The 2nd International and 11th National

IRANIAN CONFERENCE ON BIOINFORMATICS

2023

28 Feb-1 Mar 2023

Institute for research in fundamental sciences, Tehran

2nd International Iranian Conference on Bioinformatics

February 28 - March 1, 2023

یازدهمین همایش ملی
بیوانفورماتیک ایران

۱۰-۹ اسفند ۱۴۰۱

Main Topics

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

محورهای همایش

- آنالیز توالی‌های زیستی
- بیوانفورماتیک ساختاری
- زیست‌شناسی سامانه‌ای
- هوش مصنوعی و یادگیری ماشین
- طراحی و کشف محاسباتی دارو
- مدل سازی در زیست شناسی محاسباتی

Abstract submission deadline
February 4, 2023

آخرین مهلت ارسال چکیده مقالات
۱۵ بهمن ۱۴۰۱

icb11.ibis.org.ir
icb11@ibis.org.ir
IranianBioinformaticsSociety
School of Biological Sciences, Institute for Research in Fundamental Sciences, Shahid Lavasani St., Tehran, Iran.

In living memory of

Dr.Abbas Nowzari-Dalini

It is with great sadness that we must inform you of the passing of Professor Abbas Nowzari-Dalini. A faculty member of the computer science department at Tehran University, Professor Nowzari's life was devoted to education, and so his memory shall live forever in the hearts of his pupils. To his family and friends, we extend our heartfelt sympathies in this time of loss.

Organizers

Computational Biology Research Center

Iranian Bioinformatics Society

INSTITUTE FOR RESEARCH IN FUNDAMENTAL SCIENCES

Indexed by

ISC Islamic World Citation Center

CIVILICA

WELCOME MESSAGES

Dr. Fatemeh Zare-Mirakabad
Amir Kabir University of Technology

Greetings
It is an honor to announce that in cooperation with the Iranian Bioinformatics Society (IBIS) and with the Institute for Research in Fundamental Sciences (IPM) we have decided to hold the 11th National Conference and the 2nd International Conference on Bioinformatics of Iran (ICB) from the 28th of February to the 3rd of March 2023. IPM will host this series of conferences with the aim of exploring the latest topics in bioinformatics, presenting the latest scientific achievements, and creating an opportunity for students, researchers and professors to exchange ideas. In 2023, ICB will cover a broad range of topics in bioinformatics, including analysis of biological sequences, structural bioinformatics, systems biology, computational drug design, image processing in bioinformatics, big data in bioinformatics, mathematical modeling in biology, machine learning and artificial intelligence.
We hope that holding this event, in addition to promoting bioinformatics knowledge in Iran and showcasing our scientific efforts and achievements to the world, can also provide the opportunity to get acquainted with professors and researchers abroad and lead to international cooperation for the participants.

WELCOME MESSAGES

We also intend to hold training events in various areas such as drug design, teaching Python programming for bioinformatics, etc. to cultivate useful skills for the attendees.
As the scientific secretary of the 2nd International and 11th National Iranian Conference on Bioinformatics (ICB11), I invite all professors, students, researchers, and anyone interested in the field to attend and present their latest research findings. May we be able to achieve these goals with your help and bring acclaim to our beloved country in bioinformatics.

Sincerely,

Fatemeh Zare-Mirakabad

Scientific Secretary of the 2nd International and 11th National Iranian Conference on Bioinformatics

WELCOME MESSAGES

Dr. Najmeh Salehi

Institute for Research in Fundamental Sciences

Greetings
I am delighted to announce that, in collaboration with the Institute for Research in Fundamental Sciences (IPM), the Iranian Bioinformatics Society (IBIS) will hold the 11th national and the 2nd international Iranian Conference on Bioinformatics (ICB11) both online and in-person from the 28th of February to the 1th of March. This conference will draw on the 15 years' experience of the Bioinformatics Society, holding a dozen national and international events, as well as the successful track record of the IPM in promoting scholarly discourse throughout the country to achieve high scientific credibility.

WELCOME MESSAGES

Bioinformatics as an interdisciplinary science, as a result of the ever-increasing quantity of biological data, and the need to store, retrieve, and properly analyze such data, is being studied and utilized at an increasing pace. We aim to bring together researchers in the fields of computer science, biology, mathematics, genetics, chemistry, physics, and other relevant disciplines to promote dialogue and foster new perspectives for research and education as well as to publish the latest developments in bioinformatics. It is planned to invite prominent domestic and international scientists and researchers as key speakers, and to hold specialized workshops, think tanks, and panel discussions throughout this conference. Therefore, all faculty members, scientists, researchers, and students interested in the field of bioinformatics are hereby invited to attend and are encouraged to present their research. It is our hope that this conference will provide a venue where researchers from domestic as well as international institutions engaged in bioinformatics research will be able to share information, exchange ideas, and synergize their efforts effectively.

Best Regards,

Najmeh Salehi

Executive Secretary of the 2nd International and 11th National Iranian Conference on Bioinformatics

TABLE OF CONTENT

Plan Table

Detailed Schedule

28 Febraury 2023

7:45-8:15	**Welcom Receotion**	
8:15-9:00	**Conference opening Ceremony**	
Time	**Speaker**	**Title**
9:00-9:45	**Kaveh Kavousi**	**In-Silico Functional Microbiomics: From Medicine to Industry**
9:45-10:05	Masoud Arabfard	Candidate prognostic gene signatures in lung cancer risk among asthmatic and COPD patients
10:10-10:30	Nazanin Hosseinkhan	Comparative gene co-expression network analysis between adenocarcinoma and squamous cell carcinoma subtypes of esophageal cancer

Plan Table

Detailed Schedule

28 Febraury 2023

Time	Speaker	Title
10:30-11:00	**Break and Poster presentations**	
11:00-11:45	**Fatemeh Vafaee**	**Big data and Artificial Intelligence: driving personalized medicine of the future**
11:45-12:05	Farideh Bahari	A novel approach for identifying cancer drivers in coding and non-coding genomic elements
12:10-12:30	Fatemeh Ahmadi	Characterization of chemotherapy resistant triple-negative breast cancers at single-cell resolution
12:30-13:30	**Lunch and Poster presentation**	

Plan Table

Detailed Schedule

28 Febraury 2023

Time	Speaker	Title
13:30-14:15	**Burkhard Morgenstern**	**Alignment-free sequence comparison: the SpaM approaches**
14:15-14:35	Zahra Seraj	Antimicrobial resistance prediction in Acinetobacter baumannii using collaborative matrix factorization
14:35-14:55	Soheila Shabani	A bioinformatics approach to analysis UPR-related gens impact on AML
15:00-15:30	**Break and Poster presentations**	
15:30-16:15	Mohieddin Jafari	Network biology: An article's decorative element or an attampt to meet a need?

Plan Table

Detailed Schedule

28 Febraury 2023

Time	Speaker	Title
16:15-17:15	**Panel: Network in Biology** Mehdi Sadeghi, Mohieddin Jafari, Mohammad hossein Karimi-Jafari, Ali Kamandi	
17:30-18:15	**Albert Laslo Barabasi**	**Network Medicine: From Cellular Networks to the Human Diseasome**

Plan Table

Detailed Schedule

1 March 2023

Time	Speaker	Title
8:15-9:00	**Bernhard Palsson**	**Novel knowledge-enriched data analytics for transcriptomes**
9:00-9:20	Sepideh Mofidfar	Reconstruction of genome-scale metabolic model for Helicobacter pylori
9:20-9:40	Kasra Mokhtarzadeh Azar	Transcriptome-wide analysis of DNA-RNA hybrids reveals how non-coding RNAs contribute to ASDs
9:40-10:00	Sara Kerachian	Identification of Transcription factors involved in cisplatin drug resistance in ovarian cancer by integrated analysis of transcriptome and epigenomic data

Plan Table

Detailed Schedule

1 March 2023

Time	Speaker	Title
10:00-10:30	**Break and Poster presentations**	
10:30-11:15	**Ali Sharifi-Zarchi**	**Cancer Detection: How Bioinformatics & AI Can Help to Process Genomic Data?**
11:15-11:35	Mahsa Saadat	Improvement of peptide-HLA class I prediction using transformers
11:35-11:55	Edris HosseiniGol	Single-cell RNA sequencing data analysis using Explainable Artificial Intelligence identified key transcriptional factors for early COVID-19 severity prediction
11:55-12:20	Zahra Tavakol Hamedani	Feature set prioritization of metagenomes for predicting colorectal cancer based on multiple human gut microbiome datasets

Plan Table

Detailed Schedule

1 March 2023

Time	Speaker	Title
12:30-13:30	**Lunch and Poster presentations**	
13:30-13:50	Fatemeh Farrokhi	Molecular design of sgRNA based on CRISPR-Cas13a for specific detection of ORF1ab region in SARS-COV-2 genome infecting Iranian patients
13:50-14:10	Donya Afshar Jahanshahi	A Plastic-Contaminated Soil Gene Catalog Constructed by Metagenomic Approaches
14:15-15:00	**Ivan G. Costa**	**Computational medical genomics at the single cell sequencing age**
15:00-15:30	**Break and Poster presentations**	

Plan Table

Detailed Schedule

1 March 2023

Time	Speaker	Title
15:30-16:15	**Changiz Eslahchi**	**SIMPLIFY: Similarity-based method to predict drug synergies score using machine learning by integration of multiple cancer line's features**
16:15-16:35	Dr. Jamshid Pirgazi	Improving prediction of drug-target interactions based on fusing multiple features with data balancing and feature selection techniques
16:35-16:55	Parvin Mansouri	Prediction of protein druggability by supervised learning approaches
17:00-18:00	**Conference closing ceremony**	

KEYNOTE SPEAKER

Brief Research Description

Given the functional interdependencies between the molecular components in a human cell, a disease is rarely a consequence of an abnormality in a single gene, but reflects the perturbations of the complex intracellular network. The emerging tools of network medicine offer a platform to explore systematically not only the molecular complexity of a particular disease, leading to the identification of disease modules and pathways, but also the molecular relationships between apparently distinct (patho) phenotypes. Advances in this direction are essential to identify new disease genes, to uncover the biological significance of disease-associated mutations identified by genome-wide association studies and full genome sequencing, and to identify drug targets and biomarkers for complex diseases.

Albert-László Barabási Ph.D.

Affiliations

- Professor at Northeastern University and visiting professor of Central European University
- Ph.D., 1994, Boston University, in physics; advisor H.E. Stanley
- M.Sc., 1991, Eötvös Loránd University, Budapest, in physics; advisor T. Vicsek
- 1986-1989,University of Bucharest, major in physics and engineering

Main Research Areas

- **Network science**
- **Statistical physics**
- **Biological physics**
- **Medicine**

Awards and Honors

- The FEBS Anniversary Prize for Systems Biology-2005
- Jon von Neumann Medal for outstanding achievements in computer-related science and technology-2006
- C&C Prize from the NEC C&C Foundation-2008
- Cozzarelli Award from the US National academies of Science-2009
- Lagrange Prize-CRT Foundation for his contributions to complex systems-2011
- Prima Primissima Award for his contributions to network science by the Hungarian Association of Entrepreneurs and Employers-2014
- Bolyai Prize- 2019
- Elected fellow of AAAS a Fellow of the Massachusetts Academy of Sciences

KEYNOTE SPEAKER

Bernhard Palsson Ph.D.

Affiliations

- Professor of Bioengineering and an adjunct professor of Medicine at the University of California, San Diego
- The Principal Investigator of the Systems Biology Research Group in the Department of Bioengineering
- 1984-1995: Chemical Engineering University of Michigan, Ann Arbor, Ann Arbor, MI
- 1995- : Bioengineering University of California, San Diego, La Jolla, CA
- 1984: received Ph.D. in Chemical Engineering from the University of Wisconsin
- 1989: held a faculty position at the University of Michigan for 11 years and was named the G.G. Brown Associate Professor at Michigan
- 1995: a Fulbright fellow,
- 1996: an Ib Henriksen Fellow
- Member of the National Academy of Engineering
- Fellow of the AIChE, AIMBE, AAAS, and the AAM

Main Research Areas

- **The development of computational biology methods (flux-balance analysis, and modal analysis)**
- **Genome-scale models (M models, ME models)**
- **Data analytic methods (iModulons, pangenomics, alleleome, structural proteomics)**
- **The formulation of specific sysbio models of the red blood cell, E. coli, CHO cells, and many human pathogens**

Awards and Honors

- Co-authored more than 600 peer-reviewed research articles and has authored four textbooks, with more in preparation
- Inventor on over 40 U.S. patents
- co-founder of several biotechnology companies
- Several major biotechnology awards
- Clarivate Highly Cited Researcher since 2014.

KEYNOTE SPEAKER

Brief Research Description

Big data has become a ubiquitous watchword of biomedical innovation advocating the deployment of advanced data-driven artificial intelligence techniques and systems thinking to revolutionise biomedical research and practice. AI-empowered Biomedicin Laboratory led by Associate Professor Fatemeh Vafaee develops cutting-edge innovative machine-learning and network science methodologies to leverage large-scale molecular and clinical data to find hidden structures within them, account for complex interactions among the measurements, integrate heterogeneous data and make accurate predictions in different biomedical applications. In this talk, A/Prof Vafaee provides some examples of ongoing projects across three main themes in her research program, including 1) minimally-invasive biomarker discovery for personalised medicine, 2) single-cell sequencing data analysis and integration, and 3) computational drug repositioning and network pharmacology. Across all themes, Dr Vafaee's research heavily relies on multidisciplinary expertise and cross-faculty collaborations to generate translatable outcomes impacting upon biomedicine of the future.

Affiliations

- Associate Professor, University of New South Wales
- Deputy Director, Data Science Centre, University of New South Wales, Sydney, Australia.
- Health Data Science Lead, Data Science Centre, University of New South Wales, Australia.
- Founding CEO, Founder of OmniOmics.AI Pty Ltd (http://omniomics.ai).
- Research Fellow, 2017, University of Sydney, in Computational biomedicine, systems biology, bioinformatics
- Postdoctoral Associate, 2012, University of Toronto, in Computational biology, cancer informatics
- Ph.D., 2011, University of Illinois at Chicago, in Computer Science, Artificial Intelligence
- B.Sc, 2006, Sharif University of Technology, Computer Engineering

Fatemeh Vafaee Ph.D.

Main Research Areas

- **Medical informatics**
- **Computational biology and Bioinformatics**
- **Artificial intelligence**
- **Biomarker discover**
- **Drug discovery**

Awards and Honors

- Women in AI Australia and New Zealand Award-2021
- Georgina Sweet Award for Women in Quantitative Biomedical Science-2020
- Member of National Computational Merit Allocation Committee, NCMAC- 2019-2023
- Member of Women in Research Network Executive Committee, UNSW Sydney-2018-2021

KEYNOTE SPEAKER

Brief Research Description

Sequence alignment is the first step of most methods for DNA or protein sequence analysis. In particular, pairwise and multiple alignment is a fundamental step in phylogenetic tree reconstruction. With the huge amounts of sequence data that are nowadays available, however, alignment methods have become too slow. Therefore, faster alignment-free methods for sequence comparison have been developed in recent years. We propose to use so called "spaced words" and "spaced-word matches" (SpaM) for alignment-free sequence comparison. We show that phylogenetic distances can be accurately estimated using word or spaced-word matches.

Burkhard Morgenstern Ph.D.

Affiliations

- Group leader, Int. Grad. School of Bioinformatics and Genome Research, Univ. Bielefeld
- MIPS, Max-Planck-Institut, Martinsried, and GSF Research Center, Neuherberg
- Postdoc, GSF Research Center, Neuherberg
- Dr. math., Univ. Bielefeld. Thesis work on multiple sequence alignment. Supervisor: Prof. A. Dress
- Diploma in Mathematics, Univ. Munich. Thesis work on partial differential equations. Supervisor: Prof. J. Batt
- Vordiplom (comparable to BSc) in Biology, Univ. Freiburg

Main Research Areas

- **Algorithms for sequence analysis**
- **In particular alignment-free sequence comparison**

Editorial Services

- Founder and Editor of the journal `Algorithms for Molecular Biology' since 2006
- BMC Bioinformatics
- Algorithms for Molecular Biology
- The Open Applied Informatics Journal
- Advances in Bioinformatics
- PeerJ
- NAR Genomics and Bioinformatics
- Mathematical Biosciences and Engineering

KEYNOTE SPEAKER

Brief Research Description

Single cell sequencing are powerful tools to uncover epigenetic and transcriptional changes related with cellular and malignant transformations. However, this data has challenging characteristics such as high dimension, high sparsity, multi modality and low count distributions. We will describe computational methods for tackling problems as dimension reduction, denoising or for learning integrative embeddings from single cell multiomics data. Furthermore, clinical applications represent the next challenge in single-cell genomics, as we are still lacking computational methods to analyse single-cell and pathomics data at a patient level for finding patient trajectories associated with diseases. We here propose a method to perform unsupervised analysis at the sample level and to uncover trajectories associated with disease progression. These were used to dissect differentiation processes associated with fibrosis in kidney diseases and heart myocardial infarction.

Affiliations

- Professor at Institute of Computational Genomics, University Hospital RWTH Aachen
- 2004-2008: Ph.D., Max Planck Institute for Molecular Genetics, in Computational Biology
- 2002-2003: M.Sc. in Computer Science, Centro de Informatica, Universidade Federal de Pernambuco, Brazil; Supervisors: Francisco de A. T. de Carvalho, Marcilio C. P. de Souto
- 1996-2001: Bachelor in Computer Science, Centro de Informatica, Universidade Federal de Pernambuco, Brazil

Ivan G. Costa
Ph.D.

Main Research Areas

- **Computational Biology**
- **Machine Learning**
- **Personalized Medicine**

Professional Positions

- Since 2017: Professor at Institute of Computational Genomics, University Hospital RWTH Aachen
- 2012 to 2017: Group Leader in Computational Biology Interdisciplinary Clinical Research Center Aachen, University Hospital RWTH Aachen Institute for Biomedical Engineering, Helmholtz Institute for Biomedical Engineering, Aachen
- 2012 to 2016: Assistant Professor in Computational Biology (on leave), Center of Informatics, Federal University of Pernambuco, Recife
- 2009-2012: Assistant Professor in Computational Biology at Center of Informatics, Federal University of Pernambuco, Recife

KEYNOTE SPEAKER

Brief Research Description

Network biomedicine is a field of study that focuses on reconstructing, analyzing, and using various biomedical networks to depict the complexity of biomedical systems. The network model predictions, akin to other computational models, are based on our presumptions. In biomedicine, experimental and clinical observations aid in the improvement of the biomedical presumptions that were utilized to develop the models and, in turns, pinpoint anticipated model interpretations. In this talk, I'll go over a few reconstructs of biological networks and how our research team has used them.

Mohieddin Jafari Ph.D.

Affiliations

- Principal Investigator(PI) at University of Helsinki
- Senior Researcher, Helsinki, Southern Finland,2019-2022
- Docent of Bioinformatics, Faculty of Medicine, University of Helsinki, Finland, 2020
- Specialization certificate Systems Biology, Systems Biology Center New York (SBCNY), USA
- Ph.D. Applied Proteomics, Beheshti University of Medical Sciences (SBMU), Iran
- Ph.D. Sabbatical Harvard Proteomics Resource, Harvard School of Public Health (HSPH), USA
- M.Sc. Cell and Molecular Biology, National Institute of Genetics and Biotechnology (NIGEB), Iran
- B.Sc. Cell and Molecular Biology, Tehran University, Iran

Main Research Areas

- **Proteomics**
- **Systems biology**
- **Mass Spectrometry**
- **Bioinformatics**
- **Biological Data Mining**

Professional Positions

- Senior Researcher, Institute for Molecular Medicine Finland (FIMM),2018-2019
- Founder and Editor of the journal `Algorithms for Molecular Biology' since 2006
- Independent Researcher, Pasteur Institute of Iran,Iran, 2014-2018
- Short-term Research Visitor, Instituto Gulbenkian de Ciencia, Oeiras, Portugal, 2017
- • Research Scientist, Institute for Research in Fundamental Sciences(IPM), Iran, 2010-2016
- • Research Scientist, Center for Medicinal Plant and Drug Research Institute, Iran, 2012-2014

KEYNOTE SPEAKER

Brief Research Description

Combination therapies are often required to treat advanced stages of cancer. Predicting drug synergy scores can help identify potential drug combinations to enhance therapeutic benefits while minimizing adverse effects. In this presentation, we will discuss drug synergy in general and present SIMPLIFY, our similarity-based method to predict drug synergy scores using machine learning by integrating multiple cancer cell line features. Our approach involves defining similarity scores based on gene expressions, copy number variations, mutations, and dependency data and training a machine learning model to predict similarity scores between new and existing cell lines. We will discuss our evaluation of SIMPLIFY on the O'Neil and ALMANAC datasets using cross-validation techniques, and demonstrate its superior performance compared to state-of-the-art methods. Our results show that SIMPLIFY has the potential to facilitate the discovery of drug combinations for the treatment of new cancer cell lines.

Changiz Eslahchi Ph.D.

Affiliations

- Professor of algorithm in bioinformatics, Department of Computer Science, Shahid Beheshti University
- Ph.D. in Mathematics, Department of Mathematical Sciences, Sharif University of Technology, Tehran, IRAN,1998. Title of Thesis: Hall condition and list coloring of graphs. Thesis advisor : Professor E. S. Mahmoodian.
- M.S. in Mathematics, Department of Mathematics, Shiraz University, Shiraz, IRAN,1990.
- B.S. in Mathematics, Department of Mathematics, Training Teachers University Of Tehran, IRAN, 1987.
- Head of Department of Cognitive Modeling, Institute for Brain and Cognitive Science, Shahid Beheshti University, from 2014
- Member of the School of Biosciences at the Institute for Research in Fundamental Sciences (IPM), Tehran.
- Member of the Center of Excellence Discrete Structure, Algebra and Logic at Shahid Beheshti University.
- Member of the Bioinformatics Group at the Tehran University
- Chairman of Math. Dept. 2006-2008
- Manager of Computer Center of Shahid Beheshti University, 2002-2004
- Member of Iranian Mathematical Society and Iranian Bioinformatics Society
- Memeber of International Society for Computational Biology
- Visiting Scientist, Department of Mathematics, Reading university, Reading, England, July 1999- September 1999.
- Visiting Scientist, School of Mathematics, University of Science Malaysia (USM),Pinang, Malaysia, July 2010- August 2010.
- Visiting Scientist, School of Computing, National University of Singapore(NUS),Singapore, September 2010.
- Visiting Scientist, School of Computing, National University of Singapore(NUS),Singapore, September 2012.
- Visiting Scientist, School of Computing, National University of Singapore(NUS),Singapore, September 2013.

Main Research Areas

- **Combinatorial algorithms in bioinformatics**
- **Phylogenetic networks construction**
- **System biology**
- **Graph theory.**

Awards and Honors

- Second Prize Winner in Algebraic Competition for Undergraduate University Students in 1986. This Competition is held annually by the Iranian Mathematical Society, and the Participants are the Distinguished Students from the Universities throughout the Country
- First place in the Nationwide Entrance Examination for the MS. degree,1987
- Top Second Place in the Entrance Examination for the Ph.D. degree of the Department of Mathematical Sciences of Sharif University of Technology, 1995

KEYNOTE SPEAKER

Brief Research Description

The environmental microbiome consists of different microbial communities that reflect the characteristics of the surrounding environment. The information that can be extracted from this environment is not limited to the abundance of living microorganisms in this environment. The functional elements of meta-omics at different levels, including gene, transcript, protein and enzyme, metabolite and reaction, and functional ortholog groups, contain much richer information than taxonomic level information. Due to the very long history of microorganisms on the planet and the high rate of mutation in their genome, which has led to a ultra-high diversity, during the process of evolution, microorganisms have been able to adapt to different environments and extreme conditions. The development of proteins and enzymes and other bioactive molecules with extraordinary functionalities has been one of the advantages gained by microorganisms in the underlying game of life. The computational microbiomics science and its sub-branches, by employing statistical and machine learning models makes it possible to investigate the functional characteristics of the microbiome and the interactions of microbial communities by in-silico methods. The proven association between complex diseases with the microbiome of different sites of the human body has opened a wide discipline in microbiome studies applied to the field of medicine and health. On the other hand, the potential of microorganisms in initiating, facilitating and catalysis, and progression of industrial processes has caused the microbiome of different environments to be studied as a valuable resource for the discovery and extraction of bioactive biomolecules. In this lecture, the research work done or being done by the members of the "Complex Biological Systems and Bioinformatics (CBB)" laboratory, in both medical and industrial biotechnology fields that have been going on for the past 10 years will be reviewed. CBB has been established at 2014 in the bioinformatics department, Institute of Biochemistry and Biophysics, University of Tehran.

Main Research Areas

- **machine learning**
- **bioinformatics**
- **complex biological systems**

Kaveh Kavousi

Ph.D.

Professional Positions

- Associate Professor,Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran
- Member of the board of directors and vice president of the Iranian Bioinformatics Society,2019
- Member of the Board of Directors and Treasurer of the Iranian Bioinformatics Society,2016

Affiliations

- Associate Professor,Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran
- Ph.D. 2012, Computer Engineering - Artificial Intelligence and Robotics, University of Tehran
- M.S, 2001, Computer Engineering - Artificial Intelligence and Robotics, University of Tehran
- B.S, 1997, Computer Engineering, Amirkabir University of Technology (Tehran Polytechnic)

KEYNOTE SPEAKER

Brief Research Description

Accurate detection of cancer using genomic data has been a crucial challenge since decades ago. Alternations of gene expression transcription and epigenetics are among cancer hallmarks, which can be detected in high-throughput data such as tumor RNAseq and cell-free DNA (cfDNA) Bisulfite-seq. In this talk we will review how deep neural networks can be employed to identify cancer type. Furthermore, we will introduce a novel algorithmic method for highly accurate alignment of Bisulfite-seq data, which can lead to improved accuracy in early detection of cancer.

Affiliations

- Assistant Prof. of Bioinformatics, Computer Engineering Department, Sharif University of Technology
- Head of Bioinformatics Lab, Royan Institute for Stem Cell Biology and Technology
- Associate of Bioinformatics, Computer Science Department, Colorado State
- Associate of Bioinformatics, Chitsaz Lab, Computer Science Department, Colorado State University, 2014-2016
- Research Assistant, Genomics and Computational Biology Division, Max Planck Institute for Molecular Biomedicine, Muenster, Germany, 2012-2013
- Bioinformatics Researcher, Royan Institute, 2014
- Member of iPS & Epigenetic Reprogramming lab, Department of Stem Cells, Royan Institute, 2011-2014
- Ph.D. of Bioinformatics, Institute of Biophysics & Biochemistry, University of Tehran, 2007-2015
- M.Sc. of Computer Engineering (Software), Sharif University of Technology, Tehran, 2004-2006•B.Sc. of Computer Engineering (Software), Sharif University of Technology, Tehran, 2000-2004
- Advanced Programming, Data Structures and Algorithm, Graph theory, Combinatorics & Backtracking, Summer and Winter Schools of the National Olympiad in Informatics, Young Scholars Club, 1998-2000
- Diploma in Math, National Organization for Development of Exceptional Talents - Shahid Sadoughi, 1996-2000

Ali Sharifi-Zarchi Ph.D.

Main Research Areas

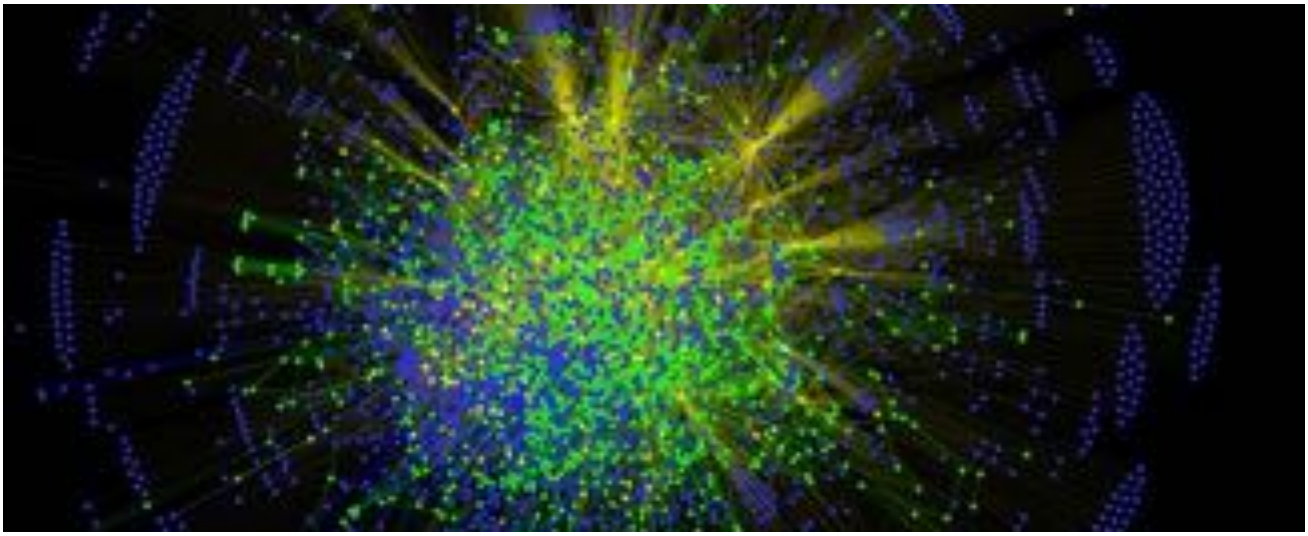
- **Biostatistics**
- **High-Dimensional Data**
- **Longitudinal and Multilevel Data**
- **Computational Biology**
- **Statistical Network Analysis**
- **Neuroimaging**

Awards and honors

- Chair of Host Scientific Committee, International Olympiad in Informatics (IOI), Tehran, 2017
- Elected Member of International Scientific Committee, International Olympiad in Informatics (IOI), 2012-2014
- First place award of the AAAI (Association for the Advancement of Artificial Intelligence) Rescue Robots Competitions, Canada, 2002
- Champion team of the ACM Collegiate Programming Contest in the Southwestern Asia Site, Tehran, 2002
- Member, National Committees of the Iranian national Olympiad in Informatics, 2000-Now
- Gold Medal of the International Olympiad in Informatics (IOI), China, 2000
- 1st place, National Team Selection Contest for participating the International Olympiad in Informatics, 2000
- Gold medal, National Olympiad in Informatics, 1999
- Silver medal, National Olympiad in Informatics, 1998
- 1st place, National Collegiate Scientific Contest, 1996

SYSTEMS BIOLOGY

Topics



Systems biology is a field of research that focuses on understanding whole biological systems, such as protein complexes, metabolic pathways, or gene regulatory networks, and it attempts to understand cells, tissues, and organisms and how they behave and function from a systems perspective. In other words, systems biology is a computational and mathematical analysis and modeling of complex biological systems, and a systemic view of biological issues means that the function of no organ is independent of the other so that the behavior of all components will affect the behavior of the whole system.
Also, systems biology is defined as a set of computational methods whose purpose is to integrate all molecular measurements taken from the elements of a biological system to create a model that describes and predicts the overall behavior of the system. The goal of systems biology is to understand the basics of designing living systems, and conducting extensive research on biological interactions with the aim of generalization is the main indicator in systems biology.

SYSTEMS BIOLOGY

For more information,please refer to the following sources:

Najarian, K., Najarian, S., Gharibzadeh, S., & Eichelberger, C. N. (2009). Systems biology and bioinformatics: a computational approach. CRC Press.

Klipp, E., Herwig, R., Kowald, A., Wierling, C., & Lehrach, H. (2005). Systems biology in practice: concepts, implementation and application. John Wiley & Sons.

According to Westerhoff and Palsson (2004), systems biology has two historical roots. The first and most commonly cited etymology relates to the structure and function of genetic material and the discovery of methods of genetic manipulation. The second route relates to thermodynamic aspects of organisms introduced into biology in the 1940s.
Here we discuss the three evolutionary stages of systems biology. The first step involved transforming molecular biology into systems molecular biology. This phase was related to the discovery of gene structure and function and genetic engineering (Westerhoff and Palsson, 200).
Since 1953, molecular biologists have discovered the structure and function of genes, and finally, at the beginning of the 21st century, they also discovered the human genome. During the postgenomic era, the purpose of molecular biology has changed. The search for an explanation of how complex molecular pathways and networks support biological structure and function has become a central question in molecular biology. The transition from a single molecule to a molecular network marked the birth of systems molecular biology. The second phase is the development of systems-mathematical biology, related to general systems theory (GSS) and the nonlinear dynamics of living organisms. The third phase, followed by the convergence of molecular systems biology and mathematical systems biology, is related to systems-based medicine, biotechnology, and drug development.
Although the term systems biology was widely used at the beginning of the 21st century with the advances in experimental methods, biological data analysis has always been a dynamic field of research, and with the emergence of so-called high-throughput approaches, This made it possible to simultaneously observe the behavior of a large number of distinct molecular species.

SYSTEMS BIOLOGY

For more information,please refer to the following sources:

https://medium.com/computational-biology/research-in-computational-biology-and-bioinformatics-121d92681aad

The main purpose of such analysis is to process the measurements obtained from the biological system to describe the functions and behavior of the system. Using high-power computing technologies, biological data were obtained at the molecular level, which requires new computational methods to process the data and generate answers to new questions in biology. The use of high-power computing technologies helps to understand complex biological systems and cellular interdependence and allows researchers to discover the function of a system. Understanding the complex relationship between the paths of biological systems is the most important challenge in biology, and the use of powerful computing technology has had a significant impact on solving this challenge.

A few major topics of research in the field of systems biology include:

- Gene regulatory networks
- Modelling metabolic interactions
- Model protective mechanisms induced by antibiotics
- Studying cell signaling pathways

According to the definition of systems biology as the ability to acquire, integrate and analyze complex experimental data sets using interdisciplinary tools, some of the main technologies of this field are as follows:

- Phenomics
- Genomics
- Epigenetics
- Transcriptomics
- Interactomics

Artificial Intelligence and Machine Learning

Topics

Artificial intelligence is one of the most widely used fields of the 21st century, and the first ideas for its emergence date back to the 17th century and the presentation of the concepts of “universal language” and “calculus ratiocinator” by Leibniz. Leibniz’ universal language aims to express all human knowledge through symbols and calculations. This will provide a precise background for the expression of logical thinking so that a machine can replicate it. He called a machine that could do such a thing a calculus ratiocinator. After Leibniz, the efforts of logicians such as John Stuart Mill, George Boole, and Gottlob Frege in the nineteenth century regarding formal logic, the use of logic in studying the laws of thinking, and the invention of a comprehensive system of mathematical logic laid the groundwork for the emergence of artificial intelligence in the twentieth century. The birth of artificial intelligence in its standard and well-known form goes back to concepts created in the fourth, fifth, and sixth decades of the 20th century: concepts such as the universal Turing machine, the logical reasoning machine of McCulloch and Pitts inspired by biological neurons and the Turing test.

Artificial Intelligence and Machine Learning

Topics

The term “artificial intelligence” (AI) was first coined by McCarthy at the Dartmouth Conference in 1956. According to his definition, “artificial intelligence is the science and engineering of making smart machines.” Since then, artificial intelligence, as a field that is concerned with building computer systems capable of learning, reacting, and making decisions in variable and complex environments, has developed and expanded into many subfields. Among these subfields, we can mention machine learning, deep learning, artificial neural networks, natural language processing, and machine vision. Since intelligence and ability to learn are closely related to each other, one of the important branches of artificial intelligence is machine learning, which aims to use specific features to recognize patterns and analyze a particular problem. In this way, the machine can learn from these patterns and apply what it has learned to similar situations.

Today, artificial intelligence and machine learning are used in various fields, including bioinformatics. In light of the ever-increasing growth of various types of biological data and advancement of the technologies needed to store them, large databases have been created for gene expression data, structures, sequences, microscopic images, and medical data, including clinical symptoms, CT scans, MRIs, FMRIs, EEGs, EKGs, and ECGs. The application of artificial intelligence and machine learning to these data can result in the recognition of patterns, the discovery of characteristics, and the solution of complex problems that could not be solved in the traditional way in the past. Also, the use of these two areas, besides the feasibility of data integration, leads to the development of tools and methods that are more efficient in terms of time and cost, the improvement of treatment methods, more accurate diagnoses, and a better understanding of the mechanisms of living organisms and biological processes.

For more information,please refer to the following sources:

Skansi, S. (2018). Introduction to Deep Learning: from logical calculus to artificial intelligence. Springer.

Singh, V., & Kumar, A. (Eds.). (2021). Advances in Bioinformatics. Springer.

Artificial Intelligence and Machine Learning

Topics

Common methods in these two fields include support vector machines, linear regression, principal component analysis, clustering, decision trees, random forests, evolutionary algorithms, image processing methods, and different types of neural networks (convolutional, auto-encoder, multilayer perceptron, RNN, DBN, LSTM, etc.). These methods can be used to solve the following problems:

- Biomarker discovery
- Promoter detection
- Gene expression analysis
- Medical image processing such as CT and MRI
- Prediction of metabolic pathways
- Protein family modeling and function prediction

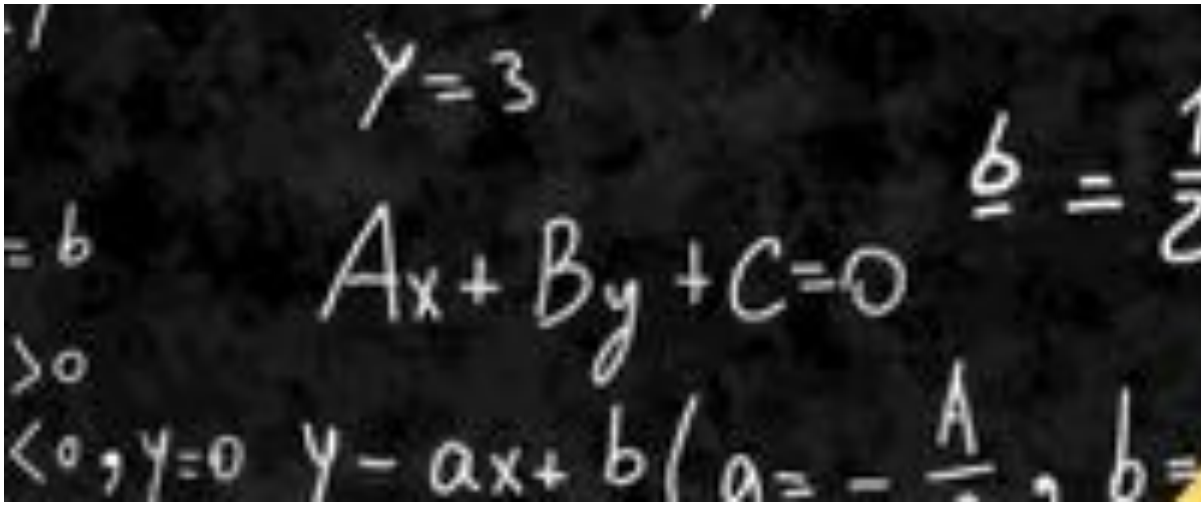
For more information,please refer to the following sources:

SSatpathy, R., Choudhury, T., Satpathy, S., Mohanty, S. N., & Zhang, X. (Eds.). (2021). Data Analytics in Bioinformatics: A Machine Learning Perspective. John Wiley & Sons.

Larranaga, P., Calvo, B., Santana, R., Bielza, C., Galdiano, J., Inza, I., ... & Robles, V. (2006). Machine learning in bioinformatics. Briefings in bioinformatics, 7(1), 86-112.

Mathematical Modeling in Computational Biology

Topics



The beginning of computational biology as a field, dates back to the 1960s and the use of computers in protein structure studies, such as performing Fourier analysis to determine the three-dimensional structure of proteins. Although the initial focus was on modeling the structure of proteins, by measuring and collecting various data, this field now covers a wide range of biological topics. The development of high-throughput technologies as well as the increase in the processing power of computers, especially since the 1990s, have enabled biology to rely more on computational methods. Mathematical approaches can be helpful in this regard.Fibonacci was one of the first to apply mathematics to biology by introducing his famous series to describe the population growth of rabbits in the 13th century. Also, Daniel Bernoulli, who used differential equations in the 18th century to analyze the mortality caused by smallpox and demonstrate the impact of inoculation.

Mathematical Modeling in Computational Biology

Topics

Since the 18th century, mathematics has been applied in various scopes of biology and medicine, such as studying the flow of biofluids (such as blood circulation), microbial colonies, bacterial and viral life cycles, and their relations with human diseases, genetics and heredity, clinical experiments, systems biology, epidemiology, immunology, instrumentation, microbiology and molecular cell biology, neuroscience, and so on.
Using mathematical approaches in the study of complex phenomena, including biological phenomena, helps us to better understand their mechanisms and enables us to predict and anticipate their behavior under different conditions. Among the mathematical methods, we can mention quantification of observations, optimization, data analysis and processing, prediction, and modeling. Mathematical modeling is the process by which a real-world problem is described by a mathematical formulation. So, the key issue in modeling is to find a suitable form of the phenomenon in mathematical language in the form of various functions, algebraic or differential equations, linear operators and matrices, combinatorial forms and graphs, etc.After this formulation is done, algorithms, theorems, and other common or new mathematical tools can be used to further study the model. Almost all fields of applied mathematics are used in these tools. Areas such as: differential equations (ODE and PDE), discrete-time and continuous-time dynamical systems, information theory and coding, graph and network theory, integral transformations, numerical analysis and calculation, delay differential equations, statistics, probability, and time series analysis.

For more information,please refer to the following sources:

Faugeras, O., & Janin, J. (2012). Modeling in Computational Biology and Biomedicine: A Multidisciplinary Endeavor. Springer Science & Business Media.

Ingalls, B. P. (2013). Mathematical modeling in systems biology: an introduction. MIT press.

Mathematical Modeling in Computational Biology

Topics

By modeling, a link between mathematics, computer science, statistics, and biology is established, which not only gives a clearer view of the phenomena, but also makes biology continuously benefit from the progress of these fields.
In addition, the use of computational and mathematical models in biology allows us to understand the mechanisms of complex diseases, improve treatment strategies and drug discovery, increase our sight of new molecular mechanisms, and reveal the rules that govern biological systems.
Among the topics that are studied in this field, the following can be mentioned:

- The spread of epidemics
- Macromolecular complexes
- The central nervous system's neural connectivity
- Validation of drug efficacy and safety
- Estimation of the survival rate for patients under a specific treatment
- Gene Regulatory Networks and Metabolic Networks

For more information,please refer to the following sources:

Bacaër, N. (2011). A short history of mathematical population dynamics (Vol. 618). London: Springer.

BIG DATA IN BIOINFORMATICS

Topics

Bioinformatics is one of the sciences that has faced massive data issues in various areas. As research in the life sciences becomes increasingly dependent on laboratory data, “bioinformatics” seeks to integrate closely related aspects of a field in order to comprehend the mechanism of a phenomenon in its entirety, as well as to comprehend the outputs and make them accessible for practical and research use. Therefore, the integration of these data makes us face a large amount of data; this increase in data will increase with time. The term “big data” refers to very large sets of data that can be analyzed by computers in a structured or unstructured way, as well as in a homogeneous or heterogeneous way, and that can show patterns, trends, and relationships.

The beginning of bioinformatics dates back more than 50 years. In fact, the foundations of bioinformatics were laid in the early 1960s using computational methods for protein sequence analysis. Research on protein sequences in 1956 led to the report of the first protein sequence related to bovine insulin. A decade later, Dayhoff created the Protein Data Bank as the first bioinformatics database.

BIG DATA IN BIOINFORMATICS

Topics

For more information,please refer to the following sources:

Branco, I., & Choupina, A. (2021). Bioinformatics: new tools and applications in life science and personalized medicine. Applied microbiology and biotechnology, 105 (3), 937-951.
Gauthier, J., Vincent, A. T., Charette, S. J., & Derome, N. (2019). A brief history of bioinformatics. Briefings in bioinformatics, 20(6), 1981-1996.
Pal, S., Mondal, S., Das, G., Khatua, S., & Ghosh, Z. (2020). Big data in biology: The hope and present-day challenges in it. Gene Reports, 21, 100869.

Since then, various databases have been gradually created, and with the increase in data, life science researchers have decided to use advanced technologies for data analysis. Many platforms with different applications, such as complete sequencing of multiple genomes, the study of gene expression profiles, study of epigenetic changes, study of mutations, etc., have been created, continuously adding to these data. Since the data in bioinformatics is usually very scattered and includes heterogeneous formats, many tools for analyzing this data type are available. Also, the growth rate of bioinformatics data is very high; for example, the total amount of sequence data generated doubles approximately every seven months. BigFiRSt and Sequence Scanner are two programs used to analyze and sequence this type of data. Among the other challenges that bioinformatics researchers are facing are substantial biological networks. Various tools are also available to analyze these networks. In addition to R software packages, we can also mention Sitescape and Netminer software.

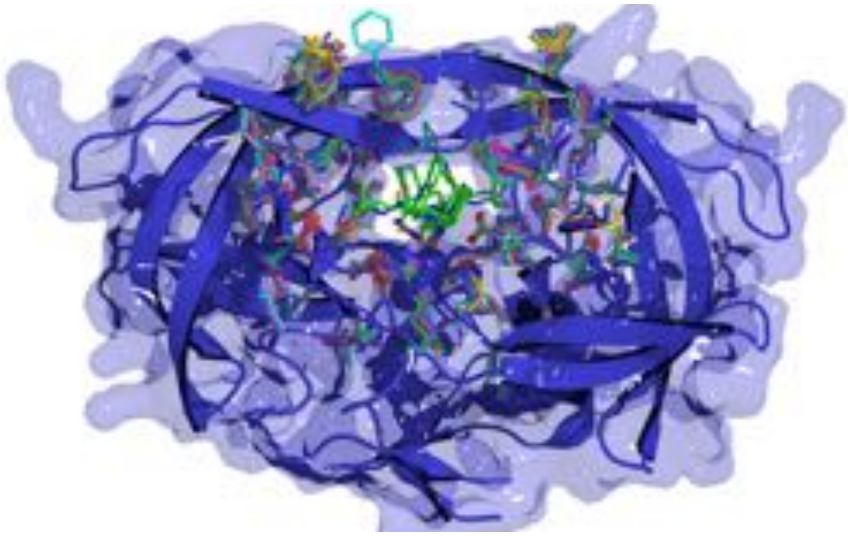
The progress of the tools for big data analysis is essential because it reduces the cost of calculations and increases computing power. Also, biologists no longer use traditional laboratories to discover and investigate biological interactions. Instead, they rely on the massive and constantly growing genomic data made available by various research groups.

Among the important areas of bioinformatics that are highly related to big data are the following:

1. Gene expression analysis
2. Sequencing
3. Determination of protein structure
4. Ontology of biological data

STRUCTURAL BIOINFORMATICS

Topics



Structural biology is a branch of biology that investigates and studies the 3D structure of biological molecules and their complexes. Investigating and studying the structure of these molecules is necessary to understand their function and role in the cell. This is because a biological molecule's structure determines and defines its function.Understanding how cell components function is crucial to developing more effective and successful treatments for illnesses. Research in the structural biology branch started in 1953, when Watson and Crick discovered the double helix structure of DNA. This was followed by the determination of the 3D structure of the myoglobin protein in 1958. Nowadays, the 3D structure of biological macromolecules is determined in laboratories using techniques like X-ray crystallography, nuclear magnetic resonance (NMR), and electron microscopy. Structures determined this way comprise structural biology data.

STRUCTURAL BIOINFORMATICS

For more information,please refer to the following sources:

- **Gu, J., & Bourne, P. E. (2009). Structural bioinformatics (Vol. 44). John Wiley & Sons.**
- **Punetha, A., Sarkar, P., Nimkar, S., Sharma, H., KNR, Y., & Nagaraj, S. (2018). Structural Bioinformatics: Life Through The 3D Glasses BT – Bioinformatics: Sequences, Structures, Phylogeny (A. Shanker (ed.); pp. 191–253). Springer Singapore. https://doi.org/10.1007/978-981-13-1562-6_10**

Due to the massive increase in the number of structural data in recent years, researchers have turned to bioinformatics approaches to organize and analyze these data. The branch of bioinformatics that collects, stores, analyzes, and interprets structural data is called structural bioinformatics. Structural bioinformatics uses computer models created based on the basic principles of biology, chemistry, physics, and computer science. These models are used to answer questions about the 3D structure of biological macromolecules like proteins, RNA, and DNA. Utilizing the speed and accuracy of computers, structural bioinformatics can play a significant role in improving health research, particularly by accelerating drug development and design.

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.

The main topics of research in the field of structural bioinformatics are:

- Protein and RNA structure predictions from sequence
- Protein function prediction from structure or sequence
- Protein classification
- Molecular movement modeling
- Binding site prediction
- Binding site modeling
- Drug design
- Functional protein design

COMPUTATIONAL DRUG DISCOVERY AND DESIGN

Topics

Drug (Medicine) is a known chemical compound that can make biological modification in living organism like activation or inhibition of a disease-related protein. It is used for the treatment, diagnosis, improvement and prevention of diseases. When there is not a suitable medicinal treatment, the de novo drug discovery project is started. Shortly after discovery of ‘miracle drugs’ such as penicillin during World War as well as adventing synthetic organic chemistry make the preparation of ‘un-natural’ compound or drug candidates economically possible on an industrial scale.

De novo drug discovery consists of three major steps: Research and development, pre-clinical studies and clinical examination. It is crucial for drug discovery to consider disease type and identify the potential biological targets such as receptors, enzymes, proteins, genes and RNA. One of the most important properties of a determined drug target is to be modulated by a therapeutic agent.

COMPUTATIONAL DRUG DISCOVERY AND DESIGN

According to latest studies, de novo drug discovery takes more than 10 years and costs about 314 million to 2.8 billion dollars. Therefore, despite the progress in biotechnology, drug discovery is a risky, costly and time-consuming process.

Drug design is the inventive process of finding new medications based on the knowledge of a biological target. drug design involves the design of molecules that are complementary in shape and charge to the biomolecular targe.

The computer systems can help and reduce the costs of drug design and discovery, due to the growing number of available databases, the computation power improvement of computer systems, expansion and development of computational methods in different branches of science. Drug design with aid of computers is called Computer-Aided Drug Design or (CADD). CADD comprises a broad range of theoretical and computational approaches that are part of modern drug design and discovery.

The computational drug design models can be divided into two main types:

- Structure- based: relies on knowledge of the three-dimensional structure of the biological target,
- Ligand-based: relies on knowledge of other molecules that bind to the biological target of interest.

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
- And...

BIOLOGICAL SEQUENCES ANALYSIS

Topics

A major application of bioinformatics is the analysis of biological sequences, which was sparked by the development of the Basic Local Alignment Search Tool (BLAST) program in 1990. This is despite the fact that its beginning dates back several years before the development of this algorithm, between 1969 and 1977. In 1969, sequence analysis of tRNAs was used to infer interactions resulting from associated changes in nucleotide sequences, which eventually led to a tRNA secondary structure model.
Another concern of these years that led to a massive global project was the Human Genome Project. Before this project, limited information about human genes had been discovered. However, not only were the volumes of this information not significant, but not all chromosomes were examined. But, preventing hereditary genetic diseases, such as Down syndrome, is impossible without studying all chromosomes. The purpose of this project was to map and locate about 25,000 human genes (located on 23 different pairs of chromosomes) and to determine the function of each of these genes. This had not previously been possible using non-computational methods.

BIOLOGICAL SEQUENCE ANALYSIS

Biological sequence analysis is the process of analyzing the sequence of RNA, DNA, or a peptide using one of the available analytical methods. Biological sequence relationships and the meaning of these relationships are studied to identify evolutionary relationships (common ancestors) and to predict unknown sequence structure and function.
These goals can be achieved by using algorithms and search tools from biological databases such as BLAST and FASTA, phylogenetic tree construction methods, repetitive algorithms like genetic algorithms, statistical models such as Markov chains, dynamic programming like Smith-Waterman and Needleman-Wunsch, scoring matrices like PAM and BLOSUM, Bayesian alignment algorithms, progressive alignment models like ClustalW and Gibbs Sampler, transformation grammars, neural networks, etc.
Using such methods to solve biological sequence problems has provided more efficient methods in terms of time and cost, as well as solving problems that were not possible with laboratory methods. 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

For more information,please refer to the following sources:

- **Durbin 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 at: https://books.google.com/books?id=R5P2GlJvigQC.**

WORKSHOPS

Drug design

Dr. sajjad Gharghani

Date : 2023-03-03 - 2023-03-04

Machine learning for drug

Dr.Changiz Eslahchi
Dr.Fatemeh Yassaie

Date : 2023-03-04

Introduction to exom data analysis and variant calling

Dr. Kaveh kavosi
Mr.Nasser Elmi

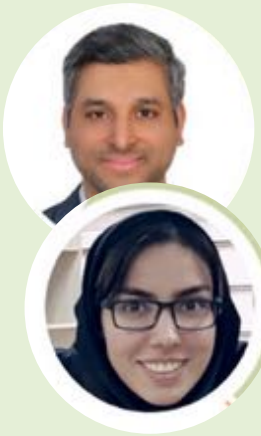
Date : 2023-03-02 - 2023-03-03

Metabolic networks modelling

Dr. Mohammad Hossein karimi-jafari
Ms.Sepideh Mofidifar

Date : 2023-03-03 - 2023-03-04

WORKSHOPS



Application of computational models in cancer studies

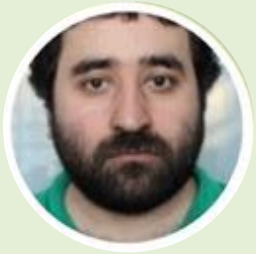
Dr. Hesam Montazeri
Ms.Mozhgan Mozaffari

Date : 2023-03-09

Biological network reconstruction

Dr.Nazanin Hoseinkhan

Date : 2023-03-04



Reconstruction of gene regulatory networks using gene expression data

Dr. Alirezai Fotuhi

Date : 2023-03-07

MICROARRAY DATA ANALYSIS

Dr.samaneh Maleknia

Date : 2023-03-03 - 2023-03-04

WORKSHOPS

RNA-seq data analysis

Dr. Amir Malekpour
Dr. Najmeh salehi

Date : 2023-03-02 - 2023-03-03

Metagenome data analysis

Dr.Mohammad Hossein Nourozi

Date : 2023-03-05

Analysis of 16s rRNA sequences using Qiime2

Dr.Mohammad Hossein Nourozi

Date : 2023-03-06

STUDENT SYMPOSIUMS

DRUG DESIGN AND REPORPUSING

Date : 2023-02-23

In this webinar, we explain the de novo drug development procedure, then we analyze available databases. Finally, we speak about drug repositioning as a strategy to hasten this process.

IMMUNOINFORMATICS

Date : 2023-02-24

In this course, we provide an introduction to immunoinformatics For this reason, we will first talk about the science of immunology and computational immunology, and after introducing immunoinformatics and its applications, we will introduce and teach working with IEDB

Single-cell RNA-seq analysis and its application in cancer

Date : 2023-02-22

STUDENT SYMPOSIUMS

DATA VISUALIZATION WITH PYTHON

Date : 2023-02-23

In this course, we will describe one of the primary data visualization libraries of Python called “Matplotlib” and draw a few basic graphics. Next, we will browse through a library called “Seaborn” which provides a high-level interface for drawing beautiful and informative statistical graphs. Lastly, we will learn about interactive data plotting.

SYSTEMS BIOLOGY

Date : 2023-02-23

- Introduction: Intro, importance, and history of systems biology
- overview of modeling in systems biology
- Graph-theory and Basic Network properties (Barabaci)

STUDENT SYMPOSIUMS

DEEP LEARNING IN BIOINFORMATICS

Date : 2023-02-24

In this workshop, we will review some basic concepts in deep learning then discuss several deep ANN architectures and their application in bioinformatics. ,Introduction to AI and Machine learning, Basic concepts in Deep Learning , Convolutional Neural , Networks (CNN) , Recurrent Neural Networks (RNN) , Autoencoders , Deep Belief Neural Networks (DBN) , Generative Adversarial Network (GAN) , Graph Neural Networks (GNN)

BIOPYTHON

Date : 2023-02-23

Installation , Bio.Seq , Accessing Entrez , Seq.IO , Protein visualization

STUDENT SYMPOSIUMS

NOVEL MOLECULAR DYNAMICS SIMULATION APPROACHES

Date : 2023-02-24

Molecular dynamics (MD) simulation is one of the widely used techniques in structural bioinformatics. This method plays an important role in understanding and investigating biological structures and their dynamic processes. From studying proteins and peptides to large-scale complexes like biological membranes, all are feasible with different types of MD. In addition to traditional and well-known MD simulations, new methods are being developed that can capture free energy, thermodynamic parameters, and chemical reaction processes. The main goal of this field is to simulate a system that is close to natural conditions. In this seminar, the basis, applications, and studies on several techniques like SMD, REMD, aMD, and Metadynamics will be discussed.

GENETIC ALGORITHM

Date : 2023-02-24

The focus in this presentation is on a deep and detailed understanding of this algorithm theoretically and implementing this algorithm in Python is also done in order to have a better and deeper understanding. Therefore, it is not necessary to learn or know Python programming language, as a prerequisite for this presentation, but, if you want to implement GA one day, being able to Python program will be beneficial.

BioGap

BioGap is an online talk and debate show featuring scholars from various disciplines discussing the fields of biology and computation. As an initiative by the executive team of the 11th Iranian Conference on Bioinformatics (ICB11), it got overwhelmingly positive feedback from biologists and computer scientists all over the country.

By leveraging its position within the Iranian Bioinformatics Society (IBIS), the organization behind ICB, and using their connections, BioGap was able to secure the participation of scientists with the highest honors and then capitalize on the notoriety of some of these scientists to attract a previously unattended Persian-speaking scientific audience. Drs. Sharifi Zarchi, Karmi Jafari, Safikhani, and Elahe Elahi are among the most popular speakers on the program.

Each episode's topic and guests were chosen to encourage a lively discussion between two persons with opposing points of view. One episode, for instance, included two highly respected guests, one from a scientific background and the other from an engineering background. In another, more nuanced instance, two bioinformaticians argued for almost three hours over the significance of genomic data in the context of polygenic diseases.

Despite the great reception we have received thus far, we still consider the show to be in its infancy. We've got some fresh ideas about the show's format, and we're open-minded to experimenting with it.

COMMITTEESS LIST

Mehdi Sadeghi

Conference president

Associate professor of cell and molecular biology at National Institute of Genetic Engineering and Biotechnology (NIGEB)
Head of School of Biological Sciences, Institute For Research In Fundamental Sciences (IPM)

Fatemeh Zare-Mirakabad

Conference scientific secretary

Assistant Professor in Mathematics and Computer Science Department at Amirkabir University of Technology, Tehran, Iran

Najmeh Salehi

Conference executive secretary

Postdoctoral Researcher at Institute for Research in Fundamental Sciences

Abbas Nowzari Dalini

Scientific committee

Full Professor; Department of Mathematics, Statistics and Computer Science, University of Tehran

SCIENTIFIC COMMITTEESS

Bahram Goliaei

Scientific committee

Professor of Biophysics at University of Tehran

Hamid Pezeshk

Scientific committee

Professor at University of Tehran

Kaveh Kavousi

Scientific committee

Faculty member at Bioinformatics department, Institute of Biochemistry and Biophysics (IBB), University of Tehran

Sajjad Gharaghani

Scientific committee

Assistant Professor at University of Tehran, Laboratory of Bioinformatics and Drug Design

COMMITTEESS LIST

Mohammad Hossein Karimi-Jafari

Scientific committee

Assistant Professor; Institute of Biochemistry and Biophysics, University of Tehran

Changiz Eslahchi

Scientific committee

Professor in the department of computer sciences, school of mathematics, Shahid Beheshti University

Mohammad Ganj Tabesh

Scientific committee

Associate Professor; Department of Mathematics, Statistics and Computer Science, University of Tehran

Hesam Montazeri

Scientific committee

Assistant Professor; Institute of Biochemistry and Biophysics, University of Tehran

COMMITTEESS LIST

Seyed Shahriar Arab

Scientific committee

Assistant Professor of School of Biological Sciences, Tarbiat Modares University.

Sayed-Amir Marashi

Scientific committee

Associate Professor; Department of Biotechnology, University of Tehran

Alireza Fotuhi Siahpirani

Scientific committee

Assistant Professor; University of Tehran

Yazdan Asgari

Scientific committee

Postdoc fellowship at INSERM U1018 CESP, Exposome and heredity; Institute Gustave Roussy, Villejuif, France

COMMITTEESS LIST

Babak Khalaj

Scientific committee

Innovative Data Science Ecosystems

Mohieddin Jafari

Scientific committee

PI at University of Helsinki

Seyed Abolfazl Motahari

Scientific committee

Assistant professor, Department of Computer engineering, Sharif University of technology

Ali Sharifi-Zarchi

Scientific committee

Faculty of Bioinformatics and Artificial Intelligence, Sharif University of Technology

COMMITTEESS LIST

Zarrin Minuchehr

Scientific committee

Associate Professor, Faculty Member of the National Institute of Genetic Engineering and Biotechnology

Rosa Aghdam

Scientific committee

Senior Post-Doctoral Research Fellow, School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM)

Reza Hassanzadeh

Scientific committee

Assistant professor at the Faculty of Advanced Technologies of University of Mohaghegh Ardabili

Pegah Khosravi

Scientific committee

Assistant Professor, City Tech

COMMITTEESS LIST

Javad Zahiri

Scientific committee

Postdoctoral Scholar Employee, Department of Neurosciences, University of California San Diego

Zahra Razaghi-Moghadam

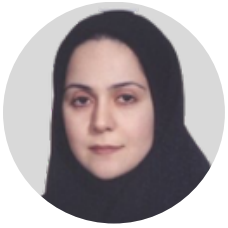
Scientific committee

Researcher at Max Planck Institute of Molecular Plant Physiology

Zaynab Mousavian

Scientific committee

Postdoctoral researcher at Karolinska Institute



Nazanin Hosseinkhan

Scientific committee

Assistant Professor at Iran University of Medical Sciences

COMMITTEESS LIST

Abbasali Emamjomeh

Scientific committee

Assistant Professor at University of Zabol

Mehdi Mirzaie

Scientific committee

Researcher at University of Helsinki

Ghasem Hosseini Salekdeh

Scientific committee

Professor in Genetics

Esmaeil Shahsavand

Scientific committee

Genomic Psychiatry

COMMITTEESS LIST

Zahra Salehi

Scientific committee

Tehran University of Medical Sciences

Saeideh Khodaee

Scientific committee

Postdoctoral researcher in the Computational Biology group (del Sol Lab)

Najmeh Khorasani

Scientific committee

Postdoctoral Researcher at School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM)

Seyed Amir Malekpour

Scientific committee

Postdoctoral Researcher at School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM)

COMMITTEESS LIST

Ali Najafi

Scientific committee

Professor at Baqiyatallah University of Medical Sciences

Mahya Mehrmohamadi

Scientific committee

Assistant professor at Department of Biotechnology, University of Tehran

EXECUTIVE COMMITTEE MEMBERS

ZAHRA GHORBANALI

TEAM LEADER

MAHSA SAADAT

LEADERSHIP TEAM

MARYAM RASTEGARI

LEADERSHIP TEAM

MOHAMMAD FARAHANCHI

TRANSLATION TEAM LEADER

MEHDI PAYKAN HEYRATI

SPONSERING TEAM LEADER

MARYAM NAZARLOO

RESEARCH TEAM

SEYYED HASSAN ALAVI

SITE TEAM

SOBHAN AHMADIAN

BOOKLET DESIGN TEAM LEADER

REYHANEH EBRAHIMI

DOCUMENTARY TEAM LEADER

EXECUTIVE COMMITTEE MEMBERS

ROZA HAMIDI

BOOKLET DESIGN TEAM

RAZIYEH SOLEIMANBEIGI

SPONSERING TEAM

AREZOU PAKSERESHT

RESEARCH TEAM

AMIR ZARGHAMI

BOOKLET DESIGN TEAM

MILAD BESHARATI

RESEARCH TEAM

MOSTAFA MESHKATI

SITE TEAM LEADER

AMIR ABBAS ALVAND

REGISTRATION

MOZHGAN MOZAFFARI

REGISTRATION

SARA MOHAMMADZADEH

GRAPHIC TEAM

Abstracts

SPIDER CHART

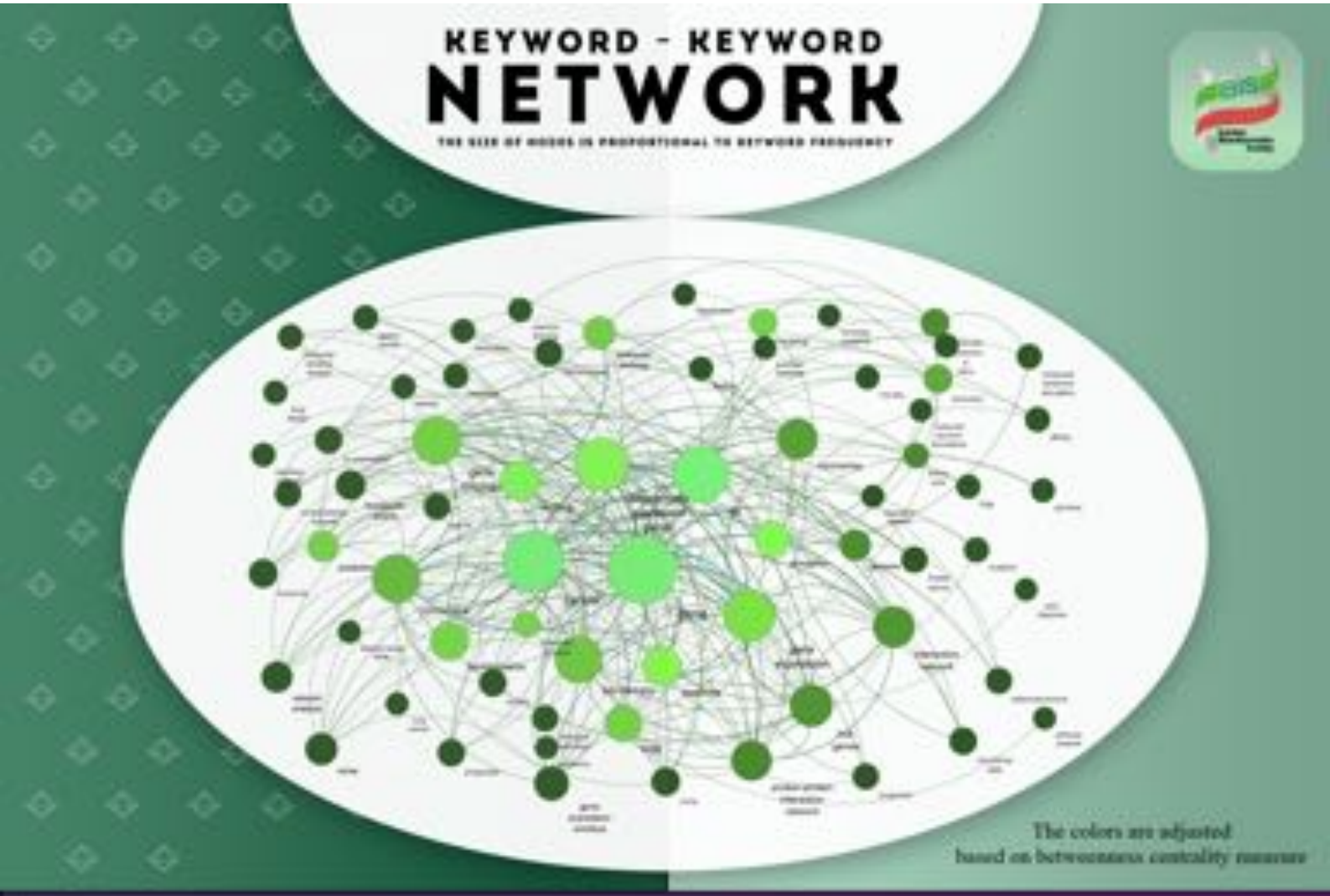
INSTITUTE-SCIENTIFIC TOPIC

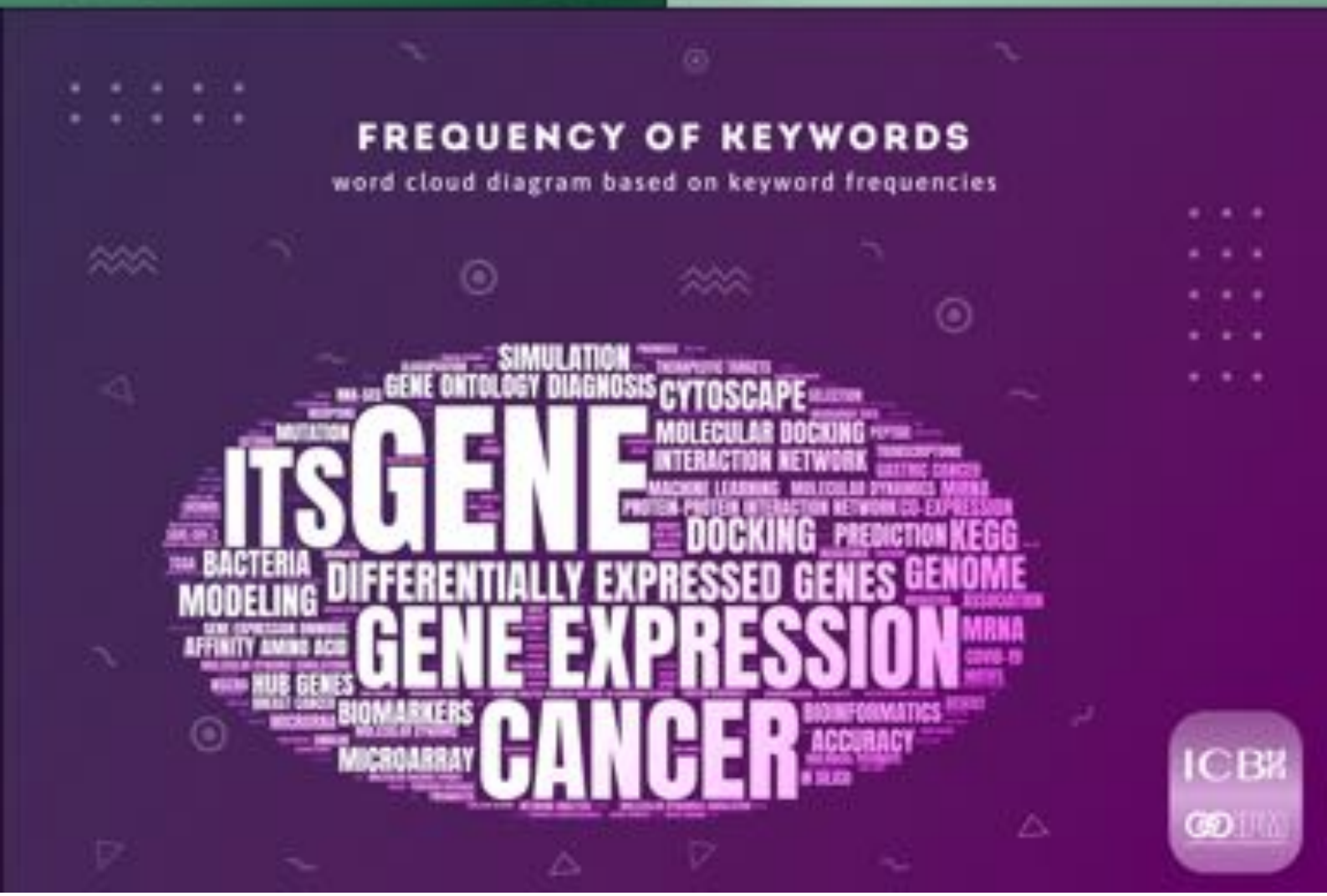
INSTITUTE-SCIENTIFIC KEYWORD

TOP INSTITUTIONS

AFFILIATION WORD CLOUD

Based on number of accepted papers


KEYWORD - KEYWORD
NETWORK
THE SIZE OF NODES IS PROPORTIONAL TO KEYWORD FREQUENCY
The colors are adjusted
based on betweenness centrality measure


FREQUENCY OF KEYWORDS
word cloud diagram based on keyword frequencies
SIMULATION
GENE ONTOLOGY
DIAGNOSIS
CYTOSCAPE
MUTATION
ITS
GENE
MOLECULAR DOCKING
INTERACTION NETWORK
MACHINE LEARNING
MOLECULAR DYNAMICS
MRNA
PROTEIN-PROTEIN INTERACTION NETWORK
CO-EXPRESSION
DOCKING
PREDICTION
KEGG
BACTERIA
MODELING
DIFFERENTIALLY EXPRESSED GENES
GENOME
AFFINITY
AMINO ACID
GENE EXPRESSION
MRNA
HUB GENES
BIOMARKERS
BIOINFORMATICS
MICROARRAY
CANCER
ACCURACY
ICBK

ICB11-1401-12-1001

A Novel Approach for Identifying Cancer Drivers in Coding and Non-coding Genomic Elements

Farideh Bahari[1], Reza Ahangari cohan[2] and Hesam Montazeri[3*]

[1]Institute pasteur of iran. [2]Institute pasteur of iran. [3*]University of tehran.

*Corresponding author(s). E-mail(s): hesam.montazeri@ut.ac.ir;

Abstract

Background:
Cataloging cancer drivers is crucial in cancer research. Current methods for identifying cancer drivers, especially in non-coding regions of the genome, mainly rely on detecting signals of positive selection using two approaches: mutational burden tests, which compare observed and background mutation rates, and functional impact tests, which incorporate functional relevance of mutations in a genomic region. However, previous studies have had limited success in identifying potential drivers in non-coding regions. This paper introduces a new statistical test combining functional impact and mutational burden tests to better prioritize driver candidates in both coding and non-coding elements.
Methods:
This study used somatic PCAWG consensus callsets of SNV/Indels of 2253 non-hyper mutated samples across 33 cancer types from the PCAWG project. The somatic variants were annotated using CADD scores. The background mutation rates were estimated using the DriverPower gradient boosting model based on a feature matrix of 1373 genomic and epigenomic features influencing mutation rates. This calculation took into account the genomic coordinates of both coding and non-coding elements and randomly generated bins, excluding blacklisted regions of the genome. The study then employed a novel statistical test based on probabilistic graphical models that incorporate both mutation recurrence and CADD score to prioritize cancer driver elements. Finally, the p-values were corrected for multiple testing with the Benjamini-Hochberg procedure.
Results and Discussion:
The performance of this approach was compared to other PCAWG driver discovery methods namely ActiveDriverWGS, DriverPower, NBR, oncodriveFML, MutSig, ncdDetect, and ExInAtor in terms of AUPR and AUC measures. The study resulted in AUC (0.679751, 0.5470776) and AUPR (0.195035, 0.080829) and outperformed other computational methods applied to the PCAWG data.

Keywords: Cancer driver, non-coding genomic elements, Somatic mutations, Mutational burden tests, Functional impact tests

ICB11-1401-12-1002

Transcriptome-wide analysis of DNA-RNA hybrids reveals how non-coding RNAs contribute to ASDs

Kasra Mokhtarzadeh azar[1], Leila Kian mehr[2*], Hassan Rajabi maham[3], Minoo Rassoulzadegan[4]

[1]Shahid beheshti university. [2*]Shahid beheshti university.[3]Shahid beheshti university. [4]Université cote d'azur, erciyes university.

*Corresponding author(s). E-mail(s): kianmehr96@gmail.com;

Abstract

Autism spectrum disorders (ASDs) are a class of neurodevelopmental disorders with a strong genetic basis. In addition to genetic factors, the role of epigenetic and environmental factors has also been reported in ASDs. Still, the pathogenesis of ASDs is not well understood. DNA-RNA hybrids are epigenetic factors whose abnormal accumulation on the genome can lead to genomic instability and DNA damage. Their various roles in neurological disorders make them promising candidates for investigating the pathogenesis of ASDs. In this study, we investigated the distribution of DNA-RNA hybrids in the blood samples of ASD patients compared to control samples using RNA-Seq data analysis. Furthermore, transcriptome assembly methods were used to discover novel non-coding transcripts involved in ASDs. The overall distribution of DNA-RNA hybrids showed that in contrast to previous studies, most of these hybrids were accumulated in exonic regions. Moreover, we found 301 genes with significant accumulation of DNA-RNA hybrids in ASD samples. Interestingly, we observed ASD risk genes like SMARCC2 and GIGYF1 among these genes. This can point to a novel mechanism that can disrupt the normal function of these genes and lead to the pathogenesis of ASDs. Furthermore, there were numerous mitochondrial genes among the genes with significant DNA-RNA hybrid accumulation, which can put an emphasis on the previously reported role of mitochondrial dysfunction in the pathogenesis of ASDs. Finally, we found a gene module containing 45 novel lncRNA genes that showed DNA-RNA hybrid accumulation in ASD samples and can be used to classify ASD and control samples. In summary, our study showed that DNA-RNA hybrids demonstrate a distinct pattern of transcriptomic profiles in ASD and control samples, thus a potential role for DNA-RNA hybrids in the pathogenesis of ASDs can be inferred. In addition, the genes introduced in this study can be potential biomarker candidates in future studies.

Keywords: DNA-RNA hybrids, ASDs, RNA-Seq, transcriptome assembly, lncRNA

ICB11-1401-12-1003

Reconstruction of genome-scale metabolic model for Helicobacter pylori

Sepideh Mofidifar[1], Mohammad hossein Karimi-jafari[2*], Abbas Yadegar[3]

[1]University of tehran. [2*]University of tehran.[3]Shahid beheshti university.

*Corresponding author(s). E-mail(s): mhkarimijafari@ut.ac.ir;

Abstract

Helicobacter pylori is a gram-negative bacteria that specifically colonizes gastric epithelium of more than half of the world population. The presents of H. pylori evokes various diseases including peptic ulcers, gastric ulcers, adenocarcinoma, MALT lymphoma, and gastric cancer . Based on the H. pylori's species and host characteristics, the type of disease and intensity may be highly variable . The genome sequence of H. pylori is well characterized as NCBI reports more than 500 different strains around the world. Despite that, these genome analyses merely supply information about the overall metabolism of helicobacter pylori. On the other hand, the survival of H. pylori in human host could reflect its unique capability of physiological response which displays the need of crucial research in this field. Genome-scale metabolic model is an approach which extensively used for analysis of metabolic pathways of different organisms. It defines gene-protein-reaction associations for metabolic genes by contextualizing different types of Big Data (e.g., genomics, metabolomics, and transcriptomics) . Therefore, in this study various omics data of helicobacter pylori 26695 were utilized to construct and contextualize metabolic model of H. pylori. In this way, the existed genome scale metabolic model of helicobacter pylori 26695, Ilt341 , was used as template model in RAVEN Toolbox, which combined draft GEM from KEGG and MetaCyc pathway databases. Based on comparison between the combined draft model and Ilt341model, the reactions that were absent in Ilt341, were incorporated in to the model. Subsequently, transport DB, UniProt and other available experimental data were utilized for improving and validating of model. According to amino acid auxotrophy, Biolog phenotype microarray and gene essentiality simulation, the model shows approximately 85%, 80% and 79% accuracy respectively. This model can be used for deeply understanding mechanistic underlying H. pylori's metabolic pathway.

Keywords: Helicobacter pylori, Genome scale metabolic model

ICB11-1401-12-1004

A bioinformatics approach to analysis UPR-related gens impact on AML

Soheila Shabani-mashcool[1*], Mojtaba Amani[2], Reyhaneh Goloazari[3], Reza Hassanzadeh[4], Soheila Shabani-mashcool[5]

[1*]University of tehran.[2]Ardabil university of medical sciences. [3]Ardabil university of medical sciences. [4]University of mohaghegh ardabili. [5]University of tehran.

*Corresponding author(s). E-mail(s): soheila.shabani@ut.ac.ir;

Abstract

Acute myeloid leukemia (AML) is a cancer of the blood and bone marrow. It is the most common type of acute leukemia in adults. This type of cancer usually gets worse or can recur quickly if it is not treated. AML is also called acute myelogenous leukemia and acute nonlymphocytic leukemia . The unfolded protein response (UPR) is an evolutionarily conserved adaptive response triggered by the stress of the endoplasmic reticulum. UPR can change cell protein homeostasis . It have pivotal signature in cancers progression like breast cancer but its role in AML have been not fully investigated. Here, UPR associated proteins were extracted from GeneCards database , transcriptional data for AML and normal blood cells were achieved from TCGA and GTEx respectively. For obtaining Differentially Expressed Genes (DEG) in AML, three R packages were used , and . The intersection of these DEGs with UPR related genes were extracted and the products of this DEGs were invested by using String database and PPI network obtained. By Cytoscape topological properties of the network were studied. To see the effect of UPR-related DEGs in survival time of the patients, two groups of AML patients extracted and named the ones with low and high survival time. DEGs were achieved by limma R package and then the intersection of theses gens by UPR related gens like BRCA1 and Cenpi were reported. Gene enrichment analysis were done to find significant terms, like cell cycle, in survival time of patients. In this work, a bioinformatics approach employed to investigate UPR role in AML. Results showed that UPR can be the promising diagnostic target in AML.

Keywords: UPR, AML, Gene expression data

ICB11-1401-12-1005

Finding a genomic pattern for classifying breast cancer patients by Random Forest and Linear Discriminant Analysis and cancer-specific biomarkers by Integrative analysis of DNA methylation and gene expression

Romina Norouzi[1], Sajjad Gharaghani[2*]

[1]University of tehran. [2*]University of tehran.

*Corresponding author(s). E-mail(s): s.gharaghani@ut.ac.ir;

Abstract

Differential methylation analysis of Illumina Human Methylation 450K from TCGA performed between 27 normal and 313 tumor samples and 27,342 methylation sites, Beta value (β) used as an index to quantify the level of methylation, The Wilcoxon rank-sum test used to determine the differentially methylated CpGs (DMCs), and the p-values adjusted using the FDR method. DMCs reported if the mean methylation difference was >0.2 with an FDR of 5%. Transcriptome data and clinical information of breast cancer patients (including 1222 samples) accessed from the TCGA database. Differentially expressed genes, reported only if the log-fold change was >1.5 and the adjusted p-value was smaller than 0.05. Then, we integrated Differential expression and methylation data to identify the breast cancer-specific markers. After that, to find a genomic pattern for classifying breast patients, we performed Random Forest and Linear Discriminant Analysis on TCGA data and compared a complex model with three hyperparameters, such as RF against a simple baseline model such as LDA. For machine learning analysis, we will separate the dataset into two groups, training, and testing (80%-20%) according to the output of the training we have 1090 samples and 13793 genes. To follow a robust experimental design, we followed Random Forest and Linear Discriminant with 5 runs of a 10-fold cross-validation approach. Implementation of the RF has the hyperparameters such as the number of features randomly selected for each tree and the minimum node's size. Implementation of the LDA has no hyperparameters. To find which model is the best, we compare these two models; we ran 5 repetitions of a 10-fold cross-validation experiment and ensured that both models worked on the same set of data and partitions and in the same order. The Accuracy of two models during the five runs for the training set of the random forest was 0.65 and for Linear Discriminant Analysis was 0.48 After finding reliable biomarkers and specific patterns, functional enrichment, and drug-gene interaction, done. To discover the main function and the information from associated enriched pathways of the differentially expressed and methylated genes, we used Gene ontology (GO), Benjamini–Hochberg method, and p-value <0.05 and q-value<0.05 criteria used for GO analysis of DEGs. For finding Drug -the gene interaction analysis of the differentially expressed and methylated genes, we only selected protein coding gene that logFC>3 and logFC < 3 for DGIdb input, export interaction from DGIdb and, and then selected only an activator and inhibitor interaction type.

Keywords: machine learning, DNA methylation and gene expression, cancer-specific biomarkers

ICB11-1401-12-1006

Candidate prognostic gene signatures in lung cancer risk among asthmatic and COPD patients

B. fatemeh Nobakht m. gh[1], Masoud Arabfard[2*]

[1]Baqiyatallah university of medical sciences. [2*]Baqiyatallah university of medical sciences.

*Corresponding author(s). E-mail(s): arabfard@gmail.com;

Abstract

The second most frequent cancer in the world is lung cancer (LC). It is clear that both inherited and environmental factors contribute to the onset and progression of various respiratory diseases. According to reports, some disorders, including asthma and chronic obstructive pulmonary disease (COPD), elevate the risk of LC. Therefore, this study aims to find common genes and biological pathways in patients with LC/COPD and LC/asthma so that perhaps we can achieve a better understanding of the progression of LC from other pulmonary disorders. In the current study, the nine microarray datasets (three for asthma, four for LC, and two for COPD) were selected from the NCBI GEO database to explore differentially and commonly expressed genes as well as unique and shared biological pathways involved in LC/asthma and LC/COPD patients. Nine datasets, including 547 individuals, were included in the analysis. Our results showed that 46 and 23 candidate signatures predicted the incidence risk of LC from asthma and COPD, respectively. Furthermore, common biological pathways involved in LC/asthma and LC/COPD patients were 28 and 22, respectively. It may be possible to predict the incidence risk of LC from asthma and COPD by looking at the specific and shared essential genes and biological pathways involved in LC/COPD and LC/asthma. The importance of the key genes for COPD and asthma in predicting LC should be investigated further.

Keywords: Lung cancer, asthma, COPD, prognostic genes, Microarray

ICB11-1401-12-1007

Identification of Transcription factors involved in cisplatin drug resistance in ovarian cancer by integrated analysis of transcriptome and epigenomic data

Sana Kerachian[1], Zahra Fadavi[2], Hadi Alizadeh[3], Bahram Mohammad soltani[4*]

[1]Tarbiat modares university. [2]Tarbiat modares university. [3]Tarbiat modares university. [4*]Tarbiat modares university.

*Corresponding author(s). E-mail(s): soltanib@modares.ac.ir;

Abstract

Ovarian cancer causes 140,200 deaths each year and is typically diagnosed at a late stage. Standard treatments for newly diagnosed ovarian cancer are surgery and platinum-based chemotherapy like cisplatin. Ovarian cancer at diagnosis is typically very responsive to platinum-based chemotherapy. However, it frequently relapses and becomes increasingly resistant to chemotherapy. Consequently, it is essential to understand the mechanisms underlying platinum resistance and find ways to overcome them. here we used integrated analysis of RNA-seq , CHIP-seq data to find transcription factors that are the core player of cisplatin resistance. First, we find differentially expressed genes by analysis of A2780 cisplatin sensitive and resistance cell lines RNA-seq data from three datasets (PRJNA724680, PRJNA384352 & PRJNA735397). In order to find pathways involved in resistance to cisplatin we performed functional enrichment analysis, including GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis. We then analyzed CHIP-seq data to find differentially acetylated enhancers. By integrated analysis of transcriptome and epigenomic data we find transcription factors that target up regulated genes in resistant cell lines. Among these transcription factors, some of them are upregulated in cisplatin-resistant cell lines, which can be further studied to design drugs that improve response to cisplatin.

Keywords: cisplatin resistance, ovarian cancer, transcription factors

ICB11-1401-12-1008

Purple Sulfur Oxidizing Bacteria: Model Reconstruction and Flux Balance Analysis

Mohadese Astaraki[1], Fereshteh Naeimpoor[2*], Panteha Pirieh[3]

[1]Iran university of science and technology. [2*]Iran university of science and technology.[3]Iran university of science and technology.

*Corresponding author(s). E-mail(s): fnaeim@iust.ac.ir;

Abstract

Background:
Sulfur containing contaminants released into environment have become an environmental issue due to their hazards and threats to human health and ecosystems. Various chemical, physical and biological methods have so far been used to remove these compounds. Application of purple phototrophic sulfur-oxidizing bacteria has shown advantages over conventional biological method. These microorganisms perform anoxygenic photosynthesis alongside utilization of various sulfur compounds as electron donors and are known for production of elemental sulphur as desired product.
Purpose:
Providing insights into metabolism of these group of microorganisms in particular their cell growth and product formation.
Method:
Firstly, a metabolic network (99 reactions) including oxidation of sulfur compounds, central carbon metabolism, synthesis of various biomolecules and the exchange reactions for nutrients and products was reconstructed for the purple sulfur bacteria (PSB). Subsequently, the model was solved using COBRA toolbox in MATLAB software and flux distributions within cell under different conditions were examined.
Result:
It was shown that the specific growth rate was optimized (0.042 h-1) when sulfide was consumed as electron donor, with 7.576 mmol/gDwh of light required for 1 mmol/gDwh of sulfide uptake and sulfate formation as the only product. However, the flux of sulfur as the desirable product was maximized at 2 mmol/gDwh of light consumption with specific growth rate of 0.011 h-1.
Conclusion:
The reconstructed metabolic model for PSB can have economic usage in the industrial removal of sulfur contaminants due to using natural light as the energy source and producing elemental sulfur as the appropriate product.

Keywords: Distribution of intracellular fluxes, FBA, Purple sulfur bacteria, Sulfur phototrophic metabolism

ICB11-1401-12-1009

Prediction of protein druggability by supervised learning approaches

Parvin Mansouri[1], Mahboobeh Zarrabi[2*], Fatemeh Ebrahimi tarki[3]

[1]Alzahra university. [2*]Alzahra university.[3]Alzahra university.

*Corresponding author(s). E-mail(s): mzarrabi@alzahra.ac.ir;

Abstract

Drug discovery includes several steps that lead from a biological hypothesis to an approved drug. Target identification is usually the starting point of the modern drug discovery process. Candidate targets may be selected based on different experimental criteria . Nowadays, most druggable targets or drug receptor molecules are proteins. As a different approach from traditional drug development, machine learning algorithms are used to predict drug targets using data mining. This method has received more attention in recent years due to its various advantages such as reducing cost and time. The primary approach of this research is to provide appropriate models for extracting and selecting the features of proteins and predicting their druggability. The first step included data extraction in such a way that druggable and non-druggable protein sequences were taken from valid databases such as DrugBank and UniProt, and a balance was established between positive and negative data. Physiochemical features such as amino acid composition (AAC), composition of k-spaced amino acid group pairs (CKSAAGP), Moran correlation, etc. are extracted using appropriate methods and functions . In the next step, by the use of support vector machine (SVM), k-nearest neighbors (KNN), multilayer perceptron (MLP), and other classification algorithms, the belonging of each protein sequence to either the druggable or non-druggable group was determined. Results were measured based on various performance metrics. The best algorithm among the methods used was SVM with a polynomial kernel, which obtained an accuracy value of 90 percent, which indicates that the selected features could be a good substitute for the features obtained from previous studies. In conclusion, results obtained from this method are promising and this approach could be expanded for more extensive data in future studies.

Keywords: Drug Development, Druggable Proteins, Druggability Potential, Supervised learning, Binary classification

ICB11-1401-12-1010

Feature set prioritization of metagenomes for predicting colorectal cancer based on multiple human gut microbiome datasets

Zahra Tavakkol-hamedani[1], Sayed-amir Marashi[2*], Kaveh Kavousi[3]

[1]University of tehran. [2*]University of tehran.[3]University of tehran.

*Corresponding author(s). E-mail(s): marashi@ut.ac.ir;

Abstract

Colorectal cancer is one of the most prevalent cancers in the world and based on various studies, this cancer is associated with gut microbiome dysbiosis. One step in developing diagnostic methods is understanding cancer mechanisms by finding the key features changing during cancer. However, independent studies have reported different results in finding the best feature sets for classifying samples into groups, due to differences in the quality of datasets, sampling method, sample size, and analyzing methodologies. To address this issue, we defined a combined classification-robustness score and then employed a genetic algorithm to find the feature set that maximizes this score. This feature set is considered robust because it gives a high classification score when applied to new, unseen datasets. The classification score measures the strength of classifying groups by calculating separability measures, and the robustness score reflects the level of consistency of selected features across various input datasets. In addition to taxonomic profiling, which provides insight into the cancer-associated species, we also achieved a functional perspective by extracting features from seven public whole-metagenome sequencing datasets of colorectal cancer at both the species and functional levels (i.e., KEGG Orthology group, Enzyme Commission number, and reaction). We used the genetic algorithm to prioritize feature sets and found that exploiting multiple datasets and different feature types resulted in higher classification and robustness scores compared to using a single dataset or feature type. In conclusion, our study emphasizes the importance of combining datasets and feature types for a more robust and effective classification of colorectal cancer.

Keywords: Colorectal Cancer, Microbiome, Genetic Algorithm, Robust Features

ICB11-1401-12-1011

Antimicrobial resistance prediction in Acinetobacter baumannii using collaborative matrix factorization

Zahra Seraj[1], Fatemeh Zare-mirakabad[2*]

[1]Amirkabir university of technology. [2*]Amirkabir university of technology.

*Corresponding author(s). E-mail(s): f.zare@aut.ac.ir;

Abstract

Background:
Antimicrobial resistance (AMR) is a phenomenon which enables bacteria to survive against antibiotics and become resistant to them. AMR is one the most serious crises of the 21st century and must be controlled quickly. In order to control this crisis, it is necessary to determine AMR phenotype of a bacterium to a specific antibiotic. It is possible to answer this question through laboratory methods, but it is often cheaper and faster to use computational methods.
Methods:
In this article we proposed a collaborative matrix factorization (CMF) model to predict AMR phenotype of 845 different strains of A. baumannii bacteria to 12 different antibiotics. Basic assumption of CMF model is that there exists a low-dimensional representation of bacteria and antibiotics which makes it possible to model AMR phenotype accurately. The purpose of CMF model is to find that d-dimensional latent feature space and map both bacteria and antibiotics to this space. To predict AMR phenotype of a pair, CMF model uses inner product of their latent feature vectors. As a feature vector to represent each bacterium, a binary vector of gene's presence/absence pattern, and for the case of each antibiotic, fingerprint representation was considered. We have used three matrices to create this CMF model; matrix of known bacterium-antibiotic phenotypes (phenotype matrix), bacteria similarity matrix and antibiotic similarity matrix. In order to create bacteria (antibiotic) similarity matrix, similarity of a pair of bacteria (antibiotics) is estimated by similarity of their corresponding feature vectors. Finally, proposed model outputs a predicted phenotype matrix.
Conclusion:
The proposed CMF model predicted a phenotype matrix, which determines AMR phenotype of each strain to each of 12 drugs. The resulting model was evaluated using 5-fold cross validation and achieved 81.5% accuracy, 87.3% sensitivity and 80% area under ROC curve (all in terms of mean cross-validation scores).

Keywords: Antimicrobial resistance, Collaborative matrix factorization, Acinetobacter baumannii

ICB11-1401-12-1012

Improving prediction of drug-target interactions based on fusing multiple features with data balancing and feature selection techniques

Hakimeh Khojasteh[1], Jamshid Pirgazi[2*]

[1]Department of computer engineering, university of zanjan, zanjan, iran. [2*]University of science and technology of mazandaran.

*Corresponding author(s). E-mail(s): j.pirgazi@mazust.ac.ir;

Abstract

Predicting drug-target interaction (DTI) is an important research area in drug discovery. It means identifying the interaction between chemical compounds and protein targets. Wet lab experiments to explore these interactions are expensive as well as time-consuming. On the contrary, a dry lab environment focusing more on computational methods of interaction prediction can be helpful to limit the search space for wet lab experiments and give clues before developing a new medicine. This paper proposes a novel drug-target interaction prediction method called SRX-DTI. First, we extract various descriptors from protein sequences, and the drug is encoded as an FP2 fingerprint. Besides, we present the One-SVM-US technique to deal with imbalanced data. We also developed the FFS-RF algorithm, a forward feature selection algorithm, and coupled it with a random forest (RF) classifier to maximize the predictive performance. This feature selection algorithm removes the irrelevant features to obtain optimal features. Finally, the balanced dataset with optimal features is given to the XGBoost classifier to identify DTIs. The experimental results demonstrate that our proposed approach SRX-DTI significantly achieves higher performance than other existing methods for predicting DTIs. The datasets and source code are available at: https://github.com/Khojasteh-hb/SRX-DTI .

Keywords: drug-target interaction, protein sequence, molecular fingerprint, data imbalance, feature selection, random forest, XGBoost

ICB11-1401-12-1013

Molecular design of sgRNA based on CRISPR-Cas13a for specific detection of ORF1ab region in SARS-COV-2 genome infecting Iranian patients

Fatemeh Farrokhi[1], Fatemeh Akbarian[2*], Mohammad hossein Sanati[3], Hanieh Motahari rad[4]

[1]Ale -taha institute of higher education. [2*]Ale- taha institute of higher education.[3]Nigeb. [4]Royan institute for reproductive biomedicine.

*Corresponding author(s). E-mail(s): f.akbarian1@gmail.com;

Abstract

COVID-19 has become a global SARS-CoV-2 coronavirus pandemic, due to high strength and transmission speed of the virus . Early and accurate diagnosis is necessary to prevent the infectiousness along with avoiding progression of the disease in patients. The emerging diagnostic tools based on Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), function through the binding of Cas proteins to a target region of the viral nucleic acid sequence, with the most conservation and the least mutation rate. The genomic sequence of RNA dependent RNA polymerase (RdRp) is a suitable option for targeting Cas protein. The aim of this study is to apply bioinformatics tools to investigate the conserved regions of the RdRp gene of SARS-CoV-2 genome and molecular design of sgRNA based on CRISPR-Cas13a, in order to specifically and quickly diagnose through the ORF1ab gene region in Iranian patients with COVID-19. Therefore, the conserved regions of RdRp sequence were identified using MAFFT multiple sequence alignment program (version 7) , Jalview (version 2.11.2.3) and BioEdit (version 7.2.5.), through multiple alignment of the reference Wuhan-Hu-1 genome (NC_045512.2) with 470 complete SARS-CoV-2 genomes in Iranian patients with Covid-19 obtained from the GISAID database. Thereafter, in order to design sgRNA, the target sequences were obtained from the conserved regions of the ORF1ab sequence using the ADAPT online software. As a results, the 28-nucleotide sequence 5'-TGGTGTACTGACATTAGATAATCAAGAT-3' was selected, showing 100% similarity with SARS-CoV-2 genome and no significant similarity with MER-CoV and SARS-CoV. Target-specificity of sgRNA template was guaranteed by the design of highly specific recombinase polymerase amplification (RPA) forward and reverse primers. In conclusion, the findings of this study resulted in the computational design of an efficient sgRNA, as the first step of a mega project leading to development of a SARS-CoV-2 diagnostic kit.

Keywords: Diagnosis, SARS-CoV-2, RdRp, CRISPR, ADAPT, MAFFT

ICB11-1401-12-1014

Characterization of chemotherapy resistant triple-negative breast cancers at single-cell resolution

Fatemeh Ahmadimoghari[1], Ioanna Mavrommati[2], Yu Zhang[3], Ioannis Roxanis[4], Syed Haider[5*], Rachael Natrajan[6]

[1]The institute of cancer research. [2]The institute of cancer research. [3]The institute of cancer research. [4]The institute of cancer research. [5*]The institute of cancer research.[6]The institute of cancer research.

*Corresponding author(s). E-mail(s): syed.haider@icr.ac.uk;

Abstract

Triple-negative breast cancer is a highly aggressive and heterogeneous disease with an average survival of less than 50% and a median progression free survival of 1.7 months for women who have therapy resistant metastatic disease . Although immunotherapy is showing promising results in early triple-negative breast cancer , many patients do not respond, and chemotherapy remains the main treatment option . High rates of recurrence in triple-negative breast cancer are in part due to its inherent molecular heterogeneity and sub-clonal diversity, in which cells present in minority sub-populations escape the pressures of therapy . Here we sought to characterise the transcriptomic sub-clonal diversity of chemotherapy resistant triple-negative breast tumors. Sixteen primary untreated triple-negative breast tumours (8 chemosensitive, and 8 chemoresistant, who died of their disease within three years of initial diagnosis) were sequenced using the single nuclei RNA-seq to characterise the inter- and intra- sample transcriptomic heterogeneity. Unsupervised meta-clustering of epithelial cells in G1 phase of cell cycle from all samples identified three clusters that were shared between chemoresistant and chemosensitive patients, six clusters that were dominated by cells from chemoresistant, and ten clusters dominated by cells from chemo-sensitive patients. Top significant markers of each cluster were tested for enrichment in GO biological process and pathway databases using the Fisher's exact test (hypergeometric distribution). The meta-clusters based on enriched biological process illustrated a clear distinction between chemoresistant-dominant and chemosensitive-dominant meta-clusters. For instance, chemoresistant-dominant meta-clusters exhibited enrichment for biological functions related to the interaction of tumour cells with immune cells. Further, some meta-clusters were enriched for known markers of disease aggression such as the EMT, ERBB2 and tumour hypoxia pathways.

Keywords: Triple-negative breast cancer, chemoresistant, single-cell, sub-clonal heterogeneity

ICB11-1401-12-1015

Comparative gene co-expression network analysis between adenocarcinoma and squamous cell carcinoma subtypes of esophageal cancer

Nazanin Hosseinkhan[1*], Mohammad ebrahim Khamseh[2]

[1*]Endocrine research center, institute of endocrinology and metabolism, iran university of medical sciences.[2]Endocrine research center, institute of endocrinology and metabolism, iran university of medical sciences.

*Corresponding author(s). E-mail(s):

Abstract

Background:
Individuals with higher Body Mass Index (BMI), are prone to developing the less aggressive subtype of esophagus cancer (ESCA) known as esophagus adenocarcinoma (EAC) . On the contrary, in normal weights, the aggressive form of ESCA, esophagus squamous carcinoma (ESCC) is more plausible. This is in contrast with what is commonly reported in other cancers. In this study, we investigated this paradox by conducting a co-expression analysis on two subtypes of ESCA; EAC, and ESCC.
Method:
Expression data in the form of FPKM, were retrieved from respectively 23 and 54 normal-weight and 59 and 18 overweight/obese EAC and ESCC individuals from the cancer genome atlas (TCGA) database. Co-expression analysis was separately conducted on EAC and ESCC subtypes using the WGCNA package in R 4.0.0.
Results:
One module of 34 co-expressed genes were identified with a correlation with BMI trait in EAC patients. Pathway enrichment of these co-expressed genes resulted in the identification of Proteasome (PSMA4; PSMB1), cell cycle (CCNB2; SMC1A), and cellular senescence (CCNB2; MYBL2) associated genes beside several metabolic associated pathways. The presence of co-expression between these genes, in correlation with higher BMI in EAC, which was not detected in the ESCC subtype, indicates their probable protective roles in EAC patients from the more severe form of ESCA.
Conclusion:
The presence of the identified co-expressed module in ESCA with the probable protective role, which was not detected in the ESCC subtype, is a partial indicator of different mechanisms of cancer development in these two ESCA subtypes. Inspecting other co-expressed modules in EAC and ESCC will give us more clues about the difference in the overall mechanisms of cancer development in these two ESCA subtypes. This can be resulted in applying distinct strategies for the clinical management of EAC and ESCC subtypes of ESCA.

Keywords: Esophagus cancer, Adenocarcinoma, Squamous Cell Carcinoma, Obesity paradox, Co-expression network

ICB11-1401-12-1016

Single-cell RNA sequencing data analysis using Explainable Artificial Intelligence identified key transcriptional factors for early COVID-19 severity prediction

Edris Hosseinigol[1], Adib Miraki feriz[2], Benyamin Ghahremani nezhad[3], Mohammad Ghasemigol[4], Hossein Safarpour[5*]

[1]Department of computer engineering, birjand university, birjand, iran. [2]Birjand university of medial sciences. [3]Amirkabir university of technology. [4]Birjand university. [5*]Birjand university of medial sciences.

*Corresponding author(s). E-mail(s):

Abstract

Early and precise diagnosis of SARS-CoV-2 infection can be critical for managing the COVID-19 pandemic. But one of the challenges with this disease's administration is the lack of standardized methods for diagnosis and the incapability to guesstimate prognosis based on clinical features. Explainable Artificial Intelligence (XAI) has been one of the most widely discussed topics in recent years. It is widely used to tackle various issues across multiple domains in the medical studies. Here, we have applied SHAP algorithm on a single cell RNA-Seq data to find hub genes in different stages of COVID-19 pathophysiology. Firstly, a single-cell RNA-seq dataset had downloaded under the GEO number of GSE165080. To evaluate data, the Scanpy pipeline was employed. This data contains the samples of normal, severe, asymptomatic, severe recovery and moderate COVID-19 patients. Single-cell RNA data processing was performed on cells that met the following criteria: (I) Each cell must contain at least 500 genes and no more than 60000 gene expression counts. (II) Less than 20% of the genes counted must be mitochondrial genes. The scran package on Bioconductor was used to calculate normalized expression. At the next step, SHAP algorithm was used to find most hub genes based on SHAP value in each status. Finally, the hub genes were imported in MsigDB to find their enrichment and DGiDB to construct their related drugs. Our study revealed some transcriptional signature measurable in blood samples, which discriminated between healthy people and COVID-19 positive patients and showed predictive value for later severity of COVID-19 symptoms. This type of approaches could, by employing standard hospital laboratory analyses of patient blood, be utilized to identify, COVID-19 patients at high risk of mortality.

Keywords: Explainable Artificial Intelligence, Single-Cell RNA-sequencing, COVID-19, SHAP

ICB11-1401-12-1017

A Plastic-Contaminated Soil Gene Catalog Constructed by Metagenomic Approaches

Donya Afshar jahanshahi[1], Kaveh Kavousi[2*], Shohreh Ariaeenejad[3]

[1]University of tehran. [2*]University of tehran.[3]Department of systems and synthetic biology, agricultural biotechnology research institute of iran (abrii), agricultural research education and extension organization (areeo), karaj, iran.

*Corresponding author(s). E-mail(s):

Abstract

Plastics are widely used worldwide for numerous goals, and polluted marine and terrestrial environment by plastics has become an intense ecological obstacle. Plastic recycling is very tardy, and polymer residues harm different environments, such as humans and animals. The roles of plastic-exposed environment microbiome are conclusive for identifying microbes and enzymes to advance plastics degradation. Combining computational methods and metagenomic approaches is crucial to increase the current plastic degradation methods. Gene, protein and metagenome-assembled genomes catalogs simplify taxonomical and functional profiling of microbiome samples . Based on our knowledge, no integrated gene catalogs from plastic-contaminated environments exist. In this study, 66 whole metagenome samples (including soil , , leachate, sludge , and plastic fabrics samples) were combined from four exiting metagenome data with our new sample. A mean library fragment size of 365 Gb of high-quality data was obtained; this data size is excellent than all of the used plastic-contaminated samples. A plastic-contaminated soil gene catalog (PCSGC) and a plastic-contaminated soil protein catalog (PCSPC) were generated by expanding a computational workflow composed of quality control of reads, trimming, assembly, and binning of contigs. These catalogs contain 53.3 million nonredundant genes and proteins. Also, glycoside hydrolases (GH) and Carbohydrate-Binding modules (CBM) were the most frequent carbohydrate-active enzymes (CAZymes) that were generated using dbCAN2 software. dbCAN2 resulted in the 205,066 CAZyme-encoding genes in the PCSGC, which belong to 51 CAZyme subclasses. Our workflow and generated catalog of plastic-contaminated samples can help researchers to expand their knowledge of plastic-associated genes and proteins with minimal time and cost.

Keywords: Gene catalog, Genome, Plastic-contaminated soil

ICB11-1401-12-1018

De novo Antibody design against PD-1 and PDL-1 as critical immune check points involve in cancer treatment

Sarah Abdolmaleki[1], Mohamad reza Ganjalikhany[2*], Mazdak Ganjalikhani hakemi[3]

[1]Department of cell and molecular biology and microbiology, university of isfahan. [2*]University of isfahan.[3]Isfahan university of medical sciences.

*Corresponding author(s). E-mail(s): m.ganjalikhany@sci.ui.ac.ir;

Abstract

Cancer control and treatment is an important issue in therapy, which recently have immune checkpoints considered by cancer researchers. Programmed cell death 1 (PD-1) is an immune checkpoint protein on T cells and it plays an important role in the interaction between tumors and T cells . The interaction of PD-1 with its ligands PD-L1 and PD-L2 on cancer cells, leading to inhibition of the effector function, and apoptosis of T cells. Furthermore, blocking the interface area of the interaction PD-1/PD-L1 results in the normalization of antitumor response . In this study, high-specific binding antibodies have been designed using de novo method to block the immune checkpoints with the aim of cancer treatment. We used computational methods such Rosetta protein design package along with developing new codes to enhance the design process. At first, thousands of de novo designed antibodies have been generated and then the models with higher stability, affinity and shape complementarity values were selected. Then, MD simulation and MM-PBSA were used to measure the dynamics and stability as well as the binding affinity of each final design. The results obtained in this study showed that the designed antibodies bind selectively and with high potency to the PD-1 receptor. We designed 25400 antibodies for different regions of PD-1 and PDL-1, docked them onto the targets, and identified %19.82 and %30.6 high-affinity antibodies, respectively. The PD-1 structure has flexibility loops that adopt variable conformations by binding to antibodies, implying that these loops contribute to the binding affinity for antibodies and provide a “hotspots” region for Immune checkpoint therapy. The results demonstrated that mutation and optimization of the residues at the interface area of the CDRs, it helps the stability and high-affinity models.

Keywords: De Novo Design, Antibody, Immune checkpoints, Protein-Protein Interactions

ICB11-1401-12-1019

Improvement of peptide-HLA class I prediction using transformers

Mahsa Saadat[1], Fateme Zare-mirakabad[2*], Nazanin Hosseinkhan[3]

[1]Amirkabir university of technology. [2*]Amirkabir university of technology.[3]Iran university of medical sciences.

*Corresponding author(s). E-mail(s): f.zare@aut.ac.ir;

Abstract

Since Human leukocyte antigen (HLA) can bind foreign peptides to present them to specialized immune cells and initiate an immune response, accurate prediction of binding between HLA and neoepitope is critical for target identification in immunotherapy . However, current algorithms for predicting neoantigens are resulting in high false positives and most of them are limited by fixing model input length. In this study, we proposed an allele-specific and transformer-based model to predict antigen presentation in the context of HLA class I alleles. This model benefits from ProtBERT which is a pre-trained transformer on proteins to encode peptides and does not need to fix peptide length. Then, we use random forest and multilayer perceptron networks as a classifier on encoded peptides. The dataset we used was obtained from the immune epitope database (IEDB) as in previous works and includes peptide and HLA pairs of 20 high-frequency HLA-A and HLA-B allotypes. Results show that our proposed model outperforms the former methods in terms of positive predictive value (PPV) (0.58 vs.0.38 on average). Our best result was obtained on HLA-A*01:01 with a PPV value of 0.872, which is 0.286 and 0.267 in APPM and netMHCpan-4.0 models, respectively. Also, using a pre-trained encoder allows the model to predict more quickly and with less computational effort. On the other hand, these results confirm that transformers can be used as an embedding that extracts structural properties from the sequence. Since our model only requires the peptide sequence of HLA-peptide binding pairs, it can be applied to other binding problems without the need for structure data.

Keywords: Neoantigen, Vaccine design, Machine learning, Bioinformatics

ICB11-1401-12-1020

Molecular docking and ADME studies of natural compounds against prime targets of HIV

Ali Afshar[1], Samad Nejad ebrahimi[2*]

[1]Shahid beheshti university. [2*]Shahid beheshti university.

*Corresponding author(s). E-mail(s): s_ebrahimi@sbu.ac.ir;

Abstract

The human immunodeficiency viruses (HIV) are two species of Lentivirus (a subgroup of retrovirus) that infect humans. They cause acquired immunodeficiency syndrome (AIDS), in which progressive immune system failure allows life-threatening opportunistic infections and cancers to thrive. AIDS is multifaceted, and this underlying complexity hampers its complete cure. The toxicity of existing drugs and the emergence of the multidrug-resistant virus worsen the treatment. The development of effective, safe and low-cost anti-HIV drugs is among the top global priority. Exploration of natural resources may give a ray of hope to develop new anti-HIV leads. Among the various therapeutic targets for HIV treatment, reverse transcriptase, protease and integrase receptors is the prime focus. In the present study, we predicted potential plant-derived natural molecules for HIV treatment using a computational approach (molecular docking, in silico ADMET and drug-likeness) to inhibit the effects of HIV. Receptor-ligand binding studies were performed using Schrodinger. Seventeen natural product-based compounds were selected from several natural compounds by pharmacophore screening from the PubChem database and docked against the HIV targets. In this study, the Glide docking program was applied and extra precision (XP) was used. Docking scores revealed that mulberroside A (-9.744, -9.868, -8.958 Kcal/-mol), chlorogenic acid (-11.566, -11.360, -8.406 Kcal/mol) and gallic acid (-7.049, -7.533, -7.193 Kcal/mol) are promising candidates that bind with multi-targets of HIV, while Caffeic acid, Curcumin, and Silymarin were target-specific candidates. From molecular docking results, we have identified few potent molecules of natural origin against identified targets, which may give new drugs to combat HIV infection after wet lab validation.

Keywords: HIV, Molecular Docking, ADMET, Natural Compound

ICB11-1401-12-1021

Investigation of Some Nanobodies against SARS-C0V-2 Spike protein by immune-bioinformatics tools

Ali Johari[1*], Mohammadmehdi Ranjbar[2], Saeideh Samareh mousavi[3], Soudabeh Niknia[4]

[1*]Kavian.[2]Razi. [3]Kavian. [4]Kavian.

*Corresponding author(s). E-mail(s): alijohari73@gmail.com;

Abstract

Antibody-antigen interactions and understanding binding affinity for predictive computational Modeling is a main part of the drug discovery process, ensuring that therapeutic antibodies that bind their targets selectively and specifically. Following the SARS-CoV-2 pandemic, Nanobodies were selected as an effective particle in the treatment of COVID-19. Small size, simple structure, ease to use and relatively low cost, low immunogenicity, and high affinity, make them special in research fields such as treatment and diagnosis. We selected 17 nanobodies from PDB and investigated their structure and function in relationship to spike glycoprotein, among them, some are synthetic antibody (sybody) and we focused on their IC50 and KD. IC50 is the concentration of neutralizing antibody that displaces 50% of the specifically bound labeled antigen and KD is the equilibrium dissociation constant, a calculated ratio of Koff / Kon, between the antibody and its antigen. KD is inversely proportional to affinity, so the lower the KD value (the lower the concentration), the higher the affinity of the antibody. Among 17 nanobodies, only 2 nanobodies pose either low IC50 or low Kd in the amount of pico mol. According to structural studies the conformational flexibility of the SARS-CoV-2 trimer S protein allows each of its RBD to exist in two main conformations: up and down. the "up" conformation is easily bound by ACE2 and most neutralizing antibodies, so these antibodies compete with ACE2 to bind to the S protein of RBD in an upward conformation thereby preventing viral infection. The "down" conformation is not easy to be accessed by ACE2 or most neutralizing antibodies. Nanobodies as a small particle with high affinity can bind to and stabilize the "down" conformation of the S protein, thereby preventing the conformational changes required for the virus to enter the host.

Keywords: nanobody, affinity, SARS-CoV-2

ICB11-1401-12-1022

Selection of the most effective SARS-C0V-2 proteins in order to design a recombinant vaccine by immune-bioinformatics tools

Soudabe Niknia[1*], Saeed Zibaei[2], Mohammadmehdi Ranjbar[3], Neda Amini[4], Ali Johari[5]

[1*]Kavian.[2]Razi. [3]Razi. [4]Kavian. [5]Kavian.

*Corresponding author(s). E-mail(s): soudabe.niknia@gmail.com;

Abstract

The largest RNA genome among all the RNA viruses is found in Coronavirus. The genome of SARS-CoV-2 encodes several proteins. Most of them are non-structural proteins (NSP) except envelop (E), membrane (M), nucleocapsid (N), and Spike (S) proteins. The M and E proteins are necessary for virus assembly. The spike glycoprotein is an essential need for attachment to host cells and the nucleocapsid protein plays a main role in the transcription and replication of viral RNA and interference with cell cycle processes of the host cell. The spike glycoprotein is located on the surface where the RBD (RNA-binding domain) mediates the interaction with angiotensin-converting enzyme 2 (ACE2). The RBD is a globular domain situated on the distal surface of the spike protein. Two conformations have been observed in the stabilized trimer. One conformation where one RBD is ACE2 accessible while the other two are not, and one conformation where all three RBDs are down. According to the in vitro studies, the spike glycoprotein and the nucleocapsid protein have high antigenic and immunogenic activity and are highly expressed during infection. The development of diagnostics, therapeutics, and precautionary tools is substantially rapid with unprecedented race as the greatest combat against the SARS-CoV-2 outbreak. We investigated some features in these proteins by immune- bioinformatics tools: self-tolerance and structural alignment by BLASTP and T-coffee; PI and MW by SWISS-MODEL and prediction of protective antigen by Vaxijen. It seems that both of them may be potential antigenic proteins for serodiagnosis of the COVID-19. Therefore, we suggested designing a novel synthetic protein vaccine consisting of multiple immune dominant B-cell epitopes from N and S proteins of SARS-CoV-2.

Keywords: SARS-CoV-2, spike, vaccine

ICB11-1401-12-1023

Predicting Adverse Drug Reactions Using Computational Methods: An Analysis of Drug and Adverse Reaction Features and Representations

Milad Besharatifard[1], Fatemeh Zare-mirakabad[2*], Zahra Ghorbanali[3]

[1]Amirkabir university. [2*]Amirkabir university.[3]Amirkabir university.

*Corresponding author(s). E-mail(s): f.zare@aut.ac.ir;

Abstract

Identifying and controlling adverse drug reactions is a complex problem in the pharmacological field. Despite the studies done in different laboratory stages, some adverse drug reactions are recognized after being released, such as Rosiglitazone. Due to such experiences, pharmacists are now more interested in using computational methods to predict adverse drug reactions. In computational methods, finding and representing appropriate drug and adverse reaction features are one of the most critical challenges. Here, we assess fingerprint and target as drug features; and phenotype and unified medical language system as adverse reaction features to predict adverse drug reaction. Meanwhile, we show that drug and adverse reaction features represented by similarity vectors can improve adverse drug prediction computational methods. This article proposes four frameworks to analyze drug and adverse reaction features and representations in drug-adverse reaction association prediction. Two frameworks are based on random forest classification and neural networks as machine learning methods called F_RF and F_NN, respectively. Rest of them apply matrix factorization methods by improving the CS and TMF models. They are extended by considering target as a drug feature and phenotype as an adverse reaction feature. However, machine learning frameworks with fewer drug and adverse reaction features are more accurate than matrix factorization frameworks. In addition, the F_RF framework performs significantly better than F_NN with ACC = %89.15, AUC = %96.14 and AUPRC = %92.9. Next, we contrast F_RF with some well-known models designed based on similarity vectors of drug and adverse reaction features. Unlike other methods, we do not remove rare reactions from the data set in our frameworks.

Keywords: Adverse drug reaction, Machine learning, Random Forest

ICB11-1401-12-1024

Reconstruction of miRNA-mRNA regulatory network in the pediatric Adrenocortical Adenoma

Pariya Mohammadyari[1], Amirhesam Pahlevani[2], Alireza Morovat[3], Maryam Talaei[4], Mohammad mehdi Naghizadeh[5], Maryam Bahmanyar[6], Maryam Maghbol[7*]

[1]Tehran university of medical science. [2]Fasa university of medical science. [3]Fasa university of medical science. [4]Fasa university of medical science. [5]Fasa university of medical science. [6]Fasa university of medical sciences, fasa, iran. [7*]Fasa university of medical science.

*Corresponding author(s). E-mail(s): maghbol.maryam@yahoo.com;

Abstract

Pediatric Adrenocortical Tumors (ACT) are a rare disease in children, and Adrenocortical Adenoma (ACA) is one of its benign forms . MicroRNAs (miRNAs) are small noncoding RNAs that have been shown to be potential biomarkers for a variety of diseases, including cancer . In this study, we aim to identify diagnostic biomarkers of pediatric ACA by transcriptome data analysis. First, based on the Gene Expression Omnibus database, gene expression profiles related to ACA were downloaded. Differentially expressed genes (DEGs) were analyzed with GEO2R tool. Then, a defined formula [-log(p.value) × | log (fold change) |] was applied to select the 1000-top significant probs. Protein-protein interaction (PPI) network was established through the STRING database. As well as using the Toppgene database for Gene Ontology (GO), we used it to select target miRNAs related to ACA hub genes. Next, the miRNA-mRNA regulatory network is reconstructed via Cytoscape software. After drawing the network, hub mRNA (FOS, SIRT1, ESR1, COL1A2, THBS1) and miRNA (miR-26a, miR-29c, miR-29b, miR-29a, let-7b) with the highest degree were determined. The most frequent top signaling pathway were IL-4, 13 and IL-18 signaling pathway. We reconstructed a miRNA-mRNA network that may show crucial functions in the ACA, thus providing potential diagnostic biomarkers and therapeutic targets for ACA.

Keywords: biomarkers, miRNA, Adrenocortical Adenoma

ICB11-1401-12-1025

The role of circRNA-miRNA-mRNA interaction network in rheumatoid arthritis using transcriptome data analysis

Ali Zeraatpishe[1], Fateme sadat Fattahian[2], Amir hesam Pahlavani[3], Fateme Yarian[4*]

*Corresponding author(s). E-mail(s): f.yarian@gmail.com;

Abstract

Background : Rheumatoid arthritis is the most common inflammatory joint disease that affects about 1% of the population. MicroRNAs (miRNAs) and circularRNAs(CircRNAs) are a group of ncRNAs that regulate gene expression at the post-transcriptional level. previous study have found the involvement of miRNAs and CircRNAs in the pathogenesis of RA. In this study, we used integrated bioinformatic analyzes with the aim of identifying genes, microRNAs, and biological pathways effective in the pathogenesis of rheumatoid arthritis.
Methods:
Gene expression profile of RA was downloaded from Gene Expression Omnibus database. Differentially Expressed Genes related to 47 healthy samples and 72 rheumatoid arthritis samples of synovial tissue were analyzed using the GEO2R tool. A defined formula [-log(p.value)×| log(fold change)|] was applied to select 3000 top significant probs. After filtering, removing duplicates and finding common probes between mRNA datasets, 20 relating miRNA‘s of Differentially Expressed Genes established through Toppfun database.Toppgene database was used for Gene Ontology and microRNA target prediction. Cytoscape software was used to identify hub microRNAs.
Results:
Total 20 Common Differentially Expressed Genes were identified in 6 datasets. By using miRNA target peridction through Toppfun database, these miRNA‘s had 54 in common with 164 Differentially Expressed miRNA‘s in our dataset. 6 hub genes (NFIB, ITGA4, CLIC5, AR, MYH11), and 7 microRNA hubs (miR30c, miR130b, miR301a, miR579, miR30d) based on degree were identified that were related to these circRNA‘s.
Conclusion:
Our study investigated circRNA-miRNA-mRNA interactions which have role in the pathogenesis of Rheumatoid arthritis that can help discovering new biomarkers for Early Diagnosis.

Keywords: rheumatoid arthritis, RA, circRNA, miRNA, mRNA, interaction network

ICB11-1401-12-1026

Investigating Eugenol-Lysozyme complex using docking and molecular dynamic simulation methods

Mohammad Basti[1], Fatemeh Rashno[2*], Saba Taheri[3]

[1]Islamic azad university. [2*]Islamic azad university.[3]Islamic azad university.

*Corresponding author(s). E-mail(s): fatemeh.rashno@iau.ac.ir;

Abstract

Amyloid fibrils are β-sheet rich protein aggregates with fibrillar morphology which are a challenge in amyloidosis and recombinant protein technology. Various small molecules inhibit fibrillation process of different proteins. In this study, the effect of eugenol, a natural ingredient of cloves, on hen egg-white lysozyme (HEWL), a model system for amyloid fibrillation, was assessed using molecular docking and molecular dynamic (MD) simulations. Blind docking, using AutoDock4, showed a binding energy of -4.99 kcal/mol. The critical amino acid residues for forming Eugenol-lysozyme complex by hydrogen bonding (Asp52 and Val109) and hydrophobic interactions (Gln57, Ile58, Asn59, Trp63, Ala107 and Trp108) were revealed. These residues constitute the aggregation prone region (APR) of the protein. Afterwards, the lowest docking energy structure was selected for more investigation using MD simulations lasting 200 ns using GROMACS2019.6 software and Amber99SB force field. Parameters such as root-mean square-deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg) and solvent accessible surface area (SASA) were evaluated. RMSD and RMSF of the enzyme in the presence of eugenol decreased in comparison to the enzyme alone. Moreover, the presence of eugenol increases the number of intramolecular hydrogen bonding of lysozyme. On the other hand, eugenol effect on SASA and Rg of the enzyme were insignificant. Thus, it seems that eugenol does not drastically alter the structure of the protein. In conclusion, eugenol showed a potential to interfere with the fibrillation process of lysozyme.

Keywords: Amyloid fibrillation, Eugenol, Lysozyme, Molecular dynamic, Small molecule inhibitors, Simulation

ICB11-1401-12-1027

Evaluation of TUBB3, PSMB7, CBFB, and COL9A1 genes involved in the development of Squamous Cell Lung Cancer (SCC)

Pouyan Asadi[1], Rabee Movagharnia[2], Armin Yahyaei[3], Danial Jafari[4], Elham Talebi[5*]

[1]Golestan university of medical sciences. [2]Varamin-pishva branch, islamic azad university. [3]Iran university of medical sciences. [4]Medicine faculty of islamic azad university of shahrood. [5*]Islamic azad university of tehran medical unit.

*Corresponding author(s). E-mail(s): elham.talebi95@gmail.com;

Abstract

Introduction:
Squamous Cell Lung Cancer (SCC) is considered one of the types of lung cancer, and this type is still the deadliest cancer in the United States and around the world. Numerous genes are involved in the development of this type of cancer. Detection of such genes can be a suitable tool for the diagnosis and treatment of squamous cell lung cancer in the future. The aim of this study was to identify the main genes (hub genes) related to SCC, whose expression can be related to the death rate of patients. Materials and Methods:
First, to study the main genes related to SCC, RNAseq data available in the TCGA database was used. Genes with significant expression changes (FDR<0.01, logFC>1) and associated with poor prognosis were selected. Enrichr database and MSigdb database were used to identify pathways related to candidate genes. To detect the hub genes, the Protein-Protein Interaction (PPI) network was used and those with the highest degree were selected among the candidate genes. The linear model method was used to investigate the difference in expression and the Cox regression test was used to investigate the relationship between gene expression and mortality rate. All pre-processing and analysis were done using R software. Results:
The expression difference findings showed 1883 genes had significantly increased expression in cancer samples compared to normal. Instead, 1421 genes were identified whose increased expression was associated with poor prognosis (HR>1, logRank<0.05). The results showed 175 genes are both significantly increased and associated with poor prognosis. The results for 175 identified genes showed that these genes are involved in the pathways of metastasis, glycolysis, and DNA repair. Conclusion:
This study suggests that TUBB3, PSMB7, CBFB, and COL9A1 genes play an important role in the pathogenesis of SCC and can be suitable diagnostic and therapeutic targets.

Keywords: Squamous Cell Lung Cancer (SCC), TCGA, Enrichr, Hub Genes, MSigdb

ICB11-1401-12-1028

Identification of the regulatory interactions related to gastric cancer via integrated bioinformatics analysis

Kiarash Zare[1], Amir hesam Pahlevani[2], Sepideh Niknejad[3], Zahra Azadpour[4], Mohammad mehdi Naghizadeh[5], Ali Ganbariasad[6*]

[1]Student research committee, fasa university of medical sciences, fasa, iran. [2]Student research committee, fasa university of medical sciences, fasa, iran. [3]Student research committee, fasa university of medical sciences, fasa, iran. [4]Student research committee, fasa university of medical sciences, fasa, iran. [5]Noncommunicable diseases research center, fasa university of medical sciences, fasa, iran. [6*]Fasa university of medical sciences.

*Corresponding author(s). E-mail(s): alighanbariasad@gmail.com;

Abstract

Gastric cancer (GC) is a significant global health issue due to being the third-leading cause of cancer deaths worldwide. MicroRNAs (miRNAs) are small non-coding RNAs that post-transcriptionally regulate gene expression and play a role in pathogenesis of many cancers, including gastric cancer. In this study, we aim to Identifying the molecular signatures and specific biomarkers of GC using integrated bioinformatics analysis. The GC-related gene expression profiles were extracted from the Gene Expression Omnibus database. The identification of the differentially expressed genes (DEGs), was performed by the GEO2R tool. Toppfun database was used to investigate the Gene Ontology and prediction of miRNAs targets mRNAs. Hub genes were identified based on the protein-protein interactions (PPI) constructed in the STRING database with Cytoscape software. Centrality metrics such as degree, betweenness, closeness, and eigenvector centrality determined hub genes in PPI. Then, miRNA-mRNA reconstructed with Cytoscape. A total of 2000 DEGs were identified between six GC datasets. ITGB1, HIF1A, SOX2, STAT3, EGFR, FN1, and BRCA1 were selected as hub genes in PPI. The majority of the DEGs were enriched in Interleukin-4 and 13 signaling, RAC1/PAK1/p38/MMP2 pathway, Signaling by Interleukins. In the miRNA-mRNA bipartite network let-7b, miR-10a, miR-144, miR-17, miR-153, miR-1290 were found as hub. The hub miRNAs (let-7b and miR-10a) in the mRNA-miRNA bipartite network played a key role in GC progression; however, these findings need further investigation.

Keywords: mRNA-miRNA bipartite network, microRNA, bioinformatic analysis, gastric cancer

ICB11-1401-12-1029

A Transformer based mutation prediction algorithm

Khashayar Ehteshami[1], Seyed amirali Ghafourian ghahremani[2], Kaveh Kavousi[3*]

[1]Sharif university of technology international campus kish island. [2]Sharif university of technology international campus kish island. [3*]University of tehran.

*Corresponding author(s). E-mail(s): kkavousi@ut.ac.ir;

Abstract

Protein mutation prediction is the process of identifying the consequences of amino acid sequence alterations in a protein. It may be used to estimate the possible influence of genetic alterations on the function of a protein, as well as to discover disease-causing mutations or mutations that may impact therapeutic efficacy. Mutation prediction is often performed using machine learning models that analyze the functional impact of mutations based on reference data from evolutionary conservation, protein structure, and population genetics. Support Vector Machines (SVMs), Decision Trees, and Artificial Neural Networks are the most frequently used mutation prediction techniques. In addition, more advanced models, such as the Transformer model, have been used to predict mutations . In this work we developed a Transformer based mutation prediction algorithm to generate new mutants protein sequences based on a sequence that have already seen in nature which can be called a parent sequence. A transformer-based mutation prediction method is a machine learning model that predicts the impact of mutations on protein sequences using a transformer architecture. Transformers are predominantly employed for natural language processing tasks including language translation, question answering, as well as text summarization. The transformer architecture is based on an attention mechanism, which enables the model to concentrate on certain input components while generating predictions. This enables the model to better comprehend the input's context, resulting in more precise predictions. From language translation to protein mutation prediction, transformers have been utilized for a variety of purposes. For the results, we discover out which sequence seems to be more probable to mutant in the future with the help of bioinformatic methods such as BLOSUM Score, and we may utilize these findings for preventing a potential pandemic.

Keywords: Mutation Prediction, Transformer Models, Deep Learning, Machine Learning

ICB11-1401-12-1030

The Possible Role of Heat-shock proteins (HSPs) on Inflammatory Bowel Disease (IBD): Peptide Motif Analysis as an Immunoinformatic Technique

Mahsa Alem[1*], Seyyed meysam Abtahi froushani[2], Nasim Hajighahramani[3]

[1*]Urmia university.[2]Urmia university. [3]Ardabil university of medical sciences.

*Corresponding author(s). E-mail(s): Mahsa.Alem@gmail.com;

Abstract

Inflammatory bowel disease (IBD) is classified as Ulcerative Colitis and Crohn's Disease, both of which affect the gastrointestinal system and are caused mainly by gut dysbiosis. There is a lack of data available regarding inflammatory bowel disease pathogenesis. This encouraged us to look into whether or not any protein of microorganisms has a potential relationship in the immunopathogenesis of inflammatory bowel disease. Therefore the bioinformatic evaluation of host-microbe interactions would be a crucial approach in identifying the mentioned target. Using the basic local alignment search tool for protein (BLASTP) tool from National Center for Biotechnology Information (NCBI) and T-coffee expresso, the proteome of the mentioned microorganisms of the literature review linked to IBD was analyzed for protein sequences with identities exceeding 35%. The phylogenetic tree and degree of relationship between the protein sequences of microorganisms and the human proteome were determined using Multiple Sequence Alignment from Protein Data Bank (PDB). In this study, Campylobacter, Clostridium, Escherichia coli, Klebsiella, Listeria, Mycobacterium, Salmonella, Shigella, and Yersinia were identified as microorganisms associated with IBD. This study found traces of molecular mimicry (molecular-level similarity between microbial antigens and host proteins). Due to their high conservation in evolution, the 60, 70, and 90 kilo Dalton heat shock proteins (HSP) of the microorganisms and humans were identified as possible molecular targets, followed by autoreactive T lymphocytes against heat shock proteins in humans. It can be a possible pathogenesis in IBD through a dysbiotic gut microbiome. Finally, cytotoxic T lymphocytes and helper T lymphocyte epitopes with high homology between 60, 70, and 90 kilo Dalton heat shock proteins were extracted with Immune Epitope Data Base (IEDB) tool. Last but not least, by using in silico immunoinformatic approach, this study supports the concept that bacteria and the human proteome likely share many cross-reactive T cell epitopes.

Keywords: Immunoinformatic, Molecular mimicry, Auto-inflammatory, Heat-shock proteins (HSPs), Inflammatory Bowel Disease (IBD)

ICB11-1401-12-1031

Differential Gene Expression in Rat Middle Cerebral Artery Occlusion Model

Zahra Maharat[1], Abolghasem Esmaeili[2*]

[1]The university of isfahan. [2*]The university of isfahan.

*Corresponding author(s). E-mail(s): abesmaeili@yahoo.com;

Abstract

Ischemic stroke is a leading cause of global mortality and morbidity. Immediately following the event, neurovascular reperfusion to the infarcted area can injure the tissue, resulting in more serious disability. The discovery of differentially expressed genes and their signal pathways between MCAO and control rats can help to a better understanding of all biochemical mechanisms involved in this process. RNA-seq expression data of the GSE154098 series were used for analysis. Differentially expressed genes (DEGs) analysis, between groups (MCAO and Sham) with three biological replicates, was performed by DESeq2 and threshold with $p<0.05$ and absolute values of log 2Fold change >1 and <0.5. the STRING and Cytoscape were used to map protein-protein interaction and to visualize the network, respectively. Pathways and ontologies were performed by Enrichr. We identified 196 DEGs: 101 upregulated and 95 downregulated ($p<0.05$) and 9 DEGs: 6 upregulated and 3 downregulated ($p.adj<0.05$). DEGs were represented in the volcano plot by plotMA. The top 10 up-and-down regulated DEGs were listed. Gene ontology analysis revealed that the majority of the most enriched and meaningful biological processes and molecular function terms were mainly involved in immune and inflammatory response as well as cytokine receptor activity. Also, enriched were related to immune and neural-related pathways. The molecular mechanisms of Differentially expressed genes in cerebral hemispheres from MCAO and control rats may involve the PI3K-Akt, IGF, Rap1, BDNF, Microglia Pathogen Phagocytosis signaling Pathways, and Neutrophil Degranulation. The interaction of proteins in these pathways may be potential key targets for a better understanding of the pathological and underlying mechanism of protein-coding mRNAs that occur after MCAO.

Keywords: Ischemic stroke, MCAO, DEGs, DESeq2, Ontologies

ICB11-1401-12-1032

Identification of key genes and signaling pathways involved in bovine tuberculosis challenged with Mycobacterium bovis: Comparison of RNA-Seq and microarray expression profiling data

Farzad Ghafouri[1], Masoumeh Naserkheil[2], Vahid Dehghanian reyhan[3], Mostafa Sadeghi[4], Herman W. barkema[5], Sonia Zakizadeh[6*]

[1]University of tehran. [2]National institute of animal science of korea. [3]University of tehran. [4]University of tehran. [5]University of calgary. [6*]Animal science research institute of iran.

*Corresponding author(s). E-mail(s): sonia_zaki@yahoo.com;

Abstract

In recent years, advances in high-throughput RNA-Seq technologies have made it possible identifying differentially expressed genes between healthy and diseased tissues, better understanding of host–pathogen interaction, increased precision and sensitivity for the quantification of lowly expressed transcripts, detection of expressed coding and regulatory DNA sequence variants, and the evolution of gene regulation in different species (Ozsolak and Milos, 2011; McGettigan, 2013). Bovine tuberculosis (BTB), caused by infection with Mycobacterium bovis, is an intracellular pathogen belonging to the Mycobacterium tuberculosis complex, which is classified as the fourth most important disease of livestock in terms of zoonotic and economic impact worldwide (Wirth et al., 2008). Hence, characterization of the host response to M. bovis infection is essential for understanding the immunopathogenesis of the disease and for developing better control strategies. In this study, we analyzed and compared gene expression profiles generated using RNA-Seq and microarray data to examine transcriptional differences between healthy cows and M. bovis-infected cows. The results of the comparative analysis demonstrated that microarray detected a larger number of differentially expressed genes (1336) relative to the RNA-Seq (327) based on fold change $> \pm 2$ and false discovery rate < 0.05, of which 68 genes were common to both technologies and demonstrated the same direction of expression. KEGG pathway analysis and Gene Ontology revealed that many differentially expressed genes were enriched in several biological processes and pathways relevant to inflammatory response, defense response, regulation of T cell activation, immune response, regulation of leukocyte activation, cytokine-mediated, TNF, IL-17, and Toll-like receptor signaling pathways as being well-known. These findings demonstrate that RNA-Seq is more accurate compared to microarray analysis in identifying differential gene expression and provides new insight into the molecular mechanisms related to immune or inflammatory responses to M. bovis infection.

Keywords: Mycobacterium bovis, Microarray, RNA-Sequencing, Pathway analysis

ICB11-1401-12-1033

Comparative transcriptome analysis reveals significant differentially expressed genes in bovine viral diarrhea virus (BVDV) infection

Masoumeh Naserkheil[1], Farzad Ghafouri[2], Vahid Dehghanian reyhan[3], Mostafa Sadeghi[4], Herman W. barkema[5], Sonia Zakizadeh[6*]

[1]National institute of animal science of korea. [2]University of tehran. [3]University of tehran. [4]University of tehran. [5]University of calgary. [6*]Animal science research institute of iran.

*Corresponding author(s). E-mail(s): sonia_zaki@yahoo.com;

Abstract

In the cattle industry, the bovine viral diarrhea virus (BVDV) is one of the most common pathogens which results in fever, diarrhea, leucopenia, reduction in milk yield and reproductive dysfunctions, concurrent infections, immunosuppression, and persistently infected (PI) cattle. The virus is a member of the Pestivirus genus in the Flaviviridae family . This study aimed to provide a better understanding of the complicated pathogenesis of this disease on a molecular basis using transcriptome analysis of cells infected with this virus. The primary objective was to investigate comparative transcriptome profiling of bovine kidney cell line (MDBK) cells infected at different time points using public repositories microarray and RNA-Seq datasets. The results showed that 2054, 161, 2431, and 1562 differentially expressed genes (DEGs) in different comparison groups MDBK cells infected with BVDV post 0h (MBV0h) vs. MBV6h, MBV6h vs. MBV12h, MBV12h vs. MBV24h, MBV24h vs. MBV72h, respectively were identified based on fold change $> \pm 1$ and false discovery rate < 0.05. Among these genes, eight genes including HS3ST2, TIMP3, MT2A, SOSTDC1, RGS14, SPARC, ST8SIA4, and CALB1 were identified as the common genes, of which five were underexpressed and three genes were overexpressed in different comparison groups of MDBK cells infected with BVDV during 0 to 72h. Consequently, the identification of disease-causing genes could contribute to a better understanding of the underlying genetic background and can be considered a starting point for further studies on mechanisms related to the host's response to BVDV infection.

Keywords: Bovine viral diarrhea virus, Gene expression, Transcriptome, Cattle

ICB11-1401-12-1034

Molecular Docking Analysis of gingerol compound of Zingiber officinale Roscoe against Breast Cancer

Tooba Abdizadeh[1*]

[1*]Shahrekord university of medical sciences.

*Corresponding author(s). E-mail(s): t.abdizadeh@gmail.com;

Abstract

Breast cancer is one of the most common malignancies in women, and its current therapy is to target the hormone receptors by the use of effective drugs . The human progesterone receptor (hPR) belongs to the steroid receptor family and is regarded as the important target for breast cancer treatment . Therefore, PR can be an attractive drug target for the treatment of breast cancer. In present study, gingerol in Zingiber officinale Roscoe and doxorubicin (reference drug) were retrieved from PubChem server as 3D structures in SDF files. Crystal structure of PR was obtained from Protein Data Bank (PDB) database (PDB ID: 4OAR). Then, molecular docking of these compounds was investigated by MOE software according to free energy binding and interaction with PR protein. The docking results showed that gingerol exhibited strong binding interaction to PR similar to the known doxorubicin inhibitor and bind highly with some of the amino acid residues in the active site of PR and thus could act as potential inhibitory compound against PR protein and require laboratory and experimental investigation.

Keywords: PR, Molecular docking, Gingerol

ICB11-1401-12-1035

Using bioinformatics tools to identify pathogenic agents of Cutaneous Leishmaniasis

Kavoos Momeni[1], Ehsan Ahmadpour[2], Rasoul Sharifi[3], Saeid Ghorbian[4*]

[1]Islamic azad university. [2]Medical university of tabriz. [3]Islamic azad university. [4*]Islamic azad university.

*Corresponding author(s). E-mail(s): ghorbian20@gmail.com;

Abstract

Cutaneous leishmaniasis (CL) is an infectious disease caused by a parasite called Leishmania Tropica. CL leaves scarring on exposed body parts from bites by infected female phlebotomine sandflies. CL is an endemic disease in some cities of our country and represents a public health challenge. Leishmania Tropica lesions tend to heal more slowly (6-15 months) and are less responsive to treatment. The molecular mechanisms underlying its pathogenesis remain unclear. We herein employed a bioinformatics analysis approach to study molecular signatures of CL. Data set GSE127831 was used to construct a co-expression network for Weighted Gene Co-expression Network Analysis (WGCNA). An attempt was made here to investigate potential biomarkers that could contribute to the pathogenesis and progression of infected skin lesions, as well as to recommend the best method of treatment. In a bioinformatics analysis of healthy skin biopsy and wounds infected with Leishmania, we identified Differentially Expressed Genes (DEGs). Gene Ontology (GO)-based analysis was performed on DEGs and WGCNA modules utilizing Cytoscape. A module with 456 genes had the strongest correlation with cutaneous wound size among almost 16,600 genes expressing differently around Leishmania wounds. According to the functional enrichment analysis of this module, three clusters of genes have significantly changed expression, including genes encoding tissue-damaging cytokines, genes encoding collagen and fibrin proteins, and genes encoding extracellular matrix. These conditions lead to cutaneous wounds or delayed healing processes. OAS1, SERPINH1, and FBLN1 are the hub genes of these three groups. Using this information, we can develop new techniques for dealing with Cutaneous leishmaniasis' adverse effects.

Keywords: Bioinformatics Analysis, Differentially Expressed Genes, Weighted Gene Co-Expression Network Analysis, Gene Ontology Analysis, Cutaneous Leishmaniasis

ICB11-1401-12-1036

A comparative structural study on two immobilized metagenomic α-amylases by molecular docking

Fereshteh Noroozi tiyoula[1], Shohreh Ariaeenejad[2*], Elahe Motamedi[3], Nasim Sadeghian motahar[4], Kaveh Kavousi[5]

[1]University of tehran,department of bioinformatics, institute of biochemistry and biophysics (ibb). [2*]Agricultural research education and extension organization.[3]Agricultural research education and extension organization. [4]Agricultural research education and extension organization. [5]University of tehran,department of bioinformatics, institute of biochemistry and biophysics (ibb).

*Corresponding author(s). E-mail(s): shariaee@gmail.com;

Abstract

α-amylases are the major types of enzymes used in industries capable of catalyze starch hydrolysis. To facilitate reusability of the enzyme and address some problems associated with the free form such as lack of stability, the immobilization techniques are employed. Furthermore, to achieve more reliable results in modern research, we need structure-based modelling studies along with experimental studies. We immobilized two ***α***-amylases named PersiAmy2 and PersiAmy3 identified from rumen metagenome on cellulose nanocrystals, reusability, and structural properties were investigated. Finally, the screening results based on the interactions between the ***α***-amylases and cellulose nanocrystals were studied using 300 runs of molecular docking simulation, using Autodock 4. Immobilized PersiAmy2 and PersiAmy3 provided higher thermal stability, and sustainability over issues of reusability compared to the free enzymes. Homology modeling, structural analysis, and molecular docking studies were also applied to simulate molecular interactions between enzymes and with one cellulose unit of Cellulose nanocrystals. Results were in agreement with the experimental findings and proved higher catalytic activity of immobilized PersiAmy3 than PersiAmy2. Based on computational results, hydrogen bond numbers were not sufficient alone, to prove which binding pose is more favorable. So, H-bonds in both PersiAmy2-ligand and PersiAmy3-ligand complexes were equal, but value of binding free energy, number of residues in interactions with ligand in the binding site, and area of binding pose, had more favorable results in PersiAmy3-ligand complex in comparison with PersiAmy2-ligand complex.

Keywords: amylase, immobilization, operational stability, homology modeling, structural analysis, molecular docking simulation

ICB11-1401-12-1037

The role of Exome Sequencing analysis in the diagnosis of patients with rare neurogenetic disorders

Paria Najarzadeh torbati[1], Mehran Beiraghi toosi[2], Mohammad Doosti[3], Najmeh Ahangari[4], Ehsan Ghayoor karimiani[5*]

[1]Department of medical genetics, next generation genetic polyclinic, mashhad, iran. [2]Pediatric neurology department, ghaem hospital, mashhad university of medical sciences, mashhad, iran. [3]Department of medical genetics, next generation genetic polyclinic, mashhad, iran. [4]Department of medical genetics, next generation genetic polyclinic, mashhad, iran. [5*]Molecular and clinical sciences institute, st. george's, the university of london.

*Corresponding author(s). E-mail(s): eghayoor@gmail.com;

Abstract

Introduction Diagnostic courses for neurogenetic disorders often require the use of substantial time and resources. Previous studies have shown Exome Sequencing increased diagnostic and clinical utility in medical genetics. Although genetic heterogeneity in neurogenetic disorders has been an obstacle to phenotype-based diagnostic testing, Exome Sequencing improved the presumptive diagnostic rate in patients from 25% to 48%. Our objective was to describe the role of Exome Sequencing in identifying novel variants in a group of patients with neurogenetic disorders. Methods A complete clinical and paraclinical examination has been done by expert specialists and clinical geneticists. The team contributed to the discovery or identifies the extremely rare disorders of disease genes using the Exome Sequencing technique followed by comprehensive bioinformatics analysis. Parents and healthy offspring were assessed for the candidate gene variants. Results We have found genetic variations in genes such as PI4K2A, OGDH, and YIF1B with novel variants based on computational prediction using bioinformatics tools. Many of these genes are the subject of novel and unexpected genetic pathways crucial for normal brain function. At the present time, powerful sequencing techniques are identifying large numbers of genetic variants associated with unique phenotypes. Conclusion We have demonstrated that Exome Sequencing as a high throughput molecular technique has rapidly become a component of the clinical approach that requires a broad search for causal variants across the spectrum of rare Mendelian disorders.

Keywords: Exome Sequencing, Genetics, Neurogenetic disorders, Bioinformatics analysis

ICB11-1401-12-1038

Computational identification of driver genes and pathways of chronic lung diseases

B. fatemeh Nobakht m. gh[1], Hadi Rezaeeyan[2], Masoud Arabfard[3*]

[1]Baqiyatallah university of medical sciences. [3*]Baqiyatallah university of medical sciences.

*Corresponding author(s). E-mail(s): arabfard@gmail.com;

Abstract

Chronic lung disease is characterized by impaired lung function. Given that many of these diseases are common in terms of clinical symptoms and pathogenesis, identifying shared pathogenesis of diseases can be helpful in designing preventive and therapeutic strategies. The aim of this study was to evaluate the proteins and pathways of four Chronic Obstructive Pulmonary Diseases (COPD), Asthma, Idiopathic Pulmonary Fibrosis (IPF) and Mustard Lung Disease (MLD). In this study, after collecting the data and determining the gene list of each of four diseases, gene expression changes were examined compared to healthy individuals. Also, protein-protein interaction (PPI) and pathway enrichment analysis were used to evaluate genes and shared pathways of cited diseases. The results showed that there were 22 shared genes, which included ACTB, AHSG, ALB , APO A1, APO C3, FTH1 , GAPDH , GC, GSTP1, HP, HSPB1, IGKC, KRT10 , KRT9, LCN1 , PSMA2, RBP4 , S100A8 , S100A9 , TF and UBE2N. The major biological pathways which these genes are involved is inflammatory ones. Also, some of these genes activate different pathways in each of mentioned diseases, that lead to induction or inhibition of inflammation. Finally, it was concluded that identification of genes and shared pathways of diseases can be effective in identifying pathogenesis pathways and designing preventive and therapeutic strategies.

Keywords: Systems Biology, COPD, IPF, Asthma, Protein–protein interaction network, Mustard Lung Disease

ICB11-1401-12-1039

Bioinformatics evaluation of CFEM gene region in some fungal species

Aydin Hassanzadeh[1*], Mohammad ali Tajick ghanbari[2]

[1*]Sari agricultural sciences and natural resources university, idin.real@gmail.com.[2]Sari agricultural sciences and natural resources university.

*Corresponding author(s). E-mail(s): idin.real@gmail.com;

Abstract

The CFEM domain is a special gene region in the genome of fungi that contains eight conserved cysteines and it is found in some proteins with proposed roles in fungal pathogenesis. The aim of this research was a bioinformatics study of the similarities and differences of this gene region in some Fusarium and Trichoderma species. In this study were evaluated 24 amino acid sequences from 9 Trichoderma species and 51 amino acid sequences from 19 Fusarium species. Sequences with a length of 150-200 amino acids were received in FASTA format from UniProtKB. The conserved domains of this gene region were checked by using COBALT software on NCBI. The identification of conserved motifs and determining the position of these motifs in domains were analyzed by using MEME and Motif Finder software. The amino acid sequences were aligned by using ClustalW algorithm and phylogeny trees were drawn with MEGA11 software, Neighbor-Joining method, and Bootstrap analysis with 1000 repetitions. The results showed that due to the small size of CFEM domain, it has a similar structure in different species of Fusarium and Trichoderma. According to the analysis of phylogenetic trees, this gene region could separate Trichoderma species, but could not help to separate Fusarium species due to the high inter-species similarity. Besides, the results showed that Trichoderma and Fusarium species were placed in separate branches due to the differences in motifs. In only one exception, Fusarium venenatum and Trichoderma asperellum were placed in a sister group, which could be due to structural similarities in the proteins that contain this domain. Considering the role of this domain in the pathogenicity of pathogenic fungal species, the presence of two species of Fusarium and Trichoderma in a sister group can indicate the existence of pathogenic potential in some Trichoderma species.

Keywords: fungal domains, amino acid sequences, conserved regions, motifs

ICB11-1401-12-1040

Comparison Of Milk Production Parameters In Different Environment Predicted By Non-Linear Model And Artificial Neural Network

Niloofar Rasti doust[1], Sonia Zakki zadeh[2], Nasser Emamjomeh[3], Ali Sadeghi[4], Reza Azizi nejad[5], Mahdi Elahi torshizi[6]

[1]Department of animal science , islamic azad university , science and research branch , tehran, iran. [2]Animal science research institute of iran, agriculture research, education and extension organization(areeo), karaj, iran. [3]Islamic azad university , science and research branch , tehran, iran. [4]Islamic azad university , science and research branch , tehran, iran. [5]Islamic azad university , science and research branch , tehran, iran. [6]Department of animal science, mashha branch, islamic azad university, mashhad, iran.

*Corresponding author(s). E-mail(s):

Abstract

A mathematical formula describing the behavior of milk yield during a lactation referred to a lactation curve. Nowadays, artificial intelligence-based models, such as artificial neural networks (ANNs), are applied to predict milk curve, as well. The main advantages of ANNs are their ability to learn and control complex non-linear data. Dairy cattle breeding programs are based on milk production, which highly be affected from different climate situations. To predict the effect of temperature-humidity-index (THI) on lactation curve and compare the ANN and some different nonlinear functions (Wood, Wilmink) a dataset of 107914 milk test days from first parity cows in 7 provinces located in humid and arid area during 2000 to 2016 was used. THI index categorized to 5 stress classes (<52, 52-61, 62-71, 72-82, >82). ANN model was developed according to 1-Multilayer Perceptron with 5 nodes in layer, Sigmoid Axon function and Levenberg Marquadt learning rule by NeuroSolution V.5 software. Nonlinear models were analyzed by NLIN procedure of SAS V.9.4. The goodness of fitted models and comparison determined by using the coefficient of determination (R2) and residual mean square (MSE). The fixed effects included record number of lactation, season of milk record, age of cow (days), herd, year of calving, stress classes, number of milking. The results revealed that ANN had the highest accuracy (R2=0.493) and the lowest error (RMSE=0.089), described the milk curve better than the Wilmink (R2= 0.402, RMSE= 6.289) and Wood (R2=0.388, RMSE= 0.232). The best condition for milk production was up to THI=71 and thereafter it was decreased. According to (Lees et al,2019), the maximum daily air temperature has a great effect on reducing milk yield. During warming, to a critical range of THI of 70–72, the performance of dairy cattle is inhibited, and deterioration in milk yield occurs at THI from 72 to 78, The results of this research are consistent with the current research.Therefore, ANN model is recommended for describing the milk production curve of Iranian Holstein cattle.

Keywords: ANN, THI, Wilmink, Prediction

ICB11-1401-12-1041

Haplotype Assembly Problem Models in Diploid and Polyploid Cases

Maryam Etemadi[1*]

[1*]Islamic azad university fuman branch.

*Corresponding author(s). E-mail(s): etemadi.maryam@gmail.com;

Abstract

The haplotype assembly problem aims at finding the haplotypes from a number of sequenced fragments. This problem can be considered in two diploid and polyploid cases. Diploid case aims at finding two haplotypes but in polyploid case the number of haplotypes is grater than two. For example, humans are diploid i.e., they have two copies of their chromosomes, one inherited from the mother and the other inherited from the father. Many plants are polyploid i.e., their somatic cells contain more than two copies of each chromosome. For instance, cultivars of potato, coffee, cotton and peanut are polyploid. The polyploid case is more complicated and challenging diploid case. In this paper, a zero-one integer linear programming (ILP) model is proposed that can be used for solving haplotype assembly problem in both diploid and polyploid cases. It is compared with the previous model for diploid case using simulated dataset. Experiments show that the new model can be solved within much shorter time by CPLEX.

Keywords: Diploid, Polyploid, Haplotype assembly, Integer linear programming

ICB11-1401-12-1042

The impact of synthetic food colors on target proteins is suspected to have negative effects on the human body, an in silico study

Mohammad hossein Morshedsolouk[1], Mohammad mahdi Aghanour ashtiani[2], Nejad Ebrahimi[3]

[1]Shahid beheshti university. [2]Shahid beheshti university,samad. [3]Shahid beheshti university.

*Corresponding author(s). E-mail(s):

Abstract

Color additives are one of the most widely used additives to foods, beverages, medicines, cosmetics, etc. Daily, many food colors enter the human body in different ways. Artificial food colors are used more than natural food colors due to their easier accessibility on an industrial scale, high color intensity, and low cost. Therefore, the need to examine the artificial food colors with human biological systems is felt highly due to their synthetically. In this In silico study, the effects of color additives on systematic proteins have been investigated.104 color additive status was obtained from the PubChem then ligands were prepared with the Schrodinger's maestro 12.9 software. Afterward, the systematic proteins were received from the Protein Data Bank, and after the preparation of these proteins, molecular docking was done by using the Schrodinger's maestro 12.9 software. Finally, the obtained results were analyzed. As a result of these studies, it was found that a large number of these artificial colors have significant interaction with the biological targets of the body. This issue can be the source and cause of many common diseases, so this finding shows the importance of limiting the usage and consumption of these artificial colors and replacing them with natural dyes. Some of these synthetic colors, for example, Hematoxylin, Acid Leather Orange G, iron gluconate complex compounds, and Disodium edetate, interact with several body systematic proteins, including thyroid receptor alpha (PDB ID: 1NAV), thyroid receptor beta1 (PDB ID: 1NAX), androgen receptor (PDB ID: 1T74), and mineralocorticoid receptor (PDB ID:2AA2) with a docking score range of -8 to -12, which cause thyroid, prostate, and hypertension-related diseases. This makes it a priority to stop using these colors that interact with multiple targets. Finally, it is suggested to use natural colors or change the structure of these colors in such a way that they do not interact with these biological targets and prevent the occurrence of related diseases.

Keywords: Food colors, Molecular docking, Thyroid, Cytochrome P450, Androgen, Glucocorticoid

ICB11-1401-12-1043

Statistical physics approach in epilepsy disease

Seyede nasim Tayyeb[1], Farinaz Roshani[2*]

[1]Alzahra university. [2*]Alzahra university.

*Corresponding author(s). E-mail(s): farinaz@alzahra.ac.ir;

Abstract

Epilepsy is a central nervous system disorder in which brain activity becomes abnormal, causing seizures or periods of abnormal behavior, emotions, and sometimes loss of consciousness. Having a seizure does not mean that a person has epilepsy. At least two seizures without a known trigger (unprovoked seizures) occurring at least 24 hours apart are usually required for a diagnosis of epilepsy. Epilepsy studies often rely on EEG signals to provide information about brain behavior during seizures. In this research, we used two network-based statistical analysis methods to classify the EEG data recorded by Bonn University of Bonn. The first method was based on correlation between two time series and the second method was based on quantile networks. In the next step, we examined the networks built in two ways, for healthy and diseased groups, according to several network topological criteria. Correlation-based networks are a complex network construction method using one-dimensional time series. Each channel is considered as a vertex of the complex network. If there is a relationship between two time series, that means there is an edge between the two vertices. The purpose of using this method is to compare the degree of connection between different brain areas.Recently, a mapping of a time series to a network has been proposed, based on the concept of transition probabilities. This series results in a"quantile graph"(QG). The purpose of using this method is to obtain the relationship of different areas of the frequency ranges of each of the signals. Based on different network criteria, i.e. clustering coefficient and betweenness centrality, the obtained results showed that the correlation method and the QG method are able to detect differences in the dynamic properties of brain electrical activity.Comparing the correlation networks of patient and healthy people, it was concluded that the networks of patient people are denser and the quantile networks are denser in healthy people. This problem can be used as a diagnostic panel for EEG signal data to diagnose epilepsy.

Keywords: Epilepsy, Electroencephalographic time series, Complex network, Network correlation, QG-method

ICB11-1401-12-1044

Network Pharmacology Integrated With Experimental Validation and Molecular Docking Revealed the Anti-inflammatory Effects of Coconut Oil

Hossein Ebrahimi[1], Hossein Asgari rad[2*], Pedram Ebrahimnejad[3], Fatemeh Shaki[4], Shervin Amirkhanloo[5], Melika Ahmadi[6], Sholeh Akbari[7]

[1]Mazandaran university of medical sciences. [2*]Mazandaran university of medical sciences.[3]Mazandaran university of medical sciences. [4]Mazandaran university of medical sciences. [5]Department of pharmaceutics, faculty of pharmacy, mazandaran university of medical sciences, sari, iran. [6]Department of pharmaceutics, faculty of pharmacy, mazandaran university of medical sciences, sari, iran. [7]Department of toxicology and pharmacology, faculty of pharmacy, mazandaran university of medical sciences, sari, iran.

*Corresponding author(s). E-mail(s): dr_asgarirad@yahoo.com;

Abstract

Skin inflammation can be caused by biological factors, chemical, thermal, and the body's immune system reactions. Virgin coconut oil (VCO) has traditionally been used as a moisturizer and anti-inflammatory agent. However, the exact molecular mechanism remains unclarified. In the present study, a network pharmacology technique was employed to uncover the active ingredients, their potential targets, and signaling pathways in coconut oil for the treatment of inflammation.
Methods:
inflammation was induced by carrageenan solution in rat paw and edema was measured. Then tissue pathology and inflammatory gene expression were evaluated. The bioactive compounds of coconut oil were screened from the databases. Coconut oil biological targets were obtained from Swiss target prediction and inflammation-related genes from Gencards and OMIM. Subsequently, The "Component-Target" network and protein–protein interaction (PPI) network were constructed, and hub genes were screened out by topological analysis. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed on genes in the PPI network and finally, Molecular docking simulation was performed to evaluate the binding activity between the predicted hub genes and the bioactive ingredients.
Result:
Coconut oil significantly reduced edema and infiltration of inflammatory cells in the tissue. A total of 18 key active components in coconut oil interacted with 72 inflammatory targets, the main one being IL-6, PGST-2, and MMP-9. main compounds related to inflammation were Linoleic acid, Oleic acid, Capric acid, and Vanillic acid. Coconut oil reduced serum levels of TNF-α and IL-6.
Conclusion:
This suggests that coconut oil has a potent anti-inflammatory action, thereby providing a scientific reference for clinical studies.

Keywords: Coconut oil, active ingredients, network pharmacology, molecular docking, anti-inflammatory

ICB11-1401-12-1045

Prediction of a G-Quadruplex Aptamer structure by mFold servers for identification histidine sequence-containing proteins

Mohammad javad Jadidi[1], Rahman Emamzadeh[2*], Mahboobeh Nazari[3]

[1]University of isfahan. [2*]Universiry of isfahan.[3]Iran university.

*Corresponding author(s). E-mail(s): r.emamzadeh@sci.ui.ac.ir;

Abstract

DNA has been extensively used for target recognition in biosensors because of its stability, cost effectiveness, and easy modification. Aside from complementary nucleic acids, DNA can also recognize other types of target molecules by using DNAzymes , and aptamers of DNA structure formed on non-canonical Hoogsteen-type base pairing . G-quadruplex (G4) structures are DNA tetraplexes that Four guanine bases associate with each other through Hoogsteen hydrogen bonds to form a guanine tetrad plane (G-quartet), and then two or more G-quartet planes stack on top of each other to form a G4 structure . Hemin/G-Quadruplex DNAzyme can quickly detect molecular goals by the naked eyes which does not require complex instruments.In this study with the aim of designing an aptosensor to detect proteins containing histidine tag, the aptamer sequences were taken from Shijia Wu et al (TTTTTTTGGCAAGAGGGTGTGCTTAAGGTGGACACGGTGGCTTAG) and G-quadruplex from Chenchen Ge et al (GTGGG TAGGG CGGGT TGGG). To check whether these two sequences interfere with each other in the structure, Mfold servers were used. (CAGTAG) linker was chosen because when the aptamer and G-quadruplex are placed next to each other, the lowest possible delta G is formed for this structure without changing the spatial structure of the aptamer and G-quadruplex from the perspective of predicting the two dimensional structure of DNA.

Keywords: G-quadruplex, Aptamer, mFold server

ICB11-1401-12-1046

Identification of differentially expressed genes in papillary thyroid cancer patients based on gender

Seyede elmira Yazdi rouholamini[1], Seyed-morteza Javadirad[2*]

[1]Isfahan university. [2*]Isfahan university.

*Corresponding author(s). E-mail(s): javadirad@yahoo.com;

Abstract

Papillary thyroid cancer (PTC) is one of the most prevalent endocrine malignancies and the incidence of this malignancy is on the rise worldwide. A majority of PTC cases occur in women and it is three times more common in women than in men. This study aims to gain a deeper understanding of the molecular mechanisms of PTC and the reasons for its high prevalence among women. It also aims to develop more effective diagnostic and therapeutic approaches Method RNA-seq expression data of PTC was extracted from TCGA database with TCGABIOLINK package. A total of 500 tumor tissues were examined, including 365 women and 135 men. The edge R package was used to detect differentially expressed genes between women and men. Based on the string database, a protein-protein interaction network was constructed. The MCC method was applied to identify hub genes. Cluster Profiler package was used to perform enrichment analysis of the Gene Ontology and Kyoto Encyclopedia of Genes and Genomes. Results 118 differentially expressed genes between two tumor groups of men and women were identified according to TCGA data analysis. A protein-protein interaction network comprised of 36 nodes and 118 edges was constructed using protein-coding genes. The MCC method identified significant genes and the first ten genes selected as hub genes. The list included ZFX, TXLNG, DDX3X, EIF1AX, EIF2S3, USP9X, RPS4X, KDM5C, KDM6A, and UBA1. According to the KEGG enrichment analysis, two pathways were identified that relate to RNA transfer signaling pathways and the RIGRIG-I-like receptor signaling pathway.

Keywords: Key words: papillary thyroid cancer, TCGA analysis, differentially expressed genes

ICB11-1401-12-1047

Effect of Salt Concentration on Structural Changes of Amyloid β

Maryam Ghorbaninia[1], Shirin Doroudgar[2], Mohamad reza Ganjalikhany[3*]

[1]University of isfahan. [2]University of arizona. [3*]University of isfahan.

*Corresponding author(s). E-mail(s): m.ganjalikhany@sci.ui.ac.ir;

Abstract

Aggregation of amyloid $\boldsymbol{\beta}$ (A$\boldsymbol{\beta}$) has been associated with Alzheimer's disease. Aggregation depends on two sets of factors, including the intrinsic features of polypeptides and the physicochemical environment (such as ionic strength and pH). The effect of ion strength on the propensity of aggregation and its morphology was supposed. here, we investigated the behavior of the A$\boldsymbol{\beta}$ (16-22) peptide as a fibril-forming core at three different salt concentrations by a molecular dynamics simulation approach. The region of 16 KLVFFAE22 peptide of A$\boldsymbol{\beta}$42 (PDB: 1Z0Q) was retrieved as a starting structure. Molecular dynamic simulations were carried out using the AMBER20 package with parameterization of ff19SB-OPC. Three systems, with 0 mM (system I), 30 mM (system II), and 150 mM (system III) NaCl, were assigned. The length of the simulations was 50 ns. CCPTRAJ has been employed for the analysis of the trajectories. Root mean square deviation (RMSD) of the peptide in system I had the least structural changes compared to the others and the lowest radius of gyration (Rg). The peptide in system I showed the least distance between K16 and E22. Moreover, this peptide had the most intramolecular hydrogen bonds that decreased by increasing the ionic strength of the solvent in other systems. Based on the RMSD, and Rg results, the peptide in system I was the most stable and compact structure since the terminal ends of the peptide had the highest tendency to interact with each other. Ions interact with the opposite charge residues and decrease hydrogen bonds and the energy of electrostatic intramolecular interaction of the peptide resulting in increasing in the flexibility of peptides. Our results are in agreement with previous studies that showed increasing ionic causes the stability of the peptide to decline. These findings could be useful for predicting aggregation.

Keywords: Keywords: Amyloid $\boldsymbol{\beta}$, NaCl, Aggregation, Molecular Dynamics

ICB11-1401-12-1048

Learning state machines on protein sequences

Seyed reza Beh-afarin[1], Kaveh Kavousi[2*]

[1]University of tehran kish international campus. [2*]Tehran university.

*Corresponding author(s). E-mail(s): kkavousi@ut.ac.ir;

Abstract

Defining signatures of known families of biologically related protein sequences (at the functional or structural level) is of great significance and may help identify conserved regions among the family of proteins, revealing the importance of the function of their structural properties. Some scholars argued for the benefit of viewing the sequences as sentences derived from formal grammar. This allows us not only to overcome the position-specific characterization of the sequences but also to benefit from the explicit modeling provided by grammar. The use of sequencing through grammar gives a prediction on whether a sequence belongs to a particular family and provides information about the reason why a sequence belongs to a particular family. Automata can be learned successfully on proteins and protein sequences. In this study, based on grammatical inference and multiple alignment techniques, a sequence-driven approach is used to learn automata on protein sequences. The approach is inspired by grammatical inference and multiple alignment techniques . The study focuses on fragment similarity to identify locally conserved regions and then improves the characterization by identifying informative positions. More attempts are required to raise prediction accuracy by developing distances taking into account the weights of the amino acids at each position with respect to the training sequences. The study attempts to identify differences and synergies between various approaches (such as learning syntactical models) converging from pattern discovery, multiple alignment, and grammatical inference to learning explicit models on proteins.

Keywords: finite-state machine (FSM), finite-state automaton (FSA), protein sequences, learning syntactical models

ICB11-1401-12-1049

FANCI/ hsa-miR-30b-3p /LINC01355 ceRNA axis can potentially promote Hepatocellular carcinoma by influencing DNA repair pathways

Behrad Kashefi[1], Fariba Dehghanian[2*]

[1]University of isfahan. [2*]University of isfahan.

*Corresponding author(s). E-mail(s): fa.dehgahnoian@sci.ui.ac.ir;

Abstract

Background:
Hepatocellular carcinoma accounts for the majority of primary liver cancers. Worldwide, liver cancers are the fourth most common cause of cancer-related death and rank sixth in terms of incident cases. Hence, we aimed to take a closer look at competitive endogenous RNA (ceRNA) networks and signaling pathways related to Hepatocellular carcinoma.
Methods:
Microarray data was retrieved from NCBI Gene Expression Omnibus (GEO) which contains gene expression profiles of 16 hepatocellular carcinoma and 16 adjacent non-tumorous liver tissue samples (GSE60502). Using the GEO2R web tool, 6316 differentially expressed genes (DEGs) were obtained. ENCORI (The Encyclopedia of RNA interaction) was employed to validate the differential expression and survival analysis. Reactome and Kyoto Encyclopedia of Genes and Genomes (KEGG) were then used for gene enrichment analysis. The miRNAs that potentially bind to the 3' UTR of the selected gene were obtained from the miRWalk database. lncRNAs interacting with miRNA were suggested using DIANA-LncBase v3. Finally, mRNA-miRNA and miRNA lncRNA coexpression analysis was done using the ENCORI Pan-Cancer Analysis Platform and the ceRNA network was constructed.
Results:
According to the GEO2R DEG analysis, the FANCI gene was significantly upregulated (| LogFC| = 2.86, adj.P value = 1.60E-05) in hepatocellular carcinoma compared to adjacent non-tumorous liver tissue samples. FANCI protein is required for the efficient repair of DNA damage, especially interstrand cross-links (ICDLs). The mRNA-miRNA coexpression analysis showed the significant coexpression of has-miR-30b-3p with FANCI gene (r = -0.180, p-value = 5.17e-04) across 370 samples. Using LncBase v3, lncRNAs that can potentially bind to FANCI gene were revealed and LINC01355 was suggested (coexpression analysis; r = -0.128, p-value = 1.34e-02).
Conclusion:
In conclusion, FANCI/has-miR-30b-3p/LINC01355 ceRNA network axis can be a probable network affecting the development of Liver Hepato carcinoma.

Keywords: Hepatocellular carcinoma, Differentially expressed genes, ceRNA, FANCI, has-miR-30b-3p, LINC01355

ICB11-1401-12-1050

In silico exploration of Neohesperidin inhibitory mechanism on TGF-β pathway and improving its bioavailability

Najme Dehghanbanadaki[1], Hossein Naderi-manesh[2*], Havva Mehralitabar[3]

[1]Tarbiat modares university. [2*]Tarbiat modares university.[3]Agricultural sciences and natural sari resources university.

*Corresponding author(s). E-mail(s): naderman@modares.ac.ir;

Abstract

Background:
Neohesperidin, a natural compound obtained from orange, which is widely used as an industrial sweetener, has proven anti-cancer and anti-inflammatory properties but has a low bioavailability and unknown functional mechanism. Some papers have shown that Neohesperidin can prevent the phosphorylation and oligomerization of SMAD proteins in the TGFβ pathway, however, the detailed molecular mechanism of the process is largely unexplored. In this research, we investigated the mechanism of the effect of Neohesperidin on the intracellular domain kinase of TGFβ membrane receptor type 2 and the inhibitory effect of this molecule on SMAD protein phosphorylation, Also, using peptide-drug conjugate technology, we designed a peptide structure to increase the bioavailability of this compound for the intracellular target. Materials and Methods:
The SMAD protein structure and, TGFβ type 2 intracellular kinase domain structure were obtained from the PDB database. The 3D conformer of Neohesperidin was also achieved from the PubChem database. The interactions between the SMAD protein and TGFβ type 2 kinase domain were investigated through the ClusPro docking web server, and the interaction between Neohesperidin and this kinase domain was explored through Molegro molecular docking software. After confirming the binding of Neohesperidin to the active site of the second TGFβ kinase and intending to increase the bioavailability of this compound, we designed a peptide-drug structure using HyperChem software. Results:
As expected, the results of docking simulation analysis between SMAD and TGFβ type 2 kinase domain, show that the SMAD carboxyl terminal is entirely placed in the kinase domain active site. The analysis of the interaction of SIS3 and Neohesperidin with the kinase domain also shows that these molecules occupy the kinase active site with relatively similar binding sites and energy. The result from the OPM server also suggests that designed Neohesperidin-peptide conjugates have a membrane binding capability. Conclusion:
The results of this study demonstrate that the Neohesperidin molecule, as a natural compound, can bind to the TGFβ type 2 receptor with the same energy and site as its chemical counterpart, SIS3, to block the kinase functional region and prevent the phosphorylation of the SMAD carboxyl-terminal. On the other hand, the peptide-drug conjugate stability and membrane binding capability show this structure's probable potential to improve the compound bioavailability. These results provide the basis for the design of new pharmaceutical compounds based on Neohesperidin. It also suggests an approach for the optimal transfer of this compound to cells.

Keywords: cancer, natural compound, peptide

ICB11-1401-12-1051

In silico study of IMPDH2 interactions with proteins involved in Wnt/β-catenin pathway

Yusuf Elhaei[1], Samira Samadieh[2], Razieh Yazdanparast[3*]

[1]University of tehran. [2]University of tehran. [3*]University of tehran.

*Corresponding author(s). E-mail(s): ryazdan@ut.ac.ir;

Abstract

The Wnt/$\boldsymbol{\beta}$-catenin pathway is a well-known signaling pathway that is involved with many physiological and developmental processes. The aberrant functions of this pathway have been reported in many cancers. $\boldsymbol{\beta}$-catenin, the prominent mediator of this pathway is regulated by a destruction complex. IMPDH2 which is the rate-limiting enzyme in guanine nucleotides synthesis and the dominant isozyme in cancer, is related to the Wnt/$\boldsymbol{\beta}$-catenin pathway, reciprocally. In this study, we sought to elucidate the interactions of IMPDH2 with different proteins of the Wnt/$\boldsymbol{\beta}$-catenin pathway at the molecular level. Hence, after retrieving and preparing the proteins structures from PDB database, the protein-protein docking was conducted using HDOCK server. Thereafter, the attained docking models were visualized and analyzed by PyMOL software. The results show that IMPDH2 interacts with $\boldsymbol{\beta}$-catenin with a docking score of -266.62. The docking scores for IMPDH2 interactions with the proteins of destruction complex such as APC, AXIN, CK1$\boldsymbol{\alpha}$, and GSK3$\boldsymbol{\beta}$ were -235.02, -230.41, -282.40, and -272.82, respectively. Next, we found that the interface of IMPDH2-$\boldsymbol{\beta}$-catenin complex is composed mainly of hydrophobic R-chain residues (38.3%). However, in the case of CK1$\boldsymbol{\alpha}$ as the most likely protein of destruction complex which interacts with IMPDH2, the uncharged polar R-chain residues are remarkable (44.6%). These results show that IMPDH2 could directly interact with $\boldsymbol{\beta}$-catenin and also may affect this protein through interaction with proteins of destruction complex such as CK1$\boldsymbol{\alpha}$. So, our findings support previous reports on the correlation of IMPDH2 with the Wnt/$\boldsymbol{\beta}$-catenin pathway and provide a better understanding of these interactions at the protein-protein level. Considering the residues involved in each interaction could be useful in future research on drug design for targeting the Wnt/$\boldsymbol{\beta}$-catenin pathway in cancer.

Keywords: IMPDH2, Wnt/$\boldsymbol{\beta}$-catenin pathway, CK1$\boldsymbol{\alpha}$, Protein-Protein Interaction

ICB11-1401-12-1052

TRHDE gene and its plausible role in multiple sclerosis

Khatereh Firouzi-farsani[1], Mina Dehghani-samani[2], Modjtaba Emadi-baygi[3*]

[1]Shahrekord university. [2]Shahrekord university. [3*]Shahrekord university.

*Corresponding author(s). E-mail(s): emadi-m@sku.ac.ir;

Abstract

Introduction:
Because of improved knowledge of mechanisms of relapsing-remitting multiple sclerosis, various disease-modifying medications have been developed during the past decades that suppress or modulate the immune system to reduce relapse rates and severity. However, the choices for treating progressive multiple sclerosis remain unsatisfactory and challenging. In this study, we employed an in silico approach using high-throughput microarray data to find dysregulated mRNA(s) and its interaction with plausible miRNA(s) in multiple sclerosis (MS) patients.
Methods:
The GSE38010 microarray dataset was downloaded and examined to find novel putative dysregulated mRNAs in MS patients in comparison with controls. The Affy package was downloaded from Bioconductor. To standardize the raw data and assess differentially expressed genes (DEGs), the limma package was applied . Pathway enrichment and gene ontology (GO) analyses were performed using the Enrichr Utilizing miRWalk, a study of microRNA (miRNA)-mRNA interactions was conducted to find miRNAs for potentially dysregulated mRNAs .
Results:
Among the top DEGs, TRHDE gene was significantly down-regulated in the MS samples (logFC: -5.34, adj. P. Val: 0.02353). Based on miRNA-mRNA interaction analysis, 3´ UTR of TRHDE mRNA has a significant interaction with hsa-miR-601 (score: 1, number of pairings: 16). Enrichment analyses showed that TRHDE was involved in peptide catabolic process (GO:0043171), cell-cell signaling (GO:0007267), and proteolysis (GO:0006508). Furthermore, TRHDE involves in metalloaminopeptidase activity (GO:0070006), aminopeptidase activity (GO:0004177), and pyroglutamyl-peptidase activity (GO:0016920).
Conclusion:
There are several studies showing the role of various peptidases in MS. TRHDE may be an important potential dysregulated gene in MS patients. We postulated that miR-601 may suppress the expression level of TRHDE and modulate its peptidase activity.

Keywords: Microarray analysis, Systems biology, Multiple Sclerosis, RNA interaction network

ICB11-1401-12-1053

TRHDE gene and its plausible role in multiple sclerosis

Mina Dehghani-samani[1], Khatereh Firouzi- farsani[2], Modjtaba Emadi-baygi[3*]

[1]Shahrekord university. [2]Shahrekord university. [3*]Shahrekord university.

*Corresponding author(s). E-mail(s): emadi-m@sku.ac.ir;

Abstract

Introduction:
Because of improved knowledge of mechanisms of relapsing-remitting multiple sclerosis, various disease-modifying medications have been developed during the past decades that suppress or modulate the immune system to reduce relapse rates and severity. However, the choices for treating progressive multiple sclerosis remain unsatisfactory and challenging. In this study, we employed an in silico approach using high-throughput microarray data to find dysregulated mRNA(s) and its interaction with plausible miRNA(s) in multiple sclerosis (MS) patients.
Methods:
The GSE38010 microarray dataset was downloaded and examined to find novel putative dysregulated mRNAs in MS patients in comparison with controls. The Affy package was downloaded from Bioconductor. To standardize the raw data and assess differentially expressed genes (DEGs), the limma package was applied . Pathway enrichment and gene ontology (GO) analyses were performed using the Enrichr Utilizing miRWalk, a study of microRNA (miRNA)-mRNA interactions was conducted to find miRNAs for potentially dysregulated mRNAs .
Results:
Among the top DEGs, TRHDE gene was significantly down-regulated in the MS samples (logFC: -5.34, adj. P. Val: 0.02353). Based on miRNA-mRNA interaction analysis, 3´ UTR of TRHDE mRNA has a significant interaction with hsa-miR-601 (score: 1, number of pairings: 16). Enrichment analyses showed that TRHDE was involved in peptide catabolic process (GO:0043171), cell-cell signaling (GO:0007267), and proteolysis (GO:0006508). Furthermore, TRHDE involves in metalloaminopeptidase activity (GO:0070006), aminopeptidase activity (GO:0004177), and pyroglutamyl-peptidase activity (GO:0016920).
Conclusion:
There are several studies showing the role of various peptidases in MS. TRHDE may be an important potential dysregulated gene in MS patients. We postulated that miR-601 may suppress the expression level of TRHDE and modulate its peptidase activity.

Keywords: Microarray analysis, Systems biology, Multiple Sclerosis, RNA interaction network

ICB11-1401-12-1054

Design of Potential Urease inhibitors via Pharmacophore and Docking Based Virtual Screening

Deniz Ajabi zareian[1], Hafezeh Salehabadi[2*], Hossein Danafar[3]

[1]Zanjan university of medical sciences. [2*]Tehran university of medical sciences.[3]Tabriz university of medical sciences.

*Corresponding author(s). E-mail(s): hsalehabadi@zums.ac.ir;

Abstract

Introduction:
Urease (urea amidohydrolase, EC 3.5.1.5), belong to the superfamily of amidohydrolases and phosphotriesterases, that hydrolyzes urea to CO2 and NH3 at a very high speed. Urease is essential for the growth and survival of bacteria such as Helicobacter pylori in acidic milieu of stomach, as it helps to increase the pH of the environment. H.pylori is a gram-negative, microaerophilic spiral bacterium is a common chronic pathogen found in humans, leading to the development of diseases such as chronic gastritis, gastric ulcer, and gastric cancer. Other pathological conditions associated with urease activity include renal lithiasis, hepatic coma, encephalopathy, and pyelonephritis. Therefore, inhibition of this enzyme can be very effective in inhibiting and controlling of H.pylori infection. Hence, there has been considerable interest in uncovering novel urease inhibitors with good bioavailability and minimal toxicity. Today, computer-aided drug design is becoming more widespread due to its swiftness, cost-effectiveness, and precision. This investigation has been embraced a hybrid approach that included docking and virtual screening to identify new urease inhibitors.
Method:
The crystal structure of Jack bean urease, with the PDB ID of 4h9m, was achieved from Protein Data Bank (www.rcsb.org). A proper pharmacophore model was generated using Ligand Scout 3.12 on the most critical area on the urease active site. Then ZINC libraries (over 35 million compounds) were applied for virtual screening.
Results and Discussion:
Obtained compounds from virtual screening were followed by molecular docking studies. Six compounds were selected based on docking score and complying with Lipinski's "rule of five". Selected compounds can be considered as a proper candidate in order to develop new urease inhibitors.

Keywords: Virtual Screening, Molecular Docking, Urease, Inhibitor

ICB11-1401-12-1055

Integrated single-cell analysis of tumor-infiltrating immune cells reveals key secretory protein-coding genes of astrocytes in glioblastoma

Mohammad amin Malekraeisi[1*], Alireza Ghaleh[2]

[1*]Iran university of medical sciences.[2]Iran university of medical sciences.

*Corresponding author(s). E-mail(s): malekraeisi.m@iums.ac.ir;

Abstract

Background:
Glioblastoma is the brain's most common malignant tumor disease, with an incidence rate of 3.23 cases per 100,000 people. Astrocytes are the most prevalent nervous system cells essential to brain structure stability. Reactive astrocytes significantly influenced the course of the glioblastoma with A1-specific gene expression and followed inflammatory effects on antigen presentation, complement activation, and enhanced neurotoxicity. Single-cell RNA sequencing allows for transcriptome-wide analyses of individual cells providing fascinating biological and medical insights.
Purpose:
In the present study, we explore critical secretory protein-coding genes in astrocyte transcriptome obtained from the single-cell transcriptome of glioblastoma.
Methods:
10x Single-cell transcriptome data of peripheral blood mononuclear cell samples and tumor-infiltrating CD45+ immune cell samples derived from glioblastoma patients were downloaded from the gene expression omnibus database under accession number GSE224090. Seurat package was used to perform analysis on 31313 cells. Filtrations based on feature counts and mitochondrial count percentages were performed. The canonical correlation analysis method was applied to integrate samples. Cell type of immune clusters was identified with SingleR package. Astrocyte cluster was chosen to extract differentially expressed genes and perform network analysis in the STRING database.
Results:
Low-quality cells and cell doublets were filtered out, and 29174 cells remained. Filtered data were integrated, and after cell type identification, 83 differentially expressed genes (27 up-regulated and 56 down-regulated) were determined for the astrocyte cluster. Obtained genes were entered into the STRING database to construct a protein-protein interaction network. Eight hub-genes with node degrees more than five and Uniprot keyword of secretion (KW-0964) were chosen as critical biomarkers for astrocytes: SPARC, HAPLN2, EGFR,CLU, APOD,GSN, ANXA1, and B2M.
Conclusion:
Eight hub-genes that encode secretory proteins in astrocytes were identified.

Keywords: Glioblastoma, Astrocyte, Single-cell RNA-sequencing, Hub-gene, Secretory protein

ICB11-1401-12-1056

Reconstruction of miRNA-mRNA bipartite network in saliva of gastric cancer patients via bioinformatics analysis

Maryam Talaei[1], Mohammad mehdi Kavoosi[2], Ali Zeraatpishe[3], Mansoureh Mohseni[4], Mohammad mehdi Naghizadeh[5], Abdolmajid Ghasemian[6*]

[1]Fasa university of medical sciences. [2]Fasa university of medical sciences. [3]Fasa university of medical sciences. [4]Fasa university of medical sciences. [5]Fasa university of medical sciences. [6*]Fasa university of medical sciences.

*Corresponding author(s). E-mail(s): majidghasemian86@gmail.com;

Abstract

Abstract Gastric cancer (GC) is the third leading cause of mortality worldwide (1). microRNAs (miRNAs) have substantial roles in the GC progression (2). miRNAs have attracted considerable research interest due to their potential applications as theranostic biomarkers. (3). In this study, the miRNA-mRNA regulatory network was uncovered among saliva proliferating GC patients using bioinformatics analyses. The gene expression profiles (GSE64951) were downloaded from the Gene Expression Omnibus database. Differentially expressed genes (DEGs) were analyzed using GEO2R tools in two groups including 31 healthy individuals and 63 GC patients. Then, a defined formula [-log(p.value) × | log (fold change) |] was applied to select the 2000-top significant probs. We established a protein-protein interaction (PPI) network of the DEGs through the (STRING) database and used Gephi software to select hub genes with a crucial role. Next, we employed the Toppgene database for Gene Ontology (GO). The Toppgene online database was also applied to select target miRNAs related to 20 hub genes and reconstructed miRNA-mRNA regulatory network by Cytoscape software. In total, 1590 DEGs were identified. Additionally, 10 hub genes (HSP90AA1, HSPA4, ESR1, JUN, CDK1, ATM, CD44, EZH2, POLR2A and CREB1) with the highest degree were determined. The most frequent top 2 GO of molecular function included cadherin binding and RNA binding; of biological process included regulation of RNA splicing and peptidyl-serine phosphorylation; and of cellular pathway included RHO GTPase Cycle and Hippo Signaling. The network five hub (highest degrees of association) mRNAs (SMAD2, ESR1, CREB1, CD44, and CXCL8) and miRNAs (miR-26a, miR-656, miR-607, miR-203a, and miR-302b) were determined. We reconstructed a salivary miRNA-mRNA bipartite network with potential functions in the saliva of GC patients, demonstrating their potential theranostic applications as biomarkers.

Keywords: biomarkers, miRNA, gastric cancer

ICB11-1401-12-1057

Application of bioinformatics in cancer diagnosis and prediction

Javad Amraei[1], Mohammad Abdolahad[2], Abolfazl Mirzapour[3*]

[1]Tarbiat modares. [2]Tehran. [3*]Tarbiat modares.

*Corresponding author(s). E-mail(s): a.mirzapour@modares.ac.ir;

Abstract

Cancer is considered to be the most dangerous disease in the world. Every year many people around the world are infected with this disease and die. Breast cancer is the most common disease in women. Whreas breast cancer includes 30% of women's cancers. If cancer diseas have been predicted in early stag, chanse of treating was increased. Machine learning techniques have a high ability to recognize and classify cancer cells and normal cells. we used these techniques to study and assesment breast tissue cancer cells. In this study, 120 breast cancer patients, all of whom underwent surgery, were examined. Several different characteristics were considered for each patient. These characteristics were analyzed in the logistic regression method of machine learning. We are trying to measure the relationship between these data and the data of the CDP device based on electrochemical mechanisms, to check the effectiveness of this device. Results showed the logistic regression method recorded an accuracy rate of 81%. This study indicates significant relationship between the selected characteristics of cancer cells and the data recorded by the cdp device. To improve the accuracy of the method, increasing the number of patients is very effective.

Keywords: machine learning , breast cancer , early diagnosis , prognosis

ICB11-1401-12-1058

Identification of Candidate Genes Associated with Cisplatin Resistance in Ovarian Cancer

Zahra Fadavi[1], Sana Kerachian[2], Hadi Alizadeh[3], Seyed javad Mowla[4], Bahram mohammad Soltani[5*]

[1]Tarbiat modares. [2]Tarbiat modares. [3]Tarbiat modares. [4]Tarbiat modares. [5*]Tarbiat modares.

*Corresponding author(s). E-mail(s): soltanib@modares.ac.ir;

Abstract

Among women's malignancies, ovarian epithelial cancer is the fourth cause of women's death among all types of cancer in the world. Poor prognosis is estimated to be the main reason for the high death rate. Most of the patients with advanced stage ovarian cancer, experience cancer recurrence, most of these tumors are resistant to cisplatin. Cisplatin is a chemotherapy drug, and differences in the DNA damage response provide a useful explanation for its early sensitivity in ovarian cancer. While the body's normal cells mostly can deal with the damage of restricted dosage of the drug, tumor cells that lack proper DNA repair system cannot and die. However, cancer cells go through several processes to become drug resistant. we analyze the differently expressed genes in RNAseq PRJNA384352 (A2780S vs A2780CP) and PRJNA735397 (chemotherapy-resistant vs. chemotherapy-sensitive A2780 cell lines) and PRJNA724680 (A2780S vs A2780CP) and PRJNA594211 (ovarian cancer cisplatin resistance vs sensitive tissues). By intersecting the differently expressed genes, 314 genes were identified. Then, further function enrichment analysis, including GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis, was performed. In this study we aimed to identify possible related pathways and genes in platinum based ovarian cancer drug resistance. Finally, we find out that TMEMs (Transmembrane protein family members) may play a vital role in chemoresistance.

Keywords: ovarian cancer, chemoresistance, pathway analysis

ICB11-1401-12-1059

In Silico Identification of Novel Natural Anti-Cancer Compounds Targeting ITGAV Protein

Adel Barahooie[1*], Parisa Zare[2], Shiva Barahooie[3]

[1*]Zahedan university of medical sciences.[2]Zahedan university of medical sciences. [3]Zabol university of medical sciences.

*Corresponding author(s). E-mail(s): adel.barahooie.y@gmail.com;

Abstract

The alpha-V (ITGAV) integrins regulate localization and activity of proteolytic enzymes that remodel the extracellular matrix during tumorigenesis and metastasis. ITGAV has been identified to have central role in promoting many cancers; such as Breast cancer, Glioblastoma and Pancreatic adenocarcinoma. Yet no chemically synthesizable drug that specifically target ITGAV protein is on the market. This study aims to identify novel natural Anti-Cancer compounds targeting ITGAV Protein. To start with, a pharmacophore model was built utilizing PharmaGist web server using a database containing eight (8) highly potent ITGAV inhibitors that were collected from ChEMBL. Different pharmacophore models were generated and best model was selected based on the high score of 28.191 and maximum number of aligned compounds of eight. Selected model was screened against the ZINC natural database using ZINCPHARMER to find potential drug candidates and resulted in 218 compounds. Retrieved database from ZINCPHARMER further evaluated by docking studies and top 2% compounds (eight compounds) showing highest binding affinity were chosen. In order to select compounds structurally different, an attempt to cluster molecules based on their origin compound was done using ZINC database. Four clusters were identified eventually, from each cluster the compound that showed the highest binding affinity was selected for proceeding the study. Later on, ADMET studies were done on selected compounds and ZINC31164979 was identified to be closest to drug-like molecules than other three compounds due to its high oral bioavailability in comparison with ZINC68601232 and ZINC02121010; meanwhile it's oral administration wouldn't be toxic in despite of compounds ZINC11865573 and ZINC02121010. To confirm stability of the selected drug candidate to the target protein; the MD simulation approach were employed, which confirmed stability of the selected compound in complex with ITGAV protein. Therefore, newly obtained molecule (ZINC31164979) may serve as lead compound for treatment of ITGAV related cancers.

Keywords: Pharmacophore modeling, Docking, Molecular Dynamics Simulation, ADMET, Integrin Subunit Alpha V

ICB11-1401-12-1060

Modeling guanylyl cyclase activating protein 1 structure

Mohammad hossein Habibi[1], Alireza Ghorbanpour khaljiri[2*]

[1]University of helsinki. [2*]Tarbiat modares university.

*Corresponding author(s). E-mail(s): alireza_ghorbanpour@modares.ac.ir;

Abstract

Guanylyl cyclase activating proteins (GCAPs) are members of the calmodulin superfamily that control the responsive activity of retinal guanylyl cyclase (RetGC) in rod and cone cells in response to Ca2+. Intracellular Ca2+ is detected by GCAPs and modulates photo-transduction in retinal rods and cones cells; GCAPs have four EF-hand motifs and a covalently attached N-terminal myristoyl. Mg2+ binds to GCAP1 in light-activated photoreceptor cells. This Ca2+-free/Mg2+-bound GCAP1 activates RetGC . Herein, we present a model structure of human GCAP1. The sequence of human GCAP1 was obtained from the Uniprot database and an alignment analysis against PDB database protein sequences using NCBI Blast (Altschul and Gish,1990) was performed in order to detect the highest similar protein structure. According to Blast's results, the highest similar structure was the Bos taurus NMR structure of GCAP1 chain A (PDB ID:2NA0) with 99% query coverage and identity of 93.63% (E-val = 4e-140). Modeler software (v 10.4) was used to model human GCAP1 from the Bos taurus based on homology modeling. Among the 5 constructed models, the first model possesses the best DOPE score, -18713.90820. To refine the modeled structure, molecular dynamic simulation was performed using Gromacs 2020.2 using AMBER99SB forcefield (100 ns, timestep: 2fs). The protein structure was solvated in TIP3P water molecules. Using the Nose-Hoover thermostat and Parrinello-Rahman barostat, the pressure and temperature were maintained constant for the systems' equilibration. The potential energy was -1.5535388e+06 and the maximum force was 9.3511957e+02 on 1079 atoms. Finally, the refined model structure was analyzed using the PDBsum database.

Keywords: Guanylyl cyclase activating proteins, Homology modeling, Molecular Dynamic Simulation

ICB11-1401-12-1061

Genetic analysis of a missense mutation in the NANOS2 gene involved in Non-Obstructive Azoospermia based on in-silico pathogenicity prediction tools

Reihane Ghorbanian[1], Mehri Khatami[2*]

[1]Yazd university. [2*]Yazd university.

*Corresponding author(s). E-mail(s): m.khatami@yazd.ac.ir;

Abstract

According to the World Health Organization, infertility is becoming a growing issue worldwide. Infertility is reported to affect 50 million couples universally, and male infertility affects approximately $\frac{1}{2}$ of the infertile cases. One of male infertility causes is Azoospermia which is characterized by the absence of sperm in male semen liquid. Clinically, Non-Obstructive azoospermia is considered the most severe phenotype of male infertility with spermatogenesis failure. Statistically, non-obstructive azoospermia includes 1% of the male population and 10% of infertile men. Among all causes leading to non-obstructive azoospermia, genetic anomalies play a remarkable role in 25% of the cases; yet, the etiology of some cases is still unknown, which is considered as an "idiopathic" condition. The NANOS2 gene is one of the genes associated with the disease. As a member of the transcription factor family, the NANOS2 gene functions by binding to mRNA and also is located on chromosome 19q13.32. In this study, a NANOS2 variant is investigated in terms of alterations in protein structure and function. For a more specific prediction, the stability, hydrophobicity, pathogenicity and conservation of the protein is investigated through bioinformatic databases like HOPE, Mutation assessor and Mutation taster. Among all 128 identified polymorphisms of the NANOS2 gene listed in the National Center for Biotechnology Information/Single Nucleotide Polymorphism database (NCBI/dbSNP), a missense variant selected (rs138351361), and assessed via bioinformatic tools such as HOPE, Mutation Assessor and Mutation Taster, and also studied in terms of protein stability, hydrophobicity, pathogenicity and conservation. By utilizing HOPE, Mutation Assessor, and Mutation Taster bioinformatic tools, useful information about protein pathogenicity and deleterious effects of an alteration in the amino acid chain caused by a missense mutation is detected. A putative G105S pathogenic mutation is predicted through bioinformatic databases.

Keywords: Bioinformatic analysis, Genetic mutation, Male infertility, NANOS2 gene

ICB11-1401-12-1062

Phytochemicals as natural anti-obesity agents: Molecular docking and computer-aided drug design

Parisa Imeni[1*], Milad Darejeh[2]

[1*]University of tehran.[2]University of gorgan.

*Corresponding author(s). E-mail(s): imeni.parisa@gmail.com;

Abstract

Obesity is a serious health problem in almost all parts of the world. Excessive weight gain leads to the occurrence of various other health problems such as diabetes, cancer, respiratory and cardiovascular diseases. Various anti-obesity chemical drugs, such as orlistat, fenfluramine, and coreaserin, have various adverse health effects, such as hypertension. Therefore, more efforts should be made to find natural anti-obesity remedies that have fewer side effects and are more efficient. Accordingly, several natural secondary metabolites found in various plants, such as polyphenols, have been reported to be effective against obesity. In this study, ten plants commonly used as anti-obesity agents, including green tea, mint, cardamom, chicory, alfalfa, urtica dioica, caraway, ginger, and thyme, were selected as study subjects based on a literature search. The major phytochemicals of these plants were collected in a ligand library by gas chromatography-mass spectrometry analysis and corresponding articles. In studying the mechanism of action of anti-obesity drugs, five enzymes such as pancreatic alpha-amylase, pancreatic lipase, fatty acid synthase, lysosomal alpha-glucosidase, and glucagon-like peptide1 receptor were selected as drug targets. Using online databases (Pubchem, PDB, drug bank, binding db) and computer-aided drug development software (Chimaera, Pyrx, Discovery Studio, Padel Descriptor, SMLR, and Chemoface), molecular docking studies, quantitative structure-activity relationship studies, and IC50 predictions were performed for the desired phytochemicals. Alfalfa, green tea and urtica dioica have the most suitable profile and their major constituents such as galangin, rutin, hesperitin and naringin had a significant role in inhibiting the target enzymes compared to the positive control drugs and they are also within the appropriate range of medical dosage. Considering the amazing effect of alfalfa plant and the less common use of this plant as food and medicine, it seems that it can be a suitable alternative to chemical drugs in the treatment of obesity.

Keywords: Obesity, Anti-obesity drugs, Natural anti-obesity, Molecular docking, QSAR, Alfalfa, Green tea, Urtica dioica

ICB11-1401-12-1063

An efficient convolutional neural network for diagnosis of Alzheimer's disease

Maryam Babaei[1], Elmira Mirzabeigi[2*], Kourosh Parand[3]

[1]Shahid beheshti university. [2*]Tarbiat modares university.[3]Shahid beheshti university.

*Corresponding author(s). E-mail(s): e.mirzabeigi@modares.ac.ir;

Abstract

Alzheimer's disease (AD) is a neurodegenerative disease that causes many brain functions to weaken. Early AD diagnosis has attracted much research attention and immediate additional medical treatment to prevent its progression. Lately, the use of deep learning for the early identification of AD has generated much interest. This research develops a deep learning-based pipeline for accurate diagnosis and stratification of AD stages. The proposed analysis pipeline involves data pre-processing and developing Convolutional Neural Network (CNN), respectively. First, for making 2D brain MRIs, 3D MRI pictures are slid at a particular location and ready for classification,. After pre-processing, MobileNet was applied to 40000 MRI images of males and females aged 50 to 100. This dataset consists of two classes, people who do not have and have AD. The MobileNet adjusts the trade-off between accuracy and computational load with the help of depth-wise separable convolution. This separable convolution decomposes the convolution layer into depth-wise and pointwise convolution for computation. This feature causes the process to require less memory space and provides good quality and faster results compared to other large models such as VGG16, inception, etc. In addition, the compact size of MobileNet makes it effective in medical issues, mobile, embedded vision applications, and IOT applications in many fields . The experimental analysis showed that the proposed model was able to achieve 94.6% accuracy in two-classes classification. The findings show that MobileNet can be used by doctors to classify and predict MRI images in neurodegenerative diseases such as AD.

Keywords: Alzheimer's disease, Early diagnosis, Convolutional neural network, MobileNet, MRI image

ICB11-1401-12-1064

Identification of key genes and regulatory mechanisms in Lupus nephritis pathogenesis

Parvin Yavari[1], Amir Roointan[2*], Alieh Gholaminejad[3]

[1]Isfahan university of medical sciences, regenerative medicine research center. [2*]Isfahan university of medical sciences, regenerative medicine research center.[3]Isfahan university of medical sciences, regenerative medicine research center.

*Corresponding author(s). E-mail(s): Roointan@res.mui.ac.ir;

Abstract

Lupus nephritis (LN), one of the main causes of end stage renal disease and mortality, is an autoimmune disease with the usual hallmark of glomerulonephritis . Glomerular repositioning of immune complexes, and T and B cells overactivation are both involved in the LN pathogenicity . The aim of the present study was to investigate key drivers of LN pathogenesis by analyzing a related expression profile (GSE112943) to discover some potential drug targets for this malignancy. Firstly, dataset quality check was performed by principal component analysis, and control and disease samples were perfectly separated. Using Linear Model for Microarray Analysis (Limma) algorithm, 7532 differentially expressed genes were identified. Next, the differentially expressed genes were subjected to Gene Ontology and Reactome pathway enrichment analyses to explore the disease related regulatory mechanisms. MAPK cascade, Neutrophil activation involved in immune system, Neutrophil degranulation, and immune system were among the top enriched terms. These results were in line with previous studies on the immune system deficiency in LN patients. Furthermore, a protein-protein interaction network was constructed, and top genes were selected based on their degree of centrality. In parallel, weighted gene co-expression network analysis (WGCNA) algorithm clustered the co-expressed genes into 13 distinct modules, and the blue module was selected as disease most correlated module. To improve reliability in hub gene selection, we explored the top genes spotted by the network analysis and WGCNA. As a result, 8 hub genes including EP300, ITGAV, NUP153, VES 1, RAF1, MAP2K1, HIF1A, and SIRT1 were identified as key players in LN pathogenesis. Using an integrative approach, this study nominated some of the main drivers of this complex disease which could be validated through wet lab trials to reach potential drug targets.

Keywords: Lupus nephritis, WGCNA, Therapeutic target, Hub genes

ICB11-1401-12-1065

Design and molecular docking studies of biuret derivatives as potential Matrix Metalloproteinases 9(MMP9) inhibitors

Seyedeh faezeh Hosseini[1], Hafezeh Salehabadi[2*], Neda Adibpour[3], Majid Pourzamani molan[4]

[1]Zanjan university of medical sciences. [2*]Zanjan university of medical sciences.[3]Zanjan university of medical sciences. [4]Zanjan university of medical sciences.

*Corresponding author(s). E-mail(s): hsalehabadi@zums.ac.ir;

Abstract

Introduction MMP-9 is an important and essential protease from the large family of matrix metalloproteinases (MMPs), which belong to the category of endopeptidases dependent on zinc metal. MMP-9 has diverse functions. It is a key role in pathophysiology, as well as an important role in cleaving extracellular matrix components . In addition, MMP9 has significant effects on various cellular pathways, including migration, invasion, angiogenesis and metastasis of cancer cells. Dysregulation of expression or overactivation of this enzyme is associated with many pathogenesis conditions such as neurological diseases, cardiovascular diseases and especially cancer. Therefore, it is considered as an attractive target for the drug design . In recent years, the application of computer aided drug design has been increasing significantly due to advantages such as reducing costs, increasing the speed of operation and high accuracy . In connection with our interest in the synthesis of biuret derivatives , and advantages of computational chemistry, MMP9 inhibitory activity of 18 biuret derivatives was investigated by molecular docking studies, and the best compounds were selected to evaluate the enzymatic assay. Method At first, swiss target prediction server was applied to identify proposed therapeutic goal for biuret derivatives. Afterward, the Crystal structure of MMP9, with the PDB ID of 4XCT and resolution of 1.3 Aº was obtained from Protein Data Bank (www.rcsb.org). Eventually, molecular docking studies was carried out and compounds with best docking score have been selected for synthesis and enzymatic assay. Results and Discussion According to the molecular docking studies data, the binding energy and main interactions between the biuret derivatives and MMP9 active site were precisely investigated. Based on docking score 4 biuret derivatives were selected. These compounds can be considered as a proper candidate in order to develop new MMP9 inhibitors.

Keywords: Molecular docking studies, Matrix metalloproteinases 9, Inhibitor, Biuret derivatives

ICB11-1401-12-1066

Machine Learning Algorithms Capable of Type 2 Diabetes Mellitus Early Diagnosis using Explored Important Features

Samin Babaei rikan[1], Ali Ghafari[2*], Reza Ghafari[3], Amir Sorayaie azar[4]

[1]Urmia university. [2*]Tehran university of medical sciences.[3]Urmia university of medical sciences. [4]Urmia university.

*Corresponding author(s). E-mail(s): aligh1101291@gmail.com;

Abstract

Diabetes Mellitus is a choric metabolic disease that according to World Health Organization (WHO), is the cause of death of more than 1.6 million people. In this study, using the Random Forest algorithm, the six most important features of twenty features of the public dataset of type 2 diabetes patients was determined and by using extracted features six machine learning algorithms, namely Logistic Regression (LR), Support Vector Machine (SVM), K Nearest Neighbors (KNN), Decision Tree (DT), Extremely Randomized Trees (ERT), and XGBoost were developed. Their performance in diagnosing diabetes was trained and tested using 4-fold cross-validation and hold-out approaches (with 25% of the data excluded from the training process for testing). Accuracy of the LR, SVM, KNN, DT, ERT, and XGBoost algorithms were 92.31%, 90.77%, 96.15%, 95.38%, 95.92%, and 96.92%, with XGBoost outperforming the rest of the algorithms. Considering the F1-Score metric, LR, SVM, KNN, DT, ERT, and XGBoost algorithms achieved 93.72%, 92.21%, 96.77%, 96.15%, 96.44%, and 97.44% results, confirming the performance of the XGBoost algorithm based on the accuracy metric. Also, in addition to results acquired with a 4-fold cross-validation approach, the XGBoost algorithm offers better performance regarding the accuracy and F1-Score metrics. Through hold-out cross-validation approach, accuracy of the LR, SVM, KNN, DT, ERT, and XGBoost algorithms were 92.31%, 93.08%, 95.38%, 95.38%, 94.62%, and 96.15% and F1-Score of the LR, SVM, KNN, DT, ERT, and XGBoost algorithms were 93.90%, 94.34%, 96.30%, 96.20%, 95.65%, and 96.86%, respectively. XGBoost algorithm was capable of diagnosing type 2 diabetes outperforming other algorithms evaluated in this study using the most important features (age, gender, polyuria, polydipsia, sudden weight loss, and partial paresis) validated using 4-fold and hold-out cross-validation methods. This algorithm can act as a supplementary tool for the faster and early diagnosis of type 2 diabetes.

Keywords: Diabetes, Artificial Intelligence, Machine Learning, XGBoost

ICB11-1401-12-1067

Diagnosing MonkeyPox from Skin Indications using Artificial Intelligence Method

Amir Sorayaie azar[1*], Ali Ghafari[2], Reza Ghafari[3], Faraz Changizi[4], Erfan Abdollahizad[5]

[1*]Urmia university.[2]Tehran university of medical sciences. [3]Urmia university of medical sciences. [4]Shahid beheshti university of medical sciences. [5]Shahid beheshti university of medical sciences.

*Corresponding author(s). E-mail(s): Sorayaie20@gmail.com;

Abstract

Following the first reported case to World Health Organization (WHO) on 7 May 2022, an ongoing outbreak of monkeypox, a viral zoonosis, has faced public health with a new challenge globally that has shown public health importance on a global scale. In this study, we aim to leverage the promise of Artificial Intelligence (AI) models to diagnose monkeypox disease using images of skin manifestations. An open-access dataset containing skin images of monkeypox, chickenpox, measles, and normal patients was used in this study. After preprocessing and image augmentation, an overall number of 1754 images has been chosen (80% for tarin and 20% for test). Then a modified pre-trained DenseNet-121 model was utilized for the four-class classification of mentioned diseases. Achieved results show that the modified pre-trained DenseNet201 model offers an average accuracy of 99.04%, precision of 97.85%, sensitivity of 97.66%, specificity of 99.21%, and F1-Score of 97.61%. Our proposed model offers promising performance for differentiating monkeypox infected patients from chickenpox, measles, and normal patients. Our proposed model can be used in hospitals and clinics to help physicians accurately diagnose monkeypox disease.

Keywords: Monkeypox, Artificial Intelligence, Skin Indications

ICB11-1401-12-1068

Identifying key modules and genes associated with Focal and Segmental Glomerulosclerosis Pathogenesis using an integrated WGCNA and PPI network analysis

Saba Shafieizadegan[1], Amir Roointan[2*], Alieh Gholaminejad[3]

[1]Regenerative medicine research center, isfahan university of medical sciences, isfahan, iran. [2*]Regenerative medicine research center, isfahan university of medical sciences, isfahan, iran.[3]Regenerative medicine research center, isfahan university of medical sciences, isfahan, iran.

*Corresponding author(s). E-mail(s): roointan@res.mui.ac.ir;

Abstract

Focal segmental glomerulosclerosis (FSGS) is a heterogeneous renal histopathological disorder that is highly prevalent globally and primarily affects the glomerulus. Despite extensive research, the molecular basis of focal and segmental glomerulosclerosis (FSGS) is yet to be elucidated. This study aimed to analyze an FSGS-related microarray dataset (GSE108109) from the Gene Expression Omnibus database in order to uncover the top pathways and molecules involved in the pathogenesis of this disorder. After quality checking, normalization, and analysis of the dataset, a protein-protein interaction network was constructed using 3315 significant differentially expressed genes (DEGs) . Hub molecules were identified in the network based on their degree of centrality. WGCNA was used to group differentially expressed genes (DEGs) into nine co-expression modules . hub molecules were selected based on module membership and gene significance values. The turquoise module, which was most correlated to the disease, was then chosen for further enrichment analysis. Gene ontology and Reactome pathway enrichment analyses revealed that Th17 cell differentiation, Apoptosis, B cell receptor signaling pathway, Th1 and Th2 cell differentiation, and Fc gamma R-mediated phagocytosis were among the top enriched terms for the turquoise module's DEGs . Nine hub molecules including CD34, ITGA5, TGFB1, GATA3, GNAI2, CYBB, LYN, NOTCH1, and CDH5 were identified as primary molecules shared between the two methods. These nine molecules and associated pathways have been proposed as potentially contributory to the pathogenesis of FSGS.

Keywords: Glomerulosclerosis, Segmental glomerulosclerosis, Weighted gene co-expression network, Drug target

ICB11-1401-12-1069

Bioinformatics-based secretome mining of four Meloidogyne species to recognize effective excretory/secretory proteins for plant pesticide targets

Peivand Heidari[1], Shadi Heidari[2], Baharak Heidari[3*]

[1]Science and research, islamic azad universiy. [2]Science and research, islamic azad universiy. [3*]Islamic azad university of kermanshah.

*Corresponding author(s). E-mail(s): b_heidari@iauksh.ac.ir;

Abstract

Root-knot nematodes are important pathogens of plants worldwide. Excretory/secretory proteins (ES), including proteases, play an important role in the invasion and pathogenicity of these nematodes. Limitations in the use of chemical pesticides increase the desire to investigate plant compounds as pesticides. Therefore, the aim of the present study was to predict the ES proteins of four nematode species as potential targets for parasite control. 76617 sequences from four Meloidogyne species were retrieved from the WormBase Parasite Web Server. The bioinformatics pipeline "SignalP, SecretomeP, TMHMM, Phobius, TargetP, PS-SCAN, PredGPI, and cd-hit" predicted 3643 non-redundant ES proteins, including 738 sequences (6/8%) for M. graminicola, 1230 (2/4%) for M. enterolobii, 834 (5/8%) for M. hapla, and 841 (3/9%) for M. incognita from the total proteomes. ES proteins were mapped to Gene Ontology (GO) terms and annotated into three categories using the PANNZER2 web server. GO term enrichment analysis was performed by WEGO by χ^2 test, using the whole proteome as the reference group. Mapping putative ES proteins to KEGG pathways using the KAAS v2.0 server revealed the most representative pathways in all classes. To identify proteases, ES proteins were scanned against the MEROPS database using BLASTP, and 15 ES protein sequences associated with pathogenicity and trophic genes were identified as potential pesticide targets. Using PEPPI server, the protein-protein interaction of these proteases was investigated on 12 families of plant protease inhibitors available in MEROPS, and molecular docking was analyzed through a chemoinformatics approach. A pathogenic protease with a known 3D structure (Cathepsin L-like protease) in two species, M. graminicola and M. incognita, interacted with a cysteine proteinase inhibitor. A trophic protease (trypsin-like protease) was identified in 4 species as a target for 3 serine proteinase inhibitors. Therefore, these compounds can be experimentally tested by direct assay or by expression in transgenic crop plants.

Keywords: Root-knot nematodes, Secretome mining, Proteases, Pesticide targets, Proteinase inhibitor

ICB11-1401-12-1070

Comparison of Sequence- Based Methods for sequence–structure alignment using the results of Picea abies L. genomics analysis

Baharak Heidari[1*], Shadi Heidari[2], Peivand Heidari[3]

[1*]Islamic azad university of kermanshah.[2]Science and research, islamic azad universiy. [3]Science and research, islamic azad universiy.

*Corresponding author(s). E-mail(s): b_heidari@iauksh.ac.ir;

Abstract

Mapping corresponding residues between the target sequence and template structure—a process referred to as "sequence–structure alignment (SSA)"—is the first step in homology modeling. In a Picea abies L. genomics analysis, 34% of the examined sequences had weak or no homology against PDB database . So we applied sequence-based methods for SSA on their results. Sensitivity and accuracy of alignments were compared to the result of DALI based on 3D structure comparison method. Results revealed 3 levels of sequence similarity based on decreasing difficulty of identifying homology from sequence: "midnight" zone (<15% sequence identity), "twilight" zone (15–25%), and "daylight" zone (> 25%). BLAST searches against PDB sequences led to an accurate alignment when sequence similarity was in daylight zone with no gaps. When there was a gap, alignment with gaps placement and boundary adjustments gave better results and MSA (multiple sequence alignment) "MAFFT" helped refine target-template alignment. For "twilight" zone and related to distant evolutionary relationships, pairwise sequence comparison was not sufficient to reliably identify homology. So comparison of sequence profiles (PSI-BLAST) and HMMs "hidden Markov model" (HMMER_jackhmmer) with database sequences was done. Also, MSA method that uses additional information (predicted secondary structure, 3D structural information) was used by 3DCoffee. Evolutionary relationships that were too distant to be detected either by pairwise sequence or by profile or HMM–sequence comparisons were identified by methods that are based on profile–profile (COMA) or HMM–HMM (HHsearch) alignments. Results showed while BLAST did not detect some proteins, such as ARC6 "response to gibberellin," COMA, HMMER, and PSI-BLAST had the best consistency with results derived from Dali. Depending on the type of protein family, alignment accuracy was different with each method, and HHsearch generated accurate and fast alignments. Protein families having more diverse and intermediate homologs resulted in more accurate alignments.

Keywords: Sequence-Based Methods, Sequence–structure alignment, Sequence similarity, Picea abies L

ICB11-1401-12-1071

Investigating The Second and Third Structure of Squalene Synthetase Transferase Gene in Dracocephalum kotschyi Plant

Maryam Mehdizadeh hakkak[1*], Masoud Tohidfar[2]

[1*]Shahid beheshti university.[2]Shahid beheshti university.

*Corresponding author(s). E-mail(s): maryammehdizadeh4969@yahoo.com;

Abstract

SQS is an unsaturated compound containing six double bonds, the richest source of which is shark liver oil, which traditionally has many uses and its main use is medicine. It is very important to predict the secondary and tertiary structure of proteins in subsequent protein studies and to study and identify the efficiency of unknown proteins. Also, prediction of protein tertiary structure can be used in molecular docking. Due to the important role of SQS in the defense system, it is important to investigate the secondary and tertiary structure of this receptor. In this study, the Phyre2 software was used to investigate the secondary structure of the SQS protein. The results indicate that six similar structures were found in the protein database for SQS, one of these structures, called the crystal structure of a d1ezfa, had a similarity of 92%. Three-dimensional structure modeling was performed based on the selection of a pattern with a high resemblance to the target protein using the Swiss Model database. Three-dimensional structure modeling was performed based on the selection of a pattern with a high resemblance to the target protein using the Swiss Model database. The model chosen for modeling SQS protein in Homo sapiens (FARNESYL-DIPHOSPHATE FARNESYLTRANSFERASE) (1EZF) contains 340 amino acids and discovered by X-RAY DIFFRACTION with a resolution of 2.15 angstroms. The Identity of 1EZF pattern with target, protein is 30.78. The results of this research can be used in future research and molecular docking and provide basic information to investigate other immune receptors.

Keywords: 3D structure, 2D structure, Swiss Model Phyre2, Docking

ICB11-1401-12-1072

Bioinformatic Investigation Of Protein Stability (SQS) In Prokaryotes and Eukaryotes

Maryam Mehdizadeh hakkak[1*], Masoud Tohidfar[2]

[1*]Shahid beheshti university.[2]Shahid beheshti university.

*Corresponding author(s). E-mail(s): maryammehdizadeh4969@yhoo.com;

Abstract

Squalane is a triterpene hydrocarbon found in olive, palm, safflower and rice bran oils. Squalane is an unsaturated triterpene that has wide applications in pharmaceuticals. In this research, squalene production and bioinformatic analysis of its gene in dracocephalum kotschyi plant as eukaryote and in Stigmatella aurantiaca bacterium as prokaryote were investigated in order to investigate the difference of this gene in eukaryote and prokaryote. Among the medicinal uses of squalene, we can mention its use as an adjuvant of vaccines, inhibitor of cancer cells and antioxidant property. GC percentage of SQS protein was evaluated by GC Content Calculator, as well as aliphatic index and instability index by Protparam. The results of the study showed that the percentage of GC was 60% in the studied eukaryote and 63% in the studied prokaryote, which is close to each other in both cases. The higher the GC percentage of the examined protein, the more stable that protein is. Also, the instability index and aliphatic index for SQS- breeding in the studied eukaryote were higher than 40 (54/24) and less than 100 (84/94), respectively, and the studied prokaryote was 48.42 and 15.62, respectively. The instability index in the two investigated groups is close to each other, but the aliphatic index in prokaryotes is much lower than in eukaryotes. The instability index in the two investigated groups is close to each other, but the aliphatic index in prokaryotes is much lower than in eukaryotes. But the results of the investigation showed that both in prokaryotes and in eukaryotes, SQS protein is one of the semi-stable proteins.

Keywords: GC Percentage, Isoelectric, Aliphatic Chains, Relative Protein Volume, SQS Protein

ICB11-1401-12-1073

Saccharum spontaneum transcriptome analysis for identification of the potential genes involved in receptor-mediated nonhost resistance

Shadi Heidari[1], Peivand Heidari[2], Baharak Heidari[3*]

[1]Science and research university of tehran. [2]Science and research university of tehran. [3*]Islamic azad university of kermanshah.

*Corresponding author(s). E-mail(s): b_heidari@iauksh.ac.ir;

Abstract

Nonhost resistance (NHR) protects plants against a wide range of non-adapted pathogens, which implicates a potential exploitation as a source for novel disease resistance strategies. Recently, promising strategies for developing durable resistant crops have been provided with the introduction of two gene groups to plants, including membrane-localized pattern-recognition receptors (PRRs) and nucleotide-binding LRR receptors (NLRs), both of which are associated with NHR. S. spontaneum (Sspon), the polyploidy progenitor of sugarcane, is considered a valuable resistance source to various stresses. However, little has been reported on the genes involved in the receptor-mediated NHR. With the aim of identifying potential genes, transcriptome data from the Saccharum project under different tissues and conditions was investigated. Amino acid sequences of known genes with grouped domains were used as queries in the search for gene homologs and analogs in Sspon transcriptome. Reverse alignments were performed using TBLASTN. Matching clusters were annotated in a non-redundant local database. The Expasy Translate Tool was used to predict the correct Open Reading Frames (ORF). Aiming to identify domain patterns, ORFs were submitted to RPS-BLAST against CDD database. Predictions of subcellular localization and transmembrane helix were inferred by TargetP and TMHMM servers. Results revealed at least seven classes of genes in Sspon. PRRs include five classes: class I with three domains "Leucine Rich RepeatXII, Protein Kinase, Transmembrane" (EFR and XA21 genes), class II "LRRXI, PK, TM" (PEPR2), class III "Legume lectin, PK, TM" (LECRK IV.1, IV.2, IX.1), class IV "LRR, TM" (EIX1, EIX2), and class V "LysM, PK, TM" (LYK). NLRs include two classes: class VI "Toll/Interleukin 1Receptor, TM" (TIR) and class VII "NB-ARC, LRR" (RPS2, RPP13). These receptors, which are homologous to previously known resistance genes, are promising candidates for conferring NHR. In addition, novel combinations of gene domains were reported in Sspon, which might represent novel gene structures.

Keywords: Nonhost resistance, Saccharum spontaneum, Transcriptome analysis, Receptors, Protein domains

ICB11-1401-12-1074

Reconstruction of the Co-expression Network in Pediatric Low-grade Glioma

Sara Tavallaei[1], Fariba Dehghanian[2*]

[1]Isfahan university. [2*]Isfahan university.

*Corresponding author(s). E-mail(s): fa.dehghanian@sci.ui.ac.ir;

Abstract

Background:
Brain tumors, in particular the central nervous system low-grade tumors, are the most common causes of death in children with brain cancers. A different genetic background was detected in children's tumors compared to similar tumors in adults (2,5). Hence, in this study, we aimed to take a closer look at gene modules and signaling pathways related to pediatric low-grade Glioma using the Systems biology approaches.
Methods:
First, the raw RNA sequencing data of 54 samples of pLGG (accession number: E-MTAB-6270) was collected from the ArrayExpress database. Based on the Galaxy online platform, a step-by-step RNA-seq analysis was done to create an expression matrix (1). The expression matrix was used as an input to the WGCNA package in R software to reconstruct the co-expression network and perform the module detection analysis. Then, hub genes of important modules were determined by the CHAT tool in Cytoscape software. At last, gProfiler was used for the functional enrichment analyses (6).
Results:
Among the 19 modules, 2 modules including brown and turquoise modules showed the most association with Glioma. These two modules contain the most important genes previously reported as critical biomarkers in pLGG (2-5). The turquoise module includes IRS1, MLST8, FOXO3, NLRC5, and TRIM6 hub genes, and the brown module contains LTBP1 and SOSTDC1 hub genes. The gene enrichment analysis regarding the signaling pathways ultimately led to the association of the Longevity regulating, mTOR signaling, PIP3 activates AKT signaling pathways with these two modules and mainly Glioma tumors in children.
Conclusion:
Our results suggest new biomarkers in association with pLGG and their relation to some signaling pathways in tumor cells. The exact investigation of these genes among a vast number of genes in each critical signaling will result in developing potential therapeutic approaches.

Keywords: Pediatric Low-grade Glioma, WGCNA, Co-expression network, Signaling pathway

ICB11-1401-12-1075

Stoichiometric Metabolic Modelling of Nitrate Reducing Sulfur Oxidizing Bacteria for Wastewater Application

Panteha Pirieh[1], Fereshteh Naeimpoor[2*]

[1]Iran university of science and technology. [2*]Iran university of science and technology.

*Corresponding author(s). E-mail(s): fnaeim@iust.ac.ir;

Abstract

Background:
Despite the vital role of S- and N-atoms in life, some reduced inorganic sulfur compounds (RISCs) and several N-compounds such as nitrate released into environment have detrimental effects on human health and ecosystems. These compounds therefore should be removed or transformed into less toxic products. Nitrate-reducing sulfur-oxidizing bacteria (NR-SOB) can effectively oxidize RISCs alongside reduction of nitrate under anaerobic condition.
Purpose:
Metabolic modeling of these bacteria provides insights into combination of oxidation, reduction and disproportionation reactions resulting in optimal removal performance of NR-SOB alongside production of desired products.
Method:
A stoichiometric metabolic model of chemolithotrophic NR-SOB was reconstructed which consisted of main metabolic pathways, sulfur metabolisms and nitrate denitrification reactions as given in KEGG, and Metacyc sources. To investigate the oxidation-reduction system under different conditions, intracellular metabolic fluxes were estimated by applying metabolic flux analysis using Cobra-toolbox v3.0 in MATLAB platform 2020a.
Result:
Results showed that sulfide (a completely reduced S-compound), among the four examined S-compounds, led to the highest specific growth rate with sulfate and molecular nitrogen as products. Increasing role of disproportionation reactions beside oxidation-reduction metabolism also resulted in more amassed biomass of NR-SOB. Nitrate uptake rate had a substantial influence on S-products distribution due to shifting of sulfur metabolism. Moreover, decreasing the flux of S-compound to nitrate ratio (S/NO3) from an optimized level in each S-substrate led to incomplete denitrification and toxic gases such as NO or N2O were produced during the process.
Conclusion:
Simulation of NR-SOB using this stoichiometric metabolic model provided insight to simultaneous oxidation of S- and reduction of N-compounds and can have generic application in biological removal of sulfur and nitrate contaminants in industrial wastewater.

Keywords: Denitrification, FBA, Metabolic modeling, Sulfur-oxidizing bacteria, Nitrate-reducing bacteria

ICB11-1401-12-1076

The effect of Alzheimer's drugs on carbonic anhydrase II and brain pH

Najmeh Shadfar[1], Masoumeh Sohrabi[2], Zeinab Tari[3*]

[1]Alzahra university(tehran). [2]Alzahra university(tehran). [3*]Alzahra university(tehran).

*Corresponding author(s). E-mail(s): Z.tari@student.alzahra.ac.ir;

Abstract

Alzheimer's disease is an important central nervous system disorder, in which beta amyloid, accumulation of hyper-phosphorylated tau protein, and oxidative stress are prominent abnormalities. Several factors may influence the process of protein accumulation, such as: increasing temperature, chemical changes, and the effect of pH. Carbonic anhydrase (CAs, EC 4.2.1.1) which catalyzes the reversible hydration of CO2 to HCO3- and protons, plays an important role in pH regulation; Therefore, carbonic anhydrase's inhibitors and activators are clinically important compounds. The Human Carbonic Anhydrase II isoform, which is expressed in the brain, red blood cells, etc., has the highest catalytic activity. AutoDock, AutoDock vina, Chimera and Chem3D software were used to investigate the interaction of carbonic anhydrase enzyme II structure (pdb: 6EQU) with some drug structures extracted from PubChem database. The grid box was set to X=36, Y=34, Z=40 and the number of runs was set to 10 to increase docking accuracy. As carbonic anhydrase is a zinc-containing metalloenzyme; Zinc is attached to Histidine 94, 96 and 119 (metal binding site) which is important for the catalytic activity of the enzyme. Also, Histidine=64, the proton shuttle, plays an important role in enzyme activity. The results showed the binding of Donepezil to Histidine=64,94, Galantamine and Tacrine to Histidine=94,96,119 and Rivastigmine to Histidine=64,94,119, which inhibit carbonic anhydrase enzyme. Memantine was the only drug that did not bind much to the active site of the carbonic anhydrase enzyme. Based on the multiple investigations point to pH reduction resulting in the buildup of amyloid plaques and tau phosphorylation, it appears to be wise when assessing the latest medications utilized to treat Alzheimer's, to also take into account the drug's impact on the carbonic anhydrase enzyme and its aptitude in pH control.

Keywords: Alzheimer's disease, Brain pH, Carbonic Anhydrase**II**, Inhibitor, Molecular docking

ICB11-1401-12-1077

The Molecular Modeling of Vitamin E Arrangement on Fibrous Cellulose Superstructures

Nima Goodarz[1], Ali Javanmard[2], Mehdi Jahanfar[3], Maryam Azimzadeh irani[4*]

[1]Shahid beheshti university. [2]Shahid beheshti university. [3]Shahid beheshti university. [4*]Shahid beheshti university.

*Corresponding author(s). E-mail(s): m_azimzadeh@sbu.ac.ir;

Abstract

Vitamin E refers to a class of fat-soluble antioxidants among which α-tocopherol has the highest bioactivity in humans. The antioxidant capacity of vitamin E is critical in neutralizing the life-threatening consequences of oxidative stress . Vitamin E also has a demonstrated role in cell signaling that distinguishes it from other so-called antioxidants . The abundance and biocompatibility of cellulose along with its exceptional binding properties make this biopolymer an ideal drug carrier . A recent molecular dynamics simulation study has indicated that the formation of two distinctive plane and cylindrical superstructures is feasible when cellulose nanofibers (CNs) aggregate. Compared to the linear cellulose, the atomistic interactions of superstructures show a closer resemblance to natural systems. Herein, computational modeling has been carried out to illustrate the positioning pattern of α-tocopherol on CN superstructures and the formed interactions. The three-dimensional conformer of D-α-tocopherol was obtained from PubChem (PubChem CID: 14985). CN superstructures were collected from a previous study. Molecular docking was performed via AutoDock Vina in the UCSF Chimera platform and figures were prepared with UCSF Chimera. In order to increase the accuracy, several docking sessions were performed with modified exhaustiveness of search to examine enough search points inside docking boxes of variable size. . Molecular docking scores were in the range of -3.5 to -3.0. The resulting structures were sorted into several groups based on the hydrogen bonding between the phenol or ether groups of α-tocopherol and hydroxymethyl or hydroxyl groups of cellulose. This study may lead to a better understanding of how vitamin E would interact with CNs. According to the results, vitamin E can be superficially emplaced at both ends and also in the middle of CNs.

Keywords: Vitamin E, α-Tocopherol, Cellulose nanofibers, Molecular docking

ICB11-1401-12-1078

Molecular docking investigation of vipirinin for colon cancer-related molecular target

Tooba Abdizadeh[1*]

[1*]Shahrekord university of medical sciences.

*Corresponding author(s). E-mail(s): t.abdizadeh@gmail.com;

Abstract

Cancer is affected by a diversity of pathways, which includes different types of enzymes. Among them, Cyclin-dependent kinase-2 (CDK2) enzyme is one of the most common enzymes that are involved in the cancer development. Therefore, CDK2 can be an attractive drug target for the treatment of cancer such as colon cancer. In present study, vipirinin and doxorubicin (reference drug) were retrieved from PubChem server as 3D structures in SDF files. Crystal structure of CDK2 was obtained from Protein Data Bank (PDB) database (PDB ID: 2A4L). Then, these compounds were investigated by molecular docking studies, in terms of free energy binding against the CDK2. The docking results revealed that vipirinin exhibited better binding interaction to CDK2 than the known doxorubicin inhibitor and bind strongly with some of the amino acid residues in the active site of CDK2 and thus could act as potential inhibitory compound against CDK2 protein and require laboratory and experimental investigation.

Keywords: CDK2, Molecular docking, Vipirinin

ICB11-1401-12-1079

A comprehensive analysis of a pseudogene/lncRNA-miRNA-mRNA ceRNA network in esophageal cancer

Milad Daneshmand-parsa[1], Sharareh Mahmoudian-hamedani[2], Parvaneh Nikpour[3*]

[1]Department of genetics and molecular biology, faculty of medicine, isfahan university of medical sciences, isfahan, iran. [2]Department of genetics and molecular biology, faculty of medicine, isfahan university of medical sciences, isfahan, iran. [3*]Department of genetics and molecular biology, faculty of medicine, isfahan university of medical sciences, isfahan, iran.

*Corresponding author(s). E-mail(s): pnikpour@med.mui.ac.ir;

Abstract

Background:
Esophageal carcinoma (ESCA) is often diagnosed at the advanced stages which has a poor survival rate and overall is one of the deadliest cancers worldwide . Recent studies have elaborated the significance of non-coding RNAs like pseudogenes, long non-coding RNAs (lncRNAs) and microRNAs (miRNAs) in cancer progression . This RNAs -if containing the same mRNA response elements (MREs)- can act as competitive endogenous RNAs (ceRNAs) and communicate with each other through networks called ceRNA networks .
Purpose:
In this study, a four-component ceRNA network associated with ECSA was constructed and a ceRNA with a prognostic potential was introduced.
Methods:
Transcriptome profiles of mRNAs, lncRNAs, pseudogenes and miRNAs were obtained from The Cancer Genome Atlas (TCGA) database. A ceRNA network was constructed based on differentially-expressed RNAs. Survival analysis was carried out on a selection of ceRNAs with the highest centrality degree ranks to discover potential prognostic biomarkers.
Results:
A four-component ceRNA network with 529 nodes and 729 edges was constructed. Based on ceRNA network centrality degree ranks, a subset of ceRNAs with a degree more than two was selected for survival analysis. Amongst the ceRNAs that were found to be associated with survival, SYT10 (Synaptotagmin 10) had the highest hazard ratio (hazard ratio=1.4697, p-value=0.026).
Conclusion:
Our study presented a four-component ceRNA network for ESCA which has a potential to introduce survival-associated genes as effective prognostic biomarkers and therapeutic targets.

Keywords: ceRNA, Esophageal cancer, Prognosis, Pseudogene

ICB11-1401-12-1080

Identification of LPIN1 and MTCO1P12 genes as gastric cancer-related potential biomarkers based on a ceRNA network

Milad Daneshmand-parsa[1], Parvaneh Nikpour[2*]

[1]Department of genetics and molecular biology, faculty of medicine, isfahan university of medical sciences, isfahan, iran. [2*]Department of genetics and molecular biology, faculty of medicine, isfahan university of medical sciences, isfahan, iran.

*Corresponding author(s). E-mail(s): pnikpour@med.mui.ac.ir;

Abstract

Background:
Gastric cancer (GC) is one of the most common causes of cancer-related death in the world . Finding new biomarkers in the detection of tumor in early stages is crucial as GC patients are mostly diagnosed at advanced stages . Different kind of RNA molecules can compete and interact with each other as competing endogenous RNAs (ceRNAs) through their microRNA (miRNA) response elements (MREs) which influence tumorigenesis in different cancers such as GC .
Purpose:
In the present study, we constructed and analyzed a three-component ceRNA network in GC for finding potential diagnostic biomarkers.
Methods:
Gene expression data of The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD) dataset including 375 tumoral and 32 normal samples were retrieved using TCGABiolinks R package. Then differentially-expressed mRNAs (DEMs), pseudogenes (DEPs) and miRNAs (DEMis) between tumoral and normal samples with adjusted $p<0.05$ were extracted utilizing DESeq2 R package. Different databases like RNAInter (RNA Interactome Database), miRDB, miRTarBase and TargetScan were utilized for finding and predicting interaction between miRNAs and pseudogenes/mRNAs. Subsequently, Cytoscape software (version 3.8.1) was utilized to construct a three-component ceRNA network including mRNAs, pseudogenes and miRNAs.
Results:
10,145 DEMs, 3,576 DEPs, and 60 DEMis were identified in TCGA-STAD tumoral vs. normal samples with the defined p-value cut off. A ceRNA network including 277 nodes (263 DEMs, 10 DEPs and 4 DEMis) and 284 edges was then constructed. A four-component axis including hsa-miR-105-5p, hsa-miR-122-5p, LPIN1 and MTCO1P12 was selected for further experimental validation.
Conclusion:
Current study gives an outline of DEMs, DEPs and DEMis in GC that communicate with each other by a sophisticated mechanism which is ceRNA network having a potential to introduce ceRNAs as effective diagnostic biomarkers and therapeutic tools.

Keywords: Competing endogenous RNAs, Gastric cancer, microRNAs, Pseudogenes

ICB11-1401-12-1081

Identification of Differentially Expressed Genes in Systemic Lupus Erythematosus Patients, Before and After Initiation of Immunosuppressive Therapy

Shima Rahimirad[1*], Samira Rahimirad[2]

[1*]Shahrekord university.[2]Mcgill university.

*Corresponding author(s). E-mail(s): Rahimirad_sh68@yahoo.com;

Abstract

Systemic lupus erythematosus (SLE) is a chronic autoimmune disorder characterized by producing anti-nuclear and anti-double stranded DNA antibodies, leading to variable clinical symptoms . Due to its genetic and phenotypic heterogeneity, diagnosing and therapies are challenging . Treatment modalities include steroidal and nonsteroidal anti-inflammatory drugs and immunosuppressive agents . In the present study, we evaluated differentially expressed genes (DEGs) in the peripheral blood of lupus nephritis patients, before and after initiation of immunosuppressive therapy. The profile GSE72747 was obtained from gene expression omnibus (GEO) database including 30 samples of SLE patients with renal deficiency, prior to (baseline), 3 months, and 6 months after initiation of immunosuppressive therapy (Corticosteroids, Cyclophosphamide or Mycophenolate or Azathioprine) and DEGs related to each time point was collected using R based packages (v4.2.2). Functional/pathway enrichment analysis was conducted using Clusterprofiler package to identify the biological characteristics of DEGs. The limma package was used to screen the GSE72747 dataset, 57 and 133 significantly DEGs were respectively obtained at 3 and 6 months after treatment as compared with baseline (-1<logFC>1/Pvalue <0.01). Two significantly down-regulated genes (IGHG1, CKAP2) were found in both treatment periods. Also, decreased expression of 14 genes and up-regulation of only one gene were detected 6 months after treatment initiation (Table 1, 2). Moreover, altered genes were highly involved in immune-related processes and neurological/immune-related diseases (e.g. positive regulation of cytokine production). This study through in-silico analysis represented the consistent down-regulation of immune-related genes during the treatment, particularly IGHG1 which is responsible for activation of immune response. The effectiveness of the immunosuppressive therapy in SLE patients may be monitored longitudinally by analyzing the DEGs.

Keywords: Systemic lupus erythematosus (SLE), Immunosuppressive therapy, Differentially expressed genes (DEGs), Gene Expression Omnibus (GEO)

ICB11-1401-12-1082

Association of miR-3916 expression with overall survival in breast cancer and identification key pathways and hub genes using bioinformatics analysis

Zahra sadat Tabatabaei jafari[1], Kambiz Ahmadi[2], Somayeh Reiisi[3*]

[1]Shahrekord university. [2]Shahrekord university. [3*]Shahrekord university.

*Corresponding author(s). E-mail(s): s.reiisi@yahoo.com;

Abstract

Background and aim:
miRNAs are small non-coding RNAs that are important in regulating gene expression using RNA degradation or translational repression . Dysregulation of miRNAs plays a role in the initiation and progression of many cancers, including breast cancer. Also, identification of the main target genes provides a miRNA-related mechanism in cancer and suggests potential treatments. The purpose of this study is to determine the relationship between miR-3916 expression and key pathways in breast cancer, based on the real-time PCR data and computational analysis.
Methods:
The expression of miR-3916 was determined using real-time PCR in breast tumor tissue compare to adjusted normal tissue. multiMiR package in R was used for miRNA target genes prediction . Gene ontology and KEGG pathway analysis were performed to identify the potential function of miR-3916 by the clusterProfiler package in R version 4.2.1 . A PPI network (Protein-protein interaction network) was constructed to display key target genes. For hub genes validation, GEPIA databases were used.
Results:
miR-3916 was upregulated in breast tumor tissues, and its increase was significantly related to a bad prognosis in breast cancer. GO analysis showed that target genes were mainly enriched in the regulation of gene expression and regulation of transcription, DNA-templated. KEGG pathway analysis suggested that target genes were enriched in the neurotrophin signaling pathway, and Wnt signaling pathway. And finally, the ten hub genes were detected from the PPI network in the following order (KRAS, HSP90AA1, MAPK1, FOXO3, IGF1, MAPK8, SMAD3, STAT5B, AR, and MAP2K7).
Conclusion:
This study proposed that miR-3916 up-regulation is associated with breast cancer progression and worse survival, and also identified ten genes associated with breast cancer, which can help to provide candidate targets for the treatment.

Keywords: miR-3916, hub gene, Breast cancer, Bioinformatics analysis

ICB11-1401-12-1083

Supervised Kohonen networks as tools for virtual screening of PubChem database: A case study with cyclin dependent kinases

Sara Kaveh[1], Marzieh sadat Neiband[2], Ahmad Mani-varnosfaderani[3*]

[1]Tarbiat modares university. [2]Payame noor university (pnu). [3*]Tarbiat modares university.

*Corresponding author(s). E-mail(s): a.mani@modares.ac.ir;

Abstract

The Cyclin-dependent kinase (CDK) is a family of serine/threonine kinases that plays essential roles in regulating the cell cycle, transcription, and cell migration. In addition, they control metabolism and apoptosis . Improper regulation of kinases had been clinically proven to be associated with different diseases including cancers, and inflammatory and cardiovascular diseases. CDKs are in eleven isoforms with specific biological roles and identifying the characteristics of their selective inhibitors is of great importance. The main aim of this project is to find a series of general structure-selectivity relationship patterns for CDKs inhibitors. To achieve this goal, 4201 active inhibitors of CDK1, CDK2, CDK4, CDK5, and CDK9 were collected from Binding DB and analyzed using machine learning techniques. A total of 3224 physicochemical properties were calculated for each molecule using DRAGON 5.5 software. As a method for selecting discriminatory molecular features, the variable importance in projection (VIP) approach was used. Counter propagation artificial neural networks (CPANN) and supervised Kohonen networks (SKN) were used for the classification of molecules based on their therapeutic targets. The developed multivariate classifiers were used for ligand-based virtual screening of two million random molecules of the PubChem database. The average values of the enrichment factor (EF10%) for the SKN and CPANN models were 7.03 and 4.70, respectively. In addition, the average values of the area under the receiver operating characteristic (ROC) curves were more than 0.65 and 0.84 for the CPANN and SKN models, respectively. The VIP-selected molecular descriptors in this work defined a well-separated subspace for discriminating molecules based on the isoforms of CDKs. The information obtained from this study is a stepping stone to help pharmacists and medicinal chemists to produce drugs with better efficacy and fewer side effects.

Keywords: Classification, Cyclin dependent kinase, PubChem database, Selective drug design, Structure-activity relationship (SAR), Virtual screening

ICB11-1401-12-1084

Identification of ALB as a potential hypoxia-stemness-associated prognostic biomarker in gastric cancer

Sharareh Mahmoudian[1], Parvaneh Nikpour[2*]

[1]Isfahan university of medical science. [2*]Isfahan university of medical science.

*Corresponding author(s). E-mail(s): pnikpour@med.mui.ac.ir;

Abstract

Background:
Gastric cancer (GC) is among the deadliest cancers worldwide. Many patients are diagnosed in the late stages of the disease, so a poor overall survival is generally observed . One of the hallmarks of solid tumors is hypoxia , which accelerates the formation of stemness characteristics in cancer cells leading to a more severe malignancy .
Purpose:
In the present study, our aim was to introduce a combined hypoxia and stemness signature with prognostic value in GC.
Methods:
We downloaded the stomach adenocarcinoma transcriptome data from The Cancer Genome Atlas (TCGA) database. We then computed the hypoxia and stemness scores using Gene Set Variation Analysis and mRNAsi methods , respectively. Based on these scores, we performed an unsupervised hierarchical clustering and carried out survival analysis for the clusters. Next, we extracted the differentially-expressed genes (log2 fold change>2 and $p<0.05$) and constructed a protein-protein interaction (PPI) network for up-regulated genes. We then performed overall survival (OS) analysis for a gene with the highest centrality degree rank (the hub gene).
Results:
We identified four clusters based on hypoxia and stemness scores (clusters 1 to 4) and by performing survival analysis, we found out that the greatest OS difference was between clusters 1 and 4. We then found 697 down-regulated and 82 up-regulated genes between the two clusters. OS analysis for the ALB gene, the hub gene with the highest centrality degree rank in the PPI network, revealed that its higher expression is associated with the poorer OS (p-value=0.03, hazard ratio=1.2).
Conclusion:
The results of our study indicated that considering hypoxia and stemness scores, as two crucial factors in cancer pathogenesis, can lead to the introduction of potential prognostic biomarkers and therapeutic targets.

Keywords: Gastric cancer, Hypoxia, Prognosis, Stemness

ICB11-1401-12-1085

A comparative meta-analysis and in silico analysis on metabolomics Profiles in Nephrotic Syndrome for Biomarker Discovery

Saba Shafieizadegan[1], Alieh Gholaminejad[2*], Amir Roointan[3]

[1]Regenerative medicine research center, isfahan university of medical sciences, isfahan, iran. [2*]Regenerative medicine research center, isfahan university of medical sciences, isfahan, iran.[3]Regenerative medicine research center, isfahan university of medical sciences, isfahan, iran.

*Corresponding author(s). E-mail(s): a.gholaminejad@res.mui.ac.ir;

Abstract

Nephrotic syndrome (NS) is a renal disorder characterized by excessive proteinuria, hypoalbuminemia, edema, and hyperlipidemia, which can lead to end-stage renal disease and require renal transplantation. Despite a growing number of NS metabolite profiling studies, there is a lack of consistency in their results. The aim of this meta-analysis was to identify a consensus hob panel of significantly dysregulated metabolites in NS that could serve as potential biomarkers. To achieve this, the Amanida package in the R environment was utilized on 10 metabolome profiles obtained from 7 human urine studies . The results panels comprised a volcano plot for quantitative results, a vote plot for the total up- or down-regulation behavior of each compound, and an explore plot of the vote-counting results. Subsequently, 27 metabolites were determined to be dysregulated across 650 samples. Among them, a panel of top meta-metabolites was introduced as biomarkers including up-regulated glucose and fumaric acid with a voting score of $\geq$4 and down-regulated Citric acid, Isobutyric acid, 3-Hydroxyisovaleric acid, and Pyruvic acid with a voting score of $\leq$-4. Subgroup analysis revealed 3, 7, and 2 specific meta-metabolites in focal segmental glomerulosclerosis, membranous glomerulonephritis, and minimal change disease, respectively. Furthermore, enrichment analyses were performed with MetaboAnalyst (Version 0.4) for metabolite set enrichment analysis (MSEA) of the NS meta-metabolites . Enrichment analyses confirmed the involvement of various effective biological pathways in NS pathogenesis, such as the TCA cycle, and amino acids metabolisms. Moreover, in the constructed gene-metabolite and metabolite-metabolite interaction network, citric acid, Pyruvic acid, and glucose were identified as hub metabolites worthy of further exploration as potential drug targets. In conclusion, the identified metabolites are potentially involved in the disease pathogenesis and could be evaluated as biomarkers or drug targets in NS.

Keywords: A comparative meta-analysis and in silico analysis on metabolomics Profiles in Nephrotic Syndrome for Biomarker Discovery

ICB11-1401-12-1086

In silico evaluation of the interactions of CLOCK-BMAL1 complex with IMPDH2

Samira Samadieh[1], Yusuf Elhaei[2], Razieh Yazdanparast[3*]

[1]Tehran. [2]Tehran. [3*]Tehran.

*Corresponding author(s). E-mail(s): ryazdan@ut.ac.ir;

Abstract

The circadian clock is an autoregulatory system that regulates various physiological processes through the generation of approximately 24-hour circadian rhythms in gene expression. The core circadian clock mechanism is mediated by a transcription/ translation-based negative feedback loop composed of CLOCK–BMAL1 transcriptional activators. In addition to regulating of circadian gene expression, the CLOCK protein has intrinsic histone acetyltransferase activity and the ability to acetylate non-histone substrates. IMPDH2, a critical metabolic enzyme in the de novo purine biosynthetic pathway, has been recognized as a substrate for acetylation by CLOCK. Therefore, the objective of this research is to evaluate the CLOCK and BMAL1 interactions with IMPDH2 at protein-protein level. By retrieving the PDB structures of the proteins, the protein-protein docking was conducted in HDOCK server. Then, the obtained docking models were visualized and analyzed in PyMOL software. The Docking Score of the best model in CLOCK-IMPDH2 interaction was -220.64 and this result was -223.40 in BMAL1-IMPDH2 complex. The percentage of the charged R-chain, hydrophobic, and polar non-charged residues were 35.2, 35.2, and 29.4 in the interface of CLOCK-IMPDH2 complex, respectively. These results were 22.8, 42.1, and 35.08 in the interface of BMAL1-IMPDH2 complex, respectively. As results shown the BMAL1 strongly binds to IMPDH2 in comparison to CLOCK. These interactions may provide a good scaffold for acetylation of IMPDH2 by CLOCK-BMAL1 complex as a post translational regulation of this enzyme. The binding of BMAL1 to IMPDH2 is predominantly through hydrophobic residues and this finding could be used in future research on the regulation of IMPDH2 and drug discovery.

Keywords: CLOCK, BMAL1, IMPDH2, Acetylation, Protein-Protein Interactions

ICB11-1401-12-1087

The effect of mating design on genetic gain and inbreeding

Reza Behmaram[1*]

[1*]University of mohaghegh ardabili.

*Corresponding author(s). E-mail(s): rezabehmaram881978@gmail.com;

Abstract

Using appropriate mating designs are important in animal breeding for higher genetic gain in next generations. Thus, the aim of this study was to evaluate genetic gain, increase of average inbreeding and accuracy of prediction using simulated data and different mating designs. QMSim software was used for simulation and creation of population. Two level of heritability (0.05 and 0.45) and five different maing designs including random mating (rnd), mating based on minimum inbreeding (mini), mating based on maximum inbreeding (maxi), positive assortative mating design based on phenotype (phen) and positive assortative mating design based on estimated breeding value (ebv) were assigned. The genetic gain after twenty generation simulation in rnd, mini, maxi, phen and ebv mating designs with heritability 0.05 were 0.756, 0.836, 0.907, 0.792 and 1.061 respectively. On the other hand, for heritability 0.45, genetic gain after twenty generation simulation in rnd, mini, maxi, phen and ebv mating designs were 3.172, 2.688, 3.029, 3.462 and 3.791 respectively. Results showed that after twenty generation simulation, increase of average inbreeding for heritability 0.05 was 0.079 in rnd, 0.026 in mini, 0.316 in maxi, 0.073 in phen and 0.223 in ebv, respectively. In contrast, for heritability 0.45 after twenty generation simulation, increase of average inbreeding was 0.069 in rnd, 0.037 in mini, 0.373 in maxi, 0.085 in phen and 0.163 in ebv, respectively. In conclusion, the genetic gain in mini design was greater than other mating designs per 1% increase of inbreeding, and had higher performance compared with the other mating designs.

Keywords: accuracy, heritability, inbreeding, mating

ICB11-1401-12-1088

The effect of genomic accuracy on interaction between genotype and environment

Reza Behmaram[1*]

[1*]University of mohaghegh ardabili.

*Corresponding author(s). E-mail(s): rezabehmaram881978@gmail.com;

Abstract

Today, thanks to availability of genomic markers and pedigree data, it is possible to investigate (G × E) interaction more than last decades. Therefore, the aim of this study was to evaluate different animal models and genomic scenarios in order to estimate breeding value and genotype × environment (G × E) interaction. Genomic data were simulated to survey variation in QTL number (200 and 400) and linkage disequilibrium (low and high) using 20K SNP panel for 50 chromosomes. The QMSim software was used for simulation and creation of population. Three levels of heritability (0.05, 0.25 and 0.45) were used respectively as low, medium and high levels. Then, low (0.25) and high (0.75) genetic correlations were created between third environments. Results showed that the accuracy of genomic prediction increased with heightening the amount of heritability, linkage disequilibrium and the genetic correlation between the traits. Compared with single trait animal model, multi trait animal model resulted in higher accuracy of genomic prediction. Surprisingly, the highest accuracy (0.44) obtained from the scenario with high linkage disequilibrium and low QTL. Simultaneously, accuracies of genomic prediction increased with higher percentage of animals from 15 to 85. On the other hand, LD level, type of animals in training set, number of phenotypic records in validation set and genetic correlation among different environments were important factors if there was an interaction G × E. In conclusion, G × E interaction can be used for identifying variation of quantitative traits and more accurate genomic selection.

Keywords: accuracy, heritability, linkage, prediction, simulation

ICB11-1401-12-1089

Bioinformatic identification of miR-4645 hub target genes and experimental validation of the miR-4645 in breast cancer

Marziyeh Naderi[1], Kambiz Ahmadi[2], Somayeh Reiisi[3*]

[1]Shahrekord university. [2]Shahrekord university. [3*]Shahrekord university.

*Corresponding author(s). E-mail(s): s.reiisi@yahoo.com;

Abstract

Breast cancer is the most common cancer among women and is the second cancer frequently occurring worldwide.Both genetic and environmental factors play a role in breast cancer development.It has been shown that breast cancer is driven by epigenetic factors such as noncoding RNAs including miRNA.The miRNAs are small non-coding RNAs;regulate gene expression using RNA degradation or translation repression . miRNAs expression profiling was shown to be associated with tumor progression and response to therapy,suggesting their possible use as diagnostic, prognostic and predictive biomarkers .This study aimed to determine the relationship between miR-4645-3p expression and clinical factors based on the qPCR data and regulatory mechanisms in breast cancer.
Methods:
twenty tumor and adjacent non-tumor sample breast tissues were examined in the study. hsa-miR-4645-3P levels were evaluated by qPCR in tumor tissues.For future explore of miRNA function, targets genes of miR-4645 were identified by the multiMiR package in R version 4.2.1.Gene ontology and KEGG pathway analysis were accomplished to identify the biological function of miR-4645-3P.APPI network was constructed to display key target genes.For hub genes validation,GEPIA databases were used.
Results:
miR-4645-3P was downregulated in breast tumor tissues,and its reduction was significantly related to a poor prognosis in breast cancer.For identification of most pathways and genes,firstly 516 target genes were predicted,KEGG pathway analysis proposed that target genes were enriched in long-term potentiation,glioma,ErbB signaling pathway,Oocyte meiosis, melanoma,and apoptosis.Finally,the ten hub genes(MAPK1, MAP2K1,RHOA,JUN, ARRB2,GNAQ,IGF1R,TOP2A, SMC2,CEP55)were detected from the PPI network.
Conclusion:
This study proposed that miR-4645-3p down-regulation is associated with breast cancer progression and worse survival.

Keywords: miR-4645, Signaling pathway, Breast cancer, Overall survival

ICB11-1401-12-1090

A survival-related competitive endogenous RNA (ceRNA) network of lncRNAs, miRNAs and mRNAs associated with gastric cancer

Basireh Bahrami[1], Maryam Hosseini[2], Parvaneh Nikpour[3*]

[1]Isfahan university of medical sciences. [2]Isfahan university of medical sciences. [3*]Isfahan university of medical sciences.

*Corresponding author(s). E-mail(s): pnikpour@med.mui.ac.ir;

Abstract

Background:
Gastric cancer (GC) is known as a highly aggressive malignancy in which environmental and genetic factors can influence its development . Among the genetic factors, competitive endogenous (ce) RNAs can cross-regulate each other through the sequestrating of shared miRNAs and affect the development of cancer .
Purpose:
In the present study, we investigate the relation of ceRNA elements with GC by construction of a competitive network based on RNA-seq data in The Cancer Genome Atlas.
Methods:
The RNA-seq and clinical data of GC patients were downloaded using TCGAbiolinks R-package , including 335 tumor and 30 non-tumor samples. Differentially-expressed long non-coding RNAs (lncRNAs) (DELs), miRNAs (DEmiRs), and mRNAs (DEMs) were extracted by R-package DESeq2 based on | Log2Fold Change| >1 and adjusted p<0.05. To screen the key survival-related DELs, DEmiRs, and DEMs, univariate Cox regression analysis was performed with a threshold of p<0.05. The multiMiR R-package and DIANA-LncBase v3.0 were used to predict the miRNA–mRNA and miRNA–lncRNA interactions. A lncRNA-miRNA-mRNA ceRNA network was then constructed and enrichment analysis was conducted for them using the Enrichr tool .
Results:
We identified 3947 DELs, 266 DEmiRs, and 4388 DEMs, of these, 187 DELs, 24 DEmiRs, and 524 DEMs were associated with the overall survival of GC patients. By integrating the relations with common miRNAs, we constructed a ceRNA network consisting of 12 DELs, 11 DEmiRs, and 70 mRNAs. Our results revealed that these genes were significantly enriched in cancer-related pathways, including Pathways in cancer, HIF-1, Hippo and p53 signaling pathways.
Conclusion:
These results might help characterize the pathogenesis of GC and provide possibilities of introducing ceRNAs as diagnostic biomarkers or therapeutic targets.

Keywords: Gene Expression Profiling, Non-coding RNAs, Stomach neoplasms, Systems Biology

ICB11-1401-12-1091

Comprehensive bioinformatics analysis for Identification of Potential Crucial Genes and Key Pathways in Early Diagnosis of Gastric Cancer: Helicobacter pylori (Follicular gastritis) Induced Chronic Gastritis with Increasing Risk of Gastric Cancer

Sare Golpasand[1], Shahrokh Ghovvati[2*]

[1]Department of Animal Sciences, Faculty of Agriculture, University of Guilan, Rasht, Guilan, P. O. Box: 41635-1314, Iran. [2*]Department of Animal Sciences, Faculty of Agriculture, University of Guilan, Rasht, Guilan, P. O. Box: 41635-1314, Iran.

*Corresponding author(s). E-mail(s): ghovvati@guilan.ac.ir;

Abstract

Chronic gastritis (CG) is an inflammatory disease in which the epithelium of the gastric mucosa is invaded by various pathogenic factors. CG can be divided into three categories: chronic non-atrophic, chronic atrophic, and "special" If not treated in a timely manner, CG can transform into gastric cancer. Microarray analysis has been used for more than 10 years as a reliable technique to probe differentially expressed genes (DEGs) and identify potential clinical biomarkers. In this study, we used integrated bioinformatics to screen for key genes associated with the development of gastric cancer and reveal their potential molecular mechanisms. The gene chip dataset GSE116312 was downloaded from the GEO datasets and R software was used to identify DEGs. Kaplan Meier was used for estimating the survival function. A total of 374 DEGs including 94 upregulated and 290 downregulated were identified. Protein-protein interaction network and module analyses were performed using Cytoscape software. The top 10 hub genes identified from the PPI network were APOB, APOA4, SLC2A2, DPP4, SI, CFTR, SLC5A, DGAT1, EPCAM and CLCA1. GO analysis showed that the biological process of DEGs mainly focused on digestion, transport, gastric acid secretion and cell adhesion. The main cellular components (CC) include basolateral plasmamembrane, apical plasma membrane, cell membrane. Molecular function (MF) include protease, aminotransferase, acyltransferase, lyase and hydrolase.KEGG analysis of the hub DEGs showed that were mainly enriched in the pancreatic secretion, gastric acid secretion, protein digestion and absorption and metabolic pathways. The results of examining the survival of key genes in patients with gastric cancer showed that there is a significant relationship between the genes and the survival of existing patients, except APOA4 and DGAT1.This study provides a new insight into the understanding of the molecular mechanisms associated with gastric cancer and suggested that the hub genes may serve as prognostic predictors.

Keywords: Chronic gastritis and Follicular gastritis, Microarray, Biomarkers, Gene Expression Ombious (GEO), Survial

ICB11-1401-12-1092

Breast milk DNA methylation-based potential prognostic biomarkers of triple-negative breast cancer

Dorsa Hayati[1], Hassan Rajabi-maham[2*], Leila Kianmehr[3]

[1]Shahid beheshti university. [2*]Shahid beheshti university.[3]Shahid beheshti university.

*Corresponding author(s). E-mail(s): H_Rajabi@sbu.ac.ir;

Abstract

Introduction:
The discovery of new biomarkers and experimental treatments is critically needed for triple-negative breast cancer (TNBC) because its prognostic biomarkers and therapeutic targets are unavailable. If abnormal DNA methylation is detected noninvasively, potential biomarkers can be provided for detecting TNBC, thereby offering more effective therapeutic choices to help clinicians. In the GEO database, the obtained information was expression microarray data (GSE65194, GSE38959) and methylation microarray data (GSE133918). The Limma package and Shinyepico package of the R program was employed to identify differentially expressed genes (DEGs) and differentially methylated genes (DMGs). The DEGs and DMGs overlapped to obtain methylation-regulated DEGs. These genes were subjected to functional enrichment analyses by g: Profiler. The Protein-Protein Interactions Network (PPI-Net) construction and sorting of hub genes were performed by STRING and Cytoscape, respectively. Finally, the findings were verified according to The Cancer Genome Atlas (TCGA) database by the survival curves of univariate Cox regression and gene expression box plots. The overlapped DEGs and DMGs identified 87 downregulated, hypermethylated genes and 132 upregulated, hypomethylated genes. The main enrichment of the detected genes was in the biological cell cycle processes, nuclear division, and the negatively regulated PI3K−Akt signalling pathway. Additionally, the identification of 23 hub genes of the PPI-Net was based on the MCODE instrument. Significant alterations in TCGA, including one hypermethylated, downregulated gene c-KIT and three hypomethylated, upregulated genes BUB1B, BIRC5, and CENPE, were observed in the expression and methylation conditions of the four hub genes. Conclusions: Crucial methylation-regulated DEGs were detected in this study, and their associated pathways and biological function were discovered in TBNC. Based on these results, promising prognostic biomarkers may be provided by using breast milk as a biospecimen to evaluate forthcoming diseases.

Keywords: Breast milk, Methylated gene, Hub gene, Biomarkers, TNBC

ICB11-1401-12-1093

Integrated bioinformatics analysis revealed that hub genes correlated with the prognosis and pathogenesis of gastric cancer

Sare Golpasand[1], Shahrokh Ghovvati[2*]

[1]Department of Animal Sciences, Faculty of Agriculture, University of Guilan, Rasht, Guilan, P. O. Box: 41635-1314, Iran. [2*]Department of Animal Sciences, Faculty of Agriculture, University of Guilan, Rasht, Guilan, P. O. Box: 41635-1314, Iran.

*Corresponding author(s). E-mail(s): ghovvati@guilan.ac.ir;

Abstract

Gastric cancer is a major health issue in many countries, with a high prevalence in Asia, Africa, South America and Eastern Europe. Helicobacter pylori is associated with the development of several lesions in the human stomach. This chronic infection produces gastritis, which can progress to intestinal metaplasia and gastric cancer. Next-generation sequencing technologies have become widely used and cancer genomic analysis has recently revealed the relationships between various malignant tumors and genomic information. We aimed to elucidate potential candidate hub genes and key pathways related to gastric cancer based on bioinformatics analysis. GSE 116312 microarray dataset including gastric cancer biopsy and follicular gastritis biopsy were selected from GEO gene expression database and analyzed using GEO2R online tool. The differentially expressed genes (DEGs) between gastric cancer biopsy and follicular gastritis biopsy were analyzed by utilizing the R software. The shared DEGs were analyzed by Gene Ontology (GO) and functional enrichment. Protein-protein interactions (PPI) were constructed by utilizing the STRING database. Hub genes were analyzed by MCODE and Cytohubba. A total of 852 shared DEGs between gastric cancer biopsy and follicular gastritis biopsy samples were identified. Gene function and KEGG pathway enrichment revealed that DEGs were mainly enriched in focal adhesion, gastric acid secretion, protein digestion and absorption and metabolic pathway. Eight modules were clustered based on PPI network analysis. The three main modules with the highest score were identified with a score of 11.52, 7.42 and 7, respectively. Ten hub genes were identified by PPI network analysis respectively include: FN1, IL6, MYC, COL1A1, CXCL8, SOX2, THY1, THBS1, SPP1 and CTGF. Taken together, the identification of these 10 hub genes and enrichment pathways might have important clinical implications for gastric cancer treatment and diagnosis. The hub genes in this research can be considered as potential and valuable prognostic and pathogenesis biomarkers in gastric cancer.

Keywords: Helicobacter pylori, Gene expression, Gastric cancer, Hub genes, Bioinformatic analysis, Microarray data

ICB11-1401-12-1094

The Application of Consensus Weighted Gene Co-expression Network Analysis in Celiac Disease: Identification of Potential Protective Genes and Pathways

Mahmood Fadaie[1], Anis Khalafian[2], Elham Ghafouri[3], Tayebeh Ranjbarnejad[4], Shiva Moein[5*]

[1]Isfahan university of medical sciences. [2]Isfahan university of medical sciences. [3]Isfahan university of medical sciences. [4]Isfahan university of medical sciences. [5*]Isfahan university of medical sciences.

*Corresponding author(s). E-mail(s): shiva.moein@res.mui.ac.ir;

Abstract

Background:
Celiac disease (CeD) is an autoimmune enteropathy triggered by dietary gluten. Almost 90% of CeD patients have HLA-DQ2 or -DQ8 haplotypes. As a high proportion of first-degree relatives (FDRs) of CeD patients have the same haplotype, it is assumed that they are at a higher risk of disease development than the general population. Nevertheless, the prevalence of CeD among FDRs is considerably low (7.5%).
Methods:
In order to figure out this discrepancy, a microarray dataset of intestinal mucosal biopsies of CeD patient, FDR, and control groups was reanalyzed, and gene co-expression network using WGCNA was constructed. Differentially expressed genes in the opposite modules from consensus analysis were obtained. The functional enrichment analysis were carried out.
Results:
WGCNA analysis identified 10 consensus modules associated with CeD and FDR groups, including 5 modules with opposite correlation. Among the genes of opposing modules, there were a total of 1369 DEGs, 159 of which were common between FDR and CeD groups. Functional enrichment analysis revealed potential pathways and candidate genes involved in the host energy metabolism, programmed cell death, actin folding, and antigen cross presentation that might play a compensatory protective effect in the intestinal mucosa of celiac FDR. Conclusions: In current study, our result determined some genes and pathways with protective functions in FDR groups. These genes and molecular mechanisms could be a matter of investigation as potential druggable targets or prognostic markers in CeD.

Keywords: Celiac Disease, WGCNA, Protein-Protein Interaction network, Transcriptome, Systems Biology

ICB11-1401-12-1095

Prediction of key responsive genes to environmental stresses in Thermus thermophilus by using SVM-RFE

Abbas Karimi-fard[1], Abbas Saidi[2*], Masoud Tohidfar[3]

[1]Shahid beheshti university, tehran. [2*]Shahid beheshti university, tehran.[3]Shahid beheshti university, tehran.

*Corresponding author(s). E-mail(s): abbas.saidi@gmail.com;

Abstract

Various environmental agents such as redox, low and high temperature, nutrient availability, heavy metal, hydrogen peroxide, and salt stresses have negative effects on the growth and development of all organisms. Bacteria adapt and illicit a quick response against these environmental changes by changing transcriptome content. They are armed with variety of receptors and signal molecules, enabling them to protect themselves against harsh environmental changes. The identification of responsive genes to these changes is one of the main steps in better understanding of defense mechanisms. To reach this aim, we performed SVM-RFE (support vector machine-recursive feature elimination) algorithms based on meta-analysis of 24 samples of microarray data from four different types of stress conditions namely cold, heat, salt, and hydrogen peroxide stresses. We modified the SVM-RFE by using bootstrapping and leave-one-out cross-validation to overcome small sample size. We analyzed 15 key genes (TTHA0902, TTHA0260, TTHA0230, TTHB150, TTHA1660, TTHB160, TTHA1285, TTHA1233, TTHA0350, TTHA1152, TTHY7080, TTHC013, TTHA1376, TTHA0718 and TTHC010) as predicted by SVM-RFE, some of which were involved in DNA repair mechanisms and in response to abiotic stresses. Our study indicates that SVM-RFE could be utilized as a suitable machine learning method to predict key responsive genes involved in abiotic stresses.

Keywords: Thermus thermophilus, machine leaning, meta-analysis, Gene expression, SVM-RFE

ICB11-1401-12-1096

Identification key responsive genes to some environmental stresses in Thermus thermophilus by using computational biology approaches

Abbas Karimi-fard[1], Abbas Saidi[2*], Masoud Tohidfar[3]

[1]Shahid beheshti university, tehran. [2*]Shahid beheshti university, tehran.[3]Shahid beheshti university, tehran.

*Corresponding author(s). E-mail(s): abbas.saidi@gmail.com;

Abstract

Bacteria like other organisms are continually exposed to many types of stresses such as redox, temperature, nutrient availability, pH, heavy metal, hydrogen peroxide, antibiotic, and salt stresses. Adaptation to these environmental changes is vital for bacterial survival and defense against various stresses, where they are able to repair damaged cellular components through their advanced regulatory systems. One of the most important methods of molecular genetics to characterize the defensive mechanisms of the bacteria in response to environmental stresses is the use of transcriptome analysis and next-generation sequencing (NGS) technologies. By using computational biology approaches, we can identify hub genes and important pathways associated with these conditions. To identify key responsive genes in thermus thermophlius, we performed meta-analysis on 24 samples of microarray datasets from four different types of environmental stresses namely cold, heat, salt, and hydrogen peroxide stresses. In this study, we identified 262 DEGs, where123 of them were upregulated and 139 downregulated. The TTHA0979 (ABC transporter substrate-binding protein) and TTHA0980 (Fructose-1,6-bisphosphate aldolase/phosphatase) had the highest expression values (LogFC= 1.65 and 1.22, respectively) in the up-regulated genes. Moreover, protein-protein interaction (PPI) analysis of DEGs identified the followings: TTHA0965, TTHA0966, TTHA0963, TTHA0964, TTHA0960, TTHA0967, TTHA0961, TTHA0861, TTHA1138, and TTHA0902 as hub genes.

Keywords: Thermus thermophilus, meta-analysis, gene expression, protein-protein interaction analysis

ICB11-1401-12-1097

An analysis of bioinformatics on the antiviral activity and effectiveness of vitamins in treating monkeypox

Donya Poy[1], Masoud Tohidfar[2*]

[1]Shahid beheshti. [2*]Shahid beheshti.

*Corresponding author(s). E-mail(s): m_tohidfar@sbu.ac.ir;

Abstract

Today, bioterrorism poses a serious threat to the world. Since ancient times, contaminated and infectious materials, corpses, and animal carcasses have been used to commit bioterrorism. We must therefore prepare for their presence. The eradication of smallpox through vaccination led to an end to vaccination worldwide, which caused the next and enlarged generation of orthopoxviruses, including Mpox, to become a major concern (Henderson,2011). In spite of very preliminary research, we are facing a widespread outbreak of the MPox virus around the world today. Several trace elements and vitamins have been shown to significantly enhance immunity to support the immune system's normal functioning . There is a wide range of minerals and vitamins involved in biosynthesis and energy generation. In the literature, vitamin treatments such as K, D, B, and B17 may be beneficial in treating some viral diseases, such as Coronavirus or Hepatitis C infection , asthma, bronchitis, diabetes, and vaginal infections . Due to the similarity of conserved regions between smallpox and MPox, the study was designed to investigate the inhibitory mechanism of drugs approved by the FDA to treat smallpox (Tecovirimat, Brincidofovir, and Cidofovir) . Additionally, to investigate the effect of vitamins on the MPox virus' cell surface protein on preventing virus entry into the cell and causing infection. So far, no studies have been reported on vitamins in the inhibition of the MPox virus. In this study, by examining various vitamins against MPox, vitamins K and B17 showed a good inhibitory effect on the MPox virus. Interestingly, in docking and molecular dynamic studies, also ADMET prediction Analysis, we observed two patterns of behavior based on approved drugs. Vitamin B17 with Brincidofovir and vitamin K with few differences with Tecovirimat showed similar patterns. These bioinformatic and in silico studies can open new doors for experts to conduct more experimental experiments.

Keywords: Vitamin K, Vitamin B17, Monkeypox, Docking, Molecular dynamic

ICB11-1401-12-1098

Power Domination Set-based Algorithm for Identification of Key Genes in Lung Adenocarcinoma

Samad Khodaei[1], Soheila Shabani-mashcool[2*], Reza Hassanzadeh[3]

[1]University of mohaghegh ardabili. [2*]University of tehran.[3]University of mohaghegh ardabili.

*Corresponding author(s). E-mail(s): soheila.shabani@ut.ac.ir;

Abstract

Lung cancer is responsible for 18% of cancer-related deaths. The most common lung cancer is Lung adenocarcinoma, whose patients are increasing especially among non-smokers, women and younger people . Despite all these facts, the factors of its development are not yet known precisely. The present study was conducted with the aim of identifying key genes in the development or prognosis of lung adenocarcinoma. So far, the concept of power dominating set (PDS) has not been used to identify or predict important genes in any type of biological network, and this study is the first research in this case. It is necessary to mention that a set of vertices that eventually monitors, using special rules, all the vertices of a graph is called a PDS . Here, we considered six microarray datasets related to lung adenocarcinoma from the Gene Expression Omnibus database , each set containing several control and tumor samples. Then, we extracted 143 differentially expressed genes using the GEO2R and considering p-value$<$0.05 and $|$ logfc$|$ $>$1. To validate these genes, we used 574 datasets of The Cancer Genome Atlas (https://www.cancer.gov/tcga), which included 515 tumor and 59 normal samples. After analyzing these samples, we reached 3744 differentially expressed genes. We calculated the intersection of these genes and 143 genes obtained from GEO and finally, 133 genes were obtained. Then, we obtained the network corresponding to these 133 genes using STRING and tried to find PDS in this network. Our algorithm identified four pairs of genes {SPARCL1, CDH5}, {CLEC14A1, CAV1}, {CDH5, VWF} and {RHOJ, VWF} as four PDSs. Upon analysis of their functions within various biological processes and pathways, we have reached the conclusion that these genes apparently play a crucial role and may be highly correlated with lung adenocarcinoma.

Keywords: Power dominating set, Lung adenocarcinoma, Gene expression, Graph theory

ICB11-1401-12-1099

Integrated Bioinformatics Analysis For Differentially Expressed Genes, Micrornas And Pathways Identification In Age-Related Macular Degeneration

Amir reza Safavinejad[1*], Ali reza Morovat[2], Maryam Talaei[3], Ali reza Fazilat[4], Ali Ghanbari asad[5]

[1*]Fasa university of medical sciences (student research committee).[2]Fasa university of medical sciences (student research committee). [3]Fasa university of medical sciences (student research committee). [4]Fasa university of medical sciences (student research committee). [5]Fasa univeersity of medical sciences (noncommunicable diseases research center).

*Corresponding author(s). E-mail(s): ar.Safavinejad@gmail.com;

Abstract

Background:
Age Related Macular Degeneration is leading cause of irreversible visual loss in developed countries. Mechanisms of Age Related Macular Degeneration pathology include Aging of Retinal Pigment Epithelium cells, Oxidative Stress, increased levels of High-density lipoprotein cholesterol, Complement system activation, Neovascularization and other mechanisms.
Purpose:
In this study we aimed to identify genes and microRNAs important in pathogenesis of Age Related Macular Degeneration.
Methods:
Gene expression profile of GSE1719 was available from Gene Expression Omnibus database. 18 samples of Dermal fibroblasts of healthy people and 18 samples from patients were selected from database. Dataset was analyzed with GEO2R tools to calculate Differentially Expressed Genes. A defined formula [-log(p.value)×| log(fold change)|] was applied to select top significant probs. After filtering and removing duplicates, Protein-Protein Interaction network of Differentially Expressed Genes established through STRING database. Gephi software was used to select hub genes. Toppgene database was used for Gene Ontology and microRNA target prediction. Cytoscape software was used to identify hub microRNAs.
Results:
Total 1756 Differentially Expressed Genes were identified of which 836 were upregulated and 920 were downregulated. 5 hub genes (TP53, ACTB, HSP90AA1, HRAS, VEGFA), 5 hub genes based on closeness centrality (ACTB, TP53, HSP90AA1, HRAS, HSPA4) and 5 microRNA hubs (mir-7, mir-495, mir-5688, mir-302b, mir-25, mir-222) based on degree were identified. Significant Gene Ontologies were extracellular matrix organization, focal adhesion and axon guidance.
Conclusion:
Our study investigated genes and microRNAs effective in pathogenesis of Age Related Macular Degeneration that can help discovering new biomarkers for diagnosis and new treatment methods.

Keywords: Age Related Macular Degeneration, Biomarkers, Transcriptomics, microRNA, Hub Genes

ICB11-1401-12-1100

In Silico Analysis of Novel VHL Germline Mutations in Iranian RCH Patients

Massod Naseripour[1], Kowsar Bagherzadeh[2], Golnaz Khakpor[3], Ahad Sedaghat[4], Reza Mirshahi[5], Hengameh Kasraee[6], Fatemeh Azimi[7*]

[1]Eye research center, the five senses institute, rassoul akram hospital, iran university of medical sciences, tehran, iran. [2]Eye research center, five senses health institute, iran university of medical sciences, tehran, iran. [3]Eye research center, five senses health institute, iran university of medical sciences, tehran, iran. [4]Eye research center, five senses health institute, iran university of medical sciences, tehran, iran. [5]Eye research center, five senses health institute, iran university of medical sciences, tehran, iran. [7*]Eye research center, the five senses institute, rassoul akram hospital, iran university of medical sciences, tehran, iran.

*Corresponding author(s). E-mail(s): Fatemeazimi80@yahoo.com;

Abstract

Von Hippel-Lindau (VHL) syndrome is an autosomal dominant inherited multisystem neoplasia disorder caused by VHL tumor suppressor gene, coding for VHL protein (pVHL), variants. Various types of VHL variants present different clinical phenotypes that later lead to events resulting in benign or malignant lesions including Retinal Capillary Hemangioblastoma (RCH). Here in, 3 novel mutation sites observed in 3 families (5 RCH patients), including c.511A>C, c.514C>T, and c.511A>T in exon 3 of the VHL gene are reported for the first time. According to ACMG classifications, c.514C>T and c.511A>T variations are likely pathogenic, and c.511A>C is a variant of uncertain significance (VUS) in accordance with autosomal dominant inheritance. The location and impact of the incidence mutations on pVHL were computed using in silico analysis. The obtained structural information and computational analysis showed that the studied mutations induce conformational changes that limit the flexibility of pVHL interaction interface with elonginB/C, elongin C/B, and cullin2, which is necessary for hypoxia-inducible factor 1-alpha binding. The recently added gene variants and their related clinical phenotypes will improve the VHL diagnosis accuracy and the patients' population carrying VHL gene mutations. These pioneering results can be used as a model for future functional studies.

Keywords: Von Hippel-Lindau (VHL), Retinal Capillary Hemangioblastoma (RCH), Molecular Dynamic Simulations, Novel Mutations

ICB11-1401-12-1101

Computational Designing the Ligands of Protein L Affinity Chromatography Based on Molecular Dynamics Simulations

Saman Rahmati[1], Kowsar Bagherzadeh[2], Masoud Amanlou[3], Behrouz Vaziri[4*]

[1]Pasteure institute of iran. [2]Eye research center, five senses health institute, iran university of medical sciences, tehran, iran. [3]Tehran university. [4*]Department of biochemistry, pasteur institute of iran, tehran, iran.

*Corresponding author(s). E-mail(s): Behrouz-vaziri@pasteur.ac.ir;

Abstract

Protein L is a multidomain protein characterized from Peptostreptococcus magnus which has the binding affinity to kappa light chain of immunoglobulin (Ig) without stimulating an immune response and is useful for purification of antibody fragments in affinity chromatography system. The advance in protein engineering and computational biology approaches lead to development of engineered affinity ligands with improved properties like high binding capacity. In affinity chromatography, increasing the binding affinity and multimerization of the ligand domains lead to high binding capacity. In this study, docking methodology, molecular dynamic simulations and Osprey software were used to design single B domains of Protein L with higher affinity to antibody fragments. Then, the modified B domains which had high binding affinity were polymerized to ligands with six and eight B domains by homology modeling methods. The results showed that single B domain mutants of Thr865Trp and Thr847Met-Thr865Trp had higher binding affinity to Fab compared to wild single B domain. Also, the polymerized Proteins L to six and eight B domains were stable.

Keywords: Protein L, Antibody fragments, Molecular docking studies, Molecular dynamic simulations, OSPREY, Affinity chromatography

ICB11-1401-12-1102

Computational engineering of Protein L ligands to achieve an optimal affinity resin for purification of antibody fragments

Saman Rahmati[1], Kowsar Bagherzadeh[2], Fatemeh Torkashvand[3], Massoud Amanlou[4], Pezhman Fard esfahani[5], Hooman Aghamirza moghim aliabadi[6], Behrouz Vaziri[7*]

[1]Department of biochemistry, pasteur institute of iran, tehran, iran. [2]Eye research center, five senses health institute, iran university of medical sciences, tehran, iran. [3]Department of biochemistry, pasteur institute of iran, tehran, iran. [4]Department of medicinal chemistry, faculty of pharmacy, tehran university of medical sciences, tehran, iran. [5]Department of biochemistry, pasteur institute of iran, tehran, iran. [6]Pasteure institute of iran. [7*]Department of biochemistry, pasteur institute of iran, tehran, iran.

*Corresponding author(s). E-mail(s): Behrouz-vaziri@pasteur.ac.ir;

Abstract

Protein L affinity chromatography can be a well-established platform for the purification of antibody fragments containing kappa light chain in biopharmaceutical downstream processing. Affinity chromatography resins usually suffer from low binding capacity and this problem increases the cost of the final product at the industrial production. Recent studies have shown that an increase in the binding affinity of resin ligands towards their target proteins leads to increased product purity, recovery and dynamic binding capacity values in affinity chromatography. The advance in genetic engineering and computational biology approaches leads to development of engineered affinity ligands with improved properties. Computational biology methods such as molecular modeling, molecular docking, molecular dynamic simulations and protein redesign software were applied to design mutated Protein L ligands. The engineered ligands were experimentally studied after being coupled to a solid matrix. The influence of the engineered ligands on the performance of affinity purification with loading prepared Fab were investigated and the dynamic binding capacity, product purity and recovery for the engineered resins were evaluated and compared to Capto™ L resin (GE Healthcare Bio-Sciences AB, Sweden). The highest dynamic binding capacity at 10% breakthrough (DBC10%) was observed for Model-1 resin; 19.5 mg/ml, as compared to the DBC10% of Model-2, Model-3 and Capto™ L resins; 7.2 mg/ml, 7.8 mg/mL and 16.5 mg/mL respectively. Also, recovered Fab fragments from E.coli lysate by Model-1 and Model-3 resins was evaluated 95% compared to 92% and 39% by Capto™ L and Model-2 resins as well as the results of SEC-HPLC analysis showed the purities of recovered Fab fragments by all of resins were over 95%. The results indicated that the Protein L Model-1 resin was an optimal resin for purification antibody fragments containing kappa light chain from E.coli host cell proteins in affinity chromatography.

Keywords: Protein L, Antibody fragments, Molecular docking studies, Molecular dynamic simulations, OSPREY, Affinity chromatography

ICB11-1401-12-1103

Prediction of interaction network of altered proteins, kinases and transcription factors, miRNAs, lncRNA in breast cancer disease

Mohadese Safabakhsh[1], Sohrab Boozarpour[2*], Kamran Ghaedi[3], Mina Lashkarboloki[4], Shaaban Ghalandarayeshi[5]

[1]Gonbad kavous university. [2*]Gonbad kavous university.[3]University of isfahan. [4]Tarbiat modares university. [5]Gonbad kavous university.

*Corresponding author(s). E-mail(s): so.boozarpour@gmail.com;

Abstract

The most common cancer type, is female breast cancer with more than 2.3 million cases diagnosed and 685,000 deaths in 2020. Eukaryotic gene expression is regulated at the epigenetic, transcriptional, and post-transcriptional levels by coordinated multi-layer mechanisms . Create multi-layered networks consisting of different biological levels is one of the appropriate methods to increase the depth of our knowledge regarding this type of diseases. In the present study, in order to identify the regulatory layers including the transcription factors and the transcription factors regulated kinase, the ChEA2016, and KEA2015 header of the EnrichR database was used respectively as well as IncHub and TargetScan headers were used to detected lncRNAs and miRNAs regulator of all three gene layers. Based on the three topological parameters such as closeness, degree and between 10 driving elements of each layer were selected. Analysis related to message transmission pathways and gene ontology was performed using MetaScape and Cytoscape. To determine the DEPs biological functions, gene ontology enrichment analysis was conducted at the following levels: biological process (BP), molecular function (MF) and cellular component (CC). The message transmission pathways and gene ontology related to all 3 layers of kinases, transcription factors and altered proteins were identified in breast cancer. Multilayer network analysis, drawing, and central element identification revealed that hsa-miR-16-5p, hsa-miR-92a-3p, hsa-miR-615-3p, and lncRNAs IGBP1-AS2, TMED2-DT were involved in breast cancer progression. In the findings of biological processes in gene ontology and molecular functions, changes in similar pathways such as mRNA binding, Cadherin binding, and other regulatory pathways such as inhibitory enzyme activity and oxidoreductase activity were shown. The pathways identified in the sub- section also detected changes in the regulation of the cellular matrix and cellular junctions.

Keywords: proteomics, breast cancer, lncRNAs, Bioinformatics

ICB11-1401-12-1104

Examining protein expression patterns in breast cancer compared to healthy group by bioinformatics

Mohadese Safabakhsh[1], Sohrab Boozarpour[2*], Kamran Ghaedi[3], Mina Lashkarboloki[4], Shaaban Ghalandarayeshi[5]

[1]Gonbad kavous university. [2*]Gonbad kavous university.[3]University of isfahan. [4]Tarbiat modares university. [5]Gonbad kavous university.

*Corresponding author(s). E-mail(s): so.boozarpour@gmail.com;

Abstract

Breast cancer is a multifactorial disease and various factors contribute to its occurrence. There is a significant difference between the occurrence of this disease and mortality rates . Despite improvements in the genetic classification, protein level is frequently used in clinical diagnostics and treatment decisions . In the molecular sciences, proteomics has become an important field. Proteomics revealed critical details in the characterization of molecular mechanisms, development, and metastasis cancer, identifying specific treatment targets as well as useful biomarkers . The aim of this research is to determine the key proteins and central molecules involved in the pathogenesis of breast cancer using protein data analysis. MaxQuant software was used to convert the raw data (ID.number PXD012431) into the protein table according to the target-decoy identification algorithm. The results were analyzed with Perseus software and their quality verified by component analysis so that infected and identified-reverse sequence proteins removed. T-test and Permutation-based FDR method were used to correct the P-values and identified expressed proteins in study groups. Cytoscape v and CluePedia V software created a comprehensive view of interactions between proteins and other obtained layers. The confidence threshold for network edges was 0.6 and considered for Inhibition, Binding, Activation and PTM.The quality control shows appropriate separation of healthy and disease samples in study groups. The results indicated presence of 1149 altered proteins in patient sample, including 191 low expression and 958 high expression in the comparison with healthy group with confidence threshold of FDR $<$ 0.05. Also, identified-central proteins in the network were included NID1 , LAMC1, COL4A1, COL1A2, COL4A2, TNS1. It is expected that these proteins will be suitable as a marker for early diagnosis or for therapeutic purposes.

Keywords: proteomics, breast cancer, protein expression, bioinformatics

ICB11-1401-12-1105

Selection of antibacterial peptides against methicillin-resistant Staphylococcus aureus by bioinformatics and laboratory integration methods

Armity Kiumarsy[1], Ali beman Jebali[2], Sarvenaz Falsafi[3*]

[1]Islamic azad university of tehran medical sciences. [2]Islamic azad university of tehran medical sciences. [3*]Islamic azad university of tehran medical sciences.

*Corresponding author(s). E-mail(s): annarhonin@gmail.com;

Abstract

Common hospital infections caused by methicillin-resistant Staphylococcus aureus (MRSA) bacteria are one of the most common causes of death in hospitalized people. MRSA is resistant to antibiotics and common treatments. However, so far, two drugs designed based on Anti-Microbial Peptides (AMP) have been effective in controlling and eliminating them. Antimicrobial peptides are immune molecules with a small and simple structure and effective function, which are often found in insects and plants, although they have also been observed and recorded in the body tissues of living organisms, for example, In mammals, AMP is found mainly in the granules of neutrophils and in the secretions of epithelial cells that cover the skin and mucosal surfaces. AMPs can be effective on destroying bacteria, viruses, fungi, parasite or even tumors. But using AMP as drug can have potential side effects. It's a new and unknown field of research. The aim of this research was to find a safe and effective AMP against MRSA. In this research, all peptides effective on MRSA were found from the AMP database and their characteristics were noted. Then their antigenicity, allergenicity and toxicity were investigated. In the second step, homology with human proteins and peptides was checked and peptides that were non-allergenic, non-antigenic, non-toxic and had high homology with human proteins were selected and their interactions with seven bacterial surface proteins and two bacterial internal proteins were investigated. Became. Among them, two peptides were selected as the main candidates for synthesis and future tests, one of which was selected with a shorter amino acid chain length.

Keywords: MRSA, AMP, Common hospital infections, Anti-Microbial Peptides

ICB11-1401-12-1106

venom gland Transcriptome analysis of Iranian yellow scorpion, "Odontobuthus doriae" derived some putative antimicrobial peptide with anti-SARS-CoV-2 effect

Maryam Naderi soorki[1*]

[1*]Shahid chamran university of ahvaz.

*Corresponding author(s). E-mail(s): maryam_naderi_soorki@yahoo.com;

Abstract

SARS-CoV-2 is from enveloped virus family responsible for the COVID-19 pandemic. No efficient drugs currently available for treatment of infection caused specifically by this virus. Therefore, searching for effective therapeutic treatments for severe illness caused SARS-CoV-2 is crucial . Scorpion venoms are significant sources of peptides with pharmaceutical potential including antivirals . Although, some studies determined the antiviral effects of some scorpion peptides on the other members of Coronaviridae family , but no anti-SARS-CoV-2 effects of these peptides have been reported until now. In this study antiviral effects of two predicted antimicrobial peptides (ODAMP4, ODAMP5) from Iranian yellow scorpion "Odontobuthus doriae" were assessed by Computational methods. Two predicted peptides with potential of antiviral activities were selected from the cDNA library that have been constructed by our research team from Iranian scorpion "Odontobuthus doriae". 3D model of peptides was designed with I-TASSER. The models were refined by a 200 ns Molecular Dynamics (MD) simulation by using Gromacs 2021.2 software. Refined models were Docked with RBD domain of SARS-CoV-2 spike protein by using HADDOCK software. Docking of human ACE2 peptide with RBD domain also assessed at the same time. The docked complexes (RBD-peptide and RBD-ACE2) were refined again by a 100 ns MD simulation and then analyzed. The results from molecular docking based on HADDOCK Score and Z-score showed that ODAMP5 peptide has a high affinity for RBD domain compared to another peptides. the results of molecular dynamics simulation which were done after docking for 100ns showed ODAMP5 has a high stability and affinity to the RBD domain of covid-19 spike protein to ODAMP4 and human ACE2. In fact, this peptide can be a good candidate for used as a factor to inhibit the RBD domain of SARS-COV2 virus in in clinical studies with pharmacological purposes.

Keywords: COVID-19, Antimicrobial peptide, Iranian yellow scorpion, Molecular Dynamics simulation, Molecular Docking

ICB11-1401-12-1107

Genomic analysis of the ASMT gene family in Cicer arietinum

Asra Karimian[1*], Bahman Bahramnejad[2], Hemn Salehi[3]

[1*]University of kurdistan.[2]University of kurdistan. [3]University of kurdistan.

*Corresponding author(s). E-mail(s): asra.karimian1997@gmail.com;

Abstract

Acetylserotonin methyltransferase (ASMT) is the last enzyme of the melatonin biosynthesis pathway and may play a rate-limiting role in melatonin production in plants. In this study, bioinformatic analysis of ASMT gene family in chickpea (Cicer arietinum) was done. At first, by using the nucleotide sequences related to the ASMT gene, which has been previously studied in alfalfa and soybean, we obtain the related protein sequence, and then by performing BLAST against the chickpea genome, 28 predicted protein sequences related to this gene family were detected. Only one sequence was identified as ASMT. All investigated sequences had O-methyltransferase protein domain distributed on all pea plant chromosomes except number 1 and 8, and 6 sequences were not located on any chromosome, also the sequence predicted as ASMT was located on chromosome number 5. The number of intron regions in the analyzed sequences was between 1 and 3. Phylogeny studies indicated that the examined sequences were placed in 6 different groups, each group corresponding to one of the classes of the O-methyltransferase family. The results of examining the gene structure and the composition of motifs showed high conservation in each group of the O-methyltransferase family in the arrangement and distribution of motifs. Examination of the regulatory region related to the upstream of the ASMT gene showed that this region has a light response regulatory element, cis-acting regulatory element involved in zein metabolism regulation, gibberellin-responsive element and cis-acting regulatory element related to meristem specific activation. Analysis of gene expression based on RNA-seq and qRT-PCR at the genome level showed that these genes are involved in the response to abiotic stresses.

Keywords: ASMT Gene, Bioinformatic, Melatonin, chickpea

ICB11-1401-12-1108

Investigating the pathway of resistance to Ciprofloxacin in Campylobacter jejuni by using microarray analysis

Mahlagha Cheraghi[1], Nastaran Asghari moghaddam[2*], Bita Bakhshi[3]

[1]Central tehran branch, islamic azad university. [2*]Central tehran branch, islamic azad university.[3]Department of medical bacteriology, faculty of medical sciences, tarbiat modares university, tehran, iran.

*Corresponding author(s). E-mail(s): nas.asgharimoghaddam@iauctb.ac.ir;

Abstract

C. jejuni is one of the main and effective Gram-negative bacteria in food poisoning and causing gastrointestinal infections in the intestine. This bacterium is a hidden microbe in food, so-called foodborne, which is able to cause infection by penetrating the small intestine, and for this reason, it has become one of the main sources of digestive and internal infections. The main challenge in dealing with C. jejuni infections is the resistance of strains of this species to various antibiotics (Iovine,2013). Ciprofloxacin resistance is common among animal and human C. jejuni isolates . Investigating gene expression profiles in antibiotic resistant strains of this disease can clarify the process of resistance and pathogenicity of resistant strains and play a role in designing new drugs, prescribing antibiotics and finally eradicating C. jejuni infections be effective. Data were loaded into the R software environment after being downloaded from the NCBI GEO database with the accession ID GSE41822. In order to normalize the data, we used the DESeq2 library. Then, with the help of examining differential gene expression by means of DESeq2 library and using statistical tools such as effect function, dispersion and curve fitting, genes with differential expression between different groups were identified, and the significant difference in the expression of each among the groups was determined a p-value was given. Hierarchical clustering (hClust) was used to investigate the difference in gene expression. We found that in general, M23/M37 protease and cytochrome genes and genes involved in the transfer of substances such as transferases can be considered the most common genes in Ciprofloxacin antibiotic and these genes will be possible candidates for inhibitor drugs.

Keywords: Campylobacter jejuni, antibiotic resistance, Ciprofloxacin, Microarray analysis

ICB11-1401-12-1109

Identification Of Hub Genes, microRNAs And Key Pathways Associated With Bladder Cancer

Alireza Morovat[1*], Kiarash Zare[2], Mohammadmehdi Naghizadeh[3], Alireza Monsef[4]

[1*]Fasa university of medical sciences.[2]Fasa university of medical sciences. [3]Noncommunicable diseases research center, fasa university of medical sciences, fasa, iran. [4]Noncommunicable diseases research center, fasa university of medical sciences, fasa, iran.

*Corresponding author(s). E-mail(s): almorovvat@gmail.com;

Abstract

Background:
Bladder cancer is carcinoma of epithelial cells of urinary bladder. It is classified into muscle-invasive bladder cancer and non-muscle-invasive bladder cancer based on alteration of genes like Hras, deletion of chromosome 9 and a point mutation in FGFR3.
Purpose:
In this study we aimed to identify genes and microRNAs important in pathogenesis of Bladder cancer.
Methods:
Gene expression and microRNA profile of GSE40355 was available from Gene Expression Omnibus database. 8 samples of bladder tissue from healthy people, 8 samples from patients with Low grade and 8 samples from patients with High grade Bladder Cancer were selected from dataset. Dataset was analyzed with GEO2R tools to calculate differentially expressed genes and differentially expressed microRNAs. P-values of < 0.05 and a $\mid \log(\text{fold change}) \mid > 2.5$ in differentially expressed genes and a $\mid \log(\text{fold change}) \mid > 2.0$ in differentially expressed microRNAs were applied to identify significant differentially expressed genes and differentially expressed microRNAs. We established a protein-protein interaction network of differentially expressed genes through STRING database and used Gephi software to select hub genes. We used Toppgene database for gene ontology and microRNA target prediction. Cytoscape software was used to identify hub microRNA.
Results:
Total 1398 differentially expressed genes were identified, 1244 nodes and 9014 edges were established based on Protein – Protein Interaction network. 5 hub genes based on degree (IGF1, VEGFA, CD44, FOS, ESR1) and 5 mirna hubs (mir-1, mir-29c, mir-29a, mir-29b, mir-576) were identified. Significant gene ontologies from Pathway included Ensemble of genes encoding core extracellular matrix including extracellular matrix glycoproteins, collagens and proteoglycans and extracellular matrix organization.
Conclusion:
In this study, we investigated genes and microRNA effective in pathogenesis of bladder cancer, which can help identifying biomarkers for early diagnosis and new treatment methods.

Keywords: Bladder Cancer, Biomarkers, Transcriptomics, microRNA, Hub Genes

ICB11-1401-12-1110

Molecular dynamic simulation of Buforin IIb anticancer peptide with POPC-Cholesterol model membrane

Fateme Bastami[1*], Somayyeh Dabbagh[2], Mahboobeh Zarrabi[3]

[1*]Alzahra unversity.[2]Alzahra unversity. [3]Alzahra unversity.

*Corresponding author(s). E-mail(s): F.bastami@alzahra.ac.ir;

Abstract

The multidrug resistance process has hindered the effect of chemotherapeutics. Demands for the design and synthesis of novel therapeutic peptides are increasing. Recent studies have shown the anticancer activity of Buforin IIb against 62 cancer cell lines . It is a 21-mer peptide (RAGLQFPVGRLLRRLLRRLLR) and has a remarkable selectivity toward different cancer cell lines. It has been indicated that the anticancer action of Buforin IIb involves the induction of cancer cell apoptosis. However, the mechanism of anticancer activity remains unclear . During this study, the interaction of Buforin IIb with the model bilayer (POPC-CHOL) was analyzed with All-atom molecular dynamic simulations. The model bilayer was constructed with Charmm-GUI membrane builder tool, and after 500ps equilibration at 310.15 ◦K, the motion of Buforin IIb on the upper leaflet surface area was investigated during 100ns of atomistic simulation by GROMACS 2018.1 software. RMSD analysis indicates the stable secondary structure of Buforin IIb at the vicinity of the bilayer. In such a case, the energy calculations examined the details of the interaction at the peptide-bilayer interface. The distance profile between the peptide center of mass and headgroups of the bilayer confirmed the adsorption of Buforin IIb on the membrane surface. These results also demonstrated the key effect of Proline residue at the adsorption state. Totally, this simulation unraveled the initial penetrating process of Buforin IIb through a rigid bilayer (POPC-CHOL), which could be informative for understanding the mechanistic features of peptide anticancer activity.

Keywords: Anticancer, Buforin IIb, Molecular dynamic simulations

ICB11-1401-12-1111

Examining the similarity of Hsp90 sequence amongst different Aspergillus species

Seyyed mazyar Sadati[1*], Mohammad ali Tajik ghanbary[2], Mohammad ali Barimani varandi[3], Mohsen Barari[4]

[1*]Sari agricultural sciences and natural resources university.[2]Sari agricultural sciences and natural resources university. [3]Sari agricultural sciences and natural resources university. [4]Sari agricultural sciences and natural resources university.

*Corresponding author(s). E-mail(s): sadati.mazyar@gmail.com;

Abstract

Aspergillus is an important genus of fungi. This taxonomic group encompasses organisms whose characteristics are of high pathological and agricultural importance as well as industrial and scientific value. An accurate method of identification and classification of this genus is based on genetic approach, in conjunction with morphological methods. Heat shock proteins are a main group of active proteins in the cells of living organisms. This research, about possible similarities and differences within the coding genes of these proteins, can reveal more details about the complexities in Aspergillus. In this survey, the similarity of Hsp90 sequences among 12 species of Aspergillus fungi from the 5.8s region was studied. For this purpose, Hsp90 coding genes belonging to different species of Aspergillus were collected and extracted from the NCBI gene information bank. These sequences were searched, and extracted in FASTA file format. Afterwards, they were entered into the MEGA 10 software through the ALIGN - Edit / Build Alignment portal to assess the similarities. For more accuracy in this research, all chosen sequences entered into the Mega software had same length and size. MUSCLE algorithm was used for alignment and a phylogenic tree was drawn using Construct/Test Maximum Likelihood method. For examining the authenticity of the tree, Penicillium rubens sequence was used as outgroup with bootstrap method 1000. By reviewing the drawn phylogenic tree, significant similarities and differences among different species of Aspergillus were observed. The results revealed that four species, A. tubingensis, A. japonicas, A. terreus and A. ruber, formed a separate branch from the trunk of the main branch. It can be discussed that Hsp90 is a potential candidate as a molecular marker for further phylogenetic studies.

Keywords: Bioinformatics, Phylogeny, Ribosomal 5.8s

ICB11-1401-12-1112

A brief bioinformatic survey on Colletotrichum species complex

Mohsen Barari[1*], Mohammad ali Tajik ghanbary[2], Mohammad ali Barimani varandi[3], Seyyed mazyar Sadati[4]

[1*]Sari agricultural sciences and natural resources university.[2]Sari agricultural sciences and natural resources university. [3]Sari agricultural sciences and natural resources university. [4]Sari agricultural sciences and natural resources university.

*Corresponding author(s). E-mail(s): mbararisc99@gmail.com;

Abstract

Colletotrichum, is the main cause of anthracnose diseases on different plants. An accurate taxonomic identification of Colletotrichum species is an integral part of disease management . Morphologically-based identification of Colletotrichum species has always been difficult due to lack of reliable morphological characters, and wide host range, making species identification confusing. Later studies show that morphological studies should be used in cooperation with other characters to determine species relationships within Colletotrichum especially molecular markers. Development in molecular phylogenic methods made it possible to recognize stable clades within Colletotrichum. In this survey, 33 valid ITS sequence belonging to different Colletotrichum species were extracted from NCBI and aligned by MEGA 11, then subjected to phylogenetic tree analysis by default parameters. The results showed ITS region can be used as a barcode for identification and classification of Colletotrichum species. The phylogenetic tree resulted by this alignment, was accurate in classifying related species in their respective species complex. By reviewing the phylogenetic tree, it can be discussed that some species complexes, such as C. graminicola and C. spaethianum complexes, can be characterized as one complex. Although further research is needed to complete hard identifying species identification by complementary markers like Actin (ACT), Calmodulin (CAL) or β-tubulin (TUB2) and Glutamine synthetase (GS) . This research shows that the ITS marker is a useful and reliable tool to give a preparatory identification for Colletotrichum species and assort them in related species complexes and species identification.

Keywords: Phylogeny, Ribosomal 5.8s, ITS

ICB11-1401-12-1113

Are Repetitive Motifs of LEAp Chaperones the Main Bodyguards of Target Proteins?: in silico Prediction

Mostafa Nikbakht ardalani[1], S.shirin Shahangian[2*], Mahmoudreza Aghamaali[3]

[1]Department of biology, faculty of sciences, university of guilan, rasht, iran. [2*]Department of biology, faculty of sciences, university of guilan, rasht, iran.[3]Department of biology, faculty of sciences, university of guilan, rasht, iran.

*Corresponding author(s). E-mail(s):

Abstract

Introduction:
Severe water stress, either through freezing or drying, leads to protein unfolding and subsequent cell damage. LEA proteins are involved in providing protection to biological macromolecules during water loss. Among animals, LEA group 1 proteins have been reported only in Artemia as a unique stress model in animals. AfLEA1 interacts with other proteins and nucleic acids . In addition, it protects the enzyme lactate dehydrogenase (LDH) during drying [1,2]. Experiments: I-Tasser (https:zhanggroup.org/I-TASSER) predicts the structure of proteins by comparing parts of known crystal structures available from the protein data bank to the investigated polypeptide and combining them into models using a hierarchical prediction. Based on TM-score, RMSD and C-score parameters, the best model was selected for the docking process and HADDOCK web server (https://wenmr.science.uu.nl/haddock2.4/) was used to investigate protein-protein interaction. Active and passive residues were predicted using the CPORT server (https:alcazar.science.uu.nl/services/CPORT).
Results:
The model provided by I-TASSER structurally includes 30% random coil and 70% α-helix. 63% of its residues are on the surface and exposed to the solvent, and 37% of the residues are buried in the structure. The target protein has a more flexible structure than other regions in regions with a random coil structure and at the N and C-terminus of the amino acid sequence, which interacts with the lactate dehydrogenase protein substrate through 14 hydrogen bonds, 4 salt bridges and 188 non-bonded interactions.
Conclusion:
LEAPs are highly specialized in confronting water-deficiency conditions. In spite of considerable research, the mechanism of action of LEAs to confer tolerance against abiotic stresses still remains obscure. Based on our findings, their highly repetitive structure are involved in binding to and protecting the target protein during desiccation to maintain its proper folding. The repetitive motifs in LEA1 may further act as multivalent sites for protein–protein, and protein–RNA interactions.

Keywords: LEAp, Dessication tolerance, HADDOCK, I-TASSER, CPORT, Chaperone activity, repetitive motif

ICB11-1401-12-1114

Identification of sodium alendronate derivatives as purple acid phosphatase inhibitors and promising candidates for osteoporosis therapeutics: in silico studies

Mitra Ghobadi[1], S.shirin Shahangian[2], Behnam Rasti[3], Hadi Adibi[4]

[1]Department of biology, faculty of sciences, university of guilan, rasht, iran. [2]Guilan, department of biology, faculty of sciences, university of guilan, rasht, iran. [3]Department of microbiology, faculty of basic sciences, lahijan branch, islamic azad university. [4]Pharmaceutical sciences research center, health institute, kermanshah university of medical sciences, kermanshah, iran.

*Corresponding author(s). E-mail(s):

Abstract

Purple acid phosphatases (PAPs) belong to binuclear metallohydrolases which catalyze the hydrolysis of a variety of phosphorylated substrates at acidic to neutral pH. In humans, increasing PAP levels shows the progression of osteoporosis. Alendronate sodium is used to treat bone diseases. In this research, we designed and synthesized alendronate sodium derivatives for the first time, as specific inhibitors against the rkbPAP enzyme. Due to its ability to predict the binding site and binding affinity of ligands to receptors, molecular docking simulation is one of the most frequently used methods in Structure-Based Drug Design (SBDD). The three-dimensional structure of PAP (PDB ID: 2QFR) was used for docking studies. Water molecules were removed from all coordinate files prior to docking, then minimized by chimera 1.16 software. The structure of compounds was constructed by ChemDraw v.17.3 and optimized using MM2 calculation available in ChemBio3D 17.3. Blind molecular docking analysis was planned. The grid box space was set on the whole space around the protein's structures (PAP 66×46×54Å and grid-point spacing of 1Å). computational docking experiments were carried out by utilizing AutoDockVina v.1.1.2. The docking results were analyzed using PyMOL, Discovery studio client, and Ligplot+ v.2.2.7 software. The results indicated that all the docked compounds (1-9) interact with the enzyme by different binding affinity energies (in the range of 6 to 7kcal/-mol). According to the the enzyme-ligand complexes, C12H18NNaO9P2 showed the highest binding affinity energy (-7 kcal/mol) to interact with the enzyme, involved in hydrogen bond interactions with G156, R188, S189, and Q193 and hydrophobic interactions with T118, Y128, I117, Q157, Y192, E150, Y149, and K155. According to in silico studies C11H15N2NaO10P2 were capable of interacting with His202 and Asn201, as critical residues involved in enzyme activity. so these compounds could be a candidate for the development of chemotherapeutics to treat osteoporosis.

Keywords: Molecular docking, Purple acid phosphatases, alendronate sodium derivatives, Osteoporosis, Enzyme inhibitors

ICB11-1401-12-1115

Molecular Docking of DEP B enzyme with DON mycotoxin chemotype (15ADON): In-silico analysis

Zahra Mousavi[1], Ali Javadmanesh[2*]

[1]Ferdowsi university of mashhad. [2*]Ferdowsi university of mashhad.

*Corresponding author(s). E-mail(s):

Abstract

Deoxynivalenol (DON), which is secreted by fusarium molds, is one of the most common mycotoxin in cereal grains and thus poses a significant risk on human health and productivity of farm animals . Although, DON cannot be controlled by common methods such as mycotoxin-binding absorbents, numerous studies have reported that enzymatic transformation seems to be the most promising method for detoxification of DON and its chemotypes . Therefore, the aim of present study was to investigate the interaction between DEP-A (DON deactivators enzyme) with DON mycotoxin chemotypes such as 15-acetyl-deoxynivalenol (15ADON) through molecular docking. The three-dimensional structure of DEP-A was predicted by SWISS MODEL server. The accuracy of the predicted structure was estimated by ERRAT and Z-score. Then, the stability of this structure was evaluated in molecular dynamic conditions with GROMACS (5.1.2) software. After that, the root mean square deviation (RMSD) curve of enzymes was drawn with Grace software. H-DOCK online server was used to docking of DEP-A with 15ADON. The results showed that, ERRAT and Z-score of predicted DEP-A structure were 91% and -8.65, respectability, which is within acceptable range. The calculation of the root mean square deviation (RMSD) verifies that the system is in an equilibrated structure during 100 ns simulation. Moreover, the results of molecular docking demonstrated that DEP-A enzymes bind to active site of 15ADON with relatively strong binding energy (-157.90). Finally, it is suggested that using DEP-A enzyme is an effective detoxification method through changing the molecular structure of 15ADON, however, laboratory studies are needed in the future.

Keywords: DEP-A, Mycotoxin, 15ADON, GROMACS, Molecular Docking

ICB11-1401-12-1116

Molecular docking of rat myostatin with human follistatin: an in-silico analysis

Elnaz Karbaschian[1], Ali Javadmanesh[2*]

[1]Ferdowsi university of mashhad. [2*]Ferdowsi university of mashhad.

*Corresponding author(s). E-mail(s):

Abstract

Myostatin (MSTN) belongs to transforming growth factor-beta (TGF- β) family as an autocrine/paracrine hormone produced mainly by muscle cells that inhibits muscle mass development. Follistatin is known to antagonize the function of TGF-β ligands, such as MSTN and activin A. Therefore, the aim of this study was to predict the 3D structure of rat (Rattus norvegicus) MSTN and investigate the interaction between rat MSTN and human (Homo sapiens) follistatin with molecular docking method. This method could provide insights into increase muscle mass in animals. The rat MSTN structure was predicted with Swiss-model server, and evaluated with SAVES 6.0 online server. Then, the interactions of rat MSTN with human follistatin (retrieved from UniProt: P19883) were performed using H-DOCK online server based on a hybrid algorithm of template-based modeling and ab-initio free docking. The Verify3D assessment of rat MSTN three-dimensional indicates that this protein having appropriate compatibility with 1D and 3D protein structures. ERRAT is an online server that could show incorrect regions of protein structures according to errors leading to random distributions of atoms, which can be distinguished from correct distributions. The ERRAT score for the predicted rat MSTN was 80% which was in the acceptable range. The Ramachandran plot is the 2d plot of the **Φ-Ψ** torsion angles of the protein backbone that provided overall view of protein conformation. Ramachandran plot indicated that in the predicted rat MSTN, around 80% residues belonged to the most favored regions. The docking result showed that human follistatin can be connected tightly to the predicted rat MSTN with a docking score of -244.73. Furthermore, the amino acids involved in the hydrogen bond including, H324, H326 and Y322 with hydrogen bond distances of 2.7, 3.5 and 1.9, respectability. The results of this study indicated possible application of human follistatin to inhibit rat MSTN, although further molecular dynamics study in addition to in-vitro experiments are required.

Keywords: Follistatin, Rattus norvegicus, Myostatin, Muscle, Molecular Docking

ICB11-1401-12-1117

In silico Molecular Docking simulation of Antiviral Peptide Cathelicidin LL-37 with Acyclovir

Behnaz Lotfi[1*]

[1*]Alzahra university.

*Corresponding author(s). E-mail(s):

Abstract

Over the last few years, peptide-based therapeutics is explored and promisingly used due to peptide's natural origin, low toxicity, and fewer side effects. Cathelicidin LL-37 is an antimicrobial/antiviral peptide that has been found in mammals and demonstrates inhibitory activity toward several enveloped viruses. Herpes Simplex Virus (HSV) types 1 and 2 are generally double-stranded DNA viruses that can cause infections, including oral and genital blisters. HSV type 1 reactivation or neo-formation may reach the Central Nervous System and trigger an acute inflammatory response contributing to severe Herpetic Simplex Encephalitis and Alzheimer's Disease pathogenesis . HSV type 2 is the cause of most genital herpes and can be sexually transmitted. Nowadays, Acyclovir is prescribed as a first-line drug and its mechanism of action is to prevent the synthesis of viral DNA by inhibiting DNA-Polymerase . Because of the poor oral bioavailability of Acyclovir as well as viral genome mutation that results in viral resistance and drug ineffectiveness , Cathelicidin LL-37 is being investigated via computational methods in this research. Due to the lack of the experimental third structures, both the receptors, DNA-Polymerase, and envelope glycoprotein of HSV, are homology modeled by Phyre2. Energy minimization by Yasara server and refinement by ModRefiner contribute to optimized structures of 93.2% in envelope glycoprotein and 90.6% in DNA-Polymerase. Molecular docking of Cathelicidin LL-37 into HSV type 2 envelope glycoprotein by HDock server results in 14 more Hydrogen bonds and 35 more non-ligand bonds in comparison to the same action on HSV type 1 DNA-Polymerase into Acyclovir. Based on the binding energy score of envelope glycoprotein and its ligand, it can be an antiviral candidate against HSV type 2 and used to develop effective drugs.

Keywords: Molecular Docking, Homology Modeling, Antiviral Peptides, Herpes Simplex Virus

ICB11-1401-12-1118

Co-expression network analysis revealed ATP6AP1-DT, and TRIM52-AS1 as potential therapeutic lncRNAs in hepatocellular carcinoma

Amirhossein Naghsh nilchi[1], Navid Kazemi[2], Fariba Dehghanian[3*]

[1]University of isfahan. [2]University of tehran. [3*]University of isfahan.

*Corresponding author(s). E-mail(s):

Abstract

Background:
Liver hepatocellular carcinoma (LIHC) is the fourth most common cancer-related cause of death globally, with the sixth highest incidence of new occurrences. Fibroblast growth factor receptor 4 (FGFR4) is a receptor tyrosine kinase and its aberrant activation is one of the carcinogenic factors in Hepatocellular carcinoma. In this study, we used differential expression and co-expression analyses to discover new therapeutic targets in LIHC based on FGFR4 gene expression.
Methods:
The expression profile of LIHC primary tumor and normal tissue were downloaded from the Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) database, respectively. Differential expression analysis was performed using the DESeq2 package (1.36) in R (4.2.1). $|\log2FC| \geq 0.58$, and adjusted p-value< 0.05 cutoff were set to identify the significant differentially expressed genes. The weighted gene co-expression network analysis (WGCNA) package (1.71) was used to construct the co-expression modules. The module containing the FGFR4 gene was identified and significantly differentially expressed long non-coding RNAs (lncRNAs) co-expressed with FGFR4 were visualized with Cytoscape software (3.9.1). Finally, survival analysis was performed using the Survival Genie web platform.
Results:
A total of 11206 genes were significantly differentially expressed in LIHC samples compared to normal samples. 16 modules were identified, among which the brown module contains the FGFR4 gene. Three lncRNAs including ATP6AP1 Divergent Transcript (ATP6AP1-DT), TRIM52 Antisense RNA 1 (TRIM52-AS1), and ZSCAN16 Antisense RNA 1 (ZSCAN16-AS1) were significantly upregulated in LIHC and were co-expressed with FGFR4. The survival analysis indicates that the overexpression of FGFR4, ATP6AP1-DT, and TRIM52-AS1 is significantly correlated to lower overall survival in LIHC patients.
Conclusion:
In conclusion, our analysis identifies potential therapeutic targets in FGFR4 signaling which may contribute to LIHC progression and lower overall survival.

Keywords: LIHC, FGFR4, Co-expression, WGCNA

ICB11-1401-12-1119

Analysis of TPD52 gene in prostate adenocarcinoma and probability analysis of rs757443831 in hsa-mir-33a-5p

Aliasghar Akbarinia[1*], Zahra Maravandi[2]

[1*]Naghshejahan higher education institute.[2]Naghshejahan higher education institute.

*Corresponding author(s). E-mail(s):

Abstract

Introduction
Prostate cancer is the second most common cancer and the fifth leading cause of cancer-related death in men.The progression of cancer in the prostate is a gland in the male reproductive system.Most prostate cancers grow slowly, but some grow relatively quickly.Cancer cells can spread from the prostate to other parts of the body, especially the bones and lymph nodes.This cancer may not show symptoms at first and in its more advanced stages can cause problems such as blood in the urine or pain in the pelvis and waist when urinating.Factors that increase the risk of prostate cancer are age.High and especially genetic. The TPD52 gene activates calcium ion binding activity and protein homodimerization activity and is involved in B cell differentiation located in the endoplasmic reticulum region and the nuclear-cytoplasmic nuclear region.
Materials and Method
The TPD52 gene,which activates calcium ion binding activity and protein homodimerization activity involved in B cell differentiation, was selected at the NCBI site.The expression of this gene was studied at the GEPIA2 site.The microRNAs that acted on this gene were then examined at the MIRWALK site. In LNCRNASNP2, these microRNAs were examined for expression in prostate adenocarcinoma,and finally,the effectiveness of SNPs on microRNAs at the MIRNASNP site was investigated.
Discussion
The TPD52 gene is much more expressed in tumors.The selected has-mir-33a-5p MIRNA had a score:1 and np:12 and this MIRNA has a great effect on the TPD52 gene.This MIRNA had an expression of 11.83 in prostate adenocarcinoma and many SNPs affect this MIRNA.
Conclusion
SNPs are high-density natural sequence variations in genomes and are considered a major genetic source of phenotypic variation within a species and are considered important genetic markers.Based on previous research,it is predicted that SNP rs757443831 is generated between the gene and MIRNA and causes functional dysfunction.

Keywords: LNC, GENE, PROSTATE, MIRNA, TPD52, CANCER

ICB11-1401-12-1120

Using Proteochemometrics models to predict interaction between isoforms of caspase and their inhibitors

Zahra Bastami[1*], Razieh Sheikhpour[2], Parvin Razzaghi[3], Ali Ramazani[4], Sajjad Gharaghani[5]

[1*]Laboratory of bioinformatics and drug design (lbd), institute of biochemistry and biophysics, university of tehran, tehran, iran.[2]Department of computer engineering, faculty of engineering, ardakan university. [3]Department of computer science and information technology, institute for advanced studies in basic sciences (iasbs), zanjan, iran. [4]Cancer gene therapy research center, zanjan university of medical sciences, zanjan, iran. [5]Laboratory of bioinformatics and drug design (lbd), institute of biochemistry and biophysics, university of tehran, tehran, iran.

*Corresponding author(s). E-mail(s): zahra.bastami@gmail.com;

Abstract

In this study we will focus on proteochemometrics modeling which is a new computational approach to the study of drug design for prediction of the interactions between Caspase isoforms and their inhibitors. Caspase is a family of aspartate proteases, that play key roles in programmed cell death and inflammation but what is essential in apoptosis is the division of caspases into two groups of initiating caspases (caspase 2, 8. 9 and 10) and executive caspases (caspase 3, 6, and 7) .
Materials and Methods:
In this project for modelling, we used protein and ligand descriptors. First, the ligands collected from Binding DB (in SDF format) were optimized with hyperchem, and 1444 1D and 2D, along with 431 3D ligand descriptors, were extracted with Padel software. Next, 1441 protein descriptors were extracted using the protr R package, yielding a total of 3317 descriptors. For feature selection we used, NCA (Neighborhood Component Analyses). For PCM modelling three models were developed such as SVR, Decision Tree, and Ensemble.
Results and Conclusion:
A model induced by a machine learning algorithm should be validated. Common methods for validation are; application of the model to a test set, k-fold cross validation and randomized shuffling of the outcomes. The analysis based on R2 and RMSE and Q2 showed that the ensemble model was the best. R2 in this model was 0.81, Q2 was 0.76 and RMSE was 0.036. The criteria for an acceptable model R2 was $>$ 0.6 and Q2 was $>$ 0.5. The results of this study demonstrated that the Ensemble model has better performance than other models in terms of R2, Q2, and RMSE criteria.

Keywords: Caspase, Inhibitor, Proteochemometrics, QSAR, Protein Descriptor, Modeling

ICB11-1401-12-1121

A mathematical model involving the effect of face masks in the control of COVID-19

Hossein Gholami chahkand[1], Mortaza Gachpazan[2*], Majid Erfanian[3]

[1]Ferdowsi university of mashhad. [2*]Ferdowsi university of mashhad.[3]University of zabol.

*Corresponding author(s). E-mail(s): gachpazan@um.ac.ir;

Abstract

In this work, we have constructed a simple model for COVID-19 disease. We have divided our model into seven groups as: First group, dynamics of low-risk susceptible individuals who follows the sanitary recommendations and wear face masks, or they don't have any underlying disease and have received at least one dose of vaccine. Second group, the dynamics of high-risk susceptible individuals includes individuals such as hospital staff members or individuals who have underlying diseases such as asthma or individuals who have not been vaccinated yet. Third group, the dynamics of asymptomatically infected individuals. Fourth group, the dynamics of symptomatically infected individuals. Fifth group, the dynamics of the confirmed asymptomatic cases. Sixth group, the dynamics of the confirmed severe cases and seventh group, the dynamics of recovered individuals. Also in this study, we have interfered the efficacy of face masks to see whether the continuous and suitable use of face masks can inhibit the spread of the COVID-19 pandemic. The positivity and boundedness of solutions are proved. The disease-free equilibrium point and endemic equilibrium point are obtained. The basic reproduction number is computed using the next-generation matrix method given in [1]. It is proved that the equilibrium points are globally asymptotically stable whenever the basic reproduction number is less than one and greater than one respectively. We have used Lyapunov functions in the proof. In addition, we have discussed the sensitivity of the basic reproduction number to each of the variables that it depends on. Finally numerical simulations verify the obtained theoretical results. Simulation for the model is done using Matlab software encoded with an ODE45 solver.

Keywords: Basic reproduction number, Lyapunov function, Disease-free equilibrium point, Endemic equilibrium point

ICB11-1401-12-1122

Molecular dynamics (MD) simulation study on holo and apo forms of the ALS-linked hSOD1 mutation

Nazanin Soleimanifard[1], Bagher Seyedalipour[2], Saman Hosseinkhani[3]

[1]Mazandaran university. [2]Mazandaran university. [3]Department of biochemistry, faculty of biological sciences, tarbiat modares university, tehran, iran.

*Corresponding author(s). E-mail(s):

Abstract

Familial amyotrophic lateral sclerosis (fALS) is a neurodegenerative disorder; approximately 20% is caused by dominant mutations in the gene encoding Cu/Zn human superoxide dismutase (hSOD1). Several mutations have been identified and linked to ALS. This study was designed to evaluate the effect of ALS-associated point mutation, namely N65S, located in loop IV, on structure properties. Molecular dynamics (MD) simulation was performed in order to investigate the possible effects of mutation on the structure of the wild-type protein and compare it with the mutant protein. According to the dynamic analysis, the effect of N65S mutation on the stability of SOD1 protein was predicted and significant structural differences between wild-type and N65S mutant were observed in terms of the radius of gyration (Rg). The average values of Rg were (1/44±0.010), (1/44±0.011), (1/44±0.017), and (1/41±0.013) Åfor apo and holo of wild-type and N65S mutant, respectively. Also, to check the second structure, which, based on the hydrogen bond patterns and some geometrical constraints, each residue is assigned to one of the possible states (α- helices, β-sheet). The secondary structure parameters of hSOD1 exhibited significant changes upon N65S mutation. There is no change of β Sheet content in the form apo wild-type and mutant, while contents of β -Sheet in the form holo Wild-type were more than that of form holo mutant. There is no change of α-helices content in the form holo wild-type and mutant, while contents of α helices in the form apo wild-type were more than that of the form apo mutant. The results of this study showed that the occurrence of mutation at the loop IV of the SOD1 mutant causes conformational changes which lead to neurodegeneration disorders in humans. These findings provided insight into the effect of mutations on the hSOD1, which leads to neurodegeneration disorders in humans.

Keywords: Amyotrophic lateral sclerosis (ALS), N65S variants, Molecular dynamics (MD) Simulation, Human superoxide dismutase-1 (hSOD1)

ICB11-1401-12-1123

Prediction of the Binding Site of Kryptofix-221 to Na-dependent β-Galactosidase: A Molecular Docking study

Maryam Alipour khomami[1], S.shirin Shahangian[2*], Navid Pourzardosht[3]

[1]Department of biology, faculty of sciences, university of guilan, rasht, iran. [2*]Department of biology, faculty of sciences, university of guilan, rasht, iran.[3]Department of biochemistry, guilan university of medical sciences, rasht, iran.

*Corresponding author(s). E-mail(s): shahangian@guilan.ac.ir;

Abstract

Due to its nutritional, biotechnological and therapeutic applications, β-D-Galactosidase (E.C.3.2.1.23) has been extensively studied. It catalyzes the hydrolysis of lactose to glucose and galactose. β-galactosidase from E.coli is a 464kDa homotetramer. Each subunit (1,023 residues) consists of five domains, the third being an alpha/beta barrel that contains most of the active site residues. Sodium ion is critical for generating the action potential in nervous and cardiac tissues, and its level in blood is an excellent biomarker to show electrolyte imbalance and provide valuable information regarding an individual's physical and mental situation Accordingly, measurement of Na+ level in blood is of great clinical importance. Measuring serum level of Na+ is based on the sodium-dependent activity of β-galactosidase. In this method, Kryptofix-221 (Na+-binding agent) is used to prevent the enzyme reaction from being saturated with excessive sodium ions. Pre-incubation of the enzyme with Kryptofix-221 leads to a decrease in enzyme stability. Regarding this problem, efforts were raised to analyze the ligand interaction with β-galactosidase. The 3D structures of the enzyme and Kryptofix-221 were obtained from PDB (ID:1D0P) and PubChem (ID:123438), respectively. It was checked by Chimera v.1.16 for hydrogen and charge addition. The active site pocket of the enzyme was evaluated using CASTp web server, and YASARA was used for energy minimization of the protein. Docking studies between Kryptofix-221 and β-galactosidase were carried out using Autodock vina v.1.1.2. The ligand interaction plots of protein-ligand complexes were illustrated by Ligplot+ v.2.2.7 program. Enzyme showed optimum binding affinity to target ligand with the binding energy of −5.6 kcal/mol. Based on docking analyses, seven residues, including Val-103, His-418, Glu-461, Met-502, Tyr-503, Phe-512, and Trp-999 are involved in non-bonding interactions with Kryptofix-221. Moreover, Kryptofix-221 can form one hydrogen bond with Lys-517.

Keywords: β-Galactosidase, Sodium sensing, Molecular docking, Kryptofix

ICB11-1401-12-1124

Phylogenetic study of L-asparaginase II (AsnB) in the bacteria domain and fungi kingdom

Mahdis Mofidi[1], Maryam Azimzadeh irani[2], Gholamhossein Ebrahimipour[3], Mohammad Yaghoubi-avini[4*]

[1]Shahid beheshti university. [2]Shahid beheshti university. [3]Shahid beheshti university. [4*]Shahid beheshti university.

*Corresponding author(s). E-mail(s): m_yaghoubi@sbu.ac.ir;

Abstract

L-asparaginase II (AsnB), an important enzyme in the pharmaceutical field, is often used in the treatment of acute lymphoblastic leukemia (ALL). Only AsnB from E.coli K12 and Erwinia has been used as a chemotherap agent. Although these enzymes have been successfully used in treatment, several medical complications like severe immune responses leading to sensitivity, and anaphylaxis. have limited their use. Thus, there is a need to search for other sources of AsnB with novel therapeutic properties and rational design to reduce side effects. Thirty full-length amino acid sequences of AsnBs from diverse bacteria and fungi were collected from the protein database of the National Center for Biotechnology Information. The multiple sequence alignment was performed through the MAFFT algorithm and curated by the BMGE method on the online server NGPhylogeny. The phylogenetic tree was created using the maximum likelihood method based on the Le Gascuel model by Mega-7 software. The phylogenetic tree of AsnBs shows three distinct clusters of enzymes distributed in Gram-negative bacteria including therapeutic E.coli K12 enzyme, Gram-positive bacteria, and a cluster consisting of B.subtilis sequences, fungi, and E.chrysantemi which indicates plausible horizontal gene transfer among these species in distance times. Due to the low bootstrap values, the exact position of the branches shows some uncertainty. Two FDA-approved AsnBs have different phylogenies. Although other asparaginases with anti-tumor properties are more related to E.coli; the enzyme resemblance of fungi and B.subtilis to Erwinia enzyme suggests that these microorganisms could serve as a suitable substitute. Also, The tree shows the divergence in the phylogeny of the Gram-positive group.Suggesting that this group's asparaginase could serve as a suitable substitute for current drugs, as their immunogenicity may differ significantly from those of commercial enzymes. The potential of other AsnB of Gram-positive bacteria and fungi should be explored to obtain an enzyme with fewer side effects against (ALL).

Keywords: L-asparaginase II, Acute lymphoblastic leukemia, Phylogenetic analysis, Immunogenicity

ICB11-1401-12-1125

Towards enhancing L-Asparaginase stability by computational-structural modeling

Leila Sarafan soleimanzade[1], Mohammad Yaghoubi-avini[2], Gholamhosein Ebrahimipour[3], Maryam Azimzadeh irani[4*]

[1]Shahid beheshti. [2]Shahid beheshti. [3]Shahid beheshti. [4*]Shahid beheshti.

*Corresponding author(s). E-mail(s): m_azimzadeh@sbu.ac.ir;

Abstract

Abstract
L-asparaginase enzymes are widely used for the treatment of lymphoblastic leukemia. The extracellular and glutaminase-free asparaginases have been used for leukemia treatment. Asparaginase folds as homotetramer with four active sites. The enzyme active sites contain five residues, Thr12, Tyr25, Thr89, Asp90, and Lys162 (according to E.coli numbering). Two therapeutically important sources are Escherichia coli and Erwinia chrysanthemi. To enhance stability, the enzymes need to be optimized by protein engineering. In this study, structural alignment and electrostatic potential maps were used to examine structural features of the L-asparaginase active site in the two most important medicinal species to enhance enzymatic stability.
Material & Methods
The structures of E.coli and Erwinia chrysanthemi enzymes (PDB ID: 3ECA, 5F52 respectively) were obtained from Protein Data Bank (PDB). Structural alignment has been done with PyMol Molecular Graphics. Molecular preparation for Poisson-Boltzmann calculations was obtained from the pdb2pqr server. For APBS analysis, the PyMol APSB plugin determined the electrostatic properties of the enzyme.
Results
APBS results show the negative overall electrostatic potential of the E.coli active site, while it shows a positive trend in Erwinia chrysanthemi. The negatively charged pockets of E.coli contain six Asp and one Glu, and Erwinia positively charged ones have two His, eight Arg, and three Lys. However, the main interactions of the active sites remain intact in both enzymes that occur through polar interactions. Ligand has 11 polar interactions in the active site with residues; Thr12, Tyr25, Ser58, Gln59, His87, Thr89, Asp90, Thr91, Lys162, Arg248, and Glu283.
Conclusions
Although the essential interactions in these two medical enzymes are the same, the differently charged binding pockets could be a promising target for further enzyme optimization designs.

Keywords: L-Asparaginase, Leukemia, Binding pocket, Protein engineering, Structural modeling

ICB11-1401-12-1126

In-silico emplacement of curcumin on higher-order assemblies of cellulose nanofibers

Akram Hamedi[1], Mehdi Jahanfar[2], Naser Farrokhi[3], Maryam Azimzadeh irani[4*]

[1]Shahid beheshti. [2]Shahid beheshti. [3]Shahid beheshti. [4*]Shahid beheshti.

*Corresponding author(s). E-mail(s): m_azimzadeh@sbu.ac.ir;

Abstract

Cellulose nanofibers (CNs) due to being biocompatible, biodegradable, and non-toxic biopolymers with low immunogenicity have been used for drug delivery .Curcumin, an active pharmaceutical ingredient in turmeric, is used in treating cancer with inflammatory conditions .To enhance the curcumin-induced medication, the interaction between the medicinal substituents with the drug raft, here CNs, needs to be resolved at the molecular level .Recent computational and simulation studies on the atomic constructions of CNs units show that the two primary forms of cylindrical and planar can be envisaged.In this work, molecular modeling and docking were performed to investigate the atomistic interactions between curcumin and planar/cylindrical forms of CNs.
Methods:
Curcumin structure was obtained from the PubChem database (PubChemCID:969516). Cylindrical and plane cellulose structures, were obtained from a molecular dynamics study on the same assemblies. Autodock Vina via the Chimera platform was used for docking.The curcumin-cellulose binding poses were selected based on the binding free energy, RMSD, and the number of polar contacts in each arrangement. Molecular visualizations and RMSD calculations were carried out by PyMOL.
Results:
In total, four binding poses were selected. Curcumin seems to interact with the middle residues of the fibers in the cylinder conformation. Docking scores are -3.7 and -3.5 for the selected binding poses. While two and one hydrogen bonds maintain the curcumin placement on the cylinder in each arrangement. The most favorable binding poses in the plane arrangement belong to curcumin residing on the surface polymers. Docking scores were -5.2 and -4.6 for the selected poses. Presenting one and two hydrogen bonds between the curcumin and the polymer unit from the plane arrangements.
Conclusion:
Docking of curcumin on cylinder and plane arrangements of cellulose nanofibers shows that the dominant placement of the curcumin is on the middle residues and surface polymers of the higher-order assemblies.

Keywords: Cylinder and plane assemblies, Molecular modeling, Molecular docking

ICB11-1401-12-1127

Identifying Potential Inhibitory Ligands Against σ2 receptor

Atena sadat Hosseini[1*]

[1*]Islamic azad university of yadegar imam.

*Corresponding author(s). E-mail(s): atena79hosseini@gmail.com;

Abstract

Studies on ligand-binding have demonstrated that the $\boldsymbol{\sigma}$2 receptor, which is known to be a membrane protein that is strongly expressed in the liver and kidney, as well as in a number of cancer cell lines and tumors that are in the process of proliferating. Drugs that target the $\boldsymbol{\sigma}$2 receptor have recently emerged as promising therapeutic targets, and these compounds are currently being tested in clinical trials for the diagnosis of breast cancer as well as for the treatment of Alzheimer's disease and schizophrenia . In this study, we identified some potential drugs that are highly possible to inhibit the $\boldsymbol{\sigma}$2 receptor. The structure of the $\boldsymbol{\sigma}$2 receptor was downloaded from the AlphaFold database (AlphaFold ID: Q5BJF2). In order to refine and minimize the structure of the protein, a 100 ns Molecular dynamic simulation was done using Gromacs 2020 with Amber99SB forcefield. In the next step, 200 drug-like ligands were downloaded from the ZINC12 database and they split and convert to PDBQT files through Raccoon software and the target file was prepared using Autodock tools . The virtual screening was performed using Autodock Vina (Eberhardt and Santos-Martins). The 5 top five ligands were ZINC00850103, ZINC00664787, ZINC00387261, ZINC00633958, and ZINC00877571 with the affinity of -10.7, -10.6, -9.8, -9.7, and -9.7 Kcal/mol, respectively. In the last step, the ADME analysis was performed using the SWISSADME database on candidate ligands. According to physiochemical properties such as Molecular Weight (MW) and log P, ZINC00850103 (MW: 487.55 g/mol and log P:3.49), and ZINC00633958 (MW: 470.42 g/mol and log P: 3.45) were selected as potential ligands that may inhibitory effect on $\boldsymbol{\sigma}$2 receptor.

Keywords: $\boldsymbol{\sigma}$2 receptor, Drug-like Ligands, Molecular Dynamic Simulation, virtual screening

ICB11-1401-12-1128

A comprehensive bioinformatic evaluation the NTRK family as prognostic biomarkers in breast cancer

Ramtin Mohammadi[1], Peyman Kheirandish zarandi[2], Shahla Mohammad ganji[3]

[1]National institute of genetic engineering and biotechnology. [2]Science and research branch, islamic azad university. [3]National institute of genetic engineering and biotechnology.

*Corresponding author(s). E-mail(s):

Abstract

Introduction:
Breast cancer (BC) is the most often prevalent and diagnosed cancer worldwide, accounting for 1 in 8 cancer diagnoses and a total of 2.3 million. In 2040, it is anticipated that BC would account for nearly 3 million new cases and 1 million mortality. The neurotrophic tyrosine receptor kinase (NTRK) genes 1/2/3 respectively encode tropomyosin receptor kinases (TRK) A/B/C. Intrachromosomal rearrangements resulting in NTRK gene fusions may lead to constitutive activation of TRK proteins, which subsequently function as oncogenic drivers by activating cellular growth pathways. Fusions of the NTRK gene occur in 0.3% of all solid tumors; meanwhile the incidence varies by cancer type. Their frequency is more than 90% in uncommon malignancies such as secretory breast carcinoma and mammary analogue secretory carcinoma of the salivary gland (MASC). The intent of this bioinformatics study was to clarify the prognostic significance of the NTRK family in BC.
Material and methods: We aimed to conduct a comprehensive bioinformatics analysis elucidating the prognostic values of the NTRK families in BC. Therefore, TCGA, UALCAN, Kaplan–Meier plotter, bc-GenExMiner, cBioPortal, STRING, Enrichr, GeneMANIA, and TIMER were utilized for analysis.
Results:
Our bioinformatics analysis indicates that high levels of NTRK2/3 better association with overall survival (OS) and recurrence-free survival (RFS) in BC patients; while high levels of NTRK1 showed applicable correlation with RFS in BC patients. Consequently, our findings indicated that NTRK3 were notably correlated with the clinical outcomes of BC patients.
Conclusion:
Our findings offer a fresh perspective that might aid in the therapeutic use of NTRK as predictive biomarkers in BC.

Keywords: NTRK, biomarker, prognostic, breast cancer

ICB11-1401-12-1129

Gene selection in classifying Alzheimer's disease using a statistical, biological, and artificial intelligence-based pipeline

Hamed Khaiiate ajami[1], Sepideh Parvizpour[2], Jafar Razmara[3*]

[1]University of tabriz. [2]Tabriz university of medical sciences. [3*]University of tabriz.

*Corresponding author(s). E-mail(s): razmara@tabrizu.ac.ir;

Abstract

Diagnosing a disease based on gene expression data extracted from microarrays is still an open field of research. Due to the availability of whole-genome data through microarrays technology, diagnosis accuracy is expected to be improved. Despite the high potential of the data prepared by the technology, their analysis on different platforms shows that they may differ for different samples with respect to biomarker status. This affects the diagnosis accuracy because of the existing bias between two different experimental conditions. To address this problem, we propose a pipeline-based approach using statistical analysis of biological data combined with artificial intelligence techniques. At first, the p-value and a new score based on a gene interaction network are used to evaluate genes. The p-value helps us to find differentially expressed genes. The new score, called the evidence score, measures the compliance level of the differentially expressed genes with past biological evidence. Next, we take advantage of artificial intelligence methods to find the subset of genes that define high separability between normal and affected samples. To this end, we employed a genetic algorithm to find the best subset. The performance of the pipeline was compared with other state-of-the-art methods. The results indicate that the proposed method can obtain fruitful predictive performance for diagnosing Alzheimer's disease.

Keywords: gene expression, gene interaction network, artificial intelligence

ICB11-1401-12-1130

Investigation miRNAs that might attenuate Diabetes and Polycystic Ovary Syndrome: an In-silico study

Mohammad javad Javid-naderi[1*], Ali Mahmoudi[2]

[1*]Department of medical biotechnology and nanotechnology, faculty of medicine, mashhad university of medical sciences, mashhad, iran.[2]Department of medical biotechnology and nanotechnology, faculty of medicine, mashhad university of medical sciences, mashhad, iran.

*Corresponding author(s). E-mail(s): Javidmj991@mums.ac.ir;

Abstract

BACKGROUND:
Polycystic ovary syndrome (PCOS) is a common infertility condition in reproductive-age females with creation of many cysts. The prevalence of PCOS in females was reported between 15-20%. Diabetes is also a prevalent malady caused by dysfunction in insulin production or insulin occupation. It was reported that 27 million people ranging from 20–79 years of age have diabetes in 2021 . Diabetes is a common PCOS-associated event that contributes to initial PCOS. Defects in the many functions of proteins or gene expression are the main reason for those disorders. Based on the bioinformatics method, we selected therapeutic miRNAs against the important proteins concerning both diseases.
METHODS:
We primarily searched the common proteins relatives to diabetes and PCOS-related proteins in several gene-diseases databases. To discover important hub genes, we build a PPI network for common proteins achieved via Cytoscape software. Besides, we conducted an enrichment examination to assess biological pathways and processes (Gene Ontology and KEGG) by Enrichr. Next, we discover the potential connections of miRNAs therapeutic against imperative hub proteins of diabetes and PCOS.
RESULTS:
We recognized 729 intersection genes connected with diabetes and PCOS. PPI network analysis exhibited 10 hub genes (INS, AKT1, ACTB, ALB, TNF, IL6, CTNNB1, EGFR, STAT3, and PPARG). The biological processes conveyed by important intersection genes are mainly concerned with positive regulation of protein phosphorylation, positive regulation of gene expression, and positive regulation of intracellular signal transduction. In KEGG enrichment, the majority got involved with the FoxO, AGE-RAGE, and Focal adhesion signaling pathways. MiRNet analysis also showed that hsa-mir-34a-5p, hsa-mir-1-3p, and hsa-mir-155-5p had the highest relation with hub proteins related to two conditions.
CONCLUSION:
We revealed has-mir-34a-5p, hsa-mir-1-3p, and hsa-mir-155-5p could be considered as therapeutic agents against diabetes and PCOS.

Keywords: Diabetes, KEGG, Gene Ontology, miRNA, PCOS, PPI network

ICB11-1401-12-1131

Genome expression profiling of the human brain affected by Mirtazapine through transcriptomics data analysis

Zeynab Aghaie[1], Reza Mohammadi[2*]

[1]Faculty of Chemistry and Chemical Engineering, Malek Ashtar University of Technology, Iran.
[2]Faculty of Chemistry and Chemical Engineering, Malek Ashtar University of Technology, Iran.

*Corresponding author(s). E-mail(s): rezamohammadi@mut.ac.ir;

Abstract

Mirtazapine is an antidepressant drug that improves mood by blocking alpha-2, HT2A-5, and HT2C-5 adrenergic receptors. In this research, the expression profile of the human brain genome under the influence of Mirtazapine is investigated through the analysis of expression data (transcriptomics). In the biomolecular field, gene expression profiling measures the activity of thousands of genes simultaneously to create an image from the function of the cell. This profile can determine how a cell responds to treatment. This study used the available data in valid databases such as GEO datasets. In this regard, we analyzed the genome of people after taking different doses of Mirtazapine by Bioconductor in R software. According to the final results, the most important genes that have changed expression levels by mirtazapine treatment and have played a central role in the recovery of patients are the MDM2 (mouse double minute 2), CCND1 (Cyclin D1), and MYCN (MYCN proto-oncogene). These genes are involved in essential pathways such as glioma formation, regulating the cell cycle, metabolic pathways, proliferation, and retinoblastoma protein activation. The Nitric Oxide Synthase 2 (NOS2) gene, along with Peptide Deformylase (PDFA) and Early Growth Response 1 (EGR1) genes, helps to prevent depression, where the expression of all three genes has increased by the use of Mirtazapine. Upregulation of the RUNX1 (RUNX Family Transcription Factor 1) gene plays a role in the forebrain transcription repressor pathway. In the down-regulated genes group, BAIAP2 (BAR/IMD Domain Containing Adaptor Protein 2) gene involves stress fiber formation, brain angiogenesis, and brain tumor formation with P53 protein in response to stress. NR4A2 gene mutation in this gene is involved in Parkinson's disorders, schizophrenia, depression, and madness. COL1A1 gene is involved in related pathways in depression and mood disorders, and TBX3 and STAT2 genes are involved in stress-related signaling.

Keywords: Mirtazapine, Genome expression profiling, Analysis of expressive data, GEO dataset, Microarray

ICB11-1401-12-1132

New Encoding Model Based on gRNA-DNA Pairs to Predict off-target Effects of Genome Editing with CRISPR/Cas9

Roghayyeh Alipanahi[1*], Leila Safari[2], Alireza Khanteymoori[3]

[1*]Zanjan university.[2]Zanjan university. [3]Universitätsklinikum freiburg.

*Corresponding author(s). E-mail(s): alipanahir@gmail.com;

Abstract

CRISPR/Cas9 is a new genome-editing technology used in biomedical applications. To make genome editing with CRISPR far more precise and practical, we must concentrate on predicting CRISPR off-target effects and try to decrease them. Although numerous computational models have been developed to predict off-target activities, the existing methods suffer from low precision for gene editing at the clinical level. In addition, the inputs of most of these algorithms are gRNA sequences in on-hot vector encoding form. However, recent research illustrated that both gRNA and DNA strongly impact on the prediction precision of off-target activity. To address these problems, we propose a novel encoding scheme of gRNA-DNA sequences, and deploy it in two deep neural network-based architectures, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to predict off-target effects. The comparison of off-target prediction results based on our proposed gRNA-DNA encoding scheme with state-of-the-art on two popular gene-editing datasets, CRISPOR and GUIDE-seq, reveals the superiority of our approach based on the area under the Receiver Operating Characteristic (AUROC) curve criteria which is the promising value of up to 98%.

Keywords: CRISPR/Cas, gRNA Design, Off-target, Encoding, Machine learning

ICB11-1401-12-1133

Identification of key genes in non-functioning pituitary adenoma using gene expression profile and protein-protein interaction network

Nahid Safari-alighiarloo[1*], Seyyed mohammad Tabatabaei[2], Nasibeh Khayer[3], Nahid Hashemi madani[4], Mohammad e. Khamseh[5]

[1*]Endocrine research center, institute of endocrinology and metabolism, iran university of medical sciences, tehran, iran.[2]Medical informatics department, school of medicine, mashhad university of medical sciences, mashhad, iran. [3]Skull base research center, the five senses health institute, iran university of medical sciences, tehran, iran. [4]Endocrine research center, institute of endocrinology and metabolism, iran university of medical sciences, tehran, iran. [5]Endocrine research center, institute of endocrinology and metabolism, iran university of medical sciences, tehran, iran.

*Corresponding author(s). E-mail(s): safari.nhd@gmail.com;

Abstract

Introduction:
The non-functioning pituitary adenoma (NFPA) is a common intracranial tumor with local invasive potential and a high recurrence rate . The only treatment options are surgery and radiation . There is an urgent need to develop novel therapeutic strategies to reduce the recurrence rate of NFPAs, either through biomarkers or medical therapies. We aimed to identify biomarkers and biological pathways associated with the progression of NFPA by construction and analysis of protein-protein interaction (PPI) network based on differentially expressed genes (DEGs).
Methods:
A microarray dataset (GSE26966) from Gene Expression Omnibus (GEO) was employed to determine DEGs between normal pituitary tissues and NFPA tissues. DEGs were identified using GEOquery and Limma packages. DEGs were mapped on experimentally confirmed PPIs retrieved from the IntAct database (02-2022). The constructed PPI network's topological analyses (Degree, Betweenness, and Closeness centrality) were performed using the igraph package. Functional clusters were identified using the MCL package. The enrichment of clusters was performed using the Enrichr database.
Results:
We identified 1135 DEGs in the NFPA dataset considering adjusted p-values <0.05 and log fold change>2. The constructed PPI network was analyzed, and high-centrality genes were identified. Using functional analysis of clusters, we identified potential candidate genes in NFPA, which were the more central genes in the PPI network and involved in biological pathways. EGFR, MET, KIT, ERBB4, and BNDF, enriched in the MAPK signaling pathway and Ras signaling pathway, CDKN1A, and CDKN2A, enriched in the cell cycle, FGFR3, enriched in the regulation of actin cytoskeleton, STAT3, enriched in JAK-STAT signaling pathway, and HSPA2, enriched in protein processing in the endoplasmic reticulum, were identified as key candidate genes.
Conclusion:
This study provides new insights into network biomarkers that may be potential therapeutic targets in NFPA.

Keywords: Non-functioning pituitary adenoma (NFPA), gene expression, protein-protein interaction network (PPI), systems biology

ICB11-1401-12-1134

Drug-Disease data integration for drug repurposing using deep neural networks

Ramin Amiri[1], Jafar Razmara[2*], Habib Izadkhah[3], Sepideh Parvizpour[4]

[1]University of tabriz. [2*]University of tabriz.[3]University of tabriz. [4]Tabriz university of medical sciences.

*Corresponding author(s). E-mail(s): razmaraj@gmail.com;

Abstract

The coronavirus disease has led to a rush to repurpose existing drugs, although the underlying evidence base is of variable quality. Drug repurposing is becoming an increasingly attractive direction in drug discovery and development because it involves potentially shorter development timeline and lower development cost than designing a new drug. According to statistics, only 30% of the drugs designed and manufactured each year are approved by the Food and Drug Administration (FDA), which proves the reasons for the necessity of using existing approved drugs for various diseases (drug-repurposing). Drug repurposing needs to consider different aspects of drugs and diseases in order to efficiently find new targets for an existing drug. In this research, we propose a model for integration of different data related to drug and disease. Then, we employ a convolutional neural network to capture similarities between data and repurpose drugs for target diseases. To this end, stratified 5-fold cross-validation technique was used for evaluation of the methods. The results of the proposed method for the parameters of accuracy, precision, recall, and F1-measure are 0.97, 0.69, 0.96 and 0.84 respectively. The results of comparative evaluations indicate the high performance and efficiency of the method compared to the state-of-the-art methods.

Keywords: drug repurposing, deep learning, convolutional neural networks

ICB11-1401-12-1135

Phylogenetic and structural analysis of native avian influenza subtype H9N2 hemagglutinin isolated

Sogand Shayesteh[1], Naser Farrokhi[2*], Farahnaz Motamedi sedeh[3], Shahrokh Ghovvati[4]

[1]Shahid beheshti university. [2*]Shahid beheshti university.[3]Nuclear agriculture research school, nuclear science and technology research institute, atomic energy organization of iran. [4]University of guilan.

*Corresponding author(s). E-mail(s): n_farrokhi@sbu.ac.ir;

Abstract

The widespread prevalence of H9N2 avian influenza infections in poultry and humans has raised concerns about the possibility of converting these viruses into strains of epidemics and the incidence of a dangerous pandemic. The study of phylogenetic relationships and the structure of the hemagglutinin protein of the native variants is deemed important in the design of effective vaccines for application in Iran and the Middle East. Blastp analysis was performed to evaluate the similarity of native avian influenza subtype H9N2 hemagglutinin isolated with others, and the phylogenetic tree was generated by DNAMAN version 10. Further analyses on hemagglutinin were carried out by Targetp, Protparam, Scanprosite, and Protscale. Hemagglutinin is a secretory protein (pI = 6.48 with an instability index of #31) that binds to host cell surface receptors, integrating within the host membrane for successful infiltration. Hemagglutinin undergoes various post-translational modifications that can affect its relationship with the pathogenicity of influenza virus. Hemagglutinin also has hydrophobicity at different places (Kyte & Doolittle hydrophobicity scale). According to Blastp analysis, aligned sequences from Pakistan and Saudi Arabia have a high degree of similarity, which can be attributed to the high similarity of H9N2 virus strains infecting the Middle East. DNAMAN was illustrative of little changes on hemagglutinin samples from sick poultries sampled in Iran during 1998–2014; calssifying H9N2 in a group of viruses that are hardly mutate. This might be due to little selection pressure on the virus for its control by any means and just root out the sick animals. Moreover, this may be helpful in the design of a promising vaccine against hemagglutinin. However, following vaccine application, the onset of the disease and the possible sequence changes need to be monitored over years.

Keywords: Avian influenza virus, H9N2, Hemagglutinin, Bioinformatics tools

ICB11-1401-12-1136

A sequence and structure analysis on the Human Arc protein

Seyyedeh hediyeh Mousaviyan sesari[1], Emran Heshmati[2*], Khosrow Khalifeh[3]

[1]University of zanjan. [2*]University of zanjan.[3]University of zanjan.

*Corresponding author(s). E-mail(s): heshmati@znu.ac.ir;

Abstract

Arc (Activity regulated cytoskeleton associated protein) is a key regulator of synaptic plasticity. Its gene was emerged from the retrotransposon Gag capsid genes . In the nervous system, it is responsible for intercellular transfer of mRNA to the cytoplasm of neurons through endocytosis via its virion-like capsid. In current work, a comperehensive sequence and structural analysis was performed on the Human Arc protein and its homologoues counterparts found in other organisms. For sequence analysis, we used Bioinformatics tools including the BLAST , ClustalW , ESpript and finally a representative sequence of programs in the Phylip package to deduce the evolutionary history of the Arc proteins. The structural features of the Arc protein have been studied using a combination of Bioinformatics and Computational tools. Currently, there is no structural report on the whole structure of protein, and the reported structural data in the PDB database is limited to specific fragments of the original sequence from various organisms (UniProt codes: Q7LC44, Q63053, Q7K1U0, and Q7JV70). We used these fragments as templates to build a compact structure for the human variant Arc protein. Due to the structural data limitation, the final structure was built and validate with a combinatorial approach using respective online servers and the MODELLER program (Ver. 10.4). We used the capabilities of the UCSF Chimera program to visualize and structural analysis on the selected models.

Keywords: Arc, Gag, Activity Regulated Cytoskeleton Associated Protein, Sequence analysis, Structure analysis, Bioinformatics

ICB11-1401-12-1137

Identification of the Genes and Pathways in Glioblastoma cell lines by bioinformatics analysis

Ali Abedini[1], Masoumeh Moslemi[2], Ehsan Sohrabi[3*]

[1]Young researchers and elite club, central tehran branch, islamic azad university, tehran, iran. [2]Department of medical genetics and molecular biology, faculty of medicine, iran university of medical sciences (iums), tehran, iran. [3*]Department of medical genetics and molecular biology, faculty of medicine, iran university of medical sciences (iums), tehran, iran.

*Corresponding author(s). E-mail(s): e.sohrabi95@gmail.com;

Abstract

Glioblastoma is the most common type of malignant brain tumor which characterized by high mortality and poor prognosis. The average survival of this cancer has estimated about 8 to 10 months. Therefore, it is important to choose the most effective strategy for molecular diagnosis and interactions to treatment. For this aim, bioinformatics analysis of microarray data was applied to find the association of decreased and increased in gene expressions and relevant miRNAs in glioblastoma cell lines.The microarray datasets GSE50173, GSE9171 and GSE15824 which was composed of gene expression data were downloaded from the GEO database. Each dataset was processed by using the log2 transformation and normalized by Package of Affy and RMA in R software. Differently expressed genes with P.value <0.05 and fold change (| FC|)>2 were selected as biomarkers in all three datasets. Functional and pathway enrichment analysis were performed by using the BioDB and EnrichR databases, although, protein–protein interaction network was constructed by Cytoscape software. A total of 35446 Genes in three datasets showed significantly differences in expression levels (P-value <0.05), among them, approximately 67 genes showed Up regulation with log Fold change more than 2 and 683 genes down regulation with log Fold change less than 2. To assess the function of the differently expressed genes, the gene ontology, Kyoto Encyclopedia of Genes and Genomes and pathway enrichment analyses from EnrichR with adjusted p-value <0.05 were used. It is showed that up regulated genes were enriched in “cytosolic part”,”ribosome”,”focal adhesion” and down regulated genes were in “RNA binding”,” insulin receptor binding”, ”double-stranded DNA binding” significantly (p-value <0.05). In this study, we used some databases to determine the variation expression of some target genes in glioblastoma cell lines.

Keywords: Glioblastoma, Glioblastoma cell line, Microarray analysis, R soft ware

ICB11-1401-12-1138

Systems Biology Assessment and Transcriptome Analysis of the Interaction of Probiotic and Host in the Anti-Inflammatory Process

Arezu Heydari[1], Farshid Parvini[2], Najaf Allahyari fard[3*]

[1]Department of biotechnology, faculty of biotechnology, semnan university, semnan, iran. [2]Department of biology, faculty of basic sciences, semnan university, semnan, iran. [3*]Department of systems biotechnology, national institute of genetic engineering & biotechnology (nigeb), tehran, iran.

*Corresponding author(s). E-mail(s): allahyar@nigeb.ac.ir;

Abstract

Inflammatory diseases in the gut are linked with high levels of inflammatory oxidants, including hydrogen peroxide (H2O2) and hypochlorous acid (HOCI), which are antimicrobial compounds. Probiotics' effectiveness in treating inflammation in the intestine depends on their ability to survive against H2O2 and HOCI stresses. In this study, by using the bioinformatics approach and the RNA sequencing data, the essential genes and pathways of resistance to H2O2 and HOCl stress in Lactobacillus reuteri (L. reuteri) was evaluated and investigated. After checking the GEO and SRA databases, the data related to L. reuteri (GSE127961) was retrieved, containing a probiotic under H2O2 and HOCI oxidative stress compared to control group (without oxidative stress). The data was analyzed by CLC software. After obtaining the differentially expressed genes, analyzes were performed, such as investigating the gene/protein interaction network, extracting key genes, and finally evaluating the biological pathways and ontology of these genes. Based on the results, the GSE127961 related to L. reuteri (HOCI and H2O2 stresses) respectively, 64 and 198 DEGs were obtained. According to the results, deoC, mfd, uvrB and grpE genes in HOCI stress and Lreu_0645 gene in H2O2 stress are considered as key genes that probably play an essential role in the resistance of these probiotics to HOCI and H2O2 stress. Also, these key genes play a role in biological pathways such as "homologous recombination", "mismatch repair", "nucleotide excision repair", and "pentose phosphate pathway" in stress conditions. The effectiveness of probiotics in the host depends on their resistance and survival against oxidative stress such as H2O2 and HOCI. Therefore, the key genes and biological pathways obtained in this study can be used to improve probiotics' efficacy in treating inflammatory diseases.

Keywords: Intestinal inflammation, Probiotics, Transcriptome, Lactobacillus reuteri, Resistance

ICB11-1401-12-1139

Inhibition of competitive pathways to increase butanetriol production in the Escherichia coli through metabolic network simulation

Haniyeh Tamjidi[1], Reza Mohammadi[2*]

[1]Faculty of Chemistry and Chemical Engineering, Malek Ashtar University of Technology, Iran.
[2*]Faculty of Chemistry and Chemical Engineering, Malek Ashtar University of Technology, Iran.

*Corresponding author(s). E-mail(s): rezamohammadi@mut.ac.ir;

Abstract

Supplying clean, renewable, and abundant energy is one of the main concerns of today's human civilization. Fossil fuels are limited and expensive, and damages caused by their use have made people turn to biofuels which are part of renewable energy. One of these valuable substances that can be obtained from microorganisms today with the advancement of metabolic engineering and bioinformatics is 1.2.4-butanediol, an unnatural chemical substance with high value, and its properties are similar to glycerol. This chemical has essential applications as a precursor in manufacturing 1,2,4-butanediol trinitrate as a product in the military industry, polymers, and medicines. In this study, the genome-scale metabolic network reconstruction of Escherichia coli (iML1515) was modeled for butanetriol production from the cheap carbon source, xylose. For this purpose, we applied the "Constraint-Based Reconstruction and Analysis" (COBRA) method by using the COBRA toolbox, an extension implemented in MATLAB software. The pathway for butanetriol production from xylose was optimized, and competing pathways that consumed xylose were blocked. The result of the FBA analysis showed that the reconstructed network for butanetriol production has flux. Hence, its output in the metabolic network of Escherichia coli was possible in silico. Finally, an attempt was made to determine the necessary and unnecessary reactions for improving butanetriol production by applying additional constraints and making some transcriptomics analysis. Considering that the complete deletion of competitive genes decreased the amount of biomass production by increasing the time required for proliferation; so by reducing the expression of competitive genes and not deletion, the amount of the butanetriol production reached 20mMol/gDW/hr, which is a significant increment compared to the time of complete genes deletion.

Keywords: Metabolic network modeling, 1, 2, 4-butanetriol, E. coli , COBRA toolbox

ICB11-1401-12-1140

CCL20 and CNTNAP3C as prognostic biomarkers and therapeutic targets in HCT116 cell lines treated with Fusobacterium nucleatum for colorectal cancer metastasis

Seyed taleb Houseini[1], Arash Sattari[2], Roya Bishehkolaei[3], Farkhondeh Nemati[4*]

[1]Islamic azad university. [2]Islamic azad university. [3]Islamic azad university. [4*]Islamic azad university.

*Corresponding author(s). E-mail(s): farkhondehnemati@gmail.com;

Abstract

Introduction:
Colorectal cancer (CRC) is the fourth most commonly diagnosed cancer worldwide and the second most frequent cause of death closely following lung cancer. Cancer metastasis, a major clinical problem of human cancer, is responsible for most colorectal cancer patient mortality. Thus, it is important to elucidate the underlying mechanisms of metastasis in CRC patients. F. nucleatum is a Gram negative anaerobic oncogenic bacterium enriched in human colorectal carcinoma compared with adjacent normal tissues. However, the potential effects and underlying mechanisms of F. nucleatum in the interactions between cells in CRC metastasis remain largely unknown. Material and Method:
SRP313780 dataset was downloaded from https://www.ncbi.nlm.nih.gov/sra , the sra file was converted to two Fastq paired files with sratoolkit.2.11.3- software. Quality control of Fastq files was performed with FASTQC software, and low quality reads were trimmed with Trimmomatic software and we mapped the trimmed file to GRCH38_genome using Hisat2 software and sam file was obtained. In the next step, we converted the sam file to count file using htseq-count software and finally, DEGs were identified using the DESeq2 package in R Studio software by threshold (FDR:0.05 and log2foldchange>1).
Results:
Based on the results, in the HCT116 cell lines treated with Fusobacterium nucleatum compared with control group, upregulated genes included: CCL20, CXCL8, CSF2, CXCL1 and CXCL2. Downregulated genes included: CNTNAP3C, NCR3LG1, LOC100128653, EGFR-AS1 and ANKRD36BP1.
Conclusion:
The overabundance of F. nucleatum in CRC tissues has been reported to promote CRC cell proliferation, recurrence and chemo resistance. Due to the important role of Fusobacterium nucleatum in cell proliferation, differentiation, apoptosis and signaling pathways, genes that up and down expression, were used for prevention, diagnosis, treatment and reduction of metastasis in patients with colorectal cancer.

Keywords: Colorectal Cancer, Fusobacterium nucleatum, prognostic biomarkers, therapeutic targets

ICB11-1401-12-1141

Phylogenetic analysis of amino acids in baculovirus enhancins

Maryam Rashki[1*], Mojtaba Mortazavi[2]

[1*]Graduate university of advanced technology.[2]Graduate university of advanced technology.

*Corresponding author(s). E-mail(s): ma_rashkigh@yahoo.com;

Abstract

Baculoviridae is a family of insect viruses and is grouped into two main groups or genera: Nucleopolyhedrovirus and Granulovirus. The amino acid analysis in baculovirus enhancin is very important and applicable because the engineered recombinant is able to express the insect selective toxin as bioinsecticide and this recombinant virus elicited a response significantly faster than the common progenitor wild-type virus . All 67 sequences obtained from GenBank have been aligned using MAFFT . To detect motifs, the multiple sequence alignment was edited using Mesquite program and was displayed using Weblogo 3 and Multiple Sequence Alignment-Viewer. The conserved sequence was declared using Clustal program. There were two different parts where the variation of the second part was more than the first one related to their functions. The conserved HEXXH sequence was found. The signature pattern is adequate to group a protein into the metalloprotease superfamily. Most baculovirus enhancins have this conserved metalloprotease zinc-binding domain and in the metalloproteases, the glutamic acid residue within the HEXXH sequence is the catalytic base that polarizes a water molecule included in the nucleophilic attack of the peptide bond to be cleaved . To compare the sequences BLASTP was used and the phylogenetic tree based on the Maximum likelihood method using IQ-tree program was created. The validation of trees has been conducted using ultra bootstrap method implemented in IQ-tree program. According to the phylogenetic tree of Nucleopolyhedrovirus and Granulovirus enhancins, three main clades were showen. One group included Lymantria dispar multiple nucleopolyhedrovirus_Viral Enhancing Factor1 and Lymantria dispar multiple nucleopolyhedrovirus_VEF2 sequences. All Granuloviruses, with the exception of Agrotis segetum granuloviruses, contained Granulovirus_Viral Enhancing Factor1, Granulovirus_Viral Enhancing Factor2, Granulovirus_Viral Enhancing Factor 3 and Granulovirus _Viral Enhancing Factor 4 sequences. The last group comprised five subgroups including Nucleopolyhedrovirus_Viral Enhancing Factor 1, Nucleopolyhedrovirus_Viral Enhancing Factor2, Nucleopolyhedrovirus_Viral Enhancing Factor3, Mamestra brassicae multiple nucleopolyhedrovirus to Mamestra configurata nucleopolyhedrovirus-A and Mythimna unipuncta nucleopolyhedrovirus. The high level of heterogeneity exhibited by the baculovirus enhancins may propose that these genes appeared in viral genomes from independent sources.

Keywords: Amino acid, Baculoviridae, Motif, Viral enhancing factor

ICB11-1401-12-1142

In silico analysis of baculovirus enhancin genes

Maryam Rashki[1*], Mojtaba Mortazavi[2]

[1*]Graduate university of advanced technology.[2]Graduate university of advanced technology.

*Corresponding author(s). E-mail(s): ma_rashkigh@yahoo.com;

Abstract

Arc (Activity regulated cytoskeleton associated protein) is a key regulator of synaptic plasticity. Its gene was emerged from the retrotransposon Gag capsid genes . In the nervous system, it is responsible for intercellular transfer of mRNA to the cytoplasm of neurons through endocytosis via its virion-like capsid. In current work, a comperehensive sequence and structural analysis was performed on the Human Arc protein and its homologoues counterparts found in other organisms. For sequence analysis, we used Bioinformatics tools including the BLAST , ClustalW , ESpript and finally a representative sequence of programs in the Phylip package to deduce the evolutionary history of the Arc proteins. The structural features of the Arc protein have been studied using a combination of Bioinformatics and Computational tools. Currently, there is no structural report on the whole structure of protein, and the reported structural data in the PDB database is limited to specific fragments of the original sequence from various organisms (UniProt codes: Q7LC44, Q63053, Q7K1U0, and Q7JV70). We used these fragments as templates to build a compact structure for the human variant Arc protein. Due to the structural data limitation, the final structure was built and validate with a combinatorial approach using respective online servers and the MODELLER program (Ver. 10.4). We used the capabilities of the UCSF Chimera program to visualize and structural analysis on the selected models.

Keywords: Baculoviridae, Codon, Enhancin, Phylogenetic relationship

ICB11-1401-12-1143

Identification of Biomarkers Involved in Adrenocortical Carcinoma by Systems Biology Approach

Shahrzad Ameri[1], Fatemeh Lotfi[2*]

[1]Shiraz university. [2*]Shiraz university.

*Corresponding author(s). E-mail(s): fatemeh.lotfi7820@gmail.com;

Abstract

Adrenocortical Carcinoma (Cancer of the adrenal cortex or ACC) is believed to be rare cancer originating in the outer layer of the adrenal gland (steroid hormone-producing tissue). However, the overall 5-year survival rate is about 50% in these patients with an incidence of one to two per million populations annually. The goal of this research was to screen and find key genes in the development of ACC based on Systems biology approaches. From the GEO database, which is considered a database for gene expression profiling and RNA methylation profiling, the GEO set GSE19776 (platform: GPL570) was downloaded. This Microarrays data contain the gene expression differences between the tumor and normal samples, which contains 44 tumor samples and 4 normal samples, totaling 48 samples. In order to screen the DEGs between tumorous tissue and normal tissue the web tool GEO2R was used. Variables with adjusted P values <0.05 and $|\log FC| \geq 2$ were regarded as statistically significant. The STRING online database is used to obtain the protein-protein interaction network information of DEGs. Then, the results were visualized by the Cytoscape software. Having produced The protein-protein interaction network (PPI Network) diagram, the Betweenness analysis method was used to calculate the top 10 nodes. Survival analysis by GEPIA (the Gene Expression Profiling Interactive Analysis web server) demonstrates that the upregulated genes CRK and COL1A2 were involved in the poor prognosis of Adrenocortical Carcinoma. Moreover, the expression of these genes was dramatically upregulated in Carcinoma tissues compared to normal tissues. CRK and COL1A2 might be new biomarkers for prognosis, diagnosis and subsequent therapy of Adrenocortical Carcinoma.

Keywords: Adrenocortical Carcinoma, Adrenal Cortex, Gene Expression Profiling, Protein Interaction Mapping, Protein-Protein Interaction Network, Systems biology

ICB11-1401-12-1144

Investigate effects of plant secondary metabolites Compared with standard drugs on SARS-CoV-2 by Molecular Docking and Molecular Dynamics

Afifeh Hajipour[1], Masoud Tohidfar[2*]

[1]Shahid beheshti university. [2*]Shahid beheshti university.

*Corresponding author(s). E-mail(s): m_tohidfar@sbu.ac.ir;

Abstract

Coronaviruses are viruses that often cause acute complications in the respiratory system with symptoms similar to a cold. Some of them, such as SARS and MERS in the last few years and Covid-19 since January 2019, have caused thousands of deaths and many pandemics. Medicinal plants are used in traditional medicine to treat various diseases, including infectious diseases. The purpose of this research is to determine the effective antiviral compounds of medicinal the genes involved in their synthesis against the coronavirus. In this study, the secondary metabolites of the plants were evaluated on the human species and the important and key proteins were determined by the meta-analysis method. In the database, interaction of proteins with respect to each other and hub genes were obtained from Cytoscape software. To study the ontology of genes from the Enrichr database and then using HeatMap, gene expression was compared. As a result of meta-analysis, 14 genes related to blood coagulation factors and complement system were determined. The chemical compounds of medicinal plants (Glycyrrhiza glabra, Echinacea angustifolia, Panax ginseng, Chicory coffee)with antiviral and antimicrobial effects were extracted from the Pubchem database for docking and then checked by the HeatMap database and the results obtained from the Kegg Pathway server were evaluated and two factors, coagulation factor X (F10), Coagulation factor II thrombin (F2) as reactivity of human cells in face with virus and Spike-protein and Main-protease as coronavirus receptor were extracted from PDB database. Receptor (1A2C, 6lu7, 4Y79, 6VXX)-ligand docking (secondary combinations) was confirmed by coach-d protein ligand server. At the end, their validation was done by performing molecular dynamics by Gromax2022 software. According to these results, by reducing the expression of thrombin factor and x factor, preventing bloodclotting and inhibiting the spike protein and the main protease of the corona virus, it was found that the virus multiplication was prevented by using the antiviral extract of these plants.

Keywords: SARS-CoV 2, Molecular docking, Molecular Dynamics, Drug design, Medicinal plants

ICB11-1401-12-1145

The effect of L813N|A814R double mutations on Taq polymerase structure

Seddigheh Borhani[1], Seyed shahriar Arab[2*]

[1]Tarbiat modares. [2*]Tarbiat modares.

*Corresponding author(s). E-mail(s): sh.arab@modares.ac.ir;

Abstract

Thermus aquaticus DNA polymerase I (Taq polymerase), similar to family A DNA polymerase other members, possess polymerization and 5'-3' exonuclease activities. This enzyme plays a role in DNA repair and Okazaki fragments ejection during DNA replication and the lack of 3'-5' exonuclease activity causes its high error rate. Taq polymerases, due to their thermostability properties, have a wide application in biotechnology techniques, namely PCR. There are several studies in order to create improved enzymes, including hot start enzymes, chimeric enzymes with 3'-5' exonuclease activity and additional thermostability. Here, we studied the effect of L813N| A814R double mutations on Taq polymerase structural stability using MD simulation methods. The comparison of three iterated simulations on wild-type and mutant enzymes demonstrated that the enzymes did not show any significant difference in RMSF, SASA, and protein-protein hydrogen bonds analysis. However, the number of protein-solvent hydrogen bonds increased in mutant ones (with seven more bonds). Substitution of aliphatic residues (L and A) with polar ones (N and R) enhances the possibility of hydrogen bond formation, which improves protein structure stability. In our case, due to the presence of 813 and 814 residues on the protein surface, this increase in the bonds was observed between protein and solvent. Therefore, it is proposed that L813N| A814R double mutations are able to raise Taq polymerase structural stability.

Keywords: Taq Polymerase, Thermus aquaticus, DNA polymerase, structural stability, MD simulation

ICB11-1401-12-1146

Identification of potential therapeutic microRNAs for gastric cancer based on bioinformatics analysis

Ali Khezrian[1], Tahereh Latifi pakdehi[2], Saeid Afshar[3*]

[1]Hamadan university of medical sciences. [2]Hamadan university of medical sciences. [3*]Hamadan university of medical sciences.

*Corresponding author(s). E-mail(s): s.afshar@umsha.ac.ir;

Abstract

Gastric cancer (GC) is the fourth leading cause of cancer-related deaths in 2021. Despite advances in detection and treatment, gastric cancer has only a 20% 5-year survival rate. Therefore, there is a need to approach the exact molecular mechanisms of carcinogenesis by using molecular biology techniques and develop therapeutic targets for this disease in the future. To conduct this study, we referred to the GEO database and selected the GSE54129 dataset on the Affymetrix platform with the accession ID GPL570, which included 111 cancer samples and 21 healthy samples (tissue surrounding the cancerous area). The analysis was performed with Log FC $\geq$ 1.5 criteria, Log FC. $\leq$-1.5, adj.P.Val $<$0.01 to establish DEGs (Differentially Expressed Genes). 1835 genes were selected and gene ontology and KEGG pathway enrichment analysis were performed using DAVID online tool (https://david.ncifcrf.gov/tools.jsp). Then STRING and Cytoscape software were used to generate protein-protein network and find hub genes. The top 10 genes according to the parameters: Degree, Eccentricity, MNC and EPC, were identified in the network and the hub of genes was determined by sharing between them. Further, referring again to the GEO database, GSE23739 dataset was selected to analyze miRNA expression in non-cancerous and cancerous stomach tissues. and DEMs (Differentially Expressed miRNA) were separated with the criteria of Log FC$\geq$1, Log FC$\leq$-1, adj.P.Val $<$0.05. By removing duplicates, 93 miRNAs were checked in the mirwalk database and their target genes were extracted. The list of these genes was compared with the DEGs obtained from GSE54129 dataset and 1255 common genes were identified. Finally, 2 hub genes (CTNNB1 and TP53) were obtained, which were common in the list of DEGs and DEMs, which are possible key genes involved in gastric cancer. Then miRNAs related to these genes (hsa-miR-103a-3p, hsa-miR-1224-5p respectively) were identified, which promises their potential in the treatment of this disease.

Keywords: Gastric cancer, therapeutic targets, GEO database, hub genes, miRNAs

ICB11-1401-12-1147

Preeclampsia: Ranking the genes and associated biological functions

Maedeh Rahrovi[1], Omid Jazayeri[2*], Mohammad Karimian[3], Ali Taravati[4]

[1]Mazandaran. [2*]Mazandaran.[3]Mazandaran. [4]Mazandaran.

*Corresponding author(s). E-mail(s): o.jazayeri@umz.ac.ir;

Abstract

Background and aims: Preeclampsia is a gestation-specific disorder and it is one of the most important causes of fetal and maternal morbidity and mortality in the whole world. The purpose of this research was to compile all the genes related to Preeclampsia and determine the top ten genes as well as determine the most important biological functions related to this disease. Methods:
A comprehensive review was performed using the PubMed database and data were taken from research articles published between 1995 and January 2022. To evaluate the association between different genes and preeclampsia, we selected the articles with significant association (p-value < 0.05) with preeclampsia. The GeneMANIA Cytoscape and Cytoscape's plugin CytoHubba apps were used to investigate these associations, visualize the gene network and discover key nodes (genes) in this biological network. Results:
After reviewing the 796 articles, 356 articles were included in our study. Our results showed that 207 genes are associated with the risk of Preeclampsia. Based on Maximal Clique Centrality, the top ten genes that have significant associations with Preeclampsia are AGTR2, CCR5, CXCR2, DRD4, GNB3, AGTR1, CXCR4, AGT, CXCL12, and CXCL8. In addition, nitric-oxide synthase biosynthetic process and regulation of nitric-oxide synthase biosynthetic process are the two most important biological functions with a 48.96 enrichment score related to Preeclampsia. Conclusions: This study provides novel insights to the pathomechanism of Preeclampsia. The outcomes of this study revealed the top ten genes that have significant associations with Preeclampsia. Additionally, the biological functions related to Preeclampsia were ranked based on their enrichment scores.

Keywords: association, bioinformatic, biological network, gene network, preeclapsia

ICB11-1401-12-1148

Enhancement of In-hospital Mortality Prediction in Emergency Department using Ensemble Machine Learning Models

Zahra Rahmatinejad[1*], Toktam Dehghani[2], Fatemeh Rahmatinejad[3], Hamidreza Reihani[4], Ali Pourmand[5], Ameen Abu-hanna[6], Saeid Eslami[7]

[1*]Mashhad university of medical sciences.[2]Mashhad university of medical sciences. [3]Mashhad university of medical sciences. [4]Mashhad university of medical sciences. [5]George washington university. [6]Department of medical informatics, amsterdam umc - location amc, university of amsterdam, the netherlands. [7]Mashhad university of medical sciences.

*Corresponding author(s). E-mail(s): Zahra.rahmatinejad@gmail.com;

Abstract

Objectives:
Early prediction of severity level in critically ill patients and their outcome, definitely can be affected on prioritizing patients and declining in-hospital mortality in emergency department. Several classical Scoring systems have been proposed for indication of severity of patient illness, such as logistic regression models. However, as an alternative approach, one of the main streams in machine learning methods, known as ensemble models, are proposed in the current study to compare their modern models with the old ones. Hence, the aim of this study is evaluation and comparison of the traditional model's logistic regression (LR) and modern ones (Bagging, AdaBoost, Random Forest (RF), and Extreme Gradient Boosting (XGB), and detection of the model with the best predictive performance for prediction of in-hospital mortality
Methods:
An observational single-center study was conducted in the Emergency Department (ED) of Imam Reza Hospital from March 2016 to March 2017 which is located in the northeast of Iran. Adult patients with one to three level of ESI acuity was defined as meeting criteria for this study. The training and validation visits from the ED were randomly divided into 80% vs 20%. After training the models using 10-fold cross-validation on the training set, their predictive performance was then evaluated.
Results:
the cohort consisted of 2,025 unique patients admitted to the ED of the hospital. There were about 19% of hospital deaths. Of the 1,476 patients in the training group, 274 (18.6%) died during hospitalization, and of the 728 patients in the validation group, 152 (20.8%) died during hospitalization. The AUCs with 95% confidence intervals (CIs) for the different models were as follows: the RF, 0.812 (95% CI = 0.742 to 0.830); XGB, 0.798 (95% CI = 0.756 to 0.840); Bagging, 0.796 (95% CI = 0.753 to 0.837); Adaboost, 0.791(95% CI = 0.745 to 0.837); and the LR, 0.786(95% CI = 0.742 to 0.837). Among all ML-based prediction models the AUC of the RF was statistically different from the LR ($p \leq 0.0275$). The AUPRC was 0.577(0.541-0.612) for the RF, 0.557(0.520-0.593) for the XGB, 0.605(0.569-0.640) for the Bagging, 0.603(0.567-0.638) for the Adaboost, and 0.592(0.556-0.627) for the LR. All models had inadequate calibration. The most accurate models belonged to the RF and XGB with lowest BS. Conclusions: the RF can be used as a screening tool to identify patients at risk of mortality. This tool could improve the management of hospital resources and patient-throughput planning, thus delivering more effective care to hospitalized patients.

Keywords: Ensemble machine learning, In-hospital mortality, Emergency department, Logistic regression, Bagging, AdaBoost, Random Forest, and Extreme Gradient Boosting, Unbalanced data, Feature selection

ICB11-1401-12-1149

Investigation effect of Pembrolizumab antibody combined with Napabucasin compared with Pembrolizumab on the gene expression profiling in colorectal cancer based whole transcriptome RNA-Seq analysis

Seyed taleb Houseini[1], Arash Sattari[2], Roya Bishehkolaei[3], Farkhondeh Nemati[4*]

[1]Islamic azad university. [2]Islamic azad university. [3]Islamic azad university. [4*]Islamic azad university.

*Corresponding author(s). E-mail(s): farkhondehnemati@gmail.com;

Abstract

Introduction:
Colorectal cancer (CRC) is a leading cause of cancer mortality worldwide, being the second- and third-most frequently occurring cancer in women and men, respectively . Napabucasin has the potential to treat certain types of cancer by inhibiting the self-renewal and survival of cancer stem cells, as well as inducing apoptosis in both cancer stem cells and heterogeneous cancer cells. Pembrolizumab is a humanized monoclonal antibody (mAb) directed against the programmed cell death protein 1 (PD-1) that has been previously approved for the treatment of a variety of solid tumors.
Material and Method:
DRP008249 dataset was downloaded from NCBI>SRA, the sra file was converted to two fastq paired files with sratoolkit.2.11.3-ubuntu64 software under Linux. Quality control of fastq files was performed with FASTQC software, and low quality reads were trimmed with Trimmomatic software and we mapped the trimmed file to hg38_genome using Hisat2 software and a sam file was obtained. In the next step, we converted the sam file to counts files using hg38.ncbi.refseq.gtf and finally, DEGs were identified using the DESeq2 package in R Studio software by threshold (FDR:0.05 and log2foldchange>1).
Results:
Based on the results, the effect of Pembrolizumab antibody combined with Napabucasin in comparison with Pembrolizumab, cause of decreased the expression of PDE10A, MALRD1, FGFR2, LOC105378116 and GABRB3 genes and increased the expression of LOC105371308, ABO, ANKRD18A, LINC00444 and E4F1 genes in tumor tissues of patients.
Conclusion:
Due to the important role of Pembrolizumab and Napabucasin antibodies in the process of apoptosis and treatment of a variety of solid tumors, identified genes can be used for prevention, diagnosis and as a therapeutic marker for patients with colorectal cancer.

Keywords: Colorectal cancer, Pembrolizumab, Napabucasin, stem cells

ICB11-1401-12-1150

In-silico docking and virtual screening of natural products as possible BACE1 inhibitors: a new hope for treatment of Alzheimer's disease?

Setareh Azarkar[1], Amir masoud Jafari-nozad[2], Sahar Asadolahizoj[3], Amirsajad Jafari[4*], Saman Yousefi[5]

[1]Birjand university of medical sciences. [2]Birjand university of medical sciences. [3]Zabol university. [4*]Shiraz university.[5]Islamic azad university of shahrekord.

*Corresponding author(s). E-mail(s): a.sajad.jafari@gmail.com;

Abstract

The most common cause of dementia and the fifth most prevalent cause of mortality in people over 65 is Alzheimer's disease (Wung, 2020). A different pharmacological target for treating AD is needed due to the ongoing failure of present therapies. BACE1 is a widely recognized prospective therapeutic target for Alzheimer's disease because it catalyzes the rate-determining step in the production of Aβ peptides (Kumar et al., 2020). COCONUT database (Collection of Open Natural Products database) was screened for creating a library of natural products with the following filters: 1. molecular weight between 400-600 2. Lipinski's role of five. One hundred best hits were downloaded as a library. We used the protein 3D structure of BACE1 (PDB ID: 7MYI; X-Ray; Chain A/B; 1.25 Å) for docking. All the ligands and the 3D structure of the BACE1 were then prepared using the Schrödinger suite (Maestro 11.5) Ligprep module and ProPrep wizard, respectively. Ligand docking in the module GLIDE of the Schrödinger suite was used for the evaluation of the interaction between ligands and chain A of the target protein (opls2005 forcefield). Best ligands according to glide E-model and docking score were evaluated for pharmacokinetic properties predictions by Swiss ADMET, pkCSM, and ADMETlab 2.0. Molecular docking analyses of our natural products library found promising ligands that had excellent protein-ligand interactions with BACE1 and reasonable docking scores (CNP0278652, CNP0014230, CNP0385754, CNP0194616, CNP0178972, CNP0251096, CNP0393956 and CNP0366016). Some of these ligands displayed acceptable pharmacokinetic and physicochemical characteristics, such as good absorption, distribution, metabolism, and excretion (ADMET). BACE1 appears to be a promising target for the development of novel medications. Future in-vivo and in-vitro research is necessary to provide additional information about this target, nevertheless.

Keywords: Alzheimer, In-Silico, Drug design, Natural Products

ICB11-1401-12-1151

Molecular screening and docking analysis of natural GSK3B inhibitors: A promising target for Alzheimer's disease

Setareh Azarkar[1], Amir masoud Jafari-nozad[2], Sahar Asadolahizoj[3], Amirsajad Jafari[4*], Saman Yousefi[5]

[1]Birjand university of medical sciences. [2]Birjand university of medical sciences. [3]Zabol university. [4*]Shiraz university.[5]Islamic azad university of shahrekord.

*Corresponding author(s). E-mail(s): a.sajad.jafari@gmail.com;

Abstract

Alzheimer's disease (AD) is a brain disorder that gradually impairs memory, reasoning, and, eventually, the capacity to preform even the most basic tasks (Knopman et al., 2021). A different pharmacological target for treating AD is needed due to the ongoing failure of present therapies. GSK3B protein is associated with a high risk of neurodegenerative conditions and is linked with the pathways of Alzheimer's disease (Wu et al., 2022). COCONUT database (Collection of Open Natural Products database) was screened for creating a library of natural products with the following filters: 1. molecular weight between 400-600 2. Lipinski's role of five. One hundred best hits were downloaded as a library. We used the protein 3D structure of GSK3B (PDB ID: 4AFJ; X-Ray; Chain A/B; 1.98 Å) for docking. All the ligands and the 3D structure of the GSK3B were then prepared using the Schrödinger suite Ligprep module and ProPrep wizard, respectively. Ligand docking in the module GLIDE of the Schrödinger suite was used for the evaluation of the interaction between ligands and chain A of the target protein (opls2005 forcefield). Best ligands according to glide E-model and docking score were evaluated for pharmacokinetic properties predictions by Swiss ADMET, pkCSM, and ADMETlab 2.0. Molecular docking analyses of our natural products library found promising ligands that had excellent protein-ligand interactions with GSK3B and reasonable docking scores (CNP0200330, CNP0384098, CNP0400076, CNP0461489, CNP0377796, CNP0048544, CNP0229054 and CNP0298746). Some of these ligands displayed acceptable pharmacokinetic and physicochemical characteristics, such as good absorption, distribution, metabolism, and excretion (ADME). GSK3B 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: Alzheimer, In-silico, Drug design, Natural Products

ICB11-1401-12-1152

Identification of hub genes and key signaling pathways of ascites-derived ovarian cancer stem cells from different origins and associated with clinicopathological features

Muhammad reza Tayebi zadeh[1], Sadra Daneshvari[2], Zeinab Dehghani-ghobadi[3], Hesam Montazeri[4], Ghamartaj Hossein[5*]

[1]Developmental biology laboratory, department of animal biology, school of biology, college of science, university of tehran, tehran, iran. [2]Developmental biology laboratory, department of animal biology, school of biology, college of science, university of tehran, tehran, iran. [3]Developmental biology laboratory, department of animal biology, school of biology, college of science, university of tehran, tehran, iran. [4]Department of bioinformatics, institute of biochemistry and biophysics, university of tehran, tehran, iran. [5*]Developmental biology laboratory, department of animal biology, school of biology, college of science, university of tehran, tehran, iran.

*Corresponding author(s). E-mail(s): ghossein@ut.ac.ir;

Abstract

Ovarian cancer stem cells (OCSCs) are responsible for many epithelial ovarian cancer-related-clinical features, including tumor initiation and progression, metastasis, and recurrence. Ascites-derived OCSC populations could be originated from the fusion of cancerous cells with immune cells or the fallopian tube fimbriae. However, the differently expressed genes between OCSC populations from different origins are not well understood. Here we sought to identify the common hub genes and key signaling pathways related to ascites-derived OCSCs from different origins. At first, comprehensive transcriptomic analysis from two datasets GSE90125 (ascites-derived OCSC populations of fimbriae origin) and GSE75036 (ascites-derived OCSC population of fusion origin) was performed to find hub genes separately and in the next step, the common hub genes and signaling pathways were determined. The up-regulated genes from GSE90125 were mainly related to the migration and epithelial to mesenchymal transition and TGFβ, PI3K-AKT, and Wnt signaling pathways. Whereas up-regulated genes from GSE75036 were mostly involved in cytokine production and inflammatory response and to NFκB and MAPK signaling pathways. We found ten common hub genes with amplification mutations associated with ascites-derived OCSC populations from different origins. Among the signaling pathways, the Wnt and TGFβ pathways are the most important pathways associated with these common hub genes. Clinicopathological evaluation of common hub genes showed that CD44, CD68, TGFBI, and ZEB2 were significantly related to the lymphatic and venous invasion. While HIF1A and VCAN were significantly related to the venous invasion. It is interesting to note that PPARG and VCAN were significantly related to new neoplasm event type and primary outcome success, respectively. In conclusion, we suggest that deciphering and targeting common hub genes between OCSCs from different origins could have an impact on epithelial ovarian cancer–related clinical features.

Keywords: Ovarian cancer stem cells, ascites, fimbriae, immune cells, Wnt signaling, TGFβ signaling

ICB11-1401-12-1153

Study of the Gene Network effective on Bovine Mastitis under Microarray data

Zahra Roudbari[1*], Saideh Eskandarinasab[2]

[1*]University of jiroft.[2]University of zabol.

*Corresponding author(s). E-mail(s): roudbari.zahra@ujiroft.ac.ir;

Abstract

Mastitis is one of the most common production diseases and economically the biggest and most important disease in the dairy industry. it causes a lot of economic loss to the dairy industry and the reduction of milk production is a small part of its effects. The immune response of the mammary gland to the presence of microorganisms causes mastitis . The development of microarray technology provides the possibility of a comprehensive investigation of known genes in certain tissues such as the mammary gland in order to increase our knowledge of the defense mechanisms involved in the mammary gland, and identify gene networks and gene expression changes. . The gene expression data was implemented with the accession number GSE50685. The criterion of Differential gene expression measurement was considered at the level of P-value $<$ 05 and LogFC $>$ 1 or LogFC$<$-1. STRING database and Cytoscape software were used to establish a gene network related to mastitis disease. In order to identify the genes with the greatest regulation effect, the genes with connected degrees above 7 were selected and studied. DAVID software was used to identify biological pathways. In this study, we found that FOXO1, ITGB1, HSP90AA1, HSP90AA1, IL2, IL6, and CXCL10 genes are the most linked genes. The enrichment results of the KEGG pathway indicated that highly connected genes were related to Mastitis. The first prominent gene is the FOXO1 gene, which plays a role in mammary gland immune and inflammatory regulation through the inhibition of EGFR-AKT signaling. The second gene is ITGB1 which researchers reported that TGF-β1 gene promoted the expression of Fn and ITGB1 genes on the surface of bovine mammary epithelial cells and contributed to mammary gland infection in vitro and in vivo. The results of this study imply that Fn and ITGB1 may be useful therapeutic targets for the treatment of mastitis in dairy cows. Another gene is HSP90AA1, which this gene promotes autophagy and inhibits apoptosis through PI3K/Akt/mTOR pathway and JNK/P38 pathway. IL2 gene is an anti-inflammatory cytokine that can inhibit natural killer cells. IL6 gene has also been reported to play a role in regulating mastitis inflammation by suppressing NF-kB signaling and CXCL10 genes have been introduced in previous studies as the "main gene" involved in the defense of mammary glands, and in another study has been proven, the relationship between the polymorphism of this gene and resistance to mastitis.

Keywords: biological pathways, Gene network, Mastitis, Dairy cow

ICB11-1401-12-1154

Evaluation of Protein Clusters in Wool Fiber Production Using Interaction Map Analysis

Zahra Roudbari[1*], Saideh Eskandarinasab[2]

[1*]University of jiroft.[2]University of zabol.

*Corresponding author(s). E-mail(s): roudbari.zahra@ujiroft.ac.ir;

Abstract

Wool is one of the most important sheep products and a source of income from sheep breeding, which has a high economic value. The growth and development of sheep's wool are controlled by hair follicles, which are small organs attached to the skin and have complex morphology and structure and periodic growth. Differentiation of hair follicles is regulated by a variety of signaling pathways including bone morphogenetic protein (BMP), transforming growth factor beta (TGF-β), and Wnt signaling pathways . However, our knowledge about the relevant molecular and cellular mechanisms is limited. Therefore, this study was conducted with the aim of investigating molecular changes at the level of protein interactions by constructing and analyzing the protein interaction network of wool production. The data analyzed in this study were extracted from the Arrayexpress database with accession number GSE37400. Cytoscape software and String source were used to construct the interaction network of proteins. Analysis of protein clusters was done by the Sytohubba plugin. The identification of biological pathways related to network complexes was investigated by David software . The results of our study showed that the genes forming protein complexes play a role in the process of the hair cycle, development of the skin epidermis, proliferation, and differentiation of epithelial cells. Related genes include 17 genes, these genes have a special relationship with the development of hair follicles, skin, epidermis, and epithelium. The results of biological pathway analysis showed that the above genes are involved in TGF-β, Wnt, Notch, MAPK, and VEGF signaling pathways, and these pathways are related to the development of skin or hair follicles. It seems that the introduced biomarker panel can include 17 identified genes that are effective in better understanding the wool fiber production process, which is one of the most important breeding goals in sheep breeding.

Keywords: Network analysis, Wool Fiber, Fiber Production, Sheep

ICB11-1401-12-1155

Molecular docking and ADME studies of natural compounds against bitter taste receptors(TAS2Rs)

Golnaz Golchin[1], Ainaz Movahed[2], Samad Nejad ebrahimi[3*]

[1]Shahid beheshti university. [2]Shahid beheshti university. [3*]Shahid beheshti university.

*Corresponding author(s). E-mail(s): s_ebrahimi@sbu.ac.ir;

Abstract

A subfamily of 25 G protein-coupled receptors called the human bitter taste receptors (TAS2Rs) mediates the perception of the bitter taste. The three TAS2Rs (TAS2R14, TAS2R39, and TAS2R46) stand out from the rest of the TAS2Rs. Since they are the most widely tuned bitter ness receptors, recognize a wide range of chemical, natural, and miscellaneous agonists and antagonists in the micromolar range with extremely low potency. In addition, the receptors are expressed in oral, and several extra-oral tissues, and they are suggested for having salient physiological roles associated with innate immune responses, cancer, and male fertility. In the present study, computational techniques (molecular docking, in silico ADMET, and predication of drug-likeness) were used to perform virtual screening on 452 natural ligands selected from the bitterDB database. Studies on receptor-ligand binding were carried out by Schrodinger drug discovery Suite. Subsequently, the Glide docking program and extra precision (XP) were applied. The best ligands based on docking score are myricetin(-11.50Kcal/mol), dihydrofisetin (-9.33Kcal/mol), and artesin(-9.19Kcal/mol) for TAS2R14, TAS2R39, and TAS2R46 receptors, respectively. The docking results depict the promising potent natural product which can interact with the TAS2Rs receptor for therapeutic approaches. Furthermore, higher potency ligands are needed to investigate the mentioned three receptors function and to modulate them for future clinical applications, such as cancer treatment or taste prediction.

Keywords: Bitter Taste Receptor, Molecular Docking, Natural Compound

ICB11-1401-12-1156

Cell Spatial Distribution Patterns of Human Distal Lung

Mahdi Pursalim[1], Kaveh Kavousi[2*]

[1]Department of Bioinformatics, Kish International Campus University of Tehran, Kish, Iran.
[2*]Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Iran.

*Corresponding author(s). E-mail(s): kkavousi@ut.ac.ir;

Abstract

Introduction
Despite the resolution of scRNA sequencing, there is no information about the location of cells within a tissue. Now by stRNA sequencing, studying cell type locations is possible to research tumor microenvironment, tumor development, invasion, and metastasis that is so functional to understand tumor progression (Wei et al., 2022). The aim of this study is to investigate the tissue structure and cell-type locations on it in Human distal lung (Kadur Lakshminarasimha Murthy et al., 2022) through bioinformatics analysis which can be effective to diagnose the disease in earlier stages.
Methods
The information related to normal adult lung tissue was obtained by GEO database with GSE178361. Spatial transcriptomics has been done using the protocol of 10x Genomics Visium Spatial Gene Expression assay (10x Genomics) for spatial RNA-seq (3' enrichment). Data analysis consisting of quality control, preprocessing, dimensionality reduction and clustering top expressed genes for normalization was performed by python software. 11 clusters were identified in the tissue surface. Afterwards in order to visualize cell-type distribution and spatial organization in the human distal lung tissue, the gene expression patterns of individual cells were specified.
Results and discussion
We visualized total counts per spot and number of genes by counts in UMAP space in human distal lung that are projected onto the spatial coordinates to gain insights into tissue structure and intercellular communication. Then, we removed the lncRNA gene MALAT1, mitochondrial genes that are among the top expressed, and hemoglobin genes (blood contamination) if they exist. According to our findings, a wide area of studied tissue is occupied by LTF, FDCSP, and SERPINA3. Also spots belonging to other clusters seem always to be surrounded by spots belonging to mentioned clusters. "

Keywords: Human distal lung, Cancer, Spatial coordinates, Spatial transcriptomics (ST)

ICB11-1401-12-1157

Unraveling the Complexity of rhizosphere microbiome affecting the quality of Crocus sativus

Masoud Ghayoumi[1*], Abbasali Emamjomeh[2], Kaveh Kavousi[3], Ali Najafi[4]

[1*]University of Zabol.[2]University of Zabol. [3]University of Tehran. [4]Baqiyatallah university of medical sciences.

*Corresponding author(s). E-mail(s): masoud.gh27@gmail.com;

Abstract

Saffron is an aromatic plant belonging to the Iridaceae family. It is produced from the dried stigmas of Crocus sativus. Microorganisms in the soil are the main factors affecting the production and quality of Crocus sativus. In this study, we present a study of the microbial relationship between Crocus sativus rhizosphere and quality parameters of Crocus sativus (crocin, picrocrocin, and safranal) using a metagenomic approach. This research was conducted in Ghayen city, the center of quality Crocus sativus production worldwide. The interesting point in this region is that despite the similar climatic conditions, the quality of Crocus sativus in the Shahik region is higher than in the other areas (Zobar, Shahabi, and Jafarabad). This can be attributed to different microbiomes in the root zone or rhizosphere of the soil. The purpose of this study was to investigate and compare the soil microbiome of Crocus sativus rhizosphere in the Shahik area of Ghayen city with other areas of this city and, as a result, to identify the microbiome affecting the quality of Crocus sativus. In this method, rhizosphere soil of different regions in Ghayen city was first sampled during the Crocus sativus harvest time. In the next step, soil DNA was extracted, and the 16srRNA gene sequence of each sample was identified based on the RNA-Seq method; finally, the microbiome of the samples was analyzed. The results showed that beneficial bacteria such as Kosakonia, Enterobacter, Salmonella, Plesiomonas, Pasteurella, Lactococcus, and Mitsuaria in Shahik farm are more than in other farms that these bacteria have a significant impact on increasing the quality of Crocus sativus.

Keywords: 16srRNA, Bacteria, Metagenomics, Crocus sativus, Miseq, Next Generation Sequencing