S, Cao X, Rashid K, Liu T. The STAT family: Key transcription factors mediating crosstalk between cancer stem cells and tumor immune microenvironment. Semin. Cancer Biol., vol. 88, Elsevier; 2023, p. 18–31.
[2]Karpathiou G, Papoudou-Bai A, Ferrand E, Dumollard JM, Peoc'h M. STAT6: a review of a signaling pathway implicated in various diseases with a special emphasis in its usefulness in pathology. Pathol Pract 2021;223:153477.
[3]Awadasseid A, Ma X, Wu Y, Zhang W. G-quadruplex stabilization via small-molecules as a potential anti-cancer strategy. Biomed Pharmacother 2021;139:111550.
[4]Zhang R, Lin Y, Zhang C-T. Greglist: a database listing potential G-quadruplex regulated genes. Nucleic Acids Res 2008;36:D372–6.
[5]Kikin O, D'Antonio L, Bagga PS. QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. Nucleic Acids Res 2006;34:W676–82. https://doi.org/10.1093/nar/gkl253.

ICB12-1402-1184

Analysis and Prediction PCOS Using Classification Algorithms

A.Azizi*, K. Mostafapour, S. Rahmany
Department of Mathematics and Computer Science, Damghan University, Damghan, Iran
Department of Biology, University of Guilan, Guilan, Iran
*Corresponding author's email address: arshiyaazizi1999@gmail.com

Abstract: Five to Ten percent of women of reproductive age suffer with pcos, the most prevalent endocrine gland illness. Predicting the occurrence of polycystic ovarian syndrome (PCOS) is therefore a helpful method to raise women's awareness of reproductive health.

In this study, machine learning and neural network algorithms are used to predict the PCOS. The six important classification models are applied to forecast PCOS model in women. Models were compared and evaluated using statistical measures such as Accuracy, Balanced Accuracy, AUC-ROC Curve, and F1 Score. Based on these evaluations, the best model was selected. AdaBoostClassifier is thought to be the best model for forecasting PCOS. The findings show that the F1 Score, Balanced Accuracy, and AUC-ROC Curve values are 0.74, 0.74, and 0.65, respectively, and that the AdaBoostClassifier's accuracy is close to 80%.

Keywords: polycystic ovarian syndrome; machine learning; Accuracy; statistical; AdaBoostClassifier.

References

[1] Hetal Bhavsar, Amit Ganatra, A survey Variations of Support Vector Machine classification Technique. International Journal of Advanced Computer Research. 2012.

[2] Hu Q, Jin J, Zhou H, Yu D, Qian W, Zhong Y, et al. Crocetin attenuates DHT-induced polycystic ovary syndrome in mice via revising kisspeptin.

[3] Balen, A., (2004)."The pathophysiology of polycystic ovary syndrome: trying to understand PCOS and its endocrinology." Best practice \& research clinical obstetrics \& gynaecology, Vol. 18, No. 5, pp. 685-706.

ICB12-1402-1185

Presbycusis: Ranking the genes and associated biological function

Omid Jazayeri[1*] , Faezeh karimnia[2] , Amirhossein soltani[1]
Department of Molecular and Cell Biology, Faculty of Basic Science, University of Mazandaran, Babolsar, Iran
Department of Biology, Faculty of Science, Gonbad Kavous University, Gonbad, Iran
*Corresponding author's email address: o.jazayeri@umz.ac.ir

Abstract: Presbycusis also known as Age-related hearing loss is a multifactorial genetic disease described as the degeneration of the auditory function by aging. This study aimed to gather all the genes associated with Presbycusis and identify the most important genes. Additionally, we determined the most critical biological function linked to this disorder.
The genes related to Presbycusis were systematically collected by literature review via PubMed advanced rearch. We then selected articles that showed a significant association between various genes and presbycusis (p-value<0.05). We utilized the GeneMANIA Cytoscape software and CytoHubba plugin within Cytoscape to explore the biological connections between these genes, create a visual gene network, identify crucial genes and also determine most important biological function. The gene network was created by GeneMANIA algorithm based on protein-protein interaction, genetic interaction, co-localization and pathway.
We found 78 genes related to Presbycusis. "Signal transduction in absence of ligand" with q-value 0.009 is the most significant biological function. The enrichment score is 28.31. The network revealed that Protein-Protein interactions among Presbycusis-related proteins is the most prevalent biological connection (78%). Additional the five top gene hubs according to the CytoHubba (MCC method) are: *PTK2, CTH, DCLK1, GRM7* and *TLR4*.
This study shows that *PTK2* gene is the most important gene in Presbycusis. Our study provides insight about the prioritization of the genes and biological function related to Presbycusis.
Keywords: Biological function; gene network; GeneMANIA; Presbycusis

References

[1] Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. 2010;26:2927-8. https://doi.org/10.1093/bioinformatics/btq562

[2] Montojo J, Zuberi K, Rodriguez H, Bader GD, Morris Q. GeneMANIA: Fast gene network construction and function prediction for Cytoscape. F1000Research. 2014;3:153. https://doi.org/ 10.12688/f1000research.4572.1.

[3] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology. 2014;8 Suppl 4:S11. https://doi.org/ 10.1186/1752-0509-8-S4-S11

[4] Fujimoto, C. and T. Yamasoba (2014). "Oxidative stresses and mitochondrial dysfunction in age-related hearing loss." Oxid Med Cell Longev 2014: 582849.

ICB12-1402-1186

Isolation and computational investigation of class 3 L- asparaginase form a native *Glutamicibacter*. SP.

Mahdis Mofidi [1], Mohammad Yaghoubi-Avini[1], Gholamhossein Ebrahimipour[1] and Maryam Azimzadeh Irani[1]*
[1]Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
*Corresponding author's email address: m_azimzadeh@sbu.ac.ir

Abstract: L- asparaginases are important therapeutic drugs. Type II asparaginases are still used to treat childhood leukemia. Although these enzymes have been successfully used in treatment, they are associated with many side effects [1], so much research directed to find alternative enzymes. The soil samples were collected from agricultural fields near Saveh, Markazi province. After screening on phenol red agar, the best isolate identified by molecular and biochemical methods as *Glutamicibacter* sp. The specific primers that cover the full length of the asparaginase gene were designed based on complete genomes of the nearest type strains. After amplification and sequencing, the translated amino acid sequence aligned with 28 full-length sequences of type I and II from class 1 and class 3 L-asparaginases through the MAFFT algorithm and curated by the BMGE method on the online server NGPhylogeny (http://www.NGPhylogeny.fr). The phylogenetic tree was created using the maximum likelihood method based on the Le Gascuel model by Mega-7 software [2]. The phylogenetic tree shows the clustering of new isolate L-asparaginase with newly found *Rhizobium etli* class 3 asparaginases of type IV and V. Bootstrap values ensure that their relationship is strong. The three-dimensional structure of the enzyme was predicted by the ColabFold v1.5.5 server [3]. The structure of asparagin was constructed by ChemDraw v.17.3 and optimized using MM2 calculation available in ChemBio3D 17.3 [4]. Autodock Vina via the Chimera platform was used for docking. The binding energy for the most favorable structure based on docking score are -4.6, -4.4, and -4.4 Kcal/mol for new asparaginase, ReAV, and ReAIV, respectively. The new isolated L- asparaginase from *Glutamicicbacter* seems a promising target for new drug development. However, future in-vivo and in-vitro studies are essential to illustrate more about this target.

Keywords: L-asparaginase; Acute lymphoblastic leukemia; Phylogenetic analysis; Molecular Docking.

References

[1] Beckett, A. and D. Gervais, *What makes a good new therapeutic L-asparaginase?* World Journal of Microbiology and Biotechnology, 2019. 35(10): p. 152.
[2]Kumar, S., G. Stecher, and K. Tamura, *MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.* Molecular biology and evolution, 2016. 33(7): p. 1870-1874.
[3]Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nature methods, 2022. 19(6): p. 679-682.
[4]Halford, B., *Reflections on CHEMDRAW.* Chemical & Engineering News, 2014. 92(33): p. 26-27.

ICB12-1402-1187

An Innovative Unbalanced ANOVA-Based Approach for Multi-Dimensional Feature Evaluation in Brain-Computer Interface Classification

[1] *Morvarid Ghafouri, [2]Ali Roozbahani, [3]Ali Fallah , and [4]Saeed Rashidi
1,3,4 Department of Biomedical Engineering, AmirKabir University of Technology
2 Department of Electrical Engineering, Shahid Beheshti University of Technology
*Corresponding author's email address: morvarid_ghafoury@yahoo.com

Abstract: Brain-computer interface (BCI) stands as a crucial tool in processing, and facilitating communication with intellectually disabled individuals by leveraging brain characteristics through EEG signal analysis. Effective classification, however, necessitates the use of suitable features. Commonly employed tools for evaluating selected features include Fisher Score, MI, and DBI. While these methods sequentially determine the best features, the BCI field predominantly relies on features derived from the Common Spatial Pattern (CSP) method, utilizing matrix decomposition under eigenvalues and vectors. Yet, the ordinal mode of CSP's feature selection may not consistently minimize classification errors for feature vectors exceeding two dimensions. To address this, we propose a measure based on ANOVA statistical analysis of variance, capable of evaluating diverse features, including multi-dimensional vectors. In this research, we applied this measure to BCI Competition III and Part Iva data for simulation. With two classes, hand and leg movement imagery unlabeled data segments were omitted. Our analysis considered variations in subjects, classes, and trials, addressing the imbalance caused by noise and motion factors in BCI data. Utilizing UF-ANOVA variance analysis, we evaluated features extracted from the CSP algorithm. Mahalanobis Distance gauged the distance of each feature vector from the first-class distribution. UF-ANOVA results yielded three F factorial values (2 each for class changes, subject changes, and interaction changes), forming the basis for our criterion and index definitions related to factorial class, factorial interaction, and the opposite of factorial subject. The obtained p-values for all three cases were remarkably low, underscoring the significance of characteristic data related to subject changes, hand or foot movement perception, and their interaction. This novel ANOVA-based measure demonstrates its versatility by accommodating multi-dimensional feature vectors, offering a robust approach to feature evaluation in the dynamic domain of BCI.

Keywords: UF-ANOVA variance analysis; Multi-dimensional feature vectors; Feature selection; Brain-Computer Interface.

References

[1]J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, "Brain-computer interfaces for communication and control, " Clin. Neurophysiol., 113 (6), 767-791 (2002).

[3] K.K. Ang, Z.Y. Chin, H. Zhang, C. Guan, "CSP optimization for EEG-based brain-computer interface (BCI) system, " in Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 4188-4191 (2008).

[4] A. Goljahani, A. Gharaviri, F. Marvasti, "Fisher Score: A local search criterion for feature selection, " Pattern Recognit. Lett., 34 (10), 1056-1063 (2013).

[5] A. Kraskov, H. Stögbauer, P. Grassberger, "Estimating mutual information, " Phys. Rev. E, 69 (6), 066138 (2004).

[7] W.C. Hsieh, K.T. Tang, "An entropy-based feature selection method for pattern recognition, " Pattern Recognit., 39 (2), 249-259 (2006).

ICB12-1402-1188

In silico evidence for the efficacy of egg peptides in type 2 diabetes treatment

Delaram Shahsavari [1], Soodeh Mahdian [1], Roya Hosseini [2 3 4 5], Maryam Shahhosseini [4 5 6 7]
[1] Faculty of Sciences and Advanced Technologies in Biology, University of Science and Culture, Tehran, Iran
[2] Department of Endocrinology and Female Infertility at Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran
[3] Department of Andrology, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran
[4] Reproductive Epidemiology Research Center, Royan Institute for Reproductive, Biomedicine, ACECR, Tehran, Iran
[5] Department of Basic and Population Based Studies in NCD, Reproductive Epidemiology Research Center, Royan Institute, ACECR, Tehran, Iran
[6] Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR , Tehran, Iran
[7] Department of Cell and Molecular Biology, School of Biology, College of Science, University of Tehran, Tehran, Iran.

Abstract: Type 2 diabetes is a chronic and prevalent disease that is associated with increased blood glucose levels and insulin resistance. A large percentage of people in society are susceptible to this disease. In this disease, inflammatory cytokines such as TNF-α play a crucial role in a large number of inflammatory processes in the body, because increased levels of TNF-α lead to inflammation and insulin resistance, which is one of the prominent features of type 2 diabetes. Scientific evidence shows that proteins found in egg whites may have a modulating effect on diabetes. Based on this, the aim of this study was to investigate the effectiveness of egg peptides in inhibiting TNF-α using computational methods. Based on this, the main egg proteins were extracted from the Uniprot database and converted into 15-amino acid peptides using the Genscript database. Subsequently, the Pepcalc and Expasy databases were utilized for screening based on the charge, solubility and stability of the peptides. Finally, we reached 45% of the egg peptides that were able to pass these three stages of screening. The peptides obtained from the screening were evaluated for docking with TNF_α using Hpepdock software. The docking results revealed that 40% of the peptides obtained from the main egg proteins had stability, favorable solubility and suitable binding energy with ΔG less than -200 KJ/mol. Our computational results showed for the first time that the peptides in egg may have the potential to inhibit TNF-α and may be effective in reducing insulin resistance and modulating diabetes. This study presents novel insights, indicating that peptides derived from egg proteins may have the potential to modulate TNF-α, highlighting a promising direction for further research in the treatment and control of type 2 diabetes.

Keywords: type 2 diabetes , egg peptides , TNF_ alpha , Computational methods

ICB12-1402-1190

Berberine and Liver Cancer: Analyzing Gene Expression for Therapeutic Clues

F. Zeinali, M. Safari-Kharkeshi, Gh. Hossein *
Department of Animal Biology, Developmental Biology group, School of Biology, University College of Science, University of Tehran, Tehran, Iran
*Corresponding author's email address: ghossein@ut.ac.ir

Abstract: Liver cancer has a substantial global impact, ranking as the sixth most frequently diagnosed cancer and the third leading cause of cancer-related deaths as of 2020 estimates [1]. Hepatocellular carcinoma (HCC), a prevalent form of liver cancer, is characterized by the excessive growth of malignant cells, presenting challenges due to high incidence rates, unfavorable prognosis, and therapeutic limitations, notably severe adverse reactions to synthetic chemotherapeutic compounds [2]. To address these challenges, ongoing experimental research explores natural herbal-based compounds. Notably, certain bioactive compounds, like alkaloids, have shown significant benefits in treating various diseases. Berberine (BBR), an isoquinoline alkaloid derived from medicinal plants like Berberis species, holds promise for investigating liver cancer [2, 3]. This study analyzed gene expression associated with HepG2 cell lines and HepG2 cells treated with BBR for 24 hours using microarray data (GSE126132) through R programming. Differentially expressed genes (DEGs) in the two cell line groups were identified, and interactions between DEGs were explored using String software. Hub genes were determined through Centiscape and Cytohubba, resulting in 20 hub genes associated with key signaling pathways like TNF, Wnt, Hedgehog, MAPK, IL-4, -3, -17, and Insulin. These genes play a crucial role in Type II diabetes mellitus, Non-alcoholic fatty liver disease, Hepatitis B, Oncostatin M signaling, and transcription factor regulation in adipogenesis, all influential in HCC. Consequently, BBR emerges as a potential candidate for future drug formulations in liver cancer treatment, especially HCC.

Keywords: Liver cancer; Hepatocellular carcinoma; Berberine; Gene expression; Signaling pathways.

References

[1] H. Rumgay, M. Arnold, J. Ferlay, O. Lesi, C. J. Cabasag, J. Vignat, I. Soerjomataram, "Global burden of primary liver cancer in 2020 and predictions to 2040," J. Hepatol., 77 (2022) 1598-1606.

[2] Y. Liu, E. O. Nwafor, Z. Li, J. Wang, X. Feng, H. Li, Z. Liu, "Effect of Berberine on Liver Cancer," J. Nat. Prod. Commun., 17 (2022) 1934578X221102032.

[3] H. Majidzadeh, M. Araj-Khodaei, M. Ghaffari, M. Torbati, J. E. N. Dolatabadi, M. R. Hamblin, "Nano-based delivery systems for berberine: A modern anti-cancer herbal medicine," Colloids. Surf. B., 194 (2020) 111188.

ICB12-1402-1191

A Computational Prediction of Novel PET-Degrading Enzymes: Pathway to Sustainable Plastic Management

Donya Afshar Jahanshahi[a,b] , Mohammad Reza Rezaei Barzani[b] , Shohreh Ariaeenejad[c *], Kaveh Kavousi[b *]
[a] Department of Bioinformatics, Kish International Campus, University of Tehran, Kish, Iran.
[b] Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran.
[c] Department of Systems and Synthetic Biology, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.
[*]Corresponding authors: kkavousi@ut.ac.ir , shariaee@gmail.com

Abstract: Polyethylene terephthalate (PET) is a commonly employed polyester owing to its favorable characteristics and affordability. Nevertheless, PET plays a substantial role in escalating plastic waste pollution [1],[2]. Despite its crucial role, this has severe environmental consequences due to the slow PET degradation rates. The rise of PET poses a growing environmental threat, endangering diverse life forms including humans. Existing methods for managing plastic waste are notably inefficient, emphasizing the necessity for creative solutions, such as PET biodegradation. Enzymes capable of degrading plastic, known as plastizymes [3], primarily target high-molecular-weight polymers, such as polyethylene terephthalate (PET). However, isolating these PET-degrading enzymes remains a challenge, primarily due to the cultivation difficulties associated with many microorganisms. This study embarks upon this environmental challenge. Our research employed an innovative computational strategy to identify novel PET-degrading enzymes that break down PET. Given the limited dataset of known PET-degrading enzymes, we employed a Generative Adversarial Network model to augment the dataset using synthetic but realistic protein sequences. For an in-depth understanding of these enzymes, we extracted feature embeddings from the evolutionary scale model to facilitate the training of multiple classifiers, including the Support Vector Machine, K-Nearest Neighbors, and Random Forest models. These results have led to the identification of novel PET-degrading enzymes. Their validation was further supported by the analysis of the active sites, crucial amino acid compositions, and 3D structure comparison. Our study establishes a stage for substantial advancements in the plastic degradation industry by employing computational methodologies for plastizyme prediction, and emphasizes the potential of metagenomic approaches in environmental remediation research.
Keywords: metagenomics; machine learning; PET-degrading enzymes; computational prediction

References

[1] V. M. Pathak and Navneet, “Review on the current status of polymer degradation: a microbial approach,” *Bioresour. Bioprocess.*, vol. 4, no. 1, 2017, doi: 10.1186/s40643-017-0145-9.

[2] G. Lear *et al.*, “Plastics and the microbiome: impacts and solutions,” *Environ. Microbiomes*, vol. 16, no. 1, pp. 1–19, 2021, doi: 10.1186/s40793-020-00371-w.

[3] C. Charnock, “Norwegian soils and waters contain mesophilic, plastic-degrading bacteria,” *Microorganisms*, vol. 9, no. 1, pp. 1–18, 2021, doi: 10.3390/microorganisms9010094..

ICB12-1402-1192

Identification of key differentially expressed miRNAs and their potential target genes in Non-small cell lung cancer using bioinformatics analysis

Sepide Chodari Khameneh[a], Sara Razi[a]*
A Vira Pioneers of Modern Science (VIPOMS), Tehran, Iran
*Corresponding author's email address: s.razi63@gmail.com

Abstract Background: Non-small cell lung cancer (NSCLC) which is responsible for over 80% of lung cancer cases, is identified by a poor prognosis. Plenty of evidence shows that microRNAs play an important role in cancerous cell development by targeting a wide range of mRNAs. In this study, we aimed to find the most important miRNAs involved in NSCLC development as well as their target genes.

Methods: A microarray dataset containing miRNA profiles from the total blood of NSCLC samples and healthy controls was downloaded from the Gene Expression Omnibus (GEO) database. R software was used to investigate the alteration of miRNAs expression and identification of their target mRNAs. Moreover, Funrich 3.1.3, functional and pathway enrichment analyses were performed for the selected target genes. Then, the protein-protein interaction (PPI) network and hub-genes were elucidated using the STRING database and Cytoscape, separately. Moreover, the interactions between hub genes and dysregulated miRNAs were established.

Results: Differentially expressed miRNAs (DEmiRNAs) were identified from GSE17681 datasets, among which 13 downregulated and 11 upregulated DEmiRNAs were recognized in NSCLC samples. There were 208 potential predicted target genes, which 129 genes were targeted by up-regulated miRNAs, and 79 were targeted via down-regulated ones. While genes targeted by up-regulated miRNAs were significantly enriched in the FoxO family signaling, down-regulated miRNAs targeted genes were highly enriched in the TRAIL signaling pathway. 10 pivotal bub genes for down and upregulated DEmiRNAs were identified in the PPI network among which Myc and HSPa8 showed the highest node degree.

Conclusions: This study may provide new perspectives for future studies in the context of NSCLC by suggesting a list of differentially expressed miRNAs and their potential target genes that may serve as novel markers in both the diagnosis and treatment of NSCLC patients.

Keywords: Non-small cell lung cancer; Network analysis; Hub gene identification; MicroRNAs; DemiRs.

ICB12-1402-1193

Missense HRC- rs3745297 Gene Polymorphisms May Correlate to Humans Cardiac Arrhythmia: A *In silico* Study

Abbas Nikooei[1], Abasalt Hosseinzadeh Colagar[1]*, Mohammad-Taghi Hedayati[2]
1 Department of Molecular and Cell Biology, Faculty of Basic Science, University of Mazandaran, Babolsar, Iran.
2 Cardiology Department, Babol University of Medical Sciences, Babol, Iran.
*Corresponding author's email address: acolagar@yahoo.com; ahcolagar@umz.ac.ir

Abstract: The histidine-rich calcium-binding protein, which is encoded by the *HRC* gene in humans, plays a crucial role in cardiac arrhythmia development, especially in cases involving calcium-handling dysfunction [1]. This sarcoplasmic reticulum (SR) binding protein in the heart muscle cells, acts as an intracellular Ca^{2+} buffer and tightly regulates Ca^{2+} binding and release [2]. *In silico* studies can help to identify effective single nucleotide polymorphisms (SNPs) in the structure and stability of HRC protein and to predict their relationship with cardiac arrhythmias [3]. In this study, missense SNPs of the *HRC* gene and their effects on arrhythmia were investigated. At first, all missense SNPs of the *HRC* gene, which is located on chromosome 17q25.3 monitored. Missense SNPs with a minor allele frequency (MAF)≥ 0.1 were selected in the NCBI-dbSNP database. The effect of each of the selected SNPs based on functional, structural, and stability aspects of the protein were investigated by eleven online software: SIFT, Polyphen-2, Mutation assessor, PROVEAN, I-mutant, iStable, MUpro, SNPs&GO, PhD-SNP, HOPE, PSIPRED, and GOR-IV. Analysis of missense SNPs performed with SIFT, Polyphen-2, and Mutation assessor software showed that rs3745297 (T>G, Ser96Ala), as a transversion mutation, could be a deleterious SNP. The prediction of the effects of this transversion by I-mutant, iStable, MUpro databases, PSIPRED and GOR-IV also showed that the substitution of Serin with Alanine at the residue region 96 may decrease the stability of the protein and change the secondary structure. Since serine 96 of this protein is a critical phosphorylation site for binding to Triadin (TRDN) [4], which is a regulatory protein for releasing Ca^{2+}, Ser96Ala substitution may prevent this binding and thus disrupt the HRC-TRDN pathway in individuals with mutant allele G. Based on this, the rs3745297 can be suggested as a molecular marker for further research in cardiac arrhythmia.

Keywords: Cardiac arrhythmia; *HRC* gene; *In silico* analysis; Missense SNPs; Triadin protein.

References

[1] Liu N, Li X, Fu Y, Li Y, Lu W, Pan Y, Yang J, Kong J. Inhibition of lung cancer by vitamin D depends on downregulation of histidine-rich calcium-binding protein. J Adv Res. 2020 Aug 27;29:13-22.

[2] Conte E, Dinoi G, Imbrici P, De Luca A, Liantonio A. Sarcoplasmic Reticulum Ca2+ Buffer Proteins: A Focus on the Yet-To-Be-Explored Role of Sarcalumenin in Skeletal Muscle Health and Disease. Cells. 2023 Feb 24;12(5):715.

[3] Yazar M, Özbek P. In Silico Tools and Approaches for the Prediction of Functional and Structural Effects of Single-Nucleotide Polymorphisms on Proteins: An Expert Review. OMICS. 2021 Jan;25(1):23-37.

[4] Pollak AJ, Haghighi K, Kunduri S, Arvanitis DA, Bidwell PA, Liu GS, Singh VP, Gonzalez DJ, Sanoudou D, Wiley SE, Dixon JE, Kranias EG. Phosphorylation of serine96 of histidine-rich calcium-binding protein by the Fam20C kinase functions to prevent cardiac arrhythmia. Proc Natl Acad Sci U S A. 2017 Aug 22;114(34):9098-9103.

ICB12-1402-1195

Transcriptomics-Based Computational Drug Repurposing Strategy For Identified Therapeutic Candidates Tuberculosis meningitis

Nayereh Abdali[a], Atena Vaghf [b], Zahra Taslimi[c], Shahram Tahmasebian[*d]
a Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
Email: na107abdali@yahoo.com
b Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
c Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
d Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
*Corresponding author's email address: stahmasebian@gmail.com

Abstract: Tuberculous meningitis is a devastating disease of the central nervous system. This disease mainly affects the meninges of the brain and spinal cord along with the adjacent brain parenchyma. Early diagnosis is important for successful treatment (1). Drug repurposing is a valuable alternative approach to discover new indications for approved or investigational drugs beyond their original indication (2). RNA sequencing (RNA-seq) is one of the effective methods to find the heterogeneous gene expression of diseases in response to certain drugs (3). Therefore, our study used a computational drug repurposing pipeline to discover candidate drugs by PD differential gene expression signatures derived from RNA sequencing data. The transcriptional profile of whole blood in children with and without tuberculous meningitis (TBM) was compared, under the accession code GSE111459 were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). A total of 43 whole blood samples including 15 TBM cases and 24 uninfected healthy individuals were sequenced (4). Differentially expressed genes (DEGs) between blood samples of people with tuberculous meningitis and blood samples of healthy and non-infected people were obtained using GEO2R. Then, the Library of Integrated Network-Based Signatures (LINCS) database was used to identify potential drugs that can reverse the expression of DEGs. Then, by reviewing the significant literature and Drug bank studies (https://go.drugbank.com), the top-ranked drugs with the highest p-value were selected. This study identified 252 genes that were generally affected by the disease, among which genes with $|\log2FC| > 1$ and P-value <0.05 were identified as DEGs: 109 up-regulated genes and 143 down-regulated genes. Lisofylline (LSF) is a synthetic small molecule with novel anti-inflammatory properties. LSF can effectively prevent type 1 diabetes in clinical models and improve the function and survival of isolated or transplanted pancreatic islets (5).

Keywords: RNA sequencing; Tuberculous meningitis; Drug repurposing.

References

[1] Garg RK. Tuberculous meningitis. Acta Neurologica Scandinavica. 2010 Aug;122(2):75-90.
[2]Stott, S.R., R.K. Wyse, and P. Brundin, Drug repurposing for Parkinson's disease: the international linked clinical trials experience. Frontiers in Neuroscience, 2021. 15: p. 653377.
[3]Momeni, K., et al., Unraveling the complexity: understanding the deconvolutions of RNA-seq data. Translational Medicine Communications, 2023. 8(1): p. 21.
[4]Rohlwink UK, Figaji A, Wilkinson KA, Horswell S, Sesay AK, Deffur A, Enslin N, Solomons R, Van Toorn R, Eley B, Levin M. Tuberculous meningitis in children is characterized by compartmentalized immune responses and neural excitotoxicity. Nature communications. 2019 Aug 21;10(1):3767.
[5]Yang Z, Chen M, Nadler JL. Lisofylline: a potential lead for the treatment of diabetes. Biochemical pharmacology. 2005 Jan 1;69(1):1-5.

ICB12-1402-1196

Transcriptomics-Based Computational Drug Repurposing Strategy For Identified Therapeutic Candidates Tuberculosis meningitis

Nayereh Abdali[a], Atena Vaghf [b], Zahra Taslimi[c], Shahram Tahmasebian*[d]
a Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
Email: na107abdali@yahoo.com
b Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
c Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
d Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
*Corresponding author's email address: stahmasebian@gmail.com

Abstract: Tuberculous meningitis is a devastating disease of the central nervous system. This disease mainly affects the meninges of the brain and spinal cord along with the adjacent brain parenchyma. Early diagnosis is important for successful treatment (1). Drug repurposing is a valuable alternative approach to discover new indications for approved or investigational drugs beyond their original indication (2). RNA sequencing (RNA-seq) is one of the effective methods to find the heterogeneous gene expression of diseases in response to certain drugs (3). Therefore, our study used a computational drug repurposing pipeline to discover candidate drugs by PD differential gene expression signatures derived from RNA sequencing data. The transcriptional profile of whole blood in children with and without tuberculous meningitis (TBM) was compared, under the accession code GSE111459 were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). A total of 43 whole blood samples including 15 TBM cases and 24 uninfected healthy individuals were sequenced (4). Differentially expressed genes (DEGs) between blood samples of people with tuberculous meningitis and blood samples of healthy and non-infected people were obtained using GEO2R. Then, the Library of Integrated Network-Based Signatures (LINCS) database was used to identify potential drugs that can reverse the expression of DEGs. Then, by reviewing the significant literature and Drug bank studies (https://go.drugbank.com), the top-ranked drugs with the highest p-value were selected. This study identified 252 genes that were generally affected by the disease, among which genes with $|\log2FC| > 1$ and P-value <0.05 were identified as DEGs: 109 up-regulated genes and 143 down-regulated genes. Lisofylline (LSF) is a synthetic small molecule with novel anti-inflammatory properties. LSF can effectively prevent type 1 diabetes in clinical models and improve the function and survival of isolated or transplanted pancreatic islets (5).

Keywords: RNA sequencing; Tuberculous meningitis; Drug repurposing.

References

[1]Garg RK. Tuberculous meningitis. Acta Neurologica Scandinavica. 2010 Aug;122(2):75-90.
[2]Stott, S.R., R.K. Wyse, and P. Brundin, Drug repurposing for Parkinson's disease: the international linked clinical trials experience. Frontiers in Neuroscience, 2021. 15: p. 653377.
[3]Momeni, K., et al., Unraveling the complexity: understanding the deconvolutions of RNA-seq data. Translational Medicine Communications, 2023. 8(1): p. 21.
[4]Rohlwink UK, Figaji A, Wilkinson KA, Horswell S, Sesay AK, Deffur A, Enslin N, Solomons R, Van Toorn R, Eley B, Levin M. Tuberculous meningitis in children is characterized by compartmentalized immune responses and neural excitotoxicity. Nature communications. 2019 Aug 21;10(1):3767.
[5]Yang Z, Chen M, Nadler JL. Lisofylline: a potential lead for the treatment of diabetes. Biochemical pharmacology. 2005 Jan 1;69(1):1-5.

ICB12-1402-1197

In-silico Strategy for the Affinity Improvement of the Anti-TNF-α Antibody Drug by Site-Directed Mutagenesis Approach

Mahsa Mirzaei[a*], Mohammad Mehdi Heidari[a]
a Department of Biology, Yazd University, Yazd, Iran.
*Corresponding author's email address: ma_mirzaei76@yahoo.com

Abstract: Rheumatoid arthritis (RA) is an autoimmune disorder and one of the most common immune-mediated diseases [1, 2]. In patients with rheumatoid arthritis, significant damage is usually caused in the joints of the hands and feet. Monoclonal antibody (mAb) is a type of targeted drug therapy for the treatment of RA. These drugs specifically target tumor necrosis factor-alpha (TNF-α) and other cytokines that are closely related to RA. This approach reduces the side effects of traditional treatments [3]. Adalimumab is an anti-TNF-α mAb and used for RA treatment [4] .Complementary determining regions (CDRs) from the Adalimumab are regions that bind to the TNF-α .The type of amino acids at the site of TNF-α and CDR interaction is important. Site-directed mutagenesis is one of the methods which is used to improve the affinity of mAbs [5, 6]. In this study, the structure of TNF-α in complex with Adalimumab was extracted from 3dw5 PDB code. SabDab server was applied to determine the sequences of CDRs. Moreover, SWISS-MODEL server was applied for modeling the three-dimensional structure of anti-TNF-α mAb, TNF-α and mutated anti-TNF-α mAbs .Pymol software was used to determine the mAb-epitope binding site. Then, HADDOCK server was used to investigate the interaction between mAbs and TNF-α .HADDOCK results showed that some of the mutations such as the substitution of serine 106 with arginine improved binding affinity. The effect of this mutation was also investigated with the mCSM-AB2 web server and the binding affinity improvement was confirmed. Strategies for optimization of therapeutic mAbs can be applied to improve the function of mAbs.

Keywords: Site-directed mutagenesis; Affinity; Monoclonal antibody; Rheumatoid arthritis; Bioinformatics tools.

References

[1] Korani S, Korani M, Butler AE, Sahebkar A. Genetics and rheumatoid arthritis susceptibility in Iran. Journal of cellular physiology. 2019 May;234(5):5578-87.
[2] Gravallese EM, Firestein GS. Rheumatoid Arthritis—Common Origins, Divergent Mechanisms. New England Journal of Medicine. 2023 Feb 9;388(6):529-42.
[3] Chi XK, Xu XL, Chen BY, Su J, Du YZ. Combining nanotechnology with monoclonal antibody drugs for rheumatoid arthritis treatments. Journal of Nanobiotechnology. 2023 Mar 25;21(1):105.
[4] Hu S, Liang S, Guo H, Zhang D, Li H, Wang X, Yang W, Qian W, Hou S, Wang H, Guo Y. Comparison of the inhibition mechanisms of adalimumab and infliximab in treating tumor necrosis factor α-associated diseases from a molecular view. Journal of Biological Chemistry. 2013 Sep 20;288(38):27059-67.
[5] Janeway Jr CA. Immunobiology the immune system in health and disease. Artes Medicas. 1997.
[6] Kim HY, Stojadinovic A, Izadjoo MJ. Affinity maturation of monoclonal antibodies by multisite-directed mutagenesis. Monoclonal Antibodies: Methods and Protocols. 2014:407- 20.

ICB12-1402-1198

Unraveling Key Pathways and Hub Genes in Oral Cancer: An Integrated Bioinformatics Analysis

Sogand Shayesteh, Naser Farrokhi*
Department of Cell & Molecular Biology, Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran
*Corresponding author's email address: n_farrokhi@sbu.ac.ir

Abstract: Oral cancer, ranking as the sixth most prevalent cancer globally, poses a significant mortality threat, particularly in developing nations. To address this challenge, researchers have turned to RNA modulation as a promising avenue for cancer gene silencing. The focus on exosomes as carriers for targeted RNA delivery has gained traction due to their remarkable attributes, including low immunogenicity and the ability to overcome biological barriers [1]. A microarray analysis, utilizing the GSE146483 gene chip dataset, was conducted to identify differentially expressed genes (DEGs) associated with oral cancer. The selection criteria for DEGs were $\log_2FC \geq 2$ or ≤ -2 and P-value < 0.001 for upregulated and downregulated genes, respectively. The resulting dataset revealed 517 upregulated and 473 downregulated DEGs, providing a comprehensive view of the genetic alterations in oral cancer. Further analysis involved constructing a Protein-Protein Interaction (PPI) network and identifying hub genes through Cytoscape. The top 10 hub genes, identified via the degree algorithm, were *CDK1*, *BUB1B*, *CCNB1*, *CCNA2*, *AURKA*, *CDC45*, *BRCA1*, *TOP2A*, *BUB1*, and *EXO1*, emerged as potential key players in oral cancer pathogenesis. Functional annotation using Enrichr online tool and ShinyGO unveiled gene ontologies, emphasizing protein kinase binding and cyclin-dependent protein serine/threonine kinase (CDK) regulator activity. CDKs, crucial in cell cycle regulation, were implicated as potential contributors to uncontrolled cell growth in oral cancer. Furthermore, the study revealed that the highest score in known KEGG pathways was associated with the cell cycle, while pathways related to cellular processes and oocyte maturation, influenced by progesterone, exhibited elevated scores. Despite these findings, the multifactorial nature of oral cancer development, influenced by risk factors such as tobacco use, excessive alcohol consumption, HPV infection, and genetic predisposition, underscores the need for continued research to unravel the intricacies of this pathway and to identify potential therapeutic interventions [2].

Keywords: Cell cycle; Cyclin-dependent protein serine/threonine kinase; Oocyte maturation; Oral Cancer; Microarray analysis.

References

[1] Y, Kase, K. Uzawa, "Engineered exosomes delivering specific tumor-suppressive RNAi attenuate oral cancer progression", Scientific Reports, 11(1) (2021) 5897.

[2] S. Irani, "New insights into oral cancer—risk factors and prevention: a review of literature," International Journal of Preventive Medicine., 11 (2020).

ICB12-1402-1201

Simulation of the Ossicular Chain of the Middle Ear for Material Assignment Purpose

S. Piretarighat[ab], P. Keyhanvar*[cbd], F. Şahin[e], R. JavadRashid[f], M. Naderpour[g], Y. Jabbari Moghaddam[g], R. Salehi[b], S. Heydari Keshel[h]
a Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
b Department of Medical Nanotechnology, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
c Stem Cell Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
d HealthNBICS Group, Convergence of Knowledge and Technology to the benefit of Society Network (CKTSN), Universal Scientific Education and Research Network (USERN), Tabriz, Iran
e Department of Genetics and Bioengineering Department, Yeditepe University, Istanbul, Turkey
f Medical Radiation Sciences Research Group, Tabriz University of Medical Sciences, Tabriz, Iran
g Department of Otolaryngology, Head and Neck Surgery, Tabriz University of Medical Sciences, Tabriz, Iran
h Department of Tissue Engineering and Applied Cell Science, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: keyhanvarp@tbzmed.ac.ir

Abstract: The ossicular chain (OC) of the middle ear (ME) is a complex structure that transfers sound vibrations from the tympanic membrane (TM) to the inner ear. Conductive hearing loss can be caused by OC abnormalities due to several reasons such as trauma, infection, and cholesteatoma (1, 2). A common treatment for these defects is the reconstruction of the OC through surgery using a prosthesis. However, this procedure has some drawbacks, including difficulty in proper sizing of the prosthesis, underlying diseases, rejection, and rupture of the TM. The existence of different types and designs of OC prostheses shows that no optimal prosthesis has been made (3). Therefore, producing a personalized OC prosthesis is a requirement. Another consideration in the construction of the OC prosthesis is the material properties of the prosthesis. The material properties of the ossicles affect the impedance-matching function and the transmission efficiency of the ME. In this study, we aim to simulate the OC of the ME for material assignment purposes. After achieving the computed tomography (CT) scan images of the temporal bone, multiplanar reconstruction (MPR) of the ME ossicles was implemented. After that, the OC was simulated using Materialise Mimics and Materialise 3-Matic. We constructed a three-dimensional (3D) model of the human middle ear OC based on anatomical data. The comparison of the ossicles' dimensions with morphometric reports of the ossicles confirmed the similarity of these two groups. This structure is utilized to assign different material properties to the ossicles. Furthermore, the personalized prosthesis can be fabricated by using these simulated structures. Our simulation model can provide insights into the biomechanics of the ME, investigation the effects of different material parameters on the sound transmission and the stress distribution in the ossicles, and help design better prosthetic devices for hearing restoration in personalized medicine.
Keywords: Middle ear; Ossicular Chain; Ossicle; Personalized Medicine; Prosthesis; Simulation

References

[1] G. Volandri, F. Di Puccio, P. Forte, S. Manetti, "Model-oriented review and multi-body simulation of the ossicular chain of the human middle ear," Medical engineering & physics., 34(2012):1339-1355.

[2] H. Gray, W. Lewis, "Anatomy of the human body," 21 Ed., Philadelphia: Lea & Febiger, 2000.

[3] R. Gurgel, HR. Harnsberger, "Imaging in Otolaryngology E-Book," Elsevier Health Sciences; 2018.

ICB12-1402-1203

Bioinformatic Analysis of P53 Mutations in Cancer

T. Mohammadpour, R. Mohammadzadeh *
Department of Cellular and Molecular Biology, Faculty of Basic Science, University of Maragheh, Maragheh, Iran
* Corresponding author's email address: rmohamadzadeh@maragheh.ac.ir

Abstract: The P53 gene plays a crucial role as a tumor suppressor, maintaining genomic stability and inhibiting tumor formation. However, P53 mutations, the most common genetic alterations in human cancers, lead to the dysregulation of cell growth and survival, contributing to aggressive tumor behavior and resistance to conventional treatments. This study aims to provide a bioinformatic analysis of P53 mutations in cancer using computational tools such as MutPred2, SNPeffect4.0, I-Mutant3.0, ConSurf, Phyre2, and project HOPE. We investigated the structural consequences of P53 mutations, including their impact on protein stability, examined structural changes, and evaluated evolutionary conservation and protein secondary structures. We obtained a dataset of P53 tumor variants from the TP53 Database, specifically focusing on variants with mutation rates exceeding 1. This dataset included 54 missense and 4 stop-gained variants that were further analyzed and assessed for their significance in cancer. Understanding the consequences of these mutations is crucial for guiding treatment decisions, including the development of targeted therapies. By restoring P53 function and targeting mutant P53, new treatment modalities can address therapeutic limitations and enhance treatment efficacy. Our findings highlight the critical role of P53 mutations in different cancer types. This study provides valuable insights into the specific P53 mutations in cancer, shedding light on their molecular and structural consequences and highlighting the potential for tailored treatment approaches. These findings have significant implications for precision medicine, as they pave the way for the development of innovative therapies and improved patient care in P53-mutated cancers.
Keywords: P53 mutations; Cancer; Bioinformatics analysis; Structural consequences; TP53 Database

ICB12-1402-1204

Investigating the Association between NID1 Gene SNPs with Mandibular Prognathism: An In silico Study

Negar Gorjian[1], Abasalt Hosseinzadeh Colagar[1]*, Valiollah Arash [2]
1. Department of Molecular and Cell Biology, Faculty of Science, University of Mazandaran, Babolsar, Iran.
2. Department of Orthodontics, Babol University of Medical Sciences, Babol, Iran.
*Corresponding author's email address: n.gorjian03@umail.umz.ac.ir; ahcolagar@umz.ac.ir

Abstract: Class III malocclusion is a special type of jaw deformity that can result mandibular prognathism (MP) [1]. Genetic factors such as single nucleotide polymorphisms (SNPs) can cause this protrusion. Nidogen-1 is an extracellular glycoprotein encoded by the *NID1* gene and plays a crucial role in stabilizing, cell adhesion, and mediating cell-matrix interactions. *In silico* studies can help to identify effective SNPs in the structure and stability of Nidogen-1 and to predict their relationship with MP. In this study, missense SNPs of the *NID1* gene and their effects on MP were investigated. At first, all missense SNPs of the *NID1* gene monitored. Missense SNPs with a minor allele frequency (MAF)≥ 0.1 were selected in the NCBI-dbSNP database. The effect of each selected SNPs based on functional, structural, and stability aspects of the protein and mRNA were investigated by twelve online software: SIFT, Polyphen-2, Meta LR, PantherDB, PROVEAN, I-mutant, iStable, MUpro, FATHMM, RNAsnp, PSIPRED, and GeneMANIA. Analysis of missense SNPs by SIFT, Polyphen-2, and PantherDB showed that rs3738531 (C>A, Gln807His), could be as a deleterious SNP. The prediction of the effects of this SNP by I-mutant, iStable, MUpro, and PSIPRED also showed substitution of Gln807His may decrease the stability of the protein. Analysis association of *NID1* by GeneMANIA showed that this protein has a co-expression and Physical Interactions with the heparan sulfate proteoglycan 2 (HSPG2), which is known as Perlecan, with a score of 0.013352894 [2]. Perlecan and Nidogen-1 impact chondrocyte and osteoblast functions [3]. Given their roles, the co-expression of *HSPG2* and *NID1* may have implications for MP. Furthermore, rs3738531 has been associated with structural changes in mRNA with a significant P-value, potentially emphasizing the translation of *NID1*. Based on this study, the *NID1* gene plays an important role in MP development and rs3738531 is expected to be effective in mandibular growth.

Keywords: In silico analysis; Mandibular prognathism; Missense SNPs; Nidogen-1; Perlecan**.**

References

[1] Martinez, Jerahme R., et al. "Perlecan/HSPG2: Signaling role of domain IV in chondrocyte clustering with implications for Schwartz-Jampel Syndrome." *Journal of cellular biochemistry* 120.2 (2019): 2138-2150.

[2] Chen, Xiaojun, et al. "Whole-exome sequencing reveals rare germline mutations in patients with hemifacial microsomia." *Frontiers in Genetics* 12 (2021): 580761.

[3] Doraczynska-Kowalik, Anna, et al. "Genetic factors involved in mandibular prognathism." *Journal of Craniofacial Surgery* 28.5 (2017): e422-e431.

ICB12-1402-1205

Tracing the controller genes related to the metastasis mechanisms in squamous cell carcinoma of the head and neck

Nasibeh Khayer [a], Samira Shabani [b], Maryam Jalessi [a,c,*], Frouzandeh Mahjoubi [b]
[a] Skull Base Research Center, The Five Senses Health Institute, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[b] Department of Clinical Genetic, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran
[C] ENT and Head and Neck Research Center and Department, Hazrat Rasoul Hospital, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
* Corresponding author's email address: dr.jalessi@gmail.com

Abstract: Squamous cell carcinoma of the head and neck (HNSCC) is an aggressive malignancy, which is characterized by high morbidity. Approximately 650,000 new HNSCC cases are reported annually, out of which only 40–50% patients will survive for 5 years [1]. Distant metastasis is a major factor associated with poor prognosis, and in turn, reduced survival in HNSCC, but its underlying molecular mechanisms are not well-known [2]. In the absence of a robust biomarker, patients at higher risk for metastasis cannot be provided with the effective treatment. Therefore, a deeper understanding of such mechanisms is a crucial unmet need to identify essential genes that could serve as effective biomarkers. The objective of the present study is a comprehensive analysis of transcriptome data associated with metastasizing and non-metastasizing HNSCC, aiming to capture controller genes that are related to the metastasis mechanisms. For such purpose, we have conducted liquid association (LA) analysis to identify dynamic co-expression relations and their corresponding LA-scouting genes [3, 4]. Then, the biologically relevant relations were detected using computational methods. Subsequently, for each significant LA-scouting gene, the survival curves were calculated according to the Kaplan–Meier method and differences between curves were assessed using the log-rank test. The results indicated that the low-expression levels of *Dmtn*, *Camk2a*, *C19orf33* and *A4galt* are associated with shorter metastasis-free survival; on the contrary, the low-expression levels of *Usp13*, *Dffa* and *Fam181b* are associated with longer metastasis-free survival.

Keywords: Squamous cell carcinoma of the head and neck; Liquid association Analysis; Distant metastasis; Metastasis-free survival.

References

[1] Bray, F., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018. 68(6): p. 394-424.

[2] Shimizu, S., et al., Identification of a novel therapeutic target for head and neck squamous cell carcinomas: a role for the neurotensin-neurotensin receptor 1 oncogenic signaling pathway. International journal of cancer, 2008. 123(8): p. 1816-1823.

[3] Khayer, N., et al., Three-way interaction model with switching mechanism as an effective strategy for tracing functionally-related genes. Expert review of proteomics, 2019. 16(2): p. 161-169.

[4] Wang, L., et al., Meta-analytic framework for liquid association. Bioinformatics, 2017. 33(14): p. 2140-2147.

ICB12-1402-1206

Investigating the Association between SNPs in *AGT* Gene with Preeclampsia: A *In silico* Study

Sajedeh Latifi[1], Abasalt Hosseinzadeh Colagar[1]*, Mohammad Karimian[1], Mohammad-Taghi Hedayati[2]
[1] Department of Molecular and Cell Biology, Faculty of Science, University of Mazandaran, Babolsar, PC: 47416-95447, Mazandaran, Iran.
[2] Cardiology Department, Babol University of Medical Sciences, Babol, Mazandaran, Iran.
*Corresponding author's email address: sajedeh.ltf1997@gmail.com; ahcolagar@umz.ac.ir

Abstract: Preeclampsia (PE) is a pregnancy disorder characterized by hypertension and proteinuria [1]. The Angiotensinogen (AGT) protein is an important component of the rein-angiotensin system (RAS), a key regulatory system of blood pressure, which could be closely related to PE susceptibility [2]. *In silico* studies can help to identify effective single nucleotide polymorphisms (SNPs) in the structure and stability of *AGT* protein and could predict their association with PE. In this study, missense SNPs of the *AGT* gene and their effects on PE were investigated. At first, all common missense SNPs of the *AGT* gene monitored. Missense SNPs with a minor allele frequency (MAF)≥ 0.1 were selected in the NCBI-dbSNP database. The effect of each selected SNPs based on functional, structural, and stability aspects of the protein and mRNA were investigated by nine following online software: SIFT, Polyphen-2, PantherDB, PROVEAN, I-mutant, iStable, MUpro, HOPE, PSIPRED. Analysis of missense SNPs by SIFT, Polyphen-2, and PantherDB showed that rs4762 (G>A, Thr198Met) could be as a deleterious SNP. The prediction of the effects of this SNP by I-mutant, iStable, MUpro, and PSIPRED also showed substitution of Thr198Met may decrease the stability of the protein. Angiotensin protein is expressed in the liver and is cleaved by renin enzyme in response to blood pressure reduction. The resulting product, angiotensin I, is then cleaved by angiotensin-converting enzyme (ACE) to produce the physiologically active enzyme angiotensin II. Dysfunction of this protein can play a role in the pathogenesis of hypertension and preeclampsia. Based on this study, the *AGT* gene may be involved in the development of PE and rs4762 is expected to be effective in function of this gene.

Keywords: Angiotensinogen; Blood Pressure; *In silico* analysis; Missense SNPs; Preeclampsia

References

[1] Duley, Lelia. "The global impact of pre-eclampsia and eclampsia." *Seminars in perinatology*. Vol. 33. No. 3. WB Saunders, 2009.

[2] Galão, Adriani O., et al. "Angiotensin-converting enzyme gene polymorphism in preeclampsia and normal pregnancy." *American journal of obstetrics and gynecology* 191.3 (2004): 821-824.

ICB12-1402-1208

Hub gene identification and correlation with tumor-infiltrating immune cells in esophageal carcinoma; an RNA-seq and protein-protein interaction analysis

Alireza Ahmadi[a], Mohadeseh Ebrahimi[a, b], Mohammad Reza Tohidkia[b], Mohammad Khalaj-Kondori[*]
[a] Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran.
[b] Research Center for Pharmaceutical Nanotechnology, Tabriz University of Medical Sciences, Tabriz, Iran.
[*] Full Professor of Molecular Genetics, Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran; Khalaj49@gmail.com

Abstract: Esophageal cancer (ESCA) is one of the most malignant cancers with high incidence and mortality and poor overall prognosis [1]. Recent advances in understanding the genomic aspects led to using specific genomic alterations in ESCA tumors as biomarkers for early diagnosis, treatment, and prognosis of this cancer [2]. The tumor microenvironment, including tumor-infiltrating immune cells (TIIC), plays a vital role in immune evasion, proliferation, invasion, and metastasis [3]. Therefore, there is an emerging emphasis on identifying key genes, and prognostic biomarkers and performing new therapeutic targets.
The RNA-seq data of 197 ESCA patients were downloaded and normalized with the "GDCRNATools" package. A volcano plot of differentially expressed genes (DEGs) with $Log_2FC > 0.5$ and p-value < 0.05 was drawn using the "TCGAbiolinks" package. Genes with log FC > 2 and p-value > 0.0001 were chosen for KEGG pathways and gene ontology analysis using the "clusterProfiler" package. STRING database was used for protein-protein interaction (PPI) analysis and CytoHubba (a Cytoscape software plugin) was utilized for hub gene identification. Finally, the correlation of gene expression with TIICs was evaluated by the TIMER database.
Among 5219 DEGs, 891 genes were filtered, of which 45, 98, 90, and 41 genes were respectively enriched in the Cell cycle (pathway), organelle fission (Biological Process), chromosomal region (Cellular Component), and tubulin binding (Molecular Function). PPI and hub gene analysis of 45 genes involved in the cell cycle, identified CCNB1, CHEK1, BUB1B, and MAD2L1 as hub genes. BUB1B and CCNB1 had promising correlations (Rho = 0.61, 0.62, and p-value = 1.54E-19, 6.80E-21, respectively) with $CD4^+$Th2 infiltration which many studies have shown that increment of Th2 could lead to the rapid growth and metastasis of ESCA [4]. Altogether, it is suggested that high BUB1B, CCNB1, and Th2 expression could be associated with a poor prognosis and clinical outcomes.
Keywords: Esophageal carcinoma; Hub gene; Tumor-infiltrating immune cells; prognostic biomarker; Protein-protein interaction.

References

[1] Liu, C.-Q., et al., *Epidemiology of esophageal cancer in 2020 and projections to 2030 and 2040.* Thoracic Cancer, 2023. **14**(1): p. 3-11.
[2] Sheikh, M., et al., *Current status and future prospects for esophageal cancer.* Cancers, 2023. **15**(3): p. 765.
[3] Chen, Y., et al., *Characterization of the immune infiltration landscape and identification of prognostic biomarkers for esophageal cancer.* Molecular Biotechnology, 2023. **65**(3): p. 361-383.
[4] Pei, Y., et al., *The expression and clinical value of tumor infiltrating dendritic cells in tumor tissues of patients with esophageal cancer.* J Gastrointest Oncol, 2021. **12**(5): p. 1996-2003.

ICB12-1402-1209

Down-regulation of the DAPK3 gene is associated with the metastasis of papillary thyroid tumors to the lungs.

E. Sameri, SM. Javadirad*
Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: javadirad@yahoo.com

Abstract: Papillary thyroid carcinoma (PTC), the most frequent type of thyroid cancer and the one with a favorable prognosis, is the most common type of endocrine-related cancer. The overall 5-year relative survival rate has been reported as high as 97.5%, and only a tiny percentage of papillary carcinomas are aggressive. The lungs and bones are the most common sites for distant metastasis in PTC, which commonly metastasizes to regional lymph nodes. In the presence of distant metastases, patients with PTC have a worse prognosis. Death Associated Protein Kinase 3 (DAPK3) has been identified as a tumor suppressor whose downregulation promotes cell survival, proliferation, cellular aggregation, and resistance to chemotherapy. This study aimed to determine the role of DAPK3 in lung metastases from PTC by performing a meta-analysis of nine datasets selected from the gene expression omnibus (GEO) databank. PTC tissues from patients with and without lung metastasis were compared to identify differentially expressed genes (DEGs). A log2 transformation and quantile normalization were applied to each dataset, and outliers were eliminated using principal component analysis. Benjamini-Hochberg was used to adjust P-values. The limma package was used to identify DEGs. Our next step will be to conduct RT-qPCR reactions and analyze melt curves to assess the amplification's specificity. The Livak method will be used to analyze relative gene expression. Microarray meta-analysis revealed DAPK3 gene downregulation in PTC tissues of patients with lung metastasis compared to patients without metastasis (logFC = - 2.06, adj.p.val = 0.000896). PTC tissues with lung metastasis expressed lower DAPK3 gene expression than those without, suggesting that this gene plays a role in lung metastasis.

Keywords: thyroid cancer; PTC; differential expression analysis; DAPK3; lung metastasis

ICB12-1402-1210

Artificial Intelligence-based Classification of Glioblastoma Lifespan: Unveiling the Role of TMEM176A

Sina Farazmandi[A], Hadi Kamkar[A], Seyed Alireza Khanghahi[A], Parviz Abdolmaleki[A,*]
a Department of Biophysics, Faculty of Biological Science, Tarbiat Modares University, Tehran, Iran.
*Corresponding author's email address: parviz@modares.ac.ir

Abstract: Glioblastoma (GBM) is a prevalent and aggressive brain tumor with a grim prognosis, despite extensive research efforts [1, 2]. The Cancer Genome Atlas (TCGA) provides a valuable database of gene expression profiles from various cancer types, including GBM, along with clinical data and drug response information [3]. Artificial Intelligence (AI)-based algorithms are extensively utilized in biomarker discovery, particularly in analyzing gene expression data.
This study utilized data from TCGA, focusing on 156 glioblastoma samples. Patients were categorized into two groups based on their overall survival (OS): those surviving less or equal to 1000 days (shorter living batch) and else (longer living batch). The primary objective was to categorize individuals into short and long-life groups and identify a biomarker to distinguish between them. Initially, the count matrix encompassed 60,488 genes, and subsequent filtration resulted in 34,225 remaining genes per sample. Variance Stabilizing Transformation (VST) method using the "DESeq2" package was then employed for data normalization. The Minimum Redundancy Maximum Relevance (MRMR) algorithm selected 50 genes, and a genetic algorithm served as the secondary feature selection method. AI algorithms, specifically the naïve Bayes algorithm, were employed using the selected features. A stratified 10-fold cross-validation was implemented to ensure robustness, leading to a model accuracy of 0.703 and an AUC score of 0.674. Subsequent analyses, including feature importance assessment and permutation plots highlighted TMEM176A as crucial, showing a slight upregulation in the short-life group compared to the long-life group. To validate its significance, expression patterns were compared between tumor and healthy samples using cBioPortal, revealing significant overexpression in tumor samples.
In conclusion, this study harnessed AI algorithms to uncover TMEM176A as a promising biomarker for the diagnosis and prognosis of glioblastoma, presenting a potential avenue for improving outcomes in the management of this formidable brain tumor.
Keywords: Glioblastoma; Gene expression; Artificial intelligence; Biomarker discovery; TMEM176A.

References

[1]A. C. Tan, D. M. Ashley, G. Y. Lopez, M. Malinzak, H. S. Friedman, and M. Khasraw, "Management of glioblastoma: State of the art and future directions," *CA Cancer J Clin,* vol. 70, no. 4, pp. 299-312, Jul 2020, doi: 10.3322/caac.21613.

[2]J. Y. Nam and J. F. de Groot, "Treatment of Glioblastoma," *J Oncol Pract,* vol. 13, no. 10, pp. 629-638, Oct 2017, doi: 10.1200/JOP.2017.025536.

[3]S. Chawla *et al.*, "Gene expression based inference of cancer drug sensitivity," *Nat Commun,* vol. 13, no. 1, p. 5680, Sep 27 2022, doi: 10.1038/s41467-022-33291-z.

ICB12-1402-1211

Investigating the Inhibitory Effect of Curcumin on FGFR4: Insights from Docking Studies

Mohammad Erfan Shanehsazpoor shooshtari[a], Amirhossein Naghsh Nilchi[a], Fariba Dehghanian[a,*]
[a] Department of Cell and Molecular Biology & Microbiology, Faculty of Biological Science and Technology, University of Isfahan, HezarJarib Street, Isfahan 81746-73441, Iran
*Corresponding author's email address: fa.dehghanian@sci.ui.ac.ir

Abstract: Liver cancer is the second leading cause of cancer death worldwide in men and the sixth most frequent cause of cancer death in developed countries. Hepatocellular carcinoma (HCC) is a primary liver cancer originating in hepatocytes, the liver's main functional cells. Curcumin, as a natural substance known for ages demonstrates a diverse impact on both preventing cancer and complementing cancer treatments. The great advantage of using nutraceuticals of vegetable origin in comparison to popular cytostatic drugs is the minimized side effects and reduced toxicity. The role of the FGFR4 receptor in HCC is known. In this study, the inhibitory mechanism of ponatinib an effective FGFR4 inhibitor was investigated, and the inhibitory effect of the natural compound curcumin was examined by exploring its mechanisms. The PDB structure was obtained from the PDB database. The curcumin compound was also retrieved from PubChem. After completing the docking preparation steps using the UCSF Chimera program, docking was performed using Autodock Vina. The residues involved in the interaction between the ligand and the protein were determined by LigPlot^{+}. Curcumin docked with the apo form of FGFR4 (PDB ID: 4TYG), exhibiting a high binding energy (ΔG=-8.1 Kcal/mol), compared to ponatinib docked with the apo form of FGFR4 (ΔG=-10.5 Kcal/mol). Additionally, Lys57 residue was found to be common with residues associated with ponatinib. Furthermore, the result of curcumin binding to FGFR4 (PDB ID: 4TYJ) as the post-drug-binding conformation, revealed that curcumin shares common interactions with the 6 residues connected to ponatinib. In conclusion, this study illustrates the inhibitory effect of curcumin on FGFR4, rendering it a potential FGFR4 inhibitor. Nevertheless, further investigations are warranted to explore the inhibitory effects of curcumin on FGFR4 more comprehensively.

Keywords: Liver cancer, FGFR4, Curcumin, Ponatinib, Molecular Docking

References

[1] Gurung AB, Ali MA, Lee J, Farah MA, Al-Anazi KM. Molecular docking and dynamics simulation study of bioactive compounds from Ficus carica L. with important anticancer drug targets. Plos one. 2021 Jul 14;16(7):e0254035.

[2] Sia D, Villanueva A, Friedman SL, Llovet JM. Liver cancer cell of origin, molecular class, and effects on patient prognosis. Gastroenterology. 2017 Mar 1;152(4):745-61.

[3] Willenbacher E, Khan SZ, Mujica SC, Trapani D, Hussain S, Wolf D, Willenbacher W, Spizzo G, Seeber A. Curcumin: new insights into an ancient ingredient against cancer. International journal of molecular sciences. 2019 Apr 12;20(8):1808.

[4] Kim DB, Lee DK, Cheon C, Ribeiro RI, Kim B. Natural products for liver cancer treatment: From traditional medicine to modern drug Discovery. Nutrients. 2022 Oct 12;14(20):4252.

ICB12-1402-1212

Deep Learning Framework for Splice Site Prediction Across Multiple Species

Mohammad Reza Rezvan[1]*, Ali Ghanbari Sorkhi[1], Jamshid Pirgazi[1]
[1]Department of Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran
*Corresponding author's email: reza.rzvn1@gmail.com

Abstract: Splice site prediction remains a pivotal challenge in bioinformatics, necessitating accurate and efficient computational methods to understand genetic regulation and expression. This study introduces a novel deep learning framework for the prediction of splice sites, leveraging the genetic sequences from three distinct datasets: Arabidopsis thaliana, Homo sapiens, and HS3D. Our methodology commences with the preprocessing of sequence data using a two-gram approach followed by one-hot encoding to transform genetic sequences into a numerical format amenable to deep learning techniques. We employ a Residual Convolutional Neural Network (ResidualConv1D) for robust feature extraction, capitalizing on its ability to learn hierarchical representations of sequence motifs. To address the high-dimensionality of the feature space, Principal Component Analysis (PCA) is utilized for dimensionality reduction, enhancing computational efficiency and model interpretability. The feature-rich, dimensionally reduced data is then classified using a Support Vector Machine (SVM), chosen for its effectiveness in handling high-dimensional data and its capacity for achieving high accuracy in binary classification tasks. Our approach showcases a significant improvement in splice site prediction accuracy, demonstrating the potential of integrating deep learning architectures with traditional machine learning techniques for bioinformatics applications. The study not only contributes to the advancement of computational genomics but also opens new avenues for the application of deep learning in genetic data analysis.

Keywords: Splice site; ResidualConv1D; robust feature extraction; Principal Component Analysis; Support Vector Machine

ICB12-1402-1213

Bioinformatics Analysis of flavanone-3′-hydroxylase (F3′H) in *Silybum marianum*

Hoda Alsadat Aghili[a]*, Azim Ghasemnezhad[b] , Hossein Moradi[c] , Valiollah Ghasemi Omran[d]
a Horticulture Science, Plant Production, Gorgan University Agricultural Science and Natural Resources, Gorgan, Iran
b Horticulture Science, Plant Production, Gorgan University Agricultural Science and Natural Resources, Gorgan, Iran
c Horticulture Science, Crop Science, Sari University Agricultural Science and Natural Resources, Sari, Iran
d Crop Science Genetic & Agricultureal Biotecnology Institute of Tabarestan, Sari University Agricultural Science and Natural Resources, Sari, Iran
*Corresponding author's email address: hoda_aghili@yahoo.com

Abstract: Milk thistle (Silybum marianum), a member of the Asteraceae family, is known for its use in the treatment of liver diseases due to its main component, silymarin. The biosynthesis of silymarin involves the enzyme flavanone-3′-hydroxylase (F3'H), which is encoded by a gene consisting of 1557 nucleotides with a 51.87% cytosine-guanine content. The mRNA of this gene comprises 1317 base pairs, encoding a protein of 349 amino acids in length. The promoter sequence of the F3'H gene contains 19 predicted regulatory elements, with the TATA-box and CAAT-box being the most prominent. Additionally, the gene exhibits several light response regulatory elements. According to the ProtParam database, the protein encoded by this gene has an aliphatic index of 99.32, a predicted isoelectric point of 7.87, and a molecular weight of 48.320 KD. The GRAVY range was calculated to be -0.059. Furthermore, the protein contains 49 negatively charged residues (Asp and Glu) and 50 positively charged residues (Arg and Lys). Its extinction coefficient, measured in water at 280 nm, is 42190 M-1 cm-1. The protein was found to have an instability index of 34.65, classifying it as stable.

Keywords: Flavanone-3′-hydroxylase; Protein; Medicinal plants; Secondary metabolite; Silybum marianum.

References

[1] S. Drouet, D. Tungmunnithum, É. Lainé, , & C. Hano, "Gene expression analysis and metabolite profiling of silymarin biosynthesis during milk thistle (*Silybum marianum* (L.) ", International Journal of Molecular Sciences, 21 (2020) 4730.

[2] M. Torres, P. Corchete, "Gene expression and flavonolignan production in fruits and cell cultures of *Silybum marianum*". Journal of plant physiology, 192(2016), 111-117.

ICB12-1402-1214

Design of a Novel Ligand Derived from Nigella_Sativa Plant to Inhibit Voltage-Dependent Calcium Channel to Diminish Blood Pressure Using Bioinformatics Tools

Negar Ghaleh Navi[1]*
[1]Department of Biology, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
*Corresponding author Email: negarghalenavi@gmail.com

Abstract: Hypertension, a significant risk factor for heart attacks and premature death, affects a large portion of the global population. Many people rely on modern medications like calcium channel blockers (CCBs) to control their blood pressure.[1] CCBs work by blocking the entry of calcium ions into smooth muscle cells, resulting in vasodilation and lower blood pressure.[1] Recent research suggests that Thymol, derived from *Nigella_sativa*, can also inhibit Voltage-dependent L-type calcium channel (7UHF), leading to vasodilation and decreased blood pressure.[2,3] Likewise, Carvacrol, a monoterpenic phenol found in various herbal oils, has demonstrated cardiovascular effects such as fibrinolysis, vasorelaxation, and the reduction of blood pressure in previous studies.[2] This study aims to lower blood pressure with the new ligand compared to two antihypertensive ligands Thymol and Carvacrol.

This research project used PubChem and PDB to analyze the structure of Thymol, Carvacrol and 7UHF. In addition, the new ligand designed by ChemDrow and Chem3D. Molecular docking was screened using iGEMDOCK version 2.1.The calculated ligand receptor (protein) interaction energy is represented by docking scores (DOS). More negative scores indicate a stronger likelihood of binding.

The binding of the 7UHF protein to thymol and carvacrol was investigated. Their energy levels were quite low (-62.5 and -61.4). However, a new ligand called 3-isopropyl-6-methylbenzene-1,2-diol was designed, which showed significantly higher energy efficiency (-77.3) and better results compared to thymol and carvacrol.
Ligand properties: Chemical Formula: $C_{10}H_{14}O_2$, Exact Mass: 166/099, Molecular Weight: 166/220

The findings revealed that novel ligand which designed for the study can be a potent inhibitor of hypertension. Among them, the protein (7UHF) derivative may be the more effective for the treatment of the disease. Based on the findings. It is recommended that in-vitro and in-vivo studies be carried out to determine the efficacy of this ligand against hypertension disease.

Keywords: Thymol; Carvacrol; Drug-design; calcium channel; Hypertension

References

[1] Lu L, Liu M, Sun RR, et al.(2015) Myocardial infarction: symptoms and treatments. Cell Biochem Biophys. 72(3): 865-7.

[2] Maideen, Naina M P et al. “Nigella Sativa (Black Seeds), A Potential Herb for the Pharmacotherapeutic Management of Hypertension - A Review.” Current cardiology reviews vol. 17,4 (2021): e230421187786

[3] Alam, Mohd Aftab et al. “Effect of Nigella sativa and Fenugreek on the Pharmacokinetics and Pharmacodynamics of Amlodipine in Hypertensive Rats.” Current drug metabolism vol. 21,4 (2020): 318-325

ICB12-1402-1216

Identification of Key Genes and Shared Metabolic Pathways in Multiple Sclerosis and Bipolar Disorder Using Systems Biology Approaches

Shahang Shamshiri [ab], Shahadeh Shamshiri[ab], Dr.Ali Kazemipour[ab]*
a Facility of Mathematics and Computer Science, Department of Bioinformatics, Bahonar University, Kerman, Iran
b Bahonar Bioinformatics Lab (BBL),
shahang.sh@gmail.com, shahande.sh@gmail.com.
*Corresponding author's email address: Ali.kazemi@uk.ac.ir

Abstract: For many years, there have been controversial arguments regarding the genetic interplay between Multiple sclerosis (MS) and Bipolar Disorder (BD). MS is a prevalent autoimmune disease characterized by the immune system targeting the myelin sheath; as a complex disease, both environmental and genetic factors contribute to its development. BD is a complex mental health disease, is believed to stem from a combination of genetic and environmental influences. Noticeably, individuals with MS exhibit an elevated risk of bipolar spectrum disorders.

This study focuses on unraveling the genetic interplay between MS and BD. We employed a combined set of genes from both diseases and constructed the protein-protein interaction network to identify hub genes and shared metabolic pathways Using the Cytohubba plugin, g:Profiler, and KEGG databases, we identified 10 hub genes that are: TNF, IL6, IL1B, TLR4, AKT1, BCL2, IFNG, APP, NFKB1, and MTOR, all implicative in both MS and BD.

Evaluation of metabolic pathways reveals that specific biological processes within gene ontology are common in both MS and BD. These included positive regulation of cell migration, regulation of calcidiol 1-monooxygenase activity, regulation of the mitotic cell cycle, and ncRNA metabolic processes. Additionally, shared KEGG database pathways encompass Tuberculosis, HIF-1 signaling pathway, AGE-RAGE signaling pathway in diabetic complications, and the Toll-like receptor signaling pathway.

The results of this study have the potential to identify common drugs and collaborative treatments for patients with MS and BD. Furthermore, the key genes identified can serve as biomarkers for future research in MS and BD.

Keywords: Bipolar disorder; Multiple sclerosis; autoimmune; immune system; hub genes.

References

[1] Mora, P. and C. Chapouly, Astrogliosis in multiple sclerosis and neuro-inflammation: what role for the notch pathway? Frontiers in Immunology, 2023. 14.

[2] Robinson, N. and S.E. Bergen, Environmental risk factors for schizophrenia and bipolar disorder and their relationship to genetic risk: current knowledge and future directions. Frontiers in genetics, 2021. 12: p. 686666.

[3] Carta, M., et al., The risk of bipolar disorders in multiple sclerosis. Journal of affective disorders, 2014. 155: p. 255-260.

ICB12-1402-1217

The Potential Therapeutic Effects of Quercetin in Type 1 Diabetes: Targeting Essential Genes and Pathways.

Sepideh Kiani, Abolghasem Esmaeili*
Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: aesmaeili@sci.ui.ac.ir

Abstract: Type 1 diabetes (T1DM) is characterized by severe insulin deficiency caused by chronic and progressive destruction of pancreatic β cells by the immune system. Quercetin (3,3,4,5,7-pentahydroxyflavone) is a dietary flavonoid and bioactive compound that is abundantly found in various fruits and vegetables. Quercetin has significant blood sugar-lowering effects, stabilization of long-term insulin secretion, and islet regeneration. In the pancreas, it is not dangerous. The purpose of this study was to use integrated bioinformatics analysis to clarify key candidate genes and pathways in T1DM.In this study, the Query and Limma package from GSE156908 of the Gene Expression Omnibus (GEO) were used to analyze differentially expressed genes (DEGs). A protein-protein interaction (PPI) network, gene ontology (GO), and REACTOME pathway enrichment analyses were constructed and analyzed. Drug metabolism, arachidonic acid metabolism, the HIF-1 signaling pathway, and the cGMP-PKG signaling pathway are the main metabolic pathways in which DEGs are enriched, according to the results of the GO and REACTOME enrichment analyses. Finally, we examined the genes related to type 1 diabetes that are targeted by quercetin XDH, PDE5A, PLG, TOP2A, INSR, GLO1, MPO, ALOX12, and we concluded that quercetin may improve type 1 diabetes by targeting these genes.

Keywords: Bioinformatics; Type 1 diabetes; Quercetin; Differentially expressed genes; Enrichment analysis

ICB12-1402-1219

A Bioinformatic study of the chalcone synthase (*CHS3*) gene among different species

Sajjad Mansouri*
Agronomy and Plant Breeding Department, Ilam University, Ilam, Iran.
*Corresponding author's email address: mansouri_65@yahoo.com

Abstract: The gene *CHS3* is one member of the chalcone synthase gene family found in plants, which play an important role in the biosynthesis of flavonoid compounds. Mutations or alterations in the *CHS3* gene can result in changes to pigmentation patterns and affect plant phenotype. To uncover the structural features of *CHS3* genes and proteins in different plant species and to investigate the phylogenetic relationships between them, a total of 21 *CHS3* gene sequences with full-length coding sequence (CDS) and corresponding amino acid sequences were analyzed for the following features: length of sequence, length of CDS, GC content, the number and molecular weight (MW) of amino acids (AA), isoelectric point (pI), and hydrophobic index. The CDS length ranged from 900 bp to 1200 bp and the protein length ranged from 299 to 399 AA. The GC content of the gene across species was 44-64%. With the exception of the species *Dryopteris fragrans*, all other species have a negative protein hydrophobicity index, which shows that the overall amino acid sequence of the gene is more hydrophilic. Moreover, the isoelectric point values ranged from 5.72 to 7.80. The results of multiple alignments of *CHS3* amino acid sequences showed that some regions in the sequences were conserved among the species. A neighbor-joining (NJ) phylogenetic tree based on the *CHS3* amino acid sequences grouped the species very well. The findings provide a basis for future functional research on the chalcone synthase gene family and the evaluation of their enzymatic activity in plants.

Keywords: Bioinformatics; CHS3 gene; Plant species

References

[1] V. Oberholzer, M.L. Durbin, and M.T. Clegg. Comparative Genomics of Chalcone Synthase and Myb Genes in the Grass Family. Genes Genet Syst. 2000;75(1):1-16

[2] Z. Wang, Yu, Q. Shen, W. et al. Functional Study of CHS Gene Family Members in Citrus Revealed a Novel CHS Gene Affecting the Production of Flavonoids. BMC Plant Biol. (2018). 18, 189

[3] X. Wu, S. Zhang, X. Liu, J. Shang, A. Zhang, et al. Chalcone Synthase (CHS) Family Members Analysis from Eggplant (solanum melongena l.) In the Flavonoid Biosynthetic Pathway and Expression Patterns in Response to Heat Stress. 2020. PLOS ONE 15(4): e0226537

ICB12-1402-1220

Designing targeted PROTAC-based therapeutics against mutant p53 proteins through structure-based virtual screening and dynamic simulations

Ehsan Mofidi Chelana,b,c, Amir Zarebkohanc, Reza Hasania,b,c, Ailar Nakhlband d, Mohammad M. *Pourseifb,c
a Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
c Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
d Research Center of Psychiatry and Behavioral Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: More than half of cancers bear TP53 mutations, that drive tumorigenesis and progression through having gain-of-function effects [1]. Despite p53 variants are considered as "undruggable" targets, efforts have been made in p53-targeted therapies, but challenges remain in clinical development. Proteolysis-targeting chimeras (PROTACs) are a rapidly evolving technology which target undruggable proteins through inducing their degradation. PROTACs are heterobifunctional molecules that consist of a ligand for the target protein, a ligand for the E3 ubiquitin ligase, and a linker connecting the two ends [2]. Herein, we used structure-based virtual screening and dynamic simulations to design a PROTAC system that targets a specific mutated form of the p53 protein known as p53(Y220C). By using molecular docking technology, we virtually screened a large number of peptides to identify those that have a high affinity for the target protein. This screening process involves predicting the binding interactions (and energies) between the peptides and the target protein based on their structural compatibility. To build our PROTAC system, we used previously used E3 ubiquitin ligase binder and linkers with our candidate peptide. The final construct evaluated for its stability efficacy and safety through 100ns molecular dynamics (MD) simulations and based on GROMOS96 54a7 force field for understanding atomic level motion and interactions of the PROTAC-based designed peptide in complex with TP53. The analyses of RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuation), radius of gyration, and hydrogen bond formation indicate that the designed lead PROTACs exhibit stable interactions throughout the simulation period. The subsequent breakdown of the p53(Y220C) mutant via the PROTAC system may aid in restoring or stabilizing the typical structure of the p53 protein, thereby presenting a potentially innovative therapeutic approach for combating cancer. However, further *in vitro* and *in vivo* studies needed to ensure its effectiveness and safety.

Keywords: Cancer; Molecular dynamic simulation; PROTAC; TP53; Virtual screening

References

[1] Lu, Yier, Meng Wu, Yang Xu, and Lili Yu. "The development of p53-targeted therapies for human cancers." *Cancers* 15, no. 14 (2023): 3560.

[2] Wang, Haolan, Ming Guo, Hudie Wei, and Yongheng Chen. "Targeting p53 pathways: mechanisms, structures, and advances in therapy." *Signal transduction and targeted therapy* 8, no. 1 (2023): 92.

ICB12-1402-1221

Homology Modeling and Molecular Dynamics Simulation of Recombinant Laccase

F.Noroozi Tiyoula [a,+], Sh.Ariaeenejad [b], K.Kavousi [a], A.Lohrasbi-Nejad [c,d,*]
[a] Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
[b] Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran
[c] Department of Agricultural Biotechnology, Shahid Bahonar University of Kerman, Kerman, Iran
[d] Research and Technology Institute of Plant Production, Shahid Bahonar University of Kerman, Kerman, Iran
+ Presenter: fereshtehnorouzi3@gmail.com
*Corresponding author's email address:: a.lohrasbi@uk.ac.ir

Abstract: Today, due to the abundant use of recombinant and engineered proteins in various fields, predicting the tertiary structure of proteins opens a new window in understanding their function. In this study, the sequence of recombinant laccase was used, and homology modeling was done by MODELLER software version 10.04 [1]. The maize laccase (PDB ID: 6KLG [2]) was considered as a template. A Python script did the prediction and construction of 100 models. DOPE and GA341 methods were employed to evaluate the quality of the models. To validate the 3D structures, the UCLA-DOE LAB SAVES tool was utilized. Following the identification of the best model, we assessed the structural stability of the enzyme through 10 ns molecular dynamics simulation using Gromacs 2022 [3]. The simulated structure exhibited notable stability, as indicated by consistent RMSD values observed during the simulation. In comparison to the reference protein, our novel enzyme displayed an RMSD value of 1.23 Å. The recombinant laccase formed a significant number of intra-structure hydrogen bonds, totaling 16,993, while the reference protein formed only 458. Additionally, the solvent-accessible surface area (SASA) of the structures differed, with values of 20,534.47 and 19,377.14 Å²/kcal for the reference protein and recombinant enzyme, respectively. The novel enzyme's 3D structure opens avenues for both experimental and computational investigations, such as predicting binding sites with ligands and engineering enhancements through residue design in future studies.

Keywords: tertiary structure; recombinant laccase; homology modeling; molecular dynamics simulation

References

[1] T. Xie, Z. Liu, and G. Wang, "Structural basis for monolignol oxidation by a maize laccase," *Nat. Plants*, vol. 6, no. 3, pp. 231–237, doi: 10.1038/s41477-020-0595-5.

[2] A. Šali and T. L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints," *Journal of Molecular Biology*, vol. 234, no. 3. pp. 779–815, 1993. doi: 10.1006/jmbi.1993.1626.

[3] M. J. Abraham *et al.*, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1–2, pp. 19–25, 2015, doi: 10.1016/j.softx.2015.06.001.

ICB12-1402-1226

Designing targeted PROTAC-based therapeutics against mutant p53 proteins through structure-based virtual screening and dynamic simulations

Ehsan Mofidi Chelan[a,b,c], Amir Zarebkohan[c], Reza Hasani[a,b,c], Ailar Nakhlband [d], Mohammad M. Pourseif[b,c]*
a Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
b Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
c Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
d Research Center of Psychiatry and Behavioral Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: More than half of cancers bear TP53 mutations, that drive tumorigenesis and progression through having gain-of-function effects [1]. Despite p53 variants are considered as "undruggable" targets, efforts have been made in p53-targeted therapies, but challenges remain in clinical development. Proteolysis-targeting chimeras (PROTACs) are a rapidly evolving technology which target undruggable proteins through inducing their degradation. PROTACs are heterobifunctional molecules that consist of a ligand for the target protein, a ligand for the E3 ubiquitin ligase, and a linker connecting the two ends [2]. Herein, we used structure-based virtual screening and dynamic simulations to design a PROTAC system that targets a specific mutated form of the p53 protein known as p53(Y220C). By using molecular docking technology, we virtually screened a large number of peptides to identify those that have a high affinity for the target protein. This screening process involves predicting the binding interactions (and energies) between the peptides and the target protein based on their structural compatibility. To build our PROTAC system, we used previously used E3 ubiquitin ligase binder and linkers with our candidate peptide. The final construct evaluated for its stability efficacy and safety through 100ns molecular dynamics (MD) simulations and based on GROMOS96 54a7 force field for understanding atomic level motion and interactions of the PROTAC-based designed peptide in complex with TP53. The analyses of RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuation), radius of gyration, and hydrogen bond formation indicate that the designed lead PROTACs exhibit stable interactions throughout the simulation period. The subsequent breakdown of the p53(Y220C) mutant via the PROTAC system may aid in restoring or stabilizing the typical structure of the p53 protein, thereby presenting a potentially innovative therapeutic approach for combating cancer. However, further *in vitro* and *in vivo* studies needed to ensure its effectiveness and safety.
Keywords: Cancer; Molecular dynamic simulation; PROTAC; TP53; Virtual screening

References

[1] Lu, Yier, Meng Wu, Yang Xu, and Lili Yu. "The development of p53-targeted therapies for human cancers." *Cancers* 15, no. 14 (2023): 3560.

[2] Wang, Haolan, Ming Guo, Hudie Wei, and Yongheng Chen. "Targeting p53 pathways: mechanisms, structures, and advances in therapy." *Signal transduction and targeted therapy* 8, no. 1 (2023): 92.

[3] Freixas, V. M., White, A. J., Nelson, T., Song, H., Makhov, D. V., Shalashilin, D., ... & Tretiak, S. (2021). Nonadiabatic excited-state molecular dynamics methodologies: Comparison and convergence. *The Journal of Physical Chemistry Letters*, 12(11), 2970-2982.

ICB12-1402-1228

CDC27 pseudogene expression in blast crisis phase of CML

GE. Kazemi-Sefat[a], S.Talebi[a], M. Keramatipour[b], K. Kavousi[c] , K. Mousavizadeh[d*]
a Department of Medical Genetics, School of Medicine,Iran University of Medical Sciences, Tehran, Iran
b Department of Medical Genetics, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran.
c Institute of Biochemistry and Biophysics (IBB), Department of Bioinformatics, Laboratory of Complex Biological Systems and Bioinformatics (CBB), University of Tehran, Tehran, Iran
d Department of Pharmacology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
**Corresponding author's email address: Mousavizadeh.k@iums.ac.ir

Abstract: The transition of CML to the blastic phase represents a critical factor leading to mortality in this cancer, and its underlying mechanisms remain poorly understood.

CDC27, a component of the APC/C complex, plays a vital role in precisely controlling transitions in the cell cycle, which is essential for normal cell division and development. When CDC27 and the APC/C complex are dysregulated, the development and progression of cancer may occur. The CDC27 gene has eleven pseudogenes.

Pseudogene is a DNA sequence that mimics a functional gene but has either lost its capability to code for proteins or remains unexpressed. Changes in the expression of pseudogenes have been identified across different cancer types, and their irregular activity can affect critical cellular processes like cell proliferation, apoptosis, and metastasis.

In this study, RNA-seq data from the EGA with accession EGAS00001003071 were employed. The GRCh38 was used as reference genome and HISAT2 as the mapper. Stringtie which is equipped with efficient algorithms for recovering transcript structure and estimating abundance, was employed to analyze the bulk RNA-Seq reads aligned to the reference genome. We utilized both DESeq2 and edgeR packages for the analysis of Differentially Expressed Genes (DEGs). Analyses were conducted in both paired and unpaired approaches.

As a result, in all blastic phase CML patients, the expression of three pseudogenes including CDC27P9, CDC27P10, and CDC27P11, with a False Discovery Rate (FDR) less than 0.01 and Log2FC>1 was higher than chronic phase of CML, while the expression of other CDC27 pseudogenes between these two groups was not significantly different. Continuous and thorough investigations are required to assess target molecules whose expression is influenced by these three pseudogenes including CDC27 gene that plays pivotal role in regulating cell cycle and other cellular processes. Furthermore, it will help explore potential therapeutic options.

Keywords: CML; Blast crisis phase; Pseudogene; RNA-Seq; CDC27

ICB12-1402-1229

Splicing Detection In DNA Sequences With A Focus On Balancing Visual Data And Deep Neural Networks

E. Keshavarzi Khouzani[1], A. Ghanbari Sorkhi[1]
[1] Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran
elahehkeshavarzik@gmail.com
*Corresponding author's email address: ali.ghanbari@mazust.ac.ir

Abstract: Accurate detection of splicing in DNA and proteins holds significant medical and vital importance. This process enhances genetic diversity and aids in the regulation of exon skipping. Maintaining the correct order of nucleotides is crucial for proper protein production in this process. Given the importance of this matter, extensive research has been conducted in recent years to identify regions related to exons and introns .Numerous challenges, such as imbalanced datasets and the absence of visual data, exist in this domain. In this regard, a novel multi-stage method for detecting these regions is proposed in this article. The splicing detection is performed using the powerful ResNet-32 architecture, recognized for its effectiveness in training deep networks. ResNet-32 utilizes residual blocks, where each residual block has a shortcut connection allowing the model to learn the difference between the input and output of skipped layers. This innovative approach mitigates issues like vanishing gradients during training .To address the issue of imbalanced data, three methods are employed in this project: Reweighting, Resampling, and DRW (Dynamic Reweighting). Resample: Oversamples the minority class during training to create a balanced representation of classes in each batch. Reweight: Adjusts sample weights to give higher importance to the minority class, preventing bias towards the majority class during training. DRW (Dynamic Reweighting): Dynamically adapts the reweighting strategy by changing beta values over epochs, allowing the model to gradually adjust to imbalanced datasets during different training stages. Alongside these techniques, to overcome the lack of visual data, each nucleotide in the sequence is transformed into a one-hot encoded vector, forming a sequence of zeros and ones. This sequence is then converted into a matrix, serving as a binary image representing visual data for the sequences .The proposed method's performance is evaluated on the standard Caenorhabditis elegans dataset. The obtained results demonstrate the satisfactory performance of the proposed approach.

Keywords: Splicing detection; ResNet-32 architecture; Imbalanced datasets; DNA sequence analysis

References

[1] Y. Yang, Z. Xu, "Rethinking the value of labels for improving class-imbalanced learning," 2006.07529, 2020.

ICB12-1402-1230

Deciphering the Role of CircRNA-miRNA Networks in Multiple Sclerosis Pathogenesis through Minimal Cut-Set Analysis

Mohammad Reza Zabihi[a], Zahra Salehi[b], Kaveh Kavousi[a],
[a] Institute of Biochemistry and Biophysics (IBB), Department of Bioinformatics, Laboratory of Complex Biological Systems and Bioinformatics (CBB), University of Tehran, Tehran, Iran
[b] Hematology, Oncology and Stem Cell Transplantation Research Center, Research Institute for Oncology, Hematology and Cell Therapy, Tehran University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Multiple sclerosis (MS) is a chronic autoimmune neurodegenerative disorder with a complex, yet poorly understood, pathophysiology. Recent studies have highlighted the role of non-coding RNAs, especially circular RNAs (circRNAs), in the regulation of gene expression and their impact on MS progression. Notably, circRELL1, circRPPH1, and circGSDMB have been identified as significantly upregulated in MS patients. This study aims to elucidate the competing endogenous RNA (ceRNA) network of these circRNAs using minimal cut-set methodology.

We employed the CircInteractome web tool and miRTarBase database to analyze the microRNAs and their associated mRNA targets for circRELL1, circRPPH1, and circGSDMB. The protein-protein interaction (PPI) network of these mRNAs was reconstructed, and the minimal cut-set was identified using Gephi network analysis tool. The key driver nodes were determined using the CytoCtrlAnalyser plugin in Cytoscape 3.9. The present analysis revealed that circRELL1, circRPPH1, and circGSDMB target 15, 14, and 6 miRNAs, respectively, with significant interactions observed in the PPI network. Specifically, we identified five proteins - AKT1, CCND2, BAX, CRKL, and EGFL7—as enriched driver nodes in the PPI network. Notably, at least five circRNA/miRNA/mRNA axes, including circ_0001400/ miR-637/AKT1, circ_0001400/ miR-126/ EGFL7, and circ_0001400/ miR-126/ CRKL, circ RPPH1/miR-663b/CCND2, circ_0106803/ miR-7-5p and miR-766-3p/BAX, were identified as key contributors to MS pathogenesis, influencing critical processes such as T cell proliferation, blood-brain barrier integrity, and oligodendrocyte apoptosis.

These findings enhance our understanding of the molecular underpinnings of MS and suggest new avenues for targeted therapeutic strategies by modulating these ceRNA networks. Importantly, the identification of driver nodes within these networks highlights the potential of network analysis in deciphering complex disease mechanisms and guiding the development of effective treatments for neurodegenerative disorders like MS.

Keywords: Splicing detection; ResNet-32 architecture; Imbalanced datasets; DNA sequence analysis

ICB12-1402-1232

Study of Lactobacillus Strain *L28* Using Bioinformatics Methods

*Zahra Farkhondeh Khoy Farda, Asal Roane Ghanadzadehb, Milad Shaghaghy ranjbarc
Department of Microbiology, Kish International Branch of Islamic Azad University, Kish, Iran.
Department of Microbiology, Kish International Branch of Islamic Azad University, Kish, Iran.
Microbiology Lecturer at Shirin Asal University of Applied Sciences, Tabriz, Iran.
Islamic Azad University, Tehran Medical Branch
Islamic Azad University, Kish International Branch
a.rohanighanadzadeh@iaukishint.ac.ir; m.shaghaghy@iaukishint.ac.ir
*Corresponding author's email address: z.farkhondehkhoyfard@iaukishint.ac.ir

Abstract: The study of *Lactobacillus* strain *L28* utilizing bioinformatics methods has provided valuable insights into its genetic and metabolic characteristics, shedding light on its potential applications in various domains [1]. highlighted that the comprehensive genomic analysis of *Lactobacillus* strain *L28* revealed over 100 genes in the Lactobacillus core genome (LCG), which displayed conserved organization and control, suggesting a common ancestry beyond the Lactobacillus group. This study identified genes associated with specific signaling to the host and conserved regulatory systems, indicating a fundamental network of environmental responses present in all *Lactobacillus* genomes. Furthermore, the presence of the FbpA gene in all sequenced *Lactobacillus* genomes was emphasized, underlining its potential significance. The study underscored the importance of complete genome sequences for detailed genomic comparisons and provided a platform for analyzing the pan genome, niche-specific genes, and specific features of conserved genes in *Lactobacillus* genomes. The insights gained from this study have significant implications for potential applications of *Lactobacillus* strain *L28* in biopreservation, probiotic formulations, and biotechnological applications. This research exemplifies the pivotal role of bioinformatics methods in unraveling the potential of probiotic microorganisms for diverse industrial and health-related purposes.

ICB12-1402-1233

Exploring the Synergistic Anticancer Effects of Iranian Medicinal Plants on EGFR-Driven Tumor Cell Lines: Integrative Computational and Molecular Dynamics Study)

Mohammad Shirinpoor 1, Shokoofeh Ghiam 2, Changiz Eslahchi 1*
1.Department of Computer and Data Sciences, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran
2.School of biological sciences, Institute for research in fundamental sciences, Tehran, Iran
*Corresponding author's email address: Ch-Eslahchi@sbu.ac.ir

Abstract: Chemotherapy, a crucial component in cancer treatment, presents considerable challenges, including high costs, substantial side effects, and variable efficacy. Despite advancements in EGFR-targeted therapies for EGFR-positive cancers, the persistent challenge of resistance necessitates innovative solutions. This study focuses on targeting the epidermal growth factor receptor (EGFR) using compounds derived from Iranian medicinal plants. To efficiently navigate the abundance of potential compounds, we employ a computational approach to streamline candidates for laboratory experimentation, optimizing resources and accelerating the identification of promising anti-EGFR agents.

inspired by the computational methods proposed in three seminal papers by (Ahmadi Moughari et al., 2021; Emdadi et al, 2021; Yassaee Meybodi et al, 2021) our study leverages their methodologies. We employ their models to generate outputs, subsequently used as inputs for our deep neural network designed to predict IC50 values of Iranian medicinal plant compounds against EGFR-associated cell lines. This integrative approach aims to enhance the predictive accuracy and efficiency of identifying potential anti-EGFR agents, contributing to the advancement of cancer treatment strategies.

Leveraging the ChemBL database, we meticulously extract the structural features of drugs using RDKit. Incorporating expression and miRNA data as cell line descriptors results in a dataset comprising 153,914 drugs, each characterized by 38 features across 23 unique cell lines. This comprehensive approach enables the merging of chemical and biological information, providing a holistic perspective on the relationships between compound properties and efficacy against EGFR-associated cell lines.

Following the IC50 prediction of natural compounds, our subsequent step involves molecular docking studies with the EGFR1-4 kinase domain. This analysis evaluates binding affinities and interactions, facilitating the identification of promising compounds. Prioritized candidates undergo molecular dynamics simulations to explore the stability and behavior of the compound-EGFR complexes over time, offering valuable insights into their structural aspects and potential as anti-EGFR agents.

This integrative computational and molecular dynamics approach thoroughly evaluates natural compounds, promising a multi-faceted understanding of their potential as effective EGFR inhibitors. Such initiatives significantly contribute to developing cost-effective and efficient strategies for discovering novel therapies tailored to EGFR-positive cancers.

Keywords: Computational methods, molecular docking, molecular dynamics, anti-EGFR agents, natural compounds.

References

[1] A. Emdadi and C. Eslahchi, "Auto-HMM-LMF: feature selection based method for prediction of drug response via autoencoder and hidden Markov model," BMC Bioinformatics, vol. 22, no. 1, p. 33, Dec. 2021, doi: 10.1186/s12859-021-03974-3.

[2] F. Ahmadi Moughari and C. Eslahchi, "A computational method for drug sensitivity prediction of cancer cell lines based on various molecular information," PLoS One, vol. 16, no. 4, p. e0250620, Apr. 2021, doi: 10.1371/journal.pone.0250620.

ICB12-1402-1234

Enrichment analysis of potential hub genes in breast cancer based on TCGA datasets

Narjes Seddighi[1,2*], Ali Salimi Asl[3]
1- Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
2- Hematology and Oncology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
3- Student Research Committee, Faculty of Medicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
Corresponding to: Narjes Seddighi; Tabriz University of Medical Sciences, Tabriz, Iran; Email:
*Corresponding author's email address: Narjes.seddighi80@gmail.com

Abstract: Breast cancer is the most common malignancy in women worldwide. Despite great advances in the diagnosis and treatment of cancer, the treatment has big challenges. Nowadays, with the advances in the bioinformatics analysis of RNAseq, it can identify potential biomarkers for diagnosis, targeting treatment, and evaluation of tumor metastasis and relapse. In the current study, we conducted the gene ontology and Kegg pathway based on TCGA datasets.
At first, we used GEPIA2 to examine the TCGA BRCA dataset to identify all DEGs linked with BRCA among high throughput RNA-Seq data. After analyzing the survival data of BRCA, a Protein-protein interaction (PPI) network of significant genes associated with BRCA was constructed in Cytoscape software, and the hub genes were identified. for enrichment analysis, we used the Enrichr website (https://maayanlab.cloud/Enrichr/) and extracted the Kegg pathway and gene ontology based on the potential hub Gene
Seven Hub genes (CXCL1, SELL, CXCL9, CD3E, CCL5, CXCR3, CCR5) were extracted from PPI based on their degree. the result showed in the 189 biological processes, the Cellular Response to Lipopolysaccharide was significantly meaningful (adj P-value=9.18E-06). In 12 cellular components, we identified just 3 components with meaningful adj p-values which included Gamma-Delta T Cell Receptor Complex, Alpha-Beta T Cell Receptor Complex, and T Cell Receptor Complex (adj P-value=0.02). In 30 molecular functions, Chemokine Activity (adj P-value=7.66E-06) and Chemokine Receptor Binding (adj P-value = 7.66E-06) have a more meaningful. At the last, in the 35 Kegg pathway, we identified Viral protein interaction with cytokine and cytokine receptor (adj P-value=2.06E-09), Chemokine signaling pathway (adj P-value=2.80E-08), and Cytokine-cytokine receptor interaction (adj P-value=1.61E-07) with meaningful P-value.
This study identifies gene ontology and keg pathways of hub genes which involved in the occurrence and progression of breast cancer. This information may hold promise as potential biomarkers and therapeutic targets.
Keywords: Breast cancer, hub genes, BRCA, enrichment analysis, Kegg pathway, Gene Ontology

References

[1] Cui H, Ren X, Dai L, Chang L, Liu D, Zhai Z, Kang H, Ma X. Comprehensive analysis of nicotinamide metabolism-related signature for predicting prognosis and immunotherapy response in breast cancer. Frontiers in Immunology. 2023 Mar 8;14:1145552.

[2] Luo W, Han Y, Li X, Liu Z, Meng P, Wang Y, Damante G. Breast cancer prognosis prediction and immune pathway molecular analysis based on mitochondria-related genes. Genetics Research. 2022 Jan;2022:e56.

[3] Identification of a five genes prognosis signature for triple-negative breast cancer using multi-omics methods and bioinformatics analysis

ICB12-1402-1235

Computational Discovery of Antimicrobial Enzymes from Environmental Microbiota

Arad Ariaeenejad[a†] , Arman Hasannejad[b†] , Donya Afshar Jahanshahi[a, c] , Mohammad Reza Zabihi[a] , Shohreh Ariaeenejad[b *], Kaveh Kavousi[a *]

[a]Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran.

[b] Department of Systems and Synthetic Biology, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.

[c] Department of Bioinformatics, Kish International Campus, University of Tehran, Kish, Iran.

† Equally contributed

*Corresponding author's email address kkavousi@ut.ac.ir , shariaee@gmail.com

Abstract: In contemporary society, the escalation of antibiotic-resistant microorganisms not only engenders key industries but also poses significant challenges across various parts, including healthcare, agriculture, and poultry [1]–[3] To address this issue, it is imperative to identify novel antimicrobial agents that can overcome resistance. One of the potential candidates for finding these agents is biomolecules in harsh microbiome environments, where microbial societies vie for the dominance and acquisition of potent antimicrobial attributes. Many enzymes encoded by the genomes of microorganisms have been shown to possess antimicrobial properties. In this study, we investigated the metagenome of the tannery waste environment using computational methods and identified antimicrobial enzymes in this environment. To this end, metagenomic samples were analyzed using approaches including assembly, gene prediction, enzyme prediction, and 3D structure prediction to discover new antibiotics that are effective in antibiotic resistance. All contigs were analyzed using the Metarenz [4] software to predict the presence of enzybiotics. Metarenz is a tool for identifying target enzymes in assembled contigs. 3D structures of some candidates were predicted using the Alphafold2 [5] and TMalighn [6] tools. The resulting metagenomic sequences were further investigated using the NCBI CDD database.

Keywords: Antimicrobial; Metagenomics; Microbiota; Computational; Enzymes Discovery

References

[1]Q. Wang, Z. Luo, Y. L. Wu, and Z. Li, "Recent Advances in Enzyme-Based Biomaterials Toward Diabetic Wound Healing," *Adv. NanoBiomed Res.*, vol. 3, no. 2, 2023, doi: 10.1002/anbr.202200110.

[2]B. Thallinger, E. N. Prasetyo, G. S. Nyanhongo, and G. M. Guebitz, "Antimicrobial enzymes: An emerging strategy to fight microbes and microbial biofilms," *Biotechnol. J.*, vol. 8, no. 1, pp. 97–109, 2013, doi: 10.1002/biot.201200313.

[3]K. M. Danis-Wlodarczyk, D. J. Wozniak, and S. T. Abedon, "Treating bacterial infections with bacteriophage-based enzybiotics: In vitro, in vivo and clinical application," *Antibiotics*, vol. 10, no. 12, pp. 1–36, 2021, doi: 10.3390/antibiotics10121497.

[4]K. K. Shahraki MF, Atanaki FF, Ariaeenejad S, Ghaffari MR, Norouzi-Beirami MH, Maleki M, Salekdeh GH, "A computational learning paradigm to targeted discovery of biocatalysts from metagenomic data: A case study of lipase identification," *Biotechnol Bioeng*, vol. doi: 10.10.

[5]O. Ronneberger *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. May, 2021, doi: 10.1038/s41586-021-03819-2.

[6]Y. Zhang and J. Skolnick, "TM-align: A protein structure alignment algorithm based on the TM-score," *Nucleic Acids Res.*, vol. 33, no. 7, pp. 2302–2309, 2005, doi: 10.1093/nar/gki524.

ICB12-1402-1238

Comparative genomics of 919 microbial genomes of the order Oceanospirillales for oil hydrocarbon degradation

F. Qasemnejadlangerodi, S.Asad*, M. Dastgheib*, M. Mehrshad*
a Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran
b Biotechnology and Microbiology Research Group, Research Institute of Petroleum Industry, Tehran, Iran
c Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, 750 07 Sweden
*Corresponding author's email address: f.qasemnejad@ut.ac.ir

Abstract: Oil pollution is a threat for the ocean ecosystem; negatively affecting the health of aquatic organisms and humans. Biological remediation provides a sustainable approach for alleviating threats of oil pollution on aquatic ecosystems and harnessing the ability of microbes for oil degradation in the ocean ecosystem has a vast potential that still remains largely untapped[1][2][3].
Order *Ocaenospirillales* representatives are first responders to release of oil hydrocarbons, where they bloom due to their enzymatic profile enabling them to degrade oil hydrocarbons. Here, we present a comparative genomics study of the capability of *Ocaenospirillales* genomes for oil hydrocarbon degradation. Also, position of key hydrocarbon degradation genes in the core or accessory fraction of the pangenome was inspected to shed light on the flow of these genes among populations.
A number of 845 *Ocaenospirillales* genomes were collected from the National Center for Biotechnology Information and 74 MAGs from oil contaminated sites were added to this dataset. After quality control, using CheckM, prodigal was used to predict genes and eggNOG-mapper were used for functional annotation. FastANI and motulizer were used For genome clustering. then pangenome analysis was done using the Bayesian approach of mOTUpan.py.
Updated taxonomy based on 120 single-copy gene markers show that members of the order *oceanospirillales* are transferred to the five orders of *Pseudomonadales*, *Enterobacterales*, *Xanthomonadales*, *HP12*, and *Nevskiiales*. 70% of all families were rich in genes capable of degrading alkanes. Moreover, 83% of genes involved in aliphatic hydrocarbons degradation were located in the core fraction of the pangenomes, reiterating the fact that the flow if these genes is mainly vertical. However, genes responsible for breaking down aromatic compounds were mostly in the accessory fraction of the pangenomes. highlighting that bacteria mainly acquired these capabilities through horizontal transfer in order to become more compatible with their environment.
Keywords: Oil hydrocarbons, comparative genomics, pangenome, Core genome, Accessory genome.

References

[1] T. Thi Kim Chi, "Numerical Simulation of Viscous Effects of Oil on the Capacity of Oil Booms," in *E3S Web of Conferences*, Dec. 2019, vol. 140, doi: 10.1051/e3sconf/201914006014.

[2] C. Campagna *et al.*, "Gulf of Mexico oil blowout increases risks to globally threatened species," *Bioscience*, vol. 61, no. 5, pp. 393–397, 2011, doi: 10.1525/bio.2011.61.5.8.

[3] S. Harayama, H. Kishira, Y. Kasai, and K. Shutsubo, "Petroleum biodegradation in marine environments," *J. Mol. Microbiol. Biotechnol.*, vol. 1, no. 1, pp. 63–70, 1999.

ICB12-1402-1239

Exploring Epigenetic Methylation Patterns in Parkinson's Disease through WGCNA Analysis

Fateme Vafaee Sharbaf[a], Zahra Salehi[b, c], Abdollah Safari[d], Kaveh Kavousi[a]*
[a] Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
[b] Hematology, Oncology and Stem Cell Transplantation Research Center, Tehran University of Medical .Sciences, Tehran, Iran
[c] Research Institute for Oncology, Hematology and Cell Therapy, Tehran University of Medical Sciences, Tehran, Iran
[d] School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran
*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Parkinson's disease (PD) involves progressive loss of dopaminergic neurons in the substantia nigra, causing motor and non-motor symptoms. Epigenetic mechanisms play a crucial role in PD pathogenesis, reflecting the interplay of genetics and the environment. This study aimed to identify differential methylated positions (DMP) and regions (DMRs) of PD patients compared to healthy samples, shedding light on the epigenetic factors underlying Parkinson's disease.
We conducted a comprehensive analysis of DNA methylation patterns in PD using data obtained from Illumina EPIC arrays. Data were downloaded from PPMI [1]. In total, 214 sporadic PD and 87 healthy control samples were analyzed. The identification DMPs was performed using the limma package. Next, WGCNA [2] was employed to construct co-methylation networks and detect modules of highly correlated DMPs and DMRs associated with PD phenotypes. Preprocessing steps, including quality control, normalization, and batch correction, were performed using the ChAMP package in R [3]. DMPs and DMRs between PD and healthy control samples were identified based on stringent statistical criteria. Hub genes derived from WGCNA analysis were subjected to enrichment analysis, to improve our understanding of the molecular pathways underlying PD pathogenesis.
Our analysis identified 14866 DMPs and 101 DMRs (adjusted *P*-values of 0.01 and 0.05, respectively). Utilizing WGCNA, we delineated 20 clusters and identified 13 hub genes associated with DMPs. Notably, BMP4 and MTHFD2 emerged as significant candidates, potentially influencing Parkinson's disease pathophysiology. However, none of the DMR modules achieved significance. These findings shed light on PD-associated methylation alterations and underscore pathways, including mitochondrial dysfunction, neuroinflammation, and synaptic transmission implicated in PD pathogenesis.
Our analysis revealed distinct methylation patterns associated with PD, characterized by widespread alterations in DNA methylation levels at specific genomic loci. WGCNA identified modules of co-methylated loci showing differential methylation patterns between PD and healthy control samples.
Keywords: Parkinson's disease; DNA methylation Analysis; Weighted Gene Co-expression Network Analysis

References

[1] Marek, K., et al., *The Parkinson progression marker initiative (PPMI).* Progress in neurobiology, 2011. 95(4): p. 629-635.
[2] Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC bioinformatics, 2008. 9(1): p. 1-13.
[3] Morris, T.J., et al., ChAMP: 450k chip analysis methylation pipeline. Bioinformatics, 2014. 30(3): p. 428-430.

ICB12-1402-1240

Enhancing Predictive Accuracy of Biomolecules Partition Coefficients in Aqueous Two-Phase Systems Using Machine Learning

Fateme Bagheri, Gholamreza Pazuki*, Fatemeh Zare-Mirakabad, Mahsa Sadaat, Zahra Ghorbanali

Department of Chemical Engineering, AmirKabir University of Technology, Tehran, Iran

Department of Mathematics and Computer Science, AmirKabir University of Technology, Tehran, Iran

*Corresponding author's email address: ghpazuki@aut.ac.ir

Abstract: Information about the pharmaceutical and biotech industries has experienced remarkable growth in recent years. Biological processes consist of two main stages: upstream and downstream processes. Downstream processes aim to recover and purify chemical products, constituting 50 to 80 percent of the total production cost [1]. Extraction, a stage within downstream processes, employs methods like liquid-liquid extraction systems. One type of liquid-liquid extraction system is the aqueous two-phase system(ATPS)[2]. As a significant portion of the aqueous two-phase system is water, it is used to purify biological molecules such as drugs, proteins, etc. This study uses machine learning methods to predict the partition coefficient of drugs in aqueous two-phase system[3]. The database utilized in this study includes data collected from previous articles that experimentally calculated information related to the components of the aqueous two-phase system, and details about the chemical structure and physical properties of drugs. This study aims to investigate how the properties of drugs affect their distribution coefficient in the aqueous two-phase system. In the investigation of the chemical structure, binary Morgan fingerprints, count-based Morgan fingerprints and Graph convolution were utilized. Additionally, physical properties such as melting point, density, log P, etc were considered. To predict the partition coefficient of drugs in the aqueous two-phase system, various machine-learning models were employed, including Random Forests, ANN, ensemble methods, etc. Results show the significant influence of drug properties on partition coefficient prediction. The best performance relates to combining the physical and chemical properties of drugs using count_based Morgan fingerprints representation. On the other hand, the performance of the model using the ensemble method is better than the other models. This model achieves an MSE of 0.0079, MAE of 0.057, RMSD of 0.0888, and an R^2 value of 0.84 for test data.

Keywords: Aqueous Two-Phase Systems (ATPS); Biomolecules; Machine learning; partition Coefficients; Morgan fingerprint

References

[1] Blacker, J. and M.T. Williams, *Pharmaceutical process development: current chemical and engineering challenges*. 2011: Royal Society of Chemistry.

[2] F. Surname1, F. Surname2, "Title of the Book", 3rd Ed., University of Tehran Publication, Tehran, 2022.

[3] F. Surname1, F. Surname2, Title of the book chapter in: "Title of the book", University of Science and Technology of Mazandaran, vol. 12, 2023, pp. 123-145.

ICB12-1402-1241

Ecotoxicological classification risk index for soil (ECRIS) prediction of some phenolic compounds using

multiple linear regressions and artificial neural network QSPR paradigm

[a] Z. Pahlavanyali, [b] M. H. Fatemi, [c] H. Ahangar Darabi, [d] Kobra Samghani
[a, b, c,d] Department of Chemistry, University of Mazandaran, Babolsar, Iran.
*Corresponding author's email address: zpahlevanyali@yahoo.com

Abstract: Ecotoxicology is study of toxic chemicals effects on biological organisms, especially at population, community, ecosystem and biosphere levels. Ecotoxicology is a multidisciplinary field integrates ecology, toxicology, physiology, analytical chemistry, molecular biology, and mathematics. The ecotoxicological effects are changes in state or dynamics of an organism or at other levels of biological organization, resulting from exposure to a chemical which these levels may include subcellular level, cellular level, tissues, individuals, populations, communities and finally ecosystems [1-2]. Ecotoxicological classification risk index for soil (ECRIS) is a new classification system specific for soil risk assessment which gives a comparative presentation of risk linked to environmental contamination by any chemical [3-4]. Data set chemicals of QSPR modeling were found in various landfills leachate of north Italy which the ECRIS values ranged from 1.32 to 58.44 for 2-Imidazolidinthyone and 4.4'-(Methylethylidene)bis-phenol. Linear and nonlinear models developed using multiple linear regressions (MLR) and artificial neural network (ANN) approaches. The MLR model between six selected descriptors and desired ECRIS values was $PECRIS = -9.797\ (\pm 2.948) + 4.619 * RDF045v\ (\pm 0.676) + 7.351 * MLOGP\ (\pm 1.117) - 10.066 * Mor13u\ (\pm 2.455) + 6.047 * HATS4u\ (\pm 2.645) + 56.753 * HATS5m\ (\pm 20.968) + 4.299 * Hy\ (\pm 1.670)$). In order to check any nonlinear relationships between selected molecular structural descriptors and ECRIS values, artificial neural network was applied by using STATISTICA software [5]. A three-layer network with sigmoid transfer function was designed; 6 descriptors and ECRIS respectively were used as inputs and outputs values. After training of network, it was used to predict of ECRIS values of data set for training, internal and external test sets. Robustness and reliability of constructed MLR and ANN models were evaluated by using leave-one-out cross-validation method that respectively was equal $Q^2_{MLR} = 0.84$ and $Q^2_{ANN} = 0.93$.

Keywords: Ecotoxicology; ECRIS; QSPR; Multiple linear regressions; Artificial neural network.

References

[1] Walker, C. H., Sibly, R. M., Hopkin, S. P., & Peakall, D. B. (2012). Principles of ecotoxicology. CRC press.

[2] Garcia-Reyero, N., & Perkins, E. J. (2011). Systems biology: leading the revolution in ecotoxicology. Environmental toxicology and chemistry, 30(2), 265-273.

[3] Boriani, E., Mariani, A., Baderna, D., Moretti, C., Lodi, M., & Benfenati, E. (2010). ERICA: A multiparametric toxicological risk index for the assessment of environmental healthiness. Environment international, 36(7), 665-674.

[4] Senese, V., Boriani, E., Baderna, D., Mariani, A., Lodi, M., Finizio, A., ... & Benfenati, E. (2010). Assessing the environmental risks associated with contaminated sites: definition of an ecotoxicological classification index for landfill areas (ECRIS). Chemosphere, 80(1), 60-66.

[5] Hilbe, J. M. (2007). STATISTICA 7: an overview. The American Statistician, 61(1), 91-94

ICB12-1402-1243

Integrative Bioinformatics Analysis of Breast Cancer-Derived Extracellular Vesicles in Mediating Cancer-Associated Fibroblast Activation

Mohammad Hosein Damirchi a, Motahareh Jahangir a, Mahboubeh sheikhbahaei a, Sadegh Babashah a*
Department of Molecular Genetics, Faculty of Biological Science, Tarbiat Modares University, Tehran, Iran
*Corresponding author's email address: babashah@modares.ac.ir

Abstract: Breast cancer, the most frequently diagnosed cancer and the second leading cause of cancer-related mortality in women, is strongly influenced by the tumor microenvironment (TME). Extracellular vesicles (EVs) secreted by cancer cells within the TME play a pivotal role in inducing various effects through diverse factors. Cancer-associated fibroblasts (CAFs) within the TME, activated by absorbing these EVs, contribute to the development of a malignant phenotype in breast cancer [1], [2]. Activation of CAFs is associated with the Akt pathway and its upstream regulators [3].
The RNA-seq data (accession code GSE106503) sourced from the NCBI database focuses on CAFs treated with EVs from the MDA-MB-231 cell line and CAFs treated with PBS. Data quality was assessed using FastQC, and trim were made with the Trimmoamtic tool. HISAT2 facilitated data mapping and alignment, with HTSeq-count analysis. Enrichr and the KEGG pathway database were used to explore pathways associated with upregulated genes, and STRING was used to examine protein-protein interaction networks.
The study aims to investigate the impact of breast cancer-derived EVs on CAFs and understand the underlying mechanisms. Among the upregulated genes influenced by EVs, ITGA11, ITGB5, THBS, and LAMB1 are implicated in the upstream regulation of AKT and PI3K pathways. Furthermore, a shared protein-protein interaction network involves ITGA11, ITGB5, THBS, LAMB1, COL1A1, COL5A1, COL6A3, and POSTN genes.
Considering the protein-protein interaction network among ITGA11, ITGB5, THBS, and LAMB1, and their role in activating AKT and PI3K signaling pathways, the increased expression of these genes due to breast cancer-derived EVs may significantly contribute to promoting breast cancer malignancy within the TME.
Keywords: Breast cancer; cancer associated fibroblast; tumor microenvironment; Extracellular vesicles

References

[1] A. Burguin, C. Diorio, and F. Durocher, 'Breast Cancer Treatments: Updates and New Challenges', *Journal of Personalized Medicine 2021, Vol. 11, Page 808*, vol. 11, no. 8, p. 808, Aug. 2021, doi: 10.3390/JPM11080808.

[2] X. Zhang *et al.*, 'Extracellular vesicles in the treatment and diagnosis of breast cancer: a status update', *Front Endocrinol (Lausanne)*, vol. 14, 2023, doi: 10.3389/FENDO.2023.1202493.

[3] Z. Fang *et al.*, 'Signaling pathways in cancer-associated fibroblasts: recent advances and future perspectives', *Cancer Commun*, vol. 43, no. 1, p. 3, Jan. 2023, doi: 10.1002/CAC2.12392.

ICB12-1402-1244

Integrative Bioinformatic analysis of differentially expressed genes and miRNA regulation in neoadjuvant chemotherapy resistance in HER2+ breast cancer

Mohammad Hosein Damirchi a, Motahareh Jahangir a, Mahboubeh sheikhbahaei a, Sadegh Babashah a*
Department of Molecular Genetics, Faculty of Biological Science, Tarbiat Modares University, Tehran, Iran
*Corresponding author's email address: babashah@modares.ac.ir

Abstract: Breast cancer is the most common type of cancer and the second leading cause of cancer-related deaths among women globally [1]. Patients who are initially diagnosed with locally advanced breast cancer are typically advised to undergo neoadjuvant chemotherapy (NAC) as a standard treatment approach. The effectiveness of therapy for breast cancer can vary among different populations. Hence, drug resistance is a significant factor contributing to the failure of cancer therapy [2], [3]. MiRNAs play a crucial role in negatively regulating genes by binding to the 3' untranslated region (3'UTR) of mRNAs and have been identified as significant contributors to drug resistance [4]. The main aim of this study is to identify Differentially Expressed Genes (DEGs) in patients sensitive and resistant to NAC to discover potential biomarkers related to drug resistance. We perform bulk RNA sequencing on the GSE162187 dataset to find hub genes. RNA sequencing was performed using the Galaxy server, and DEG was obtained with DESeq2. As a result, 63 DEG with p.value < 0.05 and adjusted p-value < 0.05 was found. Significant lncRNAs are up-regulated in resistant patients, such as LINC01291, LINC01285, and LINC01819. A candidate gene, GSTM1, had a significantly low expression in resistant patients (Log FC= -10.83). This gene belongs to the glutathione S-transferase family, which could provide resistance to several anticancer drugs by collaborating with efflux transporters and multidrug resistance proteins [5]. To discover miRNA that may have a role in down-regulating this gene in resistant patients, we used TarBase-v9.0, and hsa-miR-210-3p was found, which is shown in many studies that this miRNA upregulation is involved in cancer development and therapeutic response [6]. In conclusion, we found essential genes involved in NAC resistance and suggest an axis of GSTM1/hsa-miR-210-3p has a role in this manner.

Keywords: Breast cancer, Drug resistance, RNA sequencing, miRNAs

References

[1] B. Smolarz, A. Z. Nowak, and H. Romanowicz, "Breast Cancer—Epidemiology, Classification, Pathogenesis and Treatment (Review of Literature)," *Cancers (Basel)*, vol. 14, no. 10, p. 2569, May 2022, doi: 10.3390/cancers14102569.
[2] A. Daniyal, I. Santoso, N. H. P. Gunawan, M. I. Barliana, and R. Abdulah, "Genetic Influences in Breast Cancer Drug Resistance," *Breast Cancer: Targets and Therapy*, vol. Volume 13, pp. 59–85, Feb. 2021, doi: 10.2147/BCTT.S284453.
[3] Y. Xu *et al.*, "LncRNA MIR205HG expression predicts efficacy of neoadjuvant chemotherapy for patients with locally advanced breast cancer," *Genes Dis*, vol. 9, no. 4, pp. 837–840, Jul. 2022, doi: 10.1016/j.gendis.2021.10.001.
[4] A.-Y. Xing *et al.*, "Identification of miRNA Signature in Breast Cancer to Predict Neoadjuvant Chemotherapy Response," *Pathology and Oncology Research*, vol. 27, Apr. 2021, doi: 10.3389/pore.2021.1609753.
[5] P. Depeille *et al.*, "Glutathione *S* -Transferase M1 and Multidrug Resistance Protein 1 Act in Synergy to Protect Melanoma Cells from Vincristine Effects," *Mol Pharmacol*, vol. 65, no. 4, pp. 897–905, Apr. 2004, doi: 10.1124/mol.65.4.897.
[6] B. Pasculli *et al.*, "Hsa-miR-210-3p expression in breast cancer and its putative association with worse outcome in patients treated with Docetaxel," *Sci Rep*, vol. 9, no. 1, p. 14913, Oct. 2019, doi: 10.1038/s41598-019-51581-3.

ICB12-1402-1245

Down-regulation of the LAMB3 gene induces pulmonary metastasis in thyroid cancer

F. Goudarzi, SM. Javadirad*
Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: javadirad@yahoo.com

Abstract: The presence of distant metastases in thyroid cancer patients is only detected in 1% to 4% of patients with a poor prognosis. Metastasis sites for thyroid cancer are typically located in the lungs, bones, brain, and liver, with the lungs representing the most common metastatic site (43%). The overall survival rate for patients with lung metastasis ranges from 25% to 75%. Laminin Subunit Beta 3 (LAMB3) has been studied in lymph and liver metastasis, but no significant study has been conducted on thyroid tumor metastasis to lung tissue. In order to determine whether LAMB3 affects pulmonary metastasis, differentially expressed genes (DEGs) were identified from tissues with pulmonary metastasis. Nine datasets were selected from the gene expression omnibus (GEO) databank. Each data set was standardized using the Log2 transformation and quantile normalization. A meta-data set was constructed, and the limma statistical package was used to identify DEGs. In order to adjust P-values, the Benjamini-Hochberg procedure was used. To assess the specificity of the amplification, RT-qPCR reactions will be performed, and melt curve analysis will be performed. The Livak method would be used to assess the relative expression of LAMB3. Microarray meta-data analysis showed that the LAMB3 gene was significantly downregulated (logFC = -2.42931, adj.p.val = 1.02E-08). In conclusion, pulmonary metastases are characterized by a low expression of the LAMB3 gene, indicating that this gene plays a tumor-suppressive role.

Keywords: thyroid cancer; metastasis; pulmonary metastasis; differentially expressed genes; LAMB3

ICB12-1402-1246

Exploring Traditional Machine Learning and Deep Learning for Predicting Intracytoplasmic Sperm Injection (ICSI) Success Rates

Seyed Alireza Khanghahi[A], Seyedehfezeh HashemiKarouei[B], Seyed Gholamali Jorsaraee[B], Parviz Abdolmaleki[A,*]
a Department of Biophysics, Faculty of Biological Science, Tarbiat Modares University, Tehran, Iran.
b Fatemeh Zahra Infertility and Reproductive Health Research Center, Medical Science University of Babol, Babol, Iran.
*Corresponding author's email address: parviz@modares.ac.ir

Abstract: The utilization of Intra-cytoplasmic sperm injection (ICSI) in treating infertility among couples has significantly advanced patient care [1]. Despite its considerable expense, the success rates of ICSI remain disappointingly low, leading to significant anxiety for couples facing failed pregnancies. One of the effective ways to address this problem is to use state-of-the art machine learning algorithms [2-5]. This study endeavors to identify key predictors for forecasting the success of ICSI and enhance prediction accuracy through the application of support vector machines (SVM) and deep learning techniques.

A cohort of 345 patients undergoing ICSI treatment provided data comprising 29 numerical and nominal features. To address the issue of class imbalance within the dataset, Synthetic Minority Oversampling Technique (SMOTE) was employed. Two distinct methodologies were applied to construct predictive models. The first approach involved utilizing a genetic algorithm (GA) for feature selection in conjunction with SVM for prediction. Employing 3-fold cross-validation, the GA+SVM model identified 4 significant features and achieved an average accuracy of 0.728. The second method employed a state-of-the-art deep neural network. To ensure the robustness of predictions, the model underwent training and testing iterations five times with varying random seeds. This approach yielded an average accuracy of 0.737.

The findings indicate that deep neural networks surpassed SVM in predictive performance, underscoring the superior efficacy of deep learning models in forecasting the success rates of ICSI. This study contributes to advancing the understanding of factors influencing ICSI outcomes and underscores the potential of deep learning techniques in optimizing fertility treatment strategies.

Keywords: Intra-cytoplasmic sperm injection (ICSI); Machine Learning (ML); Genetic Algorithm (GA); Deep Learning;

References

[1]E. Hashemian Naeini, M. Zare Niyestanak, A. Mehrsay, E. Foroozandeh, M. Shokoohi, and K. Berjis, "Factors affecting pregnancy and abortion rates following intra cytoplasmic sperm injection in the obstructive and non-obstructive azoospermic men," *Medical Science Journal of Islamic Azad Univesity-Tehran Medical Branch,* vol. 22, no. 3, pp. 211-215, 2012.

[2]P. Abdolmaleki and S. G. Joorsaraei, "Prediction the success rate of intracytoplasmic sperm injection (ICSI) using artificial neural networks and logistic regression," *Razi Journal of Medical Sciences,* vol. 26, no. 6, pp. 118-126, 2019.

[3]M. Durairaj and K. Meena, "Application of artificial neural network for predicting fertilization potential of frozen spermatozoa of cattle and buffalo," *International Journal of Computer Science and System Analysis, Serials Publications,* vol. 2, no. 1, pp. 1-10, 2008.

[4]P. Malinowski, R. Milewski, P. Ziniewicz, A. J. Milewska, J. Czerniecki, and S. Wołczyński, "The use of data mining methods to Predict the Result of Infertility Treatment Using the IVF ET Method," *Studies in Logic, Grammar and Rhetoric,* vol. 39, no. 1, pp. 67-74, 2014.

[5]R. Milewski *et al.*, "Application of artificial neural networks and principal component analysis to predict results of infertility treatment using the IVF method," *Studies in Logic, Grammar and Rhetoric,* vol. 47, no. 1, pp. 33-46, 2016.

ICB12-1402-1247

Solving haplotype assembly problem by applying particle swarm optimization

Pourya Salati, Esmat Sadat Alaviyan Shahri, MohammadHossein Olyaee*
Electrical and Computer Engineering Department, University of Gonabad, Gonabad, Iran
*Corresponding author's email address: mh.olyaee@gonabad.ac.ir

Abstract: We propose the development of a sophisticated Particle Swarm Optimization (PSO) based methodology for reconstructing Single Individual Haplotypes (SIH). Haplotypes, representing an array of genetic variances on individual chromosomes, encapsulate crucial insights pertinent to the correlation between genomic sequences and pathological conditions. The identification of haplotypes in diploid organism is recognized as a computationally intensive task for which existing laboratory methods come with a high cost and depend on specialized apparatus. The task at hand entails the assembly of numerous DNA sequence fragments, each fragment partially revealing the composite haplotype. The pivot of this research is the efficient bi-partitioning of these fragments, minimizing inaccuracies as quantified by the Minimum Error Correction (MEC) metric. This problem is classified as NP-hard, a complexity level that has spurred numerous heuristic-based solution endeavors. Our proposed two-tiered method harnesses the PSO algorithm, a technique inspired by the social behavior of organisms. The initial phase involves the clustering of fragments utilizing a defined metric distance, effectively organizing the majority of the data. The ensuing phase capitalizes on PSO's rapid convergence properties to enhance the preliminary bi-partitioning achievements. The method is applied to various benchmark datasets, PSO showed significant promise. Through meticulous experimentation and analysis, it has been empirically ascertained that the PSO-based framework facilitates reliable SIH reconstruction, corroborating the algorithm's potential effectiveness in addressing this complex issue.

Keywords: Haplotype reconstruction; Particle swarm optimization; Minimum Error Correction; Graph Partitioning.

References

[1] Edge, Peter, Vineet Bafna, and Vikas Bansal. "HapCUT2: robust and accurate haplotype assembly for diverse sequencing technologies." *Genome research* 27.5 (2017): 801-812.

[2] Zamani, Fatemeh, Mohammad Hossein Olyaee, and Alireza Khanteymoori. "NCMHap: a novel method for haplotype reconstruction based on Neutrosophic c-means clustering." *BMC bioinformatics* 21 (2020): 1-15.

[3] Majidian, Sina, Mohammad Hossein Kahaei, and Dick de Ridder. "Minimum error correction-based haplotype assembly: Considerations for long read data." *Plos one* 15.6 (2020): e0234470.

ICB12-1402-1248

Microscopic Classification of Microbial Species Using Deep Learning

Mahdiar Mansouri [a], Mohsen Ebrahimi Moghadam[b], Mohammad Yaghoubi-Avini[a]*
a Department of Microbiology and Microbial Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
b Department of Artificial Intelligence, Robotics and Cognitive Computing, Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran
*Corresponding author's email address: m_yaghoubi@sbu.ac.ir

Abstract: This research presents a novel approach to the differentiation of Myxococcus species from other members of the Myxococcota phylum, leveraging the power of deep learning in microscopic image analysis. Myxococcus, known for its complex social behavior and unique predatory lifestyle, plays a crucial role in soil ecosystems and has significant biotechnological potential. However, the morphological similarity among Myxococcota makes it challenging to distinguish between species, especially at the microscopic level [1]. This study addresses this challenge by employing advanced image processing techniques and fine-tuning high-performance neural networks. The methodology involves a complex preprocessing pipeline to enhance the features of microscopic images of Myxococcus species. These images are then analyzed using a customized deep learning model. The model, initially based on existing high-performance neural networks, is fine-tuned to suit the specific textural and morphological characteristics of Myxococcus. This fine-tuning is critical in increasing the accuracy and specificity of the model in differentiating Myxococcus species from closely related taxa [2]. A significant portion of this research is dedicated to optimizing the preprocessing steps, including noise reduction, contrast enhancement, and feature extraction. These steps are crucial in ensuring that the neural network receives high-quality input data, thereby improving its learning efficiency and accuracy. Preliminary results demonstrate that this deep learning-based approach can significantly outperform traditional methods in identifying and differentiating Myxococcus species. This advancement holds immense potential for ecological studies and biotechnological applications, where accurate species identification is essential. This ongoing research is expected to contribute significantly to the field of bioinformatics by providing a robust, efficient, and automated solution for species differentiation in microscopic images. The final outcomes of this study aim to establish a new standard in microbial image analysis, opening new avenues for exploration in microbial ecology and systematics.

Keywords: Deep Learning; Myxococcota; Image Processing; Pattern Recognition

References

[1] H. Sajedi, F. Mohammadipanah, A. Pashaei "Automated identification of Myxobacterial genera using Convolutional Neural Network" , Sci Rep 9, 18238 (2019).

[2] C. Spahn, E. Gomez , R. Laine "DeepBacs for multi-task bacterial image analysis using open-source deep learning approaches", Commun Biol 5, 688 (2022).

ICB12-1402-1250

Molecular reasons for the effectiveness of Zinc in modulating Psoriasis

Siahmansouri, Siahmansouri, Siahmansouri*
Department of Genetics, University of Science and Culture, Tehran, Iran
*Corresponding author's email address: elmirasiahmansouri@yahoo.com

Abstract: Psoriasis is an autoimmune skin disease characterized by excessive proliferation and abnormal differentiation of keratinocytes (1). There is no cure for this disease. Using a study in the GEO DataSets Database that compared the expression in lesional and non-lesional skin (GSE161683), we separated the genes that had changed expression and filtered with logFC>2 and Padj<0.05 and reached 69 genes. Using the STRING database, we drew a network for these genes, finally we selected one of the hub genes named Krt16. Krt16 is a type I keratin and plays a key role in the skin as a regulator of innate immunity in response to skin damage (2). In the DrugBank database, there is the drug Clocortolone for this gene, and Zinc and Zinc Acetate have been introduced as targets for this protein. In a meta-analysis, it was seen that people with Psoriasis have low serum Zinc and high Copper, and the ratio of Copper to Zinc in these people is high (3). Overexpression of Krt16 has been observed in skin diseases characterized by proliferation such as Psoriasis (1). In addition, it has been reported that Krt16 is a marker of keratinocyte proliferation in Psoriasis in vivo and in vitro (1). In a study, by silencing the Krt16 gene with siRNA, the ERK signaling pathway, which is responsible for cell proliferation and VEGF secretion, was turned off and the disease improved (1). It seems that by balancing these two elements in the patient's body, hope for recovery can be brought to these patients.

Keywords: Psoriasis; Krt16; Zinc; Proliferation; ERK.

References

[1] Chen, Fun, "Silencing KRT16 inhibits keratinocyte proliferation and VEGF secretion in psoriasis via inhibition of ERK signaling pathway," Kaohsiung J Med Sci., 35 (2019) 284-296.

[2] https://www.uniprot.org/uniprotkb/P08779/entry

[3] Lei, Su, "Abnormal Serum Copper and Zinc Levels in Patients with Psoriasis: A Meta-Analysis," Indian J Dermatol., 64 (2019) 224-230.

ICB12-1402-1251

Performance evaluation of flexible-receptor docking tools based on the peptide length

M. H. Alizadeh[a], M. R. Ganjalikhany[a]*
[a] Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: m.ghanjalikhany@sci.ui.ac.ir

Abstract: Molecular docking tools play pivotal roles as a critical method in the drug design process in recent years [1]. There are various tools that estimate how ligands and receptors bind and how strong their affinity is [2]. Considering the widespread use of these tools, it is highly important to find out whether their results are accurate or not [3]. Moreover, taking into account the flexibility of receptors and ligands, imposes significant complexity on the docking calculations [4]. Peptides are among the most flexible ligands and have many applications e.g., as anticancer and antioxidant [5]. This study examines how the length of peptides with 3 to 8 amino acids affects the performance of three tools that account for flexibility, i.e., ADFR software, and HADDOCK and MedusaDock servers. The 190 protein-peptide complexes used for this purpose were carefully extracted from the PROPEDIA database and were subjected to necessary processing. After performing the focused-docking, the sampling power was evaluated by the lRMSD parameter and afterwards, the results were analyzed. The findings indicate that the average lRMSD declinations for each amino acid addition are 0.18, 0.71 and 0.82 Å for HADDOCK, MedusaDock and ADFR tools, respectively. Furthermore, the average success rate declinations including all distance cut-offs, for HADDOCK, MedusaDock and ADFR are 0.85%, 3.29% and 3.82%, respectively. This information demonstrates the superior and distinct performance of the HADDOCK over the other two tools. It seems that the peptide-specific calculations, similar to what we see in the algorithm of the HADDOCK, are essential for producing reliable results. This study provides insights for the selection and improvement of flexible docking tools for peptide-based drug design.

Keywords: Molecular docking; Flexible receptor; Peptide; Sampling power; lRMSD

References

[1] Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. Curr Comput Aided Drug Des. 2011;7(2):146-57.

[2] Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. Curr Top Med Chem. 2014;14(16):1923-38.

[3] Wang Z, Sun H, Yao X, Li D, Xu L, Li Y, et al. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. Phys Chem Chem Phys. 2016;18(18):12964-75.

[4] Huang SY. Comprehensive assessment of flexible-ligand docking algorithms: current effectiveness and challenges. Brief Bioinform. 2018;19(5):982-94.

[5] Muttenthaler M, King GF, Adams DJ, Alewood PF. Trends in peptide drug discovery. Nat Rev Drug Discov. 2021;20(4):309-25.

ICB12-1402-1252

Quantitative EEG Features to Diagnose Anxiety Disorders

M. Ansari* a, H. Tahayori, b
a Department of Computer Science and Engineering & IT, Shiraz University, Shiraz, Iran
b Department of Computer Science and Engineering & IT, Shiraz University, Shiraz, Iran
*Corresponding author's email address: marymansarii@gmail.com

Abstract: According to World Health Organization (WHO) report in 2019, 1 of 8 people live with a mental disorder. Anxiety disorders are the most common type of mental disorders. Anxiety disorders include disorders that share features of excessive fear and anxiety and related behavioral disturbances. Diagnosis of anxiety disorder is the first and the most important step towards its treatment.

While the clinical method is yet the most reliable method of diagnosis, use of biological signals, such as Electrocardiogram, electroencephalogram, etc. has attracted researchers' attention.

In this paper, we study the EEG signals and the information they convey to find a new way to diagnose anxiety disorders. For this purpose, we have selected different EEG features and perform 5 different classification methods on them. It is shown that EEG features can be a bio-marker of anxiety disorders. Obtained results show that functional connectivity (FC) feature sets achieved better results with Random Forest.

Keywords: Anxiety disorder; Electroencephalogram; EEG; Machine learning

References

[1] A. P. Association, Diagnostic and statistical manual of mental disorders (DSM-5 (R)), 5th ed, Arlington, TX: American Psychiatric Association Publishing, 2013.

[2] M. L. Otte, M. M. Schmitgen, N. D. Wolf and et al, "Structure/function interrelationships and illness insight in patients with schizophrenia: a multimodal MRI data fusion study," Arch Psychiatry Clin Neurosci, vol. 273, 2023.

[3] J. Oh, K. Yun, B. L. Oh, K. U. Lee, J. H. Chae and K. Yun, "Identifying Schizophrenia Using Structural MRI With a Deep Learning Algorithm," Front. Psychiatry, vol. 11, 2020.

[4] R. Chiong, G. S. Budhi, S. Dhakal and F. Chiong, "A textual-based featuring approach for depression detection using machine learning classifiers and social media texts," Computers in Biology and Medicine, vol. 135, 2021.

[5] A. Roy, K. Nikoltich, R. McGinn, S. Jinah, W. Klement and Z. A. Kaminsky, "A machine learning approach predicts future risk to suicidal ideation from social media data," npj Digital Medicine, vol. 3, 2020.

[6] D. Liu, X. L. Feng, F. Ahmed, M. Shahid and J. Guo, "Detecting and Measuring Depression on Social Media Using a Machine Learning Approach: Systematic Review," JMIR MENT HEALTH, vol. 9, 2022.

[7] D. Zarate, M. Ball, M. Prokofieva, V. Kostakos and V. Stavropoulos, "Identifying self-disclosed anxiety on Twitter: A natural language processing approach," Psychiatry Research, vol. 330, 2023.

[8] J. Kim, J. Lee, E. Park and J. Han, "A deep learning model for detecting mental illness from user content on social media," Scientific Reports, vol. 10, 2020.

ICB12-1402-1253

Elucidating the Impact of Opium Consumption on Blood Parameters Using Network Analysis and Machine Learning

Kiarash Zare [a], Zahra Salehi[b, c], Marziyeh Gholami[d], Fateme Vafaee Shaarbaf[d], Mohamad reza Zabihi[d], Ali reza Morovat[a], Mohammad Mehdi Naghizadeh[a], Azizallah Dehghan[a], Kaveh Kavousi[d]*

[a] Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran

[b] Hematology, Oncology and Stem Cell Transplantation Research Center, Tehran University of Medical Sciences, Tehran, Iran.

[c] Research Institute for Oncology, Hematology and Cell Therapy, Tehran University of Medical Sciences, Tehran, Iran

[d] Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran

*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Despite its prevalent use, the specific biochemical effects of opium on blood parameters have not been comprehensively elucidated. Utilizing data from the comprehensive Persian FASA cohort [1], our study harnesses the power of advanced statistical modeling and machine learning techniques to systematically investigate these effects. By integrating traditional substance use research with cutting-edge bioinformatics, this study aims to shed light on the intricate biochemical dynamics induced by opium consumption, thereby filling a critical knowledge gap in the field. The study involved 10,138 participants. It started with thorough data preprocessing, along with the implementation of strict inclusion criteria. This process guaranteed that individuals with pre-existing medical conditions were not included in the study. In the analysis based on machine learning, 80% of the data was used in the training phase and 20% for testing. We then embarked on a detailed correlation analysis of all available parameters. Parameters with a correlation coefficient exceeding |0.3| were selected for further network analysis. This analysis was conducted using Cytoscape software, leading to the construction of a detailed correlation network, distinctly highlighting the relationships between various blood parameters. Our results revealed unique correlations in the opium user group, such as RBC-MCHC, PLT-MPV, TG-HDL.C, and CHOL-GGT, which were absent in the normal group. To quantify the differences between healthy and opium-addicted individuals, we employed the random forest algorithm, achieving an accuracy of 82%, precision of 92%, recall of 69%, and an F1 score of 79%. The metrics demonstrate the model's effectiveness and highlight the physiological effects of opium on blood parameters. This study highlights the efficacy of combining mathematical approach to unravel the complex effects of opium on blood parameters [2]. The random forest algorithm was used to identify unique biochemical patterns in opium users, highlighting the potential of advanced computational techniques in biomedical research.

Keywords: Opium; Machine learning; Network analysis; Blood

References

[1] Eghtesad S, Mohammadi Z, Shayanrad A, Faramarzi E, Joukar F, Hamzeh B, et al. The PERSIAN cohort: providing the evidence needed for healthcare reform. Archives of Iranian medicine. 2017;20(11):691-5.

[2] Shahabinejad G, Sirati-Sabet M, Kazemi-Arababadi M, Nabati S, Asadikaram G. Effects of opium addiction and cigarette smoking on hematological parameters. Addiction & health. 2016 Jul;8(3):179.

ICB12-1402-1254

Investigation of FLS1 Enzyme Involved in the Flavonoids Biosynthesis in Carthamus tinctorius Using Bioinformatics Techniques

Homa Zarabizadeh[1], Mohammad Sadegh Abbasi[2], Mozhgan Latifi[3], Asadolah Ahmadikhah[4]
1,2,3,4 Shahid Beheshti University
Corresponding authors e-mail: helia.zarabi@gmail.com, sadegh.abasi@ut.ac.ir, m_latifi@sbu.ac.ir,
*Corresponding author's email address: a_ahmadikhah@sbu.ac.ir

Abstract: Flavonoids are a diverse group of over 8000 polyphenolic compounds found in plants, playing various roles in plant physiology. Particularly, flavonols, a crucial subclass, contribute to functions like UV protection, auxin transport regulation, flower color modulation, and signaling. These compounds, known for their antioxidant and health-promoting properties, also exhibit anti-proliferative, anti-angiogenic, and neuropharmacological effects [1], [2]. The transformation from dihydroflavonol to flavonol is facilitated by the FLS enzyme in the flavonoid pathway. This process relies on the coordinated expression of genes encoding core phenylpropanoid pathway enzymes, like PAL and 4CL, and flavonoid branch pathway enzymes such as CHS, CHI, and F3H [3]. FLS1 is the predominant FLS synthesis relevant gene in *Carthamus tinctorius*, the last step in flavonol biosynthesis [4]. In the present investigation, the nucleotide sequence of the FLS1 enzyme in the *Carthamus tinctorius*, having the accession number MF421815.1, was procured from the NCBI website. The chemical characteristics of the protein were assessed using the ProtParam program, while the domains were established by employing the Interpro server. The coding arrangement of the FLS1 enzyme in the *Carthamus tinctorius*, encompassing 1312 nucleotides, was responsible for encoding a protein comprising 410 amino acids. The molecular weight of this protein is 47512.62 Daltons, with an Isoelectric Point (PI) of 19.10, instability index of 47.53, and an Aliphatic Index, which is considered a positive factor for increasing protein stability is 70.66. Gene network analysis using the STRING database revealed that the FLS1 protein is involved in the biosynthesis pathway of flavone and flavonol, showing the highest interaction with *DFRA*, *CYP7385*, *UGT78D1* and *MYB12* genes. Peptide mapping of the protein is hypothetically and theoretically determined by expasy/peptidemass database.

Keywords: FLS1 enzyme; Carthamus tinctorius; flavonoid pathway; Domains

ICB12-1402-1255

A Novel Mechanistic Simulation Model for Single-Cell DNA Sequencing

Sajedeh Bahonar[1], Laura Tomas[2], Joao Miguel Alves[2], David Posada[2,3], Hesam Montazeri[1,*]
[1] Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran
[2] Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain
*Corresponding author's email address: hesam.montazeri@ut.ac.ir

Abstract: Single-cell DNA sequencing technology amplifies tiny DNA quantities through two primary methods: PCR-based and polymerase-based, with the latter exhibiting lower error rates, particularly advantageous for detecting single-nucleotide variations. The second category employs Multiple Displacement Amplification (MDA) [1] and Primary Template-Directed Amplification (PTA) [2]. MDA produces longer amplicons compared to PTA, resulting in a higher amplification imbalance. Simulating the amplification process is crucial for evaluating variant callers, yet there is a noticeable absence of mechanistic simulation tools for MDA and PTA. This study introduces a new simulation tool tailored for both techniques.

In the initial step, single-cell genomes are simulated, encompassing SNPs, SNVs, and CNVs. The simulated single-cell genome comprises maternal and paternal strands. Each amplicon is defined by parameters such as start position, end position, direction, and release status. The algorithm generates new amplicons through hexamer attachment and extension, introducing amplification errors. Additional parameters include hexamer attachment rate, maximum length, and the probability of amplification error. The simulation incorporates allelic imbalance, with biased selection of maternal or paternal regions as cycles progress. Subsequently, the produced amplicons FASTA file is used to generate the single-cell reads as FASTQ file. Real single-cell datasets of PTA and MDA were analyzed for simulation evaluation. Germline and somatic variant distributions were determined using the HaplotypeCaller and Mutect, respectively. We offer a computationally efficient implementation of this simulation model in Python.

The dataset consists of 10 PTA and 10 MDA single-cell DNAseq records was obtained from the SRA database under the accession code SRP178894. Analysis reveals distinct distribution patterns between PTA and MDA. The simulator's effectiveness was evaluated using the real dataset, revealing non-uniform coverage in alignment with actual single-cell datasets and Variants Allele Frequency (VAF) distribution. It facilitates the comparative assessment of various SNV callers, such as ProSolo [3] and SCAN-SNV [4].

Keywords: Simulation; Single Cell; DNAseq; MDA; PTA

References

[1] Spits, et al. (2006). Whole-genome multiple displacement amplification from single cells. Nature protocols, 1(4), 1965-1970.

[2] Gonzalez-Pena, et al. (2021). Accurate genomic variant detection in single cells with primary template-directed amplification. Proceedings of the National Academy of Sciences, 118(24), e2024176118.

[3] Lähnemann, et al. (2021). Accurate and scalable variant calling from single cell DNA sequencing data with ProSolo. *Nature communications*, *12*(1), 6744.

[4] Luquette, et al. (2019). Identification of somatic mutations in single cell DNA-seq using a spatial model of allelic imbalance. *Nature communications*, *10*(1), 3908.

ICB12-1402-1256

A platform for *de novo* peptide sequence library generation

F. Soorani[a], A.M. Zehtab[b], M.R. Ganjalikhany[a]*
[a] Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
[b] Department of Computer Engineering, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: m.ghanjalikhany@sci.ui.ac.ir

Abstract:: peptides are short chains of proteins composed of 20 amino acids. They have numerous functions in cells and organisms including antimicrobial, immune regulatory, hormone, pain control and many other instances [1]. Recently, peptides have attracted much attention due to their advantages in the clinical applications. Based on the reports, more than 60 approved peptide-base drugs are on the market now [2]. Sequence, structural and functional diversity of peptides could create millions of potentially therapeutic molecules. Non-toxicity, ease of synthesis, non-allergenic properties of peptides make them suitable as novel class of drugs [3]. Despite of their applications, there is no decent and dedicated platform for peptide sequence library generation and manipulation. Available tools for peptides are mostly designed for specific applications and they are not capable of generating and producing user defined peptide library. For example, Pepfold server [4] only predict the 3D structure of single peptide or Pepdraw [5] only calculate the net charge and draw 2D structure of a single peptide sequence. Peptidrive could design only a peptide sequence based on the 3D structure of a template [6]. Our designed platform is a python-based software which uniquely developed for peptide sequence generation and manipulation in large scale including: 1. *de novo* peptide sequence generation based on length, 2. Systematic mutation and insertion on the predefined sequence, 3. Uni- and bi-directional incremental sequence construction, 4. Net-charge calculation based on the peptide sequence list. Currently, *de novo* design module can generate sequences covering 2-10 amino acids using breadth first search algorithm. The software is still under development and other functions will be implemented such as solubility calculation of peptide library and generation of 3D structures based on a peptide sequence list. The current platform enable users to generate and manipulate peptide sequence in large scale for drug screening applications.

Keywords: peptide; drugs; python-based platform; de novo sequence generation; sequence manipulation

References

[1] Saladin KS. Anatomy & physiology : the unity of form and function. 6th ed. New York, Ny: Mcgraw-Hill; 2012.
[2] Agarwal G, Gabrani R. Antiviral Peptides: Identification and Validation. Int J Pept Res Ther. 2021;27(1):149-68.
[3] Wang L, Wang N, Zhang W, Cheng X, Yan Z, Shao G, et al. Therapeutic peptides: current applications and future directions. Signal Transduct Target Ther. 2022;7(1):48.
[4] https://openebench.bsc.es/tool/pep-fold
[5] https://pepdraw.com/
[6] https://rosie.rosettacommons.org/peptiderive/submit

ICB12-1402-1257

Identification of Genes and Pathways Associated with Depression in Ovarian Cancer Patients: An In-silico Study

Yussef Hashemi[a], Kiarash Zare[a,c], Ali reza Morovat[a], Yaser Mansoori[b], Mohammad Mehdi Naghizadeh[b,c]*
a Student Research Committee, Fasa University of Medical Sciences, Fasa, Iran
b Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran
c Department of Medical Bioinformatics, Fasa University of Medical Sciences, Fasa, Iran
*Corresponding author's email address: mmnaghi2@yahoo.com

Abstract: Depression plays an important role in causing pathophysiological changes in the tissue and increases the complexity of ovarian cancer pathology. We reconstructed and analyzed the miRNA-mRNA regulatory interactions for a better understanding of ovarian cancer-related depression. We identified the differentially expressed genes between depressed and non-depressed ovarian cancer groups and removed ovarian cancer-related genes. Then reconstruct a PPI and cluster it into modules. Inter-module miRNA-mRNA regulatory networks were studied, and analyzed their gene ontology. The ovarian cancer-related depression had 875 genes. The hubs in Regulatory modules were CTNNB1, HRAS, CDH1, CD8A, SMAD4, CCL2, and miR-200a, miR-124-3p, miR-194-5p, let-7b-5p that play a role in ovarian cancer-related depression pathways like neurotransmitter release cycle and inflammation. The results of this study can provide insight into the pathomechanism of ovarian cancer-related depression.

Keywords: microRNA-mRNA interaction; network analysis; ovarian neoplasm; molecular psychiatry; depression etiology

References

[1] Malhi GS, Mann JJ. Depression. Lancet. 2018;392(10161):2299-312.

[2] Rahman MR, Islam T, Al-Mamun MA, Zaman T, Karim MR, Moni MA. The influence of depression on ovarian cancer: Discovering molecular pathways that identify novel biomarkers and therapeutic targets. Informatics in Medicine Unlocked. 2019; 16:100207.

ICB12-1402-1258

Investigating the Molecular Docking of Herbal Compounds Quercetin, Isorhamnetin, and Thymoquinone Against Human Epidermal Growth Factor Receptor 2

E.Hoseinnia, S.Ghodrati, R.Habibi*
a Department of Biotechnology, Alzahra University, Tehran, Iran
*Corresponding author's email address: R.Habibi@student.alzahra.ac.ir

Abstract: Cancer is the second leading cause of death worldwide, affecting more than 9 million people annually. Breast cancer is one of the most common cancers among women. It has become one of the deadliest cancers due to late detection and resistance to chemotherapy. Human epidermal growth factor receptor 2 (HER2) is overexpressed in 25-30% of breast cancer cases, so targeting the HER2 receptor is an effective therapeutic strategy in patients with HER2-positive breast cancer. This study was conducted to investigate the potential of herbal compounds in treating breast cancer using the molecular docking method. The active compounds studied are Quercetin from *Foeniculum Vulgare*, Isorhamnetin from *Raphanus Sativus*, and Thymoquinone from *Nigella Sativa*. The molecular binding of these compounds to the HER2 receptor (3PP0) was investigated using AutoDock Vina software. The binding energy of these compounds and their interaction with HER2 receptors were studied and compared with Lapatinib. Lapatinib is an antineoplastic agent and tyrosine kinase inhibitor used to treat advanced or metastatic HER2-positive breast cancer in patients who have received prior chemotherapy treatments. In this study, Quercetin had the lowest binding energy compared to the other two compounds, and Isorhamnetin established the highest number of hydrogen bonds. This study showed that the herbal compounds Quercetin, Isorhamnetin, and Thymoquinone have significant potential in breast cancer treatment. Of course, more studies are needed to evaluate their effectiveness in the laboratory and clinical environment.

Keywords: Breast Cancer; HER2; Herbal Component; Molecular Docking

References

[1] National Cancer Institute. What is cancer? Cancer.gov. https://www.cancer.gov/about-cancer/understanding/what-is-cancer. Updated February 9, 2015. Accessed November 6, 2020.

[2] Thota, V., et al. (2020). "A REVIEW ARTICLE ON BREAST CANCER." World Journal of Pharmaceutical Research.

[3] DC Allred. "Ductal Carcinoma In Situ: Terminology, Classification, and Natural History." Journal of the National Cancer Institute Monographs, 2010

[4] Carrasco-García, E., M. Saceda, and I. Martínez-Lacaci, Role of receptor tyrosine kinases and their ligands in glioblastoma. Cells, 2014. 3(2): p. 199-235.

[5] Incorvati, J.A., et al., Targeted therapy for HER2 positive breast cancer. Journal of hematology & oncology, 2013. 6(1): p. 1-9.

[6] Shi, Z., et al., Discovery of novel human epidermal growth factor receptor-2 inhibitors by structure-based virtual screening. Pharmacognosy magazine, 2016. 12(46): p. 139.

[7] Sait, K.H.W., et al., Molecular docking analysis of HER-2 inhibitor from the ZINC database as anticancer agents. Bioinformation, 2020. 16(11): p. 882

ICB12-1402-1259

Transcriptomics-Based Computational Drug Repurposing Strategy Identifies Therapeutic Candidates for Parkinson's Disease

Atena Vaghf[a], Zahra Taslimi[b], Nayereh abdali[c], Shahram Tahmasebian[*]
[a] Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
[b] Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
[c] Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
[*] Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord, Iran
*Corresponding author's email address: stahmasebian@gmail.com

Abstract: Parkinson's disease (PD) ranks as the second most prevalent age-related neurodegenerative disorder which associated with degeneration of dopaminergic neurons within the substantia nigra (SN) [1]. Currently, no definitive disease-modifying treatment exists, and various treatments have been developed to manage the symptoms of PD [2]. Drug repurposing is a valuable alternative approach to uncovering new indications of approved or investigational drugs that beyond of their original indication. RNA sequencing (RNA-seq) is one effective approach to finding the heterogeneous gene expressions of diseases in response to specific drugs. Therefore, our study applied a computational drug repurposing pipeline to explore the candidate drugs by PD differential gene expression signatures derived from RNA sequencing data. The expression profiles of human post-mortem PD striatum under the *accession* code GSE205450 were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). This dataset included 40 controls and 35 PD samples. The differentially expressed genes (DEGs) between PD tissues and normal tissues were obtained by using GEO2R. Next, the Library of Integrated Network-based Signatures (LINCS) database was used to identify potential candidate drugs which can reversed the expression of DEGs. Then, through considerable literature review and drugbank (https://go.drugbank.com) studies, the top-ranked drugs with highest p-value were selected. This study identified 562 genes with $|\log_2 FC|>1$ and P-value <0.05 as DEGs: 390 upregulated and 172 downregulated genes. In drug list, there are drugs for the treatment of cancer and non-cancer diseases, among which Mitoxantrone and Alvocidib can be mentioned. Mitoxantrone is initially developed as chemotherapeutic agent and then, it used for multiple sclerosis. Alvocidib is a cyclin-dependent kinase (CDK) inhibitor that exerts antitumor activity. In conclusion, this study proposed probably candidates (Mitoxantrone and Alvocidib) for the treatment of PD progression that its can guide further repurposing studies tailored to different stages of disease progression.

Keywords: RNA sequencing; Parkinson's disease; Drug repurposing

References

[1] Bloem, B.R., M.S. Okun, and C. Klein, Parkinson's disease. The Lancet, 2021. **397**(10291): p. 2284-2303.

[2] Church, F.C., Treatment options for motor and non-motor symptoms of Parkinson's disease. Biomolecules,2021. **11**(4): p. 612

ICB12-1402-1260

Differentiation of glioblastoma multiforme solid and recurrent tumors using gene expression profile and machine learning algorithms

R. Zohreh[A], N. Ayareh[B], F. Montazeri Moghaddam[C], E. Amanzadeh Jajin*
a Department of Biology, Central Tehran Branch, Islamic Azad University, Tehran, Iran
b member of students' research committee of Shiraz University of Medical Sciences, Shiraz, Iran
c Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
*Functional neurosurgery Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: elnazamanzadehjajin@gmail.com

Abstract: The most prevalent type of brain tumor is called grade IV glioma, or glioblastoma multiforme (GBM). GBM is among the deadliest tumors, with an extremely poor survival rate. GBM patients have a 5-year survival rate of only 5% [1]. In the present research, we used bioinformatics methods to discover a group of genes that demonstrate notable differential expression between GBM patients and normal samples. Also, we applied machine learning algorithms to find potential biomarkers for GBM patients. The R language program "TCGAbiolinks" was used to download the TCGA-GBM transcripts per million (TPM) data comprising 175 samples: 13 recurrent glioblastomas, 157 primary tumors, and 5 solid tissue normal samples. Using an adjusted p-value of less than 0.05, the top 10 differentially expressed genes between GBM and normal samples in the TCGA-GBM cohort were selected with the R programming package "DESeq2." The LASSO algorithm, also known as the Least Absolute Shrinkage and Selection Operator, was utilized for feature selection, along with other methods such as random forest classifier (RFC), support vector classifier (SVC), and artificial neural networks (ANNs) to distinguish between the datasets. The differentially expressed top 10 genes (MAPK9, MIGA1, WDR7, CDKL5, C2CD2L, AAK1, AKAP5, UBR3, CPEB3, and SYNJ1) were screened by DESeq2 analysis. LASSO analysis led to the identification of 10 genes (ANK3, GCNT4, CD177P1, LINC02974, CDKL5, GAS7, ADRB3, KLHL2, LINC02500, and LDHAL6EP) as the feature selections linked to the prognosis of glioblastoma. The area under the curve (AUC) of ANNs, RFC, and SVC using raw data was 0.49, 0.94, and 0.91, as well as 0.69, 0.94, and 0.08 using selected features, respectively.

Keywords: Glioblastoma Multiforme; gene expression; machine learning; DESeq2; LASSO

References

[1] S. M. Mousavi et al., 'Non-coding RNAs and glioblastoma: Insight into their roles in metastasis', Molecular Therapy Oncolytics, vol. 24. Cell Press, pp. 262–287, Mar. 17, 2022. doi: 10.1016/j.omto.2021.12.015.

ICB12-1402-1262

Identifying key miRNAs and mRNAs involved in EMT in papillary thyroid carcinoma using bioinformatics techniques

S.yazdi rouholamini, S. Javadirad*
Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: sm.javadirad@sci.ui.ac.ir

Abstract: The most common type of thyroid cancer is papillary thyroid carcinoma (PTC), which accounts for approximately 85% of differentiated thyroid tumors. The five-year survival rate for patients with PTC is 97%; however, if the tumor is aggressive, this rate decreases significantly[1,2]. The present study examined critical miRNAs involved in PTC. A non-coding array dataset (GSE113629) was obtained from the GEO database, and miRNA-seq data for PTC was obtained from the TCGA database using the TCGAbiolinks package[3]. The Limma[4] and edgeR[5] packages analyzed the differentially expressed miRNAs (DEMs) in PTC and normal thyroid tissues. A list of all miRNAs involved in EMT was compiled using the EMT database. By comparing DEMs with EMT miRNAs, common DEMs were identified with ggvenn package. Using the miRDB and miRTarBase databases, we predicted mRNA targets for common DEMs. In order to analyze the predicted mRNAs further, those known to play a significant role in EMT were selected. A miRNA-mRNA regulatory network was constructed. The diagnostic ability of miRNAs was assessed by performing ROC analysis on tumor and normal samples. Among the six DEMs, four were associated with the EMT pathway, namely miR-221, miR-375, miR-34a, and miR-222. It was found that miR-221-3p and miR-222-3p regulate seven common target genes, including: BBC3, MDM2, ETS1, CDKN1B, DICER1, TRPS1, and ESR1. Four miRNAs also targeted the KIT gene. As a potential biomarker for PTC, there is a candidate regulatory mRNA-miRNA network whose members are all involved in the EMT pathway.

Keywords: Papillary thyroid carcinoma; Regulatory Network; EMT Pathway

References

[1] Zaballos MA, Santisteban P. Key signaling pathways in thyroid cancer. Journal of Endocrinology. 2017;235(2).

[2] Hong S, Xie Y, Cheng Z, Li J, He W, Guo Z, et al. Distinct molecular subtypes of papillary thyroid carcinoma and gene signature with diagnostic capability. Oncogene. 2022;41(47).

[3] Colaprico A, Silva TC, Olsen C, Garofano L, Garolini D, Cava C, et al. TCGAbiolinks: TCGAbiolinks: An R/Bioconductor package for integrative analysis with GDC data [Internet]. 2024. Available from: https://github.com/BioinformaticsFMRP/TCGAbiolinks

[4] Smyth MG. Package "limma" Title Linear Models for Microarray Data. 2024.

[5] Chen Y, Lun AT, Mccarthy DJ, Chen L, Ritchie ME, Phipson BL, et al. edgeR: Empirical Analysis of Digital Gene Expression Data in R. 2024.

ICB12-1402-1263

Whole Exome Sequencing of a Family with Non-obstructive Azoospermia Men did not Show the C677T Variant as an Infertility Risk Factor

Mousavi SZ[1], Ricco A[2], Totonchi M[3*] Soltani B[1]
[1]Department of Molecular Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran.
[2]Department of Environmental Biological and Pharmaceutical Sciences and Technologies, Università degli Studi della Campania "Luigi Vanvitelli", 81100 Caserta, Italy
[3]Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran.
*Corresponding author's email address: m.totonchi@royaninstitute.org

Abstract: It is widely shared that the genetic variants in the genes encoding enzymes involved in folate metabolism are important risk factors for male infertility [1]. Since the MTHFR enzyme has a significant role in the human folate and homocysteine metabolism, the folic acid supplements have reported to increase total sperm count [2]. The association between (chr1.exon5:c.C667T:p.A222V) variant in the coding region of the human MTHFR gene which leads to the replacement of the alanine number 222 by valine (exon5:c.C667T:p.A222V) is controversial. Some studies reported this variant as pathogenic with almost 30% and 70% less enzymatic activity in heterozygotes (CT) and mutant homozygotes individuals (TT) respectively, whereas other studies did not [3,4]. So, this variation ought to be the subject of more studies to determine its pathogenic association with male infertility. This study aimed to investigate the presence of the A222V variant in fertile and infertile individuals of an Iranian family. Whole exome sequencing (WES) in the family under study with three non-obstructive azoospermia and two normal individuals referred to the Infertility and Reproductive Biomedicine Research Center of Royan Institute, Tehran, Iran was performed. After bioinformatic analysis of the obtained data, the heterozygous C677T variant in both affected and normal siblings was detected and the variant was not co-segregated with phenotype in the family members. Our investigation showed that the MTHFR C677T variant is not a risk factor for male infertility.

Keywords: WES, MTHFR, rs1801133, non-obstructive azoospermia, infertility

References

[1] Al-Sammarraie HK, Nader MI, et al. Effect of C677T Mutation of Methylene Tetrahydrofolate Reductase Gene as a Genetic Risk Factor in Iraqi Infertile Men. Iraqi journal of biotechnology. (2021) 1(20).

[2] Montjean D, Benkhalifa M, et al. Polymorphisms in MTHFR and MTRR genes associated with blood plasma homocysteine concentration and sperm counts. Fertility and sterility. (2011) 1;95(2):635-40.

[3] Ni W, Li H. Lack of association between genetic polymorphisms in three folate-related enzyme genes and male infertility in the Chinese population. Journal of assisted reproduction and genetics. (2015) 32:369-74.

[4] Han LJ, He XF, et al. Methylenetetrahydrofolate reductase C677T and A1298C polymorphisms and male infertility risk: An updated meta-analysis. Medicine. (2020) 12;99(51).

ICB12-1402-1264

Screening of cyanobacterial bioactive peptides against viral infections

Amir kananian[a], Ahmad Farhad Talebi[a]*
a Department of Biotechnology, Faculty of New Sciences and Technologies, Semnan University, Semnan, Iran
*Corresponding author's email address: aftalebi@semnan.ac.ir

Abstract: The risk of spreading viral infections increasingly threatens public health. Unfortunately, the number of effective and available drugs for the prevention and treatment of these types of infections is limited [1]. Management of viral infections in the absence of prevention tools brings costs and has serious effects on the economy and social life [2]. The most effective way to rapidly identify antiviral compounds is high-throughput screening of the libraries containing existing drugs or random molecules. Such analysis provides potential drug candidates, however, these efforts are usually not cost-effective and often time-consuming [3]. Boosting immunity is a simple way to resist viral infection. The use of immune-enhancing nutrients has the potential to control viral infections by activating the immune response. Cyanobacteria are a rich biological source of bioactive compounds with therapeutic importance [4]. In this research, among 60 proteins with possible antiviral function, during 10 step by step screening, ones that had the ability to inactivate ACE2 enzyme were selected. The 3 selected proteins not only represented acceptable biochemical and pharmacokinetic properties, but also had significant binding energy compared to commercial antiviral drugs. Bioactive peptides are defined as specific protein fragments that have a positive effect on body performance and health. With the help of bioinformatics tools, it is possible to discover bioactive peptides with enhanced health-care properties, especially against viral infections. The discovery of medicinal metabolites, especially from the branch of aquatic biotechnology, has opened a promising future for us in the face of the spread of drug-resistant strains.

Keywords: drug discovery; viral infections; bioinformatics analysis; bioactive prospecting.

References

[1] H. Mazur-marzec, M. Cegłowska, R. Konkel, and K. Pyr, "Antiviral Cyanometabolites — A Review. Biomolecules, 11, no. 3 (2021) 474.
[2] M. Nicola et al., "The socio-economic implications of the coronavirus pandemic (COVID-19): A review," Int. J. Surg., 78 (2020): 185-193.
[3] G. Riccio et al., "Ten-year research update review: Antiviral activities from marine organisms," Biomolecules 10, no. 7 (2020): 1007.
[4] S. K. Ratha, N. Renuka, I. Rawat, and F. Bux, "Prospective options of algae-derived nutraceuticals as supplements to combat COVID-19 and human coronavirus diseases" Nutrition 83 (2021): 111089.

ICB12-1402-1265

A Deep Learning Approach for Peptide-based Treatment Design

Behnaz Lotfi, Mahboubeh Zarrabi, Sajjad Gharaghani
Department of Mathematical Sciences, Alzahra University, Tehran, Iran
Department of Biological Sciences, Alzahra University, Tehran, Iran
Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran
*Corresponding author's email address: b.lotfi@student.alzahra.ac.ir

Abstract: The great Human Genome Project revealed the heterogeneity of human society. Genetic variants drive the healthcare process and treatment design toward individualization and per-person treatment methods. Finding the causative agent and rational drug design, to increase or inhibit the activity of the target protein is the first stage of treatment design. Side effects and drug resistance have increased the interest in exploiting natural compounds.
The use of therapeutic peptides has been known as a therapeutic strategy for years due to their lower toxicity and immunogenicity, type of physical and chemical structure, and specificity. However, the expensive production process and the limitations of human drug testing, as well as the existence of big data, have caused computational methods based on machine learning and deep learning to be used to identify the effective features of peptide therapy or to introduce new therapeutic peptides for synthesis with generative algorithms.
In the current research, while introducing a new data library consisting of 162 pairs of therapeutic peptides and target proteins of the diseases, which were collected from bioinformatics servers and articles, feature extraction was done using the iFeature python toolkit [1], and data classification was conducted by Random Forest algorithm, which 80% precision and 78% accuracy were achieved. Next, according to the linguistic nature of peptide and protein sequences, another classification was modeled from the deep learning method based on the transformer self-attention mechanism and the embedding vector resulting from the encoder layer in the ProtBERT [2] architecture, which in comparison to the machine learning model, respectively It resulted in a precision and accuracy of 71% and 78%. This research shows that despite the limitations of the peptide data library, the classification of therapeutic peptides with the transformer model can compete with the results of the machine learning model.
Keywords: Peptide; Treatment Design; Machine Learning; Transformer

References

[1] Chen, Zhen, Zhao, Pei, Li, Fuyi, Leier, André, Marquez-Lago, Tatiana T, Wang, Yanan, Webb, Geoffrey I, Smith, A Ian, Daly, Roger J, Chou, Kuo-Chen, et al. "iFeature: a python package and web server for features extraction and selection from protein and peptide sequences." Bioinformatics, 34(14):2499–2502, 2018.

[2] Elnaggar, Ahmed, Heinzinger, Michael, Dallago, Christian, Rehawi, Ghalia, Wang, Yu, Jones, Llion, Gibbs, Tom, Feher, Tamas, Angerer, Christoph, Steinegger, Martin, et al. "Prottrans: Toward understanding the language of life through self-supervised learning." IEEE transactions on pattern analysis and machine intelligence, 44(10):7112–7127, 2021.

ICB12-1402-1266

Predicting the impact of missense mutations on the protein structure networks: a case study on human phenylalanine hydroxylase and α-galactosidase

Fateme Ghandali [a], Sayed Amir Marashi [b]*, Golnaz Bahramali,[c]*, Kolsoum InanlooRahatloo [a]
[a] School of Biology, College of Science, University of Tehran, Tehran, Iran
[b] Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran
[c] Department of Hepatitis, AIDS and Bloodborne Diseases, Pasteur Institute of Iran, Tehran, Iran
*Corresponding author's email addresses: marashi@ut.ac.ir , gbahramali@gmail.com

Abstract: Missense mutations can affect protein structure and function, and contribute to various diseases. However, predicting the impact of a missense mutation on a specific disease may be challenging [1]. Computational methods that use both sequence and structure information can help to investigate and analyze the relationship between protein genotype and phenotype. In this study, we used an approach based on protein structure networks [2] to examine the correlation between the severity of pathogenicity of missense mutations and the average changes in the vertex degree of mutated amino acids and their neighbors. We collected data from multiple databases on the mutations, disease severity, protein stability changes, protein-protein binding affinity changes, conserved sequences, functional domains, and physicochemical properties of two proteins, namely, phenylalanine hydroxylase and alpha-galactosidase. We performed statistical analysis on the data and found that there was a significant relationship between the average changes in the vertex degrees and the pathogenicity intensity of the mutations (p-value ≤ 0.001). We also found that as the vertex degree increases, the proportion of pathogenic mutations in high-degree vertices in the protein structure network increases, indicating that pathogenic mutations were more likely to occur in residues with higher vertex degrees. Moreover, we found a significant relationship between the severity of pathogenicity and various factors, including changes in stability, changes in protein-protein binding affinity, removal of salt bridges, removal of charged amino acids, replacement of an inflexible proline in an alpha-helix, and replacement of small amino acids in buried positions. Our results suggest that network features and other factors can be useful in predicting the pathogenicity severity of missense mutations.

Keywords: Missense mutation; Protein structure network (PSN); Degree centrality; Alpha-galactosidase; Phenylalanine hydroxylase; Severety of pathogenicity

References

[1] Miosge, L.A., et al. "Comparison of predicted and actual consequences of missense mutations." Proceedings of the National Academy of Sciences 112.37 (2015): E5189-E5198.

[2] Zhou, Z., Hu, G. "Applications of graph theory in studying protein structure, dynamics, and interactions." Journal of Mathematical Chemistry (2024): doi: 10.1007/s10910-023-01511-6

ICB12-1402-1267

Discovering the Secrets of Gastric Cancer: Key Genes and Interaction Networks Revealed Through Bioinformatics

Seyyed Mohammad Yaghoubi [1 2 F A], Mahdiyeh Azimi [1 2], Sevda Ghoshouni [3], Ali Emami [4], Dariush Shanehbandi* [1 5 C]

1 Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
2 Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran
3 Tuberculosis and Lung Disease Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.
4 Department of Biochemistry and Molecular Medicine, Institute for Research in Immunology and Cancer, Université de Montréal, Montreal, QC, Canada
5 Pharmaceutical Analysis Research Center, Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding authors email addresses: m.y.ard.ir@gmail.com, dariush.shanehbandi@gmail.com

Abstract: Gastric cancer (GC), with a death rate of 32.2 (95% CI: 29.1-35.3) and 16.3 (95% CI: 13.9-18.6) in northwestern Iran (Ardabil), casts a long shadow with high mortality rates both in Iran and in the world [1]. Using bioinformatics, this study identifies differentially expressed genes (DEGs) that may help understand GC involving genes.

We analyzed gene expression profiles from normal tissue and GC tissue samples using the Gene Expression Omnibus (GEO) data set GSE184336. Results were evaluated using QC measures and DESeq2 analysis, with an adjusted p-value threshold of 0.05, to identify 278 DEGs that may play potential roles in GC development. Data was further analyzed using R software, including visualization using volcano plots generated with the ggplot2 and principal component analysis (PCA) using the "prcomp" function in R.

This bioinformatics study evaluated GC gene expression, identifying 96 DEGs with potential roles in GC development. Strikingly, 71.8% of these DEGs were categorized as signal peptides, suggesting a potential focus on cellular secretion and communication. Transcription factor analysis further highlighted the significance of the NFIC family (29%), alongside the involvement of ONECUT1, RREB1, HOXB4, ARID3A, and MEF2A families. Additionally, Gene Ontology (GO) enrichment analysis revealed enrichment of "ECM structural constituents" (12.5%%, $p < 0.001$), implicating genes crucial for maintaining the extracellular matrix. Other enriched GO terms included "Metallopeptidase activity", "Cell adhesion molecule activity", and "Chemokine activity", emphasizing diverse functional roles of the identified DEGs. Finally, network analysis uncovered 278 interacting genes, visualized using Cytoscape, providing insights into potential GC-associated molecular interactions.

Bioinformatics identified GC gene expression patterns. Further investigation of the identified DEGs enriched in specific functional categories could unlock novel therapeutic targets, benefiting those battling this disease.

Keywords: Gastric cancer; DEGs; bioinformatics; gene interaction; big data.

References

[1] Abdi, E., Latifi-Navid, S., Zahri, S., Mostafaiy, B., Yazdanbod, A., & Pourfarzi, F. (2023). Risk of Gastric Cancer is Highly Dependent on Type of First-Degree Family Member Affected by Cancer: Lessons from a High-Risk Population in Iran. Iranian journal of public health, 52(9), 2004–2013. https://doi.org/10.18502/ijph.v52i9.13582

ICB12-1402-1268

Colon Cancer Detection on Imbalanced Dataset based on using SMOTE

Seyedeh Zahra Ahmadi, Zahra Farokhi, Reza JavanmardAlitappeh*
Electrical and computer Engineering Department, University of Science and Technology of Mazandaran, Behshahr, Iran.
*Corresponding author's email address: Rezajavanmard@mazust.ac.ir

Abstract: Colon cancer is the third most common cancer and the second most malignant cancer in the world, which kills many people every year. In 2020, colon cancer claimed 576,858 lives globally, prompting a deep dive into a groundbreaking machine learning project for colon cancer detection [1, 2]. Since the diagnosis of cancer with laboratory methods is expensive, the *early detection* of this disease is the main imaginable approach to increase the probability of survival of patients. In this article, machine learning techniques have been used to diagnose the disease on colon cancer gene expression data[3]. Since working with colon cancer gene expression data has a series of challenges, such as the large number of features, the small number of samples, and the imbalance between the data classes in this paper, we used the *Synthetic Minority Over-sampling Technique* (SMOTE) [4], method to balance the data. After applying SMOTE, the number of samples increased after SMOTE in the minority class, which made the classes balanced. We compared the accuracy on the data before and after balancing on different categories. In the result we found that balancing between the classes will have a higher level of accuracy, precision and recall. Another contribution in our study is applying feature selection techniques such as PCA, RFE (recursive feature elimination), after classifying the new data using Logistic Regression, Naive Bayes, and Support Vector Machines (SVM), we achieved to 100% accuracy (with F1-score=1). This was a significant turning point that greatly impacted the success of our study. This study emphasizes the importance of balancing techniques, especially in managing unbalanced data sets. This highlights the influential role of feature selection in increasing algorithmic performance and highlights the importance of machine learning in diagnosing colon cancer with the best accuracy.

Keywords: Machine learning; Colon cancer detention; Balance and imbalance data; Feature selection

References

[1] Recio-Boiles, A., & Cagir, B. (2021). Colon cancer. Florida: StatPearls Treasure.

[2] Murugesan, M., Arieth, R. M., Balraj, S., & Nirmala, R. (2023). Colon cancer stage detection in colonoscopy images using YOLOv3 MSF deep learning architecture. Biomedical Signal Processing and Control, 80, 104283.

[3] Pirgazi, J., Alimoradi, M., Esmaeili Abharian, T., & Olyaee, M. H. (2019). An Efficient hybrid filter-wrapper metaheuristic-based gene selection method for high dimensional datasets. Scientific reports, 9(1), 18580.

[4] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Int. Res. 16, 1 (January 2002), 321–357.

ICB12-1402-1269

Cancer Prediction through Advanced Gene Expression Analysis Using Machine Learning Method

Zahra Farokhi, Seydeh Zahra Ahmadi ,Jamshid Pirgazi*
Electrical and computer Engineering Department, University of Science and Technology of Mazandaran, Behshahr, Iran.
*Corresponding author's email address: j.pirgazi@mazust.ac.ir

Abstract: Cancer is a term used to describe a group of diseases where cells grow abnormally and can spread to different parts of the body. According to the World Health Organization (WHO), cancer is the second most common cause of death globally. Detecting cancer is challenging due to the complexity of early symptoms, which can be subtle or variable, necessitating better tools for accurate diagnosis.
One of the approaches for cancer diagnosis designed on RNA sequence gene expression, however understanding gene expression in cancer is also challenging; it involves decoding complex molecular details in how cancer develops, requiring advanced techniques [1].
This study investigates gene expression as a crucial path for identifying cancer biomarkers, utilizing sophisticated molecular techniques such as DNA microarrays and RNA sequencing. We worked on the gene expression RNA seq dataset from the UCI repository [2], which includes five distinct cancer types LUAD, BRCA, KIRC, LUSC, and COAD. Dataset sample number and dimension
In order to reduce the high dimension of the input dataset we carefully applied feature selection technique i.e. Mutual information, and Linear Discriminant Analysis (LDA) to extract distinct features from RNA sequencing data. Subsequently, these discerned features serve as inputs into diverse machine learning classifiers, including decision trees, k-nearest neighbors (KNN), support vector machines (SVM), and naive Bayes, facilitating the discrimination and classification of specific cancer types.
Notably, the study highlights the superior performance of the proposed classifier, achieving an impressive accuracy rate of 99.89% compared to existing approaches [3]. This underscores its contribute to cancer studies by emphasizing the crucial role of gene expression analysis and the integration of new techniques in machine learning. Potential as a robust tool in the realm of cancer classification.
Keywords: Cancer diagnosis, Gene expression, RNA-Sequence, Feature selection, Classification.

References

[1] Alharbi, F., & Vakanski, A. (2023). Machine Learning Methods for Cancer Classification Using Gene Expression Data: A Review. *Bioengineering (Basel, Switzerland)*, *10*(2), 173. https://doi.org/10.3390/bioengineering10020173
[2] https://archive.ics.uci.edu
[3] Abdul Wahid and M Tariq Banday, "Analyzing RNA-Seq Gene Expression Data for Cancer Classification Through ML Approach" International Journal of Advanced Computer Science and Applications(IJACSA), 14(9), 2023. http://dx.doi.org/10.14569/IJACSA.2023.0140984

ICB12-1402-1270

S-Adenosylhomocysteine hydrolase: an evolutionary analysis on the structure and sequence

P. A. Tabibian[a], M. R. Shams[b], M. R. Ganjalikhany[a]*
[a] Department of Cell and Molecular Biology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
[b] Department of Animal and Plant Biology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: m.ganjalikhany@sci.ui.ac.ir

Abstract: S-Adenosylhomocysteine hydrolase (AHCY), one of the most conserved proteins involved in the methionine cycle and converting S-adenosylhomocysteine (SAH), linked to cancer, heart disease, DNA damage, and prevention of DNA repair, to homocysteine (HCY) [1-4]. High levels of HCY leads to oxidative stress, endoplasmic reticulum stress, and increased cell death [5]. Evolutionary studies have suggested the lateral transfer of the gene for the enzyme [6], with its highly conserved protein present in all domains of life and its potential as a drug target, encouraged us to delve into its evolutionary path. Employing Clustal Omega and TimeTree5, comparative trees were constructed to illustrate the evolutionary emergence of both species and the enzyme. Regarding the limitations of the algorithms in similar studies, we attempt to construct two other trees, using Maximum Likelihood and Minimum Evolution methods. Using TM-align, we compared the 3D-structure of enzyme in human with 11 other organisms. At the beginning of its evolutionary journey, the enzyme's overall length increased abruptly and then gradually decreased, 432 amino-acids in Human. Multiple sequence alignment revealed the absent of two distinguished segments in the enzyme and low conservation in the N-terminal. Comparing enzyme phylogenetic trees to the Tree of Life demonstrates that their evolution does not follow the usual three-domain tree, suggesting either a potential for a lateral gene transfer between distanced organisms or high pressure in selection duo to enzyme's significance. Comparing Tm-scores reveals that organisms from archaea domain show higher score than protozoa or bacteria. 3D-structure analysis indicated changes in the C-terminal domain and a helix structure, in the NAD binding domain of *Homo sapiens*, was missed from the surface. This study provides a wide range of analysis on AHCY as a vital enzyme and its evolutionary dynamics. These novel perspectives improve our understanding of structural and, may functional evolution of AHCY.
Keywords: S-Adenosylhomocysteine hydrolase; evolutionary analysis; phylogenetic tree

References

[1] Vizán P, Di Croce L, Aranda S. Functional and Pathological Roles of AHCY. Front Cell Dev Biol. 2021;9:654344.
[2] Bujnicki JM, Prigge ST, Caridha D, Chiang PK. Structure, evolution, and inhibitor interaction of S-adenosyl-L-homocysteine hydrolase from Plasmodium falciparum. Proteins. 2003;52(4):624-32.
[3] Menezo Y, Clement P, Clement A, Elder K. Methylation: An Ineluctable Biochemical and Physiological Process Essential to the Transmission of Life. Int J Mol Sci. 2020;21(23).
[4] Boczki P, Colombo M, Weiner J, Rapöhn I, Lacher M, Kiess W, et al. Inhibition of AHCY impedes proliferation and differentiation of mouse and human adipocyte progenitor cells. Adipocyte. 2024;13(1):2290218.
[5] Wang J, Ding K, Wang Y, Yan T, Xu Y, Deng Z, et al. Wumei Pill Ameliorates AOM/DSS-Induced Colitis-Associated Colon Cancer through Inhibition of Inflammation and Oxidative Stress by Regulating S-Adenosylhomocysteine Hydrolase-(AHCY-) Mediated Hedgehog Signaling in Mice. Oxid Med Cell Longev. 2022;2022:4061713.
[6] Stepkowski T, Brzeziński K, Legocki AB, Jaskólski M, Béna G. Bayesian phylogenetic analysis reveals two-domain topology of S-adenosylhomocysteine hydrolase protein sequences. Mol Phylogenet Evol. 2005;34(1):15-28.

ICB12-1402-1271

A method to visualize pseudo-potential landscape of high-dimensional Boolean networks

Mohammad Taheri-Ledari [a], Sayed-Amir Marashi [b], Kaveh Kavousi [a, *]
[a] Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
[b] Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran
*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Boolean networks (BNs) are a successful mathematical framework to represent the dynamical behavior of molecular networks. For the BNs, a pseudo-potential function can be defined to assign a numerical value to each state such that its value is minimum for the stable states or attractors of the system. Visualization of the pseudo-potential landscape of the BNs helps us gain insight into the state space of the BNs. However, plotting the pseudo-potential landscape for the BNs in the case of having three or more variables is a challenging task [1]. Since the state space of the BNs grows exponentially with the number of variables, calculating and storing the pseudo-potential values for all states is not feasible. Here, we present a method to visualize the pseudo-potential landscape of a possibly high-dimensional BN (having three or more variables) by choosing a representative subset of the states and placing them over a 1D or 2D plane, such that the consecutive states are placed as close together as possible [2]. Consequently, the second axis in the case of a 1D plane (or alternatively, the third axis in the case of a 2D plane) shows the pseudo-potential values associated with different states. We present the results for some BNs and then discuss the performance of the proposed visualization approach.
Keywords: Boolean network; Pseudo-potential landscape; Visualization

References

[1] J. Guo, F. Lin, X. Zhang, V. Tanavde, J. Zheng, "NetLand: quantitative modeling and visualization of Waddington's epigenetic landscape using probabilistic potential", Bioinformatics, 33, 10 (2017) 1583-1585.

[2] M. Choi, J. Shi, S. H. Jung, X. Chen, K. Cho, "Attractor Landscape Analysis Reveals Feedback Loops in the p53 Network That Control the Cellular Response to DNA Damage", Sci. Signal., 5, 251 (2012).

ICB12-1402-1275

Pathogenicity Prediction Assays of Human Missense Variant (rs1484954450) in CTLA-4 Gene Based on Protein Sequence, Structure and Function

Hadis Shahrokhi Mohammad Mehdi Heidari
Department of Biology, Yazd University, Yazd, Iran
Affiliation Department of Biology, Yazd University, Yazd, Iran
shahrokhi.1997@gmail.com
*Corresponding author's email address: heidarimm@yazd.ac.ir

Abstract: Globally, cancer is the main cause of death. During the occurrence of cancer, in the microenvironment surrounding tumor cells, the process of identifying and connecting T cells by histocompatibility complex(MHC) to antigens on the surface of antigen-presenting cells occurs, in order to provide activation and cytotoxic activity of T cells. Binding of CD80 and CD82 molecules on antigen presenting cells with CD28 receptor on T cell surface leads to T cell activation. Appropriate levels of this connection cause the division of T cells and their differentiation through the secretion of cytokines and increase the expression of genes related to cell survival, and the process of inhibiting tumor cells begins. Cytotoxic T lymphocyte antigen 4)CTLA4 (is a negative regulator belonging to the immunoglobulin subfamily. This gene is located at chromosomal position 2q 32.2 and has 4 exons. Upon receiving the signal from the side of the tumor cell, it is expressed in active T cells and regulatory T cells and encodes a protein that transmits the inhibitory signal to these cells and helps the tumor spread. In this research we investigated the function and structure of this type of protein by studying and checking bioinformatics servers, I-mutant-2, Expasy.
The A11G variant (rs1484954450) pathogen polymorphism was selected from NCBI. By studying this variant in I- mutants, a decrease in protein stability was predicted. Hydrophobic index (score -1.75) increased and molecular weight (score 138) decreased in Expasy tool.
It was predicted that the missense variant replaces Alanine with Glycine can cause structural disorder and as a result protein dysfunction.
Keywords: Cancer; T- cell; CTLA-4

ICB12-1402-1276

Reveal Genetic Links Among COVID-19, A Deep Neural Network Method to Alzheimer's Disease, and Multiple Sclerosis Based on Single-Cell RNA-seq

Asiyeh Mirzaei[1], Mohammad Shirinpoor[1], Hediyeh Talebi[1], Shokoofeh Ghiam[2*] and Changiz Eslahchi[1,2*]

[1]Department of Computer and Data Sciences, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

[2]School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

*Corresponding author's email address: sh.ghiam@ut.ac.ir, ch-eslahchi@sbu.ac.ir

Abstract: The ongoing COVID-19 pandemic, driven by changing SARS-CoV-2 variants, has highlighted new concerns: the virus can affect the brain beyond the respiratory system. SARS-CoV-2's ability to invade the nervous system through ACE2 receptors raises worries about its impact on neurological diseases. We need to investigate if COVID-19 worsens cognitive decline and existing neurological conditions. Understanding how SARS-CoV-2 enters the brain, causes inflammation, and disrupts the blood-brain barrier is crucial for predicting and managing long-term neurological effects.

Early and accurate diagnosis of neurodegenerative diseases (NDs) is crucial due to limited treatment options and high costs associated with late or inaccurate diagnosis. Challenges include difficulties in making definitive diagnoses in early stages and predicting disease progression. Identifying blood biomarkers offers a promising avenue for easier, more cost-effective, and faster ND diagnosis, potentially leading to earlier intervention, better disease management, and improved patient outcomes. Real-time PCR is a standard method for gene expression measurement, limited in gene coverage. RNASeq analysis is used to identify gene expression differences but faces challenges in specific cell line analysis and requires large sample sizes. Single-cell RNA sequencing (scRNA-seq) offers unparalleled insights into cellular heterogeneity and gene expression at the individual cell level. It enables the discovery of rare cell populations, identification of biomarkers, and understanding disease mechanisms with unprecedented precision.

This study aims to identify blood biomarkers for early diagnosis of neurodegenerative diseases in COVID-19 patients. To achieve this goal, scRNA-seq data related to COVID-19, Alzheimer's Disease (AD), and Multiple Sclerosis (MS) were extracted from the GEO database, focusing on peripheral blood mononuclear cells (PBMC). After analyzing scRNA-seq data and selecting genes with adjusted p-value < 0.05 and ($logfc < -0.01 \mid logfc > 0.01$), we employed a deep neural network, particularly an adversarial autoencoder with a classifier, to accurately classify normal and diseased samples based on their cell types. Models were tailored to individual diseases and four blood tissue cell lines (monocytes, NK cells, B cells, and T cells). To evaluate the importance of genes in the classification process, we employed multiplication of the weight matrix. Subsequently, we compared the results extracted from the models for each disease and each cell line. These findings indicate that genes associated with COVID-19 may contribute significantly to the development of neurological diseases such as AD and MS by affecting inflammatory and apoptotic pathways. This discovery holds promise for advancing drug development and early detection of post-coronavirus complications in individuals.

Keywords: Blood Biomarkers, Neurological Diseases, Alzheimer's Disease, Multiple Sclerosis, COVID-19, Single-cell RNA-seq, Deep Neural Network

ICB12-1402-1277

Breast cancer early detection biomarkers identified from ductal carcinoma in-situ patients: a bioinformatics approach

Majid Fathi Kheirabad[a], Behzad Jafari[b], Karim Hasanpur[c], Yosef Masoudi-Sobhanzadeh[d], Mohammad Mostafa Pourseif[d]*

a Faculty of Pharmacy, Urmia University of Medical Sciences, Urmia, Iran
b Department of Medicinal Chemistry, Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran
c Department of Animal Sience, Faculty of Agriculture, Tabriz University of Medical Sciences, Tabriz, Iran
d Research Center for Pharmaceutical Nanotechnology, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author's email address: mmpourseif@gmail.com

Abstract: The Breast Cancer is the most prevalent cancer in the world [1]. According to WHO (World Health Organization), there were more than 2 million women diagnosed with breast cancer and more than 600,000 recorded deaths all around the world in 2020. Common diagnosis approaches for breast cancer like mammography have some limitations such as low sensitivity and specificity, high cost and time-consuming procedures so new and better diagnostic methods are presented [2, 3]. The Events Responsible for Transmission of normal breast tissue into cancer phase are poorly known. In this study, we aim to discover the pathways initiating breast cancer that can modify diagnosis and treatment methods. Four datasets (PRJNA484546, PRJNA602512, PRJNA855324 and PRJNA950392) were selected from the NCBI SRA database. These datasets included ductal carcinoma in-situ and normal breast samples. In order to ensure of the quality of these samples, a test was performed using FastQC. Using Trimmomatic software, appropriate changes were made on the samples to increase quality. After performing Various analysis, we used UMAP Plot (Uniform manifold approximation and projection) to classify the genes with different alterations of gene expression in order to identify the pathways related to breast cancer. Bioinformatic analysis on 34 samples selected from 4 datasets, showed that the presence of conserved mutations in genes including PIK3CA, TP53 and CDH1 in matched samples of both DCIS and normal tissue suggest a fitness benefit of these mutations. Studies of matched synchronous cases have also shown DCIS-specific mutations in selected cases. Discovering the pathways initiating breast cancer is a key goal that can modify diagnosis and treatment methods.

Keywords: breast cancer; diagnosis; RNA-sequencing

References

[1] Gil Del Alcazar, C.R., et al., Immune escape in breast cancer during in situ to invasive carcinoma transition. Cancer Discov, 2017. 7(10): p. 1098-1115.

[2] Kaur, H., et al., RNA-Seq of human breast ductal carcinoma in situ models reveals aldehyde dehydrogenase isoform 5A1 as a novel potential target. PLoS One, 2012. 7(12): p. e50249.

[3] Trinh, A., et al., Genomic alterations during the in situ to invasive ductal breast carcinoma transition shaped by the immune system. Mol Cancer Res, 2021. 19(4): p. 623-635.

ICB12-1402-1278

Efficient Hybrid algorithm for Gene Selection Based on Chaos Game Optimization

Pourya Salati, Parastwo tavakoly, MohammadHossein Olyaee*
Electrical and Computer Engineering Department, University of Gonabad, Gonabad, Iran
*Corresponding author's email address: mh.olyaee@gonabad.ac.ir

Abstract: Gene selection stands as a crucial preprocessing step in biomedical data mining and holds significance in various biological applications, particularly in biomarker identification. Despite the introduction of numerous gene selection algorithms, challenges persist, including prolonged convergence time, parameter tuning, and suboptimal performance. This paper introduces an efficient approach employing the Chaos Game Optimization (CGO) algorithm for the selection of relevant genes from large-scale gene datasets. Our algorithm operates in two distinct phases: during the Relief-based filtering phase, weights are assigned to genes, and in the subsequent CGO-based wrapping phase, optimal genes are identified. To assess the efficacy of our proposed method, we conduct evaluations using standard gene expression datasets. The experimental results highlight that, in comparison to analogous methods, our approach produces a more streamlined set of features without compromising accuracy.

Keywords: Gene selection; Chaos Game Optimization; High dimensionality; Bioinformatics

References

[1] Pirgazi, Jamshid, et al. "An Efficient hybrid filter-wrapper metaheuristic-based gene selection method for high dimensional datasets." *Scientific reports* 9.1 (2019): 18580.

[2] Osama, Sarah, Hassan Shaban, and Abdelmgeid A. Ali. "Gene reduction and machine learning algorithms for cancer classification based on microarray gene expression data: A comprehensive review." *Expert Systems with Applications* 213 (2023): 118946.

[3] Liu, Yun, et al. "Simulated annealing-based dynamic step shuffled frog leaping algorithm: Optimal performance design and feature selection." *Neurocomputing* 503 (2022): 325-362.

ICB12-1402-1279

Enhancing Drug Design for VEGFR2: Integrating Quantum Mechanics-Driven 3D-QSAR with Deep Learning to Predict Drug Efficacy

Amir Mahdi Zardoshti [a,c], Maral Zarvani [b], Mohammad Taheri-Ledari [c], Kaveh Kavousi [c,*]
a Department of Chemistry, Faculty of Science, University of Tehran, Tehran, Iran
b Department of Computer Engineering, Faculty of Engineering, Alzahra University, Tehran, Iran
c Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Cancerous tumors require nutrients and oxygen to grow and spread. They achieve this by promoting the formation of new blood vessels through angiogenesis. Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) is central to this process. It is a receptor on the surface of cells on which the Vascular Endothelial Growth Factor (VEGF) is banded. When VEGF binds to VEGFR2, it triggers a signaling cascade inside the endothelial cells. These signals stimulate the endothelial cells to multiply and form new blood vessels. The inhibition of VEGFR2, a key receptor in this pathway, has emerged as a critical approach to hinder tumor growth and spread by targeting the angiogenesis process essential for tumor sustenance and expansion. Traditional 3D-QSAR techniques like Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Index Analysis (CoMSIA) rely on classical mechanics. Despite their success, these approaches have certain limitations, particularly when it comes to computing the molecular electrostatic field based on atom-centered partial charges for individual ligand molecules. To address these limitations, we present a 3D-QSAR method for designing drugs targeting VEGFR2 based on quantum mechanics. This approach offers a more precise description of molecular interactions, so that it addresses some disadvantages of classical methods. We generate electrostatic potential surfaces to gain profound insights into molecular interactions with the VEGFR2 protein at the molecular level [1]. This information is pivotal to the precise design of effective drugs. We harnessed Density Functional Theory (DFT) calculations to obtain a Point Cloud (PC) representation as well as the Electrostatic Surface Potential (ESP) of a set of compounds. It led to a dataset of PCs and ESPs related to 1306 drugs. Deep learning (DL) enhances QSAR models by directly using raw PCs and alleviating the need for traditional feature extraction and dimension reduction. We have developed a unique deep learning model that uses the principles of PointNet [2] and Transformer [3] to predict the effectiveness of compounds on VEGFR2. Our model demonstrated robustness and effectiveness, achieving a Precision of 0.839, Recall of 0.824, F1 Score of 0.823, Accuracy of 0.826, Matthews Correlation Coefficient (MCC) of 0.663, and an Area Under the Curve (AUC) of 0.873.
Keywords: 3D-QSAR; Deep learning; Point cloud; VEGFR2; Drug design

References

[1] M.R Bauer, M.D. Mackey, "Electrostatic complementarity as a fast and effective tool to optimize binding and selectivity of protein–ligand complexes." Journal of medicinal chemistry 62.6 (2019).

[2] Ch. R. Qi, H. Su, K. Mo, "Pointnet: Deep learning on point sets for 3d classification and segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 652-660. 2017.

[3] A. Vaswani, N. Shazeer, "Attention is all you need." Advances in neural information processing systems 30 (2017).

ICB12-1402-1280

A Siamese neural network for immunotherapy response prediction

Reyhaneh Mortezaee[a], Fatemeh Zare*

[a, *]Department of Computational Biology Research Center (CBRC), Department of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran

*Corresponding author's email address: stahmasebian@gmail.com

Abstract: Immune checkpoint inhibitors (ICIs) have emerged as groundbreaking treatments for various types of cancer. However, only a portion of patients with solid tumors experience positive outcomes from ICI therapies, which sometimes come with significant side effects and costs. Therefore, accurate prediction of ICI response is crucial. The scarcity of patient samples treated with ICIs, which have both clinical outcomes and transcriptomic data, has posed a critical challenge in developing predictors for patient response. To address this issue, we first identify pathway-based biomarkers and estimate the pathway expression levels, effectively reducing the dimensionality of the data. We utilized Reactome pathways and conducted single-sample gene set enrichment analysis (ssGSEA) to compute the expression levels of pathways. The normalized enrichment score (NES) employed as each sample's pathway expression levels. Second, we propose a novel model that applies a homogeneous Siamese neural network, which takes the pathway expression levels of two patients as input to predict the similarity of their ICI response. As a result, we generate a new dataset and increase the sample size. This dataset contains pairs of patients and the binary similarity of their ICI response. We utilize a dataset comprising 91 ICI-treated patient samples. Remarkably, our model achieves an accuracy of ACC = 0.78, outperforming the last existing model that uses logistic regression with an accuracy of ACC = 0.73.
Keywords: Siamese neural network; immune checkpoint inhibitors; pathway expression; machine learning

ICB12-1402-1281

Transcriptomics-Based Systems Biology Analysis Suggests New Potential Biomarkers for Colorectal Cancer

Zahra Taslimi[a], Nayereh abdali[b], Atena Vaghf[c], Shahram Tahmasebian*

[1, 2 ,3] Department of medical biotechnology, Shahrekord University of medical sciences, Shahrekord, Iran
*Corresponding author's email address: stahmasebian@gmail.com

Abstract: Colorectal cancer (CRC) is the second most common cancer worldwide which is associated with significant morbidity and mortality. while, colonoscopy is the most accurate screening test for detecting CRC, it is somewhat invasive, inconvenient, and expensive [1]. The incidence rate of sporadic early-onset colon cancer (EOCC) has been surprisingly increasing worldwide, and its molecular mechanisms occurrence and development remain unclear [2]. The global burden of CRC and the annual increase in young patients highlight an essential understanding of the underlying mechanisms to identify novel diagnostic and prognostic biomarkers or therapeutic targets [3]. System biology is a biology-based computing science that focuses on the complex interactions in biological systems. In this regard, transcriptomic data analysis is a powerful approach to linking genotype to phenotype of a cell. Furthermore, it provides valuable insights to identify gene expression patterns in cancer cells [4]. To understand the expression relationships between genes to find a therapeutic target, the bulk RNA-seq dataset (GSE240623) was obtained from NCBI Gene Expression Omnibus (GEO). Then, 13 EOCC samples and 13 healthy control samples were analyzed by GEO2R to identify differential expression genes (DEGs). Next, the most significant genes were selected for further analysis. In the next step, the STRING database was used to construct significant protein-protein interactions by DEGs gene list. ESCO2 was suggested as a new target for treating CRC. Further expression analysis was validated by GEPIA and analysis of possible mRNA-miRNA interactions was performed by miRWalk. The target gene analyses showed that hsa-miR-145-5p specifically targets the ESCO2 gene. In conclusion, these results showed that ESCO2 and hsa-miR-145-5p could be suggested as potential biomarkers associated with colorectal cancer.

Keywords: Transcriptomic; Colorectal cancer, DEG

References

[1]. Sawicki, T., et al., A Review of Colorectal Cancer in Terms of Epidemiology, Risk Factors, Development, Symptoms and Diagnosis. Cancers (Basel), 2021. **13**(9).

[2]. Furuhashi, S., et al., Spatial profiling of cancer-associated fibroblasts of sporadic early onset colon cancer microenvironment. NPJ Precis Oncol, 2023. **7**(1): p. 118.

[3]. Ogunwobi, O.O., F. Mahmood, and A. Akingboye, Biomarkers in Colorectal Cancer: Current Research and Future Prospects. Int J Mol Sci, 2020. **21**(15).

[4]. Zhang, Y., et al., Methods and biomarkers for early detection, prediction, and diagnosis of colorectal cancer. Biomedicine & Pharmacotherapy, 2023. **163**: p. 114786.

ICB12-1402-1282

Enrichment analysis of potential hub genes in Acute Myeloid Leukemia based on TCGA datasets

Narjes Seddighi[1,2*], Ali Salimi Asl[3]
Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
Hematology and Oncology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
Student Research Committee, Faculty of Medicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
*Corresponding author's email address: narjes.seddighi80@gmail.com

Abstract: Acute myeloid leukemia (AML) is the most common acute leukemia in adults. AML is a malignancy of the stem cell precursors of the myeloid lineage (red blood cells, platelets, and white blood cells other than B and T cells). Like other malignancies, it is due to genetic variations that lead to neoplastic changes and clonal proliferation. Despite great advances in the diagnosis and treatment of cancer, the treatment has big challenges. Nowadays, with the advances in the bioinformatics analysis of RNAseq, it can identify potential biomarkers for diagnosis, targeting treatment, and evaluation of tumor metastasis and relapse. In the current study, we conducted the gene ontology and Kegg pathway based on TCGA datasets.
At first, we used GEPIA2 to examine the TCGA AML dataset to identify all DEGs linked with AML among high throughput RNA-Seq data. After analyzing the survival data of AML, a Protein-protein interaction (PPI) network of significant genes associated with AML was constructed in Cytoscape software, and the hub genes were identified. for enrichment analysis, we used the Enrichr website (https://maayanlab.cloud/Enrichr/) and extracted the Kegg pathway and gene ontology based on the potential hub Gene
Twelve Hub genes (MRPS10,MRPL16,NDUFS8,MRPL4,ENO1,MRPL12,MRPL2, EPRS1,SDHA,SOD1,HSPA5,SDHB) were extracted from PPI based on their degree. the result showed in the 157 biological processes, the Mitochondrial Translation (adj-P-value=1.32E-07), Translation (adj-P-value=1.32E-07) and Mitochondrial Gene Expression (adj-P-value=1.32E-07) was significantly meaningful. In 26 cellular components, we identified just 3 components with meaningful adj p-values which included Mitochondrial Inner Membrane (adj-P-value=1.11E-08), Organelle Inner Membrane (adj-P-value=1.11E-08) and Mitochondrial Membrane (adj-P-value=6.14E-08). In 34 molecular functions, RNA Binding (adj-P-value=0.002645376) and GTPase Binding (adj-P-value=0.004874789) have a more meaningful. At the last, in the 25 Kegg pathway, we identified Ribosome (adj-P-value=5.47E-07) with meaningful P-value. This study identifies gene ontology and keg pathways of hub genes which involved in the occurrence and progression of Acute Myeloid Leukemia. This information may hold promise as potential biomarkers and therapeutic targets.

Keywords: Acute Myeloid Leukemia, Gene Ontology, Enrichment Analysis, AML, Kegg Pathway

ICB12-1402-1286

Rs35107962 modifies Microtubule Cytoskeleton Organization via Frameshift in MAP7 Gene and Protein in Skin Cancer patients: in-silico investigation

Mehrafarin Nekoonam[1], Mansoureh Azadeh[1]*
Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
*Corresponding author's email address: mehrafarin.nekoonam@gmail.com

Abstract: The skin cancer known as cutaneous melanoma (SKCM) arises from melanocytes, which are cells that produce pigment. This type of cancer is particularly aggressive and often fatal [1]. Epithelial-mesenchymal transition (EMT) is a process commonly observed in malignant skin tumors [4]. MAP7, also called Ensconsin, is a protein that stabilizes microtubules and may play a crucial role in organizing the cytoskeleton and microtubules. This study sought to explore a new regulatory biomarker in skin cancer patients using an integrated approach involving bioinformatics and systems biology analysis.

Microarray profiling analysis and gene expression data assessment were conducted using the GSE73652 database with GEO2R. GEPIA2 was used to examine the correlation between the gene and skin cutaneous melanoma. Single Nucleotide Polymorphisms (SNPs) of the gene were extracted from miRNASNP and biological pathway analysis was performed using Enrichr. The protein structure changes caused by the single nucleotide variations were studied using the Hope database and the Alphafold artificial intelligence program [3]. sorting intolerant from tolerant SNPs was predicted using the SIFT database .

The microarray analysis revealed a significant downregulation of the MAP7 gene in skin cutaneous melanoma patients. Further investigation indicated that the product of the MAP7 gene is a microtubule-associated protein predominantly expressed in epithelial cells. The process of cell invasion and migration relies on the dynamic changes taking place in the cytoskeletal components such as actin and tubulin [2]. Changes in cellular architecture by internal clues will affect the cell functions leading to the formation of different protrusions that helps cell migration eventually leading to metastasis [2]. Our analysis also revealed that a single nucleotide variant(C>G) in the protein-coding sequence (CDS)region caused the mutation of an Arginine into a Proline at position 512. This mutation can disrupt an α-helix which affects the structure and function of the protein [3].

Keywords: EMT; MAP7; Skin Cancer; Microtubule Organization

References

[1] Naik PP. Cutaneous Malignant Melanoma: A Review of Early Diagnosis and Management. World J Oncol. 2021

[2] Aseervatham, J. Cytoskeletal Remodeling in Cancer. Biology 2020

[3] Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics. 2010.

[4] -Hodorogea A, Calinescu A, Antohe M, Balaban M, Nedelcu RI, Turcu G, Ion DA, Badarau IA, Popescu CM, Popescu R, Popp C, Cioplea M, Nichita L, Hulea I, Brinzea A. Epithelial-Mesenchymal Transition in Skin Cancers: A Review. Anal Cell Pathol (Amst). 2019

[5] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013

ICB12-1402-1287

AKR1B10 overexpression correlating with LINC01359 and hsa-miR-4482-3p predicts worse overall survival in hepatocellular carcinoma

Golnaz Enayat[1], Mahsa Ghandi[1], Mansoureh Azadeh[1,*]
1 Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
*Corresponding author's email address: mazadeh@phd.iaurasht.ac.ir

Abstract: Hepatocellular carcinoma (HCC) is the third cause of cancer-related death worldwide while Asia accounts for highest mortality rate among the world [1]. Also, the five-year survival rate for patients with advanced HCC remains bleak; a reality that is largely attributable to an absence of adequate diagnostic and prognostic biomarkers during HCC treatment.[2] In present study, a search of Gene Expression Omnibus (GEO) database was performed to screen and download the expression profiles of GSE14520 and GPL571 to obtain differentially expressed genes (DEGs) which results in a mRNA, a member from aldo-keto reductase superfamily, named AKR1B10 that exhibits overexpression in HCC compared with non-tumor tissue(all $P < 0.05$). Furthermore, survival analysis of HCC in this database was used for identifying AKR1B10 as a contributing factor for predicting the overall survival (OS) of HCC patients. Moreover, Protein-Protein Interaction analysis (PPI) by STRING database and KEGG pathway enrichment analysis for AKR1B10 were performed which illustrate salient role of AKR1B10 in glucuronate interconversions. Recently, increasing evidences indicate that evolutionarily conserved ncRNAs, particularly microRNA (miRNA) and long noncoding RNA (lncRNA), play a significant role in diverse pathological processes [3]. In current study, analysis has revealed AKR1B10 interacts with a miRNA named hsa-miR-4482-3p as well as a novel HCC-related lncRNA named LINC01359 which has a significant upregulation in patients with liver cancer. In conclusion, evaluation of diagnostic value of AKR1B10 in tumor tissue exhibits a prospective clinical outcome in the hope of providing useful insights into hepatocarcinogenesis and aggressiveness.

Keywords: Hepatocellular carcinoma, AKR1B10, miR-145, LINC01359, biomarker

References

[1] Z. Huang, J. Zhou, Y. Peng, W. He, and C. Huang, "The role of long noncoding RNAs in hepatocellular carcinoma," pp. 1–18, 2020

[2] J. K. Distefano and B. Davis, "Human Hepatocellular Carcinoma," pp. 1–13, 2019, doi: 10.3390/cancers11040486

[3] H. Yan and P. Bu, "Non-coding RNA in cancer," *Essays Biochem.*, vol. 65, no. 4, pp. 625–639, 2021, doi: 10.1042/EBC20200032.

ICB12-1402-1288

Assessment of the RBP4 gene and LncRNA LINCO1215 in individuals with breast cancer

Saeedeh Ghazaey Zidanloo Sanaz Zare[1], Golnaz Enayat [2], Mansoureh Azadeh[3]*
1, 2, 3 Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
*Corresponding author's email address: Email: mazadeh@phd.iaurasht.ac.ir

Abstract: Breast cancer (BC) is the most prevalent malignancy globally and the leading cause of cancer-related fatalities among women [1]. While BC is a complicated illness, its onset and progression can be influenced by both hereditary and environmental factors [2]. The purpose of the present study is to identify the critical genes for BC diagnosis and prediction. The main objective of this article's several analyses is to discover biomarkers for this specific type of cancer. GSE36295 in the BC field was subjected to microarray analysis utilizing data collected from the GEO database. The RBP4 gene has been discovered to display a significant reduction in expression after being compared to the GEPIA2 and ENCORI databases. The KEGG database was used to examine the signaling pathways in which the RBP4, which is a gene was expressed. The STRING database was used for assessing the gene's protein-protein interaction. The miRWalk database analyzed the interaction between the target gene and associated miRNAs. Following analyzing the selected gene in conjunction with its correlating miRNAs, hsa-let-7f-5p was determined as the miRNA affecting the 3UTR region. The lncRNA linked to the RBP4 gene was identified in the lncRResearch database as LINCO1215. Lastly, the study was done on the interaction between the chosen lncRNA and the related miRNAs. The RBP4 gene showed a significant drop in expression in breast cancer and could serve as a biomarker for this cancer, according to the study's microarray analysis.

Keywords: Breast cancer, Gene expression, RBP4 gene, GSE36295

References

[1] Wilkinson, L., & Gathani, T. (2022). Understanding breast cancer as a global health concern. The British journal of radiology, 95(1130), 20211033. https://doi.org/10.1259/bjr.20211033

[2] Wilkinson, L., & Gathani, T. (2022). Understanding breast cancer as a global health concern. The British journal of radiology, 95(1130), 20211033. https://doi.org/10.1259/bjr.20211033

[3] F. Surname1, F. Surname2, Title of the book chapter in: "Title of the book", University of Science and Technology of Mazandaran, vol. 12, 2023, pp. 123-145.

[4] Edgar R, Domrachev M, Lash AE.Gene Expression Omnibus: NCBI gene expression and hybridization array data repository Nucleic Acids Res. 2002 Jan 1;30(1):207-10.

ICB12-1402-1289

G-quadruplex structures in ring finger protein 1 (RING1) gene sequence

Saeedeh Ghazaey Zidanloo
Department of Cell and Molecular Biology, Kosar University of Bojnord, Bojnord, Iran.
*Corresponding author's email address: Ghazaeys@kub.ac.ir

Abstract: Recent studies have shown that RING1 is overexpressed in various types of human cancers, including lymphoma, non-small cell lung cancer, and prostate and liver cancers [1]. The Ring finger protein1 (RING1), an essential component of the poly-comb family of proteins, plays vital roles in the tumorigenesis of various cancer types [2]. However, further research is required to determine RING1 expression and prognostic value in some cancers. Therefore, regulating its gene expression is important for disease treatment. One of the ways to control the expression of this gene is to induce G-quadruplex structures, which act as a barrier to the gene expression apparatus [3]. Interestingly, there is a database that lists quadruplex-forming G-rich sequences (QGRS) (http://tubic.tju.edu.cn/greglist/) [4]. The sequence of the RING1 gene was obtained from GenBank at NCBI. We used a web-based server, QGRS Mapper, which predicts quadruplex-forming G-rich sequences (QGRS) in nucleic acid sequences (http://bioinformatics.ramapo.edu/QGRS/) [5]. RING1 was a large gene and the GC content of its sequence was high, based on the data of the online software, we found 29 sequences prone to G-quadruplex formation, among which 14 regions had a score above 20, which indicates that the probability of the structure formation in these regions is very high. According to the recent evidence for the in vivo location and role of DNA G-quadruplexes in several cellular pathways including DNA replication and gene expression, effective treatment strategies can be defined for drug design and targeted treatment of the diseases. Inhibiting this gene via G-quadruplex structures provides opportunities for shutting down pathways associated with tumor development and metastasis.

Keywords: RING1; Regulation of gene expression; G-quadruplex structure; In Silico analysis

References

[1] Gao S, Wang S-Y, Zhang X-D, Wu H, Pang D. Low expression of the polycomb protein RING1 predicts poor prognosis in human breast cancer. Front Oncol 2021;10:618768.

[2] George M, Masamba P, Iwalokun BA, Kappo AP. Zooming into the structure-function of RING finger proteins for anti-cancer therapeutic applications. Am J Cancer Res 2023;13:2773.

[3] Awadasseid A, Ma X, Wu Y, Zhang W. G-quadruplex stabilization via small-molecules as a potential anti-cancer strategy. Biomed Pharmacother 2021;139:111550.

[4] Zhang R, Lin Y, Zhang C-T. Greglist: a database listing potential G-quadruplex regulated genes. Nucleic Acids Res 2008;36:D372–6.

[5] Kikin O, D'Antonio L, Bagga PS. QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. Nucleic Acids Res 2006;34:W676–82. https://doi.org/10.1093/nar/gkl253.

ICB12-1402-1290

HSA-MIR-539-5P promots gastric adenocarcinoma by disturbing interactions of STK32A-AS1 as a lncRNA

Mahsa Ghandi[1], Sanaz Zare[2], Mansoureh Azadeh[3*]
1, 2, 3 Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
*Corresponding author's email address: mazadeh@phd.iaurasht.ac.ir

Abstract: Gastric cancer (GC) represents the fifth most common tumor and the fourth leading cause of cancer-related deaths worldwide[1]. Several genetic and epigenetic factors, including microRNAs (miRNAs) and lncRNAs, affect its initiation and progression. MiRNAs are short chains of nucleic acids that can regulate several cellular processes by controlling their gene expression, Long noncoding RNA (LncRNA) is a large class of RNA molecules with size larger than 200 nucleotides. They exhibit cellular functions although having no protein-coding capability [2]. In recent decades, miRNAa and lncRNAs have been studied and considered as impactful biomarkers in cancer.In the present study, microarray analysis was performed on GSE81948in the field of GC from the GEO database. The SERPINE1 gene was selected as a gene with significant over expression after checking with ENCORI and GEPIA2 databases. The signaling pathways in which the SERPINE1 gene was active and was checked by KEGG database. The protein- protein interaction of the gene was analyzed by using STRING database. The interaction between the desired gene and its related miRNAs was checked by miRWalk database. After analyzing the desired gene with its related miRNAa, hsa-mir-539-5p was selected as the miRNA affecting the 3UTR region. STK32A-AS1 lncRNA was selected in the lncRResearch database as the lncRNA associated with the SERPINE1 gene. This up-regulted gene (SERPINE1) has correlation with a down-regulated gene which is DNER. In conclution, the interaction between the selected lncRNA and its related miRNAa was studied. After microarray analysis on SERPINE1 gene, the result of this study showed that the gene which was mentioned has a significant increase in expression in gastric cancer.

Keywords: Gastric cancer, gene expression, SERPINE1 gene, GSE81948

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.

[2] Panda, S., Shiras, A., & Bapat, S. A. (2018). Long noncoding RNAs: insight into their roles in normal and cancer stem cells. In Cancer and noncoding RNAs (pp. 409-427). Academic Press

ICB12-1402-1291

Bioinformatic Prediction of CMS Typing in Colon Polyps: Revolutionizing for Early Detection Intervention

Haniye Rahimi Kolour[a], Negar Mehrzadeh[a], Moein Piroozkhah[a], Leili Rejali[a], Zahra Salehi[c,d], Ehsan Nazemalhosseini-Mojarad[b]*

a Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

b Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

c Hematology, Oncology and Stem Cell Transplantation Research Center, Tehran University of Medical Sciences, Tehran, Iran.

d Research Institute for Oncology, Hematology and Cell Therapy, Tehran University of Medical Sciences, Tehran, Iran

*Corresponding author's email address: e.nazemalhosseini@sbmu.ac.ir

Abstract: Numerous studies have demonstrated the influence of polyp pathways on the likelihood of benign or malignant occurrences, as well as the clinical manifestations in patients [1]. In recent years, the consensus molecular subclassification (CMS) has proven effective in predicting the prognosis of colorectal cancer and the risk of recurrence in affected individuals [2]. This study aims to investigate the gene expression signatures (GSEs) associated with CMS and polyps, particularly high-risk polyps, and to identify shared genes to identify a biomarker that can predict the likelihood of polyps becoming malignant and provide a more favorable prognosis for patients. Initially, GSEs related to CMS and sessile serrated adenoma polyps (SSA) were selected (GSE103479, GSE79462, GSE198692, and GSE45270), and microarray analysis was conducted to identify genes with differential expression across different CMS groups and SSA polyps. Subsequently, the differentially expressed genes (DEGs) were compared to identify commonalities, and a Venn diagram was constructed to visualize the shared genes. The expression levels of the identified genes were then validated using the UALCAN database, and the Gepia2 and EnrichR databases were utilized for gene enrichment analysis. Microarray analysis of these GSEs revealed 16 common genes and enrichment analysis indicated their involvement in various cancers, including colorectal, prostate, endometrial, and breast cancer. This study demonstrated the potential use of EGFR as biomarkers for CMS1, and SLC7A4, RARRES1, ACAP1, and IFI30 genes as biomarkers for CMS3. Additionally, PI15, AKT3, PTEN, KLF6, TNS1, and ACE genes were identified as potential biomarkers for CMS4.

Keywords: Sessile serrated adenoma polyps (SSA); Consensus molecular subclassification (CMS)

References

[1] Utsumi, T., et al., *Sessile serrated lesions with dysplasia: is it possible to nip them in the bud?* Journal of Gastroenterology, 2023. **58**(8): p. 705-717.

[2] Rejali, L., et al., *Principles of Molecular Utility for CMS Classification in Colorectal Cancer Management.* Cancers, 2023. **15**(10): p. 2746.

ICB12-1402-1294

Investigating DCAF10 Gene Expression and its Correlation with Azoospermia

S. Seifodini[a], G. Sadeghian Boroujeni[a], S. M. Javadirad[b]*
a Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Islamic Republic of Iran.
b Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Islamic Republic of Iran.
*Corresponding author's email address: Javadirad@yahoo.com

Abstract: Azoospermia, the complete absence of sperm in the semen, is defined as the most severe form of male infertility and accounts for 10-15% of all male infertility cases. This disorder can manifest in two forms: obstructive (OA) and non-obstructive (NOA)[1]. Discovering the genetic factors associated with infertility is a major focus in the field of andrology. However, a considerable proportion of infertility cases, particularly those involving azoospermia, remain etiologically elusive.[2] Living organisms experience a plethora of DNA-damaging agents that can influence health and disease states. Robust DNA repair and damage-bypass mechanisms play a pivotal role in preserving genomic integrity by eliminating or tolerating damage, ensuring overall survival[3]. Consequently, the dysregulation of repair genes is anticipated to be associated with health implications, including a heightened susceptibility to infertility. Utilizing the R programming language and Limma package, we conducted a detailed analysis of results from two microarray datasets in the "NCBI-GEO" database. These datasets aimed to identify gene expression signatures underlying the pathogenesis of infertility, focusing on global gene expression profiling in testis samples from patients with severely impaired and normal spermatogenesis. The analysis resulted in a carefully curated list of around 19,000 genes. Subsequently, a compilation of DNA repair genes was acquired from the "PathCards" database. The use of R to create a Venn diagram aided in identifying common genes, suggesting the significant involvement of expression changes in DNA repair genes in the development of male infertility. A key gene of interest is DCAF10, encoding a factor linked with crucial repair genes, CUL4 and DDB1. This gene shows significant downregulation in expression levels in samples from individuals with non-obstructive azoospermia (LogFC= -2.49, Adj.P.Val=1.59E-10), suggesting its potential as a reparative gene associated with male infertility. Therefore, DCAF10 may be considered a relevant candidate gene contributing to infertility in men.

Keywords: Azoospermia; DNA repair; Male infertility; Microarray; R programming

References

[1 The American Society for Reproductive Medicine, "The management of infertility due to obstructive azoospermia" *Fertil Steril.*, 2008;90(5):S121–S124.

[2] Tournaye H, Krausz C, "Novel concepts in the aetiology of male reproductive impairment.", *Lancet Diabetes Endocrinol.*, 2017;5(7):544–553

[3] Chatterjee N, Walker GC., "Mechanisms of DNA damage, repair, and mutagenesis", *Environ Mol Mutagen.* 2017;58(5):235–263

ICB12-1402-1296

Re-purposing Clove Oil in Ovarian Cancer Treatment: A Multi-Dimensional Approach Involving Bulk RNA-Seq and scRNA-Seq

Mohadeseh Sadat Meftagh
a Department of Science and Culture University
mohadesehmeftagh@gmail.com

Abstract: Ovarian cancer (OC) is currently the fifth leading cause of cancer-related deaths among the women, with approximately 140,000 fatalities globally per year. To improve the prognosis of OC patients, novel therapeutic approaches are essential. Cancer Stem Cells (CSCs) are integral to ovarian cancer's entire development process, including initiation, metastatic progression, therapeutic resistance, and disease recurrence. Thus, targeted therapy, particularly against the Ovarian Cancer Stem Cells (OCSCs), is expected to be more effective and less toxic, potentially improving patient survival and reducing tumor relapse.

Our initial analysis utilized GSE13237 dataset, which includes data from 11 pairs of primary and metastatic OV tumors. We employed an RNA-seq pipeline along with the DESeq2 package, setting a threshold of p.value <0.05 and |logFC| > 1 to identify highly variable genes. This yielded 2134 significant genes. Subsequently, using the clusterProfile package and David online enrichment tool for KEGG, we pinpointed clove oil—a mild analgesic used for toothache and suggested for Diabetic Cardiomyopathy therapy—as a potential agent targeting the COL1A1 gene. This gene plays a significant role in metastasis and stem cells regulatory pathways.

To ensure accuracy, we further analyzed pooled scRNA-seq datasets GSE184880 and GSE158937, comprising 8 samples (5 healthy and 3 high-grade serous metastatic OC). The results significantly enriched the COL1A1 gene in the cancer cell cluster, providing clear evidence of its association with OCSCs.

By understanding the genetic underpinnings and potential therapeutic targets like the COL1A1 gene, we can pave the way for more effective and personalized treatments for ovarian cancer, offering hope for improved outcomes.

Keywords: Ovarian cancer,cancer stem cells; Clove oil; COL1A1(Collagen type I alpha 1)

ICB12-1402-1298

A subspace learning aided matrix factorization for drug repurposing

Amir mahdi Zhalehfar[a], Zahra Narimani[a*]
Department of Computer Science and Information Technology, Institute for Advanced Studies in BasicSciences (IASBS), Zanjan, Iran
*Corresponding authors email addresses: amirmahdizhalefar@iasbs.ac.ir, narimani@iasbs.ac.ir

Abstract: In the rapidly advancing field of drug discovery, the repurposing of existing pharmaceuticals for novel therapeutic applications has emerged as a promising strategy. Traditionally, drug design has been a costly and time-consuming endeavor, but recent advancements in high-throughput technologies and machine learning have significantly streamlined the process. This paper presents a novel approach to drug repurposing that integrates state-of-the-art computational methods, including graph-based analysis, matrix factorization, and machine learning techniques, to enhance the predictive capabilities of an existing model, iDrug, which was based on matrix factorization. The iDrug model, a pioneering tool in drug repositioning, operates within the interconnected domains of drug-disease and drug-target networks. By harnessing the power of graph-based methods, we can analyze the intricate relationships between drugs, targets, side-effects, and diseases, uncovering patterns that may not be apparent through traditional methods. Matrix factorization techniques are employed to decompose sparse drug-target and drug-disease matrices, revealing latent features that can predict drug-target and drug-disease similarities with unprecedented accuracy. To further refine the iDrug model, we propose the integration of Sparse and Low-Redundant Subspace Learning-based DualGraph Regularized Robust Feature Selection (SLSDR). SLSDR is an efficient method that addresses the challenges of feature selection in high-dimensional datasets, extracting meaningful patterns while discarding irrelevant or redundant features. This integration enhances the model's predictive accuracy, interpretability, and scalability, making it a more powerful and versatile tool for drug discovery. By leveraging the strengths of both computational methods and machine learning techniques, we have developed a novel framework that not only expands the potential therapeutic applications of drugs but also represents a significant step forward in the field, paving the way for more efficient and effective drug discovery processes.

Keywords: Bioinformatics; Drug Repositioning; Machine Learning; Computational Biology; SLSDR; iDrug

References

[1] Shang, Ronghua, et al. "Sparse and low-redundant subspace learning-based dual-graph regularized robust feature selection." Knowledge-Based Systems 187 (2020): 104830.

[2] Chen, Huiyuan, Feixiong Cheng, and Jing Li. "iDrug: Integration of drug repositioning and drug-target prediction via cross-network embedding." PLoS computational biology 16.7 (2020): e1008040.

ICB12-1402-1300

Identification of biomarkers involved in the development of grade II meningiomas using the systems biology approach

S.Ghanaatian[a], F.Dehghanian[b]*
[a]Department of Cell and Molecular Biology & Microbiology, Faculty of Biological Science and Technology, University of Isfahan, HezarJarib Street, Isfahan 81746-73441, Iran
[b]Department of Cell and Molecular Biology & Microbiology, Faculty of Biological Science and Technology, University of Isfahan, HezarJarib Street, Isfahan 81746-73441, Iran
*Corresponding author's email address: ghanaatian10.sara@gmail.com

Abstract:

Introduction

Meningiomas are the most common primary brain tumors in adults. Most of them are identified as benign (WHO grade I) and slow-growing lesions. Some Meningiomas are classified as atypical (WHO grade II), with increased mitotic activity, including four mitoses or more, spontaneous necrosis, increased cellularity, and small cells with a high nuclear-to-cytoplasmic ratio(1). Malignant Meningioma (WHO grade III) is associated with aggressive growth patterns reflecting their clinical and histopathologic features of malignancy and can spread by metastatic dissemination. Here, we investigated the potential biomarkers of developing grade II meningiomas using systems biology approaches(2).

Methods

Microarray data, including 14 meningioma samples, were obtained from the GEO database (GSE77259). The affy (1.80.0) package in R(4.3.2) was used to indicate gene expression profiling, evaluate data quality, and extract differentially expressed genes (DEG) between grade I and grade II samples. $|log2FC| \geq 0.58$, and p-value< 0.05 cutoff were set to identify the significant differentially expressed genes. The weighted gene co-expression network analysis (WGCNA) package (1.72-5) was used to construct the co-expression network. Then, two modules were selected and visualized with Cytoscape software (3.10.1). Furthermore, Gene Ontology analyses were performed using the EnrichR database.

Results

Our results indicated 294 genes were significantly differentially expressed in grade II compared to grade I meningioma samples. Orange and lightcyan modules were selected among the identified modules. Three hub genes, including *TCTEX1D1*, *KCNC4*, and *TRUB1*, were significantly downregulated in grade II vs grade I samples. These biomarkers have a role in ubiquitin-protein transferase inhibitor activity and mRNA pseudouridine synthesis.

Conclusion

The behavior and outcome of Grade II meningiomas are intermediate. Our analysis demonstrates some biomarkers that may contribute to developing grade II meningiomas.

Keywords: Meningioma, Differentially expressed genes, Co-expression network, WGCNA.

References

[1] Gousias K, Trakolis L, Simon M. Meningiomas with CNS invasion. Frontiers in Neuroscience. 2023;17.
[2] Apra C, Peyre M, Kalamarides M. Current treatment options for meningioma. Expert review of neurotherapeutics.2018;18(3):241-9.

ICB12-1402-1302

Structural and Functional Consequences of a Novel DDX3X Frameshift Mutation: A Computational Perspective

F. Seyedeh Yasaman Mousavi [a], Arash Salmaninejad [b] A. Seyedeh Yasaman Mousavi, C. Arash Salmaninejad*
[a]Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran
[b]Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), Tehran, Iran
*Corresponding author's email address: arash.salmany@yahoo.com

Abstract: DDX3X syndrome, an X-linked neurodevelopmental disorder, is characterized by pathogenic variants in the *DDX3X* gene [1, 2]. Clinical manifestations encompass a spectrum of symptoms, including delayed motor milestones, intellectual disability, epilepsy, autism spectrum disorder, and sensory reactivity disorders [3]. Here, we report a comprehensive investigation of a novel heterozygous frameshift mutation in the *DDX3X* gene in a 5-year-old girl with DDX3X syndrome, identified via whole exome sequencing (WES). Sanger sequencing confirmed the mutation's inheritance from her mother. The 3D structure of wild-type and mutant-type DDX3X protein was modeled and investigated using the homology modeling tool called MOE. Additionally, some web servers including MutationTaster, CADD, and VarSome were employed to assess the pathogenicity of the identified variant. Molecular docking analysis, facilitated by Autodock Vina and Ligplot, highlighted alterations in the binding affinity between DDX3X and its ligand, NXF1, which contribute to the export of mRNA from nucleus to cytoplasm due to the mutation. This study underscores the pivotal role of bioinformatic tools in elucidating the genetic and structural aspects of DDX3X syndrome, emphasizing the importance of understanding the functional consequences of frameshift mutations in disease pathogenesis.

Keywords: DDX3X syndrome; frameshift mutation; WES; MOE; Autodock Vina

References

[1] Blok LS, Madsen E, Juusola J, Gilissen C, Baralle D, Reijnders MR, Venselaar H, Helsmoortel C, Cho MT, Hoischen A, Vissers LE. Mutations in DDX3X are a common cause of unexplained intellectual disability with gender-specific effects on Wnt signaling. The American Journal of Human Genetics. 2015 Aug 6;97(2):343-52.

[2] Wang X, Posey JE, Rosenfeld JA, Bacino CA, Scaglia F, Immken L, Harris JM, Hickey SE, Mosher TM, Slavotinek A, Zhang J. Phenotypic expansion in DDX 3X–a common cause of intellectual disability in females. Annals of clinical and translational neurology. 2018 Oct;5(10):1277-85.

[3] Schroeder M. Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation? Biochemical Society transactions. 2011 Apr 1;39(2):679-83.Lennox AL, Hoye ML, Jiang R, Johnson-Kerner BL, Suit LA, Venkataramanan S, Sheehan CJ, Alsina FC, Fregeau B, Aldinger KA, Moey C. Pathogenic DDX3X mutations impair RNA metabolism and neurogenesis during fetal cortical development. Neuron. 2020 May 6;106(3):404-20.

ICB12-1402-1303

scFedVI: A Privacy-Preserving Approach to Mitigating Batch Effects in Single-Cell RNA-Sequencing Data

P. Mokhber [a], A. Gargoori Motlagh[*b]
a Department of Computer Science, Sharif University of Technology, Tehran, Iran
b Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
*Corresponding author's email address: *alireza.gargoorimotlagh@ee.sharif.edu

Abstract: The growing field of single-cell RNA sequencing (scRNA-seq) has revolutionized our understanding of cellular heterogeneity. In this study, we introduce Single-Cell Federated Variational Inference (**scFedVI**), a novel federated learning-based method to address the challenge of batch effects in scRNA-seq data analysis. Batch effects, arising from variances in cell processing such as different chips, sequencing lanes, or harvest times, significantly affect transcriptome measurements, creating discrepancies within and across experiments. To mitigate these, we integrated deep neural networks, specifically Variational Autoencoders (VAEs) [1], into our federated learning framework for sophisticated batch correction, enhancing biological insights while maintaining data integrity. Our approach utilizes the inherent differences in each client's dataset as a feature rather than a limitation, enabling more robust and generalizable models**.** By distributing the learning process across clients, each possessing their unique scRNA-seq dataset with distinct batch characteristics, we employed the Federated Averaging (Fed-Avg) [2] algorithm to aggregate the learned models. This approach, prioritizing data privacy, demonstrates enhanced effectiveness in batch effect correction compared to running single-cell variational inference individually on each client's data. Performance evaluation using the k-nearest-neighbor batch effect test (kBET) and the Adjusted Rand Index (ARI) for clustering confirms that scFedVI outperforms scVI [3], a current leading method in batch correction and integration for single-cell data. Furthermore, we establish the robustness of scFedVI by testing various scenarios involving different numbers of clients, ranging from 2 to 5. Our results, validated across diverse pancreatic and nervous system scRNA-seq datasets, illustrate that the scFedVI not only effectively corrects batch effects but also utilizes these variations to enhance overall data analysis. This is a significant advancement over conventional non-private batch correction methods, which typically aim to merely eliminate these effects. This method opens new avenues for collaborative research across different laboratories without compromising data privacy or integrity.

Keywords: Single-cell RNA-seq; Batch Effects; VAE; Privacy; Federated Learning

References

[1] D. P. Kingama, M. Welling, "Auto-Encoding Variational Bayes", arXiv:1312.6114, 2013.
[2] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. Arcas "Communication-Efficient Learning of Deep Networks from Decentralized Data", arXiv:1602.05629, 2016.
[3] Gayoso, A., Lopez, R., Xing, G. et al. "A Python library for probabilistic analysis of single-cell omics data", Nat Biotechnol 40, 163–166 (2022). https://doi.org/10.1038/s41587-021-01206-w

ICB12-1402-1304

Zika virus vaccine design based on epitopes

Yasmin Alamdar*, Mohamadreza Cheraghian1, Meisam Akrami
Department of Biology, Islamic Azad University Central Tehran Branch, Tehran, Iran
*Corresponding author's email address: yasminalmdar@gmail.com

Abstract: Zika virus, a member of the Flaviviridae family, transmitted by Aedes mosquitoes, poses a global health threat, causing microcephaly in newborns. Despite extensive research, no approved vaccine for Zika virus exists.We propose a novel Zika virus vaccine design. Initial epitope extraction from the Immune Epitope Database (IEDB) was followed by assessment for allergenicity, toxicity, and immunogenicity using ToxinPred, AllerTop, and Vaxijen servers, respectively. Confirmed major histocompatibility complex class I (MHC I) epitopes were linked with an AAY linker, while MHC class II epitopes were connected using a GPGPG linker. B cell epitopes were linked with a KK linker. An adjuvant, interleukin-12 (partial), was incorporated from the NCBI database, and a His-Tag (6-H numbers) was added for purification.The designed vaccine underwent comprehensive evaluation for allergenicity, toxicity, and immunogenicity. It molecular weight was 38 K daltons, with an isoelectric point of 7.03, indicating stability. Structural analysis confirmed a favorable structure, exhibiting 99.5% similarity with c6sffA. Molecular docking simulations with the HDock server demonstrated promising interactions with MHC class I (Docking score: -252.66) and MHC class II (Docking score: -257.69) representative proteins, suggesting a promising approach to combat Zika virus infection.

Keywords: Max. Zika; Virus,Vaccine; B-cells,Epitopes

References

[1] Malone RW, Homan J,Callahan MV,Glasspool-Malone J,Damodaran L,Schneider A,et al.(March 2016)."Zika Virus:Medical Countermeasure Development Challenges".PLOS Neglected Tropical Diseases.10 (3):e0004530.doi:10.1371/journal.pntd.0004530.PMC 4774925.PMID 2693453.

[3] https://www.who.int/news-room/fact-sheets/detail/zika-viru.

ICB12-1402-1305

Deep Convolutional Neural Network Model for Predicting MHC I Binding Affinity in Peptide-Based Therapeutics

Narges Sangarani Pour *
Department of Shahid Beheshti University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: nargessangaranip@gmail.com

Abstract: The intricate interactions between human leukocyte antigens (HLAs) and peptides are fundamental to the human immune system's functionality. A key application of understanding these interactions is in the realm of peptide drug discovery and the development of therapeutic mRNA. This study introduces a pioneering deep convolutional neural network model (DCNN) designed to predict Major Histocompatibility Complex Class I (MHC I) peptide binding affinities. Notably, this model autonomously learns the encoding of MHC sequences and their binding contexts, circumventing the need for explicit MHC-peptide bound structure data.

A distinctive feature of the proposed DCNN model is its ability to adapt to peptides of variable lengths, enhancing its robustness and applicability across a diverse range of peptide sequences. This adaptability is crucial given the inherent length variance in naturally occurring peptides. The performance of the model was rigorously evaluated using a test set comprising 30% of the total data, ensuring a comprehensive assessment of its predictive capabilities.

The evaluation metrics underscore the model's high efficacy and reliability: it achieved an accuracy of 91.216%, precision of 71.499%, recall rate of 93.243%, and an F1-score of 80.936%. Moreover, the model demonstrated exceptional discriminative ability, as evidenced by an Area under the Receiver Operating Characteristic Curve (AUC) of 0.975. These metrics collectively highlight the model's potential as a significant tool in peptide-based therapeutic research.

In conclusion, this DCNN model stands as a significant advancement in computational immunology, offering a potent tool for predicting HLA-peptide interactions. Its implications extend to enhancing peptide drug discovery and the design of therapeutic mRNA, marking a noteworthy contribution to biomedical research and healthcare innovation.

Keywords: Deep Convolutional Neural Network (DCNN); Major Histocompatibility Complex (MHC) I; Peptide Binding Affinity; Therapeutic mRNA; Peptide-Based Therapeutics

ICB12-1402-1306

Reconstruction of the miRNA-mRNA regulatory network specifically in the remission and relapse phase of people with multiple sclerosis

Zahra Ahmadi[a], Bita Bakhtarnia[a], Maryam Talaei[a], Mohammad Mehdi Naghizadeh[b]*
a Student Research Committee, Fasa University of Medical Sciences, Fasa, Iran
b Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran
*Corresponding author's email address: mmnaghi2@yahoo.com

Abstract: Multiple sclerosis (MS) is a chronic, unpredictable disease of the central nervous system that is known for its periods of remission and relapse. The molecular mechanisms underlying these phases are complex and not fully understood. miRNAs are small non-coding RNAs that regulate gene expression at the post-transcriptional level. This study aims to elucidate these mechanisms by reconstructing the miRNA-mRNA regulatory network specific to these phases of MS. GSE41890 was downloaded from GEO database. In this dataset, the gene expression data was taken from the peripheral blood leukocytes. The difference in gene expression in the relapse and remission groups was analyzed compared to the control group. Then, using the miRTarBase database, we identified the microRNAs related to the mRNAs of each group and reconstructed the regulatory network using Cytoscape software. Our results showed H2AC18, H2BC21 and H3C10, hsa-miR-34a-5p, hsa-miR-192-5p and hsa-miR-5693 have regulatory role in relapse network and MCM6, SCD and GEN1, hsa-miR-124-3p, hsa-miR-192-5p and hsa-let-7b-5p have regulatory role in remission network. In this study we identified key miRNAs and mRNAs that may serve as potential biomarkers or therapeutic targets in MS.

Keywords: Relapse; remission; multiple sclerosis; miRNA

ICB12-1402-1312

New feature selection method based on random forest for cancer classification

Nima Mahmoudi Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran , nimam454@gmail.com
Mohammad Reza Jafari , Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran, mr.jafaryi9393@gmail.com
Jamshid Pirgazi , Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran,
*Corresponding author's email address: j.pirgazi@mazust.ac.ir

Abstract: Cancer diagnosis based on gene expression data is a critical area of research, offering insights into molecular mechanisms and aiding personalized medicine. In this study, we employed machine learning algorithms to classify cancer types using a gene expression dataset [1]. The high dimensionality, noise, and interpretability are challenges addressed in this research.
Understanding gene expression patterns in cancer is vital for advancing cancer biology and treatment strategies. Machine learning models, such as SVM, kNN, MLP, Decision Tree, Naive Bayes, and Random Forest, were applied for accurate cancer type prediction, showcasing potential applications in clinical settings [2].
High-dimensional gene expression data poses challenges, addressed through preprocessing and feature selection. Label encoding and dataset splitting were performed, and Random Forest identified top genes. Results indicated improved performance, particularly in SVM, MLP, and Random Forest models, showcasing enhanced classification accuracy [1][3].
The study emphasizes the potential of machine learning in cancer type classification. Models achieved high accuracy, with SVM, MLP, and Random Forest outperforming others [1][2]. Feature selection with Random Forest contributed to improved interpretability. Further research should focus on addressing interpretability challenges and validating results in clinical contexts.
Machine learning models offer promising results in cancer type classification based on gene expression data [1][2][3]. SVM, MLP, and Random Forest demonstrated enhanced accuracy. Addressing interpretability challenges is crucial for translating these findings into clinical applications. Notably, the study showcases improvements in detection rates, highlighting the significance of these models in advancing cancer diagnostics [1].
Keywords: Diagnosis , Interpretability , Classification , Random Forest , Detection rates

References

[1] Li, H., & Greene, C. S. (2019). Construction of a cancer cell line knowledge base and application of machine learning approaches for drug response prediction. Bioinformatics, 35(11), 2015-2022

[2] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[3] Cho, H., Howlader, N., Mariotto, A. B., Cronin, K. A., & Schwenn, M. (2011). Cancer survival statistics based on the Surveillance, Epidemiology, and End Results (SEER) database. Journal of the National Cancer Institute, 103(4), 1-11.

ICB12-1402-1317

IDENTIFICATION OF CANDIDATE TREATMENT TARGETS IN ESOPHAGEAL CANCER: A COMPREHENSIVE BIOINFORMATIC ANALYSIS

Amirhossein Rajabi[1], Alireza Parnian[2], Manica Negahdaripour[3,*]
Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran.
Pharmaceutical Sciences Research Center, Shiraz University of Medical Science, Shiraz, Iran

Abstract: Esophageal cancer, is among the leading causes of cancer-related mortality [1]. Despite recent progress in early diagnosis and more effective treatment options, patients have variable prognosis [2]. We conducted a comprehensive bioinformatic approach to identify novel therapeutic and prognostic targets.

The gene expression omnibus (GEO) database was used to compare transcriptomic profiles of esophageal tumor and normal tissue and identify differentially expressed genes (DEG)s. The DEGs were used to construct an interaction network using the STRING. We limited the number of our genes using centrality parameters such as degree, closeness, and betweenness. Next, pathway enrichment analysis was conducted. Hub genes of the network were selected based on the effect of genes on overall survival (OS) in the cancer genome atlas database (TCGA) and their effect on clinical staging of the disease was interrogated. Finally immunohistochemical (IHC) staining of hub proteins in esophagus tumors compared to the normal tissue was examined using Human Protein Atlas and mutational profile of the hub genes was investigated in the Gene cBioPortal database.

We identified IDO1, COL4A1, and ATF3 as the main hub genes. Four modules were identified in the network that showed the most correlation with Cytokine-cytokine receptor interaction, protein digestion, ECM-receptor interaction, relaxin signaling pathway and Epithelial cell signaling in Helicobacter pylori infection pathways in enrichment analysis. Surveying Drug Bank database for hub genes as target genes recommended Cannabidiol as a drug candidate with an inhibitory effect on overexpression of IDO1 in tumor tissues.

Using a multi-aspect systems biology approach, we identified genes with the main regulatory role and survival impact on esophageal cancer. Targeting these genes with currently available or novel synthetic agents can be of potential survival advantage.

Keywords: Esophageal Cancer, Systems Biology, Hub Gene, Drug Target

References

[1] Huang FL, Yu SJ. Esophageal cancer: Risk factors, genetic association, and treatment

[2] Kato H, Nakajima M. Treatments for esophageal cancer: a review

ICB12-1402-1318

Identification of potential prognostic biomarkers in lung adenocarcinoma: A network-based approach

Mehrdad Ameri[a,b] , Aryan Ghorbani[c] , Amin Ramezani[a,b*]
Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran
Shiraz Institute for Cancer Research, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran
Genetics Department, Faculty of Science, Shahrekord University, Shahrekord, Iran
*Corresponding author's email address: aramezani@sums.ac.ir

Abstract: Non-small cell lung cancer (NSCLC), one of the most common malignant tumor globally with an extremely high mortality rate, is classified into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma (1, 2). Adenocarcinoma is the most common type among these three (3). Long-term survival remains low because most patients are diagnosed late. Therefore, it is vital to understand the molecular mechanisms and identify biomarkers for prognosis and early detection of patients.
Two gene expression profiles related to lung adenocarcinoma (GSE32863 and GSE75037) were obtained from gene expression omnibus (GEO). To identify differentially expressed genes (DEGs) between tumor and normal samples, GEO2R was used. Enrichr was employed for enrichment analysis. Using Enrichr, GO terms and KEGG pathway enrichment analysis was accomplished. Moreover, to reconstruct the protein-protein interaction (PPI) network, STRING was used. Cytoscape 3.9.1 was utilized to visualize and analyze the network of DEGs. Using the CytoHubba plugin and based on betweenness, the top 15 hub genes were selected. Hub genes were validated using GEPIA.
Considering GO biological process, DEGs are associated with "Cellular Response To Cytokine Stimulus". GO Molecular function revealed the relationship of DEGs with "Calcium Ion Binding". GO cellular components showed that DEGs are related to "Collagen-Containing Extracellular Matrix". KEGG pathway enrichment analysis revealed the relationship of DEGs with "Complement and coagulation cascades". Based on betweenness centrality, GAPDH, IL6, IL1B, CDH1, UBB, CD44, JUN, ERBB2, CAV1, PPARG, PECAM1, CDH5, CD34, MMP9, and COL1A1 were considered as hub genes. By performing overall survival analysis using GEPIA, we observed that GAPDH, and PECAM1 have a reverse relationship with the survival of lung adenocarcinoma patients.
By analyzing microarray data and using a network-based approach, we identified GAPDH, and PECAM1 as potential prognostic biomarkers in lung adenocarcinoma.
Keywords: Lung Adenocarcinoma; Differentially Expressed Genes; DEGs; Hub Genes

References

[1] Davidson MR, Gazdar AF, Clarke BE. The pivotal role of pathology in the management of lung cancer. J Thorac Dis. 2013;5 Suppl 5(Suppl 5):S463-78.

[2] Giaccone G, He Y. Current knowledge of small cell lung cancer transformation from non-small cell lung cancer. Seminars in Cancer Biology. 2023;94:1-10.

[3] Seguin L, Durandy M, Feral CC. Lung Adenocarcinoma Tumor Origin: A Guide for Personalized Medicine. Cancers (Basel). 2022;14(7).

ICB12-1402-1319

Inference of gene regulatory network using dimension reduction methods and rotation forest

Marzieh Emadi[1], Farsad Zamani Boroujeni[2]*, Jamshid Pirgazi[3]
[1]Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
emadi.marzieh@gmail.com
[2] Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
farsad.zamani@srbiau.ac.ir
[3]Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran
Behshahr, Iran
*Corresponding author's email address: j.pirgazi@mazust.ac.ir

Abstract: Inferring gene regulation networks from gene expression data has been one of the most important challenges during the last decade. These networks are used in various fields such as disease diagnosis, gene therapy, etc. In order to infer gene regulatory networks using computational methods, there are various approaches. However, gene expression data are often noisy and gene regulation networks are very sparse. In fact, in these networks, the number of regulators of a target gene is very small. Most machine learning methods for predicting regulators of a target gene face a high false positive rate. In this article, in order to overcome this problem, we first reduce the number of regulatory genes using dimension reduction methods. Then, in the next step, the regulatory genes of each gene are identified using the rotation forest. The evaluation results show that dimension reduction methods increase the efficiency of rotation forest. In addition, the t-sne method has a better performance than other dimension reduction methods such as SVD, PCA, etc.

Keywords: Gene regulatory network; gene expression data; rotation forest; dimension reduction; machine learning

ICB12-1402-1320

Study of the Effect of Cinnamon Plant Hydrosol (Cinnamomum verum) on Peroxidase and Lipase

Z. Vahdani [a].M.Zarrabi [b]

[a] Department of Computer Science, Faculty of Mathematical Sciences, Alzahra University, Tehran, Iran
[b] Department of Microbiology, Faculty of Biological Sciences, Alzahra University, Tehran, Iran

Abstract: Peroxidase enzyme is considered as one of the important factors of biochemical spoilage. It is an iron-containing enzyme and catalyzes a large number of reactions [1]. Hydrogen peroxide plays a role as a reductant by donating an electron, and its empirical relationship with the reduction of taste and color in raw materials has been proven. Cinnamon hydrosol includes Vulgarol, Emersol, Cinnamic acid methyl ester, Methyl palmitate, and Oleic acid. One of the possible reasons for the antibacterial effect of cinnamon hydrosol can be related to the presence of cinnamic acid derivatives in cinnamon hydrosol [2]. Therefore, it seems that the extracted hydrosol of cinnamon can be used as an antibacterial agent against Bacillus cereus, anti-mold against Rhizopus oryzae, and also to prevent the activity of polyphenol oxidase enzyme. In this study, with the help of easymodeller software, the tertiary structure of cinnamon was predicted and the interaction between cinnamon and cinnamon compounds with polyphenol oxidase and lipase was investigated using cluspro server and Edock server. Also, cinnamon compounds were separately performed in PyRx software as virtual screening, and the results of docking were examined in pymol and ligplus software. The results indicate that cinnamon can be used as a fruit preservative. Therefore, deactivating polyphenol oxidase and lipase increases the shelf life of fruits and vegetables during storage, and it can be used as an organic substance for producing sprays or solutions that fruits are soaked in to prevent browning and degradation of fruits and increase the life of fruits.

Keywords: Cinnamomum verum; Anti-fungal

References

[1] N. G. Tzortzakis, “Maintaining postharvest quality of fresh produce with volatile compounds”, Innovative Food Science & Emerging Technologies, 8(1) (2007) 111-116.

[2] M. Mirfeizi, Z. Mehdizadeh Tourzani, S. Z. Mirfeizi, M. Asghari Jafarabadi, H. Rezvani, M. Shoghi, “Effects of cinnamon on controlling blood glucose and lipids in patients with type II”, 2014.

ICB12-1402-1321

Improved the performance of pathway topology based methods using Perti Net

Fatemeh Mansoori*
Department of Applied Mathematics and Statistics, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: f.mansoori@mcs.ui.ac.ir

Abstract: Disruption of the normal functioning of cell signaling pathways frequently results in diseases. Precisely identifying disrupted signaling pathways is essential for understanding diseases. Pathway analysis methods have been developed for the specific purpose of identifying significantly disrupted signaling pathways in a given condition. Among these methods, some take into account the topologies of the pathways in their analysis, known as pathway topology-based (PT-based) methods. These types of methods are superior to other types of pathway analysis methods, as they take into account the internal structure of the pathway .

PT-based methods model signaling pathways using graphs, which have limitations in capturing all types of relations within a pathway. Research has demonstrated that modeling signaling pathways with Petri nets can address the limitations of graph-based modeling. PAPET is a method that employs Petri net modeling and reports better results compared to other PT-based methods. However, the algorithm used in the analysis of this method is excessively time-consuming .

In this research, we attempted to simplify the algorithm used in PAPET. The proposed analysis can be incorporated as an additional layer preceding any PT-based method. We incorporate our proposed analysis as an additional layer to SPIA and PADOG methods which are well-known PT-based methods. The results indicate that the additional layer would enhance the performance of the aforementioned methods. Applying the target pathway technique on the benchmark of datasets from various diseases shows that the proposed strategy prioritizes the target pathways better than other methods including PAPET.

Keywords: Pathway analysis; Enrichment analysis; Petri net; Pathway topology based

References

[1] Z. Kangmei, S., Yon Rhee, "Interpreting omics data with pathway enrichment analysis", Trends in Genetics, 2023.

[2] D. Chicco, G. Agapito, "Nine quick tips for pathway enrichment analysis", PLoS computational biology, vol. 18, p. e1010348, 2022.

[3] M. Sarah, et al. "On the influence of several factors on pathway enrichment analysis.", Briefings in Bioinformatics, vol. 23, p.bbac143, 2022.

ICB12-1402-1322

In-Silico Pathogenicity Prediction of a Human Missense Variant (rs147396263) of *NANOS2* Gene In the Patients with Varicocele Disease

R. Ghorbanian, M. Khatami*
Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: m.khatami@yazd.ac.ir

Abstract: Varicoceles is defined as an abnormal, excessive dilation of the pampiniform plexus of veins within the scrotum. Varicocele is highly prevalent among infertile male and can result in deleterious effects on spermatogenesis. Recent researches illustrated that genetic mutations might play role in the onset of varicocele disease through imbalanced oxidative stress and DNA fragmentation of sperm DNA. One of the genes associated with imbalanced oxidative stress in varicocele is Nitric Oxide Synthesis Family called *NANOS* (*NOS*) gene including: *NANOS1*, *NANOS2* and *NANOS3*. NANOS2 protein which is scrutinized in the current study acts as an mRNA binding protein which can regulates protein translation leading to germ cell differentiation into male faith. The aim of this study is to investigate a genetic variant (rs147396263) of NANOS2 gene in terms of mutant protein alterations in both structure and function in comparison with the wild type by means of bioinformatic pathogenicity prediction tools such as: Mupro, iMutant, HOPE, PolyPhen-2 and Predict-SNP. The results demonstrated that the missense mutation p. 73 R>C could affect protein function by changes in protein stability, hydrophobicity and conservation as well as protein structure. All in all, identifying genetic biomarkers by scrutinizing molecular mechanisms underlying varicocele might shed light into a new perspective of varicocele prognosis, diagnosis and treatment which leads to a better management of the disease for health-care professionals.

Keywords: Bioinformatic Tools; NOS; Oxidative Stress; Point Mutation

References

[1] R. Finelli, K. Leisegang, "Oxidative Stress: A Comprehensive Review of Biochemical, Molecular, and Genetic Aspects in the Pathogenesis and Management of Varicocele", J. Men's Health 2022 40(1): 87-103.

[2] D. Arya, N. Balasinor, "Varicocoele-associated male infertility: Cellular and molecular perspectives of pathophysiology", J. Andrology 2022 10: 1463-1483.

[3] A. Hassanin, H. Ahmed, "A global view of the pathophysiology of varicocele", J. Andrology 2018 6: 654-661.

ICB12-1402-1323

Applying various strategies in machine learning models to predict Type II diabetes by using the risk factors

Fatemeh Mansoori*, Nima Chelongar, Ali Zarrinkhat
Department of Applied Mathematics and Statistics, University of Isfahan, Isfahan, Iran
*Corresponding author's email address: f.mansoori@mcs.ui.ac.ir

Abstract: Diabetes is one of the most common chronic diseases in the world including in Iran. The prevalence of type II diabetes in Iran in the age range of 25 to 64 is 7.7 percent. Despite the increasing prevalence of the disease, its treatment and effective control are suboptimal. There appear to be significant gaps in disease detection and treatment. Early detection can lead to lifestyle modification and more effective treatment. Hence, this highlights the significance of developing tools for predicting diabetes using common health risk factors.

The Behavioral Risk Factor Surveillance System is an annual survey that gathers healthcare information from Americans regarding behaviors that pose health risks. We used the datasets of 253,680 survey responses to the CDC's BRFSS2015. The dataset has 21 feature variables and it is not balanced. The features include high blood pressure, high cholesterol, Body Mass Index, and so on. The target feature Diabetes has 2 classes the diabetes class, and no diabetes class. We explore to answer the following research question, whether survey questions from the BRFSS provide accurate predictions of whether an individual has diabetes.

Due to the imbalance in the dataset, we employed data balancing techniques such as ADASYN and SMOTE to address the issue. Subsequently, we built multiple machine learning models to predict type II diabetes using potential health risk factors.

Out of all predictive models, XGBOOST demonstrates the highest accuracy 81% but has a 61% recall for diabetics. On the other hand, Random Forest has an accuracy of 73% and a 72% recall for diabetics. Consequently, the Random Forest model is preferred for initial screening for type II diabetes, because it has the highest sensitivity and therefore, detection rate. This finding can be used for the early detection of individuals with diabetes based on their general health information.

Keywords: Type 2 Diabetes; Matching Learning; Healthcare;

ICB12-1402-1324

DeepSiPPI: Enhancing Protein-Protein Interaction Prediction through Siamese Neural Networks and Sequence-to-Image Transformation

Ali karimi, Jamshid Pirgazi
Department of Computer Science,University of Science and Technology of Mazandaran, Behshahr, Iran
Department of Computer Science,University of Science and Technology of Mazandaran, Behshahr, Iran
Ali.karimi@mazust.ac.ir
*Corresponding author's email address: j.pirgazi@mazust.ac.ir

Abstract: This study focuses on the analysis and prediction of Protein-Protein Interactions (PPI) within biological systems. We propose an innovative methodology employing a Siamese Neural Network architecture with a Convolutional Neural Network (CNN) as its underlying framework. Notably, we introduce a novel preprocessing step involving the conversion of protein sequences into images through the application of a robust statistical method. Subsequently, these transformed representations are utilized as input data for the Siamese Neural Network, a choice motivated by its intrinsic capacity for effective feature extraction. The discerning ability of the Siamese Neural Network proves instrumental in discerning subtle patterns and features crucial for the identification of protein interactions. The presented approach not only showcases its utility in refining interaction predictions but also underscores the potential for advancing the comprehension of intricate protein networks, thereby contributing to the broader landscape of bioinformatics research.
Keywords: Protein-Protein Interaction; Siamese Neural Network; Protein Representation; Sequence-to-Image Transformation; Bioinformatics

References

[1] Mahapatra, Satyajit, and Sitanshu Sekhar Sahu. "Improved prediction of protein–protein interaction using a hybrid of functional-link Siamese neural network and gradient boosting machines." *Briefings in Bioinformatics* 22, no. 6 (2021): bbab255.
[2] Chen, Muhao, Chelsea J-T. Ju, Guangyu Zhou, Xuelu Chen, Tianran Zhang, Kai-Wei Chang, Carlo Zaniolo, and Wei Wang. "Multifaceted protein–protein interaction prediction based on Siamese residual RCNN." *Bioinformatics* 35, no. 14 (2019): i305-i314.
[3] Hu, Lun, Xiaojuan Wang, Yu-An Huang, Pengwei Hu, and Zhu-Hong You. "A survey on computational models for predicting protein–protein interactions." *Briefings in bioinformatics* 22, no. 5 (2021): bbab036.

ICB12-1402-1325

Enhancing Precision for Background Mutation Rate Estimation in Cancer Genomes

Farideh Bahari[1, 2], Reza Ahangari Cohan[1], Hesam Montazeri[2*]

1 Department of Nanobiotechnology, New Technologies Research Group, Pasteur Institute of Iran, Tehran, Iran
2 Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Iran
*Corresponding authors email addresses: hesam.montazeri@ut.ac.ir

Abstract: Accurate estimation of the background mutation rate (BMR) is crucial for identifying cancer drivers, significantly impacts the precision of prioritizing both coding and non-coding cancer drivers. This study investigates the influence of neighboring regions on BMR prediction by applying a novel transformer-based neural network using self-attention mechanisms. Additionally, it systematically compares various machine learning algorithms, investigates the impact of the number of genomic elements in the training set, and explores different bin generation methods for BMR prediction.

The analysis was performed on 2253 whole cancer genomes of 33 cancer types from non-melanoma-lymphoma donors. Somatic mutations were used from the PCAWG project, with the TCGA portion of data obtained from the Genotypes and Phenotypes (dbGap) database (project #32607). A total of 971290 variable-size intergenic genomic coordinates were derived by removing functional genomic elements based on PCAWG definition and retaining callable genomic regions. A set of 1372 genomic context and epigenomic features affecting mutation rate were extracted for each genomic region. This study compared a transformer-based model to XGBoost, random forest, classic neural network, and binomial generalized linear model. Model performance was compared using Spearman correlation and mean squared error.

We comprehensively analyzed the performance of various models for BMR estimation, examining the effects of binning strategies, sample sizes, and model types on the accuracy of predictions. The results of our analysis indicate that the transformer model outperformed traditional deep neural networks and the binomial generalized linear model. The performance of the transformer model was similar to the XGBoost method.

Keywords: cancer genomics; transformer neural networks; background mutation rate

ICB12-1402-1326

Enhancing Model Evaluation in Single-Cell Perturbation Response Prediction

Mina Karimpour[1,2] and Hesam Montazeri[2*]
[1] Department of Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran
[2] Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran
*Corresponding author's email address: hesam.montazeri@ut.ac.ir

Abstract: Various computational approaches such as variational autoencoders [1] and neural optimal transport [2] methods have been used to predict single-cell perturbation responses. Correlation or distance metrics, which compare observed and predicted gene expression profiles, have previously been used for model assessment. However, we have observed that these metrics might not accurately represent the performance of various prediction models due to their sensitivity to the variable ranges of individual genes. To enhance model evaluation, this study introduces several novel evaluation metrics intended to provide more accurate measures of prediction accuracy.

This study introduces a range of evaluation metrics to assess the accuracy of model predictions in comparison to observed data. Alongside correlations across samples, we evaluated the normalized root mean square error, R-squared, and the average Pearson correlation coefficient across genes, which are insensitive to the variable ranges of genes. Moreover, we employed the Frobenius norm to measure the difference between the observed and predicted correlation matrices as an additional metric. This metric can be applied across both samples and genes. Additionally, we calculated the Area Under the Curve (AUC) value of the Receiver Operating Characteristic (ROC) curve for a discriminative model, aiming to predict the observed correlation matrix by adjusting the thresholds of predicted correlations.

Two single-cell RNA sequencing datasets were used including PBMC cell types under interferon-beta stimulation [3], and a subset of the SCIPlex3 dataset [4] comprising three cell types and 10 single-dose drug conditions. Following necessary preprocessing using *Scanpy*, the CPA and cellOT models, along with their baselines (no-perturbation, vector arithmetic, and PCA + vector arithmetic), were applied to predict gene expression profiles of different conditions across cell types. Evaluation of predictions against observed values using the aforementioned metrics indicated that relying solely on assessment across samples (without considering the range of each gene) might be misleading, as baseline models demonstrated better accuracy using gene expression range and graph-based methods. Permutation test approaches were employed to establish a foundation for comparing the evaluations of model performance against random models. We demonstrated that our introduced evaluation metrics provide valuable insights into the accuracy of model performance.

Keywords: perturbation responses; single cell RNA sequencing; evaluation metrics

References

[1] M. Lotfollahi, A. Susmelj, et al. "Predicting cellular responses to complex perturbations in high-throughput screens." Mol Syst Biol., (2023): e11517.
[2] C. Bunne, S. Stark, et al. "Learning single-cell perturbation responses using neural optimal transport." Nat Methods., 20.11 (2023): 1759-1768.
[3] H. Kang, M. Subramaniam, et al. "Multiplexed droplet single-cell RNA-sequencing using natural genetic variation." Nat Biotechnol., 36.1 (2018): 89-94.
[4] S. Srivatsan, J. McFaline-Figueroa, et al. "Massively multiplex chemical transcriptomics at single-cell resolution." Science 367.6473 (2020): 45-51

ICB12-1402-1329

Cancer Gene Expression Classification Using Machine Learning Methods

HessamOddin Pourrostami, Dr. Mousa Nazari*, Dr. Mohammad M. AlyanNezhadi*
Department of Computer Science, University of Science and Technology of Mazandaran, Behshahr, Iran
*Corresponding author's email address: Hessam.pourrostami@gmail.com

Abstract: Cancer, a diverse group of diseases, and a major malignant, remains a massive challenge in medical research and its related subject. Using gene expression data, especially real-world data to predict the outcome of a patient's tumor makes discovery a reliable tool for reducing death arising from this disease. The PANCAN dataset included random gene expression records from patients with 5 classes of tumor types: Breast Invasive Carcinoma (BRCA), Kidney Renal Clear Cell Carcinoma (KIRC), Colon Adenocarcinoma (COAD), Lung Adenocarcinoma (LUAD), and Prostate Adenocarcinoma (PRAD), which helps as a remarkable resource for solving the complexities of cancer gene expression and its challenges in medicine and other related majors. Our study aimed to show machine learning techniques and their quality for tumor classification, exploring the potential of several algorithms such as Support Vector Machines (SVM), k-nearest Neighbors (KNN), and Multi-Layer Perceptron (MLP) algorithms.

To explore the dataset, we encouraged comprehensive research using Machine Learning and Deep Learning algorithms such as SVM, KNN, MLP, and 1D-CNN algorithms for tumor classification. we employed different approaches with feature selection techniques and 1D-CNN models together, to refine the classification process. The effort with this approach did not yield a well-done result because of unbalancing in the LUAD class and lack of data which affects the training process of deep learning models. Remarkably, our machine learning models were well-accurate and reliable in tumor classification, especially the SVM model gained 100% accuracy in predicting between the five classes of tumors. This remarkable achievement emphasizes the resilience of Support Vector Machines (SVM) in effectively managing the complexity of multi-dimensional gene expression data.

Keywords: Cancer; Gene Expression; SVM; Tumor Classification

ICB12-1402-1330

Translation Initiation Site Prediction based on the TITER and Machine Learning Method

Ali Fathalizadeh[1], Ali Ghanbari Sorkhi[1], Jamshid Pirgazi[1]
[1] Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran
Fathali.9080@gmail.com
*Corresponding authors email addresses: Ali.ghanbari@mazust.ac.ir, j.pirgazi@mazust.ac.ir

Abstract: Translation initiation is a crucial step in the regulation of gene expression. The selection of the translation initiation site (TIS) is critical, as it establishes the correct open reading frame for mRNA decoding. Annotated translation initiation sites initiate the translation process and may commence from several alternative TISs, including AUG and non-AUG codons, posing a significant challenge.

Nucleotides flanking the repeat region, particularly those in close proximity to the start site, are believed to enhance translation initiation. Consequently, extensive research has been conducted on this issue. In this paper, we propose a machine learning-based method to aid in the identification of ATGs and translation start sites that are nearly identical. Kozak, random forest classification (RFC), and Translation Initiation siTE detector (TITER) similarity score algorithms are employed in this article, with the RFC algorithm yielding the most favorable results.

The final method is evaluated on ATG RFC and Near-Cognate RFC datasets. The RFC model demonstrates an accuracy of 87.79% in 344 balanced cases. The experimental results confirm that the proposed approach, in comparison to similar methods, achieves a more concise set of features while maintaining high accuracy.

Keywords: codon; Nucleotide; Translation mRNA; machine learning; TITER.

References

[1] Gleason, A. C., Ghadge, G., Chen, J., Sonobe, Y., & Roos, R. P. (2022). Machine learning predicts translation initiation sites in neurologic diseases with nucleotide repeat expansions. PloS one, 17(6), e0256411.

ICB12-1402-1332

AI-powered PET imaging analysis for enhanced Parkinson's disease diagnosis

A. Fallah[a], Y. Zandiyevakili[a], Z. Salehi[b, c], K. Kavousi*[d]
[a] School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran
[b] Hematology, Oncology and Stem Cell Transplantation Research Center, Tehran University of Medical Sciences, Tehran, Iran.
[c] Research Institute for Oncology, Hematology and Cell Therapy, Tehran University of Medical Sciences, Tehran, Iran
[d] Institute of Biochemistry and Biophysics (IBB), Department of Bioinformatics, Laboratory of Complex Biological Systems and Bioinformatics (CBB), University of Tehran, Tehran, Iran
*Corresponding author's email address: kkavousi@ut.ac.ir

Abstract: Parkinson's disease (PD), a prevalent neurodegenerative disorder characterized by striatal dopamine deficiency, poses challenges in early and accurate diagnosis, often leading to misclassification with atypical parkinsonism. While conventional imaging techniques like PET and SPECT detect changes in striatal dopamine levels and functional changes and correlate them with motor responses, they necessitate expert interpretation. This study addresses these diagnostic complexities by leveraging machine learning (ML) and convolutional neural networks (CNN), to automate and expedite PD accurate diagnostics. Our research focuses on analyzing 400 PET images from visits at months 12, 36, and 48 from the Parkinson's Progression Markers Initiative (PPMI) dataset. The goal is to identify similarities between brain images of PD patients for enhanced diagnostic accuracy. In the first step, we resize our 3D images (128x128x128) to 2D images (128x128) by applying average pooling to reduce complexity. Subsequently, we further resize these images to 2D (64x64). Following the preprocessing steps, we employ a convolutional autoencoder to extract similar features, thereby enhancing the clustering operation. After clustering using MeanShift, we analyze each cluster for every pair of original images. For each pair, we calculate the correlation matrix and average all correlation matrices to obtain a single correlation matrix. In the last step, we scrutinize the final correlation matrix to identify cells with values closer to one. We map these cells to the corresponding pixels, and by aggregating high-correlation pixels, we identify the regions that are repeated in our images. Employing advanced AI algorithms, our methodology trains models to recognize recurring similar regions across images. This enables the detection and computation of specific brain regions, potentially indicative of PD. Automated and rapid diagnosis by integrating ML and CNN into PET imaging analysis has the potential to enhance personalized treatment strategies and assist doctors in facilitating the diagnostic process.

Keywords: Convolutional neural networks; PET; Machine learning; Parkinson's disease; Correlation matrix; Mean-shift

ICB12-1402-1334

Identification and Evaluation of Neoantigens in Melanoma Patients Using Data Fusion

E.Erfani[1], K.Kavousi, PhD[2], M.Mazlomi, MD, PhD[3], B.Negahdari MD, PhD[3] , G.A.Kardar, PhD[1]
[1] Immunology Asthma and Allergy Research Institute, Tehran University of Medical Sciences, Tehran, Iran
[2] Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
[3] Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Tehran University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: gakardar@tums.ac.ir

Abstract: Melanoma is a deadly type of skin cancer, and conventional therapies like chemotherapy have many side effects and are often ineffective due to melanoma cell resistance. Neoantigens, specific peptides derived from tumor mutations, offer hope for cancer treatment, but predicting neoantigens is challenging. Machine learning algorithms are recommended to accurately predict neoantigens. Therefore, the aim of this study is to create a new model based on data fusion on different immunogenic features to predict melanoma cancer neoantigens and test the predicted neoantigens in vitro.
Methods consist of three phases. Phase 1: Construction of machine model in three stages: a. Training: Using experimentally confirmed peptides. b. Validation. c. Test: by mutated melanoma s peptides. Phase 2: Identification of genetic changes in melanoma (WES and RNA-Seq from TCGA database). Phase 3: in vitro evaluation of candidate peptides.
The model was constructed and fitted on 263 features. Different tools have been used to evaluate and find peptides features. The netCTLpan was used to find the features related to the binding affinity of MHC and peptides (like MHC_prediction, TAP prediction score, Cleavage prediction score, percentile rank, netMHCstab). physicochemical features of peptides gained by Peptide master package in R (PI, MW, aIndex, Boman, Charge, Hydrophobicity, instaIndex and 9 Features related to composition of amino acids). Also, we obtained the features related to the peptide sequence using the iLearnPlus and p feature packages. Different classifiers such as Logistic Regression, Random Forest, Gaussian NB, XGBoost, Support vector machine, Linear SVC, KNeighbors, SGD, Gradient Boosting, Extra Tree Classifier, Decision Tree, MLP Classifier were evaluated and tested in this machine. Finally, on nine classifiers, weighting and voting methods of data fusion applied. The accuracy of this machine was 65 percent.
Keywords: Data Fusion; Machine Learning; Melanoma; Neoantigens; Immunotherapy

References

[1] Sahin U, Türeci Ö. "Personalized vaccines for cancer immunotherapy". Science. Mar (2018) ; 359 (6382) : 1355–60.
[2] Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. "An immunogenic personal neoantigen vaccine for patients with melanoma". Nature. July (2017); 547 (7662) : 217–21.
[3] Schumacher TN, Schreiber RD. "Neoantigens in cancer immunotherapy". Science. Apr (2015) ; 348 (6230):69–74.Sahin U, Türeci Ö. Personalized vaccines for cancer immunotherapy. Science. 2018

ICB12-1402-1335

Computational Modeling for Predicting and Designing Melanoma Neoantigens

A. Bakhshian[a], M.A. Mazlomi[a]*, K. Kavousi[b], G.A. Kardar[a,c]
a Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Tehran University of Medical Sciences (TUMS), Tehran, Iran
b Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
c Immunology, Asthma and Allergy Research Institute (IAARI), Tehran University of Medical Sciences, Tehran, Iran
*Corresponding author's email address: ma.mazlomi@gmail.com

Abstract: Melanoma is one of the deadliest skin cancers and its global death rate is increasing. Among the different treatment methods, immunotherapy plays a significant role in the field of cancer treatment. In recent years, with the discovery of neoantigens, specific peptides to cancer cells, hope for cancer treatment has increased. Accurate prediction of neoantigens is one of the most challenging issues in immunoinformatics due to the high diversity, multi-step biochemical processes and the random nature of the T cell immune response, and no efficient algorithm has been provided for this until now. On the other hand, the complexity of the computational processes and the need for experts in this field to accurately analyze the analysis, the use of a machine learning algorithm is suggested. We collected a dataset by retrieving experimentally validated epitopes from The Immune Epitope Database (IEDB). Nonameric HLA-A*02:01 peptides were selected. After deletion of duplicates and common peptides between positive and negative, we reached a unique list of epitopes (positive = 1,672; negative = 1,792). We quantified some features for peptides. Final dataset with 263 features was labeled as 1(immunogene) and 0 (non immunogen). Our model is a Support Vector Machine (SVM) trained on data from T cell assays of antigen-specific immune response. For a test set area under the receiver operating characteristic curve (AUROC) is 0.80. This indicates a reasonably strong ability to discriminate between different classes in the dataset. While further optimizations and fine-tuning could potentially enhance predictive capabilities, the achieved AUROC of 80% signifies a promising foundation for the model's efficacy in the given context.

Keywords: Neoantigen; Machine learning; Immunotherapy

References

[1] P. A. Ott, Z. Hu, "An immunogenic personal neoantigen vaccine for patients with melanoma," Nature., 547.7662 (2017): 217-221.

[2] U. Sahin, E. Derhovanessian, "Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer," Nature., 547.7662 (2017): 222-226.

ICB12-1402-1336

Theoretical modelling of surface plasmon resonance biosensor employing graphene and SnSe heterostructure for Pseudomonas-like bacteria detection

Elham Jalili Tazekande Janze[a], Bahar Meshginqalam[b]. Sholeh Alaei[a,*]
[a] Department of Physics, Urmia Branch, Islamic Azad University, Urmia, Iran
[b] Faculty of Physics, University of Tabriz, Tabriz, Iran
*Corresponding author's email address: sh.alaei@iau.ac.ir

Abstract: In this study, using graphene with the addition layer coating of α-SnSe as a two-dimensional (2D) material, a novel SPR biosensor is proposed computationally for Pseudomonas-like bacterial detection. The performance parameters like sensitivity and detection accuracy are calculated using the Frensel equation and transfer matrix method.

For investigating the presence effect of α-SnSe monolayer on the sensitivity improvement, the performance of conventional structure consists of BK7 prism, silver (Ag) thin film, graphene monolayer, three affinity layers (toluene, poly (trifluoroethyl methacrylate) and nicotine) and sensing medium has been compared with the proposed structure, BK7/ Ag/ graphene/ α-SnSe monolayer/ affinity and sensing medium. The thickness of metal layer Ag and graphene are 50 nm and 0.34 nm, respectively [1]. Moreover, the SnSe is employed with an optimized thickness of 0.575 nm and affinity layers with thickness of 3 nm [2,3]. According to the results, the performance advancement of SPR biosensor is observed by covering the graphene with α-SnSe single layer. The biosensor yields the maximum sensitivity value of 165.85 deg/RIU and detection accuracy equal to 1.82 deg^{-1} while the refractive index (RI) of sensing medium varies between 1.33 to 1.40 and for the affinity layer,toluene, with RI of 1.49368. Considering the performance of the proposed structure, we believe that the inclusion of graphene with the α-SnSe monolayer can find potential use in Pseudomonas and Pseudomonas-like bacteria detection for medical and biological detection applications.

Keywords: SPR Biosensor; Pseudomonas bacteria; Modeling; Graphene; SnSe

References

[1] A. Verma, A. Prakash, R. Tripathi, "Performance analysis of graphene based surface plasmon resonance biosensors for detection of pseudomonas-like bacteria", *Optical and Quantum Electronics*, *47* (2015) 1197-1205.

[2] Z. Y. Hu, K. Y. Li, Y. Lu, Y. Huang, X. H. Shao, "High thermoelectric performances of monolayer SnSe allotropes", *Nanoscale*, *9*(41) (2017) 16093-16100.

[3] A. S. Kushwaha, A. Kumar, R. Kumar, M. Srivastava, S. K. Srivastava, "Zinc oxide, gold and graphene-based surface plasmon resonance (SPR) biosensor for detection of pseudomonas like bacteria: A comparative study" *Optik*, *172* (2018) 697-707.

ICB12-1402-1337

Investigating the Prediction of the Pathogenicity of p. Pro146Arg Missense Mutation (rs17718883) in the *PD-L1* Gene with a Bioinformatics Aproach

E. Jafari Nejad Bidgoli, M. M. Heidari*, M.Khatami
a Department of Biology, Yazd University, Yazd, Iran
b Department of Biology, Yazd University, Yazd, Iran
c Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: heidarimm@yazd.ac.ir

Abstract: *CD274* gene is also known as *PD-L1*. This gene is encoded on immune inhibitory receptor ligand that is expressed by hematopoietic and non-hematopoietic cells, such as Tells and B cells, and is overexpressed on the surface of various tumor cells. Interaction of this ligand with its receptor inhibits T cell activation and cytokine production. During infection or inflammation of normal tissue, this interaction is important for preventing autoimmunity by maintaining homeostasis of the immune response. In tumor microenvironments, this interaction provides an immune escape for tumor cells through cytotoxic T-cell interaction. Expression of this gene in tumor cells is considered to be prognostic in many types of human malignancies, such as small-cell lung cancer (SCLC). SCLC is an extremely aggressive subtype of cancer. The disease is characterized by rapid growth and early metastasis to distant organs. In this study, we have assessed a missense variation c.437C>G (p. Pro146Arg; rs17718883) in the *CD274* gene. We evaluated this variation in prediction SNP servers and our results show pathogenicity score in MAPP is 86%, PhD-SNP is 55%, PolyPhen-1 is 59%, PolyPhen-2 is 45%, SIFT is 79%, SNAP is 62% and PANTHER is 56%. We also examined this variation in meta-SNP the results are as follows PANTHER: 0.534, PhD-SNP: 0.412, SIFT: 0.100, SNAP: 0.575, and Meta-SNP: 0.448. The hydrophobicity of protein measured by PEPTID-2 changed from 41.03% to 40.69% with this nucleotide change. Mutation taster also reported that this variation is disease-causing. According to this in-silico study and prediction results, the rs17718883 mutation might have pathogenic effects on the CD274 protein. This theory should be tested with experimental studies to confirm pathogenic effects.

Keywords: PD-L1 Gene; CD274 Gene; Inhibitory Receptor Ligand; Tumor Cell; SCLC

References

[1] J. George, M Saito, "Genomic Amplification of CD274 (PD-L1) in Small-Cell Lung Cancer" 2017 Mar 1; 23(5):1220-1226.

[2] B. Wen, L. Zhao, "Nanobodies targeting the interaction interface of programmed death receptor 1 (PD-1)/PD-1 ligand 1 (PD-1/PD-L1)", 2020; 50(3):252-259.

[3] Y. Han, D. Liu, PD-1/PD-L1 pathway: current researches in cancer 2020; 10(3): 727–742.

ICB12-1402-1338

In-Silico pathogenicity prediction of rs886045862 and rs144788254 in the LAMB3 gene in Junctional epidermolysis bullosa (JEB)

Fatemeh Rezaei
Department of Biology, Yazd University, Yazd, Iran
*Corresponding author's email address: ieazer08@gmail.com

Abstract: Junctional epidermolysis bullosa (JEB) is a major form of epidermolysis bullosa, a group of genetic conditions that cause the skin to be very fragile and to blister easily [1]. LAMB3 gene mutations are the most common, causing about 70 percent of all cases of junctional epidermolysis bullosa [1]. The product encoded by this gene is a laminin that belongs to a family of basement membrane proteins [2]. This protein is a beta subunit laminin, which together with an alpha and a gamma subunit, forms laminin-5 [2]. Consequently, this study was undertaken to find the pathogenic SNPs in the LAMB3 gene. Among all SNPs, we studied two SNPs, rs886045862 and rs144788254: rs886045862 affects protein function with a score of 0.02 by the SIFT and is predicted to be possibly damaging with a score of 0.918 by the PolyPhen2. This variant's MetaRNN score is 0.22674635 by the Hope and it is more likely to be pathogenic. Also, thers144788254 affects protein function with a score of 0.01 by the SIFT, and this mutation is predicted to be probably damaging with a score of 1.000 by the PolyPhen-2. His variant's MetaRNN score is 0.9777092 by the Hope, and it is more likely to be pathogenic. These residues are located on the surface of the protein; mutation of these residues can disturb interactions with other molecules or other parts of the protein. In conclusion, this study suggested that G1003S and G287S missense variants of the LAMB3 gene would affect the protein function.

Keywords: Junctional epidermolysis bullosa (JEB); LAMB3 gene; SNP; PolyPhen-2; SIFT

ICB12-1402-1341

Deciphering the molecular interaction between lysozyme and two commonly utilized pesticides: a comprehensive in-silico and spectroscopic study

Sanaz Asgharzadeh [1], Behzad Shareghi *, Sadegh Farhadian*
sanaz_msh72@yahoo.com
*Corresponding authors email addresses: b_shareghi@yahoo.com; sadeghfarhadian@gmail.com

Abstract: Recently, the widespread application of pesticides such as chlorpyrifos (CPY) and Amitraz (AMZ) in agriculture has been linked to harmful impacts on biomolecules, presenting considerable hazards to both environmental health and public well-being. In this regard, our study aimed to explore the potential binding interaction of lysozyme (LSZ) with CPY and AMZ through molecular modeling and different spectroscopic methods. Through computational docking, specific and highly favored binding sites for CPY and AMZ within the structure of lysozyme (LSZ) were identified. Molecular dynamic simulations showed that the complexes formed between CPY and AMZ with LSZ remained stable over time, and the binding process led to structural changes in the enzyme. The UV-vis absorption and fluorescence experiments showed the complex formation and static quenching of the intrinsic fluorescence intensities. According to circular dichroism and Fourier transform infrared investigations, binding interactions changed the secondary structure of LSZ through increasing α-helix presence and decreasing the levels of β-sheet and β-turn content. On the other hand, complexation with CPY and AMZ decreased the stability and biological activity of LSZ. These results provide critical insights for future studies focused on understanding the possible dangers that pesticides could present to human health.

Keywords: Lysozyme, Chlorpyrifos, Amitraz, Docking, MD simulation, Spectroscopy

ICB12-1402-1342

CDK1 as a Hub Gene: Promising Biomarker for Bladder Cancer Diagnosis

M.Afshari Kave*
Royan Institute, Tehran, Iran
Department of Basic Science, Science and Culture University, Tehran, Iran
*Corresponding author's email address: mobinaafshari@gmail.com

Abstract: Bladder cancer is the tenth most prevalent cancer globally and is characterized by distinct molecular profiles at different stages [1]. Early diagnosis is crucial for effective treatment selection, recurrence prevention, and identification of specific disease subtypes. Biomarkers provide essential information about cancer stage, recurrence risk, and disease etiopathology. This study aims to identify biomarkers and a hub gene for bladder cancer. The study utilized expression profiling by an array (GSE7476) from the GEO Dataset, which included 12 samples of bladder tissue categorized into normal and cancer groups. Differentially expressed genes (DEGs) were identified using GEO2R, resulting in the identification of 145 upregulated genes with adjusted P-values<0.05 and logFC≥2. Gene Ontology (GO) analysis and KEGG pathway analysis were performed, and PPI networks were established using STRING and Cytoscape. The GO analysis revealed that the DEGs were primarily enriched in biological processes such as cell division and the apoptotic process. These DEGs were predominantly located in the nucleus, cytoplasm, and cytosol, and were found to be involved in protein binding at the molecular function level. Pathway enrichment analysis indicated that the DEGs were significantly associated with cell cycle and oocyte meiosis. Additionally, hub genes such as CDK1, TYMS, CCNB1, TOP2A, and CCNB2 were identified. CDK1, in particular, was highlighted as a potential biomarker due to its crucial role in regulating the G2/M phase transition in the eukaryotic cell cycle and its overexpression in various cancers, including bladder cancer and lung cancer [2]. Given its presence in urine and plasma, CDK1 shows promise as a potential non-invasive biomarker for these cancers, offering an alternative to the invasive diagnostic method of Cystoscopy.

Keywords: bladder cancer; CDK1; biomarker; systems biology; bioinformatics

References

[1]Dyrskjøt, L., et al., Bladder cancer. Nature Reviews Disease Primers, 2023. 9(1): p. 58.

[2] Li, M., et al., CDK1 serves as a potential prognostic biomarker and target for lung cancer. Journal of International Medical Research, 2020. 48(2): p. 0300060519897508.

ICB12-1402-1343

Design of guide RNAs for genome editing of *Yarrowia lipolytica* using deep learning

Z. Vahdani [a], F. Darvishi [b,*], F. Zare-Mirakabad [c]
[a] Department of Computer Science, Faculty of Mathematical Sciences, Alzahra University, Tehran, Iran
[b] Department of Microbiology, Faculty of Biological Sciences, Alzahra University, Tehran, Iran
[c] Department of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran
*Corresponding author's email address: f.darvishi@alzahra.ac.ir

Abstract: The yeast Yarrowia lipolytica is capable of producing important natural and recombinant products with industrial and therapeutic commercial value. The CRISPR/Cas system is used for genome editing of this yeast to achieve effective metabolic engineering and increase the production efficiency of valuable products. In this system, a protein called Cas combines with a short RNA called guide RNA (sgRNA) and makes a double-strand break precisely at the desired location in the genome. If the sgRNA is not designed with high precision and without off-target, unexpected off-target edits can occur and potentially cause harmful effects. Hence, the aim of the current study is to design precise and less off-target sgRNA by computational and deep learning [1]. In the new computational model, with the help of new functions and articles, chemical and physical features were extracted for each sequence. Influential features were identified using an encoder and decoder, and a model was obtained using CNN. With the help of cutting score (CS) that Baisya et al. [2] obtained in the laboratory for each sgRNA and using computational learning models. They obtained Spearman value of 0.37% and Pearson value of 0.43%. In the current study, Spearman value and Pearson value reached to 0.44% and 0.52%, respectively. Furthermore, a library was designed to produce sgRNA sequences for the Cas9 protein using the alignment method for the genome of each organism so that the effectiveness of the sequences could be evaluated with the sequence obtained in the laboratory. A computational model was also presented to obtain the epigenetic value for each sgRNA with the least off-target. The findings can lead to a better and more accurate design of sgRNA with the least off-target which will greatly help the genome engineering of yeasts and other organisms.

Keywords: Guide RNA; Machine learning; Deep learning; Yeast; Genome editing.

References

[1] F. Darvishi, "Synthetic Biology of Yeasts: Tools and Applications", Springer, 2022.
[2] D. Baisya, A. Ramesh, C. Schwartz, S. Lonardi, I. Wheeldon, "Genome-wide functional screens enable the prediction of high activity crispr-cas9 and-cas12a guides in *Yarrowia lipolytica*", Nat. Commun., 13 (2022) 922.

ICB12-1402-1344

Evaluating the Induction of Apoptosis via Caspase 3 Activation: A Comparative Study of Mefenamic Acid and Aspirin

Melika Haghdadian[a], Mahsa Soufian[a], Fateme Sheykh Mohamadi Fazli[a], Shiva Moosavi[a] ,Reza Sahebi[a]*
[a]Department of Cellular and Molecular Biology, Kavian Institute, Mashhad, Iran
Corresponding author's email address Reza_Sahebi2001@hotmail.com

Abstract: This study investigates the therapeutic effects of Mefenamic acid and Aspirin, specifically their anti- inflammatory, anti-pyretic properties and their role in reducing stomach and colon cancer [1]. Previous research indicates that these drugs induce apoptosis by positively expressing Caspase3[3]. Caspase 3, in turn, cleaves and activates sterol regulatory element binding proteins (SREBPs), inducing apoptosis [2]. The activation of Caspase 3 can occur in three ways, either dependent or independent of cytochrome C release from mitochondria and Caspase 3 activation [2]. The objective of this study is to compare the binding affinity of Mefenamic acid and Aspirin to Caspase 3, which is associated with apoptosis in stomach and colon cancer [3].
Target proteins for Mefenamic acid and Aspirin were identified using PubChem, Drugbank (target section), and target.net.scbdd.com site by entering the SMILES of each drug [3]. Protein- ligand docking was performed using the ChemSpider database (properties > predict-Mcule) and the results were analyzed using Pymol software version 2.5.5 [3].
The binding of Caspase 3 with two FDA-approved anti-inflammatory drugs was investigated. Docking results showed binding affinities of -5 for Aspirin and -6 for Mefenamic acid to Caspase 3 [3][1].
The docking score analysis revealed that Mefenamic acid (-6) has a higher binding affinity to Caspase 3 than Aspirin (-5) [3]. Therefore, Mefenamic acid may have a greater impact on the production and regulation of Caspase3 expression, thus potentially enhancing the induction of apoptosis in gastric and intestinal cancers [1] [2] [3].
Keywords: Mefenamic acid; Aspirin; Caspase3; Apoptosis; Molecular Docking

References

[1] F. Woo DH, Han IS, Jung G. Mefenamic acid-induced apoptosis in human liver cancer cell-lines through caspase-3 pathway. Life sciences. 2004 Oct 1;75(20):2439-49.

[2] F. Porter AG, Jänicke RU. Emerging roles of caspase-3 in apoptosis. Cell death & differentiation. 1999 Feb;6(2):99-104.

[3] Savjani JK, Mulamkattil S, Variya B, Patel S. Molecular docking, synthesis and biological screening of mefenamic acid derivatives as anti-inflammatory agents. European Journal of Pharmacology. 2017 Apr 15; 801:28-34.

Beihaghi M, Tehrani H, Akbari N, Beihaghi MR, Sahebi R. In Silico Analyzes for the Inhibition of HIV Protease by Ritonavir and Indinavir. Research in Biotechnology and Environmental Science. 2022 Sep 1;1(1):16-22.