Indian Conference on Bioinformatics 2023 – Inbix'23 THEME : Application of Bioinformatics in Healthcare BOOK OF ABSTRACTS Organized by SCHOOL OF BIO SCIENCES AND TECHNOLOGY In asso…

Conference Paper · December 2023

CITATIONS
0

READS
778

10 authors, including:

Sajitha Lulu
VIT University
56 PUBLICATIONS **323** CITATIONS
SEE PROFILE

Vino Sundararajan
VIT University
34 PUBLICATIONS **219** CITATIONS
SEE PROFILE

Premkumar Krish
VIT University
3 PUBLICATIONS **1** CITATION
SEE PROFILE

Devi Soorya Narayana S
VIT University
4 PUBLICATIONS **20** CITATIONS
SEE PROFILE

VIT

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

Inbix'23

Indian Conference on Bioinformatics 2023 -

Inbix'23

THEME : Application of Bioinformatics in Healthcare

November 24-26, 2023

BOOK OF

ABSTRACTS

Organized by

SCHOOL OF BIO SCIENCES AND TECHNOLOGY

In association with

b! Bioclues Organization

Indian Conference on Bioinformatics 2023 - Inbix'23

THEME: Application of Bioinformatics in Healthcare

November 24-26, 2023

Book of Abstracts

Organized by

SCHOOL OF BIO SCIENCES AND TECHNOLOGY

In association with

ORGANIZING COMMITTEE

Chief Patron	Dr. G. Viswanathan, Founder & Chancellor, VIT
Patrons	Mr. Sankar Viswanathan, Vice President, VIT
	Dr. Sekar Viswanathan, Vice President, VIT
	Dr. G. V. Selvam, Vice President, VIT
Co-Patrons	Dr. V. S. Kanchana Bhaaskaran, Vice-Chancellor-in-Charge, VIT
	Dr. Partha Sharathi Mallick, Pro-Vice Chancellor, VIT
	Dr. T. Jayabarathi, Registrar, VIT
Organising Chairs	Dr. R. Siva, Dean, SBST, VIT
	Dr. Jayaraman K Valadi, Distinguished Professor, FLAME University, Pune and Adviser-In-Chief, Bioclues.Org
Co- Chairs	Dr. Jayanthi Abraham, Associate Dean, SBST, VIT
	Dr. Gyaneshwar Chaubey, President, Bioclues.Org
Convenor	Dr. Vino S, Professor, SBST, VIT
	Dr. K Sri Manjari, GenepoweRx, Secretary, Bioclues.Org
Organising Secretaries	Dr. Sajitha Lulu S, Associate Professor, SBST, VIT
	Dr. Anshuman Dixit, ILS, Bhubaneshwar
	Dr. Anil Kumar S, Outreach Chair, Bioclues.Org
BIRD Award Chairs	Dr. Pushpendra Singh, Research Chair, Bioclues.Org
	Dr. Prashanth N Suravajhala, Founder, Bioclues.Org
Program Committee Advisers	Dr. P. B. Kavi Kishor, Professor Emeritus, Adviser, Bioclues.Org
	Dr. Nirmal K Lohiya, Professor Emeritus, Adviser, Bioclues.Org
	Dr. Anshuman Dixit, Principal Scientist, ILS, Bhubaneshwar
	Dr. V.S. Sundararajan, Professor Emeritus, Adviser, Bioclues.Org
	Dr. Arnold Emerson I, Professor, VIT
	Dr. Venkat Kumar S, Professor, VIT
	Dr. Rameshpathy M, Professor, VIT
	Dr. Devi Rajeshwari V, Associate Professor, VIT
Outreach Chairs	Dr. Indrani Biswas, Joint Secretary, Bioclues.Org
	Dr. Rajul Jain, Joint Secretary, Bioclues.Org
Local organizing	Dr. Pragasam V, Professor, SBST, VIT

Committee	Dr. Jayaraman G, Professor, SBST, VIT
	Dr. Gothandam K M, Professor, SBST, VIT
	Dr. Ramanathan K, Professor, SBST, VIT
	Dr. Rajasekaran C, Professor, SBST, VIT
	Dr. Shanthi V, Professor, SBST, VIT
	Dr. Asha Devi S, Professor, SBST, VIT
	Dr. Mohanapriya A, Associate Professor, SBST, VIT
	Dr. Subathra Devi C, Associate Professor, SBST, VIT
	Dr. Abilash V G, Associate Professor, SBST, VIT
	Dr. Suthindhiran K, Associate Professor, SBST, VIT
	Dr. Sudhakaran R, Associate Professor, SBST, VIT
	Dr. Gayathri M, Associate Professor, SBST, VIT
	Dr. Nalini E, Assistant Professor, SBST, VIT
Student Coordinators	Ms. Malavika J
	Ms. Kavya P
	Ms. Anagha Renjitha
	Ms. Goldy Singh
	Ms. Devi Soorya Narayana S
	Ms. Tejaswini B
	Mr. Premkumar T
	Mr. Girishwaran M

VIT®
UNIVERSITY
(Estd. u/s. 3 of UGC Act 1956)

Dr. G. VISWANATHAN
Founder & Chancellor
Former Member of Parliament
Former Minister, Govt. of Tamil Nadu
President, Education Promotion Society for India, New Delhi

MESSAGE

I am happy that the School of Bio Sciences and Technology (SBST), VIT, is organizing the Indian Conference on Bioinformatics 2023, Inbix'23, in collaboration with the Bioclues Organization, during November 24–26, 2023.

With excellent facilities and highly qualified teachers, SBST is among the top 10 bioschools in the nation. The Inbix series of conferences are globally recognized for their high scientific quality. I understand that the conference features a stellar line-up of world-class bioinformaticians from India and countries such as Japan, USA, UK, and Singapore, who will be presenting valuable keynote addresses and invited talks.

I welcome the participants from more than 40 institutions across India, as well as delegates from prestigious organizations.

I appreciate the members of the Organizing Committee for their meticulous planning and able execution of the Conference.

I wish the event great success.

Dr. G. Viswanathan
Founder & Chancellor

November 21, 2023

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Sankar Viswanathan
Vice President

24/11/2023

Message

I am delighted to greet all delegates and registered participants to the Indian Conference on Bioinformatics 2023, Inbix'23, which will be held from November 24-26, 2023, in collaboration with the Bioclues Organization and will bring together academic professionals, researchers, and industrialists from around the world.

I hope every one of you has enlightening and life-changing days in VIT, Vellore from November 24-26, 2021.

Sankar Viswanathan

Vellore - 632 014, Tamil Nadu, India; Phone: 91 - 416 - 2243091 (10 Lines) Fax: 91- 416 - 2243092
E-mail: registrar@vit.ac.in
www.vit.ac.in

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. G.V. Selvam
Vice President

21.11.2023

Message

I am pleased to welcome all delegates and registered participants to the Indian Conference on Bioinformatics 2023, Inbix'23, which will be held in collaboration with the Bioclues Organization from November 24-26, 2023, promises to bring together eminent researchers, and industrialists from all over the world. I believe that, at the end of the conference, attendees will be able to exchange research collaborations, hold brainstorming sessions for current and future publications, discover future research areas, and publish high-quality peer-reviewed articles.

I wish this conference to be both stimulating and productive. Personally, I hope you enjoy it and benefit from it!

Dr. G. V. Selvam
Vice President, VIT

Vellore - 632 014, Tamil Nadu, India. ☎ Office : +91- 416 - 220 5555 Email : gvselvam.vp@vit.ac.in
www.vit.ac.in

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. T. Jayabarathi
Registrar, VIT

Message

I'm delighted to inform you that the School of Biosciences and Technology, in collaboration with the Bioclues Organization, is hosting the Indian Conference on Bioinformatics, Inbix'2023. This event centres around the theme "Applications of Bioinformatics in Health Care". This Multidisciplinary Conference is one such step towards motivating students and researchers to present their innovative ideas and latest discoveries in the field of Bioinformatics. I would truly appreciate the team for devoting their time and creativity to organizing such a wonderful event.

In this conference, the team has curated an enriching agenda featuring keynote speakers from India and other nations, interactive workshops and thought-provoking discussions in the field of bioinformatics. This program will serve as an excellent environment for participants to exchange valuable insights and advancement in the field of bioinformatics.

I wish the conference a great success.

T. Jayabarathi

Dr. T. Jayabarathi
Registrar, VIT
24/11/2023

REGISTRAR
Vellore Institute of Technology (VIT)
(Deemed to be University under section 3 of UGC Act, 1956)
Vellore-632 014, Tamil Nadu, India

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. V. S. Kanchana Bhaaskaran
Vice Chancellor i/c

24-11-2023

Message

I am extremely happy to know that the School of Bio Sciences and Technology is organizing an Indian Conference on Bioinformatics, Inbix'2023, in association with the Bioclues Organization. The theme of the conference "Applications of Bioinformatics in Health Care" with its various subthemes that focus on pioneering areas like Machine Learning, Next Generation Sequencing, Functional Genomics and Proteomics etc., which are more pertinent in today's context of research.

During this conference, the team has an exciting lineup of keynote speakers, panel discussions, and workshops. This forum will establish collaboration, fostering knowledge sharing and research exploration, infusing novel ideas critical to advancing cutting-edge research in Bioinformatics and allied domains.

My best wishes for the resounding success of the conference.

Dr. V.S. Kanchana Bhaaskaran
Vice Chancellor i/c

Vellore – 632 014, Tamil Nadu, India. Phone: +91- 416 - 224 9720 / 220 2921; Fax: +91- 416 - 224 3092
Email: vc@vit.ac.in
www.vit.ac.in

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. Partha Sharathi Mallick

Pro-Vice Cancellor

MESSAGE

I am delighted that the School of Bio Sciences and Technology (SBST) will host the Indian Conference on Bioinformatics, Inbix'23, during November 24-26, 2023, in collaboration with Bioclues Organization, India's largest Bioinformatics Society that hosts the Inbix series of conferences in collaboration with multiple reputed organizations of the country.

I am delighted to read that several notable Professors from prestigious institutes in India and overseas have been invited to participate in the conference, as have participants from 45 institutions across the country.

I welcome everyone and hope the Inbix'23 a great success.

With Best wishes,

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. Siva Ramamoorthy
Dean, SBST

Message

On behalf of the School of Bio Sciences and Technology, it is my great pleasure and honor to welcome all the delegates and participants of the Indian Conference on Bioinformatics, Inbix'2023. I want to express my heartfelt appreciation to the team for devoting their time to put on a fantastic event for us to enjoy. I thank all the delegates for this incredible opportunity to learn and share from some of the field's most eminent scientists.

We have interesting keynote sessions as well as invited talks from renowned bioinformaticians and Industrialists from India and a few other nations. These speakers have dedicated their time to enhance our scientific understanding and ultimately to reach our full potential. Researchers, scientists, educators, and students throughout the country have given a positive response.

I am delighted to welcome you all for this three-day scientific feast.

Dr. R Siva
Dean, SBST
24/11/2023

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. Jayanthi Abraham
Professor and Associate Dean, SBST

Message

I am pleased to welcome all the delegates and participants to the Indian Conference on Bioinformatics 2023, Inbix'23. I hope that the acquaintances and networks formed as a result of this conference will lead to several fruitful collaborations that will pave the way to great advancements in our field.

We have organized several interesting keynote sessions and invited talks by professionals from renowned industries and leading academic institutes in the field of bioinformatics both from India and abroad.

This conference has interactive sessions with eminent scientists, which will pave the way to new therapeutics, diagnostics, and new products for the betterment of society.

I am sure this conference will lay the foundation for productive scientific interactions leading to useful and sound scientific outcomes in the thrust areas of the conference.

I am overjoyed to welcome you all to this three-day exploration of scientific knowledge.

Dr. Jayanthi Abraham,
Professor and Associate Dean, SBST
Associate Dean
Scheol of Bio Sciences and Technology (SBST)
Vellore Institute of Technology (VIT)
(Deemed to be University under section 3 of UGC Act,1956)
Vellore- 632 014, Tamil Nadu, India

Gyaneshwer Chaubey
President, Bioclues Organization

Message

The Bioclues probably stands as South Asia's largest non-clinical Biological society, boasting 9400 members (414 life members). What draws such a significant following? The dedicated core members and, of course, the field of Bioinformatics, which itself exhibits boundless qualities- interdisciplinary and pulsating with the vigor to embrace movements fostering new domains, consistently connecting the threads, and reaching new boundaries.

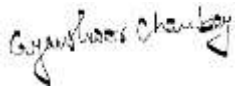
Attending our annual meetings (INBIX) ensures exposure to state-of-the-art research, instills excitement, and imparts immediate, applicable knowledge. Moreover, presenting at these yearly conferences offers the chance to connect with numerous individuals and share one's research. I attribute these expectations to our annual gatherings' unrestricted and open atmosphere. Regardless of age or professional standing, participants engage in uninhibited discussions and occasionally vigorous debates in the universal language of science. Alongside the Organizing Committee, we aim to nurture and promote this sense of exhilaration and stay at the forefront.

Our society has addressed research ethics issues previously considered taboo and challenging to confront directly. While recognizing the diversity of opinions on these matters, I firmly believe that the educational initiatives for young scientists, particularly in training on handling statistics and data appropriately, hold significant implications for the future of Science. The society has undertaken a distinctive initiative involving Bioinformatics education for school students. Over 100 lectures have been conducted online and offline, supported by the volunteer spirit of our members. Engaging in such grassroots activities sets our academic society apart, as such endeavors are relatively uncommon. Our commitment extends to other activities, including collaborative efforts for internationalization with overseas educational institutions. I extend my heartfelt gratitude to all those who have actively participated in our society, contributing to these endeavors passionately.

India lags in gender equality, securing the 127th position out of 146 countries in the Global Gender Gap Index 2023. While actively spearheading measures to combat this issue in our country, we should recognize the formidable challenge of achieving immediate improvement. This realization underscores the need for sustained and dedicated efforts to reinforce the positive trends that numerous individuals in our society have conscientiously pursued. Beyond gender and LGBTQ+ concerns, diversity encompasses broader dimensions such as nationality and regional disparities. We seek collaborative brainstorming and convergence on potential actions an academic society can take to collectively address these multifaceted diversity issues.

The Bioclues Society aims to keep exciting science at its core while actively addressing challenges, fostering an inclusive environment, and promoting diversity in Bioinformatics. Members are encouraged to contribute their opinions and advice to shape the future direction of the society.

In the end, I congratulate the honorable Dean and the organizing committee of VIT for hosting INBIX2023! During this congress, maintaining a foundation rooted in captivating scientific pursuits, we aspire to address diverse challenges with your valuable guidance and counsel.



Yours in Bioclues
Gyaneshwer Chaubey

Dr. K. Sri Manjari
Secretary,Bioclues Organization.

Bioclues, an acronym for BIOinformatics CLUb for Experimenting Scientists, has emerged as a pivotal force in the realm of bioinformatics, especially within the Indian scientific community. Established in 2005, this non-profit virtual organization has grown into one of the fastest-growing bioinformatics societies in India, boasting a membership of over 4600 individuals from nearly 30 countries.

In a landscape where bioinformatics has flourished with the mantra of 'sequence predicts structure predicts function,' the early bioinformaticists in India were primarily computational chemists and cell biologists exploring molecular modeling and protein function prediction. The need for a society focused on mentoring, outreach, research, and entrepreneurship became evident with a surge in bioinformatics programs at both undergraduate and graduate levels. This led to the inception of Bioclues in 2005.

Bioclues operates with a focus on four key avenues: Mentoring, Outreach, Research, and Entrepreneurship (MORE). These avenues represent the pillars of Bioclues' mission, guiding its efforts to foster a vibrant and collaborative bioinformatics community. These four avenues collectively form the foundation of Bioclues' activities and initiatives. The organization's commitment to MORE reflects its holistic approach to building a strong and dynamic bioinformatics community that excels in academic and research pursuits and embraces practical applications and real-world challenges.

Bioclues' commitment to excellence extends beyond national borders, as evidenced by its diverse membership and collaborations with international universities. The society's emphasis on open access and its role in bridging the gap between real-time professionals and academics highlight its relevance in the dynamic field of bioinformatics.

Bioclues stands as a testament to the collaborative spirit of bioinformaticians. Overcoming challenges, fostering innovation, and nurturing the next generation of scientists, Bioclues continues to play a vital role in shaping the future of bioinformatics in India and beyond.

As Bioclues steps into its 19th year of service in October 2023, the organization has set its sights on Vision 2030. Having successfully organized conferences across various locations, including Jaipur, Jalandhar, Shillong, and Guntur, the society remains committed to fostering interactions and collaborations among practitioners. The upcoming Indian Conference on Bioinformatics 2023 - Inbix'23, organized by Bioclues and Vellore Institute of Technology, signifies a continued dedication to providing a platform for researchers to engage with top professionals in the field. We eagerly anticipate engaging in future endeavors with VIT and Bioclues, fostering further successful collaborations.

K. Sri Manjari (PhD)
Secretary, Bioclues

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Dr. Vino Sundararajan
Professor, SBST, VIT
Convener, INBIX'23

Message

I feel delighted to welcome our international and national presenters, delegates from Japan, USA, U.K and across the world, and our generous sponsors to Indian Conference on Bioinformatics 2023, Inbix'23, hosted by School of Bio Sciences and Technology, VIT. Inbix'23 has over 300 members ranging from PhD students to significant research program heads from a variety of disciplines. Our conference theme is 'Applications of Bioinformatics in Health Care,' which recognizes the rapid breakthroughs in basic research and the necessity for their translation from the bench to the bedside, as well as the translation of the evidence created into clinical practice.

We would like to extend a warm welcome to our distinguished speakers who will carry forward each of the conference's themes in their talks. Invited posters will be on view during the conference with the main viewing session during lunch on Day1 and Day 2 at Foodys. Our Preconference workshops on 'Machine Learning' and 'Galaxy Server' further complement our program. Early career researchers, Graduates and Postgraduate students will be in the running for oral and poster awards in recognition of quality research.

A panel discussion on Saturday afternoon will address the topic 'Millets for Millennium' – will enlighten us with the nutritional and health benefits promised by millets. We aim through this forum, with active delegate participation, to debate key translational research questions and thereby help direct our efforts towards greater impact.

It has been our privilege to convene this conference. Our sincere thanks, to the conference organising committee for their dedication and hard work in creating an excellent scientific program. Inbix'23 is being supported by a number of sponsors to whom we are very grateful. Please take the time to visit their exhibition stalls and gather the information from our supporters.

We welcome you to VIT, Vellore and hope that this year's conference will challenge and inspire you, and result in new knowledge, collaborations, and friendships.

Dr. Vino S

INDEX

Keynote

How Can We Interpret the Genome Language?

Kenta Nakai, The Institute of Medical Science, the University of Tokyo, Japan

knakai@ims.u-tokyo.ac.jp

Abstract

It is often said that the genome is a blueprint of life written in an unknown language (i.e., the genome language). In this analogy, genes may be regarded as sentences and motifs as words. To master a language, we need a dictionary and a grammar book. In biology, dictionaries correspond to databases, and thus databases containing known transcription factor binding sites have been constructed. However, if we use such a database to interpret an unknown genome sequence, we instantly face the problem of many false-positive hits. Thus, motifs must be interpreted in an appropriate context, though there must be various molecular reasons underlying such contextual effects. In other words, an appropriate grammar book is needed. In contrast to dictionaries, however, it is not obvious how to organize a grammar book in biology. It may be constructed as a probabilistic model, like hidden Markov models, or a rule-based knowledge base. Recently, there have been great advances in the field of natural language processing. Surprisingly, such an approach, the large language model (LLM) approach, seems to be rather useful in interpreting the genome language, too. Then, maybe we can understand what contexts are important by reverse-engineering which positions in the genome those LLMs learned important to pay attention to. In this talk, I will explain what I have tried so far with the above thoughts as well as our recent efforts to interpret the grammar of enhancers and RNA splice sites.

LTA1

Historical Evolution of Protein Function Annotation Methods

Jayaraman Valadi, FLAME UNiversity

jayaraman.vk@flame.edu.in

Abstract

Overwhelmed with recent advances in sequencing methodologies, the volume of pure sequence and structure data along with its diversity are growing rapidly. Fueled by this, newer function annotation methods and paradigms are evolving continuously. In this lecture I will briefly outline the historical evolution of Machine learning and Deep learning-based Protein function annotation methods.

LTA2

Functional Genomic Approaches to Breed Sorghum bicolor (L.) Moench, the Great Millet for Climate Resilience

Kavi Kishor P B (Department of Genetics, Osmania University, Hyderabad, 500 007).

pbkavi@yahoo.com

Abstract

Sorghum bicolor or jowar is also known as great millet and used as a staple food in Asia including India, Africa and Middle East countries. It is used as a fodder and an important source for the production of ethanol (biofuel). It is rich in many phenolic compounds and anthocyanins that can act as antioxidants and hence prevent cancer. S. bicolor has been shown to reduce inflammation which could be due to the presence of anthocyanins and phenolics. It is gluten-free millet with high fiber, being used for weight loss and also preferred by the diabetics. It is claimed that it is safe for humans suffering from celiac disease and gluten intolerance. Ever increasing population coupled with climate change and water scarcity are the major threats to our food and nutritional security in future. This needs to be addressed on war footing utilizing our current understanding of the gene function identification, and genome-editing technologies. The genomic sequence of S. bicolor is known, but the functional validation of many of its genes is not yet over. Functional validation of the genes by overexpression's or suppressions will help to identify candidate genes that can be deployed in future in breeding programs of S. bicolor aimed at its superior agronomic performance under saline, water and high temperature stress conditions. In the present study, genes implicated in salt, drought and temperature stresses such as sodium porters, potassium porters, sodium porter-like proteins (SbNHXLP), and transcription factors like SbAP37 have been identified and validated. The details will be discussed during the presentation.

LTA3

Transcriptomics and Epigenetics of Leaf Rust Resistance in Wheat

Pushpendra K Gupta (CCS University, Meerut)

pkgupta36@gmail.com

Abstract

Transcriptomics and epigenetic modifications in wheat involving resistance against leaf rust due to two different genes, namely Lr28 for seedling resistance and Lr48 for adult plant resistance were examined. Using NILs, differentially expressed transcripts (DETs) and the sequences with differential epigenetic modifications (DNA methylation, histone modifications and ncRNAs) were identified. Methylation of DNA was studied using MASP, ChIP-PCR, MeDIP and BiS-Seq, histone modifications were examined using ChIP-PCR and ChIp-Seq and ncRNAs were identified using RNA sequencing. In transcriptome analysis >100 genes were found to be differentially expressed in resistant and susceptible NILs. Transcriptome data was also utilized for identification of few putative effector molecules of the pathogen. In susceptible line, a large number of genes were activated due to hypomethylation and fewer genes were repressed due to hypermethylation, suggesting that many genes that are active in S cultivar are silenced in R NIL. Among differentially expressed genes, two genes encoding N-acetyltransferase and peroxidase12 (examined using ChIP-PCR), largely matched with changes in H3K4/H3K9 acetylation patterns of their promoter regions. Methylation context was also important, such that mCG methylation was abundant in S cultivar and that of mCCG methylation in R NIL. Similarly, methylation of CHH, which is generally uncommon, was found to be abundant among differentially methylated regions. Among different regions of the genes also, level of methylation was generally abundant in intergenic regions followed by that in promoters, transcription termination sites (TTSs) and exons/introns. Using RNA-seq data, a number of miRNAs and lncRNAs were also identified to be differentially expressed.

Invited Talks

Decoding The Hypes and Hopes of Drug Discovery and Precision Medicine in Schizophrenia: A Genomic and Epigenomic Perspective

Moinak Banerjee (DBT-Rajiv Gandhi Center for Biotechnology)

moinak@gmail.com

Abstract

Schizophrenia is known to be influenced by both gene and environment. Majority of the drugs used to treat target the neurotransmitters. Based on this hypothesis several investigators had focused on genetics and epigenetics of neurotransmission in Schizophrenia. We wonder are altered neurotransmission the real targets of disease pathogenesis. In epigenetics DNA methylation, histone modifications and microRNAs are the crucial molecular signatures for determining the epigenetic influence. Several reports suggest differences in the pattern of DNA methylation both at global and gene specific level in Schizophrenia. However, many of these observations could not be replicated unanimously. The reason could be due to environmental differences, medication effects or ethnic (genomic) differences. Besides most of the studies have considered the genetic and epigenetic events as two independent events. None of these studies have implied the role of genomics of methylome in Schizophrenia. We demonstrated that the DNMTs which are responsible for maintaining the methylations, are themselves associated with Schizophrenia which might possibly explain the discrepancies in global or gene specific methylations. Methylation pathway can also be influenced by folate cycle, methionine cycle and transsulfuration cycle genes. Therefore, in continuation to methyltransferases which are regulated intrinsically, we also find that the methylation can also be influenced by extrinsically modulated genes. Majority of the studies have investigated on the role of epigenetics from the perspective of Schizophrenia pathogenesis but none of them have investigated whether these epigenetic changes could also be induced by the therapeutic drugs itself. This compelled us to identify if antipsychotic drugs can impact the host epigenome and if so, how does it impact treatment response. We evaluated the epigenetic response of the drugs using 850K genome wide methylation, 60K genome wide miRNA screening. Interestingly we observe similar direction for methylation and miRNA signature, which indicate that the epigenetic observation needs a careful evaluation in pathogenesis and how it impacts treatment response.

Pistachio Adaptation to Salt Stress as Revealed by Physiological and Proteomics Studies

Ramesh Katam (Florida A&M University), Mohammad Akbari (Pistat Research Center, Nazari Business Group), Rakesh Singh (Georgia Institute of Technology), Dalia Vishnudasan (Amrita School of Biotechnology, Amrita Vishwa Vidhyapeetham) and Elena Andriunaite (Lithuanian Research Centre for Agriculture and Forestry).

ramesh.katam@famu.edu

Abstract

Pistachio (*Pistacia vera* L.) is an economically important tree nut that commonly thrives in semi-arid and arid environments. P. vera is a highly adaptable to various abiotic stresses, and it can tolerate drought and salinity stresses, which makes it suitable for reforestation of arid and salinized zones. However, the mechanisms underlying the salinity tolerance of this plant are not well understood. The present study was aimed at physiological and molecular investigations to unravel the metabolic pathways associated with the salt tolerance mechanisms in UCB-1 cultivar. Five one-year-old pistachio rootstocks were treated with four saline water regimes for 100 days. The rootstocks adopted Na+ exclusion strategy to resist the salinity stress. Total proteins were isolated from the roots and treated with different NaCl concentrations. The proteins were characterized using high throughput LC-MS/MS spectrometry searched against the Citrus database. Over 1600 protein IDs were detected, among which the comparative analysis revealed 245 more abundant and 190 low abundant proteins to three stress levels. The proteins associated with amino acid metabolism, cell wall organization, protein homeostasis, response to stress, signal transduction, TCA cycle, and vesicular trafficking were constantly over expressed at all stress levels. At low and moderate stress levels, the chromatin and cytoskeleton organization lipid metabolism proteins were over expressed, while at higher salt concentrations, they were unaffected. Transcription and translation processes were affected by all stress levels, as the proteins showed down-regulation in response to all stress levels. Transcription proteins were downregulated at low and moderate stress while over expressed at high salt stress treatment. Protein interaction network with all the orthologous proteins mapped to Arabidopsis thaliana and the clusters associated with these proteins revealed the cytosolic, carbohydrate, and amino acid metabolism are associated with

salinity stress tolerance. The proteome data were validated with corresponding changes in transcripts.

Tools for Molecular Design and Activity Profiling: A Reductionist Approach

Vibin Ramakrishnan, Indian Institute of Technology Guwahati

vibin@iitg.ac.in

Abstract

Translation of basic science to novel diagnostic and therapeutic solutions for prevention, diagnosis and treatment of diseases is an exciting area of research. This process however, is complex in its design and execution. Here we present a blend of three computational and experimental methodologies that can be employed in the design and profiling of functional molecules in the early stages of the drug discovery.

In the first part, we discuss the prospects of a reductionist approach in converting protein structure to a Barcode. In the second, a 'clock model', for virtual activity profiling of drug candidates will be discussed. This algorithm may be extended to drug promiscuity and supplementing fragment-based drug design efforts.

In the third part, we present the design and development of a minimal blue fluorescent protein. We have employed two design tools developed by us; AR-SAMD and IDeAS. This molecule to the best of our knowledge, is the first blue fluorescent artificial protein designed with diversified chain stereochemistry.

Newborn Screening by Integration of Metabolomics and Genomics for Second-Tier Analysis

Sunil Kumar Polipalli, MAMC & Associated Lok Nayak Hospital

sunilpkumar18@gmail.com

Abstract

Newborn screening for treatable "hidden" hereditary metabolic disorders was introduced almost 60 years ago. Since, its inception in 1963, newborn screening has traveled a part from a targeted analyte assay to a program of immense public health. From inception, being a microbiologic test conducted on Guthrie paper, it currently uses fluoroimmunoassay and liquid chromatography mass spectrometry. As envisaged originally, in context to a good public health program, it is designed to have a good sensitivity at a trade off with tolerable specificity. As a result, all the analytes that do not fall into a definite presumptive positive flagging but may suggest a series/cause of IEM. As a protocol despite repunch from another spot recalling the infant for an ambiguous diagnosis, brings with it both the emotional trauma and anxiety besides additional costs for the program. Last 2 decades have seen the emergence of second tier biochemical screening on the residual blood spots to increase the positive predictive value of these analytes. The performance matrix has improved by adding ratios, adding second tier biochemical tests like alloisoleucine for the diagnosis of Maple syrup urine disease, methyl citric acid for propionic acidemia in a neonate with increased propionyl carnitine amongst many others. Genomics in NBS is emerging not only due to improved techniques of DNA extraction but also the widespread use of next generation sequencing. Using DNA based technology as a second-tier testing modality will not only complement its biochemical counterpart but also significantly improve the specificity of the program. In a country like India, with limited medical literacy, conceptual understanding of the informed consent form and denial of recall due to cultural taboos makes second tier genomic testing of immense benefit. There are other challenges that need to be addressed using second tier genomics. The genetic basis of IEMs is very heterogeneous and can involve abnormalities such as point mutations, deletions or insertions, or more complex genomic rearrangements. Introduction of molecular genetics techniques have made it possible to identify molecular defects and confirm diagnoses in IEMs. However, disease phenotypes are not always explained

by the detected variants. Indeed, the level of IEM complexity requires an integrated understanding of perturbations in genetic and biochemical networks. The aim of this study is to integrate expanded newborn screening using metabolomics and genomics for the second-tier screening method to assist clinical diagnosis. IEM phenotype characterization potentially provides the clinician better information for personalized care. In combination with genomics, large-scale semi targeted metabolomics is expected to reveal genotype-phenotype correlations and the overall effect of drugs and dietary interventions. However, for the successful translation of global metabolomics from the bench to bedside, some gaps need to be bridged and quality control challenges need to be addressed. This will usher a new area where positive predictive yields will improve without increasing the recall and precision neonatology will improve neonatal mortality rates with a move towards precision wellness, the aim of this decade.

Can Glycan Alphabet Provide Clues to Strain-Level Variations in Gut Microbiome?

Jaya Srivastava (Indian Institute of Technology Bombay) and Petety V Balaji (Indian Institute of Technology Bombay)

balaji@iitb.ac.in

Abstract

Proteins, nucleic acids, carbohydrates, and lipids are the four major classes of biomolecules. The alphabets of proteins and nucleic acids have been extensively studied and are found to be well conserved in evolution. In contrast, characterization of the alphabets of carbohydrates and lipids has considerably lagged behind due to inherent structural complexity. There are no known "templates" (vis-a-vis those for proteins and nucleic acids) for the biosynthesis of glycans and lipids. Our analysis of completely sequenced prokaryotic genomes showed that monosaccharides may be grouped as common, less common, and rare based on their prevalence in Archaea and Bacteria. In addition, we found substantial variations in the set of monosaccharides used by organisms belonging to the same phylum, genera and even species. This can be exploited to identify strain-level variations in microbial communities.

Human Housekeeping Cis-Regulatory Elements and Their Involvement in Tumor Suppression

Martin Loza (The University of Tokyo), Alexis Vandenbon (Kyoto University) and Kenta Nakai (Institute of Medical Science, University of Tokyo)

knakai@ims.u-tokyo.ac.jp

Abstract

This study explores housekeeping cis-regulatory elements (HK-CREs) in the human genome. Through extensive multiomics analysis, we highlight the unique epigenetic features of these elements and explore their importance in vital biological processes beyond the regulation of housekeeping genes. Notably, we observe reduced activity of HK-CREs in cancer cells, particularly those near the telomere region of chromosome 19 and associated with zinc finger genes. Further analysis, including cancer samples, suggests the importance of these genes in housekeeping tumor suppressor processes. Overall, our findings highlight the importance of HK-CREs within the cells for preserving cellular integrity and stability.

Advances in Plant Phenotyping for Climate Smart Agriculture: Applications of Hyperspectral Imaging

Pawan Kulwal (Centre for Advanced Agricultural Science and Technology for Climate Smart Agriculture), Sunil Kadam* (Mahatma Phule Krishi Vidyapeeth, Rahuri, MS), Anjali Pundkar (Mahatma Phule Krishi Vidyapeeth, Rahuri, MS) and Vishal Pandey (Mahatma Phule Krishi Vidyapeeth, Rahuri)*

pawankulwal@gmail.com, sunil21075@gmail.com

Abstract

Majority of the agriculturally important traits are dependent upon many physiological and biochemical parameters which are ultimately responsible for biotic and abiotic stresses. However, these traits are often expensive, destructive or slow to score and also there is large gap between plant physiology, genetics and phenomics investigations. Moreover, in order to identify the desirable plants, breeder often needs to take repeated observations in the field and make careful selection. This not only requires lot of time, but skill and experience of a breeder. It has now been realized that in any plant breeding program, rapid and precise phenotyping for the desired trait is very essential. Since this involves recoding thousands of data points in shorter time, in recent years a shift from traditional way of phenotyping to use of sensor-based phenotyping has been seen. For instance, digital images or sensor-based images of standing crop in the fields are taken from the surface or through air with the help of unmanned aerial vehicles. This not only saves time in recording the data but also reduces the error associated with the manual way of recording observations. This makes it necessary to use techniques which are high throughput and non-destructive in nature. Hyperspectral imaging-based canopy reflectance is one of the recent and promising techniques. The spectral signatures reflected from the plant canopy at different wavelengths provide different types of information on specific plant characteristics responsible for phenotyping. The spectral signatures are closely related to biotic and abiotic stress induced changes in several biochemical and biophysical traits. These can be related to genotypic differences and stress levels and can be detected through the changes that take place in the spectral signatures of the canopy measured in the visible, near-infrared, and short wave-infrared regions.

Based on the experiments carried out under controlled as well as field conditions, we observed that spectral signatures captured using hyperspectral imaging system can efficiently distinguish tolerant/resistant and sensitive genotypes for temperature stress in wheat as well as leaf blast in rice. We observed that the canopy spectral signature can efficiently be used for phenotype sensing for breeding purposes. The amount of high-dimensional data which can be generated in any such large experiment can effectively be analyzed using the technique of artificial intelligence and can be used for identification of genes/QTLs governing different traits. This will provide plant breeders with important information to increase the chances of recognizing genotypes/ varieties that are well-adapted to biotic and abiotic stresses by identifying indirect non-destructive traits.

Millets for Millenium

Pawank Kulwal (Vignan University) and Anil Kumar S (VFSTR, Guntur)

pawankulwal@gmail.com

Abstract

The United Nations General Assembly at its 75th session in March 2021 declared 2023 the International Year of Millets (IYM). As part of these celebrations, several events in the form of workshops, seminars, symposia, brain storming sessions, etc. related to promotion of millets were organized earlier this year while some more are planned. Millets because of their resilient ability to grow under arid conditions and with minimal inputs can play an important role in food and nutritional security.

Discussions taking place as part of IYM at different forums provide an opportunity to create awareness about several aspects of millet production, processing, consumer preference, marketing, crop improvement, health benefits and many other things. The panel discussion is thus planned to discuss following aspects

i) are millets potential answer for providing food and nutritional security for the growing population,

ii) how advances in genomics can play important role in understanding millets better,

iii) how bioinformatics can help in identification of important genes in millets,

iv) what more need to be done to promote millets,

v) what does the future hold for millets in India?

vi) can celebrating IYM will change people's perspective towards consuming millets?

Model-Informed Drug Discovery and Development: From Bioinformatics to Quantitative Systems Pharmacology (QSP)

Viji Chelliah, Certara UK

viji0312@gmail.com

Abstract

A large proportion of drug development projects fail in phase II and phase III clinical trials mainly because of the lack of efficacy and unacceptable safety profile. One of the notable contributing factors contributing to this failure is an inadequate understanding of the underlying disease biology and target-disease linkage. This results in poor target choice, suboptimal target modulation, unanticipated structure-based or mechanism-based toxicity, inappropriate patient-population selection, and the absence of decision-making biomarkers. Therefore, finding novel, druggable targets associated with high confidence in rationale for therapeutic efficacy and safety remains a major challenge. Adoption of a discovery pipeline based on in-depth understanding of disease biology and mechanisms is an absolute need for identifying potential targets for clinical success. It has been reported that the implementation of the so-called 5R framework (right target, right tissue, right safety, right patient, right commercial) increased the trial success rate from 4% to 19%. Model-informed drug discovery and development involves the use of mathematical models in exploring disease mechanisms, biomarker predictions, drug-dose predictions and decision-making in pharmaceutical R&D contributing majorly to 5R framework. Quantitative systems pharmacology (QSP) models, in particular, is now increasingly being employed to translate the rapidly growing understanding and mapping of complex biology and pathophysiology into a solid foundation for the efficient and rational development of novel modalities, combination therapy and innovative treatment regimens to treat diseases in high medical need areas. However, developing a QSP model is an extremely laborious process which limits exploring multiple targets and mechanisms. Here, I will discuss about how bioinformatics approaches can inform the development of QSP models to enable safe and effective new therapeutics to advance more efficiently through the different stages of drug discovery and development pipeline.

Journey Of Developing Reversible Inhibition of Sperm Under Guidance (Risug®) As an Injectable Male Contraceptive with Special Reference to Seminal Proteomics

A.S. Ansari (Centre for Advanced Studies, Department of Zoology, University of Rajasthan, Jaipur), Barkha Khilwani (Centre for Advanced Studies, Department of Zoology, University of Rajasthan, Jaipur) and N. K. Lohiya (Centre for Advanced Studies, Department of Zoology, University of Rajasthan, Jaipur)

ansari3808@gmail.com

Abstract

Reversible inhibition of sperm under guidance (RISUG®) is an intravasal injectable male contraceptive intend to be developed as an alternative to vasectomy to provide safe and long-term contraception. It is a polymeric gel consisting of styrene maleic anhydride (SMA) dissolved in dimethyl sulfoxide (DMSO) in 1:2 ratio when injected into the vas deferens blocks sperm passage that traverse through and destroys them thereby rendering the subjects sterile. The non-invasive and invasive reversal techniques, successfully demonstrated in langur monkeys, and rats and rabbits, respectively. Extensive toxicological investigations carried out in different animal species demonstrated safety of the procedure. Based on phase I, II, extended phase II and limited phase III clinical trials data and on genotoxicity, mutagenicity and carcinogenicity study data, the Indian Council of Medical Research, New Delhi conducted the phase III clinical trials at five centers. The phase III clinical trial on RISUG injected subjects at our center indicated early onset of contraception. The sperm functional tests showed a drastic reduction and marked sperm deformities prior to azoospermia. The neutral α-Glucosidase, GPC and L-Carnitine concentration was gradually and markedly reduced. No marked alteration was found in circulatory levels of hormones, PSA and anti-sperm antibodies. Two-dimensional gel electrophoresis analysis revealed a total of 235 protein spots in RISUG injected human subjects. When they compared with fertile subjects 110 protein spots were matched. Out of which 57 spots were down-regulated and 53 spots were up-regulated. Following MALDI TOF – TOF MS analyses Prolyl endopeptidase-like (Fragment) (PREPL), Focal adhesion kinase 1 (Fragment) (PTK2) and two different spots of Prolactin-inducible protein (PIP) were identified through MASCOT protein database with NCBI blast search.

It is concluded that RISUG, a safe and quicker contraceptive method, hopefully available for mass application for human soon. The identified differentially expressed proteins following RISUG® administration may be useful for management of infertility.

Candida Sterol 14α-demethylase – Patterns of Amino Acid Substitutions and Azole Resistance

R. Shyama Prasad Rao (Center for Bioinformatics, NITTE deemed to be University, Mangaluru 575018)

drrsprao@gmail.com

Abstract

Drug resistance in bacteria and fungi is a major concern, and the azole-resistant Candida infections are on the rise. One key mechanism is the emergence of resistant amino acid substitutions in sterol 14α-demethylase, the target of azole drugs. While many such substitutions are reported, it is unclear how prevalent they are or how they exert differential effects on different azoles. We performed sequence analyses, molecular dynamics simulations, and free energy calculations to understand the nature of substitutions and how they alter the binding free energy and interactions of azoles with the protein. Based on a set of 2,222 instances, Y132F/H, K143R, D116E, G464S were some of the frequent azole-resistant substitutions. While substitutions were found at 133 residue positions, only a third of the azole-binding sites had any known substitutions. The ligand-binding free energy for fluconazole, a short-tailed azole was far higher (-13.81 kcal/mol) than for VT1, a medium/long-tailed azole (-35.04 kcal/mol). There were differences in the ligand-binding free energies after substitutions compared to the wild type protein. Multiple substitutions also showed incremental differences in ligand-binding free energies. Concomitant alterations in the residue orientations, and the distances between the residues and the ligand were also observed. The results provide valuable insights into azole resistance and antifungal drug discovery and optimization.

Zebrafish as a Powerful Model System to Investigate Cellular and Developmental Mechanisms

Anil Kumar Challa (Shiv Nadar Institution of Eminence (Deemed to be University), Delhi NCR)

anil.challa@snu.edu.in

Abstract

The zebrafish, once a lesser-known tropical fish mentioned in Western scientific literature as early as the 1820s, has emerged as a powerhouse in modern biomedical investigations. Initially discovered for its striking appearance, it wasn't until the 1930s that researchers recognized its potential as a valuable experimental system. However, it wasn't until the 1970s that the zebrafish truly captured the spotlight, being adopted as a model organism for the study of neural development and genetics. The zebrafish's rapid external development, transparency of embryos, and genetic tractability have since propelled it to the forefront of scientific research. Its utility in large-scale genetic screens and its biological similarities to humans make it an invaluable tool for unraveling mysteries in cellular and developmental biology, contributing significantly to advances in biomedical research. The talk will focus on a few success stories where zebrafish, as a model system, played an important role in advancing our knowledge of cellular and developmental biology, with far-reaching implications for both basic science and clinical applications.

New Drug Development - Lab to Launch (An Industry Perspective)

Manivannan Boomi (Chief Brand Promotion & VP - Plant Operation, Genaxy Scientific Pvt Ltd., New Delhi/Solan (HP)

bmani@genaxy.com

Abstract

Discovery of new drugs for human application involves many years, many processes, often failures with much uncertainty. There are many hits and trials to identify a new molecule from basic research in to discovery and development of new drug in application point of view. The process is too complex, time consuming, expensive and to resolve many operational issues to target the human application. Developing a new medicine in regulatory point of view, i.e., an Investigational New Drug (IND) primarily of its mode of action different from the approved medicine intended for an indication that is not addressed yet, involves several stages, viz., target identification, mode of action, process development, process standardization and validation, proof of potency, safety and convenience to use in mass application. Development of a new molecule towards commercial application that is from target identification to market authorization approximately takes over 12 years many times, even more, costing about $ 1-2.5 billion and even more, an estimate based on analysis across several therapeutic development. The developmental activities of an IND involve basic research, preclinical studies in animal models, clinical studies in human participants (Phase I, Phase II and Phase III). The developmental phases of new drug development in an Industry Perspectives from the developmental stage to Launch of the product shall be discussed.

Bioinformatics Applications for Solving the Mystery of Complex Human Diseases: An Aid in Healthcare

Tiratharaj Singh (CEHTI, Department of Biotechnology and Bioinformatics, JUIT)

tiratharaj@gmail.com

Abstract

Alzheimer's Disease (AD) and various kind of cancer, both are very complex and multifactorial diseases and a serious global burden to our society, where multiple enzymes simultaneously activate due to the combination of genomics, interactome, and environmental factors. Due to the complex nature of these diseases, combination of bioinformatics, genomics, systems biology and molecular evolution can play a central role to identify the potential targets as well as the disease mechanisms. We have used computational genomics and the systems biology approaches to find the disease progression mechanisms. Quantitative systems biology approach was applied to decipher the drug inhibition as well as disease progression mechanisms. A virtual AD cell model was reconstructed using the systems biology graphical notation and then the concentration of targets (key enzymes) and drugs were increased to observe the effects. We have predicted eight key targets which can play a key role in AD progression. In another work we have created an AD network and then heuristic approaches were used to find the novel information. We have calculated several statistical parameters such as shortest path length, node betweenness, degree, clustering coefficient and predicted the important nodes (hub nodes) which can act as a target for AD pathogenesis. On the other hand, we have performed combination of these studies on various human malignancies such as colorectal cancer, endometrial cancer and prostate cancer. Structure based small molecular studies were also done for AD as well as cancer. From all the results we have concluded that bioinformatics, computational genomics, structure-based drug design and systems biology approaches are useful to analyze cancer and AD-like complex diseases data to provide a meaningful piece of information to the experimental scientists for lab-based validations. It is anticipated that this kind of analyses will provide a systematic protocol for the analysis of complex diseases data and will help the biomedical community and healthcare sector for further progressing in the positive direction.

Bioprospecting of Halotolerant Microorganisms by Metabolomic Approach

Gurunathan Jayaraman (School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore)

gjayaraman@vit.ac.in

Abstract

Microorganisms can be found residing in nearly any type of environment, even in the harshest chemical and physical conditions. These microbes have evolved through time to become adapted to these "unfriendly" conditions. Due to this natural phenomenon, they are a valuable target for the discovery of novel macromolecules as well as small compounds, which may have appealing uses in numerous industrial processes or in therapeutics. Over the past few years, we have been working to understand the adaptive traits of the microorganisms found in various salt pans along India's coastal region, as well as their ability to produce proteins and metabolites of significant industrial value. These microbes produce thermostable, halotolerant hydrolases with distinct substrate specificity and are tolerant to heavy metal ions. Some of these organisms can survive even at 3–5 M NaCl owing to alterations in their general morphology and enhanced exopolysaccharide synthesis. Comparative metabolomics approach, using nuclear magnetic resonance spectroscopy as well as mass spectrometry, has helped us to identify the changes in the biochemical pathways which occur under salt stress and to discover new and rare small molecules which are of industrial / therapeutic importance.

In Silico Secretome Prediction and Expression Analysis of Potential Effector Candidates of Fall Armyworm (*Spodoptera frugiperda*)

Sundaram Shilpi (Central University of Rajasthan), Vivek Verma (Central University of Rajasthan) and Jayendra Shukla (Central University of Rajasthan)

jayendrashukla@curaj.ac.in

Abstract

Effector proteins, one of the major insect salivary gland components, alter host defense mechanism(s) and facilitate pests for successful infestation of host plant. Fall armyworm, *Spodoptera frugiperda* is a polyphagous lepidopteran insect infesting a wide range of agricultural crops. Despite being one of the world's deadliest pests, no information about the effector proteins of S. frugiperda is available, till date. In view of the secretory nature of effectors, an in silico secretome of S. frugiperda was generated. For this, we performed an *in-silico* analysis of interproscan-annotated protein sequences of S. frugiperda (derived from its transcriptome) using established secretome prediction pipelines. Out of 21,779 protein sequences of S. frugiperda, 821 proteins were predicted to be secretory in nature, leading to the generation of an in silico secretome database of S. frugiperda. The proteins of S. frugiperda secretome were categorized into different functional groups as per their annotated functions. The expression of 40 selected candidates was analyzed in different tissues (head, gut, salivary gland and fat body) of S. frugiperda, which revealed 14 candidates to be exclusive to a single tissue. In addition, expression of 13 candidates were found to be exclusive to gut or salivary glands or to both the tissues indicating that they may be secreted out from the insect's body and serve as potential effector proteins. Further, the expression (in the gut and salivary gland of S. frugiperda) of potential effector candidates will be compared between the insects fed on artificial diet versus the insects fed on plants which will help in the identification of effector proteins of S. frugiperda.

Biogenic Iron Oxide Nanoparticles and CRISPR as a Panacea for Combatting Global Warming

K Sangeetha (School of Biotechnology, Amrita Vishwa Vidhyapeetham), S Induja (School of Biotechnology, Amrita Vishwa Vidhyapeetham), S Ajeesha (School of Biotechnology, Amrita Vishwa Vidhyapeetham), M.K. Reddy (nternational Centre for Genetic Engineering and Biotechnology, New Delhi, India), V. Mohan M. Achary (nternational Centre for Genetic Engineering and Biotechnology, New Delhi, India), Nair Bipin (School of Biotechnology, Amrita Vishwa Vidhyapeetham), Panicker Manoj (TKM College of Arts and Science, Kollam) and Vishnudasan Dalia (School of Biotechnology, Amrita Vishwa Vidhyapeetham).

daliavishnudasan@gmail.com

Abstract

Biogenic iron oxide nanoparticles (FeNp) are synthesized using green chemistry approaches and are hence very eco-friendly and sustainable. These FeNp have numerous applications, particularly in the field of agriculture and wastewater remediation. FeNp is used as a fertilizer to improve the growth and yield of crops. FeNp is known to activate the oxidation defense system, scavenging reactive oxygen species (ROS) and adsorbing heavy metals. Studies in our lab indicate the potential of FeNp in enabling the plants to overcome abiotic stresses such as - Drought stress, Salinity Stress, and Arsenic stress. These nanoparticle-based iron fertilizers have enormous potential in agriculture because of their low cost and low toxicity.

Wastewater remediation using Biogenic iron nanoparticles (FeNp) encompasses the removal of organic pollutants, heavy metals, microplastics, and the catalytic degradation of pollutants. Biogenic iron oxide nanoparticles (FeNp) can be used to adsorb contaminants from wastewater and decolorize dyes in wastewater. Studies in our lab indicate using FeNp for wastewater remediation and microplastic removal. Magnetic iron oxide (FeNp - Fe3O4) nanoparticles were effective in removing microplastics, including polyethylene terephthalate, high-density polyethylene, LDPE, polyvinyl chloride and polypropylene. Hence FeNp has potential applications in wastewater treatment. Furthermore, we addressed the safety issue of the FeNp that would remain in the aquatic environment. The aquatic plant Azolla was exposed to microplastics, and its amelioration was initiated with FeNp. Interestingly, Azolla and its symbiont are negatively

impacted by microplastics (Mp), but use of FeNp enables the Azolla plant to overcome Mp pollution.

Rice (Oryza sativa) is a dietary staple for half the world's population. To meet the growing demand for rice and adaptability to changing environmental conditions, the CRISPR technology is being employed to alter its traits. Phospholipase D beta 1 (PLDβ1) a key enzyme in lipid metabolism was altered using CRISPR Cas9. By manipulating the PLDβ1 gene, adaptability to adverse conditions, such as drought and salinity was noticeably achieved in Samba Masuri BPT-5402. Moreover, PLDβ1 is responsible for the modulation of other PLD isoforms expression in rice. Knock out of PLDβ1 gene elevates expression of stress-inducible genes, osmotic biosynthesis genes, lignin biosynthesis genes, and enhances ROS scavenging activity, which together contribute to improved abiotic stress tolerance. This study provides novel insights into the function of the PLDβ1 gene in rice. Further analysis of their regulation under abiotic stress will promote the molecular breeding of stress-tolerant rice.

A Computational Study of Conformational Transitions in Intrinsically Disordered Regions on Complexation

*Prachi Bhargava[1], Madhabendra Mohon Kar[1], Paramveer Yadav[1] and Amita Barik[1]**

Department of Biotechnology, National Institute of Technology, Durgapur, India

*amita.barik@bt.nitdgp.ac.in**

Abstract

The structure–function paradigm has been considered as the central dogma of structural biology. In contrast to it, a different group of proteins has been given recognition that do not adopt their defined three-dimensional structure but plays important role in different biological functions. This group of proteins is referred as Intrinsically Unstructured or Intrinsically Disordered Proteins (IDPs) and the regions are termed as Intrinsically Disordered Regions (IDRs). IDRs play a pivotal role in modulating cellular processes and signaling pathways and can provide valuable insights into drug design as their structural disorder can aid in ligand selection for drug development. IDRs often act as hubs in protein-protein interaction (PPI) networks due to their conformational flexibility, allowing them to interact with multiple biomolecules. They can adopt a well-defined tertiary structure upon binding, retain some degree of disorder, or remain completely disordered, forming fuzzy complexes. We analyse the conformational changes in IDRs upon complex formation using unbound proteins and the protein complexes. IDRs are enriched in polar charged amino acids and are depleted in aromatic residues. A study of IDRs in unbound proteins which get ordered upon complex formation after binding to suitable partners such as other proteins, DNA and RNA was carried out. An analysis of secondary structures reveals that 79.78% residues in unbound polypeptides form coils upon binding to suitable partners. We also observe that such residues are located at the interface of protein complexes suggesting that they contribute to the stability of the complexes. Amino acids that undergo transitions also contribute in hydrogen bond formation in the protein complexes. Our findings provide fundamental insights into the underlying principles of molecular recognitions by disordered regions. There are some structured regions in the unbound proteins which upon complexation become disordered and it will be intriguing to learn more about these ordered-to-disordered transitions upon complex formation.

Keywords: Intrinsically Disordered Regions (IDRs), protein complexes, unbound proteins, conformational changes, interface, hydrogen bonds

Design and Validation of CRISPR-Cas13a-Based Tool for Detection of K. pneumoniae

Gargi Bhattacharjee, Vijai Singh

Department of Biosciences, School of Science, Indrashil University, Rajpur, Mehsana, Gujarat, India

vijaisingh15@gmail.com

Abstract

Every year millions of lives lost due late and inadequate diagnosis. A rapid, specific and sensitive diagnosis play a vital role in infection control and public health initiatives to restrict disease transmission in highly equipped medical centre. An ideal diagnostic procedure is one which is rapid, affordable, error-free, and enables point-of-care (POC) operation without the necessity of technical trained person, expensive instrumentation, or power supply. A test that endows these features could aid in the early detection of highly virulent pathogens, isolation to prevent disease spread, and facilitate prompt medical attention and timely cure. In an attempt of addressing the unmet need for reasonably priced, logistically feasible and distributable detection devices for the rapid detection of Klebsiella pneumoniae that can function without the assistance of sophisticated laboratory equipment, a reliable CRISPR-assisted lateral flow assay-based detection platform was devised. In this study, we established a rapid and sensitive CRISPR-Cas13 diagnostic assay using a lateral flow-based platform by amalagamating recombinase polymerase amplification (RPA) for sequence-amplification and the Cas13a orthologue from Leptotrichia wadei (Lwa)Cas13a for detection of hypervirulent phenotypes of K. pneumoniae. We employed CRISPR detection platform with RPA to first demonstrate the detection of *Klebsiella spp*. by aiding species-specific detection of K. pneumoniae by targeting housekeeping rpoB gene, and then proceeded to detect the hypervirulent strains of K. pneumoniae. The capsular polysaccharide regulating gene rmpA was opted as the target. Out of 18 K. pneumoniae strains, the devised tool detected *K. pneumoniae* M59 and *K. pneumoniae* KP109 strains with the presence of rmpA. In the long-run, the idea is to generate an instrument free platform for routine diagnosis of K. pneumoniae from serum, urine and saliva samples from patients. This would enable the healthcare personnel to encourage proper and timely treatment of infections caused by *K. pneumoniae*.

Keywords: *Klebsiella pneumoniae*, CRISPR-Cas13a, Hypervirulent, Diagnostics, Multidrug resistance

Oral Presentation

MicroRNAs and Gene Expression Analysis for their Regulatory Role in Alzheimer's Disease

*Vinamrata Sharma [1], Saumya Tyagi [1], Vikram Singh[2] and Tiratha Raj Singh[3]**

[1]Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT), Solan, India

[2]Centre for Computational Biology and Bioinformatics, Central University of Himachal Pradesh (CUHP)

[3]Centre of Excellence in Healthcare Technologies and Informatics (CHETI), Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT), Solan, India.

tiratharaj.singh@juitsolan.in

Abstract

Rationale/Motivation: To identify the key genes and miRNAs that are connected to molecular events in Alzheimer's disease (AD) and to examine how they interact and regulate these biological entities. AD is a progressive neurodegenerative disease that is characterized by changes in neuropathology and a decline in cognitive function. There is no cure for AD, and current treatments are only able to slow the progression of the disease. miRNAs are small, non-coding RNAs that play an important role in gene regulation. They can bind to mRNAs and either promote or inhibit their translation into proteins. miRNAs have been shown to be involved in a variety of diseases, including AD. This extensive study about miRNAs can be utilized in the future to familiarize with the basic mechanisms underlying the over- and under-expression of genes in AD. As a result, by looking and evaluating the miRNAs targets in the future molecular research we can tune their therapeutic and regulatory role. We hope to identify the key genes and miRNAs that are involved in AD so that they can develop new diagnostic and therapeutic strategies for the disease. By understanding how miRNAs regulate gene expression in AD, they may be able to develop new drugs that can target these miRNAs and alter the course of the disease.

Objective: To evaluate which target genes and microRNAs are implicated during Alzheimer's disease (AD), as well as which of them are often overexpressed and under expressed in AD, and to comprehend the regulatory network of these genes. Using a range of datasets, this study

performed a bioinformatic analysis on miRNA-AD studies in order to find 250 expressed genes and 67 expressed miRNAs. The study then looked at the regulatory networks and pathways associated with the genes that govern AD. The TAU and APOE genes were found to have a major effect in AD. The goal of the study is to discover miRNAs and the target genes that are both frequently overexpressed or under expressed in AD, as well as to comprehend the regulatory network behind these miRNAs. This information can be used in future research to develop new therapeutic and regulatory strategies for AD.

Materials and Methods: In this study, we conducted a bioinformatic analysis of miRNA-AD studies using a variety of databases, including GEO Database, STRING, miRBase, KEGG, TargetScan, and Database for Annotation, Visualization to identify the miRNAs that are frequently overexpressed and under expressed in AD. We also analyzed the genes and miRNA expression profiles associated with AD, for which we identified the miRNA target genes.

Data collection: We collected data from the GEO Database, STRING Database, miRBase, KEGG, TargetScan. Identification of differentially expressed miRNAs: We used miRBase to identify the target genes of the differentially expressed miRNAs. Identification of miRNA target genes: We used the GEO2R tool to identify differentially expressed miRNAs in AD patients compared to healthy controls. Pathway analysis and Gene ontology analysis: We used the KEGG and STRING Database to analyze the pathways of the genes that control AD.

Results and Discussion: We conducted a bioinformatic analysis of miRNA-AD studies to identify miRNAs that are frequently overexpressed and under expressed in Alzheimer's disease (AD). We found a total of 250 expressed genes and 67 expressed miRNAs associated with AD. The GEO2R software provides the logFC values for each gene and miRNA, which measure the fold change in expression between the two age groups. The logFC (log fold change) value in GEO2R is a measure of the fold change in expression of a gene or miRNA between two groups of samples. It is calculated by taking the log base 2 of the ratio of the average expression values in the two groups. The logFC value is a useful metric for identifying differentially expressed genes (DEGs) in GEO2R. A positive logFC value indicates that the gene or miRNA is upregulated in the older age group, while a negative logFC value indicates that it is downregulated. After gathering this dataset from the GEO Dataset and utilizing GEO2R to analyze it, we identified some common genes that are either overexpressed or under expressed together with the miRNAs that are linked to them. This suggests that these common genes may be primary in AD. It concludes that this extensive

study about miRNAs can be utilized in the future to familiarize with the basic mechanisms underlying the over- and under-expression of genes in AD. They suggest that by looking and evaluating the miRNA targets in the future molecular research, we can tune their therapeutic and regulatory role.

Conclusions: This study provides valuable insights into the molecular mechanisms underlying AD. The identification of common genes that are either overexpressed or under expressed together with the miRNAs that are linked to them is particularly noteworthy. This suggests that these common genes may be key targets for future therapeutic development. The study also highlights the importance of miRNA research in AD. miRNAs are small molecules that play a critical role in gene regulation. By understanding how miRNAs regulate gene expression in AD, we can develop new and more effective treatments for this disease. Overall, this study is a significant contribution to the field of AD research. The findings could have a major impact on the development of new diagnostic and therapeutic tools for this devastating disease.

Keywords: Alzheimer's Disease, microRNAs, Pathways, Bioinformatics Tools, Expression studies

Ancient Antimicrobial Resistance Genes Unearthed: Insights from Pleistocene Permafrost and Ice Core Metagenomes

*Sankaranarayanan Gomathinayagam, Kodiveri Muthukaliannan Gothandam**

School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore.

*gothandam@gmail.com**

Abstract

Motivation: The rationale for this study is rooted in the ancient origins of antibiotic resistance, predating the clinical use of antibiotics. Previous reports have identified antimicrobial resistance genes in Pleistocene permafrost sediment and ice cores, primarily through PCR amplification of targeted loci associated with potential antimicrobial resistance genes. However, in the broader natural context and considering the period when antibiotics were not extensively used, there may exist numerous unidentified antimicrobial resistance genes. Therefore, this study aims to unveil the presence of such genes in metagenome data extracted from ice cores and permafrost samples.

Objectives: To identify and retrieve nucleotide sequences of putative antimicrobial resistance genes from early Pleistocene to late Pleistocene permafrost and ice-core metagenome reads. To computationally model the protein structure of the retrieved ARGs and temporally analyze the evolution of identified putative ARGs.

Materials and Methods: A total of 2225.4 Giga base pairs of data were obtained from publicly available sequence read archives and repositories, encompassing samples aged from 19,000 years to as old as 1,100,000 years. Simultaneously, a redundant database of antibiotic resistance genes was created, combining several publicly available databases. The following databases were employed: Comprehensive Antibiotic Resistance Database (CARD) ARG database (v3.0.9) MarillynR Tetracycline Database (http://faculty.washington.edu/marilynr/, accessed in November 2022), MEGARes database (v3), NCBI Refgene Catalogue for Antibiotic Resistance Genes (v2020.10.12), ResFinder (v4.1.0), Beta-Lactamase Database (http://bldb.eu, accessed in November 2022), NDARO (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA313047,

accessed in January 2023), CBMAR (http://proteininformatics.org/mkumar/lactamasedb/index.html, accessed in November 2022), MUSTARD (http://mgps.eu/Mustard/, accessed in November 2022), and ResFinder FG 2.0. Subsequently, the metagenome reads were mapped to the newly created redundant ARG database using a k-mer alignment (KMA) tool. The reads that mapped to the genes in the database, meeting specific threshold criteria, were subjected to scrutiny, and protein models were constructed.

Results and Conclusions: Antimicrobial resistance is indeed of ancient origin. The study focused on identifying antibiotic resistance genes within the 'antibiotic inactivation' class, revealing the presence of vancomycin resistance genes, beta-lactamase genes, tetracycline resistance genes, and notably, tigecycline resistance genes. It was observed that the abundance of antimicrobial resistance genes (ARGs) was influenced by the depth of metagenomic sequencing and not correlated with the age of the retrieved samples. Additionally, protein modelling demonstrated significant similarities to contemporary ARGs. However, the phenotypic activity of the identified ancient ARGs remains untested.

Keywords: Ancient metagenome, Permafrost, Ice-core, Antimicrobial resistance genes, Pleistocene

Exploring Antimicrobial Compounds in *Streptomyces* Strain VITGV100 (MCC 4961) with Chemical Elicitor

*Madhuri Mukindrao Moon and John Godwin Christopher**

[1]School of BioSciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India

*godwinj@vit.ac.in **

Abstract

Streptomyces strains have long been recognized as prolific producers of bioactive compounds, including antibiotics. *Streptomyces* strain VITGV100 is exhibiting promising antimicrobial compounds. Genome analysis of *Streptomyces* strain VITGV100 reveals thirty-five cryptic biosynthetic gene clusters that remain dormant under standard laboratory conditions which includes peptide, polyketide's, terpenes, siderophores indole and melanin biosynthesis. This study explores the activation of these cryptic biosynthetic genes within *Streptomyces* strain VITGV100, using nutrient broth with 0.5% dimethyl sulfoxide as chemical elicitor at different incubation periods such as 7, 14 and 21 days of culture. Through a systematic screening process, we identified the 7th day of crude extract exhibited significant antimicrobial activity, against selected human pathogens *Bacillus subtilis* (MTCC 2756), *Staphylococcus aureus* (MTCC 737), *Escherichia coli* (MTCC 1687) and *Pseudomonas aeruginosa* (MTCC 3541). Among these pathogens maximum zone of inhibition of 29 mm at 100 µg/ml was recorded against *Staphylococcus aureus* and minimum inhibition zone of 20 mm at 25 µg/ml against *Pseudomonas aeruginosa* were recorded. The extract was analyzed in GC-MS for a detailed chemical characterization. The results revealed that there were 45 distinct peaks from each extract on different days. Only 30 peaks recorded for their controls. The GC–MS data showed the presence of unique compounds such as cyclopentathiazole, thymol lacthydrazide, tyrosol, acetate, trimethylsilyl derivative and few more compounds responsible for antimicrobial activity. Thus, the results of the present study reveal *Streptomyces* strain VITGV100 is an excellent organism for synthesizing antibacterial compounds.

Keywords: *Streptomyces*, secondary metabolites, elicitor, GCMS analysis, dimethyl sulfoxide, antimicrobial compounds.

Network Pharmacology Based Study on the Mechanism of Aloe Vera for Treating Psoriasis

*Vignesh B, Aadil Ahmed Irshath, Thomas Theodore, Anand Prem Rajan**

Vellore Institute of Technology

*janandpremrajan@vit.ac.in**

Abstract

This study aims to analyze the targets of the effective active ingredients of Aloe vera in psoriasis by network pharmacology and molecular docking and to explore the associated therapeutic mechanism. Psoriasis is a condition in which skin cells build up and from scales and itchy, dry patches. There is no cure for psoriasis, but treatments can help to reduce the effect of psoriasis. There are more than 8 million cases per year. In that way there is lot of research in treatment of psoriasis using natural plants. Nature has a source of medicinal plants. Plants such Aloe vera as have adverse effects against psoriasis. The phytochemical constituents of the plants include alkaloids, flavonoids, saponin, phenol, glycosides and tannins. The effective active ingredient of Aloe vera was determined from the TCMSP database and the drug ingredient target network was constructed using the cystoscope software. String database was used to analyse the PPI. Then GO and KEGG analysis was done by the bioinformatics tool and the top 20 key signalling pathways were obtained. The drug ingredient target key pathway was determined using cystoscape software. Finally docking was performed to finalize the binding efficiency of the active ingredient and the target. The results of the study show the active ingredients against the psoriasis disease.

Keywords: Aloe Vera, Psoriasis, TCMSP, Protein-Protein Interaction, Cystoscope, Docking

Targeting Mitochondrial Dynamics: An In-silico Approach for Repurposing Antifungal Drugs in OSCC Treatment

Neha Sivakumar[1], Rohith Raali[1], Harsh Vardhan J[1], and Suresh P.K[2].

[1]Department of Biotechnology, [2]Department of Biomedical Sciences,

School of Bio Sciences and Technology, Vellore Institute of Technology

p.k.suresh@vit.ac.in *

Abstract

Rationale: Drug repurposing for cancer treatment is a valuable strategy to quickly identify existing drugs with known safety profiles that could effectively restrain tumorigenesis and by potentially reducing preliminary screening costs. One in ten cancer patients within the Indian subcontinent suffer from OSCC, primarily due to incessant chewing of betel plant derivatives. Among the different therapeutic modalities, concomitant administration of the chemotherapeutic agent (cisplatin/paclitaxel) is the treatment of choice. Multiple studies have also pointed out to the inefficiency of these chemotherapeutic agents to stun the growth of the neoplasm. Analysis of the oral mycobiome of OSCC patients has projected the intrinsic role of Candida albicans in potentiating OSCC. In the context of treating OSCC, repurposing antifungal drugs emerges as a promising approach, as these drugs could target both the cancer cells and the fungal infections. Cancer cells often have heightened energy requirements, and targeting mitochondrial proteins to disrupt mitochondrial division and induce dysfunction which can lead to cell death, offering a potential method for treating OSCC and improving clinical outcomes for affected individuals. It's also imperative to note that in most cases tumour relapse is mainly owed to mitophagic flux associated with the presence of cancer stem cells (CSC) within the hypoxic niche of the solid neoplasia. Mitochondrial quality control i.e., Biogenesis, fission, fusion and mitophagy are essential to maintain a healthy mitochondrial population. Thereby, deregulation of this essential homeostasis by targeting mitophagy and promoting apoptosis points to Anti CSC therapeutic capabilities.

Objective: The objective of this study is to investigate the potential of antifungal ligands in targeting mitochondrial proteins for the treatment of OSCC. To achieve this, a comprehensive research methodology has been employed, involving docking studies, ligand profiling, molecular dynamics simulations, and with a focus on identifying the most promising candidates for the development of a novel nano-formulation.

Materials and Methods:

A thorough literature search identified 18 mitochondrial proteins associated with mitochondrial dynamics. When experimental structures were available, 3D X-ray crystallography data from the RCSB Protein Data Bank was gathered. In cases where experimental structures were absent, computationally annotated structures from the AlphaFold Protein Structure Database and molecular models from the SWISS-MODEL repository were used. A dataset of 125 unique antifungal ligands was collected from sources like PubChem and the zinc database. Ligand pharmacokinetics and toxicity properties were assessed using SwissADME, ProTox-II, and Prediction of activity spectra for substances tool (PASS). Computational Atlas of Surface Topography of proteins (CASTp) was utilized to identify binding pockets in target proteins. These pocket values were used to create grid boxes for AutoDock Vina and AutoDock 4, estimating the most likely binding locations. Protein structures obtained from RCSB PDB were prepared for molecular docking using CHIMERA and PyMOL to ensure the absence of ligands and additional chains. AutoDock Vina was initially employed to assess the binding affinities of selected ligands with the target proteins. A reference chemotherapeutic drug, paclitaxel, was included for comparison, as it is known for its anti-mitochondrial properties. Ligand-protein combinations with binding affinities lower than that of paclitaxel were eliminated, ensuring the inclusion of only high-affinity interactions. The best ligand for each protein was selected, and AutoDock 4 was used to calculate Estimated Inhibition Constant (Ki) values. A lower Ki value indicates a higher binding affinity to the protein, and Ki values within the picomolar range or below were considered significant. For the ligand-protein combinations with Ki values in the picomolar range or below, GROMACS (2023.2) was employed to conduct molecular dynamics simulations. These simulations provide insights into the dynamic behavior of ligand-protein complexes, further validating the binding affinities and interactions observed in the docking studies.

Results: AutoDock Vina revealed promising binding affinities for various antifungal ligands compared to the reference drug paclitaxel. Notable binding affinity values were obtained for a range of mitochondrial proteins, with the top values being reported for MID51, MID49, VDAC, and others. Molecular docking studies with AutoDock 4 identified five ligand-protein combinations with Ki values in the picomolar range or below, indicating high binding affinity. The proteins MID51, MID49, VDAC, DRP1, and PDK1 exhibited Ki values ranging from 63.60 pM to 715.31 fM. The free energy of binding for these interactions was also assessed, reinforcing their significance. ADME prediction using Swiss-ADME favored itraconazole as a promising drug candidate due to its favorable properties. Certain parameters like molecular weight in the context of Lipinski's rule were not considered highly significant in the context of developing a liposome-encapsulated drug delivery system. The PASS analysis highlighted the antifungal properties of both natamycin and itraconazole. Natamycin exhibited anti-neoplastic and immune stimulant properties, while itraconazole was predicted to have multiple effects, including Cytochrome P450 inhibition and Lanosterol 14 alpha demethylase inhibition with a probability margin over 0.7. Predicted LD50 values for antifungal drugs natamycin and itraconazole were reported, indicating their safety profiles. Both drugs were predicted to be inactive against the tumor-suppressing protein p53 and were not considered carcinogenic or mutagenic. Itraconazole was predicted to be hepatotoxic, although this effect could be managed through dosage control. Natamycin was not predicted to be hepatotoxic, while and both drugs were forecasted to cause immunotoxicity. In order to refine our screening and selection strategy based on the stability and temporal nature of the binding of the ligand to the receptor, MDS based experiments are underway.

Conclusion: In this study, we analysed protein ligand interactions through docking simulations and comprehensive ligand profiling was performed using SwissADME, ProTox-II and PASS. Highest docking score was conferred by MID51 and itraconazole with binding affinity of -16.57 Kcal/mol and a low KI value of 715.31 fM. Taken together with the positive ligand profiling of itraconazole, we believe this potent drug to target the Achilles heel in dysregulating mitochondrial biogenesis and promoting apoptosis of the neoplastic cells. Further validation of these results is warranted for through in-vivo analysis.

Keywords: Oral squamous cell carcinoma, Mitochondrial dynamics, Antifungal drugs, In-silico, Drug repurposing, Molecular docking

Stratifying Breast Cancer Subtypes Using DNA Methylation Markers

Sri Lakshmi Bhavani Pagolu[1], Suba S[2]*, Nita Parekh[2]*

[1]IIIT Hyderabad, [2]International Institute of Information Technology

nita@iiit.ac.in *

Abstract

DNA methylation aberrations are common in cancers and are known to affect. Understanding how these affect the transcriptome can provide insights into subtype specific variations and treatment outcomes. In this study we carried out multi-omics profiling (DNA methylation and gene expression) in breast cancer patients from TCGA-BRCA dataset and propose a novel set of 35 methylation-based prognostic markers for subtype-specific disease stratification. Gene-set enrichment and pathway analysis of the predicted markers using MSigDB and DAVID revealed their role in mammary gland development pathway, various signaling pathways (ERBB2, NOTCH, etc.), and other cancer pathways, and showed clear association with genes affected by hormone receptor status. We show that the reported DNA methylation signature has high discriminative power in classifying breast cancer samples into three molecular subtypes, viz., Luminal, HER2-enriched and Triple Negative, by using six machine learning approaches. An accuracy of 94.12% and MCC of 0.87 is obtained in stratified 5-fold cross-validation for the three-class classification using SVM-RBF.

Keywords: Breast Cancer, Subtype classification, Machine learning, DNA methylation

Renal Sensing of Gut Microbiota Derived- Metabolites in Diabetic Chronic Kidney Disease: An Integrated Approach Using Network Pharmacology and Molecular Docking

G.R. Saranya[1], Pragasam Viswanathan, Renal Research lab, Pearl Research Park, School of Biosciences and Technology, Vellore Institute of Technology-632014*

*pragasam.v@vit.ac.in **

Abstract

Rationale: Metabolites from the gut microbiota serve as defining molecules in the gut-kidney crosstalks. However, active renal sensing gut metabolites and its mechanism in the context of diabetic chronic kidney disease is still unknown.

Objectives: This study employs the computational network pharmacology framework to examine the primary metabolites and mechanistic action of gut microbiota against diabetic chronic kidney disease (DCKD).

Materials and Methods: (1) For this study, the selection of hum

an gut microbial targets was retrieved from the gutMgene database, based on the SMILES format, each metabolite was identified using the PubChem database. (2) Using DisGeNET, GeneCard, NCBI and OMIM database Diabetes CKD targets were identified for Homo sapiens (3) Computational analysis like protein-protein interaction (PPI) networks, gene ontology and Kyoto encyclopedia of genes and genome pathway analysis, identification of metabolites for the key target using molecular docking, evaluation of drug-likeness properties for the key targets were done using ADMET lab.

Results and Discussion: A total of 205 gut metabolites were retrieved from the gutMgene database, and 1304 targets were obtained from Swiss Target Prediction (STP) database, and 1470 targets from Similarity Ensemble Approach (SEA). We obtained 203 targets for DCKD and identified 574 overlapping targets. Following the retrieval of 203 and 222 targets from the gutMgene database, twenty-seven (27) targets were identified as final DCKD targets of metabolites by microbiome. Based on enrichment analysis, host/microbiome protein-protein interaction, gene-disease association results predicted NFK-b1, AKT1, EGFR, JUN and RELA via MAPK signalling pathway would facilitate the progression of DCKD. The gut microbiota-

metabolite-substrate-sample source (GMMSS) network analysis indicates that metabolites originating from the gut microbiome significantly regulate NFkb1 and EGFR gene expression under DCKD conditions through MAPK signalling pathway. Specific bacteria intestinal epithelia, colonic region including Bacteroides distasonis, Bifidobacterium adolescentis, Faecalibacterium prausnitzi A2-165, and Bacteroides vulgatus, build a community that inhibits renal NF-kb1 a target protein expressed at higher glucose concentrations. These bacteria metabolize the unknown substrate to produce Indole 3 propionic acid which were reported to improves the blood glucose level, insulin sensitivity and maintains the intestinal barrier integrity in host. Furthermore, we discovered that the genus Lachnospiraceae prevents the activation of EGFR, a tyrosine membrane protein expressed after renal damage in CKD patients with diabetes mellitus. When NF-Kb1 and EGFR target proteins were docked with Indole 3-propionic acid, a good binding energy affinity was observed.

Conclusions: This work showed that the probiotic gut bacteria Faecalibacterium prausnitzi A2-165 from Firmicutes phylum could enhance the production of the metabolite indole-3 propionic acid. It is also known that these metabolites have renal sensing properties and that they could potentially be used to treat chronic kidney disease in diabetic individual. In addition, it provides thorough insights that could serve as the basis for further research and supports gut health.

Prediction, Design, Molecular Docking and Dynamics Simulation of Novel Antimicrobial Peptides from Aegle Marmelos Against *Staphylococcus aureus*

Rudra Awdhesh Kumar Mishra, Kodiveri Muthukaliannan Gothandam School of Bio-Sciences and Technology, Vellore Institute of Technology, Vellore - 632014, Tamil Nadu, India*

*gothandam@gmail.com **

Abstract

Rationale: Staphylococcus aureus is a nosocomial pathogen responsible for the cause of various range of infectious diseases such as skin infections (acute and sometimes chronic), soft tissue infections; and even life-threatening systemic disease. The emergence of antibiotic-resistant strains such as methicillin-resistant Staphylococcus aureus and multidrug-resistant S. aureus poses a significant challenge to traditional antibiotic therapy. To overcome antibiotic resistance among patients infected with S. aureus infections, there is an immediate need to find alternative therapeutic agents. Several molecules have shown promising results in combating S. aureus infections among resistant strains. However, antimicrobial peptides have emerged as a prominent alternative for combating infections because of their desired characteristics and properties. Antimicrobial peptides have garnered increasing attention as potential alternatives due to their broad-spectrum antimicrobial activity and low propensity for developing resistance.

Objectives: The study's objective is to predict, and design a novel antimicrobial peptide from the Aegle marmelos proteome sequences and validate its anti-MRSA property through molecular docking & dynamics simulation against Staphylococcus aureus.

Materials and Methods: Antimicrobial peptides were predicted from Aegle marmelos using in–silico digestion using five different enzymes such as Chymotrypsin (low & high specificity), trypsin and Pepsin (pH 1.3 and >2) through the EXPASY peptide cutter tool. Then, digested shorter peptide sequences were analyzed for their antimicrobial property by the DBAASP server. The physicochemical property of the predicted AMPs was analyzed using two different tool APD3 (Antimicrobial peptide Database 3) and PROTPARAM where characteristics such as half-life time, instability index, charge, pI, hydrophobicity, and GRAVY value were analysed. Further, ADMET properties were predicted for the predicted AMPs using the ADMETLAB 2.0 tool, before peptide sequences were converted to SMILE characters for the prediction of ADMET property. The structure of the AMPs was predicted using the PEPFOLD3.0 server where the best peptide model was taken for the molecular docking study against the Staphylococcus aureus target protein using the HADDOCK tool. Then molecular dynamics simulation was performed for the complex having the maximum binding affinity for 50 ns using the Desmond package - 2018 version.

Results and Discussions: Different protein sequences were digested using different enzymes to obtain linear shorter peptide sequences, which were further subjected to the prediction of AMP property using the DBAASP server. From the server, we obtained 50 peptide sequences, for which physicochemical properties were characterized using different servers and the non-toxic peptides, and stable were further utilized for structure prediction and molecular docking studies. Molecular docking studies for the peptide were performed against the two-target protein of S. aureus (PDB ID – 2W9S and 7O4M). Protein-peptide docking was performed using the HADDOCK server, and from the docking study, we have two peptides (PI - WGQPKSKITH and P II – GKEAATKAIKEWGQPKSKITH) that were able to interact with the target protein of S. aureus giving maximum binding affinity at the binding pocket of the protein. Protein-peptide interactions were further validated using molecular dynamics simulations for 50 ns using the Desmond package. Dynamics simulation was performed for peptide, protein alone and protein-peptide complex and simulation results were analysed using the Simulation Interaction diagram tool from the Desmond package. From the studies, it has been observed that these two peptide sequences can be further used for in-vitro and in-vivo studies against drug-resistant S. aureus infections.

Conclusions: In summary, the findings from this study highlight the potential of the identified peptide sequences (PI and PII) to serve as valuable candidates for further in-vitro and in-vivo studies targeted at combating drug-resistant S. aureus infections. These peptides are promising therapeutic agents in the fight against antibiotic-resistant strains, offering new avenues for addressing the growing challenges posed by S. aureus infections.

Keywords: Aegle marmelos, Antimicrobial peptides, Drug-resistant Staphylococcus aureus

Structure Activity Relationship Studies of Anacardic Acid Derivatives: Implications in Cancer Biology

*Sandhya Padmakumar *, Geetha Kumar*, Bipin Nair* Amrita Vishwa Vidyapeetham, India*

sandhyap@am.amrita.edu, gkumar@am.amrita.edu, bipin@am.amrita.edu

Abstract

SUMOylation is a post-translational modification pathway that is implicated in the manifestation of various diseases such as cancer, viral infections, diabetes, and neurodegenerative disorders. SUMOylation influences different hallmarks of cancer, such as cell senescence, carcinogenesis, cell differentiation, and apoptosis. Therefore, an imbalance in SUMOylation could affect metastasis, angiogenesis, invasion, and proliferation, making it an important therapeutic target. Hence, the effort to find new anti-cancer agents with better efficacy and fewer side effects by inhibiting the expression of SUMOylation is of major interest. Earlier studies had demonstrated that Anacardic acid (AA), a natural compound from cashew nut shell liquid, exhibited an IC50 of 2.1 µM against small ubiquitin-like modifier E1 (SUMO E1) but had limitations of high lipophilicity and low bioavailability. The present study is therefore focused on understanding the Structure Activity Relationship studies (SAR) of AA derivatives against SUMO E1 in order to identify potential chemotherapeutic agents with enhanced pharmacokinetic properties. The binding energy, inhibition constant, and binding modes of AA derivatives have been studied using molecular docking. Select 129 derivatives of Anacardic acid were screened utilizing molecular docking with AutoDock 4.0. Of these, 61 potential compounds were selected based on their binding energy. Secondary screening was subsequently carried out using the SwissADME virtual platform, where pharmacokinetic properties such as solubility, logP and toxicity were evaluated. This process resulted in the identification of 24 lead compounds, of which the top hits included 2-(Carboxymethyl)-6-hydroxybenzoic acid, 2-(7-Carboxyheptyl)-6-hydroxybenzoic acid, 2-[3-(2,5-Dimethyl-phenyl)-propyl]-6-methoxy-benzoic acid, 2-Methoxy-6-(5-phenyl-pentyl)-benzoic acid and 2-Hydroxy-6-(2-naphthalen-2-yl-ethyl)-benzoic acid with enhanced binding energy and Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties against SUMO E1.

Keywords: Anacardic acid, SUMOylation, SAR studies, Chemotherapeutic agen, SUMO E1

Computational Resources for Understanding and Predicting the Binding Affinity of Protein- Nucleic Acid Complexes

*Harini K and Michael Gromiha**

Indian Institute of Technology, Madras, Tamil Nadu,

gromiha@iitm.ac.in *

Abstract

Rationale/Motivation: Protein-nucleic acid interactions are inevitable in maintaining the homeostasis of cells. It is vital to have a quantitative understanding of these interactions, generally described in terms of the dissociation constant or free energy change of protein-DNA and protein-RNA complexes. With the increase in the experimental data, there was no well-curated database available specific for protein-nucleic acid-binding affinity. In addition, the availability of the experimental data also affects the performance of the binding affinity prediction method. Computationally, protein-DNA binding affinities are predicted using molecular dynamics simulations, statistical methods, and machine learning techniques. Most of these methods are focused on a specific protein-DNA complex or a small set of data. In addition, the performance of the available methods is not uniform in different structural and functional classes of protein-DNA complexes.

Objectives: Hence, we developed a database, ProNAB, which contains more than 20,000 experimental data for the binding affinities of protein-DNA and protein-RNA complexes. Further, we developed a web server, PDA-Pred (Protein-DNA Binding affinity predictor), for predicting the affinity of the protein-DNA complexes.

Materials and Methods: We obtained experimental binding affinity data from a detailed survey of literature and existing/ obsolete databases. We have retrieved research articles and reviews on the binding affinity of protein-nucleic acid complexes using keyword searches. From each article, we manually curated the information about the name of the protein, nucleic acid, complex, experimental conditions, measurement, method, thermodynamic data, literature information, and location of the data in the research article. Each entry in ProNAB is cross-linked with GenBank, UniProt, PDB, ProThermDB, PROSITE, DisProt and Pubmed. It provides a user-friendly web interface with options to search, display, visualize, download and upload the data.

Further, we filtered the binding affinity of protein-DNA complexes from the developed database using the following criteria: (i) experimentally known binding affinity (ΔG), (ii) known 3D structure, and (iii) non-redundant complex structures and obtained 391 protein-DNA complexes. We obtained several structure-based features such as interaction energy, contact potentials, volume, surface area of binding site residues, base step parameters of the DNA, and contacts between different types of atoms. We developed multiple regression equations for predicting the binding affinity of protein-DNA complexes belonging to different structural and functional classes of protein-DNA complexes.

Results and Discussion: Our analysis of the relationship between binding affinity and structural features revealed that the critical factors mainly depend on the number of DNA strands and functional and structural classes of proteins. Specifically, binding site properties such as the number of atom contacts between the DNA and proteins, volume of the protein binding sites, and interaction-based features such as interaction energy and contact potentials are essential to understand the binding affinity. Our method showed an average correlation and mean absolute error of 0.78 and 0.98 kcal/mol, respectively, between the experimental and predicted binding affinities on leave-one-out cross-validation (jackknife test).

Data availability: ProNAB, which is freely available at https://web.iitm.ac.in/bioinfo2/pronab/ (Harini et al., 2022). PDA-pred is available at https://web.iitm.ac.in/bioinfo2/pdapred/. (Harini et al., 2023)

Conclusions: We have developed a database that would aid researchers in gaining insights for understanding the relationship among binding affinity, structure, function, and diseases. Further, we have developed a web server for predicting the binding affinity of protein-DNA complexes, and it will be helpful for large-scale analysis and devising strategies for therapeutic targets.

Keywords: protein–DNA complex, binding free energy, contact potentials, structure-based features

Tackling Drug Resistance in Glioma by Targeting mIDH2^{R140Q} Protein: A Computational Repositioning Strategy

*Poornimaa Murali and Ramanathan Karuppasamy**

Vellore Institute of Technology,

kramanathan@vit.ac.in *

Abstract

Rationale: The Isocitrate Dehydrogenase (IDH) mutation is a hallmark of early gliomagenesis that significantly impacts various human malignancies. Presently, the first-in-class IDH2R140Q inhibitor, AG-221 (enasidenib), has a remarkable selectivity against the target with an IC50 value of 100 nM. However, acquired resistance due to the standard of care treatment and its inability to cross the blood-brain barrier has restricted its use as a mIDH2 inhibitor. Additionally, indirect hyperbilirubinemia and differentiation syndrome were the most frequent side effects of the treatment interventions.

Objective: The prevailing obstacles necessitate the need to develop more potent and selective inhibitors against the mIDH2 protein.

Materials and Methods: In the current study, an integrated virtual screening pipeline was adopted to scrutinize effective compounds from the approved subset of the DrugBank library containing 2715 compounds. The binding characteristics of the compounds were estimated using molecular docking, MM-GBSA analysis and mutational analysis. Further, compounds with enhanced binding affinity were revalidated using machine learning-based (ML) scoring functions. Finally, the conformational sturdiness of the lead molecule was reassured by performing a molecular dynamics (MD) simulation study for 100 ns.

Results and Discussions: A total of 27 compounds were scrutinized from an integrated pipeline with appreciable binding affinity, ΔGbind against the mutational variants. The binding energy of the screened compounds varied from -62.27 kcal/mol to -38.55 kcal/mol. Of note, DB00872 (Conivaptan) exhibits better binding affinity (-42.98 kcal/mol) and a satisfactory toxicity profile. The compound also displayed score values of 6.123 pK units, 0.49 µM, 6.93 kcal/mol and -8.85 CNN affinity in RF, NN, KDEEP and GNINA respectively. The MD simulation study also affirmed the conformational sturdiness of the hit molecule.

Conclusion: We are certain that the outcome of our study will be of immense importance for managing enasidenib resistance in gliomagenesis.

Keywords: Mutated Isocitrate Dehydrogenase2, MM-GBSA calculation, machine learning, molecular dynamic simulations, drug repurposing

Drug Repurposing Strategies for the Management of Triple Negative Breast Cancer: Focus on Indoleamine 2, 3-Dioxygenase and Tryptophan-2, 3 Dioxygenase Targets

*Priyanga Paranthaman and Shanthi Veerappapillai**

Vellore Institute of Technology,

*shanthi.v@vit.ac.in **

Abstract

Rationale: The kynurenine pathway (KP) plays a pivotal role in the dampening of the immune response in many types of cancer, including TNBC. The intricate involvement of tryptophan degradation via KP serves as a critical regulator in immune-privileged regions through the aberrant expression of key enzymes such as indoleamine 2,3-dioxygenase (IDO1) and tryptophan 2,3-dioxygenase (TDO). Despite the availability of navoximod, its poor bioavailability and inadequate efficacy in clinical trials have hampered its utility.

Objectives: In the present study a novel and potent dual-target inhibitor was developed against the vital enzymes, IDO1 and TDO.

Materials and Methods: The investigation employed a comprehensive pipeline of molecular docking and dynamic simulations to evaluate the binding stability of the lead compounds.

Results and Discussion: A total of 2588 compounds from the approved subset of the DrugBank database were proclaimed and subjected to a preliminary evaluation of their toxicity and pharmacokinetic properties. Subsequently, hierarchical molecular docking, prime MM-GBSA and integrated machine learning algorithms precisely identified the potential lead compounds. The antineoplastic activity of the hit compounds was also estimated using the PaccMann server, with values ranging from 0.203 to 24.119μM. Collectively, the results of the hit compound DB06292 strongly reinforced its candidature as an effective anti-cancer agent. Finally, the reliability of the results was corroborated through a rigorous 100 ns molecular dynamics simulation, ensuring the stable binding of the hit against the target proteins.

Conclusions: Considering the favorable outcomes, experimental studies on the proposed hit compound hold promising capabilities in the treatment and management of TNBC.

Keywords: Indoleamine 2 3-dioxygenase (IDO1), Tryptophan 2 3-dioxygenase (TDO), TNBC, Molecular docking, MM-GBSA, Machine learning, Molecular dynamics simulation

In-Silico Design of Antimicrobial Peptides from Bungarus Caeruleus and Molecular Docking & Dynamics Simulation Against Mycobacterium Tuberculosis

Priyanka Singh (Vellore Institute of Technology) and Jayaraman Gurunathan (Vellore Institute of Technology)

gjayaraman@vit.ac.in

Abstract

Rationale: Antimicrobial resistance is the key threat to global health due to high morbidity and mortality. The alteration of bacterial proteins, enzymatic degradation, and change of membrane permeability towards antimicrobial agents are the key mechanisms of antimicrobial resistance. Based on the current condition, there is an urgent clinical need to develop new drugs to treat these bacterial infections. Several studies related to snake venom components have shown to be promising molecules having the ability to interfere with the biofilm formation of bacteria., a study citing the effect of viper snake venom PLA2, an enzyme that is involved in the hydrolyzing of phospholipids of membrane acting as a major protein component in the *Bothrops erythromelas* has shown to inhibit the formation of biofilm against *Acinetobacter baumanii*. Considering the importance of snake venom proteins, we have designed a snake venom PLA2-derived antimicrobial peptide against the drug-resistant Mycobacterium tuberculosis.

Objectives: (1) Prediction, design and characterization of antimicrobial peptides from Bungarus caeruleus using an in-silico approach. (2) Molecular docking and dynamics simulation for the predicted AMPs against Mycobacterium tuberculosis target protein.

Materials and Methods: Prediction of antimicrobial peptides from B. caeruleus by digesting the proteome sequences using different enzymes from the EXPASY PEPTIDE CUTTER tool. Then, digested peptide sequences were subjected to the DBAASP server for the identification of being antimicrobial peptide properties. Then, the physicochemical properties of the peptides were analyzed using the PROTPARAM and APD3 database. Further, the ADMET property of the predicted AMPs was evaluated using the ADMETLAB2.0 server, before the prediction, the SMILES character of peptides was predicted using the PEPSMI server. Then, the structure of the AMPs was predicted using the PEPFOLD3.0 server where the best peptide model was taken for the molecular docking study against M. tuberculosis protein (PDB ID - 5NIO), an EthR transcriptional repressor protein using the HADDOCK tool. Then molecular dynamics simulation was performed for the complex with the maximum binding affinity of protein-peptide complex for 100 ns using GROMACS-2023.1 version. Finally, MMPBSA analysis was analyzed for the enthalpy energy and other interaction energy.

Results and Discussions: Protein sequences were retrieved from the UniProt database for the Bungarus caeruleus (Indian Krait) and subjected to expasy peptide cutter tool for cleavage of larger protein sequences into shorter peptide sequences using different enzymes such as Chymotrypsin (high and low specificity), Tryspin, and Pepsin (pH 1.3 and >2). From the server, we predicted 41 peptide sequences to have antimicrobial properties, for which prediction was performed based on the machine learning algorithm and used the Moon and Fleming scale. Based on the physicochemical properties, peptides were further narrowed down to 11 numbers based on stability and desirable ADMET properties. Docking analysis revealed that the peptide HGATVAVKQVNRCSKNHL effectively binds to the target protein of M. tuberculosis. Docking results were validated using molecular dynamics simulation using GROMACS-2023.1 version and MMPBSA analysis was also performed.

Conclusions: The results of this computational approach support the evidence of the efficiency of these AMPs as potent inhibitors of the specific proteins of Mycobacterium tuberculosis. However, further in-vitro validations are required to fully evaluate the potential of selected AMPs as drug candidates against Mycobacterium tuberculosis.

In-Silico Approach to Explore Anticancer Properties in Gloriosa Superba Derived Compounds Against Prostate Cancer

*Aswathi P and Renuka Suravajhala**

Amrita School of Biotechnology,

*renus@am.amrita.edu**

Abstract

Background: Prostate cancer develops in the prostate gland; all men are at risk of developing prostate cancer. About one in nine men will be diagnosed with it in their lifetime, but only one in 39 men will die from the disease. About 80 percent of men in their 80s have prostate cancer cells. Recent studies show an increase in the incidence of prostate cancer in young men. Prostate cancer in the younger age group is usually undifferentiated and has a poor prognosis.

Objectives: In the present study, 14 pathogenic prostate cancer targeted proteins and considered to investigate molecular interactions with gloriosa superba derived phytochemicals such as (Gloriosine, Colchicine,3-Demethyle gloriosine,3-Demethyle colchicine) and control as 5-Fluorouracil (5-FU).

Materials and Methods: Autodock4 is used for the docking of prostate cancer proteins with gloriosa superba derived compounds. The hydrogen bond and pi-pi interactions of the targeted proteins with ligand is observed by discovery studio and further complex visualizations are observed by Chimera.

Results and Discussion: We have identified Gloriosine, Colchicine,3-Demethyle gloriosine,3-Demethyle colchicine are having –6 to –8 binding energy towards targeted protein (pdb id: 2b2h,3lmy,5ctg and 7m81). We identified the possible hydrogen bonding (LYS A: 254, GLN A:5, GLY A;4, GLN:C207, GLY A:114) with respect to the proteins 7M81 and 5CTG.

Future perspective: We will further go for Insilco studies and Molecular Dynamic Simulations with the selected significant proteins.

Keywords: Prostate cancer, Gloriosa superba, Molecular Dynamics

Proteome-Wide Scanning Approach to Detect rpIE as a Novel Therapeutic Target of M.ulcerans

*Harshini Senthilkumar, Tamil Barathi P, Mohanapriya Arumugam**

Department of Biotechnology, School of Biosciences and Technology, Vellore Institute of Technology, Tamil Nadu, India,

*mohanapriyaa@vit.ac.in**

Abstract

Buruli ulcer is caused by Mycobacterium ulcerans, the third most prevalent bacterial disease after leprosy and tuberculosis. The rise of bacterial strains resistant to treatment raises severe concerns and highlights the need for improved therapies because the current treatment options are limited. Recent developments in whole- genome sequencing combined with chemotherapy, computational biology, and experimental research represent a compelling alternative strategy for identifying deserving therapeutic candidates for treatment. The KEGG Genome database is employed to comprehensively map the metabolic pathways of both the pathogen and the host. Additionally, NCBI BLAST is used to compare essential biological sequence information, while the DEG database assists in identifying dominant proteins within unique and common pathways. The String database aids in the exploration of protein associations and interactions, and the CytoHubba software is exploited to identify the most critical pathogenic drug targets. Finally, rpIE is the most prominent and important node, ranking highest for its centrality measure. This gene has been identified as the main hub gene that could represent the target of a future medication. The hub protein (rpIE) of the Mycobacterium ulcerans has been docked with phytochemicals, which have demonstrated strong inhibitory potential energies. This information can lead to significant advances in testing the efficacy of existing antibiotics compared to new antibiotics, manufacturing drugs with minimal host toxicity, and developing vaccines.

Keywords: Mycobacterium ulcerans, Buruli ulcer, Genome analysis, Hub proteins, Drug-target.

Computational Identification of Biomarker Genes for Hormone-Sensitive Cancers Considering Treatment and Non-Treatment Studies – A Meta-Analysis Approach

*Tamizhini Loganathan and George Priya Doss C**

Laboratory of Integrative Genomics, Department of Integrative Biology, SBST, Vellore Institute of Technology,

georgepriyadoss@vit.ac.in *

Abstract

Breast, ovarian, and endometrial cancers majorly impact women's mortality. These tumors share hormone-dependent mechanisms in female-specific cancers, which support tumor growth differently. Integrated computational approaches may allow us to better detect genomic similarities between these female-specific cancers, helping us deliver more sophisticated diagnoses and precise treatments. This study aims to computationally identify biomarker genes for hormone-sensitive cancers that can aid their diagnosis and treatment. The gene expression profiles of two different types of studies, namely non-treatment and treatment, are considered for discovering biomarker genes. In non-treatment studies, healthy samples are control, and cancer samples are cases. In treatment studies, controls are cancer cell lines without treatment, and cases are cancer cell lines with treatment. The Differentially Expressed Genes (DEGs) for hormone-sensitive cancers were isolated from the Gene Expression Omnibus (GEO) database, and the datasets were analyzed using the online tool IMAGEO. Two Cytoscape apps, CytoHubba and MCODE, were used to identify the hub genes from functional networks using overlap genes from different meta-analysis using MCC and particular hub gene method. Most of the biomarker genes from non-treatment studies are part of mitosis and play a vital role in DNA repair and cell-cycle regulation. In contrast, most of the biomarker genes from treatment studies are associated with cell cycle and cellular senescence. This study discovered a list of biomarkers to help experimental scientists design a lab experiment to further explore the detailed dynamics of female-specific cancer development.

Keywords: Breast Cancer, Ovarian Cancer, Endometrial Cancer, Meta-analysis

Investigation of the Impact of R273H And R273C Mutations on the DNA Binding Domain of P53 Protein Through Molecular Dynamic Simulation

Hephzibah Cathryn.R and Dr. George Priya Doss C Laboratory of Integrative Genomics, Department of Integrative Biology, SBST, Vellore Institute of Technology, georgepriyadoss@vit.ac.in **

Abstract

The P53 protein, a cancer-associated transcriptional factor and tumor suppressor, houses a Zn2+ ion in its DNA-binding domain (DBD), essential for sequence-specific DNA binding. However, common mutations at position 273, specifically from Arginine to Histidine and Cysteine, lead to a loss of function as a tumor suppressor, also called DNA contact mutations. The mutant (MT) P53 structure cannot stabilize DNA due to inadequate interaction. To investigate the conformational changes, we performed a comparative molecular dynamic simulation (MDS) of 1000 ns to study the effect of the P53-Wildtype (P53-WT) and the DNA contact mutations (R273H and R273C) on the DBD. Our research indicated that the DNA binding bases lose Hydrogen bonds (H bonds) when mutated to P53-R273H and P53-R273C during the simulation. We employed tools such as PDIviz to highlight the contacts with DNA bases and backbone, major and minor grooves, and various pharmacophore forms of atoms. The contact maps for R273H and R273C were generated using the COZOID tool, which displayed changes in the frequency of the amino acids and DNA bases interaction in the DNA binding domain. These residues have diminished interactions, and the zinc-binding domain shows significant movements by Zn2+ ion binding to the phosphate group of the DNA, moving away from its binding sites. In conclusion, our research suggests that R273H and R273C each have unique stability and self-assembly properties. This understanding might assist researchers in better comprehending the function of the p53 protein and its importance in cancer.

Keywords: P53, cancer, DNA contact mutations, DNA-binding domain, Zinc Binding domain, Molecular Dynamic Simulations

Evaluation of Ocimum Basilicum for its Antifibrotic and Drug-Likeness Properties – A Computational Pharmacology Approach

K Gayathri, N Malathi, V. Gayathri, J. Jino Blessy and P. A. Abhinand

Sri Ramachandra Institute of Higher Education and Research

gayathri.researcher@gmail.com

Abstract

Oral submucosal fibrosis (OSMF) caused by areca nut chewing is an insidious disorder with the potential to become malignant. The fibrosis of oral mucosa leads to trismus which affects the quality of life. The high occurrence of OSMF in India coupled with the non-availability of a complete cure poses a significant health challenge. Traditionally used herbs possess excellent pharmacological properties and could be re-purposed to treat incurable diseases. Bioinformatics helps expedite the identification of potential drug candidates for disease. The present study aims to identify drug-like phytocompounds of Ocimum basilicum with antifibrotic activity by pathway analysis and gene ontology (GO) analysis. Seven ligands were identified by virtual screening of compounds present in the herb. 520 potential targets were identified for the drug-like compounds and 354 targets were identified for OSMF. Among these targets, 49 common proteins were identified. GO analysis and KEGG pathway analysis of the common proteins reveal various pathways involved in the pathogenesis of OSMF. Network analysis of the PPI network comprising of 49 common proteins identified four key proteins involved in OSMF pathogenesis viz; transforming growth factor- β1 (TGF-β1), epidermal growth factor receptor (EGFR), caspase-3 (CASP-3) and hypoxia-inducible factor-1 (HIF-1A). These findings demonstrate the antifibrotic activity of Ocimum basilicum.

Keywords: Oral submucous fibrosis, Ocimum basilicum, Pathway analysis, GO analysis, Computational pharmacology

Systems and Computational Screening Identifies SRC and NKIRAS2 As Baseline Correlates of Risk (Cor) for Live Attenuated Oral Typhoid Vaccine (TY21a) Induced Protection: An In-Silico Pipeline

Akshayata Naidu [a,1], Varin Garg [a,1], Deepna Balakrishnan [a,2], Vinaya C.R [a,2], Vino Sundararajan [a], and Sajitha Lulu S* [a],*

[a]Integrative Multiomics Laboratory, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, 632014, Tamil Nadu,

*ssajithalulu@vit.ac.in *, svino@vit.ac.in **

Abstract

Rationale: Typhoid is a major public health concern, and the live attenuated typhoid vaccine (Ty21a) is one of the most widely used vaccines against this disease. However, there is significant variation in immune responses to Ty21a, and it is important to identify biomarkers that can predict vaccine responsiveness at baseline.

Objectives: (1) To investigate the molecular basis of variance in immune responses to Ty21a by exploring the baseline immune landscape. (2) To identify potential biomarkers associated with Ty21a vaccine responsiveness at baseline using two distinct computational approaches: Knowledge-based approach: retrieval of differentially expressed genes (DEGs), functional enrichment analysis, protein-protein interaction network construction and topological network analysis of post-immunization datasets before gauging their pre-vaccination expression levels. Data-driven approach: unsupervised machine learning algorithm for data-driven feature selection on pre-immunization datasets. (3) To computationally validate identified biomarkers using supervised machine learning classifiers.

Methods:

• Knowledge-based approach

o Retrieval of differentially expressed genes (DEGs) from post-immunization datasets using the GEO database (GSE100665).

o Functional enrichment analysis of DEGs to identify enriched biological pathways and processes.

o Protein-protein interaction (PPI) network construction of DEGs.

o Topological network analysis of the PPI network to identify hub genes.

o Gauging the pre-vaccination expression levels of hub genes.

• Data-driven approach

o Unsupervised machine learning algorithm for data-driven feature selection on pre-immunization datasets.

o Supervised machine learning classifiers to validate identified biomarkers.

Results: The knowledge-based approach identified three genes (NKIRAS2, SRC, and LOC100134365) that were differentially expressed between vaccine responders and non-responders at baseline. The data-driven approach also identified these three genes as potential biomarkers. Supervised machine learning classifiers using the three identified genes were able to accurately distinguish vaccine responders and non-responders, with 88.8%, 70.3%, and 85.1% accuracy for NKIRAS2, SRC, and LOC100134365, respectively.

Conclusions: This dual-pronged novel analytical approach provided a comprehensive comparison between knowledge-based and data-driven methods for the prediction of baseline biomarkers associated with Ty21a vaccine responsiveness. The identified genes shed light on the intricate molecular mechanisms that influence vaccine efficacy from the host perspective while pushing the needle further toward the need for the development of precise enteric vaccines and the importance of pre-immunization screening.

Keywords: Ty21a vaccine, Immunological profiles, Differentially Expressed Genes (DEGs), Machine learning classifiers, Vaccine responsiveness

Association of CTLA-4 Signal Peptide (T17A) Polymorphism with Rheumatoid Arthritis in the Indian Population: A Case-Control Study and In Silico Analysis

Shamala V And Asha Devi S,

Vellore Institute of Technology

sashadevi@vit.ac.in

Abstract

Motivation: The Cytoplasmic T-lymphocyte-associated antigen 4 (CTLA-4) Rs231775 signal peptide single nucleotide polymorphism (SNP) was targeted and screened for the association of Rheumatoid arthritis (RA) in the Indian population. An insilico approach to predict that CTLA-4 Rs231775 SNP in signal peptide has been efficiently affecting the transportation of CTLA-4 polypeptide chain into the endoplasmic reticulum (ER) by disrupting SRP 54 (Signal Recognition Particle - M domain) protein interaction.

Materials and Methods: The CTLA-4 Rs231775 SNP were genotyped by High-Resolution Melting Analysis (HRMA). The confirmation of SNP was done by Sanger's sequencing. Various in silico tools SignalP 6.0, ConSurf, InterPro, ProtParam, Project HOPE, RNA fold, RNA 3D Composer, NetSurfP 3.0, and SOPMA were used to predict the signal peptide structure and functions. ClusPro 2.0 for Protein-Protein docking and Google Colab for Molecular Dynamic (MD) Simulation.

Results and Discussion: The CTLA-4 Rs231775 SNP AG and GG genotypes are associated with RA susceptibility in the Indian population. An insilico study reveals CTLA-4 Rs231775 SNP (G allele) significantly affects the stability and folding pattern of mRNA structure. In addition, molecular docking and MD simulation reveal that Rs231775 SNP disturbs the SP-SRP recognition pattern, which affects the translocation of CTLA-4 nascent polypeptide chains into ER via the RAPP pathway.

Conclusion: Despite Rs231775 SNP found on signal peptide rather than mature protein, it has a significant impact on CTLA-4 gene expression. We conclude that SP-SRP interaction is important for successful mRNA stability and CTLA-4 protein translation, preventing various autoimmune diseases, especially RA.

Keywords: Rheumatoid arthritis; CTLA-4 gene; Rs231775 SNP; HRMA; Signal Recognition Particle.

Logical Modelling of Gene Regulatory Circuits Involved in CCL20 Induction in Human Organoids Using Systems and Computational Analysis of RNA seq Dataset

Akshayata Naidu, Mohak Bhattacharyya, Dhanya N, Arah Paras Bora, Tejoram Sambrani, Vino Sundararajan and Sajitha Lulu S[1]**

[a]*Integrative Multi Omics Laboratory, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, 632014, Tamil Nadu*

ssajithalulu@vit.ac.in, svino@vit.ac.in

Abstract

Rationale: Protective immune responses against mucosal infections are extremely difficult to evaluate due to inaccessibility of clinical mucosal samples and gap in the understanding of the cross-talk between systemic and mucosal immune responses. Hence, establishment of robust co-relates of protection (CoP) becomes crucial to study vaccine induced immune responses during pre-clinical and clinical trials.

Objective: • To retrieve highly influential mediators involved in the cross-talk between mucosal and systemic immune responses upon infection. • To derive key upstream and downstream regulators of the highly influential genes for the construction of dynamic logical models.

Method: A two-tiered comprehensive analysis of gene expression dataset where gene expression values were derived post-stimulation with different microbes and microbial components. Differentially expressed genes (DEGs) were retrieved and functional enrichment analysis was performed. Thereafter, a curated list of genes was used to construct protein-protein interaction (PPI) static network using STRING database for topological network analysis for identification of hub genes. Mediators associated with the hub genes were identified using Pearson correlation co-efficient and were further validated using multi-variant regression analysis. The derived multi-variant regression models were translated as logical model in order to establish and analyze dynamic network models associated with induction of mediators involved in the transition from mucosal to systemic immune responses in the human organoids.

Results: Topological Network analysis along with literature-based validation revealed CCL20 as a prominent lympho-attractant and hence was taken ahead for the analysis as a CoP candidate. The second tier of the study unveiled major determiners of CCL20 induction which includes interferon alpha and interferon gamma receptors along with IL7, IL17C, IL1B, IL1A, IL20 and IL10RB-DT. Moreover, key regulators of the derived features were found to be RELA, NFKB1, JUN, SP1, STAT3, E3F1, STAT1, HMGA1. The identified mediators (30) and their regulators were further used for the construction and analysis of dynamic network after the construction of the truth table.
Conclusion: Systems and computational analysis of RNASeq dataset reveal key mediators and associated regulatory modules involved in the crosstalk of mucosal and systemic immune responses which can act as potential correlates of protection (CoP) against mucosal infections.

Keywords: CCL20, Logical modelling, Co-relates of protection, RNASeq data, Mucosal immune response

Does Metabacillus Halosaccharovorans Possess Inherent Radiation Resistance? A Comparative Genome Approach

*Sowptika Pal and Anilkumar Gopinathan *, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India,*

*ganilkumar@vit.ac.in **

Abstract

VITHBRA001, a strain identified as *Metabacillus halosaccharovorans*, was isolated from Chavara-Neendakara placer deposit- a high background radiation area situated in Kerala, India. The organism was experimented with induced gamma radiation and was found to be radiation resistant with D10 value of 2.42 kGy, surviving 5 kGy of gamma radiation. Whole genome sequencing study was performed to analyze the radiation resistant genes present in this bacterium and so far, there is no report of this bacterium being radiation resistant. It makes us curious to know if the strain VITHBRA001 has acquired the properties of radiation resistance or this species inherits the resistance capacity. In order to achieve this understanding, we tried to compare the genome of its 4 strains with the help of functional characterization using COG analysis for the entire genome, and with segregated core, accessory and unique proteins of these strains. It could be observed that strains MET-TA-181 and B410 had higher number of proteins in almost all the COG categories including carbohydrate transport and metabolism (G), replication, recombination and repair (L), nucleotide transport and metabolism (F), inorganic ion transport and metabolism (P), cell cycle control, cell division, chromosome partitioning (D) and cell wall/membrane/envelope biogenesis (M) which are reported to be important in radiation resistance. VITHBRA001 and DSM25367 showed comparable gene number in these categories but in-dept analysis shows that DSM 25367 has more of accessory genes while others housing more unique genes. The present assessment encourages us to hypothesize that VITHBRA001, in spite of having the smallest genome of the 4 and lesser number of resistance genes as compared to others, has shown radiation resistance to the tune of 5 kGy. Thus, we are tempted to pose the question if the members of this species may have inherent radiation resistance capabilities.

Keywords: Metabacillus halosaccharovorans, Radiation resistant bacteria, High Background Radiation Area, COG

A Multi-Objective Hybrid Machine Learning Approach-Based Optimization for Enhanced Biomass and Bioactive Fucoxanthin Production in Isochrysis Galbana

Janani Manoch Kumar [1], Ashwani Kumar Cherukuri [2], Annapurna Jonnalagadda [3], and Siva Ramamoorthy [1]**

[1] Professor and Dean, School of Bio Sciences and Technology, Vellore Institute of Technology, rsiva@vit.ac.in

[2] School of Information Technology and Engineering, Vellore Institute of Technology, aswani@vit.ac.in

[3] School of Computer Science & Engineering, Vellore Institute of Technology, jannapurna@gmail.com

Abstract

Rationale of the study: To develop a multi-objective hybrid machine learning-based optimization approach that is fast, robust, scalable and provides automated analytical method for enhanced cell biomass and fucoxanthin production simultaneously in Isochrysis galbana. This study aims to identify the cost-effective optimal production of microalgal pigment within limited time using machine learning algorithms.

Objectives: 1. To analyse and understand the behavioral characteristics and pattern of data using exploratory data analysis. 2. To optimize the concentration of phytohormone for enhanced fucoxanthin production in Isochyris galbana culture. 3. To predict the yield of fucoxanthin when the concentration and number of days was given as input. 4. To assess and validate the multi-objective hybrid machine learning algorithm for enhanced biomass and pigment yield.

Materials and Methods: Development of a machine learning based model by feeding the experimental responses of spectrophotometric data depicting the growth rate, biomass and fucoxanthin production from microalgal cultures supplemented with various concentrations of phytohormones including Salicylic acid, Methyl jasmonate, Giberrelic acid, and Indole Acetic Acid.

Results and Discussion: Fucoxanthin is a valuable carotenoid with a high market value in the pharmaceutical and nutraceutical industries. Thus, the variables for enhanced fucoxanthin production and biomass in microalgae were optimized using statistical models and a feed-forward neural network-based machine learning algorithm with two hidden layers. This study highlights the advantages of using a machine learning approach to achieve optimized pigment production and serves as the foundation for future efforts to convert microalgae as an economically viable source for large-scale production of pigment.

Conclusion: A hybrid machine learning model can accurately and precisely predict the optimized biomass and fucoxanthin production in microalgal culture. The proposed model would aid the rapid estimation of pigment yield from microalgal culture that may reduce laborious, expensive and time-consuming laboratory trial experiments as well as making it a reliable and standardized method.

Keywords: Machine learning, Enhanced fucoxanthin, Microalgae

Phylogenetic Status of a Field Crab (Brachyura: Decapoda) From Pedavedu, Thiruvannamalai (Tamil Nadu): An Integrative Approach Through Molecular Taxonomy, Barcoding and Coding Matrix

Nishita Lal and Anilkumar Gopinathan School of Biosciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu,*

*ganilkumar@vit.ac.in**

Abstract

This paper reports the results of evaluation of the phylogenetic status of a field crab which has been referred to as Oziotelphusa species Muller, 1887, inhabiting the rice farms of Pedavadu, Thiruvannamalai District (Tamil Nadu, India), using morphological and molecular parameters, including bioinformatics softwares. There have been multiple descriptions of this newly discovered Oziotelphusa species in recent times, however there is no evaluation for determining the correct taxonomic classification of this brachyuran crab, especially in light of modern genetic tools. The current work aims to review the taxonomic position of this field crab utilizing morphological criteria, molecular approaches (COI and 18S rRNA gene sequences), barcoding and coding matrix. Physiognomy of carapace and major appendages (including the cephalic appendages, walking legs and abdominal appendages) was considered for creating the morphological coding matrix and subsequent construction of phylogenetic tree. COI and 18S RNA primed sequences were PCR amplified and characterized through DNA sequencer ABI 3130. The sequence information was subjected BLAST and CLUSTAL alignments, construction of phylogenetic tree and barcoding analysis with a view to obtain precise phylogenetic status of the candidate specimen. This comprehensive study reveals the candidate specimen's closer affinity to superfamily Gecarcinucoidea, family Gecarcinucidae, sub-family Parathelphusinae and genus Oziotelphusa. Comparisons of morphological and molecular characteristics with its closest phylogenetic neighbors such as Oziotelphusa aurantia, O. bouvieri, O. stricta, O. biloba, O. wagrakarowensis and O. kerala further suggest that the candidate specimen is a taxonomically distinct entity, qualified to be considered an unreported species. Further, the paper views the relevance of this taxonomic study from conservational angles as well.

Keywords: Brachyuran crabs, Gecarcinucidae, Oziotelphusa sp, molecular taxonomy, barcoding

In Silico Subtractive Proteomics Analysis to Identify the Novel Therapeutic Drug Targets Combating Antibiotic Resistance in Neisseria Gonorrhoeae

*Tamil Barathi Palanisamy, Nirmaladevi Ponnusamy, Faraz Ahmad and Mohana Priya Arumugam**

Department of Biotechnology, School of Biosciences and Technology, Vellore Institute of Technology, Tamil Nadu, India,

mohanapriyaa@vit.ac.in

Abstract

Gonorrhea, a sexually transmitted infection caused by the Gram-negative bacterium Neisseria gonorrhoeae, remains a significant worldwide health concern, persisting despite ongoing endeavors to eliminate it. In both genders, gonorrhea may manifest as urethritis in men and as cervicitis or urethritis in women. Gonorrhea is manageable and can be effectively treated using certain antibiotics. Nevertheless, the increasing prevalence of antibiotic-resistant strains of N. gonorrhoeae is posing a growing challenge in treating gonorrhea, heightening the risk of it becoming resistant to treatment. The current in-silico investigation seeks to discover potential novel drug targets for combating N. gonorrhoeae infection, utilizing bioinformatics methodologies. The foundational gene set encompasses 851 scrutinized proteins sourced from seven distinct strains of N. gonorrhoeae, retrieved from the UniProt database. Furthermore, the CD-HIT analysis identified unique sequences from the complete proteome. Subsequently, these non-redundant proteins underwent a standalone Blast against the human proteome, resulting in the identification of 232 proteins as non-homologous. Additionally, non-homologous and essential proteins underwent scrutiny in the KEGG Pathway Database, resulting in the identification of six distinct proteins. Furthermore, the subcellular localization of these specific proteins was verified, and cytoplasmic proteins were selected for the analysis of druggability. Subsequent to that, molecular docking was employed using PyRx software to assess a collection of FDA-approved drugs from the DrugBank database. This screening aimed to evaluate their binding affinity with newly identified druggable targets and receptor proteins. The top two compounds for each receptor protein were chosen based on criteria such as binding affinity and the most favorable conformation.

Lastly, analyses of absorption, distribution, metabolism, excretion, and toxicity (ADMET) were executed using the SWISS ADME and Protox tools.

Keywords: gonorrhea, antibiotic resistance, CD-HIT, non-redundant proteins, KEGG pathway, Binding affinity.

Zero Inflated Conway-Maxwell Poisson Model: An Application to Cross-Sectional Microbiome Data

Abhiram D. B.[1], B. Binukumar[1] and Gokulakrishnan Kuppan[2]

[1]Department of Biostatistics, [2]Department of Neurochemistry, NIMHANS, Bengaluru

Abstract

This study investigates the efficacy of two zero-inflated regression models, namely the Zero Inflated Negative Binomial Model (ZINB) and the Zero Inflated Conway-Maxwell Poisson Model (ZICMP), in handling count-based microbiome data characterized by zero inflation. Utilizing a secondary dataset from a NIMHANS study involving 60 subjects categorized into drug-naïve or risperidone-treated Schizophrenia patients and healthy controls, the research focuses on 16s ribosomal RNA (16s rRNA) gene sequence-based exploration of gut microbiome differences as a potential non-invasive biomarker for Schizophrenia. To address overdispersion and excess zeroes in the count data, both ZINB and ZICMP models were applied and evaluated based on criteria such as Akaike Information Criteria (AIC), Vuong's Test, and Rootogram plots. The results indicate comparable performance between the two models, with similar AIC values of 190.5, and negligible differences in Root Mean Squared Error (RMSE) values (288.56 for ZINB and 288.67 for ZICMP), suggesting similar predictive accuracy. In conclusion, both ZINB and ZICMP models exhibit comparable fits and predictive performance for the examined microbiome dataset.

Keywords:16S ribosomal RNA, Microbiome, ZINB, ZICMP, Vuong's Test, Rootogram, RMSE

Insight Into the Bacterial Gut Microbiome of Penaeus vannamei Fed with Functional Feed Additives Lactiplantibacillus plantarum by Amplicon Sequencing

N. Lalitha, K. Ambasankar, T. Sivaramkrishnan and P. N. Suganya ICAR- Central Institute of Brackishwater Aquaculture, Chennai–600 028*

lalitha.n@icar.gov.in, lalithaciba@gmail.com

Abstract

Rationale/Motivation: International shrimp aquaculture industry is booming rapidly, with emphasis on minimizing antibiotic usage on shrimp farming. Aquaculture productivity is intrinsically linked to health and the gut microbiota is rapidly emerging as a key an indicator of shrimp health. Microbes which colonize constantly, known as the gut microbiome, communicate to the host they inhabit and support a variety of essential host activities such as enzymatic digestion and competitive exclusion of pathogens, thus enhancing the host immunity. Hence, gut microbiome modulation is a promising idea for aquaculture and has been presented as a potential replacement to the use of broad-ranging antibiotics in disease management.

Objectives: The present study was carried out to investigate the gut microbiome of the shrimp following dietary feed additive of the L. plantarum probiotic and paraprobiotic in P. vannamei.

Materials and Methods: The experiment comprised of three treatment group, Group I fed with L. plantarum probiotic 1011 CFU/g of the feed (LLP), Group II fed with L. plantarum paraprobiotic (DLP) and Group III Control (CON) fed with basal diet without feed additive, in P. vannamei for the period of 45 days. At the end of the experiment gut of the shrimp (n=3) was collected from each group, DNA extracted and gut microbiome analysis was carried out using Illumina Miseq sequencing of 16S RNA V3-V4 hypervariable regions.

Results: Alpha diversity (observed OTUs, Chao1 index, Simpson and Shannon index) were evaluated and it was found that there exists no significant (P >.05) difference in diversity between two diet groups and control (Observed OTUs ranged from 298 to 386.66 OTUs; Chao1 index ranged from 307.42 to 416.05; Simpson index range from 0.93 to 0.97; Shannon index 6.18 to 7.26). There were no significant changes between the two diet groups for any of the indicators (P >.05). The principal coordinate plots based on beta diversity index by unweighted unifrac analysis

of gut of P. vannamei depicted distinct bacterial community profile between different treatments with one sample in each group out layered. Current finding, regardless of diet, Proteobacteria have been demonstrated to be the most prevalent phylum in the gut microbiome. Tenericutes, Bacteroidetes, Planctomycetes, and Verrucomicrobia were also found but were not influenced by the diet. Core microbiome analysis revealed, there have been 72 OTUs, 11 OTUs and 59 OTUs in probiotic, paraprobiotic and control, respectively. Rhodobacteraceae 37.69% and Flavobacteriaceae 15.74 %were beneficial core microbiome bacterial signatures observed predominantly in the probiotic supplemented diet. Pseudoalteromonadaceae 10.64% observed predominantly in the paraprobiotic group. Unique OTUs observed in the probiotic supplemented diet Lutimonas1.68%, Ruergia 0.76%, Cellulomonadacea 0.44%. The number of OTUs at species level specific to the probiotic, paraprobiotic, control, are 17(9.5%), 40 (22.3%) and 30 (16.8%), respectively.

Discussion: In the current study, beta diversity depicts distinct gut bacterial community between two different diets. Proteobacteria is the dominant phyla in the all the diets. Probiotic fed diet showed dominance of Rhodobacteraceae and Flavobacteriaceae core microbiome whereas it was Pseudoalteromonadaceae in paraprobiotic diet fed shrimp. Probiotic supplemented shrimp depicts unique OTUs Lutimonas, Ruergia and Cellulomonadacea. The feed additive in the shrimp diet has manipulated the gut microbiome with dominance of the beneficial bacteria and presence of unique OTUs (Albores et al., 2017; Cheng et al., 2019; Xie et al., 2019; Landsman et al., 2019) compared to diseased shrimp microbiota (Huang et al., 2020; Hou et al., 2018)

Conclusions: It was found that the probiotic supplementation modulated the host gut microbiome with relative abundance of beneficial bacteria and unique bacterial taxonomic signatures in this species.

Keywords: Gut Microbiome, Penaeus vannamei, Amplicon sequencing, Lactiplantbacillus plantarum

Poster Presentation

Evaluating the Performance of Machine Learning Methods for Predicting Mortality in Intensive Care Unit Patients

Alexander (Young-Joon) Ahn[1]

[1]*University of Wisconsin-Milwaukee*

aahn5035@gmail.com

Abstract

Machine learning methods are increasingly being used for building diagnostic models in clinical settings to identify patients who are at a higher risk of mortality. Recent studies have shown that ensemble tree-based learning methods, provide an alternative non-parametric approach compared to traditional methods for building predictive models in high-dimensional datasets. In this study, we evaluated the performance of logistic regression, random forest, XGBoost, and LGBM (leaf-wise tree-based learning algorithm) for identifying ICU patients with a 28-day mortality risk at the time of hospital admission. The case study data originates from a subset of publicly available data from the Medical Information Mart for Intensive Care (MIMIC) II database. The performance of different methods was evaluated using prediction error curves. The results show that the XGBoost classification method achieved the best prediction accuracy for classifying survivors vs. non-survivors with (cross-validation area under the curve; AUC=0.86). The top features for predicting death at the time of ICU admission included age, simplified acute physiology score (SAPS), and serum sodium levels at admission. These results can help predict which patients are likely to die within 28 days of ICU admission so that healthcare professionals can design & implement optimal treatment strategies to improve patient outcomes. All analyses were conducted using the AutoAI tool in IBM Watson Studio.

Keywords: Machine learning, mortality, predictive modeling, ICU

Structure-Based Drug Designing Towards the Identification of Potential Anti-Bacterial for Acinetobacter Baumannii by Targeting Penicillin-Binding Protein

Soumya Ranjan Mahapatra, Jyotirmayee Dey, Namrata Misra, Mrutyunjay Suar

School of Biotechnology, KIIT Deemed to be University, Bhubaneswar-751024, India

soumyaranjanmahapatra685@gmail.com *

Abstract

According to the world health organization (WHO) reports, Acinetobacter baumannii is a nosocomial bacterial pathogen and is responsible for a wide range of diseases including pneumonia, necrotizing fasciitis, meningitis, and sepsis. The enzymes involved in the peptidoglycan biosynthetic pathway are critical for the survival of this bacterium. PBPs remain attractive targets for developing new antibiotic agents because they catalyse the last steps of the biosynthesis of peptidoglycan, which is unique to bacteria and lies outside the cytoplasmic membrane. The objective was to identify natural molecules that fit best at the substrate binding pocket of the protein and interact with functionally critical residues. Here, we utilized the structure-based virtual screening (SBVS) technique to identify the promising lead molecules against PBP protein using computational approaches. During the High-throughput Virtual screening (HTVS) analysis, we started with 25,372 molecules against the PBP model, among these; only 3284 molecules could be considered suitable for further steps. Finally, only sixty molecules were able to pass Lipinski's and ADMET properties. The four best hits were chosen based on docking scores. Selected top-ranked each four compounds underwent molecular dynamics (MD) simulations for 100ns each to validate the docking interactions where all four compounds with zinc database id: ZINC27742, ZINC30015, ZINC30764, and ZINC44583 are highly supported by root-mean square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), and Hydrogen bond analysis. Further, the MM/PBSA binding free energy analysis was performed for four ligands bound PBP structure. From the study, we have found ZINC27742, ZINC30015, ZINC30764, and ZINC44583 to be a potential inhibitor having all the characteristics of a promising drug candidate.

Keywords: Acinetobacter baumannii, Drug, Penicillin-Binding Protein, Virtual screening, Molecular Dynamics simulations

Immunoinformatics Aided Designing of a Next Generation Poly-Epitope Vaccine Against Pseudomonas Aeruginosa Targeting Needle Tip Protein

Jyotirmayee Dey, Soumya Ranjan Mahapatra, Namrata Misra, Mrutyunjay Suar School of Biotechnology, KIIT Deemed to be University, Bhubaneswar

jmdey1997@gmail.com

Abstract

According to a World Health Organization report, Pseudomonas aeruginosa is one of the world's most deadly organisms, producing hospital-acquired pneumonia, surgical infections, bacteremia, and other potentially fatal diseases. However, no effective treatment or countermeasure to treat the infection has yet been identified. The needle tip PcrV protein is an important protective antigen against Pseudomonas infection. In this research, we used an immunoinformatics and molecular docking approach to construct a multiepitope vaccine of eight conserved, highly antigenic, non-allergenic, and non-toxic epitopes from needle tip protein to provide treatment against P. aeruginosa infection. The selected epitopes were then joined together using suitable linkers, and an adjuvant was added to the N-terminal to boost the immunogenicity of the vaccine. The vaccine protein was further tested for allergenicity, antigenicity, and physiochemical characteristics, and it was found to be safe and immunogenic. The best 3D model of the subunit vaccine was generated using Robetta software. Disulfide engineering in a location of high mobility was used to improve the stability of the vaccine protein. The modeled structure was successfully docked to antigenic receptor TLR-4. To determine flexibility and conformational changes, a molecular dynamics simulation of 100 ns was run. Increased levels of antibodies, INF-γ, IL-2, TGF-β, B-cells, CD4+, and CD8+ cells were seen in immune simulation experiments, indicating the induction of primary, secondary, and tertiary immune responses. Finally, to ensure vaccine expression and translation efficiency within an expression vector, an in-silico cloning approach was used. The proposed polyepitope subunit vaccine may be able to trigger both cellular and humoral immune responses. However, experimental validation of the suggested construct is required to confirm its safety and immunogenic profile.

Keywords: Pseudomonas aeruginosa, Vaccine, Bioinformatics, Molecular docking, MD simulation

Unlocking the Molecular Landscape of Hepatocellular Carcinoma Arising from Non-Alcoholic Fatty Liver Disease: WGCNA and Multi-Omics approach

Subhajit Ghosh and Subarna Thakur

University of North Bengal

subho7jit@gmail.com

Abstract

Motivation: The prevalence of Non-alcoholic fatty liver disease (NAFLD) has surged in recent years, closely paralleling the escalating rates of obesity and diabetes. NAFLD condition is characterized by the accumulation of fat within liver cells. Though initially perceived as benign, it is now considered to harbor the ominous potential to evolve into non-alcoholic steatohepatitis (NASH), liver cirrhosis, and, ultimately, hepatocellular carcinoma (HCC). Significantly, the associated risk of hepatocellular carcinoma (HCC) stemming from NAFLD has been chronically underestimated, primarily due to its insidious progression and poor prognosis. In light of this, there exists a compelling need to identify the key modulators that drive the transition of HCC from NAFLD. By identifying these pivotal factors, we aspire to enable early diagnosis of this perilous progression. Such early detection can help in prevention and treatment, potentially offering a lifeline to countless individuals at risk of developing HCC as a consequence of NAFLD.

Objectives: This study aims to utilize Weighted Gene Co-Expression Analysis (WGCNA) to pinpoint co-expressed genes linked to the transition from Non-Alcoholic Fatty Liver Disease (NAFLD) to Hepatocellular Carcinoma (HCC), with a specific focus on identifying core hub genes and unraveling their potential roles in HCC development through investigations of mutation patterns, epigenetic modifications, miRNA-mRNA interactions, and tissue-specific protein expression data. Additionally, this work intends to assess the prognostic potential of these core hub genes by conducting survival analysis utilizing multivariate cox regression analysis.

Materials and methods: Gene expression datasets of three studies were utilized, comprising five different stages – Control, healthy obese, steatosis or fatty liver, NASH, and HCC. The common genes among these studies were used in the final data matrix, and after initial data normalization, Weighted gene co-expression analysis (WGCNA) was employed for the identification of co-

expressed genes. Subsequently, a protein-protein interaction network (PPI) was utilized to identify twenty-five hub genes in the co-expressed gene module for HCC. Next, Tissue-specific expression information of the hub genes was analyzed. Concurrently, an exploration of mutation-associated data and miRNA-mRNA interactions was carried out to unveil the underlying mechanisms responsible for changes in gene expression. Additionally, to examine the prognostic potential of these hub genes, survival analysis was done based on the multivariate cox regression method.

Results and Discussion: From WGCNA analysis results, the module with the highest positive correlation with HCC showed that most of the genes are linked to vital metabolic pathways such as Carbon metabolism, Oxidative phosphorylation, Diabetic cardiomyopathy, Glyoxylate and dicarboxylate metabolism, and Fatty acid metabolism. The protein-protein interaction network analysis was based on genes from the HCC module, which yielded 25 core hub proteins. Among the 25 core hub proteins within the HCC module, genes such as ATP5F1A, ATP5F1B, ATP5F1C, AGXT, EHHADH, UQCRC1, PCCB, ALDH9A1, ACADM, HSD17B4, FH, NDUFS2, MYC and SDHB exhibited significant levels of expression in hepatic cells. Furthermore, epigenetic, mutation-associated data and miRNA-mRNA interactions indicated ACADM, MYC, NDUFS2, PDHB, ALDH9A1, SUCLG1, EHHADH, HSD17B4, and SDHB are implicated in dysregulated cellular metabolism, oxidative stress resulting from mitochondrial beta-oxidation dysfunction, aberrant peroxisomal activity, and the potential activation of proto-oncogenes, cell growth signaling, prolonged survival, and tumorigenesis. Survival analysis indicates the worst overall survival in the case of ACADM, PDHB, and SUCLG1 genes.

Conclusion: In this study, the WGCNA and PPI network analysis approach identified a total of 25 hub genes that are potentially involved in HCC development from NAFLD. Out of these, nine hub genes were pinpointed, displaying elevated expression levels within hepatic cells. Additionally, these genes displayed a heightened occurrence of mutations, along with epigenetic modifications and post-transcriptional changes. Furthermore, survival analysis highlights the potential of ACADM, PDHB, and SUCLG1 as promising candidates for prognostic biomarkers in the context of NAFLD-mediated HCC development.

Keywords: NAFLD, NASH, HCC, WGCNA, Mutation analysis, PPI, Hub genes, Biomarkers

Genomic Investigation of the Heat Shock Transcription Factor Gene Family Leveraging the Secrets of Drought Resistance in Black Pepper.

S. Mukesh Sankara Sona Charlesa Hibbah Na Soumya SLb Shivakumar MSa K.V. Sajia T. E Sheejaa aICAR-Indian Institute of Spices Research Kozhikode, Kerala, India bCentre for Post Graduate Studies and Research in Botany, St. Joseph's College (Autonomous), Devagiri, Kozhikode, Kerala, India

sonayuv@gmail.com

Abstract

Black pepper (*Piper nigrum* L.; 2n = 52; Piperaceae), known as the "king of spices," is cultivated in more than 30 tropical countries and holds global significance due to its extensive use in diets, medicines, and preservation. The productivity of black pepper in traditionally cultivated areas has been declining primarily due to drought stress, which is exacerbated by global warming and climate change. Black pepper is naturally sensitive to drought, and studies reveal that the crop in its reproductive stage requires up to 3,000 mm of water, making it highly susceptible to water deficits that often result in plant mortality. Various mechanisms have been proposed to address this challenge to protect plants from drought stress, including the induced systemic tolerance (IST) process. Within this context, the HSF gene family, encoding specific chaperones, plays crucial roles in multiple abiotic tolerance processes. In our study, we comprehensively analyzed the Hsf gene family in black pepper through whole-genome identification and characterization, which is otherwise poorly understood. A total of 41 Hsf genes were identified in the P. nigrum genome, and these genes were unevenly distributed on 19 chromosomes. Detailed annotation using the HEATSTER database for the different domains and motifs indicated 19 belong to HsfA class, 21 belong to HsfB class, and the remaining one got included in HsfC class. Afterward, the neighbor-joining method constructed a phylogenetic tree indicating a similar clustering pattern. The Hsf genes in the same group had similar gene and protein structures. Selection pressure analysis indicated that duplicated genes underwent purification selection during evolution. This is the first insight into this gene family and the results provide some gene resources for future gene cloning and functional studies toward the improvement in stress tolerance of the crop.

Role of CTC1 and DNA Repair Proteins at the Telomere in Cancer

Apurwa Mishra and Trupti Patel

Vellore Institute of Technology, Vellore

tnpatel@vit.ac.in

Abstract

The CTC1 protein functions as a part of the CST protein complex, a heterotrimer consisting of CTC1–STN1–TEN1 which promotes telomere DNA synthesis and is required for multiple steps in telomere replication and synthesis of the complementary C-strand. CTC1 is an important component in regulating proper telomere length and preserving telomeric integrity. The study aims to learn more about the way CTC1 interacts with proteins involved in double-strand break repair facilitating in maintaining genomic homeostasis. Interactomes were created between CTC1 and various repair proteins. Strong interactions were found between proteins of Non-homologous End-joining and Homologous Recombination. Non-synonymous SNPs were mined from COSMIC and detailed mutational profiling was carried out. Out of 379 non-synonymous mutations, 2 were found to be highly detrimental. HOPE and CONSURF were used to analyze the mutations to better understand their effects on protein structure and evolutionary conservation. CTC1 detrimental SNPs were docked with wild-type repair proteins using various in-silico tools to further understand how CTC1 mutations affect gene repair. Shift in binding energy hints at the altered dynamics between CTC1 and its interacting partners. Because genomic instability is a major driving force in the development of cancer, aging-related diseases, and other complex diseases, these changes may result in disrupted interactions with repair proteins and the downstream pathway, leading to the onset of Alternative lengthening of telomere (ALT) and genomic instability, hallmarks of cancer.

Keywords: telomere, CTC1, DNA repair, cancer, in silico profiling

Mutational Profiling of PPARGC1A and Its Role in Diabetes, Obesity and Cancer

Shrikirti Anand and Trupti Patel

Vellore Institute of Technology, Vellore

tnpatel@vit.ac.in

Abstract

Type 2 Diabetes mellitus (T2DM) and Obesity are two metabolic disorders caused by a range of genetic and environmental factors. Currently we have many research evidences suggesting that the Peroxisome proliferator activated receptor gamma (PPARG) gene is significantly involved in the regulation of T2D as well as in the growth and differentiation of adipocytes. Due to this, PPARG has gained popularity as an important therapeutic target for diabetic and obesity therapies. A transcriptional coactivator of PPARG, the Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PPARGC1A) plays an essential role in cellular energy metabolism. PPARCIA acts on the nuclear receptor PPARG, thus interacting with transcription factors. PPARGC1A gene, owing to its dual role in lipid and glucose metabolism has often been associated with cancer as modifications and adaptations in cellular metabolism are hallmarks of cancer cells. In this study, we attempt to establish that mutant PPARGC1A alter its interaction with DNA repair genes, thus resulting in mutagenesis eventually leading to increased cellular proliferation and genomic instability. Using in silico approach, interactome of PPARGC1A and various DNA repair proteins suggested its interaction with repair proteins involved in Nucleotide Excision Repair and DNA damage response mechanism. Detailed mutational profiling of PPARGC1A gene suggested that 2 out of 439 non-synonymous Single Nucleotide Polymorphism (nsSNPs) were detrimental for protein function. Many of these mutations were evolutionary conserved making them more impactful at the site of mutation. Docking of mutant ppargc1a with wild type repair proteins showed modified dynamics which may cause long term adverse impact on genome integrity. Further investigation on PPARGC1A regulation will help us in identifying novel mechanisms underlying the relationship between T2D, obesity and cancer.

Keywords: Cancer, Type 2 Diabetes, Obesity, Mutational profiling, DNA repair

Establishing a Computational Screening Framework to Identify Environmental Exposures Using Untargeted GC-HRMS

Hayoon Kim, Radnor High School

hayoonk481@gmail.com

Abstract

E-waste exposure to humans has been an issue for both developing countries. As technology has advanced, the production of waste has increased, and the toxicity of this e-waste has been giving workers adverse health effects. Liquid crystal monomer (LCM) is one of the toxic organic compounds within e-waste and has been researched extensively to figure out the degree to which it affects human health. Because this is a global concern, countries are currently trying to implement effective solutions.

We have identified previously written research papers related to the topic of e-waste and its effect on human health primarily by searching them in Pub Med, using specific search terms. Then, we filtered out the non-related papers by filtering out the preprints, retracted publications, and other animals (excluding humans). Moreover, we filtered out the papers by dividing them based on the categories of include, exclude, and review. This way, we were able to have a list of only the related papers for our review. After developing this list, we looked over the research papers in our list and reviewed what has been discovered about e-waste exposure and the harm and also about what should be done to further solve the issue.

The research papers proved the detrimental effects of e-waste on the human body. Specifically, they proved that LCM plays a major role in being the toxic component inside e-waste. Organizing sources based on search terms and filtering them through two different filtering methods allows relevant research papers to be gathered efficiently. Moreover, future studies would have to reveal further details of the e-waste management and LCM from the e-waste.

Keywords: E-waste, Electronic waste, Human health, Toxicity, Liquid crystal monomer (LCM)

Triphala-Induced Oxacillin Sensitivity of Methicillin Resistant Staphylococcus Aureus Mu50 Strain: Insights from In Silico Studies

Sharadha S, Aathithya Diaz, Subbiah Thamotharan, and Vigneshwar Ramakrishnan

SASTRA Deemed to be University, Thanjavur

vignesh@scbt.sastra.edu

Abstract

Finding innovative strategies to prevent bacteria from developing drug resistance is crucial since antibiotic-resistant bacteria are constantly evolving. Triphala, an ayurvedic formulation, has been shown to have synergistic effects with antibiotics. Specifically, it has been shown that the triphala increases the sensitivity of Methicillin Resistant Staphylococcus aureus (MRSA) to oxacillin. In this study, we attempted to delineate the molecular mechanism by which triphala increased the sensitivity of MRSA using a computational approach involving molecular docking of the phyotchemicals of triphala against the proteome of the MRSA Mu50 strain. A total of 66 phyotchemicals from the fruits of Phyllanthus embilica, Terminalia belerica and Terminalia chebula that constitute triphala were obtained from the IMPPAT database. The proteome of MRSA Mu50 strain was obtained from the UniProt database. The proteins which had the corresponding structure in the PDB database were identified resulting in 127 protein targets. Molecular docking was performed against these targets using Schrödinger software suite. Docking studies revealed that the phytochemicals 1,3,6-tri-O-galloyl-beta-D-glucose (IMPHY000783) and trigalloylglucose (IMPHY013560) had the highest binding affinity to the proteins dehydropantonate 2-reductase (3G17) and aldehyde dehydrogenase (3TY7), respectively. Pathway analysis showed that these proteins are involved in pantothenate biosynthesis pathway and aldehyde metabolism, respectively. Taken together, our results show that triphala increases sensitivity to oxacillin by modulating proteins that are crucial to the survival of the bacteria and not directly on the antibiotic resistance conferring proteins.

Keywords: Ayurvedic formulation triphala, Antibiotic resistance, Synergistic effect with antibiotics, Increase sensitivity, Molecular docking, Binding affinity, Pathway analysis.

In Silico Determination of Multidrug-Resistant (MDR) Genes in 2023 Sequenced Klebsiella Pneumonia's Genome

Sandeep Kumar C

Jain (Deemed-to-be University)

drsangene18@gmail.com

Abstract

Klebsiella pneumoniae is a Gram-negative enteric bacterium that causes urinary tract and nosocomial infections; widespread of multidrug-resistant (MDR) strains of K. pneumoniae are reported across the globe. MDR in K. pneumoniae are becoming more common in clinical settings, which is a serious worldwide health concern. Both extended-spectrum -lactamases (ESBL) and carbapenemases capable of hydrolyzing newer carbapenem medications can be found in K. pneumoniae. Resistance to other antibiotics, such as fluoroquinolones, aminoglycosides, trimethoprim, and sulfamethoxazoles, is frequently associated with ESBL-producing K. pneumoniae. Clinical isolates of K. pneumoniae have been shown to exhibit all three broad mechanisms of drug resistance in Gram-negative bacteria: the acquisition of novel antibiotic catalytic genes, mutations of antibiotic targets and membrane proteins, and differential expression of specific genes such as those for efflux pumps that mediate drug effects.

The application of bioinformatics in the study of antimicrobial resistance in microorganisms involved in human pathology has been enabled by the implementation of certain technologies such as whole genome sequencing (WGS) or mass spectrometry, as well as creation of national and international databases that include and gather data on MDRfrom around the world. In the present study, VRprofile web server was used to identify the different MDR genes in the pan-genome of K. pneumoniae.

Keywords: Klebsiella pneumoniae, multidrug-resistant (MDR) strains, whole genome sequencing, antimicrobial resistance, VRprofile web server, Horizontal gene transfer

RNA-seq and sRNA-seq Analysis in Black Pepper Reveals Potential Regulatory Transcripts in Drought Tolerance

Sona Charles, Sheeja Te, Priyanka Behera, Fayad A, Krishnamurthy Ks and Johnson K George,

ICAR-Indian Institute of Spices Research, Kozhikode

sonayuv@gmail.com

Abstract

Black pepper (Piper nigrum), also known as "King of Spices" is widely known for the panoply of metabolites with potential medicinal and biological properties. Transcriptome-wide studies in black pepper can uncover key genes, small RNAs and pathways that contribute to its stress tolerance. In this study, rooted cuttings of black pepper (genotype: IC 317179) were grown under normal as well as water stressed (15 days stress by withholding water) conditions. RNA was isolated from the leaves of both normal and water stressed plants and Illumina HiSeq 2000 platform was used for the paired-end sequencing. Raw reads of RNA as well as sRNA were pre-processed using FASTQC. De-novo and reference-based assembly of drought stressed transcriptome were performed. Thirty-one differentially expressed miRNAs (log fold change |>1| and p-value<0.05) were filtered. Identified miRNAs depicted stable stem-loop structures and high sequence conservation among other plant species such as Arabidopsis. The minimum free energy index of the sequences ranged between 0.63 to 0.80 and AU composition of pre-miRNA ranged between 32% to 64%. Putative target transcripts of miRNAs were also predicted. The differentially expressed miRNA were found to target genes specific for stress response. Functional annotation of the targets revealed that the miRNAs regulate drought responsive genes such as ribosomal protein S27a, catalase isozyme 1, etc. The target genes were mainly associated with stress and were found to be involved in pathways related to carbohydrate metabolism, translation and ribosome biogenesis. The findings of the study provide new insights into miRNA mediated regulatory networks of drought response in black pepper. The insights gained from transcriptome analysis will be validated to uncover key interactions between miRNAs and genes in pathways that contribute to its stress tolerance.

Keywords: drought, Stress, miRNA, black pepper, transcriptome

Design of a Multi-Epitope Based Vaccine using Spike Glycoprotein for Effective Protection Against COVID-19 through Bioinformatics Approach

Puja Jaishwal and Satarudra Prakash Singh

Mahatma Gandhi Central University, Motihari, Bihar

sprakashsingh@mgcub.ac.in

Abstract

The novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is the fundamental agent of coronavirus disease 2019 (COVID-19), which has been spread worldwide since it was first identified in Wuhan, China, at the end of 2019. With the global transmission of the virus, many SARS-CoV-2 variants have a high rate of mutations which affect the epitope conservancy and create obstacles in vaccine design. Conserved epitopes are the desired target in peptide-based vaccine design to enhance the efficacy of vaccine. This study aims to establish an efficient multi-epitope vaccine construct that could elicit both T-cell and B-cell responses and neutralize the SARS-CoV-2 virus. In present work, B- and T-cell epitopes of spike glycoprotein were extracted from IEDB database, and the suitable epitopes were systematically screened using crucial parameters of vaccine effectiveness through immunoinformatic tools involving the assessment of the HLA class I and II binding efficiency, population coverage, along with conservancy among SARS-CoV-2 variants of concern. Results of the screening finally identified 3 overlapping B-cell and T-helper cell epitopes, 1 unique T-helper cell epitopes and 8 unique cytotoxic T-cell epitopes. These epitopes were used in the final vaccine construct design through appropriate linkers and showed 99.45% world population coverage. In future, a tertiary structure of the designed vaccine construct will be developed along with the suitable adjuvant and its interaction with toll-like receptors (TLRs) as well as stability will be evaluated through molecular modelling and simulation studies.

Keywords: COVID-19, Vaccine, Bioinformatics, Modelling, MD Simulation

Gut Metagenomic Analysis of Gastric Cancer Patients Reveals Akkermansia, Gammaproteobacteria, Veillonella Microbiota as Potential Non-invasive Biomarkers

Anju R Nath and Jeyakumar Natarajan

Data mining and text mining lab, Department of bioinformatics, Bharathiar university

anjuraas@gmail.com

Abstract

The goal of the study was to investigate the changes in the gut microbiota during the advancement of gastric cancer (GC) and identify pertinent taxa associated with the disease. We used a public fecal amplicon gastric cancer dataset from the Sequence Retrieval Archive (SRA), of patients with GC, gastritis, and healthy individuals. We did sequence pre-processing, including quality filtering of the sequences. Then, we performed a diversity analysis, evaluating α and β-diversity. Next, taxonomic composition analysis was performed and the relative abundances of different taxa at the phylum and genus levels were compared between GC, gastritis, and healthy controls. The obtained results were subsequently subjected to statistical validation. To conclude, metagenomic function prediction was carried out, followed by correlation analysis between the microbiota and KEGG pathways. α analysis revealed a significant difference between male and female categories, while β analysis demonstrated significant distinctions between GC, gastritis, and healthy controls, as well as between sexes within the GC and gastritis groups. The statistically confirmed taxonomic composition analysis highlighted the presence of the microbes Bacteroides and Veillonella. Furthermore, through metagenomic prediction analysis and correlation analysis with pathways, three taxa, namely Akkermansia, Gammaproteobacteria, and Veillonella, were identified as potential biomarkers for GC. Additionally, this study reports, for the first time, the presence of two bacteria, Desulfobacteriota and Synergistota, in GC, necessitating further investigation. Overall, this research sheds light on the potential involvement of gut microbiota in GC pathophysiology; however, additional studies are warranted to explore its functional significance.

Keywords: Gastric Cancer, Metagenomic Analysis, Gut Microbiota, Diversity Analysis, QIIME2

In Silico Predictive Homology Modeling of PKHD-1 Protein: A Comparative Study among Three Different Species

Arunannamalai S

St. Joseph's College of Engineering

arun03bt@gmail.com

Abstract

PKHD-1 (Polycystic Kidney and Hepatic Disease-1) gene encodes a vital protein critical for renal and hepatic functions. Mutations in PKHD-1 lead to a severe type of disorder in early infancy called Autosomal Recessive Polycystic Kidney Disease (ARPKD). The PKHD-1 protein structure remains unavailable in databases such as PDB, with only a few low-resolution structures accessible in the Swiss Model Template Library. Therefore, Homology Modeling was employed to generate structural models of PKHD-1 proteins derived from three different species [Homo sapiens (Human), Mus musculus (Mouse), Canis lupus familiaris (Dog)]. The mouse PKHD-1 protein was structurally predicted by employing the AlphaFold DB model based on the PKHD1 ciliary IPT domain of fibrocystin/polyductin from Rattus norvegicus as a reference template. Additionally, the human and dog PKHD-1 proteins were modeled using the AlphaFold DB model of the G8 domain-containing protein from Marmota monax as the template for the prediction process. In addition, we employ GOR4 for analyzing secondary structure, ProtParam for assessing physicochemical properties, QMEAN for evaluating the quality of protein structure, and MolProbity for validating protein structures along with obtaining the Ramachandran plot. The binding pockets were also predicted using P2Rank tool (PrankWeb web server).

Keywords: PKHD-1 protein, Autosomal Recessive Polycystic Kidney Disease, In silico analysis, Bioinformatic tools, Homology Modeling

Genetic Variations and their Impact on Diabetic Retinopathy Pathogenesis: A Genomic Evolution Perspective

Shafna Asmy V S and Jeyakumar Natarajan

Datamining and text mining lab,Dept of bioinformatics,Bharathiar university

shafnaasmy.v.s@gmail.com

Abstract

Diabetic retinopathy (DR) is an eye disorder that can cause vision loss and eventually blindness in people who have had diabetes for more than a decade. This study investigated the history of genetic changes within the human genome, tracing the path from normal retinal tissue through diabetic retinal tissue and, eventually, to the formation of DR. We investigated three types of genetic variants to accomplish namely SNPs (single nucleotide polymorphisms), SSRs (simple sequence repeats), and InDels (insertions and deletions). These genetic changes were systematically studied in samples representing normal, diabetic, and DR-affected retinal tissue. SNP-influenced genes were discovered and segregated from the rest of the genetic material after a thorough study of these changes. Furthermore, the analysis identified genes with different levels of expression among the SNP-affected genes. These genes were thoroughly investigated in order to acquire insight into the evolution of SNPs and SSRs. SNPs were found in all three samples' unigenes: normal (19,97,007), diabetes (6,16,892), and DR-affected retinal tissue (11,43,374). The study resulted in five important genes—SPDYA, CACNA1C, NXF2, DNAH12, and LOC105378947—that they have a substantial role in the development of DR. This determination was made through a meticulous analysis of SNPs and InDels, complemented by studies on differential gene expression.

Keywords: SNP Analysis, Evolution, NGS, SSR analysis, Diabetes

Machine learning for T-cell epitope prediction

Riddhi Wankhade, University of Pune

riddhi_mehta74@yahoo.com

Abstract

Precise delineation of lymphocyte antigenic determinants is crucial for advancing therapeutic strategies in the domain of immunomodulation. Computational prediction is a lot cheaper than experimental validation for T cell epitope prediction. Researchers have previously predicted T cell epitopes in both a pan-specific and allele-specific fashion with varying degrees of success. In this manuscript, I leverage computational methods, specifically Kernel-based classifiers, to anticipate antigenic regions recognized by T-cell receptors. My findings are compared to the leading method, NetMHCII. While my results closely align with those obtained by established techniques, it is apparent that the robustness of my predictions relies significantly on the input data. This underscores the imperative for additional research endeavours to position my approach competitively in novel predictive applications.

Keywords: Machine learning, Major Histocompatibility Complex, Net MHC- II, Polymorphism, PREDIVAC, AUC, Margin Based classifiers, Interspersed, Orthogonal Vector, Skewed, Multi RTA, Quadratic Programming

Bioinformatics for Drug Discovery Against Drug Resistant Candida Species

Akshay Kisan Mundhe, Premanand Adaikalasamy and Reena Rajkumari Baskaran

School of Bio Sciences and Technology, Vellore Institute of Technology

b.reenarajkumari@vit.ac.in

Abstract

Infectious diseases caused by fungi contribute significantly to global mortality rates. Among the leading causes of invasive fungal infections are various Candida species, with Candida albicans being the primary causative agent in invasive candidiasis. Developing effective antifungal drugs is challenging due to the limited availability of distinct biochemical targets shared between fungi and their human hosts. Currently, only three primary drug classes are approved for treating Candida infections, and the emergence of drug resistance poses a substantial threat to their efficacy. In this study, we have selected CLB2 as a target to inhibit Candida albicans growth based on literature. We performed molecular docking of this target with nearly fifteen thousand ligand molecules using AutoDock Vina software. Molecular docking results were subsequently subjected to molecular dynamic simulations. From this extensive screening, we identified 17 out of 14,965 ligands that formed covalent bonds with active site residues of the selected target molecule. Further in-vitro experimental studies are being conducted to test the inhibitory potential of the selected compounds against wild and drug resistant Candida albicans strains.

Keywords: pathogenic fungi, Candida albicans, candidiasis, drug resistance, in-silico drug screening, virtual screening, molecular docking, structure-based drug discovery, herbal medicine

3'UTR SNP rs4709267 Associated with Rheumatoid Arthritis Risk and Enhances TAGAP Gene Expression

Shri Preethi M[1], Balaji Nandagopal[2], Raja N[3], Raja Sudhakaran[1], Anbalagan M[1], Karthikeyan K[1] and Asha Devi S[1]

Vellore Institute of Technology[1],

Sri Narayani Hospital and Research Center, Vellore[2],

Sri Padmavathy Rheumatic Care Center, Vellore[3]

ashaselvaraj74@gmail.com

Abstract

Rheumatoid arthritis (RA) is a chronic inflammatory disease. Single-nucleotide polymorphisms (SNPs) in T-cell activation Rho GTPase activating protein (TAGAP) gene is associated with rheumatoid arthritis (RA). In the present study, High-Resolution Melting Analysis (HRMA) technique was employed for SNP genotyping, and the frequency of the genotype association with RA was statistically correlated. To investigate the impact of SNP rs4709267 on the TAGAP gene expression level, TAGAP 3´UTR sequences with SNP rs4709267 and wild type were cloned into the pGL3 SV40 vector. The cloned vectors were transfected to HeLa cells and luciferase expression were compared and internal control pRL_SV40 vector was used. Significant increase ($p<0.05$) in the expression of luciferase activity, suggests that the SNP rs4709267 might be associated with RA and affect the TAGAP gene expression level.

Introduction: Several genetic variants linked to RA susceptibility have been discovered in studies examining the relationship between polymorphisms in human genome sequences and RA case-control characteristics.

Background: A polymorphism causes many autoimmune disorders in the TAGAP gene, namely, rheumatoid arthritis, Crohn's disease, celiac disease, and multiple sclerosis.

Materials and method:

SNP genotyping

The study received hospital ethical clearance and the genotyping of rs4709267 (A/G) at 3'UTR of the TAGAP gene was accomplished employing High-Resolution Melting Analysis (HRMA).

Cloning of the TAGAP 3´UTR sequences in pGL3 -SV40 vector

The wild and polymorphic variants of the 3'UTR region (rs4709267 A/G allele) of the TAGAP gene were obtained. The luciferase assay was performed according to the kit's protocol (Dual-Glo® Assay system Promega, USA).

Results:

SNP genotyping using HRMA

The frequency of the A and G alleles among the controls were 0.87(87%) and 0.13(13%), respectively, whereas, among RA patients, it was 0.6 (60%) for the A allele and 0.4 (40%) for the G allele. The statistical analysis proved that the odds ratio of AG (2.5926) and GG (1.0532) genotype is greater than the odds ratio of AA (0.1012).

Impact of rs4709267 on luciferase gene expression

A significant ($p<0.05$) increase in luciferase activity was observed in 3'UTR_ G allele construct than in the 3'UTR_A allele construct (figure 1).

Figure 1: depicts the pGL3-SV40/pRL-SV40 relative ratio obtained after vector normalization. ANOVA analysis: p value <0.05

Conclusion: The SNP rs4709267 G allele frequency was higher among RA patients than control by the HRMA technique. The SNP rs4709267 3'UTR G construct showed a significant increase in the luciferase activity than 3'UTR with A allele, thereby showing that the SNP rs4709267 (A/G) affects the TAGAP gene.

Keywords: Rheumatoid arthritis, TAGAP, HRMA, Luciferase assay, pGL3 SV40, SNP

Studies on microRNAs and Their Target Genes Through Expression Analysis in Breast Cancer

Saumya Tyagi[1], Vinamrata Sharma[1], Vikram Singh[2] and Tiratha Raj Singh[1]

Jaypee University of Information Technology[1]

Central University of Himachal Pradesh[2]

vikramsingh.jnu@gmail.com

Abstract

Rationale/Motivation: To elucidate the key genes and miRNAs related to breast cancer and their interactions with each other and with other molecular or cellular processes. Breast cancer is the most common cancer among women worldwide. It is a complex and heterogeneous disease, characterised by a diverse range of genetic and molecular alterations. Understanding which is essential for developing more effective diagnostic and therapeutic strategies. One of the recently recognized factors to breast cancer-causing pathways is microRNA (miRNA). These are small non-coding RNAs that play important roles in regulating gene expression at the post-transcriptional level. microRNAs are dysregulated in breast cancer, with some miRNAs being overexpressed and others underexpressed. This dysregulation can lead to changes in the expression of genes that are involved in a variety of cellular processes, including cell proliferation, apoptosis, and migration. miRNAs have been shown to play a role in all stages of breast cancer development, from initiation to progression and metastasis. For example, some miRNAs promote tumor growth by inhibiting apoptosis or promoting cell proliferation. Other miRNAs promote metastasis by increasing cell migration and invasion. It is essential to understand the role of miRNAs in breast cancer in order to develop more effective strategies for the diagnosis and treatment of this disease. This includes identifying the miRNAs that are dysregulated in breast cancer and the genes that they target. It is also important to understand how miRNAs interact in order to unravel the mechanisms behind the development of breast cancer.

Objective: To identify target genes and microRNAs involved in breast cancer (BC), as well as those that are often overexpressed and underexpressed, and to understand the regulatory network between these genes. This study performed a bioinformatic analysis of miRNA-BC studies using a variety of datasets. This study focuses on identifying and analysing differentially expressed genes that are common across different age groups, one from young women (age ≤ 40 years) and one

from older women (age ≥ 40 years). This study also explores the regulatory networks of the genes which regulate breast cancer, with a focus on important genes that interact with each other in breast cancer development and progression. Additionally, we will carefully study the gene ontology of each gene we identify, giving us a comprehensive understanding of how they work in the context of breast cancer.

Material and Methods: In this study, we used bioinformatics analysis to identify common overexpressed and underexpressed genes and miRNAs associated with breast cancer. We analysed gene expression data from the GEO Dataset platform using GEO2R to identify differentially expressed genes that are common across different age groups. We also explored the regulatory networks and gene ontology of the identified genes and miRNAs using the KEGG and STRING databases.

Data retrieval: We retrieved gene expression data from the GEO Database, miRBase, TargetScan, GeneCards. Differential expression analysis: We used GEO2R to identify differentially expressed genes that are common across different age groups in the breast cancer datasets. Gene ontology analysis: We used the KEGG and STRING databases to explore the regulatory networks and gene ontology of the identified genes and miRNAs.

Results and Discussion: In this extensive study, we used bioinformatics analysis to identify common overexpressed and underexpressed genes and miRNAs associated with breast cancer in young women (age ≤ 40 years) and older women (age ≥ 40 years). We found a number of overexpressed genes that are common across both age groups. These findings suggest that these genes may play a key role in the development and progression of breast cancer, regardless of age. We also explored the regulatory networks and gene ontology of the identified genes and miRNAs using the KEGG and STRING databases. These findings provide new insights into the molecular mechanisms underlying breast cancer in young and older women. Additionally, the overexpressed genes that we identified may be potential targets for new diagnostic and therapeutic approaches.

Conclusions: Our study identified common overexpressed genes and miRNAs associated with breast cancer in young and older women, providing new insights into the molecular mechanisms underlying the disease. These genes and miRNAs may play a role in gene regulation and other cellular processes essential for breast cancer development and progression. This information could be used to develop new targeted therapies and diagnostic tools for breast cancer. Therefore, our

study provides a basis for the development of new and more effective strategies for the prevention, diagnosis, and treatment of breast cancer.

Keywords: Breast Cancer, microRNAs, Pathways, Expression analysis, Bioinformatics.

Nephroprotective Role of Syzygium Cumini ZnO Nanoparticles in Streptozotocin - Induced Diabetic Nephropathy Rats

Arul Daniel John[1], Ragavee Ambalavanan[2], Asha Devi Selvaraj[3]

University of Nebraska Medical Center[1]

Vellore Institute of Technology[2]

sashadevi@vit.ac.in

Abstract

Diabetic nephropathy (DN) is one of the major microvascular complications of diabetes and the cause of end-stage of renal failure. Stern control of glucose levels and blood pressure are the therapeutic regime to be followed for the prevention and progression of DN. Previous studies have elucidated the involvement of Advanced Glycation End products (AGEs) and Transforming Growth Factor Beta (TGF-β) in DN. In this study, we evaluated the effect of green synthesized zinc oxide nanoparticles (ZnO NPs) on the expression of receptor for AGEs (RAGE), TGF-β and its subsequent involvement in the treatment of DN in Wistar rats. DN was induced using streptozotocin, which was then treated with green synthesized ZnO NPs and later checked for its physiological and pathological changes. Treatment of DN rats with green synthesized ZnO NPs significantly reduced ($p<0.05$) the blood glucose level, serum creatinine, Blood urea nitrogen (BUN), urine protein and Urine albumin excretion rate (UAER). ZnO NPs also reduced the mRNA levels of RAGE and TGF- β in kidney tissue, which was correlated with pathological improvements such as reduced mesangial expansion and interstitial inflammatory cell infiltrates. The results obtained imply that administration of green synthesized ZnO NPs improved kidney function in DN rats by regulating RAGE and TGF- β in the kidney.

Keywords: Diabetic nephropathy; Zinc oxide nanoparticles; TGF-β; Syzygium cumin; serum creatinine

Identification of Potential Phytochemicals as Inhibitors of CXCR2 and CXCR4 in Glioblastoma using Molecular Docking and Molecular Dynamic Simulation Studies

Inayah Suheb, Shubhada Naik, Premanand Adaikalasamy and B. Reena Rajkumari

Vellore Institute of Technology, Vellore

b.reenarajkumari@vit.ac.in

Abstract

Glioblastoma represents nearly 45% of all primary CNS tumors. Its aggressive nature often makes conventional treatments such as radiation, chemotherapy, and surgery ineffective. The variability in patient responses during clinical trials for glioblastoma can be attributed to factors like tumor heterogeneity, immunological resistance, the presence of the blood-brain barrier, and glioma stem cells in the tumor microenvironment. Research indicates that human proteins, namely CXCR4 and CXCR2, play significant roles in processes like angiogenesis, inflammation, and metastasis. Inhibiting these proteins has been shown to suppress oncogenesis. This study utilized an in-silico approach to screen 14,962 phytochemicals from Indian plants, aiming to identify natural inhibitors of CXCR4 and CXCR2. The ligands were docked against the target protein using AutoDock Vina. Based on the lowest binding energy and the number of hydrogen bonds formed between the ligands and their target in the docked complex, the top ligands were selected for further molecular dynamic simulation studies. Molecular Dynamic Simulation of these protein-ligand complexes was performed using Gromacs 2022.6 to assess the protein-ligand complex stability. From these findings, we chose five ligands that showed promising results to be investigated in in-vitro experiments. The intent of this research is to identify potential phytochemical inhibitors that can be used against the target proteins for glioblastoma treatment which has to be further validated through pre-clinical studies.

Keywords: Glioblastoma, CXCR4, CXCR2, Phytochemicals, Ligands, Molecular Docking, Molecular Dynamic Simulations

Docking and Molecular Dynamics Simulation Revealed the Potential Anti-Oncogenic Activity of Sesamolin in Breast Cancer Therapeutics Targeting the E2F8, A Cell-Surface Receptor Protein

Satarupa Banerjee and Sohini Chakraborty

Vellore Institute of Technology

satarupa.banerjee@vit.ac.in

Abstract

Cell-surface proteins (CSPs) have been employed extensively in cancer research as diagnostic and prognostic markers as well as targets for the creation of anticancer drugs. Few attempts have been made so far to describe the surfaceome of breast cancer (BC) patients. For enabling effective BC therapy, the identification of novel druggable biomarkers is an earnest need. Publicly available databases are utilised to identify CSPs associated with BC. We also foresee significantly altered receptor-ligand interactions in cancer, and we pinpoint significant CSPs and druggable polyphenols (DPs) with therapeutic potential for the disease using systems biology methods. Modern computer-aided drug designing techniques, thus aim to design a cost-effective DP, a natural agent for BC prevention and diagnosis. Here, 56 polyphenols with druggable properties, are initially docked with 3172 CSPs. Finally, duplicate docking was done for the five DPs against the nine significant CSPs identified in BC and proceeded for simulation. The preliminary result of the analysis, reports the highest binding energy scores of E2F8-Sesamolin to be the best-docked protein-ligand complex with a binding energy of -12.22kcal/mol which was then simulated and compared with an approved drug for BC treatment, Olaparib. A comparable binding energy score of -10.62kcal/mol was obtained by docking Olaparib with E2F8. A 100 ns MD simulation revealed that Sesamolin formed more H-bonds (1 to 5), providing a more stable and compact protein-ligand complex with E2F8 as compared to Olaparib (1 to 2). The result was also supported by solvent-accessible surface area, radius of gyration alongside MM-PBSA interaction energies of ΔG values of -51.160+/-18.054 KJ/mol (-12.22 kcal/mol) for Sesamolin which is much better as compared to -44.441+/-18.127 KJ/mol (-10.62 kcal/mol) for the approved drug, Olaparib. Expression, oncoprint, survival and functional enrichment profiles of the significant CSPs are also analysed in

the study. Thus, our findings suggest that the role of Sesamolin can further be studied in detail for BC therapeutics, which was found to target the E2F8, a CSP receptor in a stable manner.

Keywords: Breast cancer (BC), Cell surface proteins (SPs), druggable polyphenols (DPs), sesamin compounds, binding energies, therapeutics

An In Silico Study on Analysis of the Interaction Between Microplastics and Aquatic Pathogens

Amrutha V.M and Sudhakaran Raja, Vellore Institute of Technology

sudhakaran.r@vit.ac.in

Abstract

Plastic plays a vital role in all sectors they are durable, cost-effective, and easily carriable. Due to various aspects such as UV exposure, climatic changes, oxidation, and mechanical stress; plastics get shredded into small fragments. Plastics less than 5mm are considered to be microplastics some commercially available microplastics were manufactured for aircraft, cosmetics, and microfibers for clothing. These microplastics have a great impact on both marine and terrestrial environments. There are many scientific reports that indicate the presence of microplastics in marine organisms is so threatening and tropic transfer of microplastics is also noticed. There are many possible ways shrimp may be contaminated with viral and bacterial diseases due to the interaction of microplastics and pathogens. White spot syndrome virus (WSSV) VP28 is a viral infection of penaeid shrimp that causes 100% mortality in 3-5 days in commercially available farms. Vibrio parahaemolyticus is a gram-negative bacterium, rigid and comma-shaped, and mass mortality is seen in Vibrio species-affected shrimp. This study implies microplastics may act as a vector for marine pathogens so the interaction between microplastics such as polyamide, polyethylene, polypropylene, polyvinyl alcohol, and polyethylene terephthalate was docked with the target protein, stimulation carried out and further in vivo studies will be performed for confirmation

Keywords: Microplastics, Docking, WSSV, Vibrio parahaemolyticus

A Novel Cuproptosis related miRNAs to predict the prognosis of Renal Cell Carcinoma

Subeka Abraham Gnanadass and Pragasam Viswanathan

Vellore Institute of Technology

pragasam.v@vit.ac.in

Abstract

Background: Renal Cell Carcinoma (RCC) is one of the most devastating genitourinary cancers worldwide. The treatment strategies for tackling RCC could be better as the recurrence rate is high, and they need more tools for early detection. Further, the prognosis of RCC still needs improvement. Thus, new strategies must be explored to predict the prognosis of RCC. All organisms require copper as a cofactor to sustain life activities since it is essential to several biological processes, including iron intake, antioxidant/detoxification, and mitochondrial respiration. Cuproptosis is a unique form of copper and proteotoxic-induced cell death, closely associated with the TCA cycle and mitochondrial respiration. Recent research states that copper accumulation is highly associated with cancer progression and development. The regulatory role of cuproptosis in RCC is still a mystery.Micro-RNAs are tiny non-coding RNAs that regulate gene expression transcriptionally and translationally. They are key players in biological processes like cell differentiation, cell death, and cell proliferation. Dysregulation of the miRNAs is highly related to the hallmarks of cancer.

Motivation: This research aims to explore the regulatory role of cuproptosis-related genes and miRNAs in RCC development and progression.

Methods: The miRNA and mRNA transcriptomic data were obtained from the TCGA database. The miRNAs related to the cuproptosis gene signatures were identified by the TargetScan database. The differentially expressed R package identified cuproptosis-related genes and miRNAs. Further, their functional enrichment analysis was done by David and KEGG pathway analysis. The Cuproptosis-related miRNA gene signatures' survival analysis has been constructed using Kaplein-Meier analysis.

Results: An extensive literature survey identified the mRNAs associated with the Cuproptosis process. Out of 29 cuproptosis-related genes, six genes (FDX1, CDKN2A, DBT, NLRP3, PDHA1, and PDHB) were differentially expressed in RCC transcriptomic datasets (logFC >1 and p-value

< 0.05). The differentially expressed miRNAs related to cuproptosis are miR-215-5p and miR-223-3p. The functional enrichment analysis revealed that they play a potential role in cancer development and progression. The identified miRNAs have great prognostic and diagnostic value. **Conclusion:** Thus, the constructed novel Cuproptosis-related miRNA gene signature predicts the prognosis of RCC, and this might pave the way for new treatment strategies in RCC.

Keywords: Renal Cell Carcinoma, miRNA, Cuproptosis, Cancer therapy, Non-coding RNAs

Immunoinformatic-based Vaccine Candidate Development for Edwardsiellosis

Sikhakolli Yeswanth, Aadil Ahmed, Thomas Theodore and Anand Prem Rajan

Vellore Institute of Technology, Vellore

janandpremrajan@vit.ac.in

Abstract

Aquaculture is one of the fastest-growing food production industries which provides income to a lot of people, but some fish pathogens have become an adversary to the production and are causing a lot of economic loss to the fisheries industry. One such pathogen causing the most prevalent disease of Edwardsiellosis is Edwardsiella tarda a facultative anaerobic bacterium. Throughout the period these bacteria got resistant to so many antibiotics which makes it difficult to treat this disease hence vaccination is the best way to prevent this disease. This study has been designed to predict in-silico peptide vaccine candidates by finding out the T-cell epitopes present on the genes(TolC, OmpA, TamA,TamB) of Outer membrane protein sequence(OMPs) and their binding interaction with MHC class I alleles using in-silico immuno-informatics which uses Bioinformatics tools & servers like Kolaskar and Tongaonkar antigenicity tool, NetCTL.1.2 server, Galaxy-pep dock for docking & modelling of protein-peptide structures, Flex-pep for refining those modelled structures, chimera & pymol for interaction. These tools are freely available and approved by CAPRI (Critical Assessment of Predicted Interactions). Such interaction between peptides and MHC class I alleles helped in finding out the potential peptide vaccine candidates. Finally, it's found that two, six & five, four epitopes from TolC, OmpA and TamA, TamB genes respectively have good interaction with MHC class I alleles. The present study can be validated by performing in-vitro studies which help in the development of potential vaccine candidates for Edwardsiellosis.

Keywords: edwardsiellosis, T-cell epitopes, peptide vaccine, OMP

Comparative transcriptome analysis reveals insights into the growth traits of Penaeus monodon

Preety Sweta Hembrom, Mottakunja Deepthi and Tony Grace

Central University of Kerala

tonygrace@cukerala.ac.in

Abstract

The black tiger shrimp, Penaeus monodon, holds significant commercial importance in the marine industry, contributing to about 9% of the total crustacean production and being extensively farmed for human consumption. During the culture of Penaeus monodon, there tend to be notable differences in growth within the same family under the same food, water quality, and environment, which has a significant impact on cultivation efficiency. To explore the molecular mechanism of this growth difference, this study used RNA-seq technology to compare the transcriptomes of P. monodon individuals with significant growth differences from the same family. A total of over 74 million and 72 million paired-end reads were generated from large and small P. monodon samples respectively. Furthermore, we annotated 16,767 putative protein sequences derived from the assembled transcripts, obtaining 15,305 BLAST hits. Additionally, differential gene expression analysis revealed a diverse array of transcript sequences showing similarity to genes associated with growth and development in large P. monodon. Some of these genes, like cuticle protein AMP4, tubulin alpha-1 chain, and myosin heavy chain, may be linked to muscle growth. These genes were classified into 31 KEGG pathways, with the top pathways including carbohydrate metabolism, energy metabolism, and amino acid metabolism. Our findings contribute to a clearer understanding of the genes involved in the molecular mechanisms governing growth traits in black tiger shrimp, offering valuable insights for future research in shrimp development.

Keywords: Penaeus monodon, Growth traits, High-throughput sequencing, Transcriptome

Prediction of Antimicrobial Resistance Profiles from Salmonella Variants using ML and DL Approaches

Gopichand Boinapalli , Athulya V Varma, Anupama Anil, Kiran Manoharan, Joshy Alphonse and Nidheesh Melethadathil

Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, Kerala

ammasnidhi@gmail.com

Abstract

Motivation: Antimicrobial resistance (AMR) is one of the most critical global health challenges of this century. AMR occurs when an organism adapts itself to the antibiotics to the point where the medicines are no longer effective. Antimicrobial Susceptibility Testing (AST) is employed to determine the level of resistance demonstrated by microorganisms against antimicrobial agents. Traditional AST can be time-consuming with low-throughput and only feasible for bacteria that can be grown. Machine learning techniques could pave the path for automated AMR prediction based on bacteria genomic data.

Objectives: Developing ML and DL algorithms to predict antimicrobial resistance phenotypic profiles of Salmonella variants.

Testing and validating the models specific to particular antibiotics and salmonella variants.

Methods and materials: In this study, we mainly focus on the salmonella variants which includes Salmonella Typhi and Salmonella Non-Typhi as they showed high rate of AMR, multiple drug class resistance and vast data availability when compared to other species. Data driven approaches such as Artificial intelligence (AI) and machine learning (ML) methods can be utilized to predict whether a particular variant sample exhibits resistance to a specific antibiotic drug or a drug class. Since most of the previous have employed Whole Genome Sequencing (WGS) data for predicting resistance profiles of Salmonella isolates, WGS data for both salmonella typhi and non-typhi variants has been collected. This includes 5000+ salmonella non-typhi samples, and 133 salmonella typhi samples. The AMR genes of salmonella typhi and salmonella non-typhi were annotated using BacAnt tool. This analysis revealed that both the variants contain common AMR specific genes, indicating common resistance mechanisms in both variants. In addition, this suggests that there is a potential likelihood that both variants share resistance to the same category of drugs. Consequently, a subset of 250 salmonella non-typhi isolates were taken as training set,

and complete set of 133 isolates from salmonella typhi were used as external testing set. Next, XGBoost, SVM and Feed Forward neural network models have been trained using the train set and were evaluated using salmonella typhi testing set.

Results and discussion: The Algorithms developed in this study include SVM, XGBoost, Feed-Forward Network. The algorithms were trained with the k-mer sequences obtained from of salmonella variant Whole genome sequence data, and antimicrobial resistance labels specific to antibiotics. These models were trained on salmonella non-typhi data and validated on salmonella typhi variant data. These models were able to achieve high accuracy rates when tested using salmonella non-typhi variants for Ampicillin, Tetracycline, and Ciprofloxacin antibiotics. Out of the three models that were developed in this study, ML models XGBoost and SVM demonstrated acceptable performance, but Feed forward neural network has outperformed both the ML models with an accuracy of 82%, 85%, 92% for Ampicillin, Tetracycline, and Ciprofloxacin respectively. These results suggest that deep neural network architectures were more effective in capturing relevant features for the classification task, making it the preferred choice for reliable predictions in this study.

Conclusion: From our study it's clear that deep neural network architectures can effectively handle intricate and non-linear data from Whole Genome Sequencing. Further this study can be expanded to utilise all the 5000+ samples of salmonella non typhi to predict AMR profiles which can also pave a potential opportunity for cross-species learning in addition to inter-species transfer learning.

Keywords: Antimicrobial resistance, Whole Genome Sequence, Salmonella Typhi, Machine Learning, Deep Learning

Validation of Sequences from Protease Producing Bacteria with Unknown Sequences

Anchita Prasad and Vinod K Nigam

Birla Institute of Technology, Mesra, Ranchi, Jharkhand,

anchita204prasad@gmail.com

Abstract

During annotation of hypothetical proteins in a few of the bacteria like *Bacillus subtilis, Bacillus anthracis, Streptomyces coelicolor, Mycobacterium tuberculosis, Neisseria meningitidis*, and *Streptococcus*, we identified domains with proteolytic in nature and therefore, protease activity was performed from the different bacteria such as Bacillus species, *Halobacillus dabanensis*, *Halobacillus profundi*, *Halobacillus trueperi* and tried to annotate known 16SRNA sequences from them to examine the functional domains. These sequences were translated first into amino acid sequences followed by annotation using different computational tools (CDD, SMART, Scanprosite and Superfamily). The results did not offer presence of domains in amino acid sequences of known protease producing bacteria. We considered both top hit blasts with about 90% similarity as well as lowest similarity of 50% during domain analysis. The results of finding will be presented at the time of conference.

Keywords: Annotation, CDD, Functional Domain, Hypothetical, SMART, Scan prosite, Superfamily

Drug Repurposing for Non-Alcoholic and Alcoholic Fatty Liver Diseases based on Omics Data

Dharshana R, Vishnu Varadhan S and Dr. Selvi Subramanian

Department of Biotechnology, PSG College of Technology, Coimbatore

selvi.bio@psgtech.ac.in

Abstract

Rationale: Non-Alcoholic Fatty Liver Disease (NAFLD) can lead to liver cirrhosis and hepatocellular cancer, as well as death from liver disease. The primary treatment for NAFLD/NASH is currently lifestyle adjustment through diet and exercise. However, pharmaceutical therapy is required since obese people with NAFLD frequently struggle to maintain healthy lifestyles. The pathogenesis of NAFLD/NASH is not well understood. However, insulin resistance, inflammatory cytokines, and oxidative stress are thought to play a role in the disease's genesis and/or progression. The analysis of Genome-Wide Association Studies (GWAS), Phenome-Wide Association Studies (PheWAS) and Transcriptome data for drug repurposing from Alcoholic Fatty Liver Disease (AFLD) to NAFLD is a rational and efficient strategy to address the unmet therapeutic needs of individuals afflicted by liver diseases. This approach identifies shared genetic variants associated with common pathogenic mechanisms in both conditions. It also expedites the development of new therapeutic options by identifying existing drugs that may be effective in treating these liver diseases. Repurposing drugs is often more cost-effective and less time-consuming than developing new drugs from scratch, making it an attractive strategy for diseases with limited treatment options.

Objectives: To identify the genes which are present in both NAFLD and AFLD disease conditions. Functional enrichment of those genes which could be used as potential targets. Identification of the current drugs which could be repurposed for NAFLD and AFLD

Materials and methods: Disease associated genetic variants and their mapped genes were collected from surveying the data available in the GWAS catalog portal by using the keywords “NAFLD” and “AFLD”. PheWAS catalog was used to discover all associated genes corresponding to variant alleles. The data was collected by setting the phenotype as "alcoholic liver damage" and "non-alcoholic liver disease". The genetic variants and the mapped genes for these phenotypes along with the associated diseases were retrieved. Transcriptome data of both NAFLD and AFLD

disease conditions were retrieved from GEO database. These data were analysed using DESeq2 package in R to identify the differential expressed genes (DEGs). The DEGs were then sorted into up and down regulated based on their logF2 value. All differentially expressed genes of both disease conditions collected from GWAS, PheWAS and Transcriptome data were used in venn diagram to identify the genes which are common for both conditions, for drug repurposing. Pathway enrichment analysis was done for the genes which were discovered to be common between both conditions using 'Enrichr tool'. DrugBank 5.0 and DGIdb were used to identify the FDA approved drugs in the market which targets the genes that were identified and that could be repurposed.

Results and Discussion: Three hundred and twenty-eight common gene targets between NAFLD and AFLD were identified using GWAS and PheWAS and transcriptome data through venn diagram. Pathways implicated in the pathophysiology of AFLD and NAFLD were also identified by integration of all the above data. From the KEGG pathway analysis, it was found that these genes were involved in pathways which were found in other metabolic disorders such as diabetes, obesity and cardiovascular diseases. Findings were used to identify potential targets for drug repurposing. Since these disorders overlaps with other metabolic disorders, FDA approved drugs from DrugBank 5.0 common for the above conditions were chosen for drugs repurposing. For each potential gene target such as PNPLA3, drug repurposing was done through Cmap tool and the those which were having a score of above 99 were chosen as drug candidates which could be repurposed for fatty liver disorders.

Conclusion: Non-Alcoholic and Alcoholic Fatty liver disease represent a significant global health burden, with prevalence of 32.4% and 5.1% respectively. Insulin sensitisers and lipid lowering drugs are currently used along with diet and lifestyle modification. Despite the rising incidence, there is no FDA approved medications for these conditions. Results from this study, could provide a valuable resource for drug repurposing efforts in the field of liver diseases, potentially accelerating the development of new treatment options for NAFLD and AFLD. This work underscores the importance of utilizing large-scale genomics data to uncover novel therapeutic approaches and improve treatment options in the context of these prevalent and challenging liver diseases. These should be validated clinically for their effectiveness.

Keywords: NAFLD, AFLD, GWAS, PheWAS, Drug Repurposing

Drug Repurposing Analysis for Psoriatic arthritis using PheWAS and Transcriptome Data

Jayasree N, Dharuunya H V and Selvi Subramanian

Department of Biotechnology, PSG College of Technology, Coimbatore

selvi.bio@psgtech.ac.in

Abstract

Rational: The conventional method of drug development requires more time and cost. Drug repurposing reduces drug development timeline as various existing compounds have already demonstrated safety in humans and it does not require Phase 1 clinical trials. With the advancement of bioinformatics/cheminformatics tools and availability of huge biological and structural database, drug repositioning has significantly decreased the time and cost of the drug development with reduction in risk of failure. It confers reduced risk of failure where a failure rate of 45% is associated due to safety or toxicity issues in traditional drug discovery program with additional benefit of saving up to 5–7 years in average drug development time. Since psoriasis and psoriatic arthritis are associated the drugs for psoriasis could also be repurposed and used for psoriatic arthritis

Objective: To identify potential target genes for psoriatic arthritis and psoriasis. To identify drugs from drug bank against each target gene**.** Drug repurposing analysis for target genes

Methodology: In this study we used wide range of data retrieved from GWAS, and PharmgKB datasets for identification of reported gene targets for Psoriatic arthritis and Psoriasis. In addition, comparative gene expression analysis was performed on the transcriptomic dataset obtained from PsA and psoriatic conditions. The potential gene targets from the above study were subjected to logical distribution analysis using Omics box tool. Functional enrichment analysis was performed for the target genes using enrichr tool to understand the disease pathogenesis. The drug targets for these candidate genes were analyzed using drug bank database for drug repurposing

Results and Discussion: Target genes were identified using the available pharmgkb, GWAS and transcriptomic data for PsA and psoriasis. Forty-nine common gene targets associated with the disease were identified using Venn diagram. The pathway enrichment analysis led to a better understanding about the disease pathogenesis. Involvement of DEGs in the KEGG pathway was explored further. The targets genes of Psoriatic arthritis and Psoriasis involved in various

pathways. Hence, we have a wide option in treatment of the disease. The drugs for the target genes were identified through drug bank and drug repurposing was performed. For the target gene TNFAIP3, drug repurposing was performed using Cmap tool and the selective match with the score of above 99 were chosen to be the potential drug candidates. These could be further analyzed and implemented for drug usage. Likewise, drug repurposing was done for each reported gene targets which would enable wide range of treatment options.

Conclusion: Psoriasis is a multisystemic, chronic inflammatory skin illness characterized by scaly erythematous plaques on the extensor surfaces of the elbows and knees and 20-30% of psoriasis patients develop psoriatic arthritis (PsA). As the duration for drug development for a disease is high, treating the disease can be costly and painful. Drug repurposing is an alternative to reduce the duration and cost of drug development. Using the available pharmgkb, GWAS and transcriptomic data for PsA and psoriasis, 49 common gene targets associated with the disease were identified. The pathway enrichment analysis led guided to acquire knowledge about the disease pathogenesis. The drugs for the target genes were identified through drug bank and drug repurposing was performed for the target TNFAIP3 which could be an alternative drug candidate for PsA enabling wide range of treatment options.

Keywords: Psoriatic arthritis, Psoriasis, Repurposing, Gene expression analysis, GWAS, Transcriptomics

Multi-Omics Analysis to Unravel the Molecular Etiology of Migraine-Related Psychiatric Disorders

Hema Palanisamy, Jayavishnu Elanchezhian and Selvi Subramanian

Department of biotechnology, PSG college of Technology, Coimbatore

selvi.bio@psgtech.ac.in

Abstract

Introduction: Migraine (MIG) is neurological disorder that is prevalent among ~15% of general population. It is characterized by recurrent attacks of headaches along with symptoms like nausea, vomiting, hypersensitivity to light and sound. It can occur at any stage of life span including children, adult and elder individuals. Both genetic and environmental factors contribute to the development of MIG attack. Multigenetic variants influence the development MIG when compared with the variants in individual genes which is identified through Genome-wide association studies (GWAS). It is hypothesized that there is a shared genetic association between MIG and other psychiatric disorders. But the etiology of the MIG and associated disorders remain elusive. Advancements in omics technologies and availability of omics data could help to unravel their etiology and association for better therapy. In this study we implemented multi-omics approach to study the etiology of MIG associated two most common mental disorders schizophrenia (SCN) and depression (DEP).

Methodology: Variants and genes associated with MIG and SCN, MIG and DEP were retrieved from literature. The gene expression data for SCN and DEP was collected from published source through Gene Expression Omnibus (GEO) datasets. The studies which utilized expression quantification through microarray across various tissue samples were selected. The datasets were analyzed to identify Differentially Expressed Genes (DEGs) through 'GEO2R' for SCN and DEP separately. DEGs were filtered which have p-value <0.05 were selected further. Genes which had positive logFC value were up regulated genes and genes which had negative logFC were down regulated genes. Expression of genes identified through variant analysis for SCN and DEP were filtered and taken for enrichment analysis using 'Enrichr' web tool.

Results and Discussion: From literature we identified 37 genes from 14 loci associated with MIG and DEP and 298 genes for 36 loci associated with MIG and DEP. Since these genes are jointly

associated with MIG, they can be involved in MIG associated SCN and DEP. Hence, these genes were considered as candidate genes.

Expression of candidate genes involved in MIG associated SCN and DEP were identified through expression profiling datasets of SCN and DEP. For SCN, we identified ten DEGs in prefrontal cortices. ARMC6 was highly upregulated and rs11668203 variant was associated with the gene. Twenty-one DEGs in olfactory neuronal cells and UBALD1 gene associated with rs4786505 variant was downregulated. Eighteen DEGs in superior temporal cortex among which CNNM2 was moderately upregulated and NACA was downregulated. Thirteen DEGs in BA10 region from which NUP160 was upregulated and EFNA was downregulated. ARMC6 associated with rs11668203 was upregulated across all tissues during SCN. It is reported that ARMC6 is involved in regulating wnt signaling pathway that could lead to nervous system disorders. So, the rs11668203 variant could be involved in SCN during migraine. GNE, DESI1 were upregulated in DEP. rs71327107 variant was involved in both DEP and SCN. Involvement of DEGs in KEGG pathway was explored further. Pyruvate metabolism was altered by MIG associated SCN genes. Among peripheral metabolites, pyruvate is a significant substrate for brain and peripheral energy metabolism. Thus, it could be associated with the pathophysiology of SCN. Amino acid and nucleotide metabolism was altered during DEP. It is reported that the dysregulation of amino acids has been consistently correlated with psychopathology.

Conclusion: We identified candidate genes and variants that could be involved in migraine associated SCN and DEP. We further identified DEGs from the candidate genes and identified their functional involvement. Hence these targets could be evaluated for their therapeutic efficacy to treat migraine associated SCN and DEP. Further proteome and metabolome level analyses are yet to be performed to validate their involvement at a higher resolution.

Keywords: Migraine, Depression, Schizophrenia, multi-omics, transcriptome

Molecular Docking and Molecular Dynamic Simulation (MDS) Investigation of Actinobacterial based Bioactive Compounds against Fusobacterium nucleatum Aggravated Oral Squamous Cell Carcinoma (OSCC)

Zarin Taj and Indranil Chattopadhyay,

Department of Biotechnology, School of Life Sciences, Central University of Tamil Nadu, Thiruvarur, Tamil Nadu

indranil@cutn.ac.in

Abstract

Oral squamous cell carcinoma (OSCC) is the prevailing cancer impacting the oral cavity, characterized by a dismal prognosis and a relatively low probability of survival. There exists a strong association between the microbes with the OSCC. Among the periodontopathogenic bacteria, Fusobacterium nucleatum (F. nucleatum) is considered a potential risk factor in the progression of OSCC. Many natural bioactive compounds explored for their antimicrobial and anticancer potentiality in a variety of microbiome-aggravated cancers but have rarely been studied in oral squamous cell carcinoma (OSCC). The key virulence proteins of F. nucleatum including FadA and Fap2 has been identified to bind host cell and activates various oncogenic pathways in oral carcinogenesis. The current study employs High-Throughput Virtual Screening, ADME/T profiling, Molecular docking, and Molecular Dynamic Simulation techniques to identify potential actinobacterial bioactive secondary metabolites that can target therapeutic targets of both pathogenic virulence proteins and host cell proteins. In this study, among 179 bioactive secondary metabolites, AM-158 labelled metabolite exhibited multi-protein targeted with highly acceptable binding affinity with drug-likeness property and passed level of toxicity. Comprehensive docking interaction of the best top-ranked metabolite AM-158 with OSCC-related protein targets illustrated greater binding affinity towards E-cadherin and p38 proteins. The molecular dynamic (MD) simulation has been executed for the metabolite AM-158 for both bacterial virulence proteins and cancer therapeutic targets showing stable intermolecular binding with both hydrogen and hydrophobic interactions. In conclusion, this study states that bioactive secondary metabolite AM-

158 from Actinobacteria could be a powerful therapeutic compound to treat Fusobacterium nucleatum aggravated oral squamous cell carcinoma (OSCC).

Keywords: Actinomycetes, Bioactive metabolites, antimicrobial, anticancer, oral squamous cell carcinoma, Fusobacterium nucleatum, molecular docking, molecular dynamics

Computational Screening and Docking analysis of Phytochemicals from Senna auriculata and Trachyspermum copticum against Multiple targets of Mycobacterium tuberculosis

Neeraja Karumarathodi, Shreya Manjusha Binukumar, Subhankar Das, Sadhana Sundararajan, Keerthana Karunakaran and Rajiniraja Muniyan.

School of Biosciences and Technology, Vellore Institute of Technology, Vellore

rajiniraja.m@vit.ac.in

Abstract

Tuberculosis is an infectious disease caused by the potential pathogenic bacteria *Mycobacterium tuberculosis* which causes more than ~1.5 million deaths every year. The reason for its successful infection rate is owed to its resilient, tough mycolic acid-rich cell wall that makes the antibiotics hard to penetrate into the cell and its ability to manipulate the immune system. In addition, drug resistance has become a major concern. For the above-mentioned reasons, incessant attempts are being made to identify novel drug targets and newer natural anti-tubercular drugs to control the spread of TB. In the previous study, ethnobotanically important medicinal plants such as *Trachyspermum copticum* and *Senna auriculata* were evaluated for anti-mycobacterial potential against M. smegmatis. The ethyl acetate and methanol extracts of the selected plants that had anti-TB activity were analyzed in Gas Chromatography-Mass Spectrometry (GC-MS) to identify the compounds responsible for the activity. In the current study, a total of 53 phytochemicals identified and mentioned in literature from medicinal plants *Trachyspermum copticum* and *Senna auriculata* in addition to the phytochemicals obtained from the GC-MS analysis were subjected to multi-step filtration protocol against eight drug targets of *Mycobacterium tuberculosis*. Important proteins that serve as targets for front-line drugs owing to their importance in arresting the growth of the organism are considered in the study as targets for natural compounds from *S. auriculata* and *T. copticum*. Various proteins play significant roles in different pathways contributing to cell wall metabolism which makes them possible drug targets. A multi-targeted approach of natural plant compounds against the front-line drugs is attempted in this study. The compounds selected for the study were filtered using Lipinski's rule of five and the docking procedure was validated using co-crystallized ligands and the drugs, in case of the absence of co-crystallized structures. The proteins

for the study were taken from PDB based on resolution, absence of mutation, and residues present. The process of docking was done using Autodock 4.2. From the exhaustive docking analysis, stigmasterol was identified as potential smulti-targeting compound that was found effective against multiple emerging targets such as EmbC, FbiB, and MmpL3 of *Mycobacterium tuberculosis*. Subsequently, the molecular dynamic simulation was done to study the stability of the complex in comparison with apoprotein.

Keywords: Multiple target approach, Molecular docking, Senna auriculata, Trachyspermum copticum, Exhaustive docking.

Exploration of Phytochemicals as Prospective Drug Candidates for Tuberculosis: A Computational Approach

Faaizah Afsheen[1], Srinivasan M.R[2], Sujatha P.L[3], Vijayarani K[4], Devendran P[5]

Bioinformatics Centre, Madras Veterinary college, TANUVAS[1],

Laboratory Animal Medicine Unit, CAHS, TANUVAS[2],

Madras Veterinary College, TANUVAS[3],

TANUVAS, Madhavaram Milk colony[4],

Bioinformatics Centre, Madras Veterinary college, TANUVAS[5]

seenubioinfo@gmail.com

Abstract

Motivation: Tuberculosis remains a significant global public health challenge with drug resistance threatening the use of the first-line anti-TB drugs, ethambutol, isoniazid, rifampicin and pyrazinamide. This requires the development of novel drugs with fewer side effects.

Objectives: This study aims to assess the potential of phytochemicals as drug candidates for the treatment of tuberculosis through the integration of systems biology, molecular docking, and molecular simulation methodologies.

Materials and Methods: Phytochemicals of nine plants exhibiting in vitro anti-tubercular activity—*Curcuma longa, Ocimum basilicum, Opuntia ficus indica, Mangifera indica, Vitex negundo, Tinospora cordifolia, Justicia adhatoda, Acacia catechu*, and *Mukia maderaspatan* were obtained from documented references, and their structures were retrieved from PubChem. Around 75 ligands with logP values between 0.5 to 5, zero PAINS alerts, predicted non-mutagenicity, solubility, and intestinal absorption were selected following ADMET prediction in Discovery Studio 2020 and SwissADME. Mycobacterial proteins spanning the pathogenic genome were acquired from the Mycobrowser data repository and documented sources. A protein-protein interaction network was constructed in STRING with the 900 proteins. An interaction network with 634 nodes was created and analyzed in Cytoscape 3.0 and NetworkX. The proteins with the highest betweeness scores in the entire network and the proteins with the highest betweeness scores in each of the isolated networks were identified. The 3D structures of these proteins were obtained from PDB and AlphaFold databases. Binding energies of the identified compounds and mycobacterial protein complexes were calculated in Discovery Studio 2020 after docking with

CDOCKER protocol and were compared with that of the first-line drugs. The functions of the proteins were confirmed with enrichment analysis in DAVID. The compound- protein complexes for molecular dynamics were selected based on their relevance to mycobacterial pathways and their binding energies and were simulated in Discovery Studio 2020 for 20 ns to identify the most favorable compounds and interactions for subsequent drug development endeavours.

Results: The compounds showed higher affinity to the drug targets rpoB, inhA and embC which were among the identified 45 central proteins in the network. The compounds, curcumin, 1,2-dihydrocurcumin, tetrahydrobisdemethoxycurcumin, ferulic acid, isorhamnetin, xanthomicrol, tolycaine, bisdemethoxycurcumin, 3-(2,4-dihydroxyphenyl)-acrylic acid, iriflophenone, N-trans-Feruloyltyramine, 4-hydroxybenzoic acid and 4-hydroxybenzaldehyde exhibited the lowest binding energies across all the proteins. The conformational change of the highest scoring complex was observed to be >6 A.

Keywords: Network analysis, Phytochemicals, Tuberculosis

To Identify a Novel Target Compound to Inhibit the OXA ß-Lactamases Causing Extremely Drug-Resistant Hospital-Acquired Infections

Mohanraj Gopikrishnan[1], Sriroopreddy Ramireddy[1], Yamuna Devi Bakathavatchalam[2], Thirumal Kumar D[3], Binesh Lal Yesudoss[4], Abi Manesh[5], Kamini Walia[6], Balaji Veeraraghavan[4], George Priya Doss C[1]

Department of Integrative Biology, School of Biosciences and Technology, VIT, Vellore, Tamil Nadu, India[1],

Department of Clinical Microbiology, Christian Medical College, Vellore, Tamil Nadu, India[2], Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India[3], Department of Clinical Microbiology, Christian Medical College, Vellore, Tamil Nadu, India[4], Department of Infectious Disease, Christian Medical College, Vellore, Tamil Nadu, India[5], Division of Epidemiology & Communicable Diseases, Indian Council of Medical Research, New Delhi, India[6]

georgepriyadoss@vit.ac.in

Abstract

Carbapenem-resistant *Acinetobacter baumannii* (CRAB) and Carbapenem-resistant *Enterobacteriaceae* (CRE) is an important nosocomial infection in healthcare sectors, and it leads to higher rates of morbidity and mortality rates. It causes mostly bacteremia and ventilator-associated pneumoniae. The presence of Oxacillinases leads to a high drug-resistant profile, and it is mostly resistant to β-lactams. Currently, there is no proper drug regimen against the carbapenem resistance. This necessitates the need to search for newer approaches. Recently, small molecules have shown more potent activity against multi-drug resistance strains. To this aspect, we collected nearly 28,831 compounds in the Antibacterial Library from Enamine to design for the development of novel antibacterials against the Oxacillinases. So, we computationally performed the high throughput virtual screening of these compounds through Schrodinger platforms against the targeted Oxacillinases such as (OXA-23/ 24/58 like) expressing *A. baumannii* and OXA-181/232 from *E. coli* and *K. pneumonia* isolates. To further validate our findings, molecular dynamic simulations were performed using GROMACS followed by molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) analysis. To further prove our hypothesis, the potent compounds were synthesized, and we determined the minimum inhibitory concentration (MIC)

using Microbroth dilution methods. In the virtual screening, we selected the top two compounds of Z1738535794, and Z863012955 which showed the higher binding energy of -6.0 to -7.0 kcal/mol against all the variants of Oxacillinases and these compounds passed the ADMET properties. Then we performed the molecular dynamic environment for 100 ns all the complexes were stabled in the dynamic environment. Then we did an MMPBSA analysis it has a total binding affinity ranging from -81.03 to -94.06 KJ/mol. Whereas in MIC testing, the obtained two compounds were effective against Oxacillinases. The obtained potent compound is a promising drug candidate for treating Oxacillinases expressing resistant bacterial strains.

Keywords: Oxacillinases, Betalactam, Antibacterial library, Schrodinger, Molecular dynamics

Differential Gene Expression Analysis to Unveil Potential Targets and Pathways in Atherosclerosis Under Prolonged Hyperglycemic Conditions: A Bioinformatics Study

Soumya Jal[1], Sangeeta Chhotaray[2], Kanika Verma[3], Nitesh Dhar Badgayan[1]

Dean, School of Paramedics and Allied Health Sciences, Centurion University of Technology and Management, Odisha[1]

Research Scholar, SoPAHS, Centurion University of Technology and Management, Odisha[2],

Scientist B, Department of Molecular Epidemiology, ICMR-National Institute of Malaria Research[3]

soumya.jal@cutm.ac.in

Abstract

Background and Aim: Persistent elevation of blood glucose levels can result in impairments to vascular and neural systems, contributing to the development of cardiovascular and multi-organ complications. Atherosclerosis, being one such complication, is characterized by the accumulation of fatty deposits within the arteries. The weakening of the myocardium, which is caused by prolonged hyperglycemia, is a major contributor to global mortality. However, the precise mechanisms by which diabetes leads to cardiac complications are poorly defined. This study aims to employ bioinformatics methods to screen and identify molecular targets that correlate proliferative Diabetic Retinopathy (DR) with atherosclerosis.

Methods: The analysis utilized transcriptomic data obtained from the GEO database, specifically the dataset with accession number GSE94019. The GEO2R statistical tool was employed to identify differentially expressed genes (DEGs) in the Endothelial Cells from Fibrovascular Membranes of nine individuals with diabetes and four individuals without diabetes. The identification of the interaction between the differentially expressed genes (DEGs) was performed by utilizing the STRING tool, followed by visualization using the Cytoscape software. In order to identify the gene cluster within the interactive networks, Cytoscape was utilized. Functional annotation of the identified DEGs was performed using the DAVID web server and Shinygo tool. This included gene ontology (GO) and enriched molecular pathway analysis of DEGs.

Result: We examined the 414 most significant DEGs (p-value < 0.05) out of a total of 3765 DEGs. DEGs that exhibited significant differences in GO analysis were identified as being implicated in molecular pathways and critical biological processes, including calcium ion binding, cell adhesion, and vascular smooth muscle cell proliferation. The correlation between DEGs and the MAPK signaling pathway was identified through the examination of enriched KEGG pathways. It has been demonstrated that the genes implicated in the molecular pathways can be selectively regulated through the activation or inhibition of genes that are indispensable for the canonical signaling pathways. Ten hub genes (TP53, JUN, CD4, PTEN, ICAM1, GRB2, CREBBP, TFRC, EWSR1, and CDKN1A) were identified in our research as being significantly associated with DR and a heightened susceptibility to atherosclerosis.

Keywords: Atherosclerosis, Diabetic Retinopathy, STRING

Antiproliferative Activity of Prodigiosin Derived from Serratia Marcescens VITSD2: An In Vitro & In Silico Approach

Sreelakshmi R Nair and Subathradevi C,

Vellore Institute of Technology, Vellore

csubathradevi@vit.ac.in

Abstract

Prodigiosin is a potent anti-oxidant, red pigment produced by different strains of Serratia marcescens and other bacteria. The bio pigment demonstrates many hopeful impending bioactivities. It was found to be an active proapoptotic agent against multiple cancer cell lines. In the present study, pigment produced from soil isolate Serratia marcescens VITSD2 was characterized and identified using UV, FTIR, GC-MS and NMR analysis (1H NMR and 13C NMR). Prodigiosin pigment produced from Serratia marcescens VITSD2 showed potent cytotoxicity on HepG2 cancer cells. The anti-proliferative activity of prodigiosin pigment from Serratia marcescens VITSD2 was evaluated on cancer cell lines. The active sites and binding patterns of molecular marker survivin was analysed on docking against prodigiosin. A strong antioxidant potential was noticed at 5mg/mL concentration with 70±0.08% scavenging activity (2,2-diphenyl-1-picrylhydrazyl)-DPPH. The dose dependent inhibition of HepG2 cell proliferation was observed maximum with 67±0.08% cytotoxic activity at 50 µg / mL. When compared to other cell lines, A549, HL 60 and MCF-7, prodigiosin had a strong inhibitory activity on HepG2 cells. Serratia marcescens VITSD2 showed potent cytotoxicity on HepG2 cancer cells A single band with an Rf value of 0.45 was observed after chromatography. Maximum absorbance was observed at 535 nm. The pigment revealed the characteristic functional properties of the prodigiosin. On docking, the lowest binding energy exhibited was found to be -6.78 kcal/mol. The RMSD analysis indicated that the backbone structure converges at 18ns before it attains stability. The study outcomes specified that the bio pigment prodigiosin extracted from Serratia marcescens VIT SD2 is a promising drug candidate appropriate for therapeutic applications.

Keywords: Prodigiosin, Serratia marcescens VITSD2, antioxidant, pigments, Anti-cancer, Hep-G2

Virtual Screening, Molecular Docking and Dynamics Simulations for Identifying Potential Natural Inhibitors for Managing Colorectal Cancer against PDZ Domain-Containing Protein GIPC2

Karishma Sahoo and Vino Sundararajan,

Vellore Institute of Technology

svino@vit.ac.in

Abstract

Motivation: Colorectal cancer (CRC), affects the population with prevalent malignancies, high mortality rate, and poor prognosis of patients accounting for the targeted therapies that can successfully increase the overall survival rate of patients. Furthermore, conventional cancer therapies have presented numerous challenges, such as toxicity, resistance to multiple drugs, and high financial burden. Conversely, there's a growing interest in the world of complementary alternative medicine, specifically in bioactive phytochemicals. These compounds have captivated our attention by their ability to influence a wide range of molecular processes while minimizing harmful side effects. Moreover, natural products have played a critical role in medicine showing the anti-apoptosis, anti-oxidative, pro-apoptotic, and anti-metastatic activity and their ability to bind and modulate cellular targets involved in disease. Therefore, this study was planned to screen phytochemicals against PDZ domain-containing protein GIPC2 which has the potential to modulate the signaling pathways involved in cancer and may be used for developing an effective and broad-spectrum strategy for increasing the overall survival, progress-free interval, TNM staging of CRC patients, in coming future.

Objectives: The identification of potent inhibitors from the large diverse natural compound library using the target based virtual screening process. Quantify the binding affinity along with the prediction of the potential bioactivity and pharmacological properties of the screened compounds, including ADME (Absorption, Distribution, Metabolism, and Excretion) properties. Perform molecular dynamics simulations to assess the binding stability, dynamic behavior and interactions between the shortlisted plant compounds and the targeted protein.

Methodology:

3.1. Datasets construction of natural compounds:

Almost 5006 natural anticancerous compounds were collected through a literature survey and databases such as PubChem, and IMMPAT (Indian Medicinal Plants, Phytochemistry And Therapeutics) and where identified showing the inhibitory properties against PDZ domain-containing protein GIPC2. Few of the FDA approved inhibitors were retrieved by the same procedure and used for docking against the protein.

3.2. Target-based Virtual screening and Determination of ADMET, Lipinski's rule, and pharmacokinetics of screened compounds:

The ligands and the proteins were prepared for Target-based Virtual screening and the top 15 screened compounds were subjected to ADMET, Lipinski's rule, and pharmacokinetics analysis.

3.3. Molecular docking and Molecular Dynamic simulation:

Molecular Docking for single ligand and protein complex can validate the results of Auto dock vina results which can be further subjected to Molecular dynamics simulations using the GROMACS package to determine structural stability and protein properties followed by the free binding energy calculation. This simulation can be carried out for 500ns using CHARMM36 force field in GROMACS. Furthermore, the post simulation analysis such as Solvent Accessible Surface Area (SASA), Radius of Gyration, Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and visualization of all graphs can be generated using Xmgrace.

Results: To decipher the anticancerous activity of the selected natural products against the receptor protein (PDB ID- 3GGE), high throughput virtual screening was performed via Autodock Vina with a library comprising 5008 natural compounds showing anticancerous activity. The top 10 screened compounds were seen to have good docking scores and were considered to be showing better inhibition protein GIPC2 which are Oleanolic Acid, alpha-Amyrenone, Ursolic acid, alpha-Amyrin, Cimigenol, Lupeol, Bismorphine B, Gochnatiolide A, Irinotecan, and 9-Nitroamino-camptothecin. Also, the pharmacokinetic data of the screened compounds should demonstrate that all chosen natural products exhibit improved pharmacological characteristics, such as low molecular weight, compliance with the Lipinski rule of five (not exceedingly more than one violation or showing no violation), favorable absorption profiles, oral bioavailability, excellent gastrointestinal absorption, and minimal toxicity risk. Furthermore, these compounds predicted to possess reasonably good pharmacological profiles should also surpass the insignificant toxicity. Thereafter, alpha-amyrenone and oleanolic acid shows the highest binding energy, good human intestinal absorption, good water solubility, and plasma protein binding affinity which is further

selected as the best docked compounds. Finally, Molecular dynamics simulations for 500 ns including molecular mechanics Poisson-Boltzmann (MMPBSA) calculations are expected to provide strong evidence of the stability and integrity of the ligand-target protein complexes, further affirming the validity of their binding mechanism.

Conclusion: Our research highlights the role of alpha-Amyrenone and Oleanolic acid as a promising candidate for inhibiting the overexpression of GIPC2 protein, which also demonstrates their action on inappropriate activation of WNT-signaling cancer pathways.

Keywords: Virtual Screening; Molecular docking, MM-GBSA, ADMET, Molecular dynamics, Natural inhibitors, Colorectal cancer.

An Integrated Bioinformatic Approach to Identify Potential Genes in Colorectal Cancer

Soundharya G, Vyshali G and Shobana S,

Department of Biotechnology, PSG College of Technology

sbs.bio@psgtech.ac.in

Abstract

Colorectal cancer is one of the most frequent cancers in the world. Only a small percentage of patients are being cured by current treatments, which work best for those with early-stage disease. Therefore, treatment like gene therapy could be a possible way in the future to cure several types of cancer. Identifying the origins and pathology as well as developing novel biomarkers are of considerable significance and are urgently required due to the high incidence and mortality of colorectal cancer. Henceforth, to identify key genes of colorectal cancer, we have performed a gene expression analysis using a dataset (GSE223118) from Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were identified through the GEO2R analysis tool. The network of protein–protein-interaction (PPI) was established by using the STRING database and visualized by Cytoscape. A total of 1000 significantly differentially expressed genes were obtained, which consisted of 362 up-regulated genes and 638 down-regulated genes. The functional categories of the genes were identified using DAVID server. Most of the corresponding genes were involved in the process such as signaling receptor activity, G protein-coupled receptor binding activity and cytokine mediated signaling pathway. This study attempted to find further molecular changes in colorectal cancer, requiring the development of novel diagnostic biomarkers and therapeutic targets.

Keywords: Colorectal cancer, GEO, Biomarkers, protein-protein interaction

Flux Balance Analysis of Exopolysaccharide Biosynthesis in Methylobacterium mesophilicum

Vineeth L, Soundharya G and Shobana S

Department of Biotechnology, PSG College of Technology

sbs.bio@psgtech.ac.in

Abstract

Methanol, a petrochemical industry building block and a widely accessible and potent single carbon liquid molecule that can be produced from methane present in natural gas, has emerged as a potential candidate with low cost and high purity. Methylobacterium mesophilicum are pink pigmented facultative methylotrophic (PPFM) methanol utilizers that can grow on single carbon compounds such as methanol and formamide. They produce the exopolysaccharide (EPS), which forms highly viscous aqueous solutions even at low concentrations, which is appealing to the food, cosmetic, and pharmaceutical industries. Bacterial EPSs' high production capacity, low resource-intensiveness, physicochemical and structural properties make them appealing for a wide range of industrial applications. In this study we have developed a model for the biosynthesis pathway of EPS in M. mesophilicum and performed flux balance analyses (FBA) for maximizing the production of EPS using COBRA tool box. The model for EPS biosynthesis through colonic acid pathway involves 16 metabolites and 12 reactions catalysed by 11 proteins. This study demonstrates the usage of FBA for modelling the metabolic pathway thereby increasing the production of exopolysaccharides.

Keywords: Methanol, FBA, COBRA, Exopolysaccharides, Methylobacterium

Evaluation and Benchmarking of de novo Assembly Tools for Prokaryotic Long Reads from Oxford Nanopore Technologies

Megha S Kumar, Suneesh P V, Aarathi Pradeep and Satheesh T G

Amrita Vishwa Vidyapeetham

sk_megha@cb.amrita.edu

Abstract

Rationale/Motivation: Reliable genome reconstruction from the sequenced data is vital for several downstream genome analyses as it forms the base for studies such as gene annotation, comparative genomics, and identification of variations and regulatory elements. Long-read sequencing technology has enabled real-time generation of reads spanning thousands of nucleotides with the maximum reported length exceeding 2 Mb. However, the error rate of the long reads is generally higher when compared to the Illumina short reads. Over the years, several long-read assemblers have been developed, based on different assembly algorithms, to overcome this deficit and generate improved and complete assembly. With this increase in the choice of computational tools, researchers face the dilemma of opting for the most suitable tool for their study. Therefore, it is imperative to compare the performance of the tools using a common dataset and certain evaluation metrics to identify a reliable approach for a specific study.

Objectives: The main aim of this study was to assess the assembly tools in the context of prokaryotic genome assembly utilising long-read data generated by Oxford Nanopore Technology (ONT) Sequences. The evaluation is based on accuracy, structural contiguity and performance metrics. Additionally, the study aims to devise an optimised and reproducible assembly pipeline informed by the outcomes of the evaluation, with the goal of achieving accurate, comprehensive genome assemblies for prokaryotic organisms.

Materials and Methods:

Data description: The evaluation study focused on assembling two different strains of Escherichia coli using reads generated by ONT. The first strain, E. coli ST131, is a pathogenic strain known for causing extraintestinal infections associated with urinary and bloodstream infections. The reads for this strain were obtained from the European Nucleotide Archive (ENA) and had the accession number ERR3284704. For the second strain, E. coli DH5alpha, a non-pathogenic strain, the

genomic DNA was extracted in-house and sequenced using our facility's ONT's portable sequencer, MinION.

Methods: The data was assembled using ten long-read assemblers, namely Canu, Flye, Hinge, Miniasm, NECAT, Nextdenovo, Raven, Smartdenovo, Shasta and wtdgb2, with their default parameters to assemble the data. The assemblies were evaluated using many criteria like contiguity, genome completeness and performance attributes like run-time and resource usage. Furthermore, the impact of preprocessing and polishing of the data on assembly was examined.

Results and Discussion: While we evaluated the tools based on factors such as contiguity, accuracy and computational requirements, all the assemblers chosen were open-access, easily installable and relatively user-friendly. Computational resource requirements are essential for performing assemblies. More extensive run time and high computational demand can limit the tool's usability. Under limiting resource assignment, Canu had the longest run time, whereas Minisam had the least. Shasta, Redbean, NextDenovo and Miniasm were in the fastest category, while Hinge, NECAT, Raven and Flye were middling, and Canu and SMARTdenovo were in the slowest.

Contiguity analysis of the genome was based on the metric N50, though it does not provide an accurate assessment of the performance of the assembler. The N50 of the data was more or less the same for different assemblers. The assembly with the largest N50 may still have misconnections, resulting in low-quality assemblies. Therefore, another metric, gene completeness using BUSCO, was used to evaluate the quality of the assemblies. NECAT, Canu, NextDenovo and SMARTdenovo showed excellent results in this regard.

It is important to note that preprocessing and post-processing of the assemblies significantly impact the quality. Preprocessing would include eliminating low-quality reads for assembly and error correction based on consensus. Some tools like Canu, NECAT and NextDenovo included an error correction module, that allowed preprocessing of data. On the other hand, certain tools like Flye and Smartdenovo provide consensus polishing feature that involves correcting the assemblies with the raw reads, a post-assembly process.

Keywords: ONT long-reads, genome assembly, prokaryotes, pipeline, benchmarking

Targeting Active PAK1 and γH2AX to Evade Radiation Resistance in Head and Neck Squamous Cell Carcinoma

Swetha Rajendran[1], Vishnupriya K[2], Micheal Gromiha[2], Suresh Kumar Rayala[2] and Ganesh Venkatraman[3]

Department of Human Genetics, FBMS&T, Sri Ramachandra Institute of Higher Education and Research[1], Department of Biotechnology, IIT-Madras[2], Department of Biomedical Sciences, SBST, VIT-Vellore[3]

rayala@iitm.ac.in

Abstract

Head and Neck Squamous Cell Carcinoma (HNSCC) is the sixth most common cancer worldwide. HNSCCs are usually diagnosed at later stages and are given with multimodal treatment (Surgical resection and CRT). Still, HNSCCs are highly aggressive cancers showing higher rates of distant metastasis and recurrence due to various reasons such as formation of multiple primary tumors at multiple sites (second primary tumors) via field cancerization, resistance to chemotherapy and radiation therapy. With literature evidence, we are aware that the oncogenic serine threonine kinase called p21 activated kinase 1 (PAK1) is altered in multiple cancers and it plays a major role in chemoresistance in breast, lung, and pancreatic cancer. Hence, we attempt to unravel the role of PAK1 upon radiation exposure in HNSCC. All the data generated were statistically analysed using student's T test in GraphpadPrism software. We confirm that PAK1 is overexpressed in HNSCCs cell lines compared to normal using western blot. It has been validated in patient's samples, where PAK1 expression is significantly increased in HNSCC tissues (p value < 0.0001) compared to normal. Upon exposure to ionizing radiation, PAK1 gets activated in HNSCC. The cells exposed to different doses of ionizing radiation and tissue samples of patients underwent radiation therapy were used to analyze the activated PAK1 levels using western blot, IHC and in-vitro kinase assay. It is evident that the pPAK1 levels were significantly increased in patients with radiation therapy (p value for pPAK1-S199 and pPAK1-T212 are 0.0012 and 0.0024 respectively) compared to the naïve tissues. PAK1 being a central hub for the activation of many oncogenic signaling pathways, and considering its role in cytoskeletal and ECM remodeling, EMT, cell cycle and apoptosis, targeting active PAK1 helps in improving the radiosensitivity. This has been proved by various functional assays in WT and PAK1 KO cells post irradiation. PAK1 knock out in SCC131 cells

decreased the aggressive nature of cancer cells and increased the sensitivity to radiation therapy. Since, we confirm PAK1 plays a major role in radiation response, we developed radioresistant SCC131 cells (RR) to study the PAK1 status. pPAK1 levels are higher in RR cells and is associated with aggressive phenotype of cells. We also confirm that the activated PAK1 is associated with poor survival (p value = 0.016 for pPAK1-S199) in HNSCC patients using KM plot. Interestingly, we found that the DNA damage marker – γH2AX is also elevated in radioresistant SCC131 cells. This molecule acts as a docking nexus for many DNA damage repair proteins and promote non-homologous end joining (NHEJ) or homologous recombination (HR). It is evident that elevated PAK1 activity induces γH2AX via MORC2, thereby increasing the DNA repair process, ultimately resulting in the survival of cells after exposure to ionizing radiation. Hence, we developed small molecule peptide inhibitors against γH2AX that mimics its interacting partners to block the recruitment of DNA repair proteins, thereby pushing the cells into radiation induced apoptosis. A peptide called PL-8 that mimics MCPH1 was found to be highly potential in-silico. SCC131 cells showed reduction in γH2AX upon treatment with PL-8 after exposure to ionizing radiation. As an outcome of γH2AX inhibition, significantly increased DNA damage was observed using comet assay in these cells indicating the effectiveness of γH2AX inhibition. Hence, along with PAK1 inhibition, γH2AX inhibition will help in sensitizing the cells to ionizing radiation, thereby improving radiation response. Targeting active PAK1 and γH2AX ultimately decreases the rate of recurrence and metastasis and improve the disease-free survival in HNSCC.

Keywords: Head and Neck Squamous Cell Carcinoma, Radiation resistance, PAK1, γH2AX, Peptide inhibitors

Unlocking New Avenues: Exploring Azoles, Statins, and Anti-cancerous Compounds Against Mucormycosis for Drug Repurposing through Molecular Docking

Nischay Biligi and Suresh P.K,

Department of Biomedical Sciences, School of Biosciences & Technology, VIT, Vellore, Vellore

p.k.suresh@vit.ac.in

Abstract

Background: Mucormycosis is a life-threatening fungal infection caused by ubiquitous, filamentous fungi from the Mucorales order, primarily Rhizopus oryzae. It poses a significant threat to life as it invades blood vessels. The major risk factors associated with Mucormycosis include COVID-19 infection or cancer treatments resulting in neutropenia, organ or stem cell transplantation, immunosuppressive therapies, corticosteroid treatment, diabetic ketoacidosis [DKA], deferoxamine therapy, and high levels of free iron in the bloodstream. Despite the availability of treatment options such as surgical debridement, antifungal therapy, and adjunctive therapies, the mortality rates for Mucormycosis remain alarmingly high. This study delved into various innovative therapeutic approaches to combat the challenges posed by Mucormycosis.

Results: The present study focused on the most crucial protein, Lanosterol 14 alpha-demethylase, which plays a significant role in the survival of Rhizopus delemar. The investigation began by analyzing the physiochemical, structural, and functional characteristics of this protein. Subsequently, molecular docking utilizing AutoDock Vina was employed to explore the interaction between the selected protein and a diverse dataset of compounds. The compounds encompassed various categories such as FDA-approved drugs, FDA-unapproved drugs, investigational-only drugs, and biologics, including azoles, anticancerous compounds, and statins. Computational analyses were performed to estimate the ADMET parameters and biological activity of these compounds.

Conclusions: Through our computational investigations, we identified nine primary ligands with potential inhibitory properties against Rhizopus delemar. These ligands, namely Nilotinib, Conivaptan, Atorvastatin, Lapatinib, Idarubicin, Irinotecan, Simvastatin, Saperconazole, and Opelconazole, demonstrated inhibitory effects on Lanosterol 14 alpha-demethylase. These

findings suggest that these compounds could be repurposed as potential drug candidates for the treatment of mucormycosis. Further optimization, formulation techniques, and subsequent in vitro and in vivo studies are warranted to explore the therapeutic potential of these compounds in greater detail.

Keywords: Mucormycosis, Covid-19, Molecular docking, Virtual screening, Drug repurposing

A Bio Inspired Sentiment Analysis of Social Media for Early Detection of Negative Emotions in Young Adults

Neha Tyagi[1], Deepshikha Bhargava[1] Anil Ahlawat[2]

Amity University[1], KIET Ghaziabad[2]

nhtyagi190@gmail.com

Abstract

This paper describes the reviews of deep learning techniques on Sentiment Analysis (SA). Sentiment Analysis (SA) becomes more prominent development of Natural Language Processing. It is mainly processed the sentiment of a text in an efficient manner and also it can perform social media analytics. In recent times, the usage of DL approaches solves a variety of issues in sentiment analytics. It can be directly applied on live data given that the feature set is large whereas, in the DL method, the classifier requires to be initially nourished or "trained" with the raw datasets and tune to cluster the sentiments into predefined classes. But it works efficiently on large texts with large feature support. SA refers to the computational study that assesses people's emotions and opinions towards an entity. Sentiments are based on your emotions in any situation.

The objective of this paper to explore different methodologies used on sentiments and how can we detect or comparison of emotions on online reviews by using different technologies. NLP (Natural language process) technique is used to access classify or mining the text NLP has various types of technologies by which text can be accessed and can be changed into the vector form by using word2Vec.

Keywords: Deep learning, Neural network Natural Language Processing, social media

Insights into the Role of Potassium Channels in Migraine

Girishwaran M and Sajitha Lulu S,

Integrative Multiomics Lab, School of Bio Sciences and Technology, Vellore Institute of Technology

ssajithalulu@vit.ac.in

Abstract

Rationale/Motivation: Migraine is a primary headache disorder characterized by unilateral pain usually with aura, that affects approximately 1 in 5 individuals in India (as per the Global Disease Burden Survey, 2019). The underlying biomechanical processes of migraine are still poorly understood, but novel discoveries and research are constantly being published. One of the major factors in susceptibility to migraine is the dysfunction of ion channels in the trigeminal nuclei and sensory cortices of the brain. While calcium channels and the sodium-potassium pump channels are well described in relation to familial hemiplegic migraine, the current understanding of the role of inheritable channelopathies and channel dysfunctions in familial and sporadic migraine with aura is still obscure. Potassium channels are well known in neurology as modulators and regulators of neuronal signaling and conductance, playing an important role in maintenance of the membrane potential and in the passing of electric currents through nervous cells. Therefore, potassium channel dysfunctions are potential causative or exacerbating factors in migraine pathogenesis, and could provide a valuable novel target for specific antimigraine prophylaxis.

Objectives: The review focus to provide an extensive review of the current literature surrounding the role of potassium channels in migraine

Methods: The authors collected articles from the PubMed literature database using the following search terms: (potassium channels) AND (migraine), ((potassium channels) AND (migraine)) AND (treatment), (potassium channels) AND (channelopathies), (voltage gated potassium channels) AND (channelopathies), (calcium gated potassium channels) AND (channelopathies), (ATP sensitive potassium channels) AND (channelopathies), (two-pore domain gated potassium channels) AND (channelopathies), and (inward rectifying potassium channels) AND (channelopathies).

Results: A total of 1821 non-unique articles were found, which were reduced to 116 articles after removing duplicate results, citations, irrelevant articles, and articles without a valid DOI index.

The findings reported in these 116 articles were then summarized and discussed. Articles were excluded for irrelevancy based on the following criteria:

1. Article Text mentioned neither migraine nor potassium channels in detail
2. Article text discussed migraine but did not make specific mention of potassium channels in relation to migraine
3. Article text discussed potassium channelopathies but did not mention migraine in detail

Conclusion: This review reveals that potassium channels play a significant role in migraine prognosis. Dysfunctions in KIR channels, K2P channels including TRESK and TREK-1, small and large conductance calcium-sensitive potassium channels (SKCa and BKCa), and voltage-gated potassium channels (KV) are known to affect the incidence and progression of migraine in the general populace. KATP openers can induce migraine like phenotype, but KATP blockers have so far not been effective in reducing the intensity of migraine headache. Potassium channels are a potential druggable target for migraine prophylaxis with several compounds currently in preclinical trials.

Keywords: Migraine, potassium channels, druggable targets, glibenclamide, levcromakalim, acrylamide (S-1)

Identification of CYP51 From Leishmania Major, Strain Friedlin as A Potential Drug Target and Repurposing of FDA Approved Drugs Against It to Uncover Drug Against It

Chandra Kanta Bhusal[1], Upninder Kaur[2] and Rakesh Sehgal[1]

Aarupadai Veedu Medical College and Hospital, Post Graduate Institute of Medical Education and Research[1], Post Graduate Institute of Medical Education and Research[2]

c.k.bhusal71@gmail.com

Abstract

Introduction: Cutaneous Leishmaniasis, mainly caused by Leishmania major, lacks reliable treatment options. This study aimed to identify potential drug targets in Leishmania major (strain Friedlin) and repurpose FDA approved drugs for treatment. Comparative metabolic pathway analysis and selection criteria, including interactome, druggability, essentiality, and gene ontology, identified lanosterol 14-alpha demethylase (CYP51) as a promising target.

Methodology: We conducted virtual screening of FDA approved drugs against CYP51 using PyRx, assessed complex interactions in Biovia Discovery, and redocked for validation. Isavuconazole and posaconazole emerged as top candidates for CYP51 binding. Fluconazole, a known inhibitor of CYP51 was docked against the same as control. Molecular dynamics (MD) simulations over 100 nanoseconds assessed complex stability. Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and hydrogen bond analysis measured atomic distances. Principal component analysis explored essential system dynamics on a low-dimensional free energy landscape.

Results: Our study has identified CYP51 as an ideal drug target. Isavuconazole and posaconazole emerged as the most promising candidates for CYP51 binding, exhibiting binding scores of -10.4 and -9.3, along with 3 and 5 hydrogen bonds, respectively. In contrast, the control, Fluconazole, demonstrated a binding energy of -7.1 with only 1 hydrogen bond. Molecular dynamics (MD) simulations subsequently confirmed the stability of the complexes, with root mean square deviation (RMSD), root mean square fluctuation (RMSF), and hydrogen bond analysis supporting the structural integrity. Principal component analysis provided valuable insights into the essential dynamics of the system.

Conclusion: This study suggests CYP51 as a promising target and identifies isavuconazole and posaconazole as potential treatments for Cutaneous Leishmaniasis. MD simulations and analysis indicate complex stability and dynamics, highlighting avenues for effective treatment development.

Keywords: Virtual Screening, Docking, MD Simulation, Computer-aided Drug Design

Unmasking Alzheimer's Central Players: A Network Perspective

Sanga Mitra[1], Carlos Viesi[1], Marcus Seldin[2] and Manikandan Narayanan[1]

Indian Institute of Technology (IIT) Madras, Chennai, India[1]

Department of Biological Chemistry & the Center for Epigenetics and Metabolism, University of California, Irvine, CA[2]

nmanik@cse.iitm.ac.in

Abstract

Alzheimer's disease (AD) is a complex neurodegenerative condition with global health implications. Using computational approach like differential correlation (DC) analysis1 and MultiCens2 network analysis, we have demonstrated how these genes reshape gene correlation networks across various brain regions in AD and highlighted the genes with high centrality, signifying their importance. One key discovery in this study is the identification of previously unrecognized hubs of dysregulation in AD, represented by genes like ZKSCAN1, SLC5A3, RCC1, GPD1, PLK4, PPDPF etc. To delve into the roles of these hub genes, the research employed a comprehensive approach that combined computational analysis and experimental methods. MultiCens' query-set centrality approach revealed that hub genes, including SLC5A3, RCC1, GPD1, and PLK4, are central players in gene networks related to synaptic signaling genes (SSG) and plaque-induced genes (PIGs), emphasizing their importance in AD pathogenesis. To assess the impact of key genes like ZKSCAN1 on microglial (HMC3 cell line) and oligodendrocyte (HOG cell line) cell cultures, these human brain cell lines were cultured in conditioned media collected from the transfection assay. Candidate genes were incorporated into plasmids and transfected into HEK293 cells. Conditioned media collected from these cells were then applied to HMC3 and HOG cell lines, leading to notable changes in the expression of Alzheimer's disease (AD) biomarkers, including APOE, PSEN1, TREM2, and SORL1. This demonstrated the influence of specific genes on the modulation of AD-related gene expression in these cell lines. The research's future plans include RNA-sequencing and comprehensive bioinformatics analysis to uncover the molecular players and pathways affected by hub genes like RCC1 and GPD1. By using MultiCens for gene significance assessment, the study aims to reveal the mechanistic pathways associated with these hub genes and their downstream targets identified through RNA-Seq. In summary, this study enhances our understanding of AD's molecular basis.

Keywords: Inter-brain-region analysis, Differential Correlation, Rewired network, Hub genes, Conditioned Media

Machine Learning Heuristics for Oral Cancer Datasets to Ascertain Pathogenesis

Sai Bharath Natte[1], Raja Pavan Karthik[1], Sneha Kh[1], Adhithya Sudeesh[1] and Kavi Kishor Pb[2]

Amrita vishwa vidyapeetham[1]

Osmania university[2]

pbkavi@yahoo.com

Abstract

Oral cancer, an extensive and life-threatening disease, poses a significant global health challenge due to its late-stage diagnosis and limited treatment options. Detection is critical for improving patient outcomes and reducing mortality rates. In recent years, machine learning has emerged as a promising tool in the field of medical diagnostics. We aimed to explore the application of machine learning models in the detection of oral cancer, offering a novel approach to address this pressing healthcare issue. This study presents a comprehensive investigation into the development of a machine learning-based model for the detection of oral cancer. Creating and using a diverse dataset of oral cancer images, our research focuses on feature extraction and classification algorithms to identify subtle changes associated with cancerous and non-cancerous.

Keywords: Oral Cancer, Challenges, Bioinformatics, Machine Learning

Exploration of Cannabis Constituents as Potential Inhibitors Against Gamma-Secretase to Manage Alzheimer's Disease: A Structural Bioinformatics Approach

*Premkumar T, Radhika G, Sajitha Lulu S**

Integrative Multiomics lab, School of Bio Sciences and Technology

Vellore Institute of Technology, Vellore – 632014, Tamil Nadu, India

ssajithalulu@vit.ac.in

Abstract

Rationale/Motivation: Alzheimer's disease (AD) is the most common form of dementia and, is characterized by irreversible and progressive neurodegeneration. The breakdown of Amyloid Precursor Protein (APP) plays a crucial role in AD development and three proteolytic enzymes are found to be critically involved in the APP breakdown process alpha-secretase, beta-secretase, and gamma-secretase. Among these three enzymes, the gamma-secretase is highly significant as it is involved in cleaving the APP and generating beta-amyloid plaques in the extracellular neuronal membrane. Gamma-secretase dysfunction has been reported in AD, and several gamma-secretase inhibitors, including natural compounds and synthetic analogs, have been developed to treat AD. However, there is currently no treatment for AD, as most drug-like compounds have failed in clinical trials.

Objectives: The primary objective of our study is to understand the structural behaviors of gamma-secretase. Identification of novel inhibitors from cannabis plant sources

Methods: This study focuses on screening compounds from cannabinoid plant sources, against gamma-secretase to identify inhibitors to managing AD. The cannabis sativa plant contains a wide range of phytochemicals that can be used to treat various diseases. A total of 120 compounds from cannabis plants were studied using molecular docking methods. Furthermore, we analyzed and compared the interaction profiles of a cohort of medicinal plant compounds that have been recognized for their neurological relevance with those of cannabis compounds. Molecular Docking calculation was carried out using Auto dock vina and Molecular Dynamics Simulation was performed for a time scale of 200ns for the identified complexes using CHARMM36 force field in GROMACS. The trajectory analysis of Root Mean Square Deviation and Fluctuation

(RMSD/F), Solvent Accessible Surface Area (SASA), Hydrogen bonds, and Radius of Gyration (Rg) analysis was done by GROMACS and the ggplot Tidyverse package used to interpret the results.

Results: In this study, the virtual screening was performed on a set of 120 compounds from cannabis plants and 35 compounds from medicinal plants by using molecular docking to examine their binding ability. The docking results revealed that screed compounds had strong binding affinity and interactions with gamma-secretase. Notably three analgesic compounds JWH-200, JWH-018, and JWH-166, had binding energies of -10.4 Kcal/mol, -9.9 Kcal/mol, and -9.0 Kcal/mol respectively. The control compounds including GinsenosideRd, Withanone, withanolide_A, and GinsenosideRd had binding energies of -10 Kcal/mol, -10 Kcal/mol, and -9.4 Kcal/mol respectively. Furthermore, the 200ns molecular dynamics simulation results also suggested that the complexes are stable throughout the simulation. According to our study, the identified cannabinoid compounds could be a prime compound against gamma-secretase activity.

Conclusion: In conclusion, our analysis has revealed the potential of JWH-200 and GinsenosideRd as potential inhibitors against gamma-secretase activity. The results demonstrate both compounds had favorable binding affinities and were stable during simulation.

Keywords: Alzheimer's disease, Drug target, cannabinoids, Molecular docking, Molecular dynamics simulations

An Immunoinformatic Approach to Assess the Cross Reactivity of Indian Cobra Venom to Viper Venom

Vishnu Priya and Rani P

Department of Biotechnology, PSG College of Technology

rani.bio@psgtech.ac.in

Abstract

Rationale/ Motivation: The primary limiting step in the immunodiagnostic strip will likely be overcome by predicting the venom's cross-reactivity and taking measures towards it.

Objectives: To predict the cross-reactivity of Indian cobra (Naja naja) venom to viper venom by means of common epitopes.

Background: In tropical and subtropical nations, snake envenomation is a neglected public health issue that frequently results in life-threatening circumstances. The World Health Organization (WHO) reports that India records the highest number of snake envenomation cases annually roughly 81,000 with between 35,000 and 50,000 fatalities. Death due to cobra bites are mainly due to the inefficiency of existing polyvalent anti-snake venom's (ASV). Successful antivenom therapy depends on accurately identifying the snake species that a patient has been bitten by and the presence of venom in their bodily fluids. The availability of an immunodiagnostic strip as a point-of-care test would therefore aid in identifying the offending snake; but venom is a highly cross-reactive as it is a mixture of several proteins from several families, which could lead to a false positive during the diagnostic process. The purpose of this research is to forecast whether Indian cobra venom cross reacts to that of vipers.

Methodology: Sequences from the NCBI were used to build the datasets for Indian cobra venom, which contained 44 sequences from eight different protein families, and viper venom, which contained 68 sequences from ten different protein families. The dataset's B cell, MHC I, and MHC II epitopes were predicted using IEDB tools. The B cell epitopes were predicted using Bepipred-2.0, Kolaskar and Tonganskar, Emini surface accessibility tools; the T cell MHC I and MHC II epitopes were predicted using NetMHCpan 4.1 and MHC II-NP, respectively. The common epitopes between the two snake's datasets in each epitope group were then found using the Interactivenn tool. Kolaskar and Tonganskar and Class I immunogenicity tools were utilized to determine antigenic and immunogenic epitopes from the common epitopes

Results and Discussion: Using the previously listed tools, 749 B cell epitopes, 1406 MHC I epitopes, and 215 MHC II epitopes were predicted for the Cobra dataset, and 1125 B cell epitopes, 2256 MHC I epitopes, and 335 MHC II epitopes were predicted for the Viper dataset. The two snake datasets were then found to share 15 B cell epitopes, 10 MHC I epitopes, and 0 MHC II epitopes. Form the common epitopes antigenic and immunogenic epitopes were filtered out. This resulted in four antigenic B cell and immunogenic T cell (MHC I) epitopes that belong to the families Phospholipase A2 (PLA2), Snake venom serine-protease (SVSP), Snake venom metalloproteinase (SVMP), cysteine-rich secretory proteins (CRISP), and L-amino acid oxidase (LAAO). Based on the venom composition of the abovementioned families, it is determined that cross reactivity may occur due to Phospholipase A2 (PLA2) family protein. But from the literature it is known that three finger toxin is predominant in cobra venom which is 63.3% whereas phospholipase is only of 11.4% contributing very less for the antibody development. So, it is concluded that the immunodiagnostic strip developed specific for cobra venom detection may not produce false positive results with viper venom

Conclusion: It is suggested from the results that there might not be a cross-reactivity between the venom of cobras and vipers.

Future work: The above-mentioned methodology can be developed into a pipeline and utilized for the prediction of cross reactivity between venoms prior to wet lab work.

Keywords: Immunoinformatic approach, cross reactivity, Indian cobra venom, viper venom

Waste Segregation using Deep Learning Neural Network

Anjana S Nair, Lakshmi A R, Samvrutha Prasad, Sona Gokul, Joshy Alphonse and Nidheesh M,

Amrita school of biotechnology

nidheesh@am.amrita.edu

Abstract

Waste segregation is the process of categorizing waste, such as biodegradable and non-biodegradable waste. This is critical for decreasing garbage's environmental impact since biodegradable waste can be composted or recycled, but non-biodegradable waste must be disposed of in a landfill. Using machine learning to classify trash photos is one method for automating waste segregation. The goal of this research is to create a machine learning model that can categorize garbage photos as biodegradable or non-biodegradable. A dataset of labelled waste photos will be used to train the model. The dataset will be classified as biodegradable or non-biodegradable. Each image in the collection will be labelled with the category to which it belongs. A number of machine learning approaches will be used to train the model such as convolutional neural networks (CNNs). CNNs are a sort of neural network that excels at picture classification. Once trained, the model can be used to classify new garbage photos. To accomplish this, the model will take a garbage image as input and estimate if the waste is biodegradable or not. The model can be used to create a waste segregation system that can automatically classify waste photos and categorize the waste. This technology has the potential to be employed in a variety of situations, including residences, businesses, and waste management facilities. The model can also be used to educate people about the importance of trash separation. The approach, for example, might be used to create a smartphone app that allows individuals to shoot images.

Keywords: Waste Segregation, Convolutional Neural Networks (CNNs), Image Classification, Training and Validation

Unlocking Insights in CNS Cancer through Metabolomics and Multi-Omics Integration

Megaswana Guruprasad and Mohanapriya Arumugam

Department of Biotechnology, Vellore Institute of technology (VIT), Vellore, Tamil Nadu, INDIA

mohanapriyaa@vit.ac.in

Abstract

Central nervous system (CNS) cancers, while rare, impose a significant burden on patients and have high mortality rates. The scarcity of data historically associated with these malignancies has driven innovative approaches to maximise information from diagnostic tests, leading to the adoption of high-throughput technologies, including metabolomics in the field of systems biology. Metabolomics focuses on profiling small-molecule biochemicals (metabolites) within biological systems, offering insights into the complex metabolic alterations in neuro-oncology. It plays a vital role in identifying potential quantitative metabolic biomarkers for early cancer detection and evaluating treatment effectiveness.

The journey to discover these biomarkers begins in a preclinical setting, involving the use of animal models and human cell cultures, such as the glioblastoma-derived U87MG cell line, to identify candidate biomarkers. Subsequent translational validation efforts ensure their clinical relevance by confirming them in biofluids or tumour tissues. Neuro-oncology is inherently multidisciplinary, spanning domains like imaging, histology, and molecular data, including genomics, epigenomics, proteomics, and metabolomics. Techniques such as nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS) are powerful tools for characterising the complex metabolic landscape in neuro-oncology.

A major advancement in neuro-oncology research is the consolidation of big data resources through open-access initiatives, enabling integrated multi-omics analysis to discover novel biomarkers and therapeutic interventions. Metabolomics, in conjunction with other "omic" disciplines, deepens our understanding of neuro-oncology, offering prospects for improved diagnosis, treatment, and patient outcomes. Metabolomic databases and tools like MetaboAnalyst and the Human Metabolome Database (HMDB) are instrumental in metabolomics research, facilitating data analysis and interpretation. For brain cancer research, oncology data sources include repositories of brain cancer patient data, tissue samples, and cell lines, such as the Cancer

Genome Atlas (TCGA) and the Human Glioblastoma Cell Atlas. These resources contribute to a comprehensive understanding of the molecular and metabolic complexities in CNS cancers.

Keywords: Metabolomics, Central nervous system (CNS) cancers, Neuro-oncology, Biomarkers, Multi-omics analysis, Nuclear magnetic resonance spectroscopy (NMR), Mass spectrometry (MS)

Comparative Atomistic Insights on Apo and ATP-I1171N/S/T in Non-Small-Cell Lung Cancer

Ambritha Balasundaram and George Priya C Doss

Laboratory of Integrative Genomics, Department of Integrative Biology, SBST, VIT

georgepriyadoss@vit.ac.in

Abstract

Anaplastic lymphoma kinase (ALK) rearrangements occur in about 5% of non-small cell lung cancer (NSCLC) patients. Despite being first recognized as EML4-ALK, fusions with several additional genes have been identified, all of which cause constitutive activation of the ALK kinase and subsequently lead to tumor development. ALK inhibitors first-line crizotinib, second-line ceritinib, and alectinib are effective against NSCLC patients with these rearrangements. Patients progressing on crizotinib had various mutations in the ALK kinase domain. ALK fusion proteins are activated by oligomerization through the fusion partner, which leads to autophosphorylation of the kinase's domain and consequent downstream activation. The proposed computational study focuses on understanding the activation mechanism of ALK and ATP binding of wild-type (WT) and I1171N/S/T mutations. We analyzed the conformational change of ALK I1171N/S/T mutations and ATP binding using molecular docking and molecular dynamics simulation (MDS) approach. According to Principal component analysis (PCA) and Free energy landscape (FEL), it is clear that I1171N/S/T mutations in Apo and ATP showed different energy minima/unstable structures than WT-Apo. The results revealed that I1171N/S/T mutations and ATP binding significantly supported a change toward an active state conformation, whereas WT-Apo remained inactive. We demonstrated that I1171N/S/T mutations are persistent in an active state and independent of ATP. The I1171S/T mutations showed greater intermolecular H-bonds with ATP than WT-ATP. The molecular mechanics poisson-boltzmann surface area (MM/PBSA) analysis revealed that I1171N/S/T mutations binding energy were similar to the WT-ATP. This study shows that I1171N/S/T can form stable bonds with ATP and may contribute to constitutively active kinase. Based on the Y1278-C1097 H-bond and E1167-K1150 salt bridge interaction, I1171N strongly promotes constitutively active kinase independent of ATP. This structural mechanism study will aid in understanding the oncogenic activity of ALK and the basis for improving the ALK inhibitors.

Keywords: Anaplastic lymphoma kinase, ATP, NSCLC, Mutation, I1171X, oncogenic activation

L. acidophilus Mv Hampers the Gtfb Enzyme and Impedes the Formation of Dental Plaque by S. Mutans: An In Vitro and In Silico Approach

Venkatramanan M and Nalini Easwaran,

Vellore Institute of Technology

nalini.e@vit.ac.in

Abstract

Rationale: Oral infections mediated by Streptococcus mutans pose a serious challenge to human health across the globe. This opportunistic pathogen is implicated in aggressive periodontitis and leads to various systemic disorders, which result in a significant socioeconomic burden. The conventional treatment of oral infections with antibiotics has led to the emergence of resistant strains due to selective pressure. Furthermore, the efficacy of synthetic chemicals as antimicrobials may decline over time due to inconsistent usage. Hence, natural agents may provide a superior alternative. We aim to investigate the efficacy of membrane vesicles (MV) derived from Lactobacillus acidophilus, a commensal microorganism of the human gut and oral niche, to inhibit S. mutans infection. Membrane vesicles are nanosized spherical entities that originate from the outer membrane. They harbor antimicrobial peptides (Bacteriocin), protein, DNA and RNA. The antimicrobial peptides in membrane vesicles may exhibit a narrow or broad spectrum of activity, which could be harnessed for innovative therapeutic approaches.

Materials and methods: We isolated and characterized L. acidophilus membrane vesicles using a system bioscience kit, DLS, NTA and FESEM. We assessed the impact of membrane vesicles on S. mutans viability, EPS production and biofilm formation in vitro. We performed homology modelling of bacteriocin using the Swiss model online tool and targeted it against gtfB, a key protein involved in biofilm formation by S. mutans, using molecular docking and dynamics approach. We confirmed the modulation of gtfB expression by membrane vesicles using RTPCR in vitro.

Results and discussion: L. acidophilus, a commensal bacterium of the human gut, possesses immunomodulatory, antitumor and antimicrobial properties. However, the role of its membrane vesicles (MVs), which are nanoscale spherical structures released from the bacterial membrane,

remains elusive. In this study, we aimed to isolate and characterize MVs from L. acidophilus and investigate their antimicrobial potential against S. mutans, a major causative agent of dental caries. We employed Dynamic Light Scattering (DLS) and Nanoparticle Tracking Analysis (NTA) to determine the size distribution and particle concentration of the MVs, which ranged from 100 to 300 nm and reached 2.7 X 109 particles, respectively. We also visualized the morphology of the MVs by FESEM, which revealed spherical structures with smooth surfaces. We then assessed the antimicrobial and antibiofilm efficacy of the MVs against S. mutans in vitro using plate and broth assays. The MVs exhibited remarkable activity against S. mutans, which was abolished by proteinase K treatment, indicating that bacteriocins, which are antimicrobial peptides, were responsible for the activity. To identify the bacteriocins involved, we retrieved their sequences from L. acidophilus based on a previous study by Dean et al and performed molecular modelling and Ramachandran plot validation to predict their structures. We also analyzed their physicochemical properties and toxicity using DBAASP and Toxinpred online tools. We selected gtfB as the target protein of S. mutans from string analysis, as it is a key virulence factor that mediates adhesion, aggregation and biofilm formation of S. mutans in the oral environment. We conducted molecular docking of the bacteriocins with GtfB using Haddock online server and found that Lactacin B had the highest affinity for GtfB with a binding energy of -15.8 kcal/mol. We validated the interaction stability by GROMACS simulation. We also performed RT PCR to measure the expression of gtfB in S. mutans treated with MVs and found that it was significantly downregulated. These findings demonstrate that L. acidophilus MVs exert antimicrobial and antibiofilm effects by modulating gtfB expression in S. mutans.

Conclusion: The potential of membrane vesicles (MVs) derived from Lactobacillus acidophilus as a novel therapeutic strategy against oral pathogens is demonstrated in this study. This approach could help mitigate the adverse impacts of oral diseases on human health and well-being, as well as the economic costs associated with their treatment. The antimicrobial activity of the peptides contained in the MVs was not directly assessed in this work, but it will be investigated in future experiments.

Keywords: Membrane vesicles, Oral pathogen, In silico, Biofilms

Exploring Different Traits of the Isolate Priestia megaterium VIT For Maintaining Equilibrium in Energy Expenditure Between Systemic Resistance and Plant Growth

Huldah Lazarus and Nalini Easwaran, Vellore Institute of Technology

nalini.e@vit.ac.in

Abstract

Rationale: Population growth leads to higher demand for food crops, which also face more disease challenges. However, the technologies that enhance crop yield also make them more vulnerable to various biotic and abiotic stresses, especially diseases. Plants have to balance their energy allocation between growth and immunity when they encounter pathogens, resulting in trade-offs. Therefore, using biocontrol agents may help plants to improve both growth and immunity and contribute to sustainable development. PGPR (Priestia megaterium) is a promising biocontrol agent that can provide plant health protection through different mechanisms, such as quorum quenching, antimicrobial secondary metabolite production, and osmoprotectant production. The PGP-related genes can enhance plant growth and health. Moreover, the genes encoding for bacteriocin and rhizobactin can prevent phytopathogenic infections. Additionally, the presence of heat and cold shock proteins allows them to survive in harsh environments and assist plants in disease prevention and growth promotion. Therefore, exploiting Priestia megaterium can have dual benefits: one is by boosting plant growth and the other is by inhibiting phytopathogenic infections.

Materials and Methods: To assess the PGPR potential of isolated strains, we performed biochemical assays to measure their antibiotic and antimicrobial resistance, salt and pH tolerance, temperature range, and phosphate solubilization capacity. We sequenced the genomic DNA of the strains using the Illumina platform with 150 bp paired-end reads. We assembled the reads using SPAdes software and validated the assembly using barrnap version 0.9. We completed the genome using the MeDuSa web-based tool.

Results and Discussion: Biochemical tests identified one of the isolated bacteria from biofertilizer as PGPR. The strain was a Gram-positive, rod-shaped bacterium with glossy white, entire, and umbonate colonies. It showed tolerance to a pH range of 3-10, temperature range of 20-40, and salt range of 2-8%. It was resistant to ampicillin but sensitive to other antibiotics tested. It also

exhibited anti-microbial activity against the plant pathogen P. syringae. Whole genome sequencing of the isolate using the Illumina platform showed 99% similarity with Priestia megaterium. The whole genome sequencing revealed a total read length of 5,415,392 bp and an N50 value of 431,205. The GC content was 38% and the genome size was within the reported range. Several genes related to plant growth promotion and anti-microbial properties were predicted in the genome, such as glucose dehydrogenase Zwf, trehalose metabolism genes like treP,C,R, heat shock proteins GroeS, GroeL, cold shock proteins CspA,B,D,E, glycine-betaine OpuD, thiol peroxidases Bcp-1,2, glutathione peroxidase BsaA, catalases CotJC,E,H,JB,NE,R,S, SodAC, GABA production GabP,D. Glycine-betaine is an osmoprotectant that helps plants cope with abiotic stresses by osmoregulation. Trehalose is another stress protectant and the heat and cold shock proteins help bacteria survive in harsh environments. Peroxidases protect bacterial cells from oxidative and osmotic stresses. Bacteriocins and rhizobactins identified through sequencing are antimicrobial peptides and siderophores, respectively, encoded by genes that have been widely reported to inhibit bacterial growth by competing for essential nutrients. Thus, both growth as well as disease control can be achieved through this strain P. megaterium VIT.

Conclusion: Current research aimed to achieve sustainable development in the face of increasing disease severity and crop loss due to population growth. There is a global shift in the agricultural paradigm towards sustainability. Current research goal considers disease control and growth promotion as interrelated aspects of plant health. WGS analysis of the isolate identified genes that enabled the isolate to adapt to harsh conditions and enhance plant health by phosphate solubilization, siderophores production, osmoprotection, stress tolerance and phytopathogen inhibition by anti-microbial activity.

Keywords: Priestia megaterium, WGS, PGPR, sustainable development

Assessment of Anti-inflammatory Potentials of Rosa indica

Santhi N, Sahana Prajwala M, Raja Dharani A and Nithya S

Avinashilingam institute for Home Science and Higher Education for Women

23pbi002@avinuty.ac.in

Abstract

Inflammation is a multifaceted process implicated in numerous pathologies, including male infertility, a condition affecting couples worldwide. Rosa indica, known for its traditional medicinal use, has garnered interest for its potential anti-inflammatory and reproductive health benefits. In this study we examined the anti-inflammatory and antioxidant properties of fresh and dried Rosa indica petals by analysing 20 to 100 µl of 1 mg/ml concentration of extract using albumin denaturation and DPPH (2,2-diphenyl-1-picryl-hydrazyl-hydrate) assays. The extracts were characterized using GC-MS analysis to identify their bioactive constituents. To evaluate the efficiency of the extracts, Schrodinger LLC., Software (GLIDE SP and GLIDE XP) were employed and validated using MM-GBSA module. The ethanolic extract exhibited significantly high anti-oxidant activity (83.18±0.03 and 90.63±0.02 %) and anti-inflammatory effects (59.96±0.49 and 57.27±0.16 %) for fresh and dry petals respectively at p-vlaue <0.001. Molecular docking studies identified Kaempferol as a top hit compound with a docking score of -9.048 for cyclooxygenase 2; -9.226 for androgen receptor; and -5.56 for AKT1. The compound 2,4-DTBP showed the better docking score of -6.379 than Kaempferol with -4.846. Least binding energy was recorded for the best docked compound Kaempferol. Molecular Dynamics simulation studies were further performed to see the stability of the complex structure.

This research offers insights into the molecular interactions of Rosa indica extracts with inflammatory proteins and the target proteins involved in the process of steroidogenesis shedding light on potential mechanisms for their properties. These findings not only support the traditional use of Rosa indica in herbal medicine but also lay the groundwork for further experimental validation and the development of novel anti-inflammatory interventions. Furthermore, this study underscores the utility of computational methods in the initial screening of natural compounds, potentially accelerating the discovery of therapeutic agents for male infertility and related conditions. The discoveries made in this work provide a promising avenue for advancing clinical research and therapeutic development.

Keywords: Rosa indica, Antioxidant, Male infertility

Transcriptomic Analysis Reveal Tissue-Specific Gene Regulatory Circuits Associated with Systemic Lupus Erythematosus: A Systems and Computational Study

Blessy Kiruba, Akshayata Naidu, Sajitha Lulu S and Vino Sundararajan, Vellore Institute of Technology,

ssajithalulu@vit.ac.in

Abstract

Rationale: Systemic lupus erythematosus (SLE) is a complex autoimmune disorder characterized by immune system dysregulation, leading to widespread inflammation in multiple organs. Loss of immune tolerance against auto-antigens that are produced by poor clearance of apoptotic bodies remains causes tissue damage. Molecular intricacies involved in the process are not fully characterized and need further elucidation for the development of more effective and precise treatment options.

Objective: To deduce gene regulatory circuits associated with the pathophysiology of SLE using RNASeq data from PBMC samples. To deduce gene regulatory circuits associated with the pathophysiology of SLE using RNASeq data from specific cells/tissues (Neutrophils, Dendritic cells, Macrophages). To construct a molecular atlas of inter and intra-cellular gene regulatory circuits involved in SLE for identification of molecular mediators and pathways as therapeutic targets

Methods: Differentially expressed genes (DEGs) were retrieved from gene expression datasets derived from PBMC samples, Neutrophils, Dendritic cells and Macrophages

A curated list of genes was used as input for the construction of Protein-Protein Interaction (PPI) network and topological network analysis was performed for the identification of Hub Genes Network Modules. Associations between identified hub genes/modules were further characterized using corelation, clustering and regression analysis using an integrated dataset of SLE patients (>120 samples) for the establishment of gene regulatory circuits. The generated circuits were integrated and compiled and validated from literature for the construction of a molecular atlas associated with the pathophysiology of SLE.

Results: Several genes induced by interferons (IRF7, IFI35, IFT1, OAS2, etc.) were identified as hub genes along with genes involved in DNA-damage (PARP9, PARP14). Hierarchical clustering identifies several gene clusters associated with the pathophysiology of SLE, prominent among which were clusters with interferon induced mediators and with microRNAs involved in RNA silencing (MIR553, MIR3173, MIR644A, MIR199A1, etc). Hereafter regression analysis revealed key gene regulatory circuits associated with the regulation of interferon mediated signalling pathways involved in the pathophysiology of SLE in PBMC samples and in different leukocytes.
Conclusion: Systems and Computational analysis of transcriptomics dataset reveal prominent gene regulatory circuits and key regulators associated with the molecular pathogenesis of SLE

Keywords: Non-coding RNA, Systemic Lupus Erythematosus, Tissue-specific gene expression, RNASeq data, Clustering, Regression

In Silico Screening of Phytochemicals from Indian Medicinal Plants for the Identification of Potential Antibacterial Activity Against Mycoplasma Penetrans

Kandavelmani Angamuthu, Princy Caron, Sanjana Sunil and Surega Madhavan,
Department of Biochemistry, Biotechnology and Bioinformatics, Avinashilingam university, Coimbatore, India
kandavelmani_bc@avinuty.ac.in

Abstract

Mycoplasma penetrans, a species of Mycoplasmataceae, is a free-living, gram-positive bacterium having a reduced genome size of 1.3Mb. M.penetrans is a primary pathogen responsible for urinary tract infections and breathing disorders. Current treatment method involves the use of antibiotics and antiviral drugs to prevent the disease progress. Due to the increasing mutation rates, many microorganisms have obtained antimicrobial resistance. To prevent this ,we identified phytocompounds from 20 Indian medicinal plants which can be used to combat the activity of M.penetrans. Several attempts have been made to investigate Indian medicinal plants recognised for their health benefits and immune-boosting properties as well as to explore the possibility of repurposing current medications known for their antibacterial activities. The current study is focused on insilico screening of phytochemical compounds against carbamate kinase and arginine deiminases. Research studies have identified carbamate kinase and arginine deiminase enzymes of M. penetrans as potential drug targets. In addition to their functional significance, these enzymes have also been found to possess high chemical and proteolytic stability. In the present study, we have selected 114 phytocompounds from numerous Indian medicinal plants. These phytocompounds were subject to a series of systematic bioinformatics analyses. Based on the combined results of these analyses, a set of 59 phytocompounds were shortlisted. Inhibitory activity of these compounds against carbamate kinase and arginine deiminases was studied using virtual screening approach. Phytocompounds with highest binding efficiencies were then identified. These compounds could be used as promising lead molecules for the development of novel antibiotics to combat infections caused by M. penetrans.

Keywords: M Penetrans, Urinary infections, Antibacterial activity.

Comparative Analysis of Physicochemical Features of Structured and Disordered Proteins in Humans

Prachi Bhargava, Bishal Samanta and Amita Barik

National Institute of Technology, Durgapur

amita.barik@bt.nitdgp.ac.in

Abstract

Proteins play a crucial role in a myriad of biological processes ranging from enzymatic catalysis to structural support, underscoring their significance in the living world. The present study focuses on the comparative analysis of two distinct classes of proteins: structured and intrinsically disordered proteins (IDPs) present in humans. IDPs lack a stable three-dimensional structure and possess high conformational flexibility, serving as critical hubs in protein interactions. They are involved in various cellular processes and in the onset and progression of diseases such as cancer, neurodegenerative disorders, and cardiovascular conditions. The study employs a non-redundant dataset of 1090 disordered proteins containing 1637 intrinsically disordered regions (IDRs) of Homo sapiens curated from DisProt database. The length of the IDPs in our dataset varies from 24 residues to 34,350 residues while the IDR length range from 9 to 2152 amino acids. We find that the IDRs are enriched in polar and charged residues (29%) and depleted in aromatic amino acids (5%) with tryptophan being the least (0.7%). Structured polypeptide chains curated from UniProt on the other hand have higher content of aromatic residues (8%) and have relatively low polar charged residues (23%). Cysteine content, however is more in case of structured regions as compared to the disordered regions. Overall content of non-polar residues in IDRs are less (40%) as compared to structured proteins (42%), however, proline (8.3%) and glycine (7.9%) are more preferred in IDRs than in the structured proteins (Pro 5.8%, Gly 6.4%). Also, the contribution of individual residues is not dependent on the length of IDRs. This comparative analysis highlights notable distinctions in the behaviours of structured and disordered proteins, unveiling factors that govern their sequence-structure-function relationships. These findings offer valuable insights into the complexity of cellular processes and developing therapeutic strategies for diseases involving protein misfolding and dysfunction.

Identification of the Common Differentiating Genes Among Various Cancers on the Basis of RNA Expression

Krishna Sayantika Deb, Kuntal Pal and Rahul Shaw

Vellore Institute of Technology

kuntal.pal@vit.ac.in

Abstract

Various types of cancer prevalence worldwide make it the second biggest killer disease. The gene expression analysis was conducted from the public expression profile data of five cancer cells (breast, lung, glioma, cervical and Thyroid cancers). The comparative RNA-seq data analysis of expressing and regulatory genes in cancer cells may provide a detailed insight into their role in cancer metabolism. The study was done to identify the common upregulated and downregulated genes and their pathways among multiple types of cancers, based on their gene expression profile. The gene expression quantification from multiple RNA-seq data obtained from Gene Expression Omnibus (GEO) for Human was performed using Salmon. Thereafter, Differential Gene Expression (DEG) analysis among multiple cancers done by edgeR DEG analysis pipeline in R. The comparative gene expression analysis identifies common up-regulated and down-regulated genes among five cancers. The differential expression patterns of various genes among these five cancers show PRSS23 and LINC commonly up-regulated in glioma, breast and lung cancers. PSR23, belongs to serine protease family via estrogen receptor pathway regulates the proliferation of cells. The commonly down-regulated genes are SNHG and LINC, among breast, lung, and thyroid cancers. The common up-regulated and down-regulated genes can be further explored as common potential biomarkers. The DEG analysis for regulatory genes in five cancers was further explored by STRING analysis and the KEGG pathway for their role in the progression of cancers while regulating various metabolic pathways involved in cancer metabolism.

Keywords: cancer, gene expression, RNA-seq data, cancer metabolism, upregulated downregulated, Gene Expression Omnibus, edgeR DEG analysis

In Silico Analysis of Atropine as a Potential Inhibitor of NS5 Protein of Japanese encephalitis Virus

Toufeeq Ahmed A[1], Priya V[1], Swagatika Priyadarsini[2], Faraz Ahmad[1] and Mohanapriya Arumugam[1]

Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, India[1],

Scientist, Biochemistry unit, ICAR- National Research Centre on Camel, Rajasthan[2].

faraz.ahmad@vit.ac.in

Abstract

The most common viral encephalitis in the world is caused by Japanese encephalitis virus (JEV), which spreads through mosquito bites. This fatal brain infection occurs at an average of 70000 cases annually. In spite of the efforts made to identify and select several targets which are essential in progression of JE, there is yet no licensed drug available against this deadly pathogen as it has a high mutation rate. Therefore, it is necessary to find a suitable drug against JEV, and this can be done by inhibiting one of its target molecules non-structural protein 5 (NS5), which will inhibit the replication of JEV, as the NS5 protein is a multi-enzymatic protein that plays a vital role in viral RNA replication. NS5, like other flavivirus non-structural and cellular proteins, has a methyl transferase domain in the N-terminus and an RNA-dependent RNA polymerase (RdRp) domain in the C-terminus. The NS5 methyltransferase domain is mostly responsible for viral genomic RNA 5′ capping, whereas the RdRp domain directly participates in RNA replication. Although the antiviral properties of hyoscyamine and scopolamine, the two active compounds from Atropa belladonna have been studied in JEV, the effect of its third active component, atropine, has not been investigated yet. In this study, we observed that atropine binds with NS5 protein of JEV with a binding energy of -7.00 (kcal/mol) which is similar with that of the positive control curcumin (-7.84 kcal/mol). Though in silico docking and simulation suggests that atropine could be a promising therapeutic option against JEV replication, further in vitro and in vivo experiments are to be carried out in the future course of time to validate the current results.

Keywords: Japanese encephalitis virus, NS5, natural compounds, molecular docking and simulation, antiviral

Computational Analysis of Structure and Function of Siglec1 and Prediction of Potential Pharmaceutical Agents for Rheumatoid Arthritis

Lakshmi K and Vino S,

Vellore Institute of Technology

svino@vit.ac.in

Abstract

Siglec1 (Sialoadhesin/CD169) is a novel sialic acid-binding Ig superfamily type 1 membrane protein, a molecule that binds glycoconjugate ligands on cell surfaces in an alpha (2, 3)-linked sialic acid-dependent manner. This protein is involved in mediating cell-cell interactions, is only expressed by a subpopulation of macrophages, and plays an important role in autoimmune disease and rheumatoid arthritis (RA). This study aims to find a potential natural inhibitor for the siglec1 biomarker. We can better understand the function of this class of proteins by using computational techniques to predict the structure and binding location of siglec1. We used the computerized method of protein modeling and small molecules collected from different sources, structure-based virtual screening, ADMET property predictions, molecular docking, and molecular dynamic simulation studies to screen potential siglec1 inhibitors. Rheumatoid arthritis has been effectively and safely treated using herbal therapies, which have low toxicity and fewer side effects. A total of 1600 natural compounds were collected from 150 anti-inflammatory and anti-RA medicinal plants from the literature and the supernatural database 3.0 was used for extracting natural compounds and was used for molecular docking, which evaluates the binding free energy and binding affinity of H-bond interaction score-based 10 hit list compounds. Re-docking with a 10-hit list and Type-A Procyanidins has the highest binding affinity (-8.5 Kcal/mol) to siglec1. The configuration stability of siglec1 and the ligand complex was predicted by 100 ns molecular dynamic simulation (MDs) and conformed complex stability. Our study's final outline was found to be a novel natural potential inhibitor against the target siglec1.

Keywords: Siglec1, Docking, ADMET, Modeling, Molecular dynamic simulation (MDs), Rheumatoid arthritis (RA)

Transcriptomic Analysis of Stress Response Pathway in HEK293 Cells During Recombinant AAV Production

Anju Ganesan and Balaji Balakrishnan

Department of Integrative Biology, SBST, Vellore Institute of Technology, Vellore, Tamil Nadu.

balaji.balakrishnan@vit.ac.in

Abstract

Background: Recombinant Adeno-associated viral (AAV) vectors have been successfully used in gene therapy to treat rare disorders such as spinal muscular dystrophy, hemophilia A and hemophilia B etc. However, optimizing the productivity and vector quality of rAAV is essential to meet the increasing clinical and commercial demand for gene therapies. Recombinant AAV vectors are packaged by standard triple transfection methods in producer cell lines. Here we hypothesize that during rAAV production, the producer cell lines are subjected to huge endoplasmic reticulated stress which ultimately impacts vector production qualitatively and quantitatively.

Aim: To evaluate cellular stress response during recombinant Adeno-associated viral vectors (rAAV) production in HEK293 cells.

Methods: Raw sequencing data (GSE224405) in the form of sequence read archives (SRA) files were retrieved using galaxy, FASTQ files were then subjected to a quality check using FastQC. The reads were then mapped to the reference human genome using Kallisto. Normalization and differential expression analysis was carried out in R using edgeR package and limma package. The interactions of DEGs and the related partners were curated from the STRING and visualised using Cytoscape. The DEGs subjected to Gene enrichment analysis. qPCR was performed to validate the differentially expressed genes.

Results and Conclusion: We analysed 145 significant DEGs (p-value < 0.05; fold 2 change ≥ 2 or ≤ -2). The Gene ontology analysis of DEGs were categorised for stress response pathways, which determined the involvement of Unfolded Protein Response (UPR) and DNA Damage Response (DDR) pathways. Interestingly, STRING analysis shows that DDIT3 (induces cell cycle arrest and apoptosis) followed by HIF1A (transcriptional regulator of the adaptive response to hypoxia) protein had direct interaction with XBP1 protein. Among the DEGs analysed 31 Unfolded Protein Response genes and 7 DNA Damage Response genes were differentially

expressed. UPR genes such as IRE1α, XBP1 and HSPA9 showed a significant expression. Whereas, UPR pathway could be a significant stress response pathway impeading for the viral vector production in HEK293 cells because during production host cell protein machinery is used and ongoing validation work showed the involvement of UPR gene (IRE1α) pathway. Therefore, HEK293 cells will be subjected with UPR which will be a greater impact for rAAV production, accordingly UPR pathway can be regulated in HEK293 cells for improved rAAV production.

Keywords: Differential gene expression analysis, DEGs, Unfolded Protein Response, rAAV production

Single Nuclei Transcriptomics Datasets of Brain Tissue Analysis by In-House Integrated Cross Species Platform

Abhishek Pandey[1], Aravind Easwar[1], Debasmita Paul[2], Rakshika Raveendran[3] and Manikandan Narayanan[1]

Computer Science and Engineering, Indian Institute of Technology Madras[1]

Sudha Gopalakrishnan Brain Centre, Indian Institute of Technology Madras[2]

Electrical Engineering, Indian Institute of Technology Madras[3]

nmanik@cse.iitm.ac.in

Abstract

Various computational methods have been developed over the last decade to handle the rapidly expanding single cell transcriptomics data. However, there is lack of unified platform for the biologists to use for analysis of the data sets. We, in our facility, have developed and are using an in-house constructed integrated pipeline (presented in INCOB23 conference) for analysis of cross species single nuclei datasets in our pursuit to understand the brain evolution of various animal groups. We have isolated nuclei from flash frozen brain tissue and using 10X genomics technology, we generated single nuclei cDNA libraries. It was sequenced with depth of 120 million reads and we obtained FastQ files as output. These files were put through the pipeline which is built based on Cell Ranger v7.1 software, Seurat 4.1.1, and Clustermole 1.1.0 packages in a unified interface that supports cell-type labelling for multiple different species. As output we obtained barcode rank plots, gene count matrix and operations like normalization, dimensionality reduction (PCA, tSNE, UMAP) and clustering and cell type labels using differentially expressed genes. We also obtained various quality control plots like violin, elbow and jackstraw plots. This study is preliminary in nature and the datasets are obtained in-house and then passed through the integrated platform to assess the functionality of the pipeline. These operations are performed and results are made available to the user for download. In future, we aim to produce more single nuclei datasets and create gene-gene homology maps and lineage trees between various species. Our interface acts as a bridge for a biologist to overcome the hurdles of prior computational expertise in using these tools individually. The code is available at: https://github.com/BIRDSgroup/mdn.

Keywords: integrated pipeline single, nuclei RNA sequencing data, Cell ranger, Seurat, clustermole, UMAP, tSNE

Investigation of Phe-tRNA interaction with EF-Tu in GDP/GTP Nucleotide bound states: A molecular dynamics simulation study

Kavya Kallahalli Mohan Kumar[1], Upendra Nagarajachari[2], Shuchika Devaraj Biligere[3] and Krishnaveni Sannathammegowda[4].

Department of Studies in Physics, University of Mysore[1],

Regional Institute of Education, Mysuru[2],

Department of Studies in Physics, University of Mysore[3],

Department of Studies in Physics, University of Mysore[4]

sk@physics.uni-mysore.ac.in

Abstract

Elongation factor Tu (EF-Tu) is an important class of translational GTPases (trGTPases) involved in the elongation process. EF-Tu functions by forming the tight complex with the aminoacyl-tRNA (aa-tRNA: transfer RNA carrying an amino acid) in GTP-bound active state and subsequently transports the aa-tRNA to the A-site (aminoacyl-site) of the ribosome. The correct interaction between mRNA's codon and aa-tRNA anticodon triggers GTP hydrolysis, resulting in the transition of EF-Tu from the active state to the inactive state (EF-Tu bound to GDP). This, in turn, results in the dissociation of EF-Tu:GDP from the tRNA and ribosome. However, the conformational changes and interactions responsible for the tight complex formation between EF-Tu and aa-tRNA in the GTP-bound state and the disruption of the interaction between EF-Tu and aa-tRNA in the GDP-bound state are not well understood. Therefore, to explore the conformational changes in the EF-Tu:Phe-tRNA complex and to get insight into the interaction between Phe-tRNA (tRNA carrying Phenylalanine) and EF-Tu in GDP/GTP nucleotide bound state, 200 ns MD (molecular dynamics) simulation have been carried out. The RMSD, RMSF, cluster, and DSSP analyses suggest that the GTP-bound state attains a more favorable conformation, which may facilitate EF-Tu's interactions with the tRNA and ribosome. Further investigation of the non-bonded interaction energy calculation revealed the significance of domainII in interaction with the Phe-tRNA in a GTP-bound state. In addition, the H-bonds calculated between the tRNA's Phe (Phenylalanine attached to the tRNA is considered) and domainII highlights the contribution of Val285 in recognizing tRNA's Phe by forming an H-bond throughout the simulation time in the

GTP bound state. Overall, these results provide insight into how GTP nucleotide influences the interaction between EF-Tu and Phe-tRNA.

Keywords: Molecular dynamics simulation, GTPases, EF-Tu, tRNA, Ribosome

Comprehensive Analysis of Amyotrophic Lateral Sclerosis Gene Expression Data to Ascertain Candidate Biomarkers

Deboral E and Mohana Priya A,
Vellore Institute of Technology, Vellore
mohanapriyaa@vit.ac.in

Abstract

Amyotrophic lateral sclerosis (ALS) is a complex neurodegenerative disease characterized by the progressive degeneration of motor neurons, leading to muscle weakness, paralysis, and respiratory failure. Understanding the molecular mechanisms underlying ALS pathogenesis is crucial for the development of effective therapeutic strategies. In this study, we conducted a comprehensive analysis to identify key genes involved in ALS pathology. Through differential gene expression analysis of the GEO dataset GSE3307, we identified a set of up-regulated and down-regulated genes in ALS. Among the up-regulated genes, CTNNB1, EP300, PIK3R1, EGFR, ESR1, RHOA, CDC42, MAPK14, and MDM2, along with the down-regulated genes UBB and UCC, emerged as important candidate genes implicated in ALS. These genes have been associated with various cellular processes and pathways relevant to ALS, such as Wnt signaling, transcriptional regulation, PI3K signaling, neuronal survival, inflammation, and protein homeostasis. To elucidate the interactions and functional relationships among these genes, we constructed a protein-protein interaction (PPI) network using the STRING database. The PPI network analysis revealed significant clustering and identified hub genes, including CDC42, MDM2, and RHOA, which play crucial roles in ALS pathology.

Keywords: amyotrophic lateral sclerosis, gene expression, protein-protein interaction network, hub genes, pathway enrichment analysis

Isolation, Characterization, and Genome Analysis of Lytic Bacteriophage Vb_SalP_1 against Food-borne Pathogen *Salmonella enterica*

Ramya Juliet, Oishi Mitra, Bershiyal S and Ramesh Nachimuthu

Vellore Institute of technology, Vellore, Tamil Nadu

ramesh.n@vit.ac.in

Abstract

Salmonella causes gastrointestinal infections and is widespread in developing countries like India due to poor and unhygienic environments. The development of fluoroquinolone resistance and multi-drug resistance in Salmonella has been a major concern. Having more than 2600 serovars in Salmonella species it is important to find a targeted solution for each serovar when antibiotic treatment fails. Phage therapy comes to the rescue as an alternative therapy. In this study, vB_SalP_1 phage was isolated from a sewage sample, and tested for its host range, morphological characteristics, stability in different temperatures and pH, life cycle, and phage-inhibition assay. Since Salmonella is a food-borne pathogen, its stability and lytic activity in food sources like egg was also determined. vB_SalP_1 had a broad host range killing up to 80% (n=10) of tested isolates, with high tolerance to temperatures and pH, showing maximum lytic activity in vitro in all MOIs from 100 to 0.0001 up to 6 h. 95% of the phages got adsorbed on its host in 3 minutes with a latent period of 5 minutes and a burst size of 50 phage particles per host cell. To determine the stability in the food source, the phage (1010) was incubated with egg white and egg yolk separately for 1 h and observed that the phages remained more stable in egg white than egg yolk and were able to inhibit bacterial growth in egg white effectively. TEM analysis revealed the morphology as a T7-like Podoviridae phage from Caudoviricetes order with a head size of 45±5.0 nm and a short tail. vB_SalP_1 consists of double-stranded DNA with 37,099 bp with a G+C content of 51.12%, with predicted 44 open reading frames by PHASTER and with no tRNAs found by ARAGORN. Blastn revealed the closest similarity with Salmonella phage ST38 (OQ860974.1; query coverage, 98%) and likely belongs to Autographiviridae; Studiervirinae; Kayfunavirus based on preliminary analysis. Therefore, the isolated phage can be effectively used as a promising agent to prevent and control Salmonella infections.

Keywords: Salmonella, food biocontrol, fluoroquinolone, resistance, bacteriophages vB_SalP_1, phage therapy, Salmonella genome

Exploring Salidroside and its Derivative Compounds as Potential DAPK1 Inhibitors: An In Silico Study for Reducing Alzheimer's Disease Risk

Tanish Prashar, Shreshtha Bhattacharya, Pritha Saha and Priti Talwar

Vellore Institute Technology

priti.t@vit.ac.in

Abstract

Salidroside, a naturally occurring phenylpropanoid glycoside found in Rhodiola rosea, has shown potential as a bioactive compound for enhancing cognitive function specifically memory, and learning by promoting neurogenesis and protecting neural cells against damage. It is also known to inhibit Death-Associated Protein Kinase 1 (DAPK-1), an enzyme critically implicated in regulating apoptosis and autophagy. It is a multifunctional serine/threonine kinase that is highly up regulated in Alzheimer's disease (AD) and is also associated with several pathological hallmarks of AD. In this investigation, we conducted a molecular docking study, which demonstrated the substantial affinity and selectivity of salidroside towards DAPK-1. Subsequently, we identified synthetic analogs of salidroside using PubChem databases, which were subjected to docking studies to ascertain their binding affinities and ligand-target interactions. Three salidroside analogs, CID-101530110, CID-101530061, and CID-101530061, were selected for further evaluation. These analogs were subjected to comprehensive assessment, including Absorption, Distribution, Metabolism, and Excretion (ADME) profiling using Swiss ADME and toxicity analysis using the ProTox II server. The top hits were further analyzed using molecular dynamics studies to elucidate their binding patterns. The findings of this study provide crucial insights into the underlying molecular mechanisms and neuroprotective effects of salidroside and its analogs. This knowledge holds promise for the development of DAPK-1 inhibitors as a potential therapeutic approach for Alzheimer's disease.

Keywords: Salidroside, Alzheimer, DAPK1

In-Silico Identification of the Potent Carotenoids Targeting the Kinase Receptors for Cancer Therapy

Ressin Varghese, Kuntal Pal, Ashwni Kumar Cherukuri and Siva Ramamoorthy

Vellore Institute of Technology

siva.ramamoorthy@gmail.com

Abstract

Rationale: Cancer is considered the most malignant disorder with a predicted global burden of about 30 million new cases by 2040, with the greatest increases happening in low- and middle-income countries. Amidst of the sophisticated therapeutic strategies, the side effects caused by the colossal use of cancer chemotherapeutics is a major concern. A rapid upsurge was observed in the research of natural therapeutics for alleviating cancer during the past decades. In this scenario, carotenoids are one of the most potent secondary metabolites derived from various natural resources explored for targeted cancer therapy. We chose tryosine kinase receptors as our targeted molecules owing to their key roles in signaling cascade regulating the growth, differentiation, metabolism of cells, and apoptosis in response to various stimuli.

Materials and methods: A dataset was developed comprising of the FDA-approved standard drugs and anticancer potential carotenoids categorized as carotenoids, apocarotenoids, xanthophylls, and synthetic carotenoids through literature mining. A combined workflow of virtual screening, molecular docking, and a supervised machine learning algorithm was applied to identify the most potent carotenoid for kinase receptor-targeted cancer therapy. A predictive model was developed employing ML to analyze the chemical feature similarity of selected carotenoids to clinically approved positive cancer drugs. The best candidate identified for each receptor was subjected to molecular dynamics simulation to assess the interaction stability.

Results: About fourty-two potential carotenoids were identified through literature mining and were further screened by analyzing ADMET and toxicity. The potential carotenoids targeting receptors HER2, HER3, EGFR, FGFR2, PDGFR, VEGFR2, ALK, RON kinase, MET compared to FDA-approved drugs were further recognized through molecular docking and machine learning approach. Four potential carotenoids including beta-carotene, canthaxanthin, neoxanthin, peridinin, and fenretinide were identified through the combined molecular docking and machine learning approach.

Conclusion: A collective application of the carotenoid mixture for curing cancer will be a ground-breaking and highly effective approach for treating cancer.

Keywords: Carotenoids, Cancer, Kinase receptors, Machine learning, Molecular docking

Cigarette Butts: Tiny Nocuous particles inciting ecotoxicity via oxidative stress in Pila virens

Koigoora Srikanth, Siksha Sharavani and Burri Ganesh

Dept of Biotechnology, VFSTR, Vadlamudi, Guntur District, Andhra Pradesh, India

koigooras@gmail.com

Abstract

Cigarette butts are exceedingly toxic substances in the aquatic environment. Snail's are the major food source. In the aquatic (marine region) and also exposed to the chemical toxicity of the cigarette butts. In order they can straightly affect the human being through food and contamination take place in aquatic environment. In this novel study we evaluate the toxicity of cigarette butts on gastropods and oxidative stress response. The cigarette butts solutions acquire by adding the five butts in 1liter and soked it for 2 hrs and exposed the snail in different concentrations (15%. ,25%,50%) control after exposing the snail for 48 hrs the. Mortality rate were recorded. More over the 100% of organisms /snails were vanished in 50% concentrations in rest of solutions the survival rate is longer compared to each different solution. Cigarette butts accumulation in snail causes oxidative stress due to production ROS oxidative strees by Cigarette butts can damage the digestive gland in snail. In snail and break down of dna and cause mutations consequently the motive of this novel study was to inspect numerous amounts of toxicity effects of Cigarette butts exposure including with oxidative stress

Keywords: Pila virens, Cigarette butts, Antioxidants, Oxidative stress

Molecular Docking Study of Anti-Arthritic and Anti-Oxidant Phytochemicals Identified from Indian Medicinal Plants Against RANKL In the Treatment of Bone Degradation in Rheumatoid Arthritis

Devi Soorya Narayana S and Vino S,

Vellore Institute of Technology

svino@vit.ac.in

Abstract

Background: Rheumatoid arthritis (RA) is a chronic inflammatory disease causing the destruction of joints. Recent studies have found that osteoclasts play a significant role in bone degradation in the disease. The tumor necrosis factor superfamily protein, osteoclast differentiation factor receptor activator of NF-B ligand (RANKL) is important in osteoclast differentiation and bone degradation in RA. RANKL activates and differentiates pre-osteoclasts and mature osteoclasts by binding to their RANK receptors and the degree of growth and activity of the osteoclasts is determined. Bone erosions in RA are caused by osteoclastic bone resorption in synovitis sites, where RANKL expression is also noted. Furthermore, MRI bone edema in RA suggests the existence of active inflammation inside bone as well as the presence of osteitis, which is likewise related to the RANKL expression. Radiographic studies show that bone destruction in RA occurs early and continues throughout the disease's duration. Bone erosion causes deformities of the afflicted joints and inhibits patients' normal activities. Therefore, one of the most difficult objectives in the therapy of RA is to prevent bone destruction.

Methodology: 3829 ligands were collected from literature, and IMPPAT (Indian Medicinal Plants, Phytochemistry And Therapeutics) databases, which are phytocompounds with antioxidative properties, and anti-arthritic properties. The structures were docked into the X-ray crystal structure of the human RANKL-OPG complex (PDB ID: 3URF, R = 2.7 Å) after removing the OPG to free the binding site, and chain A of sRANKL (162aa) was taken for further docking studies. The CASTp server was utilized to identify the active site and the site-specific docking studies were performed with AutoDock vina. The compounds were screened virtually by molecular docking and further filtered using ADME properties. The CHARMM36 was used to generate the ligand and protein coordinates of the best conformation. Concurrently, the protein complex was created

using the Gromacs package. Protein-ligand complex trajectories were compared to confirm the stability of the protein trajectory. The MD simulation for 200 ns was performed in a cubic water box with TIP3 water molecules. The stability of the protein and protein-ligand complexes was confirmed by comparing their RMSD, RMSF, hydrogen bond, Radius of Gyration. Additionally, to evaluate the strength of the intermolecular interaction we calculate the binding free energies of the complexes using the molecular mechanics Poisson Boltzmann surface area (MM-PBSA) method. We performed MM-PBSA calculations through the gmx_mmpbsa python tool then the results were analyzed using gmx_mmpbsa_ana.

Result and Discussion: 3829 ligands with inhibitory properties of RANKL, antioxidative, and anti-arthritic properties were collected from the literature survey and IMPPAT database. The AutoDock vina docking results revealed strong interaction between the phytocompounds and RANKL and the best-docked compounds were 4-(P-Methoxyphenyl)-2-(4-phenyl-2-pyridyl)-6-(2-pyridyl)-pyridine (-9.8 kcal/mol), Bismahanine (-9.2 kcal/mol), and Tingenone (-9.1 kcal/mol). The MDS simulation study confirmed the stability of the complex.

Conclusion: The main goal of RA treatment is to avoid bone and joint deterioration and to keep patients active on a daily basis with no treatment-associated side effects. Recent research has shown that osteoclasts play a role in the etiology of bone and joint degradation in RA and can be a powerful treatment target for the illness. Therapeutics that target osteoclast production or function can, at the very least, slow the advancement of these bone alterations. Thus, the phytocompounds retrieved in our study may act as a promising therapeutic option for RA with great efficacy.

Keywords: Rheumatoid Arthritis, Bone degeneration, Phytochemicals

Bioaugmentation Of Phenanthrene- Polycyclic Aromatic Hydrocarbon Using Sphingomonas Species Isolated from The Petrol Bunk Soil

Prashanth Rajan A[1], Biswanath Mahanty[2]

Department of Biotechnology, School of Bio Sciences and Technology Vellore Institute of Technology, Vellore, Tamil Nadu, India[1]

Department of Biotechnology Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore, Tamil Nadu, India[2]

prashanth07052001@gmail.com

Abstract

Anthropogenic activities based on fossil fuels, coal gasification, coking industry, waste incineration, oil spills, and other industrial processes have burdened the environment with recalcitrant pollutant Polycyclic aromatic hydrocarbons (PAHs). Present worldwide abundant distribution and long-term environmental persistence of PAH has become a global concern. PAH are highly ecotoxic agents, also major mutagens and carcinogens in the present era, which persuaded the United States Environmental Protection Agency (USEPA), to enlist phenanthrene in the priority chemicals. Phenanthrene has fused rings and lacks a terminal surface for enzyme activation making phenanthrene resistant to biodegradations resulting in accumulation through the food chain. Phenanthrene reduces the growth and reproductive potential of aquatic life forms. It causes respiratory problems, skin irritation, and cancer when seafood contaminated with phenanthrene is consumed. The physical and chemical methods are highly expensive and ineffective in removing PAHs in bioaugmentation. Based on these challenges the present study was designed to isolate soil microbes from the contaminated sites. Soil samples from the Petrol Bunks were collected to isolate Polycyclic Aromatic hydrocarbondegrading bacteria. The bacterial isolates namely Isolate- M1S1, Isolate-M1S2, Isolate-M1S3, Isolate-M1S4, Isolate-M2S1, Isolate-M2S2, Isolate-M2S3, Isolate-M2S4 were found to tolerate and degrade the PAHs. The isolates showed good growth in phenanthrene-rich solution. Isolate-M1S3 had the highest efficiency in phenanthrene degradation. This chemoheterotrophic yellowish bacterial colony was rod-shaped, Gram-negative, aerobic, nonmotile, and non-spore-forming, which resembled the characteristics of Sphingomonas species. This microbe has great potential to degrade the phenanthrene to

nontoxic or low-toxic products. Bioaugmentation using this bacterium is economical and safer compared to other contemporary existing technologies

Keywords: Polycyclic Aromatic Hydrocarbon, Phenanthrene, Bioaugmentation, Sphingomonas species

Nicotine Deter Women's Reproductive Health: A Molecular Docking and Dynamic Simulation Studies

Saranya G M, Gomathi. S, Poornimaa Murali, Ramanathan Karuppasamy, Rameshpathy Manian,

Vellore Institute of Technology

mrameshpathy@vit.ac.in

Abstract

Over decades there is not much change in infertility and it still remains an on-going reproductive problem. Lifestyle factor like smoking have been observed to be higher across the globe and found to have a negative impact on fertility. Fertility in women is overseen by various hormones secreted by different receptors found in the HPG (hypothalamus, pituitary gland) axis. Either inactivation or activation of binding domains of these receptors by the ligand nicotine (NIC) will affect the receptor stability or the signal transduction machinery. Molecular docking simulation was achieved to NIC with receptors like LEPR, KISS1R, GnRHR, FSHR, LHR and ESTR by the computational method. Almost in all the receptors interaction of NIC ligands shows the lowest binding energy and hence confirms the effect of ligands in increasing the stability of the complex. Among the selected six receptors, FSHR docked with NIC was found to be more endangered. It shows the lowest binding energy (-7.21 kcal/mol), lowest inhibitor constant (0.005 μM), with hydrogen bond interaction (ILE 111) and hence said to be the more stable complex with good binding affinity. At the end, the conformational stability of the receptor-ligand complex in the binding sites was examined using molecular dynamic (MD) simulation study for 50 ns simulation. The results from MD simulation study highlights the stable binding of NIC with target receptor. Thus, our in-silico results provide insights in the roles of these ligands, their possible binding mode, their target contacts, efficiency and their specific interactions with the binding sites and bridges the gap between theory and research in ways that were earlier impossible.

Keywords: FSHR, HPG axis, Infertility, Nicotine

Decoding the Conformational Impact of PTEN Mutant R130Q on Binding to the PIP2 Enriched Cell Membrane in Breast Cancer: A Computational Approach

Rohit Karn, Prarthana Chatterjee, Arnold Emerson,

Vellore Institute of Technology

iarnoldemerson@vit.ac.in

Abstract

The phosphatase and tensin homolog (PTEN) is a tumor suppressor gene which encodes a dual-specificity phosphatase. It serves as the key regulator of PI3K/AKT/mTOR signalling by dephosphorylating the lipid phosphatidylinositol (3,4,5)-triphosphate (PIP3) to phosphatidylinositol (4,5)-bisphosphate (PIP2). It is one of the most frequently mutating genes In breast cancer. Here, we investigate the interaction of breast cancer PTEN mutant protein with the cell membrane, in the presence of PIP2 and PIP3 molecules, using molecular dynamics simulations. We also studied the role of C2 domain and C-terminal tail (CTT) in membrane association. Two MD simulations (wild-type and mut_R130Q) of 500ns were performed using Charmm-GUI and GROMACS. The full-length PTEN was modelled and subjected to mutation during system preparation. Our study provides a mechanistic understanding of the interactions and structural consequences of PTEN C-terminal tail and a comparison of mutant structure with wild-type which broadens our insights into PIP2 and membrane interaction of PTEN. The Wild-type PTEN protein interacts with the membrane in a better way than compared to the mutant which possibly explains its tumor suppressor activity. This study provides insight into the function of PTEN and its mutants towards disease pathogenesis.

Keywords: Breast cancer, PTEN, PIP2, Molecular dynamics simulation

In Silico Analysis of Metabolites Produced by Lactic Acid Bacterial (LAB) Cultures Against Aquatic Pathogens

Manishkumar[1], Sudhakaran Raja[*], Aquaculture Biotechnology Lab, Pearl Research Park, School of Biosciences and Technology, Vellore Institute of Technology

*2sudha@gmail.com , sudhakaran.r@vit.ac.in **

Abstract

Rationale: Aquatic pathogens pose a significant threat to aquaculture, causing substantial economic losses and jeopardizing the sustainability of this important food production sector. These pathogens can cause a range of diseases in farmed aquatic organisms, leading to mass mortalities, reduced growth rates, and compromised product quality. *Vibrio parahemolyticus* and *Vibrio harveyi* are Gram-negative bacteria, which is a major concern for aquaculture, specifically shrimp farming, due to their ability to cause luminous vibriosis, a disease that affects a wide range of marine animals including fish, crustaceans, and mollusks causing substantial economic losses and posing a risk to global food security. White spot syndrome virus (WSSV), a double-stranded DNA virus, is a devastating threat to the global shrimp aquaculture industry, causing catastrophic economic losses which include 100% mortality within 3 to 7 days. In response to these threats, researchers are exploring various strategies to control aquatic pathogens and protect aquaculture production. These strategies include the use of probiotics, such as lactic acid bacteria (LAB), which have been shown to have antimicrobial properties against aquatic pathogens.

Objectives: This study discusses the *In-silico* approach using metabolites produced by LAB cultures against *V. Parahemolyticus, V. Harveyi,* and WSSV.

Materials and Methods: (1) Metabolites were extracted from LAB isolates, **(2)** GCMS analysis was performed to confirm the metabolites present, **(3)** All the metabolites were docked against the aquatic pathogens (*Vibrio parahemolyticus, Vibrio harveyi*, and White spot syndrome virus).

Results and Discussion: Results suggested that there are 11 metabolites produced by these LAB isolates and Tris(2,4-di-tert-butylphenyl) phosphate produced by 2 out of 3 LAB cultures which has better binding affinity compared to all other metabolites produced by all three LAB cultures. This compound is also considered a common antioxidant in previous literature. This can help in the future for a better understanding and proceeding to the animal study.

Keywords: Aquaculture, Aquatic pathogens, LAB cultures, Metabolites, and Molecular docking.

BIXS Talks

Nipah Henipavirus Proteins Database

Basanth Geereddy[1], Hiranyada Chintakrinda [1,3], Bhavya S Suravajhala [1,2]

[1]. Bioclues.org, India.

[2]. Grade XI, Amrita Vidyalayam, Karunagappally, Kollam, Kerala. 690544

[3]. Grade XI, Kendriya Vidyalaya, Gachibowli, Telangana, India 500032.

sirinishu116@gmail.com[2], bhavyas02007@gmail.com[3]

Abstract

Nipah virus (NiV), an emerging pathogen with a high fatality rate, poses a significant threat to humans and livestock. We created the Nipah Henipavirus Proteins Database to address the need for comprehensive information on this virus. Animals (such as bats or pigs) or contaminated foods can transmit the Nipah virus to humans. Neither humans nor animals have been treated or vaccinated against this virus. Researchers have developed experimental monoclonal antibodies to treat Nipah virus disease under compassionate use. Supportive care is the primary treatment for humans. Intensive supportive care is recommended for patients with severe respiratory and neurologic complications. The RNA genome of Nipah henipavirus contains 9 proteins comprising 6 structural and 3 non-structural proteins. Studying and analyzing different virus proteins will help us understand their mechanism and how they interact with our cells, which will further help us develop antiviral drugs and practical tools to fight the virus. The Nipah henipavirus, which is the causative organism of the Nipah virus, is a phosphoprotein belonging to the paramyxoviridae family, which includes species that cause several other chronic diseases like measles, mumps, and other respiratory tract infections. Nipah henipavirus is a single-stranded RNA virus made of six consecutively arranged genes. The six polynucleotides/polypeptides found in a single strain of this virus are Nucleocapsid, Phosphoprotein, Matrix, Fusion glycoprotein, Attachment glycoprotein and long polymerase. The nucleocapsid, phosphoprotein, and polymerase are responsible for forming the ribonucleoprotein of the virus. The fusion glycoprotein and attachment glycoprotein are attached to the virion and play a vital role in the entry into the host cell. When the virus enters the host's body, the proteases present in the host body's immune system attack the fusion glycoprotein and cleave it into two subunits called F1 and F2. The F1 subunit contains a

fusion peptide that binds to the host's cellular membrane and causes entry into the cell. The matrix protein then initiates morphogenesis and budding. This causes virus replication in the host body, and slowly, the virus takes control of the host's immune system. Later on, due to the interactions between the viral receptors of the host cell and the glycoprotein, some conformational changes are triggered in the virus, activating the fusion glycoprotein that fuses to the cell membranes. The pathogenicity of this virus is directly related to the intensity with which the virus fuses with the receptors.

Keywords: Nipah virus, Bioinformatics, Database, Proteomics

SCVDB: A Database of SARS-CoV-2 Annotated Proteins

Hiranyada Chintakrinda[1,3], Bhavya Sri Suravajhala[2,3]

1. Grade XI, Kendriya Vidyalaya, Gachibowli, Telangana, India 500032.

2. Grade XI, Amrita Vidyalayam, Karunagappally, Kollam, Kerala. 690544

3. Bioclues.org, India.

sirinishu116@gmail.com[1], bhavyas02007@gmail.com[2]

Abstract

SCVDB stands for SARS-CoronaVirus Database and Analytics (https://bioclues.org/SCVDB/index.php). SCVDB is a unique dynamic database that curates the records that have been sourced primarily from UniProtKB and then manually curated. For any internal conflicting or inconsistent annotation, SCVDB is aimed to be an integrated platform to develop a database and analysis/tools. SCVBD is a bioclues project created during the lockdown with the aim of inculcating and including school children in the subject area of bioinformatics. Their enthusiasm warranted us to create a helpful resource Database with Sequences (Cov: 15 and Cov2: 17) as a One-Point Portal (dated 07-Nov-2023). The importance of developing an integrated platform with Sequences, Citations, Publications, and Clinical data warranted us in developing this database. SCVDB contains 32 proteins; SARS-CoV-2 Data, NCBI got 5,113,401 SRA runs; 6,004,532 Nucleotide records; 8,097 ClinicalTrials.gov; 281,614 PubMed; and 376,649 PMC presently. The SCVDB database mainly contains Catalogues that are lists of items. An integrated database catalogue is a function embedded within a server that allows administrators to view information for every database installed there. Additionally, the database catalogue stores metadata on every database, like the number of rows and tables each database has. The catalogues in the SCVDB database are: -

● Taxonomy: This catalogue spans through a set of Taxonomy Classification index

● Advance search: This catalogue allows users to search the SCVDB database through any set of combinations of one or more fields with values.

● Keywords: This catalogue allows an end user to pick a set(s) of Keywords from the Index.

● Species: This catalogue enables the user to pick a set(s) of Species index

● Families: This catalogue allows users to choose a set(s) of Family index

● Citation: This catalogue enables users to search the database through a Citation index [By Title, Author, Journal, Year] in any combination(s).

Keywords: SARS-CoV-2, Acute respiratory syndrome, COVID-19, Lung infections, Database

Video Abstracts

In Silico Analyses of Protein-Ligand Interactions Associated with Stress Pathways in Plants

Praneet Prabhanjan B[1], Nandni Kumari[1] , Abilashni Arthiswaran[1] , Dalia Vishnudasan[1] , PB Kavi Kishor[2] , Ramesh Katam[3] and Renuka Suravajhala[1]

Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana, Kerala, India[1]

Department of Genetics, Osmania University, Hyderabad, India[2]

Florida Agricultural and Mechanical University, Gainsville, FL, USA[3]

ramesh.katam@famu.edu and renus@am.amrita.edu

Abstract

A vital biological mechanism that keeps life on earth alive is photosynthesis. Green plants, algae, and some bacteria use this process to transform light energy into chemical energy. Proteins D1 and D2 are essential elements of the Photosystem II (PSII) complex, a crucial part of the thylakoid membrane found in chloroplasts. PSII is in charge of photosynthesis's first stage, which involves absorbing light energy and utilizing it to start the electron transport process. In this work, we performed a comprehensive investigation into the genetic variations within *Taxus brevifolia*, *Manihot esculenta*, and *Azadirachta indica*. Our study focused on predicting gene functions, exploring protein-protein interactions, and analyzing closeness, betweenness, and clustering coefficients specifically related to chloroplast genes. Using a combination of various software tools and biological databases. We identified significant factors associated with Photosynthesis II (PSII) and explored the impact of herbicides and their analogs on PSII. By applying ADME-based Lipinski rule properties screening with KNIME nodes and conducting molecular docking experiments, we identified potential herbicide lead molecules that we firmly believe would help enhance crop yield and mitigate weed growth, further contributing to advancements in agriculture and sustainable food production.

Unraveling the Genetic Basis of Prostate Cancer Phenotype in the Indian Population

Bhargavi R[1], Devendra Sharma[2], Maneesh Vijayvargiya[3], Rajaguru Aradhya[1], Bipin G Nair[1] and Prashanth Suravajhala[1]

Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana, Kerala, India[1],

Department of Urology, CK Birla Hospitals, Jaipur, Rajasthan, India[2],

Department of Pathology, Mahatma Gandhi University of Medical Sciences and Technology, Jaipur, Rajasthan, India[3]

prash@am.amrita.edu

Abstract

This study aims to explore the genetic variations associated with prostate cancer (PCa) in the Indian population through Whole Exome Sequencing (WES) Methods: 39 Malignant PCa FFPE samples with Gleason grade between 7-9 were collected from CK/Rukmani Birla hospital, Jaipur with institutional ethics committee (IEC) clearance and informed consent from all participants was duly obtained prior to the commencement of the study as part of our Cancer Prostate consortium of India (CAPCI) efforts. WES was performed on an Novaseq 6000 Platform with the raw reads then run through our in-house benchmarked variant calling pipeline, CONsensus Variant EXome (CONVEX) which employs four different variant callers, viz. VarScan, bcftools call, vt and Freebayes to obtain consensus variants of significance, after which gene annotation was done using ANNOVAR/SnpEff. The list of common variants among all samples was then compared to the missense deleterious genes obtained after benign subtraction earlier reported by Ravindran et al., 2023 to obtain a list of common genes representative of both north and south indian population. Furthermore, the list was compared to that of the OncoArray chip project from Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium which genotyped 140,000 PCa patients of European ancestry for 600,000 SNPs associated to PCa and other cancers. Results: Our analysis revealed a spectrum of genetic alterations with 2561 variants found to be consensus to all malignant samples of which 1669 were identified as exonic variants, 724 intronic and 28 ncRNA variants. Among the exonic variants, 754 were identified as non-synonymous using ANNOVAR/SnpEff. Key cancer-related genes that were consensus in our samples include CYP11B2, COL6A1, MYO15A, HEXB whose role in tumour

invasiveness, angiogenesis and metastasis has been extensively studied. The comparison of our list of consensus variants to the missense deleterious genes' list reported by Ravindran et al., 2023 revealed six common genes, viz. ERV3-1, GPRIN2, MUC16, PHC1, UMODL1 and ERV3-1-ZNF117 readthrough. Notable among these are MUC16 which is found to be overexpressed in various cancers and promotes migration and invasiveness of cancer cells and ERV3-1, an endogenous retroviral protein which could possibly serve as a prognostic marker. The comparison of consensus variants from our data to OncoArray data from PRACTICAL consortium revealed 118 SNPs common between the datasets representing common SNPs between Indian and European populations. Conclusions: The WES performed on PCa samples in our study revealed a diverse genetic landscape of the Indian PCa phenotype, emphasizing the need for population-specific studies to better understand the disease. The genetic variations identified in this study may open new avenues for further research into their functional significance and potential as novel diagnostic and prognostic markers.

Keywords: Prostate cancer, systems genomics, next generation sequencing, exome sequencing, variants.

Comparative Analysis of RNA-Seq Results in Parkinson's Disease: Unraveling Molecular Insights into Disease Pathogenesis

Anu Sasidharan and Ekanath R

Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana, Kerala, India

anusasidharanmay112000@gmail.com

Abstract

Parkinson’s disease (PD) is a neurodegenerative disorder that affects the motor system due to the loss of dopaminergic neurons in the basal ganglia, a structure in the midbrain.This is an age related diseases which has no complete cure. One of the reasons is that the disease mechanisms are not fully understood, especially from a subcellular level. This study focuses on an in-depth analysis of RNA-Seq data to compare the gene expression profiles between Parkinson's disease-affected and normal conditions. We explore the vast RNA-Seq datasets, examining the gene expression profiles in affected brain regions, and elucidate the dysregulated pathways and biological processes associated with PD.This study contributes to a deeper comprehension of the molecular underpinnings of Parkinson's disease and may facilitate the development of innovative treatments and diagnostic tools for this devastating condition. By leveraging advanced bioinformatics tools we aim to identify differentially expressed genes that could pinpoint potential biomarkers and therapeutic targets that may aid in the development of innovative diagnostic tools and treatment strategies. Additionally, we probe into the regulatory networks and noncoding RNAs that may play pivotal roles in the pathogenesis of PD.

Repurposing Targeted Therapeutics Against Monkeypox: In Silico Modeling for the Upcoming Battle

Anguraj Moulishankar, Indrani Biswas, Rajul Jain, Kavyasudha Kanuparthi, Arsha Gopan, Sivani Dinesh, Sneha Saji, Sneha Baiju, Sravyasree Gubbala, Renuka Suravajhala,*

Department of Pharmaceutical Chemistry, College of Pharmacy, Madras Medical College, Chennai, India

Mahatma Gandhi Medical Advanced Research Institute (MGMARI), Sri Balaji Vidyapeeth, Pondicherry,

India Bioclues.org, Hyderabad, India

Amrita School of Biotechnology, Kerala, India

renus@am.amrita.edu

Abstract

Background: The Monkeypox virus is considered the next pandemic by the World Health Organization. The evolutionary history of Monkeypox has similarities to COVID-19. There is no correlation between Monkeypox and SARS-COVID, but there is evidence of smallpox disappearing and Monkeypox re-infections emerging. This suggests a common phylogenetic origin. The study aims to screen antiviral drugs that target cellular and molecular factors important for Monkeypox virus survival and maintenance.

Materials and Methods: Homology modeling was used to predict the structure of two proteins, cell surface envelope protein (CSEVP) and DNA-dependent RNA polymerase (DpRNAp). The predicted structure of the target protein was then used for molecular docking studies with antiviral compounds. Additionally, the stability of potential lead complexes was simulated using the GROMACS software collection on the two targets - CSEVP and DpRNAp - in complex with the potential lead molecules - Rilpivirine and Tibo.

Results and Discussion: 70 compounds were tested against two targets, CSEVP and DpRNAp. Idoruxidine had the highest binding affinity for both targets and exhibited significant anti-MPV activity. Molecular dynamic simulation was also performed to assess the stability, conformational changes, and binding interactions of the ligands with the protein.

Matrix Metalloproteinases: Master Regulators of Tissue Morphogenesis

Sreesada P, Vandana*, Bhagath Krishnan*, Amrutha R*, Yash Chavan*, Anjali*, Tanha*, Parvathy Venugopal, Rajaguru Aradhya#, Bipin G Nair# **

Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Clappana PO 690525, Kerala, India

rajagurua@am.amrita.edu and bipin@amrita.edu

Abstract

Matrix metalloproteinases (MMPs) belong to a class of zinc proteases that facilitate the degradation of various components of the extracellular matrix (ECM). Beyond ECM degradation, MMPs also influence processes such as inflammation, cell development, proliferation, and more. In vivo genetic studies of Drosophila MMPs, specifically Mmp1 and Mmp2, have revealed their pivotal role in tissue remodelling while not being critical for embryonic development (Page-McCaw et al., 2003; Wen et al., 2020a). Both Drosophila MMPs exhibit the canonical and conserved MMP domain organization (Lafever et al., 2017; Llano et al., 2002). A notable distinction between Mmp1 and Mmp2 lies in their cellular localization; Mmp2 appears to be membrane-anchored, while Mmp1 is released into the extracellular environment. This localization difference underscores the significance of their roles within this small MMP family. MMPs typically feature various domains, including the signal sequence, propeptide, catalytic domain, and hemopexin-like domain. The signal sequence, also known as the pre-domain, guides MMP production and directs it from the endoplasmic reticulum to the extracellular space. MMPs can be broadly categorized into secretory MMPs and membrane-type MMPs (MTMMPs), with specific subgroups characterized by their structure and function. The intricate functional diversity of human MMPs poses challenges in analysing their intracellular activities, as the human genome encodes approximately 23 distinct MMPs with overlapping functions. In contrast, the Drosophila melanogaster genome encodes only two MMPs, dMMP1 and dMMP2 (Lafever et al., 2017).

Unraveling the Clotting Cascade in Vitamin K Deficiency Associated Comorbidities

Shalini Rajagopal, Jalaja Naravula and Pb Kavi Kishor

shalini.irajagopal@gmail.com

Abstract

The coagulation mechanism can be affected by vitamin K (VK) deficiency, a severe nutritional imbalance that is associated with a number of comorbidities. In this study, we incorporate data from the Gene Expression Omnibus (GEO) archive to provide a study into the unraveling of the clotting cascade in the context of VK deficiency. Maintaining hemostasis is crucial, and VK is an essential cofactor for the synthesis of coagulation factors. A lack of VK has been associated with a higher risk of bleeding problems and other comorbidities. In the present investigation, we utilized publicly accessible GEO datasets to examine the transcriptional alterations in genes linked with the clotting cascade when deficiency of VK and its associated conditions are involved. The results of our study suggested that there were significant variations in the expression of genes associated with fibrinolysis, vascular homeostasis, and coagulation in response to VK deficiency. These alterations were significantly connected to the risk of thrombotic and hemorrhagic episodes, as well as various diseases that have been associated with VK deficiency-associated disease phenotypes, including diabetes, placenta, prostate, hepatic, renal, and cardiovascular disorders.

Inferring Variants of Uncertain Significance (VUS) in Rare Disease Genetics: An India-Centric Study

Anjali Krishna A, Praveen Mathur, Krishna Mohan Medicherla, Renuka Suravajhala and Bipin G Nair

anjalikrishnasree04@gmail.com

Abstract

Background/Research gaps: Genetic variation was believed to be associated with exonic regions only. While much emphasis has been placed on coding regions of the genome, non-coding RNAs (ncRNAs) have emerged as essential players in gene regulation and disease mechanisms. Furthermore, genetic variation is beginning to be understood in ncRNAs. However, not much studies have taken place on genetic variation attributing to ncRNAs in rare diseases. There are approximately 6000 rare diseases in the world and over 1000 are reported in India. Many missense and nonsense mutations are known, and variants of uncertain/unknown significance (VoUS) are emerging to be understood for causing the rare diseases.

Objectives: In this work, we exploit meta-analyses and aim to identify genetic variation attributing to rare diseases of the Asia/world and India. We also contemplate identifying variants specific to Congenital Pouch Colon (CPC), an Anorectal malformation (ARM) that was largely studied and further identify variants specific to syndromic/nonsyndromic ARM from samples sequenced and analyzed in our lab.

Materials and Methods: The genetic diversity inherent in human populations plays a pivotal role in health and disease susceptibility. This study embarks on a comprehensive exploration of genetic variants within three overarching population categories: Indian, Asian, and worldwide. The initial step involved data acquisition from these databases, ensuring that the selected datasets provide an accurate representation of genetic variation within each population category. The data were collected from well-established genetic databases, ensuring inclusion of population-specific information by employing Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Filters were applied to segregate variants present in the Indian, Asian, and worldwide populations.To this end, we harness the data from two prominent sources, the "indiGenomic VCF(Indian Genomic DB)" and "NCBI Rare Disease" databases, to illuminate the genetic tapestry that unites individuals across diverse backgrounds. Functional impact assessment was then employed to variants with potential biological significance and subsequently variants were analyzed based on their minor allele frequency (MAF), with thresholds set to capture common, rare, or population-specific genetic diversity. Population-specific MAF data were utilized whenever available, offering insights into genetic variation tailored to each population.
Results and Discussion: We delved upon deciphering the candidate genes and their variants from the Asian/world and Indian sub-population and mapped the non-coding variants precisely. The common variants between Indigenomes and our annotation were searched and we found that there are 155 common variants with 15 of them attributing to ncRNA variants, and 3 of those to be pathogenic in Asian/world datasets. Whereas, 64 were commonly found variants with 5 of them attributing to ncRNA and 2 of those classified as pathogenic. We also sought to check the common variants or similar alleles attributing to rare diseases from our cohort. The results of this study present a rich and diverse landscape of genetic variants within these three population categories. The variants were visualized and interpreted with an emphasis on functional impact and allele frequency distribution.

Conclusions: In conclusion, we explored common genetic variants across Indian, Asian/ worldwide populations underscoring the need for finding candidate variants attributing to pathogenesis and ncRNAs. We believe an investigation into shared genetic diversity offers an insightful glimpse into the human genetic commonality that transcends geographical boundaries and ancestry.

Inferring Recombination Events in SARS-CoV-2 Variants In Silico

Nihal Najeeb, Aparna B. Murukan, Anagha Renjitha, Malavika Jayaram, Ayisha A. Jabbar, Haripriya Haridasan, Akshara Prijikumar, Sneha Baiju, Adrial Ann Nixon, Ponnambil Anantha Krishnan, Sunu Rodriguez, Somesh Kumar, Sunil K. Polipalli, Keshav K. Singh, Bipin G. Nair, Sudeep D. Ghate, R. Shyama Prasad Rao, Polavarapu Bilhan Kavi Kishor, Arya Aloor, Renuka Suravajhala, Gyaneshwer Chaubey, and Prashanth Suravajhala

aparnamurukan448@gmail.com

Abstract

Over the last 34 months, at least 10 severe acute respiratory syndrome coronavirus 2 (SARSCoV-2) distinct variants have evolved. Among these, some were more infectious while others were not. These variants may serve as candidates for identification of the signature sequences linked to infectivity and viral transgressions. Based on our previous hijacking and transgression hypothesis, we aimed to investigate whether SARS-CoV-2 sequences associated with infectivity and trespassing of long noncoding RNAs (lncRNAs) provide a possible recombination mechanism to drive the formation of new variants. This work involved a sequence and structure-based approach to screen SARS-CoV-2 variants in silico, taking into account effects of glycosylation and links to known lncRNAs. Our inquiry commences from the genesis of the alpha variant in December 2019, aiming to discern the occurrence and mechanisms of recombination events. Additionally, we explore the impact of these events on the transcriptional repertoire. An intriguing observation emerges as not all variant sequences exhibit uniform length, prompting an inquiry into the regression pathway of long noncoding sequences. This investigation seeks to unravel why SARS-CoV-2 demonstrates a proclivity towards specific lncRNAs, pinpointing their precise locations and assessing the likelihood of recombination events. Furthermore, we postulate that when a host is exposed to diverse SARS-CoV-2 variants, co-infection may transpire, potentially resulting in the amalgamation of genetic material. This phenomenon, intrinsic to the virus's self-defense mechanism, fosters resistance and genetic diversity. Through the amalgamation of modern bioinformatics tools, including Pangolin, ITOL, NCBI, NONCODE, and Datawrapper, we aim to glean compelling insights. Our pilot study has yielded key candidate sequences associated with SARS-CoV-2, employing similarity and dissimilarity approaches to elucidate vivid sequences pertinent to recombination and transgression pathways. Taken together, the findings suggest that

transgressions involving lncRNAs may be linked with changes in SARS-CoV-2–host interactions driven by glycosylation events.

Applications of Explainable AI in Bioinformatics

Nandini Gandhi, Flame University,

nandi2608@gmail.com

Abstract

Explainable AI (XAI) involves techniques and methods aimed at providing clear, human-understandable explanations for AI and machine learning model decisions. It addresses the opacity of black-box models, offering transparency in how AI systems reach conclusions. Interpretable Machine Learning models can clarify their prediction process and the factors influencing their outcomes. Artificial intelligence (AI) systems, using advanced neural networks and machine learning (ML) algorithms, are crucial for solving complex issues in bioinformatics, biomedical informatics, and precision medicine. However, the complexity of ML models, often viewed as intricate and opaque, can make it hard to understand how they make decisions. This lack of transparency is a challenge for users, decision-makers, and AI developers. In healthcare, where lives are at stake, it's not only important but also legally required for AI systems to be clear and accountable. Another concern is fairness, as algorithmic decisions should be unbiased and not discriminate based on sensitive attributes. Explainable AI (XAI) steps in to address this, aiming to uncover how black-box models work and reveal how AI systems reach their conclusions. Interpretable ML models can clarify their prediction process and identify the factors that influence their results. Bioinformatics has seen a rise in the use of Explainable Artificial Intelligence (XAI), which offers solutions to complex issues and increases decision-making process transparency. Protein structure prediction, genomic sequence analysis, gene-finding, DNA sequencing, gene expression analysis, genome annotation, computer-aided drug design, disease biomarker identification, biological system modeling, drug interaction prediction, and patient phenotyping are just a few of the applications for XAI. However, since biomedical data is complicated and nonlinear, using XAI in bioinformatics poses unique challenges. Despite these difficulties, case studies in text mining, bioimaging, and cancer genomics have demonstrated the potential of XAI techniques to increase transparency. There are many models of explainability in XAI. For instance, two popular models include LIME (Local Interpretable Model-agnostic Explanations) and SHAP (Shapley Additive Explanations), play a crucial role in demystifying the decision-making process of complex machine learning models. LIME operates by generating locally faithful explanations,

providing insights into a model's predictions for a specific instance by approximating it with a simpler, more interpretable model. This allows humans to grasp the reasoning behind a particular prediction. On the other hand, SHAP employs Shapley values from cooperative game theory to attribute a portion of the model's prediction to each feature, highlighting their individual impact. By assigning a quantitative measure of importance, SHAP enables a comprehensive understanding of feature contributions, aiding in identifying critical factors influencing model outputs. Both LIME and SHAP are instrumental tools in building trust and transparency in AI systems, making them invaluable for applications where model interpretability and accountability are paramount. In this work, we plan to review the existing literature on Explainable AI and also provide our own case studies employing explainable AI methodology in bioinformatics

Development of a Web-Resource for Prioritizing Mutations from Next Generation Sequencing Data Using an Agnostic Framework

Anandhu Presannan, Nidheesh M and Bipin G Nair

anandhupresannan@am.amrita.edu

Abstract

Background/Research: Next generation sequencing (NGS) technologies have enabled the identification of genetic variants associated with various diseases. However, the annotation and prioritization of these variants is challenging, especially for non-coding regions and rare diseases. Existing databases and tools rely on external information and may not capture the pathogenicity of novel or population-specific variants. Therefore, there is a need for a web resource that can prioritize mutations from different NGS data using an agnostic framework, which is independent of predictions from generic databases.

Objectives: The main objectives of this proposal are: • To develop a web-resource with an agnostic framework for ascertaining candidate mutations from vivid NGS data. • To validate the agnostic/prioritization framework with various diseased phenotypes using machine learning heuristics.

Materials and Methods: The proposed methodology consists of the following steps: • Data collection: The data for this study will be taken from the already existing WES/NGS datasets available in public domain and from the in-house NGS datasets from various human diseased phenotypes that have been studied by the authors. • Data screening: All mutations with minor allele frequency (MAF) <=0.05 and a minimum depth of coverage of 5 will be judiciously screened as a training dataset. A set of mutations for the test will be taken so that validation will be done followed by precision and recall. • Data integration: The screened mutations will be integrated with other genomic and transcriptomic data, such as long non-coding RNAs (lncRNAs), microRNAs (miRs), and differentially expressed genes (DEGs), to identify the regulatory elements and interactions associated with the diseases. • Data analysis: The integrated data will be analysed using systems genomics and machine learning approaches, to develop a prediction model for prioritizing mutations using an agnostic framework. The model will be based on various features, such as conservation, functional impact, network properties, and subcellular localization. • Data validation: The prediction model will be validated using the test dataset and compared with the

existing databases and tools, such as Clinvar, CADD, and GERP. The validation will also follow the American College of Medical Genetics and Genomics (ACMG) and Kaviar guidelines. • Web server development: The prediction model will be implemented as a web server using Apache/CGI generic with custom options and instantiation of codes. The web server will allow the end-users to upload their NGS data and obtain the prioritized mutations with detailed annotations. Results and **Discussion:** The expected results of this study are: A comprehensive marker catalogue specific to mutations and from end-users whether or not any mutations are pathogenic in nature. The standard operating procedure (SOP) for validating mutations using an agnostic framework. A user-friendly web server for prioritizing mutations from varied NGS data.

Conclusions: The proposed study aims to develop a web-resource for prioritizing mutations from different NGS data using an agnostic framework. In this study we employed integrated systems genomics and machine learning approaches to validate the agnostic/prioritization framework with various diseased phenotypes. The study will provide a useful tool for the researchers and clinicians to identify the pathogenic mutations and their regulatory mechanisms.

Keywords: web-resource, mutations, next generation sequencing, agnostic framework, machine learning

A Holistic Perspective on Host-Pathogen Interactions in Different COVID-19 Severity Levels Through an Integrated Omics Analysis

Mairembam Stelin Singh[1], Mairaj Ahmed Ansari[2], Sailu Yellaboina[3]

Department of Biochemistry, Jamia Hamdard, New Delhi, India[1],

Department of Biotechnology, Jamia Hamdard, New Delhi, India[2],

All India Institute of Medical Sciences, Bibinagar, Hyderabad, Telangana, India[3]

mairembam.stelin@rgu.ac.in

Abstract

The global SARS-CoV-2 pandemic, responsible for millions of deaths worldwide, has raised complex questions about the pathogenesis of COVID-19. The diverse clinical outcomes, ranging from mild to severe disease, remain inadequately understood. To shed light on these mysteries, a novel approach integrating host-pathogen protein interactions and virus-induced host gene expression data was employed. A thorough examination of RNA-Seq data from 1960 samples, originating from 12 projects encompassing various disease severities, revealed genes with differential expression in mild, moderate, and severe COVID-19 cases. Remarkably, when delving into the pathways influenced by the 49 SARS-CoV-2 proteins within the host, a strong connection emerged with processes linked to ribosomal biogenesis, translation, and translocation. An intriguing pattern unfolded: these pathways and related cellular components, such as ribosomal biogenesis and translation, were upregulated in mild cases but downregulated in severe conditions. This suggests that in the face of severe COVID-19, the host's response involves the shutdown of translation pathways targeted by the virus, potentially inhibiting viral replication. However, this strategy might come at a cost, possibly compromising vital host cellular functions, including protein synthesis and the host's capacity to mount an effective antiviral response, with broader implications for human health and well-being.

Semantic Retrieval of Antimicrobial Resistance Information using Natural Language Processing and Deep Learning

Swattik Biswas and Smriti Kandhadai

swattik881@gmail.com

Abstract

Antimicrobial resistance (AMR) is the ability of microorganisms to withstand the effects of antimicrobial drugs, such as antibiotics. It is a growing public health concern that threatens the effectiveness of these drugs and can lead to serious infections that are difficult to treat. Traditional methods of retrieving information from research articles on AMR are often time-consuming and rely on manual curation. To overcome these limitations, researchers are exploring ways to use natural language processing (NLP) and machine learning techniques. The objective of this project is to develop a system to automatically extract and retrieve antimicrobial resistance information from research articles. The model will use NLP techniques to identify and understand the semantic meaning of the text, and then use this information to retrieve relevant information. This approach leverages a combination of NLP techniques, including named entity recognition, entity linking, and semantic similarity measures. By analyzing the textual content of research articles, the system is capable of identifying key entities such as antimicrobial agents, resistant strains, and associated genes. Furthermore, it can establish relationships between these entities, providing valuable insights into the underlying mechanisms of antimicrobial resistance. By automating the process of information extraction from research articles, the project can also save time and resources for researchers and healthcare professionals.

Keywords: Deep Learning, Antimicrobial resistance, Natural Language Processing

AI-Generated Synthetic Genes for Investigating Novel Biomarkers of Systemic Lupus Erythematosus

Peehoo Gaur

IIS School,Jaipur

amita.1983@iisuniv.ac.in

Abstract

Systemic Lupus Erythematosus (SLE) is a complex autoimmune disorder that manifests with a wide range of clinical symptoms and immune system irregularities, largely impacting women from 15-45 years. The timely identification and intervention are paramount in effectively managing SLE. Unfortunately, the deficiency of robust research materials often hampers this process. The intricate interplay among genetic predisposition, environmental influences, and hormonal factors contributes to the onset and progression of the disease. Some genes exhibit a latent risk, like TNIP1, TNFAIP3, etc, akin to a ticking time bomb, necessitating their early identification for proactive measures in preventing SLE. Genomics, a branch of omics, has significantly contributed to the accumulation of large-scale data that fuels these investigations. Machine learning has further enabled the integration and analysis of omics data, leading to the discovery of new biomarkers with potential applications in disease prediction, patient stratification, and precision medicine. In past few years, Generative AI, including pre-trained models, is gaining traction in omics studies, drawing parallels between language constructs and cellular biology for transformative research in genetics and cellular biology. These models have been employed to create artificial protein and drug molecules in various research studies. In 2019, Periasamy and Byrd conducted a phenotype study utilizing Generative Adversarial Network (GAN) technology to analyze images of butterfly rashes associated with SLE. However, there is currently no evidence of the application of such techniques for the generation of synthetic genes in the context of SLE.

Methodology: We aimed to develop a prototype model for gene generation using (GAN) model by employing gene sequences. Eight genes TNIP1, TNFAIP3, RASGRP3, TNXB, TLR7, ATG5, ITGAM and STAT4, known as key contributors to SLE in both Asian and European populations, were selected for designing of the GAN model. The gene sequences were obtained from UniProt. The sequence length varies between 279 and 4314 in the dataset. The GAN model was

implemented in Python on the Google Colab platform. The study made use of essential libraries, including TensorFlow, Pandas, and Keras.

Results and Future Work: Through this model, we successfully generated 5 synthetic gene sequences. These synthetic genes can be subjected to further examination, as they may present a valuable opportunity for conducting comprehensive studies on protein synthesis stemming from these genes. Additionally, they may lead to the discovery of novel biomarkers and facilitate the exploration of their interactions with pharmaceutical compounds.

Editorial Team
Devi Soorya Narayana S
Tejeswini B
Karishma Sahoo
HemaNandini. R. K
Poornimaa. Mu
Premkumar T

INBIX' 23

Discover the forefront of bioinformatics in healthcare at our Inbix-2023 at VIT, where a convergence of experts, researchers, and students from across the globe promises a dynamic exchange of knowledge. Esteemed speakers will illuminate the latest advancements, providing invaluable insights into the transformative role of bioinformatics in shaping the future of healthcare. This gathering serves as a nexus for professionals to connect, fostering collaborative endeavors that transcend geographical boundaries. From groundbreaking research presentations to interactive sessions, the conference aims to unravel the complexities of bioinformatics, offering a platform for intellectual exchange and the cultivation of innovative ideas.

SPONSORS