www.genomicsindia.co.in

Genomics India

1 - 3 FEBRUARY, 2024
SHIV NADAR INSTITUTION OF EMINENCE, DELHI-NCR

ABSTRACT BOOK

Hosted By

SHIV NADAR
INSTITUTION OF EMINENCE DEEMED TO BE UNIVERSITY
DELHI NCR

Powered By

Genotypic
genomics simplified

Co-Organizers

dhiti omics TECHNOLOGIES

qtlomics
TRAITS TO GENES

Media partner

BioSpectrum

Knowledge Partner

IBioM

Welcome to GIC 2024

Join us at GenomicsIndia Conference, GIC2024 as we unlock genomics potential for a more sustainable future. The three-day genomics conference is set to be a dynamic gathering, uniting key players from diverse sectors, including industry, academia, and research. This event is designed to serve as a knowledge-sharing hub, offering a platform for insightful discussions on the latest trends and breakthroughs in the field of genomics.

Throughout the conference, attendees can anticipate a deep dive into cutting-edge research that showcases the forefront of genomics technologies. The primary focus will be on disseminating groundbreaking discoveries and innovations in this rapidly evolving domain.

GIC 2024 promises an unparalleled opportunity for engagement & networking. Attendees will have the chance to connect with industry luminaries, visionaries from the academic sphere, as well as representatives from both startups & established profit and non-profit organizations including Mincorns, Soonicorns, Unicorns, Decacorns & Hectocorns within the biotech sector.

Shiv Nadar Institution of Eminence, Genotypic Technology, Dhitiomics Technologies and QTLomics Technologies looks forward to welcoming you to this Gateway of “Harnessing Genomics for a Sustainable Future“

HARNESSING GENOMICS FOR A SUSTAINABLE FUTURE

Table of contents

Message from Prof. Samir K Brahmachari

I welcome all the delegates to the 5th edition of Genomics India Conference 2024. Twenty five years back, India made a humble beginning in Genomics by establishing in Delhi, a functional Genomics Unit which became full-fledged Institute of Genomics. Same time at Bangalore, the first genomics company "Genotypic Technologies Pvt. Ltd" was established as a Genomics Service Company, empowering laboratories across India to use Genomics techniques to participate in "Genomics Revolution in India". As we look back, we have come a long way from a humble beginning to a thriving Genomics Community as GIC2024 is held in Shiv Nadar University, Greater Noida. As higher education and Scientific research in India has got expanded from public funded institutions to Private funded Institutions, we have new opportunities to attract new talents to the frontiers of Genomics.

Science has moved from a single PI driven small laboratory initiative (Science 1.0) to mega science, involving a large number laboratories like the Human Genome Project, or the Indian Genome Variation project (Science 2.0) to crowd-sourcing of knowledge and expertise for projects like Mapping the Human Brain or Open-Source Drug Discovery (Science 3.0). We stand in the threshold of new frontiers Science 4.0, wherein AI driven tools like Alfa fold 2.0, ChatGPT4.0 and Google Gemini or IBM Watson will be a part of the research team.

Hence the future of Genomics will move from data collection to data integration; from understanding gene function to simulating gene-function Networks; from developing predictive DNA markers for diseases to whole genome scale development of predictive medicines. We will move from preventive medicine to wellness genomics. AI tools and natural language processing will be an integral part of data curation and analysis. Predictive AI will completely transform Genomics Research of the future. It will not only have a profound influence on healthcare, but also on Agri-genomics. AI tools will be developed to analyse soil metagenomics to be compatible with the crop selection to increase productivity with a sustainable ecosystem.

Let us hope, on the 10th Genomics India Conference, some of the young generation genomics scientists of India will lead India to the forefront of Science 4.0 in the field of Integrative Genomics.

We look forward to a productive and vibrant Genomics India Conference 2024.

Prof. Samir K Brahmachari

GIC 2024 - NGSA Workshop (C Block) C021 ground floor		
8:00 AM	**Registration - C021 registration desk**	
Feb 1	**Topic**	**Trainer**
8:30 - 9:15pm	**Inaugural Lecture** Navigating the Genomic Frontier: Unveiling Next-Generation Sequencing Essentials	Dr. Raja Mugasimangalam
9.15 - 10:00pm	Basics of NGS Data Analysis: Denovo assembly and Gene prediction	Mr Nihar Bachan Das
10.00 - 10.45pm	NGS data analysis without commandline	Dr Raja Mugasimangalam Dr Meenal Vyas
10:45 -11:00pm	**Tea Break**	
11:00 - 11:45pm	Unfolding an Epigenome with the revolutionary native molecule Nanopore Sequencing Technology	Dr Meenal Vyas
11.45 - 12.15pm	Insights into transcriptome and small RNA data analysis	Dr Rohini Garg
12:15 - 12:45pm	The Long and Short of sequencing	Dr Vipin Singh
12:45 - 1:15pm	Unveiling Genomic Complexity: Insights into Precision Alignment, Variant Calling And Variant Annotations	Dr Deeksha Bhartiya
1:15 – 2:00 pm	**Lunch**	
2:00 - 2:45pm	NGS For Biopharma and Vaccine Industry	Dr Sudha Rao
2:45 - 3:30pm	Alignment to Variant Calling and Reporting	Mr Nihar Bachan Das
3:30 - 4:00pm	Insights Into The Microbiome Analysis Using NGS Data	Ms Garima Sanoria
4:00 - 4:30pm	NGSA Test, Q&A and NGSA Result	Dr Raja, Dr Rohini Team of Trainers
4:30 - 5:15pm	Solving Biological Problems through Protein and Genome Sequence Analysis (Open to all)	Prof Samir K Brahmachari
5.15 - 5.30pm	NGSA Workshop Prize Distribution	Dr Raja Mugasimangalam, Prof Samir K Brahmachari
5:30 - 6:00pm	**High tea at G Block**	

GIC 2024 - AMR SATELLITE PROGRAM (HALL NUMBER D-128) TITLE: AMR IN THE GENOMICS ERA GENOMIC SURVEILLANCE AND PREDICTION OF ANTI MICROBIAL RESISTANCE		
8:00 AM	**Registration - C021 registration desk**	
Feb 1	**Topic**	**Speaker**
2:00 – 5:00 PM	**Genomic surveillance and prediction of anti microbial resistance**	Organizer: Dr Krishna Jayadev, SNIoE Chair: Dr Manoj Kumar, CSIR - IMTech Co-Chair: Dr V Udhayakumar Program Director, Emory University and Task Force for Global Health (Atlanta), USA Dr Shraddha Karve, Ashoka University Dr Praveen Bharti NIMR , New Delhi Dr KH Reeta, AIIMS, New Delhi Dr Ashish Das, BITS, Pilani Dr Neeraja Venkateswaran, Tetracore, Inc, Maryland Dr Kanchan Bhardwaj, Manav Rachana, Faridabad Mr Praveen K S, PATH South-Asia
6.15 – 7:30 PM	Inauguration Genomics India Conference	Welcome address: Dr Sanjeev Galande, Dean School of Life Sciences, SNIOE Lighting of the Lamp: Prof Samir K Brahmachari, Former Director General CSIR Prof. Krishnamurthy Kannan, Proprietor KK Biotech Patron's Address: Mr Shikhar Malhotra, Chancellor SNIOE, Director HCL Corporation Dr Ananya Mukherjee, Vice-Chancellor, Shiv Nadar Institution of Eminence, Delhi NCR GIC Evolution: Dr Raja Mugasimangalam, Founder & CEO Genotypic Technology Keynote Address by Chief Guest: Dr Rajesh S Gokhale, Secretary DBT Vote of thanks Dr Sudha Rao, Co Founder and Executive Director, Genotypic Technology

Feb 2 CONFERENCE	G Block	C Block
9:30- 11:00 AM	**Session 1- Revolutionizing Genomic Medicine**	**8:30 - 9:15 AM** **Nanopore live sequencing run** **(OPEN TO ALL)**
	Dr Sanjeev Galande - Chair	
	Dr Michael Sagner Director - European Society of Preventive Medicine	
	Dr Vijay Tiwari Prof & Head, University of Southern Denmark	
	Dr Dasaradhi Palakodeti Principal Investigator InStem, Bengaluru	
	Dr Jawahar Swaminathan *Director, Informatics, Systems Integration & Interpretation - Production Scale Genomics - OXFORD NANOPORE TECHNOLOGIES*	
11:00 - 11:20 AM	**TEA**	
11:20 - 12:50 PM	**Session 2- Precision Oncology: Targeting Cancer with Genomic Insights**	**Session 3- Better Crops with Genomics based breeding**
	Dr Sudha Rao - Chair **Co Founder and Executive Director, Genotypic Technology, Bengaluru**	**Dr Jagadish Mittur - Chair** **Director Operations - Rico Winery Pvt ltd, Bengaluru**
	Dr B S Ajai Kumar Executive Chairman - HCG Enterprises Ltd, Bengaluru	**Dr Akhilesh K Tyagi** Senior Professor, DPMB, DU - South Campus, Delhi
	Dr Natarajan Muthusamy Associate Director of Academic Affairs, The Ohio State University, USA	**Dr Ajit K Shahsany** Director, CSIR- NBRI, Lucknow
	Dr Jugnu Jain Co-Founder & CEO , Sapien Biosciences, Hyderabad	**Dr Wolfram Weckwerth** Professor - University of Vienna, Austria
	Mr Mainak Chakraborty Senior Solutions Architect Amazon Web Services (AWS)	**Dr Ravi Kumar Chilukoti** Head – India Technical Support & PM, MGI TECH CO.LTD.
12:50 - 2:20 PM	**Poster & Networking**	
1:30 - 2:20 PM	**Lunch**	
2:30 - 4:00 PM	**Session - 4 Synergies of AI ML and genomics:transforming data into discovery**	**Session -5 Lightening Talks selected from poster abstracts**
	Mr Periasamy Kaliyannan - Chair **Director QTLOmics, Financial Systems Consultant** **Panalpina Asia Pacific Management Pte Ltd, Singapore**	Speaker 1-8
	Dr Shobha Vijayaraghavan Associate Director- Bio-Innovation, Accenture, Bengaluru	
	Dr Vidur Mahajan CEO - CARPL.ai, New Delhi	
	Dr Sudhakaran Prabhakaran CEO and CSO, NonExome Inc, Massachusetts	

	Mr Nagendra Kumar S Maroor Managing Director IKA INDIA PRIVATE LTD	
4:00 - 4:20 PM	**Tea Break**	
4:20 - 5:50 PM	**Session 6: Population genomics**	**Session 7- Microbial Genomics: from diversity to diseases (Chair Dr Manickam)**
	Dr Ashish Das (Chair) Senior Professor, BITS, Pilani, Rajasthan	**Dr Natesan Manickam - Chair** Chief Scientist & Professor CSIR-Indian Institute of Toxicology Research, Lucknow
	Dr K Thangaraj J C Bose Fellow, CSIR - CCMB Hyderabad	**Dr Niyaz Ahmed** Professor - University of Hyderabad, Hyderabad
	Dr Mrinalini Watsa Research Scientist, SDZWA, San Diego, USA	**Prof Royana Singh** Institute of Medical Sciences - BHU, Varanasi
	Dr Vinod Scaria Senior Consultant, Vishwanath Cancer Care Foundation & Adjunct Prof IITK	**Dr Luis Montaner** Vice President, Scientific Operations Wistar Institute
	-	**Dr Nitya Nand Sharma** Strategic Business Unit Head - Applied PREMAS LIFESCIENCES
	-	**Dr. Ashwani Kumar Kamal** Technical Sales Manager (NEB) NEW ENGLAND BIOLABS
6:00 - 6:45 PM	**Keynote - Dr Eric Green** **Director, NHGRI, USA** **Chaired by Prof Samir K Brahmachari** **Founder Director, CSIR-IGIB and Former Director General, CSIR India**	**XXXXXXXXX**
7:00 PM	**Buses will leave for Gala Dinner venue**	
7.30 PM	**Dinner**	

Feb 3 CONFERENCE	G Block	C Block
9:00 - 10:30 AM	**Session 8- Emerging Next generation Sequencing data analysis methods and computational Genomics**	
	Dr Raja Mugasimangalam (Chair) **Founder & CEO Genotypic Technology, Bengaluru**	
	Dr. Chirag Jain Assistant Professor, Indian Institute of Science, Bengaluru	
	Dr Shailesh Kumar Staff Scientist, NIPGR, New Delhi	
	Prof Mukesh Jain Professor - JNU, New Delhi	
	Prof Dr Shuhua Xu Distinguished Professor & Principal Investigator, Fudan University, China	
10:30 - 11:00 AM	**Keynote - Dr Shiv K Sarin** Director, Institute of Liver and Biliary Sciences, New Delhi **Chaired by Dr Rangaraj Selvarangan** Director, Emerging Infections Research & Microbiology Laboratory Children's Mercy Hospital, Kansas City, USA	
11:00 - 11:20 AM	**Tea**	
11:20- 12:50 PM	**Session 9- Integrative -Omics approaches - unveiling complex biological systems**	**Session 10- Biopharma Vaccines and Diagnostics**
	Dr Rakesh Laishram Scientist F, RGCB Thiruvananthapuram	**Dr Radhakrishnan Nair (Chair & Speaker)** Scientist F, RGCB, Kerala
	Dr Mehdi Totonchi Asst. Professor, Royan Stem Cell Institute, Iran	**Dr Kar Muthumani** Chief Scientific Officer, GeneOne Life Science, United States
	Dr Paresh Sharma Scientist E, NIAB, Hyderabad	**Mr Luan Nguyen** APAC Business Manager, Vizgen SPINCO BIOTECH
	Dr Prashant Srivastava National Heart and Lung Institute, Imperial College, London	**Dr Divyank Mahajan** Product Manager - Genomics SPINCO BIOTECH
12:50 - 2:30 PM	**Poster session + Lunch & Networking + Exhibits**	

2:30 to 4:45 PM	**Session 11- Epigenomics shaping genetic destiny - 1**	**Session -12 Genomics Start ups and MSME**
	Dr Esteban Ballestar Professor, Josep Carreras Research Institute, Spain	**Ms Gunjan Sharma – Moderator** Biospectrum
	Dr Kundan Sengupta Professor, IISER Pune	**Dr Gopalakrishna Ramaswamy** Founder and CEO at theraCUES, Bengaluru
	Dr Beena Pillai Scientist, CSIR - IGIB, Delhi	**Dr Deeksha Bhartiya** Founder - Genomiki Solutions, Delhi
	Prof Sanjeev Shukla Professor - Biological Sciences IISER, Bhopal	**Dr Manoj Gopalkrishnan** Founder & CEO Algorithmic Biologics
	Prof. Rakesh Mishra (Chair & Speaker) Director - TIGS, Bengaluru	**Dr Anirvan Chatterjee** Co-Founder & CEO, Haystack Analytics
	Dr Simona Ferraioli Application Scientist, Lexogen, Vienna, Austria	**Dr Aakanksha Pant** Scientist and Product Manager Indus Health Plus
	-	**Mr Sandeep Bhatia** Founder, Helo Health, Mumbai
4:45 to 5:00 PM	**Tea**	
5:00 to 5:30 PM	**Keynote - Dr Rangaraj Selvarangan** Director, Emerging Infections Research & Microbiology Laboratory Children's Mercy Hospital, Kansas City, USA	
5:30 to 6:15 PM	**Parting thoughts:** **Dr Sudha, Co Founder and Executive Director, Genotypic Technology** **Awards Distribution**	

Speakers

Speaker Abstracts - Table of contents

Session 1- Revolutionizing Genomic Medicine

Session 2- Precision Oncology: Targeting Cancer with Genomic Insights

Session 3- Better Crops with Genomics based breeding

Session 4 - Synergies of AI ML and genomics: transforming data into discovery

Session 6: Population genomics

Session 7- Microbial Genomics: from diversity to diseases

Session 8- Emerging Next generation Sequencing data analysis methods and computational Genomics

Session 9- Integrative -Omics approaches - unveiling complex biological systems

Session 11- Epigenomics shaping genetic destiny

Session -12 Genomics Start ups and MSME

Speaker Abstracts

Session 1- Revolutionizing Genomic Medicine

1. **Scalable Bioinformatics Solutions for Human Rare Diseases and Cancer with Oxford Nanopore sensing.**

Speaker: Dr Jawahar Swaminathan, Director, Informatics, Systems Integration & Interpretation - Production Scale Genomics - Oxford Nanopore Technologies

Abstract

Oxford Nanopore Technologies provides a rapid and scalable platform that can sequence any length of DNA or RNA, short to ultra-long in unmodified form. Nanopore sequencing captures more genomic variation across all variant categories including SNPs/InDels, Copy Number and Structural Variants, Short-Tandem Repeats and Methylation (5mC, 5hmC) with unprecedented accuracy and speed. This presentation will provide a brief introduction of the technology, its capabilities, and use in cutting edge applications for detection of causal variation in rare diseases and cancer. I will briefly touch upon our readily accessible and packaged bioinformatics workflows that remove informatics barrier and enable a fully functional and automated sequence-to-answer pipeline. The main focus of my talk will be, using exemplar studies from cancer and rare disease, to be able to show how Oxford Nanopore Technologies sensing provides a deeper insight into mechanisms of disease across different variant types.

2. Leukemia Initiating Cells in Chronic Lymphocytic Leukemia Do they exist and can we target them? .. Genomic Insights…

Speaker: Dr Natarajan (raj) Muthusamy, Associate Director of Academic Affairs, OSU, USA

Abstract

The existence of rare 'leukemia initiating cells' (LICs) in chronic lymphocytic leukemia (CLL) remains controversial and understudied due to the difficulty in isolating and identifying the tumor initiating cells. We have developed a novel platform to introduce molecular beacon probes into single live cells that facilitates identification, isolation, imaging and characterization of heterogeneous LICs. Using limited-cell fluorescent activated cell sorter sequencing (LC-FACSeq), we are able to detect, and monitor rare LICs during leukemogenesis and characterize their differential drug sensitivity. Disease-associated mutation accumulation in developing B lymphoid but not myeloid lineage in CLL patient hematopoietic stem cells (CLL-HSCs), and development of independent clonal CLL-like cells in murine patient-derived xenograft models, suggests the existence of CLL LICs. Furthermore, we identify differential protein ubiquitination and unfolding response gene signatures in GATA2^{high} CLL-HSCs that exhibit differential drug sensitivity responses compared to GATA2^{low}CLL-HSCs. These results highlight the existence of therapeutically targetable disease precursors in CLL.

Session 2- Precision Oncology: Targeting Cancer with Genomic Insights

3. Understanding Indian cancer genomics for improving diagnosis and treatment

Speaker: Dr Jugnu Jain, Co-Founder & CEO, Sapien Biosciences, Hyderabad

Abstract

NGS technology has emerged as a powerful tool to understand the underlying biology of the initiation and progression of cancers. Identification of targetable driver mutations and mutations leading to resistance to treatment have enabled the stratification of patients into specific subgroups and expansion of the use of FDA-approved targeted therapies and immunotherapies based on the molecular profile of individual cancer patients. However, omics data of Asian and African patients is under-represented in publicly available databases. To address this gap, we have used the 46 cancer-relevant oncogenes OncomineDx panel, approved by the FDA as a companion diagnostic to aid in the selection of lung and other solid cancers for treatment with approved targeted therapies. Lung adenocarcinoma is the most prevalent form of non-small cell lung cancer (NSCLC), and remains a leading cause of cancer-related death globally and in India, with a 5-year survival rate being below 10%. Using NSCLC as an example, our data on the prevalence of molecular alterations in Indian NSCLC patients will be presented, focusing on the actionability of the identified gene variants using OncoKB.

Session 2- Precision Oncology: Targeting Cancer with Genomic Insights

4. Accelerate Clinical Diagnostics and Genomics Research using AWS cloud

Speaker: Mr Mainak Chakraborty, Senior Solutions Architect, Amazon Web Services (AWS)

Abstract

AWS for Genomics matches the needs of research organizations with innovative technologies to provide scalable, secure, and cost-effective tools that accelerate genomic discoveries. From solutions to migrate and securely store genomic data on AWS, to tools that accelerate secondary and tertiary analysis, to services that integrate genomic data into multi-modal datasets, AWS for Genomics offers solutions across the genomics workflows. For almost a decade, industry leaders such as Illumina, Genomics England, Ancestry, UK Biobank and GRAIL have leveraged AWS to do more with their data. In this presentation, I will show you how to effectively use AWS cloud to innovate faster.

Session 3- Better Crops with Genomics based breeding

5. Diversifying Crop Genome

Speaker: Prof Akhilesh K. Tyagi , Senior Professor, DPMB - DU - South Campus, Delhi

Reaching the targets of sustainable food availability depends on the success of agricultural sector. To achieve more output from crops, humans have been selecting natural diversity and genetically modifying them for long. Success of such practices can be seen in an increase in food production world-wide. In recent years, marker-assisted breeding has taken lead in combination with conventional breeding due to crop genome sequencing and trait-associated marker development. More diversity and potential of crop genomes is being unraveled by pan-genome approach. Knowledge about genes and their regulatory elements has allowed the transfer of trait-associated genes across taxa. The potential of genome/gene editing is being explored to create new alleles and knockdown of undesirable genes. These approaches promise rapid diversification of crop genomes to create benefits to society.

6. PANOMICS meets large germplasm collections for sustainable agriculture and food security

Speaker: Prof Wolfram Weckwerth, Professor - University of Vienna, Austria

Abstract

Since the last 80 years after the start of the green revolution, agriculture has intensified and focused on meeting human food demands. To produce more food, agriculture has used excessive nitrogen fertilizers and at the same time has to cope with declining resources such as soil vitality, water, arable land and increasing environmental pressures such as climate change and pollution. Exclusive selection of high productivity in crop germplasm has decreased stress resilience and nitrogen use efficiency (NUE). Integration of multiomics analysis using metabolomics, proteomics, RNAseq, metabolic modelling, AI, machine learning applied to natural genetic variation of large germplasm collections, in short PANOMICS technology [1,2] goes beyond classical genomic techniques and offers valuable tools to help meeting these multidimensional challenges from sustainable agricultural processes to environmental and human health. In this lecture, I explore the possibilities of using these breakthrough technologies. Application of PANOMICS technology to specific traits provides information-driven agriculture that leads to improved and sustainable agricultural production systems [1, 2].
[1] Weckwerth W et al (2020) PANOMICS meets germplasm. Plant Biotechnol J 18: 1507-1525
[2] Ghatak, A., et al. (2023) PANOMICS at the interface of root-soil microbiome and BNI. Trends Plant Sci 28, 106-122

Session 4 - Synergies of AI ML and genomics: transforming data into discovery

7. **MGI's Novel Solution for Single-Cell Biology that can be Integrated with Spatial Transcriptomics -STOmics.**

Speaker: Dr Ravi Kumar Chilukoti, Head – India Technical Support & PM, MGI Tech Co.Ltd

Abstract

This talk will focus on MGI's novel single-cell solution, DNBeLab C4 Station that is relatively inexpensive, simple to use and empowering cancer researchers to expand the knowledge of cellular heterogeneity in tumor microenvironment. DNBeLab C4 Station is a pocket-sized portable device streamlines the process of droplet formation with a very low multiplet rate of <5%, >50% of cell recovery and >4000 genes/cell. The performance of DNBeLab C4 Station has shown a high degree of consistency in Cell Clustering, Annotation, and Cell Ratio among different libraries. Further, the performance of DNBelab-C4 V2.0 platform has resulted in high concordance to the competitor's mainstream platform. The performance of Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) from MGI has resulted in multiplet rate of <4% which is lower to that of than competitor (10x = 8%, 10K of nuclei), high capture of nuclei (>60%) and high consistency among multiple libraries. DNBelab C4 station has been used in many studies covering research areas of tumor biology, neurobiology, germline transmission, organogenesis, clinical diagnostics, and so on. This C4 station has been utilized on more than 40 types of species, 300 types of tissues and the generated data has been published in more than 40 articles.

Stereo-seq (SpaTial Enhanced REsolution Omic-Sequencing) from BGI group is revolutionizing the spatial temporal systems biology research by providing high degree of resolution (500 nm that is far higher than competitor having low resolution of 100 µm), increased Capture area from 12 cm (commercial) –to 132 cm (BGI R&D) and Cell segmentation to sub-cellular resolution, which is not being provided by any competitor. BGI has algorithms such as SPOTlight that can link the two datasets from single cell and Stereo-seq (spatial biology) enabling deep gene expression coverage and the annotation of cell identify at single-cell level in spatial biology studies.

Session 6: Population genomics

8. Population Genomics and Public Health

Speaker: Dr K Thangaraj, J C Bose Fellow, CSIR - CCMB Hyderabad

Abstract

Modern India is a region of remarkable cultural, linguistic, and genetic diversity with over 4,500 anthropologically well-defined groups. We have been studying various Indian/South Asian populations to understand the origin, migration, and impact of consanguinity and endogamy on our health and diseases. To assess the impact of endogamy, we have analysed samples from more than 2,800 individuals from over 275 distinct South Asian groups and found that 81 out of 275 groups, have a strong founder event than the one that occurred in both Finns and Ashkenazi Jews. Further, we went back to the populations that have strong founder event, and found some of the populations have high prevalence of population-specific diseases. Notably, Kallar population from Tamil Nadu has high frequency of Junctional Herlitz Epidermolysis Bullosa disease, characterized by vesicobullous skin lesions, oral mucositis, congenital heart disease, and premature death. Subsequently, exome sequencing found a homozygous 11 base pair deletion in the *LAMB3* gene of the patients, whereas the parents were heterozygous for this deletion. Likewise, large number of population-specific diseases exist in endogamous populations. We are now analysing populations with strong founder event to understand the genetic basis of diseases, which would help us to provide prenatal and premarital counseling to avoid such diseases in the family/population in the future.

Session 6: Population genomics

9. **Conservation Genomics on the Go: A Changing Landscape**

Speaker: Dr Mrinalini Watsa, Research Scientist, SDZWA, San Diego

Abstract

Conservation genomics has long existed alongside a sometimes-colonial paradigm of extractive resource utilisation: biologists enter a habitat, collect specimens, and remove them for analysis elsewhere. Historically, the largest and most egregious examples of this practice have included the extraction of such resources from the Global South to laboratories in the Global North. But even within countries, sample collection takes place in resource-poor, yet highly biodiverse regions, but analysis is restricted to resource-rich urban landscapes. But when samples move in this way, they take with them also "opportunity" - the right for those who collect samples from the animals they know best, to participate in the downstream analyses of these materials. Conservation genomics, therefore, has had to reckon with two major upheavals that have disrupted this long-supported system. First, a global recognition of the harms and colonial history behind such extractive helicopter science. And second, the appearance on the sequencing landscape of a small, affordable, and now accurate sequencing technology. In this talk, I will discuss how the SDZ Alliance is applying portable nanopore sequencing technology alongside our partners within the umbrella of the In Situ Laboratory Initiative - a movement towards a global decentralized and community-run One Health laboratory network.

Session 7- Microbial Genomics: from diversity to diseases

10. Genomic of antimicrobial resistance in major human pathogens: implications for global infection control priorities

Speaker: Niyaz Ahmed, Pathogen Biology Laboratory, Department of Biotechnology & Bioinformatics, University of Hyderabad, Hyderabad, Telangana State

Abstract

Policies relevant to public health have a great deal of dependence on the dimensionality of infection burden in communities. The next major challenge for public health epidemiologists is to understand the interactions and functioning of some of the major human and animal pathogens at the community level, both inside the host species and the outside environment. Recent genomics-driven research, including our own studies has provided insights into the transmission and evolutionary dynamics of major human pathogens such as *Mycobacterium tuberculosis*, *Escherichia coli, Klebsiella pneumoniae, Vibrio cholerae, Helicobacter pylori* and *Salmonella* spp. Especially, our studies based on identification of various dominant lineages of some of these organisms based on AI and machine learning point to the possibility of a system for prediction of antimicrobial resistance (AMR) as some lineages have higher propensity to ensure survival through novel pathways and lifestyles. This is pertinent in the light of emerging AMR being one of the immediate threats posed by pathogenic bacteria in the form of a multi-layered fitness advantage which manifests as phenotypic drug resistance at the level of clinics and field settings. Further, it will also be possible to present and discuss our recent analyses based on an extensive trajectory of AMR in clinical and environmental isolates over a period exceeding 14 years. Discussion of these findings would enlighten us of the sheer prowess of bacterial organisms to evolve and spread with new fitness advantages such as AMR phenotypes of different types. It is possible that approaches based on high resolution genomic epidemiology, as mentioned above, would likely be successful in targeting the spread of multiple drug resistant pathogens and guiding the global infection control priorities and policies.

Session 7- Microbial Genomics: from diversity to diseases

11. High throughput nanopore sequencing of SARS-CoV-2 viral genomes from patient samples in India

Speaker: Dr Royana Singh, Institute of Medical Sciences - BHU, Varanasi

Abstract

Year 2019 saw the emergence of a novel coronavirus that caused a potentially fatal respiratory viral infection in Wuhan, China. The COVID-19 pandemic was brought on by the new coronavirus, known as SARS-CoV-2, which had spread around the world by early 2020. Viral genome sequencing can be used to track the introduction, dissemination, and evolution of SARS-CoV-2 due to its high rates of mutation and infection. Monitoring SARS-CoV-2 viral mutations is also anticipated to be beneficial for efforts aimed at creating efficient public health regulations, medications, or vaccinations to treat or prevent COVID-19. Here, we present a set of thorough working protocols for high throughput sequencing of SARS-CoV-2 viral genomes utilising a MinION and GRIDION device. These protocols include everything from viral RNA extraction to analysis using well-established visualisation tools. Using ARTIC primer sets and long-read nanopore sequencing technology, this collection of methods should be a dependable "how-to" guide for producing high-quality SARS-CoV-2 genome sequences. Many of the activities involved in preparation, quality control, and analysis will also usually apply to other sequencing platforms.

12. BEAT-HIV: 2024 Progress in harnessing NK function by gene modification and bi-specific antibodies in HIV cure-directed strategies

Speaker: Dr Luis J. Montaner, Vice President, Scientific Operations Wistar Institute

Abstract

Presentation will highlight 2024 priorities in HIV-cure-directed efforts with emphasis on the BEAT-HIV Martin Delaney Collaboratory program (beat-hiv.org) . Current data on pre-clinical cure-directed NK-based strategies against HIV will be reviewed.

Objectives:

#1: Outline the current priorities in HIV cure-directed efforts

#2: Highlight new data from BEAT-HIV pre-clinical cure-directed efforts with NK-based strategies

Session 7- Microbial Genomics: from diversity to diseases

13. **Speaker: Ashwani Kumar Kamal**

Abstract:

Successful Next Generation Sequencing (NGS) hinges on several factors, but one of the most impactful of them is the choice of library preparation workflow. As NGS matures into an essential analysis method, libraries of the highest quality are required, in yields sufficient for a wide range of applications and throughputs, from single-cell and cfDNA to PCR-free whole-genome analysis. Automation compatibility is a major consideration, too, as liquid-handling robots are becoming more accessible and commonplace.

14. An improved method to compute telomere-to-telomere genome assemblies using long reads

Speaker: Dr Chirag Jain, Assistant Professor, IISc, Bengaluru

Abstract

Genome sequences of several species, including humans, comprise challenging repeats. Despite four decades of research, *de novo* genome assembly remains one of computational biology's fundamental algorithmic problems. Recent breakthroughs in long-read sequencing technologies promise automated assembly of complete haplotype-resolved human genomes for the first time. However, there are open algorithmic challenges that need to be resolved before complete genomes can be computed on a routine basis. In this talk, I will discuss recently our new algorithms and mathematical analysis that address some of these challenges. Using Oxford Nanopore and PacBio HiFi sequencing data of the HG002 human genome, we achieved a twofold increase in the assembly contiguity and the number of telomere-to-telomere sequences compared to the state-of-the-art methods.

Session 8- Emerging Next generation Sequencing data analysis methods and computational Genomics

15. Identification of Novel RNAs in high throughput sequencing datasets

Speaker: Dr Shailesh Kumar, Staff Scientist, NIPGR, New Delhi

Abstract

My laboratory is exploring novel components of genome regulatory circuits. Our choice of molecules includes Transfer RNA (tRNA)-derived non-coding RNAs (tncRNAs) rRNA derived fragments, and fusion transcripts. Since the tncRNAs and fusion transcripts belong to novel classes of non-coding RNAs, and are not fully explored in plants, so, the development of databases and methodologies for their detection and classification was the first step in our research. Now, we are focusing on pinpointing the tissue and stress-responsive fusions and tncRNAs, and their mechanism of genome regulation so that those specific molecules can be used to improve plant traits. For this purpose, multi-omics datasets will be used for the *Arabidopsis*, chickpea, and rice by combining both Insilco and experimental approaches.

Session 9- Integrative -Omics approaches - unveiling complex biological systems

16. Unveiling the genetic tapestry: Integrated omics approaches for divulging molecular signatures associated with agronomic traits

Speaker: Prof Mukesh Jain, Professor - JNU, New Delhi

Abstract

Next-generation sequencing (NGS) technologies provide a revolutionary tool with numerous applications. Further, advances in multi-omics approaches have accelerated efforts for gene discovery and understanding the molecular basis of important agronomic traits for undertaking translational research. I shall illustrate the use of NGS technologies in generation of genomic resources and connecting genes and genomic variations to agronomic traits giving examples from a non-model (chickpea) crop plant. We sequenced the genome/transcriptome of chickpea to reveal the gene space and genetic variations associated with important agronomic traits. A comprehensive analysis of transcriptome dynamics during seed development in two chickpea cultivars with contrasting seed size identified a significant proportion of the genes exhibiting stage- and cultivar-specific expression patterns. The transcriptional changes in cell cycle, endoreduplication, carbohydrate metabolism and hormone signal transduction pathways were found to determine seed size/weight in the chickpea cultivars. Further, we revealed the regulation of candidate genes via differential DNA methylation and non-coding RNAs. Recently, we have developed a comprehensive gene expression atlas and demonstrated its applications in functional genomic studies and candidate gene discovery. Our studies provide insights into the molecular signatures and regulatory mechanisms underlying agronomic traits and will surely facilitate research in various areas of functional and translational genomics in crop plants.

Session 11- Epigenomics shaping genetic destiny

17. Epigenomic Reprogramming of Immune Responses in Inflammation and Cancer

Speaker: Prof Esteban Ballestar, Professor, Josep Carreras Research Institute, Spain

Abstract

Monocytes, macrophages, and dendritic cells (DCs) collectively constitute the mononuclear phagocyte system within innate immunity. These myeloid cells exhibit significant plasticity, showcasing a shared spectrum of functions that includes phagocytosis, cytotoxicity, and antigen presentation, among others. The immunological characteristics of these cells depend on the surrounding environment, encompassing distinct tissue types and pathological niches, such as inflamed tissues or the tumor microenvironment. Various cytokines, hormones, and other molecules released by cells modulate the properties of macrophages and DCs, either promoting pro-inflammatory or anti-inflammatory (or tolerogenic) responses through epigenetic control, including DNA methylation. In solid cancers, the hypoxic conditions of the tumor microenvironment and other niches can reprogram macrophages, giving rise to distinctive sets of features. In this lecture, I will discuss our work investigating the diverse effects of extracellular factors in shaping the properties of macrophages and DCs, exploring the interplay between these environmental cues, transcription factors, and the epigenomic remodeling, and assessing the impact on different pathological contexts, including inflamed compartments and the tumor microenvironment.

Session 11- Epigenomics shaping genetic destiny

18. Does Mechano-epigenomics shape genetic destiny?

Speaker: Prof Kundan Sengupta, Professor, Chromosome Biology lab, IISER Pune

Abstract

Cells are constantly subjected to mechanical stress and strain. We attempted to recapitulate the stiffness regimen of tissues by creating softer matrices with stiffness ranging from ~2 to ~55 kPa. Cells exposed to the softer polyacrylamide substrates, showed an increase in surface area along with a markedly altered transcriptome as revealed by RNA-Seq analyses. Interestingly, these cells on the softer matrices also showed a significantly altered configuration of chromosome territories in the 3D space of the interphase nucleus. Remarkably, nuclear envelope proteins - lamins and LINC (linker of nucleoskeleton and cytoskeleton) complex, mislocalized into the nuclear interior. Interestingly, the nuclear organization of active and inactive histone marks was also altered along with the activation of Emerin phosphorylation - the inhibition of which attenuated the relay of mechanical signals into the nucleus. Taken together, our studies highlight a remarkable plasticity, adaptability and resilience of the epigenome and genome to counter and protect genome organization and function from external mechanical forces.

Session 11- Epigenomics shaping genetic destiny

19. Inherited lncRNA Durga regulates neurodevelopmental gene expression pathways with persistent behavioural effects

Speaker: Dr Beena Pillai, Scientist, CSIR - IGIB, Delhi

Abstract

Long non-coding RNAs are a diverse group of transcripts that presumably do not code for proteins due to the absence of ORFs longer than 200nt. Poor conservation and low expression, pose challenges in assigning function to thousands of long non-coding RNAs. Their role in most biological phenomena and disease conditions remain under studied. Our lab has previously shown that several hundred lncRNAs, for instance, the well studied lncRNA Cyrano, are inherited in zebrafish, an ideal model for studying early development.

One such novel inherited RNA named Durga arises from the genomic locus that codes for the Kalirn gene. We studied the role of this RNA, uncovering its cellular and molecular role in the regulation of neuronal genes in a narrow developmental time window. The transcript is inherited at the single cell stage at low levels but ectopically injected Durga RNA can induce the expression of its zygotic counterpart, in specific brain regions. This auto-regulation of the gene leads to its highly localized expression in the zebrafish Habenula. Knockdown of the RNA leads to a reduction in the expression of critical neurodevelopmental genes, for instance Auts2b, the transcription factor that controls a large group of neurodevelopmental genes In keeping with its expression in the habenula, abnormally high durga expression during early development leads to persistent behavioural effects on the fear response in adult fish.

20. Decoding the Molecular Symphony of Gene Regula8on and Splicing in hypoxic Cancer cells

Speaker: Prof Sanjeev Shukla, Professor - Biological Sciences IISER Bhopal

Abstract

In this talk, I will delve into the fascinating world of hypoxia-mediated regulation of gene expression and alternative splicing in cancer. Hypoxia, a state of low oxygen, plays a crucial role in tumor progression and metastasis. Our laboratory inves3gates the molecular mechanisms by which hypoxia alters gene expression and splicing patterns, leading to significant changes in cellular behavior and phenotype. In one of our studies, we have demonstrated that the non-canonical pyruvate kinase M2 (PKM2) regulates the hypoxia- responsive element (HRE) mediated upregula3on of PFKFB3 in hypoxic cancer cells by facilita3ng HIF-1α and p300 enrichment, while the absence of PKM2 leads to poised chroma3n states that restrict HIF-2α-induced PFKFB3 expression, and the inhibition of PKM2 nuclear transloca3on by Shikonin demonstrates poten3al therapeu3c implica3ons for breast cancer. In another story, we have discovered an intricate interplay between the splicing factor SRSF2, DNA methylation, CTCF recruitment, and RNA polymerase II occupancy to regulate the angiogenesis-related isoforms of VEGFA-165 under normoxic and hypoxic conditions. Additionally, we have also discovered an intricate interplay between CTCF-mediated promoter-upstream looping and CTCF-mediated RNA polymerase II pause at exon, regulating alternative splicing outcome of EMT-associated gene. My talk will give you an overview of the intricate connections between hypoxia and the dynamic landscape of gene regulation and alternative splicing in cancer.

Session -12 Genomics Start ups and MSME

21. Tumor and Tumor Micro Environment : How to monitor them? Theracues experience in spatial transcriptomics with case studies

Speaker: Gopalakrishnan Ramaswamy

Abstract

Tumor and Tumor Micro Environment play an important role in the progress or inhibition of tumor ultimately metastasis and cancer progression. Understanding at cellular level in their environment is key for therepeutic decision. In this regard, recent introduction of spatial transcriptomics and spatial protein profiling are playing key roles and theracues is in leading front to facilitate this.

Theracues is a startup company, coded for signal for targeted therapy and facilitating translational research in India and globally. Being the initial innovator in this area theracues team has learnt the details of technology and their useful applications and pitfalls. Theracues team has collaborated with key researchers in this area. Speaker will provide insights regarding these studies.

Session -12 Genomics Start ups and MSME

22. Compressed Next-Gen Sequencing

Speaker: Dr Manoj Gopalakrishnan, Founder & CEO Algorithmic Biologics

Abstract

Despite significant reductions in the cost of sequencing the genome over the past 25 years, genomic testing remains inaccessible for most people. I will describe "Compressed Sequencing," a novel software solution developed in-house at Algorithmic Biologics that saves up to 90% in library preparation costs, thus dramatically bringing down the cost of genomics sequencing. Our pioneering solution brings down costs while allowing labs to keep using tried-and-tested reagents, kits, machines, and bioinformatics pipelines.

Posters

Cancers and inherited disorders

	Pharmacogenetic profiling of ADME-related genes in an Indian cohort predicts variability in drug response and implications for personalized medicine	
A17	*Mudassar Ali, Priyanka Majumder, Pallavi Rao T, Swasti Raychaudhuri & Koyeli Mapa* Elucidation of mitochondrial sub-compartment specific proteotoxic stress response in mammalian cells	67

Infectious diseases

Microbial Genomics and Metagenomics

Ecological and evolutionary genomics

Plant biology

A58	*S. Srivignesh, Poornima Reddy, Nihar Bachan Das, Raja Mugasimangalam & Rajashekhar Ballari* First telomere to telomere Indian rice genome	113

Computational and statistical genomics

Methods in genomics

Poster Abstracts

A01

UNRAVELING HIGH-RESOLUTION CHROMOSOMAL LANDSCAPES TO INVESTIGATE GENOTYPE-PHENOTYPE ASSOCIATIONS IN GIRLS MANIFESTING A CLINICAL PHENOTYPE OF SHORT STATURE THROUGH CHROMOSOMAL MICROARRAY ANALYSIS

Abhay Kumar Yadav[1], Ankur Singh[2] & Royana Singh[1]*

[1]Department of Anatomy, Institute of Medical Science, Banaras Hindu University, Varanasi
[2]Department of Paediatrics , Institute of Medical Science, Banaras Hindu University Varanasi
*Corresponding author: royanasingh@bhu.ac.in

BACKGROUND: Short stature is defined as a child's height being less than the 3rd percentile or 2 SD (standard deviation) below the chronological age.The height assessment can be done based on growth velocity. The height trait is mostly associated with a genetic factor, but various aspects come from a combination of genetic and environmental involvement. Eighty percent of the cause of short stature is under the influence of genetic control, and genome-wide association has revealed that 20% of the variation in adult human height is due to the effect of 697 familiar variants assembled in 423 loci .Many copies number of variations(CNV) involve the development of short stature.

METHODS: A short-stature patient sample was collected from the Department of Pediatrics, IMS BHU. The selected sample of patients were based on respective parameters like physical and clinical features such as Height, Weight, BMI ,Z-score were calculated basis of WHO Peditools Growth CDC 2-20 year whose heights were less than 3rd percentile or 2 standard deviation below. Blood samples were collected from Pediatrics Department of IMS BHU in EDTA and heparinized vial for DNA and karyotype. DNA was isolated from Invitrogen kit (catalog no- K1820-01). The quantification of DNA was done using a Nano-spectrophotometer and an Invitrogen qubit. A quality check of DNA was done using agarose-gel electrophoresis. Array-based comparative genomic hybridization was done using the Agilent protocol.

RESULTS: The copy number variation found in short stature was mostly on the autosome chromosome rather than the sex chromosome. CNV was detected on chromosomes 14 and X in short stature girls who have karyotype-turner syndrome, where non-coding copy number variation was found. Mostly CNV found turner short stature gain at 14q32.33 and loss at Xq11.2. Forty samples of short-stature girl's karyotypes were done; 50% were short-stature turners and 50% were normal.

CONCLUSION: High-resolution karyotype analysis reveals that the clinical phenotype of Short stature Turner Syndrome may be linked to copy number variation (CNV) defects in autosomes, particularly on chromosome 14 or the X chromosome, or both. Cases of CNV defects on chromosome 14, with or without accompanying X-chromosomal defects, exhibit clinical manifestations resembling short stature turner syndrome. These conditions share a common genomic network, emphasizing the need for further investigations in understanding their genetic underpinnings.

A02

UNRAVELING CELLULAR AND MOLECULAR MECHANISMS IN VITILIGO-DIFFERENTIATED MELANOCYTES

Sathvik Upadhya[1], Melisa J. Andrade[1], Vaibhav Shukla[1], Keshava Prasad[1], Raghavendra Rao[2] & Kapaettu Satyamoorthy[1,3*]

[1]Department of Cell and Molecular Biology, Manipal School of Life Sciences, Manipal Academy of Higher Education, Manipal, Karnataka, India
[2]Department of Dermatology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, Karnataka, India
[3]Shri Dharmasthala Manjunatheshwara (SDM) University, Sattur, Dharwad, Karnataka, India
*Corresponding author: ksatyamoorthy@yahoo.com

SUMMARY: In this study, we aimed to use induced pluripotent stem cells (iPSCs) as a tool to explore the disease pathogenesis of vitiligo. iPSCs generated from fibroblasts isolated from vitiligo and normal subjects were differentiated into patient-specific melanocytes. These melanocytes were then characterized using melanocyte-specific markers and assays. Whole exome and transcriptome sequencing was performed to identify gene variants and differentially expressed genes (DEGs) associated with vitiligo, which may shed light on the vitiligo disease pathogenesis.

BACKGROUND: Vitiligo is a chronic, acquired hypopigmentary condition characterized by the loss of functioning melanocytes, resulting in depigmented regions of skin. The exact cause of the disease pathogenesis remains elusive, limiting the efficacy of existing treatments. Melanocytes are severely affected in the disease condition; this increases the complexity of isolation of the melanocytes from the lesional areas to study the disease pathogenesis. iPSCs can be a promising alternative source of generating functional patient-specific melanocytes to study and understand disease pathogenesis.

OBJECTIVES: To generate patient-specific iPSCs (Vitiligo and normal) and characterize their differentiation into melanocyte cells. Further, understanding vitiligo disease pathogenesis through cellular and molecular analysis.

METHODS: Fibroblasts were isolated from the vitiligo and healthy normal subjects and iPSCs were generated using reprogramming factors. Functional characterization of these iPSCs were performed using pluripotent markers. Further, whole exome and transcriptome sequencing of the patient-specific iPSCs was performed. These iPSCs were then differentiated into melanocyte-like cells and characterized using melanocyte-specific markers such as TYR, TRP1, TRP2, and gp100. Functional characterization assays such as tyrosinase activity assay, melanin pigment quantification assay, Fontana-Masson staining assay, melanocyte dendrite measurement assay, keratinocyte-melanocyte interaction assay, and melanosome transfer assay were performed for these patient-specific melanocytes. To understand the disease pathogenesis at the molecular level, transcriptome analysis was performed from the RNA isolated from the normal and vitiligo-differentiated melanocytes.

RESULTS: Patient-specific melanocytes differentiated from iPSCs expressed melanocyte-specific markers. The differentiation of iPSCs into functional melanocytes was further validated by the positive L-DOPA and Fontana-Masson staining. Extracellular melanin pigment quantification assay in the differentiated melanocytes displayed reduced melanin content in vitiligo-derived melanocytes compared to the normal sample. Further, melanocytes derived from vitiligo demonstrated shorter dendritic projections and exhibited reduced dye transfer and interaction with primary keratinocytes compared to melanocytes derived from normal subjects. Melanosome transfer assay demonstrated defects in the transfer of melanosomes in vitiligo-derived melanocytes. The whole exome sequencing analysis of patient-specific iPSCs identified genetic variations in the genes associated with vitiligo. Transcriptome sequencing analysis of both iPSCs and differentiated melanocytes identified many

DEGs in vitiligo samples that are associated with melanogenesis, proliferation, migration, and survival.

CONCLUSION: Our findings suggest the potential impairments in the vitiligo-derived melanocyte functions that may provide further insights into vitiligo disease pathogenesis.

A03

EXPLORING THE ROLE OF NUCLEAR AND MITOCHONDRIAL R-LOOP DYNAMICS IN DNA DAMAGE-INDUCED SENESCENCE IN BREAST CANCER

Dinesh Reghunathan[1], Pradyumna Jayaram[1] & Kapaettu Satyamoorthy[1,2]*

[1]Department of Cellular and Molecular Biology, Manipal Academy of Higher Education, Manipal, Karnataka, India
[2]Shri Dharmasthala Manjunatheshwara (SDM) University Sattur, Dharwad, Karnataka, India
*Corresponding author: ksatyamoorthy@yahoo.com
[#]Presenting author: dineshgene89@gmail.com

SUMMARY: The study investigates R-loop accumulation in senescence induction via doxorubicin treatment in MCF7 cells. DRIPc-seq analysis was performed to identify the distribution of R-loop peaks in control and senescence-induced MCF7 cells. Significant accumulation of R-loop peaks was observed in senescence-induced groups compared to the control.

BACKGROUND: R-loops are three-stranded nucleic structures with a stable RNA:DNA hybrid and displaced ssDNA. Replicative senescence refers to a stable cell cycle arrest characterized by hallmark events such as irreversible cell cycle arrest, senescence-associated phenotypes (SASPs), altered metabolism and macromolecular damage. Doxorubicin intercalates into DNA, inhibits TOPII-mediated DNA repair and damages DNA. Unscheduled and unresolved R-loop accumulation disrupts transcription machinery and replication, leading to transcription replication collisions (TRCs) and genome instability. TRCs are a major cause of R-loop-induced replication stress and associated DNA damage. Mitochondria DNA (mtDNA) is more prone to damage than nuclear DNA. Mitochondrial dysfunctions are considered a hallmark of cellular senescence. However, limited evidence suggests the role of the R-loop in regulating senescence.

OBJECTIVES: To identify whether DNA damage-based senescence induction leads to R-loop accumulation in normal and tumour cells and elucidate the role of senescence-associated R-loops in breast cancer.

METHODS: Senescence-induction in MCF7 cells using 50nM doxorubicin, and senescence-associated β-gal assay was used to assess cellular senescence in 24, 48 and 72h. DNA replication in cells post DNA damage-induced senescence was performed via EdU-based replication assay. To map R-loops in the genome, DNA-RNA immunoprecipitation (DRIP) was performed using S9.6mAb, which specifically targets RNA:DNA hybrids *in vivo* and *in vitro*. DRIPc library was sequenced using an Ion Proton semiconductor sequencer with a 200bp read chemistry. The DRIPc-seq reads obtained were quality checked using FASTQC and R-loop peaks were mapped and annotated using RLbase and open-source pipelines. The obtained reads were also mapped to the 16.5 Kb mitochondrial genome, and R-loop abundant and poor regions were mapped to the heavy and light strands in the mitochondrial genome.

RESULTS: We observed a significant difference in the percentage of β-gal positive cells in doxorubicin-treated MCF7 cells compared to control after 48 hours of treatment. Immunofluorescence assay to study R-loop localization was performed using S9.6mAb, and it was observed that S9.6mAb signals were localized in the nucleus. Two to nine million reads were obtained with a mean read length of 100bp for DRIPc-seq and genomic distribution analysis of data identified 17.3% of R-loops were shown to reside in promoter proximal region in senescence-induced MCF7 and 14% R-loops were located in promoter region of control. We observed high signal intensity for R-loop peaks in TSS regions for senescence-induced MCF7 compared to control. DRIPc-seq data mapped to the mitochondrial genome revealed a higher abundance of peak regions in senescence-induced MCF7 compared to control, and peak abundance was identified in the H-strand of the mitochondrial genome.

CONCLUSION: Our study unveils a significant association between DNA-damage-induced senescence and R-loop accumulation leading to alteration of mitochondrial dynamics and senescence in breast cancer cells.

Keywords: R-loop, Senescence, DNA damage, DRIPc-seq, Replication, MCF7

Acknowledgement: The authors thank Dr. TMA Pai fellowship and ICMR (2021-13215) for project funding and fellowship, MSLS-MAHE and TIFAC - Core, Government of India for infrastructure.

A04

EMPHASIZING THE POTENTIAL OF CHROMOSOMAL MICROARRAY TO ADVANCE OUR UNDERSTANDING OF THE GENETIC FACTORS CONTRIBUTING TO RECURRENT PREGNANCY LOSS AND IMPROVE CLINICAL OUTCOMES

Shivani Misra[1#], Chetan Sahni[1], Sangeeta Rai[2], Ashish[3] & Royana Singh[1]*

[1]Department of Anatomy, Institute of Medical Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India
[2]Department of Obstetrics and Gynecology, Institute of Medical Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India
[3]Multidisciplinary Research Unit, ICMR-DHR, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India
*Corresponding author: royanasingh@bhu.ac.in
[#]Presenting author: Shivanimishra147@gmail.com; 7052875119

BACKGROUND: It is estimated that 15–20% of pregnancies that are clinically recognized end in miscarriage, and around 1% of couples experience recurrent pregnancy loss (at least two occurrences). The primary cause of early pregnancy loss, particularly in the first trimester, is the presence of a significant fetal chromosomal abnormality, contributing to over 50% of such losses. The predominant etiology involves whole-chromosome aneuploidies, representing the most common type of abnormalities. Fetal chromosomal anomalies are accountable for approximately 8–10% of instances involving intrauterine fetal demises beyond 20 weeks of gestation or stillbirths occurring in the second or third trimester.

OBJECTIVE: To assess various diagnostic approaches for array genetic analysis in cases of fetal loss across all trimesters at medical hospital S.S hospital, IMS BHU department of obstetrics and gynaecologicy OPD between 2022-2023.

METHODS: DNA isolation retained product of conceptus (RPOC) cases utilized the Salting out and PureLink Genomic DNA isolation kit. Molecular etiological analysis was conducted on DNA samples through the oligonucleotide microarray-CGH method (aCGH, 60 K ISCA design, Agilent, Germany). Hybridized probe correlations between case and reference DNAs were assessed using databases, specifically the Database of Genomic Variants Analysis, focusing on 54 functional genes associated with CNVs related to intrauterine fetal losses. Initially, Quantitative Fluorescent Polymerase Chain Reaction (QFPCR) testing was conducted followed by, array Comparative Genomic Hybridization (aCGH) was employed for the QFPCR-normal samples to analyze and identify potential copy number variations (CNVs), gene deletions, and/or duplications.

RESULTS: Among the 20 fetal samples analyzed, CNVs were detected in 16 (80%). Of these CNVs, 60% were duplications, and 40% were deletions. Evaluation revealed deletions in 11 (55%) of the 49 genes, duplications in 8 (40%), and both deletion and duplication in1 (5%) cases. While CNVs were identified in autosomal chromosomes (chromosomes 1, 2, 3, 7, 8, 10, 11, 12, 13, 18, 20, 22), the X and Y chromosome exhibited the highest frequency. Notably, CNVs associated with CFHR3, UGT2B17, LINC00280, PSG1, PSG2, PSG3, PSG4, PSG5, PSG8, PSG11, CD24, MSR1, SSX5, and DUSP22 genes were found to be more prevalent in terms of fetal loss etiology in this study.

CONCLUSION: These ongoing studies have the potential to contribute significantly to scientific clinical consultation, offering enhanced diagnostic precision and prognostic insights into subsequent pregnancies associated with embryonic cell division, organogenesis, and tissue differentiation. The findings can aid healthcare professionals in increasing awareness of chromosome-related issues linked to miscarriage and establishing a theoretical foundation for rational treatment approaches.

A05

COMPREHENSIVE PAN-CANCER INVESTIGATION OF ZNF471: INSIGHTS, PATTERNS, AND THERAPEUTIC IMPLICATIONS

Poonam Bhandari[1#], Bobby Paul[2] & Kapaettu Satyamoorthy[1,3]*

[1]Department of Cellular and Molecular Biology, Manipal School of Life Sciences, Manipal Academy of Higher Education, Karnataka, India
[2] Department of Bioinformatics, Manipal School of Life Sciences, Manipal Academy of Higher Education, Karnataka, India
[3] Shri Dharmasthala Manjunatheshwara (SDM) University Sattur, Dharwad, Karnataka, India
*Corresponding author: ksatyamoorthy@yahoo.com
[#]Presenting author: poonam121bhandari@gmail.com

SUMMARY: This study unveils the multifaceted role of ZNF471 across diverse cancers. The analysis highlights the downregulation of ZNF471 across multiple cancers, impacting patient prognosis. Genetic variations vary significantly among cancers, notably in cervical (20%) and prostate (0.73%) cancers. ZNF471 expression correlates with immune infiltration suggesting its potential as a therapeutic target in cancer treatment.

BACKGROUND: Zinc finger proteins (ZFPs) constitute the most extensive family of transcription factors within the human genome, marked by diverse DNA binding domains (DBDs) which can identify specific target DNA sequences, thus engineered for various biomedical applications. ZFPs as prevalent DBDs have been widely utilized in genome engineering, regulation of gene expression, and diagnostic procedures. Various findings have revealed the importance of ZFPs in cancer as tumor suppressors as well as oncogenes. Thus, pan-cancer investigation of ZNF471 is a comprehensive study conducted to understand the role of ZNF471 in various types of cancers.

OBJECTIVES: To provide insights into the expression patterns, genetic alterations, and functional roles of ZNF471 across different cancer types, and evaluate its potential as a prognostic marker or therapeutic target.

METHODS: We utilized various public databases such as TIMER, GEPIA, The Cancer Genome Atlas (TCGA), Genotype-Tissue Expression (GTEx), and Human Protein Atlas (HPA) to investigate the expression of ZNF471 at both mRNA and protein levels across diverse cancer types. Additionally, ZNF471 expression was assessed in different cell lines through qRT-PCR analysis. The impact of ZNF471 on the overall survival (OS) of patients with tumors was determined using Kaplan–Meier plotter. Furthermore, alterations in the ZNF471 gene were investigated across different cancers via cBioPortal, assessing the impact of mutations on the OS of patients. To investigate further, we explored the correlation between ZNF471 expression and clinical characteristics, immune infiltration, and DNA methylation by utilizing TCGA, TIMER, and SMART tools. Finally, Gene Set Enrichment Analysis (GSEA) was employed to unravel the molecular mechanisms associated with ZNF471.

RESULTS: Bioinformatic analysis confirmed that ZNF471 expression is downregulated in most cancers at both mRNA and protein levels. ZNF471 low-expression group showed a poorer prognosis and shorter OS in certain cancers. Among the cancers studied, cervical cancer demonstrated the highest occurrence rate of genetic variation at 20%, while prostate cancer exhibited the lowest rate at 0.73%. A correlation was observed between the OS and the mutated group in breast and lung cancer. Furthermore, an association between ZNF471 and immune infiltration was observed, revealing both strong positive and negative relationships with various immune cells in specific cancer types. Additionally, through GSEA, it was identified that proteins interacting with ZNF471 were enriched in categories associated with metabolic processes, membrane-enclosed lumens, and nucleic acid binding within the domains of biological processes, cellular components, and molecular functions, respectively.

CONCLUSION: This study revealed the significance of a holistic understanding of the involvement of ZNF471 in cancer, providing a foundation for future investigations aimed at elucidating its precise mechanisms and exploiting its therapeutic potential in the oncology landscape.

A06

SOMATIC MUTATION PROFILING AND CLINICAL PATHOGENIC RELEVANCE OF VARIANTS AT GENE LEVEL IN GALLBLADDER CARCINOMA

Supriya Awasthi[1], Dibyabhaba Pradhan[2], Nihar Ranjan Dash[3], Sunil Kumar[4], Prasenjit Das[5] & Pranay Tanwar[1]

[1]Laboratory Oncology Unit, Dr. B.R.A. Institute-Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India.
[2]Scientist II (Bioinformatics), Centralized Core Research Facility, All India Institute of Medical Sciences, New Delhi, India.
[3]Department of GI Surgery, All India Institute of Medical Sciences, New Delhi India
[4]Department of Surgical Oncology, Dr. B.R.A. Institute-Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India
[5]Department of Pathology, All India Institute of Medical Sciences New Delhi, India.
*Corresponding author: supriya.awasthy@gmail.com; 8826004940.

BACKGROUND: Gallbladder cancer (GBC) is a rare form of cancer and have a fatal malignancy with median survival of less than year. GBC incidence is significantly higher in certain regions of India as compare to western countries. Due to the absence of noticeable symptoms in the initial stages, gallbladder cancer often progresses undetected, leading to a worse prognosis at later stages. Here, we aimed to study the mutational profile along with significance of variants at therapeutic level using whole exome sequencing.

METHODS: Total 30 patients were enrolled for study. DNA was extracted from 30 Gallbladder tumor tissues and matched blood as normal samples. Somatic variant calling and annotation of the identified variants at oncogenic level was done using in-house developed pipeline. Tumor mutational burden was studied using TCGAmutations package, Mutational signature analysis and oncogenic signaling pathways were analyzed using maftools R package. Clinical impact of the genes associated with somatic variants was studied on basis of cancer driver activity using MutSigCV. Variants were classified using ACMG 2015 guidelines, oncogenic relevance and their therapeutic implications was confirmed on the basis of oncogenic effects of cancer genes using OncoKB.

RESULTS: This study identified total 5562 somatic mutations in 4277 genes. Missense mutations were predominantly occurred (84.43%) followed by Frameshift deletions (5.54%) and Nonsense mutations (4.8%). Overall mutation rate was low (median 1.58 Mb). Mutational signatures were matched to COSMIC Database and revealed the similarity with COSMIC (2,13-corr-0.811), (1,6-corr-0.858), (4-corr-0.583) among 28 samples. RTK-RAS (20 samples) and WNT pathways (21 samples) were identified as major oncogenic signaling pathways in our study. Total 58 variants were identified as pathogenic and 59 as likely pathogenic, among 96 genes comprising of 117 mutations. Considering the clinical impact on the basis pathogenicity, TP53 variants (30%) were mutated in 9 samples, SMAD4(23%), CTNNB1(23%) in 7 samples and ERBB2(20%) in 6 samples. These genes have cancer driver as well as pathogenic and likely pathogenic classification and their mutation is labelled as oncogenic and likely oncogenic in OncoKB.

DISCUSSION: We demonstrated the mutational profile of GBC patients, and identified the most frequently mutated genes with pathogenic and oncogenic relevance. GBC mutational burden was compared with 33 TCGA cohorts and it was found similar level of LIHC and comparatively higher than that of cholangiocarcinoma. Among the Mutational signatures, Cosmic 2,13 usually observed in subset of liver cancers and samples with known exposure to aflatoxin. Cosmic 1,6 correlates with age of diagnosis and in our study median age of the patients were 50. Cosmic 4 is usually with smoking and exposure to tobacco carcinogens which also supported the fact as maximum patients had tobacco addictive history. Among these four genes, their variants have disruptive to nondisruptive impact on

protein structures and mutational effect associated with loss of function (TP53, SMAD4) and gain of function (CTNNB1, ERBB2).

CONCLUSION: Our study provides the clinically relevant interpretation of somatic mutations at gene level that proved to be helpful in therapeutic interventions.

A07

IDENTIFICATION OF NATURALLY OCCURRING SUBSTITUTIONS IN THE HEPATITIS C VIRUS RDRP GENE

Raina Das[#], Moumita Majumdar, Supradip Dutta, Sagnik Bakshi, Shreyasi Nath, Anwesha Ghosh, Shanta Dutta & Provash Chandra Sadhukhan*

ICMR- National Institute of Cholera and Enteric Diseases, Kolkata, West Bengal, India.
*Corresponding author: sadhukhan.pc@icmr.gov.in
[#]Presenting author: raina2201@gmail.com

BACKGROUND: Hepatitis C Virus (HCV) is a common transfusion-transmitted infection (TTI) and a leading cause of liver cirrhosis and hepatocellular carcinoma. Its error-prone replication mechanism has led to the development of 8 genotypes and 64 subtypes. Direct-acting antivirals (DAAs) are the gold standard for HCV treatment. The HCV NS5B gene encoding the RNA-dependent RNA polymerase (RdRp) is one such target of a potent DAA- Sofosbuvir. The generation of viral quasi-species aids immune escape and the pre-existence of resistance-associated substitutions (RASs) interferes with DAAs therapy. Hence it is imperative to understand its transmission dynamics and evolutionary history. Moreover, identifying naturally occurring resistance-associated substitutions and understanding their impact on therapy outcomes would help benefit future therapeutic interventions.

OBJECTIVES: This study aims to understand the viral evolution of HCV Gen-3a isolates in West Bengal, to identify naturally occurring substitutions in the HCV RdRp (NS5B) gene, and further, conduct in-silico prediction studies for its effect on drug interaction.

METHODS: Viral RNA was extracted from HCV Gen 3a infected plasma samples and cDNA synthesis was performed. PCR amplification of the full-length NS5B gene (1773bp) was done with a high-fidelity polymerase. The amplicons were purified and Sanger sequencing was performed. MEGA X software was used for a phylogenetic analysis using the Maximum Likelihood method. Bayesian analysis mediated phylogeographic diffusion was estimated with Maximum Clade Credibility (MCC) trees using the BEAST package and SpreadD3. To understand selection pressures and estimate non-synonymous to synonymous substitution distributions (dN/dS), the dataset was analyzed using the Datamonkey webserver. Naturally occurring and Resistance-associated substitutions (RASs) in the NS5B gene were identified using the geno2pheno online resistance analysis tool. Finally, homology modeling and molecular docking were used as a prediction tool to compare interactions of wild-type and mutant forms of NS5B protein with Sofosbuvir.

RESULTS: The phylogeographic analysis revealed the migration of Gen-3a from India to European and South American countries back to circulation in India. The dN/DS calculations revealed mostly purifying/negative selection acting upon NS5B whereas few diversifying/positive selections were observed. Mutational analysis revealed the circulation of a few substitutions such as – N307G, D330E, A356P, R379K, and N535K as compared to the reference strain. Resistance – associated substitutions such as A150V, L159F (Sofosbuvir resistance), K206E, C316F (Sofosbuvir resistance), Y448D, P495L and G554D (atypical substitutions) were also noted. Molecular docking showed altered binding of mutant forms, incorporating fewer active site residues leading to weak interactions with Sofosbuvir.

DISCUSSIONS AND CONCLUSIONS: This study helps understand the evolutionary lineage of HCV isolates from the Eastern part of India based on its NS5B gene. Moreover, the identification of various baseline mutations and prediction of their effects on drug interaction enables to provide insight into the imminent risk they pose for HCV therapy.

A08

CANCER-ASSOCIATED MUTATION AT GLYCINE 400 DISRUPTS TIP60 ASSOCIATED FUNCTIONS

Himanshu Gupta[1], Ashutosh Singh[2] & Ashish Gupta[1]*

[1]Epigenetics and Human Disease Laboratory, Centre of Excellence in Epigenetics, Department of Life Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, Uttar Pradesh, India
[2]Translational Bioinformatics and Computational Genomics Research Lab, Department of Life Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, Uttar Pradesh, India
*Corresponding author: ag315@snu.edu.in

Mutations can alter protein structure, resulting in changes in protein function. TIP60, a recognized tumor suppressor and epigenetic regulator, has lysine acetyltransferase (KAT) activity, permitting chromatin to be accessed by acetylating histone. TIP60 participates in a wide range of biological functions, including DNA repair, apoptosis, cell cycle control, and cell migration. Even though the impact of cancer-associated patient mutations on TIP60 structure and activities is still unclear. In the present study, we dissected the consequences of mutations on TIP60 structure and its functions. We identified 94 cancer-associated patient mutations using the CBioportal database in the KAT5 gene. Further, the structural impact of selected mutations was checked on protein using Molecular Dynamic (MD) Simulation. The 500 ns MD Simulation results suggest mutation-induced structural changes in protein structure at the atomic level. Further, in silico docking of proteins and acetyl coenzyme A shows changes in docking energy in the case of mutations and no docked pose in the G400W mutation. Further, live cell imaging data shows altered protein localization in the G400W mutation. Moreover, the G400W mutation was functionally inactive and lost the TIP60-associated DNA damage repair functions. Also, G400W mutations alter TIP60 phase separation. Overall, our research opens new avenues for understanding the molecular foundations of cancer formation and development, perhaps opening the door to therapeutic approaches.

Keywords: TIP60, point mutation, tumor suppressor, DNA damage repair

A09

DECIPHERING REGORAFENIB-INDUCED RESISTANCE IN COLORECTAL CANCER: INSIGHTS FROM EXOME AND METHYLOME PROFILING

Fayyaz Rasool, Rohit Kumar Verma, Ashutosh Singh, Puli Chandramouli Reddy & Sri Krishna Jayadev M*

Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
*Corresponding author: jayadevmsk@snu.edu.in

BACKGROUND: Colorectal cancer, the most prevalent cancer globally, poses a significant challenge due to developing resistance to existing anticancer drugs. Overcoming drug resistance is a formidable task, necessitating innovative approaches. Regorafenib, an oral multi-kinase inhibitor, is employed in treating advanced stages of colorectal cancer (CRC). However, early clinicalimplications of resistance emerge with prolonged Regorafenib use, and the underlying mechanisms remain unexplored. This study aims to uncover the genetic and epigenetic alterations employed by Regorafenib-resistant CRCs (Reg-R-HCT116) to ensure their survival and proliferation. Exome sequencing captures the genetic information encoded in all the protein-coding regions of a gene, while methylome reveals distinctive patterns of DNA methylation variations. Over the last two decades, research has solidified mTOR's pivotal role in governing crucial cellular processes, from cell growth, and protein synthesis to autophagy. Rapamycin (mTOR inhibitor) and Torin 1 (ATP competitive mTOR kinase inhibitor) were applied to HCT116 and Reg-R- HCT116 cells to check the cytotoxicity and pathways involved.

OBJECTIVES: Evaluating Exome and methylome profiling of drug resistance cells.

METHODS: Regorafenib-resistant (Reg-R-HCT116) cells were established by continuous treatment of HCT116 cells with regorafenib over several passages till they attained a two-fold tolerance for the drug. The DNA was isolated and exome and methylome sequencing was performed by Agilent Sure Select 150x2 paired-end Illumina platform. Bioinformatics pipelines were used to extract readable data, revealing common pathway alterations. MD simulation studies were performed to study the impact of mutations on protein structure and function. IGV plots depicted hyper/hypomethylation. MTT assay was performed for cytotoxicity of Rapamycin and Torin 1. Western blot analysis was done to check the impact of these mutations on their associated pathways.

RESULTS: PI3K-AKT-mTOR, Ras/Raf, and lipid metabolism-associated signaling pathways are found to be the most prominent pathways. EIF2 signaling downstream of mTOR (mTOR1) is highly upregulated in proteome data as well (data not shown). Mutations in Raptor are observed which lead to its stabilization in the MD simulation study. Rapamycin fails to inhibit Raptor in Reg-R-HCT116, unlike in HCT116 control. Torin 1 effectively inhibits mTOR kinase activity, causing cell death in both control and resistant cell lines at similar concentrations. ACACA serves as a crucial enzyme in the synthesis of fatty acids. Deleterious mutations are observed in both ACACA and FASN which in turn are reported to be regulated by mTOR. The role of ACACA is to hinder excessive fatty acid synthesis, a vital process for meeting the high energy requirements of rapidly dividing cancer cells. Additionally, analysis of methylome data reveals hypomethylation in ACACA. Mutant NF-1 shows loss in secondary structure and loss of function Mutation in NF1 is known to hyperactivate Ras which subsequently activates the PI3K-AKT-mTOR pathway.

CONCLUSION: The HCT116 cells employ the PI3K-Akt-mTOR pathways as a mechanism to develop resistance to regorafenib. The combined application of Torin 1 with regorafenib proves to be a potent approach against regorafenib-resistant HCT116 cells.

A10

CANONICAL VS. NON-CANONICAL 3'-END FORMATION – ROLE IN HYPERTROPHY GENE PROGRAM IN THE HEART

Diksha Singh* & Rakesh S. Laishram

Cardiovascular and Diabetes Biology Program, Rajiv Gandhi Centre for Biotechnology, Trivandrum, Kerala, India
*Corresponding author: dikshasingh@rgcb.res.in

All precursor mRNAs undergo processing at the 3'-end that endows stability and is required for efficient translation of the mRNA. 3'-end processing involves in two steps - cleavage at the 3'-untranslated region (UTR) followed by addition of a poly(A) tail on the cleaved RNA carried out by a multi-subunit protein complex called cleavage and polyadenylation (CPA) complex. CPA includes cleavage and polyadenylation specificity factor, CPSF (30, 73, 100, 160, FIP1 and Wdr33 subunits), Cleavage stimulatory factor, CstF (64, 77, and 50 subunits), Cleavage factors, CF (Im and IIm), poly(A) polymerase (PAP), poly A binding protein (PABPN), and Symplekin. While CPSF components Wdr33 and CPSF30 recognize the PA-site, CPSF73 acts as an endonuclease that cleaves the pre-mRNA. CFIm and IIm help in the assembly whereas symplekin act as a scaffold. PAPs then catalyze the polyadenylation reaction followed by binding of the PABPN that stabilizes the poly(A) tail. Two major nuclear PAPs, canonical PAPα/g and non-canonical Star-PAP polyadenylates nuclear mRNAs. While the two PAP complexes share most of the 3'-end processing factors, the non-canonical Star-PAP is dispensable of some of the of canonical factors but requires additional protein RBM10 for the complex assembly. However, the exact mechanism how Star- PAP or canonical PAP complex specifically selects targets mRNA is still undefined. Our immunoprecipitation and mass spectrometry sequencing data of two complexes reveal unique processing factors in both complexes with PAPα, CstF64, and WDR33 in canonical complex, and Star-PAP, RBM10, and PIPKIα in non-canonical complex. To understand functional significance of the two complexes, we carried out transcriptomic analysis of canonical components (WDR33 and PAPα) and non-canonical Star-PAP. Comparison of differentially expressed genes of WDR33 shows >80% overlapped with that of PAPα reinforcing its function in the canonical processing complex. Intriguingly, comparison with Star-PAP (overlapped genes between WDR33 and PAPα) shows <5% of the WDR33 regulated genes are common and >95% were different revealing distinctness of target mRNA selection. Functional pathway analysis of down regulated genes (on the depletion of WDR33 or Star-PAP) shows most gene are involved in cardiovascular disease and/or cancer. Strikingly, among the cardiovascular related genes, WDR33-controlled genes mostly encode pro-hypertrophy factors up regulated in cardiac hypertrophy while Star-PAP regulated genes shows anti-hypertrophy factors, down regulated in cardiac hypertrophy. To further understand the role of these complex components in the hypertrophy, we induce hypertrophy in animal heart (wistar rat) and cardio-myoblast cell (H9C2) by Isoproterenol treatment. Analysis of processing factors shows up regulation of the canonical components (WDR33, CstF64, CPSF73, and CPSF30) while the non-canonical components (Star-PAP, RBM10, or CPSF160) were down regulated. However, other factors such as PAPα, CPSF100, PABPN1, and symplekin are not affected by hypertrophy induction in both animal and cellular model. We also observed differential expression of respective target mRNAs and proteins in hypertrophy by qRT-PCR and Western analysis. Thus, our study reveals involvement of a balance of two complexes in controlling overall hypertrophy gene program in the heart.

A11

CORRELATION BETWEEN METHYLENETETRAHYDROFOLATE REDUCTASE GENE POLYMORPHISM (RS1801133) AND HOMOCYSTEINE LEVELS IN MALE INFERTILITY PATIENTS FROM THE EASTERN UTTAR PRADESH POPULATION

Pooja Dubey[1], Royana Singh[1], Sameer Trivedi[2] & Chetan Sahni[1]*

[1]Department of Anatomy, Institute of Medical Science, Banaras Hindu University, Varanasi
[2]Department of Urology, Institute of Medical Science, Banaras Hindu University, Varanasi
*Corresponding authors: chetansahni@bhu.ac.in
[#]Presenting authors: poojadubeybiochem@gmail.com, +91-9450570339

BACKGROUND: Infertility poses a widespread concern, affecting 15%-20% of couples, with almost half of the cases attributed to male factors. Despite diagnostic advancements, the etiology of male infertility remains elusive, primarily characterized by spermatogenic failure. Genetic abnormalities, accounting for 15%-30% of male factor infertility, involve Y chromosome microdeletions, translocations, chromosomal aberrations, and single-gene mutations. This study explores the potential role of the MTHFR C677T polymorphism, a key regulatory enzyme in folate metabolism, in male infertility. Folate deficiency and related hyperhomocysteinemia have been linked to various diseases, including infertility.

OBJECTIVES: This study aims to investigate the association between the MTHFR C677T polymorphism and homocysteine levels of male infertility patients while exploring the potential impact on sperm concentration in an Eastern Uttar Pradesh population.
Methodology: A cross-sectional study involving 148 male infertility cases (oligospermia and azoospermia patients) and 130 fertile males as controls. Genomic DNA was extracted, and MTHFR C677T genotypes were determined by real-time PCR. Homocysteine levels were measured using Elisa methods. Statistical analyses included Chi-Square tests, odds ratios, and Mann-Whitney U tests.

RESULTS: The chi-square analysis revealed a significant association between the MTHFR C677T polymorphism and the case and control groups ($\chi^2 = 8.07$, $p = 0.018$). In the case group, the allelic frequency for the C allele was 81.76%, while in the control group, it was 72.31%. The odds ratio for the C allele was 1.72 (95% CI: 0.97-3.03) with a p-value of 0.062. Conversely, the allelic frequency for the T allele in the patient group was 18.24%, compared to 27.69% in the control group. The concentrations of homocysteine level (µmol/mL) were analyzed according to genotype distribution, case group (n=148), the mean concentrations for CC, CT, and TT genotypes were 9.5±4.7, 21.8±10.5, and 33.8±14.4, respectively, with a significantly higher concentration in the TT genotype ($p < 0.05$). In the control group (n=130), the corresponding concentrations were 8.3±4.7, 12.4±4.5, and 20.4±7.3, respectively.

DISCUSSION: The findings underscore the relevance of the MTHFR C677T polymorphism in male infertility, shedding light on potential genetic factors influencing sperm quality. The association between allelic frequencies and disease risk suggests a need for further investigations to validate these preliminary findings.

CONCLUSION: The presence of the MTHFR C677T mutation, coupled with elevated Hcy levels, emerges as significant risk factors associated with the occurrence of oligospermia and azoospermia patients. The discerned allelic frequency patterns and concentration variations of homocystein among MTHFR C677T genotypes underscore the importance investigations to validate these findings.

A12

ASSOCIATION BETWEEN THE MTHFR (RS18033) GENE VARIATION AND SERUM TRACE ELEMENTS LEVELS (COPPER AND ZINC) IN INDIVIDUALS DIAGNOSED WITH NEURAL TUBE DEFECTS

Nitish Kumar Singh[1#], Sarita choudhary[2], Sangeeta Rai[3], and Royana Singh[1*]

[1]Department of Anatomy, Institute of Medical Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India
[2]Department of Pediatric Surgery, Institute of Medical Science, Banaras Hindu University Varanasi, Uttar Pradesh, India
[3]Department Of Obstetrics & Gynaecology, Institute of Medical Science, Banaras Hindu University Varanasi, Uttar Pradesh, India
*Corresponding author: royanasingh@bhu.ac.in
#Presenting authors: nitishbiochem@gmail.com, +91-9559052600

BACKGROUND: Neural Tube Defects (NTDs) are central nervous system disorders, often caused by folic acid deficiency and influenced by genetic factors. This study explores the association between prenatal copper (Cu) and zinc (Zn) exposure, 5,10-methylenetetrahydrofolate reductase (MTHFR) gene single nucleotide polymorphisms (SNPs), and NTD occurrence.

OBJECTIVES: Association between prenatal exposure to trace elements (Cu and Zn) and the single nucleotide polymorphism (SNP) of the MTHFR gene involved in folate metabolism pathways in neural tube defects in children and their mothers.

METHODS: A case-sectional study from the North Indian Population Cross-Sectional Study. Patients are recruited from maternal and pediatric care during antenatal check-ups in Obstetrics and Gynecology and Pediatric Surgery for Neural Tube Defects in the Outpatient Department (OPD) and Inpatient Department (IPD). The total of 331 participants (90 NTD cases, 88 healthy mothers, 85 NTD children, and 68 healthy children) assessed Cu and Zn concentrations and their associations. This study was carried out at the Cytogenetic Lab and Multidisciplinary Research Unit, Department of Anatomy, Institute of Medical Sciences, Banaras Hindu University, Varanasi, India. Genomic DNA was extracted, and real-time PCR determined genotypes based on the taqman probe used. Atomic absorption spectrophotometry measured trace elements. Statistical analyses included Chi-Square tests, odds ratios, and Mann-Whitney U tests.

RESULTS: Significant associations were found between MTHFR C677T genotypes and NTD risk in mothers ($p = 0.0491$) and children ($p = 0.0297$). Allelic frequency analysis indicated a T allele association with NTD risk in children ($p = 0.0107$). Recessive models showed significant associations in mothers ($p = 0.0169$) and children ($p = 0.1678$). Cu levels differed significantly between NTD cases and controls ($p < 0.0001$), with MTHFR genotypes influencing Cu levels. Zinc levels also varied significantly ($p < 0.0001$).

CONCLUSION: This study reveals complex associations between MTHFR C677T genotypes, trace element concentrations, and NTD risk in both mothers and children. This targeted approach allows healthcare providers to identify at-risk pregnancies early, enabling personalized interventions like folic acid supplementation and counseling to moderate neural tube defect (NTD) risk in future pregnancy.

A13

BEYOND BOUNDARIES: INVESTIGATING DISORDERS OF SEX DEVELOPMENT THROUGH ADVANCED GENETIC TECHNIQUES

Manpreet Kaur*

Banaras Hindu University, Varanasi, Uttar Pradesh, India
*Corresponding author: mann199419@gmail.com

SUMMARY: Disorders of sex development (DSDs) are congenital medical conditions that remain unassociated at chromosomal, gonadal and phenotypical level resulting in ambiguous genitalia and require a range of tests for diagnosis. These conditions can be caused by mutations in genes and hormones involved in sexual development. A broad spectrum of tests, including endocrinological tests, radiological images, and genetic tests, are required for an overall diagnosis of DSDs.

BACKGROUND: DSD pose significant challenges to individuals and their families, impacting not only physical health but also psychosocial well-being. To unravel the complexities of DSD, this study employs a multidisciplinary approach, integrating molecular genetics and cytogenetic techniques to explore the underlying causes.

OBJECTIVES: (1) To characterize chromosomal abnormalities in individuals with DSD through karyotyping. (2) To assess the presence of common chromosomal anomalies using QF-PCR. (3) To correlate genetic findings with clinical and hormonal manifestations in individuals with DSD.

METHODS: Individuals with DSD will undergo through genetic analysis such as karyotyping being the initial test for structural anomalies followed with QF-PCR for common numerical anomalies. Other comprehensive clinical and hormonal assessments are also done to complement genetic analysis.

RESULTS: Preliminary findings have revealed a spectrum of chromosomal abnormalities, including numeric and structural anomalies. Other genetic findings and clinical presentations are under investigation.

DISCUSSION: The integration of these techniques has enabled a holistic exploration of DSD, shedding light on the underlying genetic landscape. These findings contribute to the refinement of diagnostic and therapeutic strategies for individuals with DSD, emphasizing the importance of personalized approaches.

CONCLUSIONSS: To unravel the complexities of DSD, this study represents a significant step by employing comprehensive array of genetic techniques. The insights gained will not only enhance our understanding on the genetic basis of DSD but also pave the way for improved clinical management and counseling for affected individuals and their families.

A14

PROTEIN-PROTEIN NETWORK ANALYSIS BASED DISCOVERY OF PROGNOSTIC HUB BIOMARKERS LINKING CANCER RISK IN TYPE 2 DIABETIC PATIENTS

Rajveer Singh Shekhawat[1]*, Harshita Kasera[1], Pankaj Yadav[1, 2] & Priyanka Singh[1]

[1]Department of Bioscience and Bioengineering, Indian Institute of Technology, Jodhpur, Rajasthan, India
[2]School of Artificial Intelligence and Data Science, Indian Institute of Technology, Jodhpur, Rajasthan, India
*Corresponding author: shekhawat.1@iitj.ac.in

BACKGROUND: Type 2 diabetes mellitus (T2DM) and cancer are highly prevalent diseases imposing major health burden globally. Despite the availability of extensive evidence from epidemiological and meta-analysis that links cancer risk to T2DM, a systematic study of the shared genetic markers possibly predisposing this risk in T2DM patients is lacking for the common cancer types, namely pancreatic (PC), liver (LC), and breast (BC) cancer.

OBJECTIVES: A genetic level causal association linking T2DM to these cancers involving identification of important linked genes and biological processes.

METHODS: A gene expression analysis was performed to identify predominant differential expressed genes (DEGs) from the peripheral blood mononuclear cell (PBMC) samples of T2DM patients, posing a risk towards three common cancer types (PC, LC, and BC). The functional enrichment analysis was performed to get an understanding of enriched biological processes and pathways. The protein–protein interaction (PPI) network was retrieved for common hub genes between T2DM and the three cancer types. Further, validation of the identified prognostic hub biomarkers was done through survival analysis.

RESULT: The functional enrichment analysis indicated the involvement of gene expression, cell transport, and oxidation pathways. The protein–protein interaction (PPI) network provided common hub genes between T2DM and the three cancer typesA genetic association was identified between T2DM and three common cancer types, i.e., PC, LC, and BC. In total, the no. of common prognostic hub biomarker were 7, 6, and 5 hub genes showing a correlation between T2DM and the three cancers, i.e., PC, LC, and BC, respectively. We identified TGFB1 as a common hub gene between T2DM and PC/LC, significantly affecting survival in cancer patients.

DISCUSSION AND CONCLUSIONS: The analysis provides insight into biological and molecular events that could link T2DM with three common cancer types (PC, LC, and BC). Further, the identified genetic markers hold the potential to predict the chances of cancer onset in T2DM patients. Notably, such markers in T2DM patient PBMC samples predisposing to increased cancer risk could help diagnosis at an early stage and provide benefits for developing personalized therapeutic strategies.

A15

CLINICAL EXOME SEQUENCING (CARRIER SCREENING) IDENTIFIES DE NOVO SPLICING VARIANT IN INPPL1 IN A SPORADIC CASE OF OPSISMODYSPLASIA

Ashish Ashish[1], Shivani Mishra[2], Royana Singh[2] & Sangeeta Rai[3]*

[1]Multidisciplinary Research Unit- Department of Anatomy, Institute of Medical Sciences- Banaras Hindu University (IMS-BHU), Varanasi, Uttar Pradesh, India
[2]Anatomy, Institute of Medical Sciences (IMS) Banaras Hindu University (BHU) Varanasi, Uttar Pradesh, India
[3]Obstetrics and Gynaecology, Institute of Medical Sciences (IMS) Banaras Hindu University (BHU), Varanasi, Uttar Pradesh, India
*Corresponding author: drsangeeta1977@yahoo.co.in ; +91-8840622942

BACKGROUND: Opsismodysplasia is a rare skeletal dysplasia that interferes with bone maturation and micromelia (a bone disease that obstructs bone development).

CASE DESCRIPTION: The parents of the patients married non consanguineously with a history of two previous instances of infant mortality. Both children exhibited global developmental delay, microcephaly, bilateral congenital talipes equinovarus, and overlapping fingers. Magnetic Resonance Imaging (MRI) results for both children indicated generalized cerebral atrophy, a simplified gyral pattern, and volume loss. Furthermore, the first female child displayed retrognathia and hypotonia, while the second male child had non-cyanotic heart disease. The mother of the patients experienced a miscarriage, leading to a genetic carrier assessment for potential pathogenic gene variations.

METHODS: Blood-derived DNA was isolated for targeted gene capture through a custom capture kit. The subsequent sequencing of libraries achieved a mean coverage of >80-100X on the Illumina sequencing platform. Validation of the variant(s) by Sanger sequencing to rule out false positives

RESULT: A heterozygous one base pair deletion in exon 15 of the INPPL1 gene (chr11:71943766delC; c.1809del) results in a frameshift and premature truncation of the protein 17 amino acids downstream to codon 604 (p.Trp604GlyfsTer17) was detected in the Patients mother by Next Generation Sequencing (NGS). A heterozygous missense variation in exon 14 of the INPPL1 gene (chr11:71943374C>T; c.1706C>T) that results in the amino acid substitution of Methionine for Threonine at codon 569 (p. Thr569Met) was detected in Patient's Father by NGS.

DISCUSSION: The identified INPPL1 gene mutations in the parents provide a plausible explanation for the recurrent infant mortality and shared clinical features in their children. The frameshift mutation in the mother (p.Trp604GlyfsTer17) likely leads to a dysfunctional INPPL1 protein, impacting neurodevelopment and contributing to the observed phenotypes. The missense variation in the father (p.Thr569Met) may also play a role in the pathology, potentially influencing the protein's function.

CONCLUSION: These findings highlight the importance of genetic carrier assessments in cases of consanguinity and recurrent infant mortality, aiding in the understanding of underlying genetic factors and facilitating informed reproductive decisions. Further research on the role of INPPL1 in neurodevelopmental disorders is warranted to deepen our understanding of the molecular mechanisms involved.

Keywords: Opsismodysplasia, Rare skeletal dysplasia, Exome sequencing, INPPL1 gene, Mutations

A16

PHARMACOGENETIC PROFILING OF ADME-RELATED GENES IN AN INDIAN COHORT PREDICTS VARIABILITY IN DRUG RESPONSE AND IMPLICATIONS FOR PERSONALIZED MEDICINE

Aakanksha Pant*, Nikhat Khan, Chainika Gupta, Sahana S, Yogesh Joshi & Amol Naikawadi

Genetics Department, Indus Health Plus, Pune, Maharashtra, India
*Corresponding author: aakanksha.pant@indushealthplus.com

BACKGROUND: Pharmacogenetics is pivotal in understanding the impact of genetic variations on an individual's drug response, laying the foundation for personalized medicine. The Clinical Pharmacogenetics Implementation Consortium (CPIC) has provided comprehensive guidelines for numerous genes, elucidating their effects on diverse medications. Nearly 98% of individuals carry at least one pharmacogenetic variant, according to CPIC.

OBJECTIVE: To assess drug response patterns in the Indian population and explore implications for personalized medicine. Our objective was to establish a robust foundation using relevant data for personalized medicine in India. Recognizing the dearth of pharmacogenetics studies specific to the Indian population, we aimed to bridge this gap by conducting a comprehensive study within our diverse Indus cohort.

METHODS: A thorough retrospective microarray-based genotyping study involving 3139 individuals from a diverse Indian cohort was undertaken, focusing on ADME-related genes crucial in drug metabolism.

RESULTS: Our findings underscore significant genetic variations in drug responses among individuals in the Indian population. Specifically, 49.5% of participants carry the *2 star allele of the CYP2C19 gene, impacting Clopidogrel metabolism. According to our findings, approximately two-thirds of this population (67.65%) may be categorized as Poor or Intermediate metabolizers of CYP2C19, suggesting a potential need for an alternative drug to Clopidogrel. Similarly, 62.28% of the cohort possesses the *3 star allele of CYP2C9, influencing Fluvastatin and Warfarin metabolism, while 21.92% exhibit the variant T allele of VKORC1, crucial for Warfarin metabolism.

DISCUSSION: These findings highlight the significance of pharmacogenetic profiling in refining drug selection and mitigating adverse drug reactions (ADRs). The integration of pharmacogenetic principles into clinical decision-making holds the potential to markedly improve therapeutic outcomes and enhance patient care within the Indian population. Customizing drug therapies based on individual genetic variations not only optimizes therapeutic efficacy but also minimizes the likelihood of therapeutic failure and ADRs.

CONCLUSION: Our research underscores the practical value of pre-emptive pharmacogenetic information and emphasizes the imperative to embrace personalized medicine. With a deep understanding of genetic factors influencing drug responses, clinicians can make well-informed treatment decisions, leading to enhanced patient outcomes. The integration of pharmacogenetics into clinical practice holds the potential to transform drug prescription practices and elevate the standard of personalized healthcare in the Indian population.

A17

ELUCIDATION OF MITOCHONDRIAL SUB-COMPARTMENT SPECIFIC PROTEOTOXIC STRESS RESPONSE IN MAMMALIAN CELLS

Mudassar Ali[1], Priyanka Majumder[1], Pallavi Rao. T[2], Swasti Raychaudhuri[2] and Koyeli Mapa*[1]

[1]Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Greater Noida, Uttar Pradesh
[2]Centre for Cellular & Molecular Biology, Habsiguda, Uppal Road, Hyderabad
*Corresponding author: koyeli.mapa@snu.edu.in
[#]Presenting author: ma267@snu.edu.in

Mitochondria are major intracellular sites of proteotoxic stress apart from Endoplasmic Reticulum and cytosol. Our current understanding about cellular stress response due to proteotoxic stress specific to sub-mitochondrial compartments (matrix and Inter-membrane space or IMS) is only limited. In this work, by specifically targeting exogenous model misfolded and aggregation-prone amyloid forming proteins to mitochondrial matrix or IMS in mammalian cells, we find that the stress response elicited due to localized stress in matrix or IMS are distinct in many aspects. Matrix being more voluminous and highly enriched with canonical chaperone machineries, is expected to be well equipped to handle the proteotoxic stress more efficiently than IMS. Initially we found that targeting of aggregation prone proteins to IMS or to matrix of mitochondria is causing severe mitochondrial fragmentation and thus reduction in branch length. However, no significant cell cytotoxicity was observed in any of the cases, stating cells are utilizing some survival mechanism against these proteotoxicity. Our ongoing study points towards enhanced mitophagy flux in case of IMS proteotoxic stress, however similar increase of mitophagy flux is not found in case of matrix stress. PINK1-Parkin dependent mitophagy is highly upregulated in case of IMS proteotoxic stress. However, the well-known UPR^{mt} marker Hsp60, Hsp70 and Clpp are not significant upregulated suggesting UPR^{mt} is not activated, in any of the above cases. Thus, enhanced mitophagic flux in IMS is eliminating the damaged mitochondria therefore helping in the cell survival. However, in matrix neither mitophagy is significantly enhanced nor UPR^{mt} is activated, still cells efficiently survive suggesting the involvement of any novel survival mechanism. Next, a quantitative proteomics study was conducted to check the stress response at protein level. In case of matrix proteotoxic stress, many proteins participating in various pathways namely ER to Golgi vesicle transport, protein trafficking, anti-apoptotic and oxidative stress related proteins significantly and exclusively upregulated. whereas the levels of these proteins remained unaltered in case of IMS stress. In conclusion, we show that imparting same proteotoxic stress to mitochondrial IMS or matrix elicits distinct stress responses which is not much explored so far in literature.

A18

UNDERSTANDING THE LIFESTYLE OF *MYCOBACTERIUM ABSCESSUS* THROUGH PANGENOME ANALYSIS

Mistu Karmakar* & Saubashya Sur

Postgraduate Department of Botany, Ramananda College (Bankura University), Bishnupur, West Bengal, India
*Corresponding author: mistu.karmakar321@gmail.com

SUMMARY: *Mycobacterium abscessus*, an opportunistic pathogen, causes a range of infections and pulmonary diseases. Pangenome analysis was conducted to explore genomic variation, pathogenicity, and virulence patterns. Phylogenetic analysis revealed cluster-based relationships, highlighting factors influencing variability and adaptive capabilities. The presence of numerous biosynthesis genes suggested potential for investigating gene clusters. In silico screening identified a vaccine candidate from hypothetical proteins in COGs. Functional analysis of the pangenome components unveiled key factors associated with virulence, pathogenicity, infection establishment, immune evasion, persistence, and resistance.

BACKGROUND: *M. abscessus* is considered as a major concern owing to its resistance to antibiotics and disinfectants. Pangenome analysis is crucial for understanding lifestyle, genomic diversity, antimicrobial resistance, virulent phenotypes, and identifying potential drug candidates.

OBJECTIVES: (1) Characterization of pangenome architecture. (2) Mapping of these genes in COGs and KEGG categories for comprehending the lifestyle of *M. abscessus*.

METHODS: Pangenome of 75 *M. abscessus* complete genomes was analysed using BPGA v. 1.3, employing its functional module for pan genome profiling, sequence extraction, exclusive gene family analysis, species phylogenetic analysis, and pan genome functional analysis. tvBot visualized pan and core phylogeny. Sequence mapping against COG and KEGG databases using ublast assessed functionalities and associations with biological processes. COG-associated hypothetical proteins were investigated for functional and immunological capabilities.

RESULTS: The pangenome analysis showed 58% core genes, 41% accessory genes, and 1% unique genes. Phylogenetic tree showed that some strains were clustered together while the others were interspersed. Functional categories indicated a prevalence of genes related to amino acid transport, metabolism, lipid transport, metabolism, secondary metabolite biosynthesis, transcription, replication, recombination, repair, with a significant number of genes poorly categorized. KEGG categories highlighted association with various metabolic pathways. Abundant virulent, toxic, and antibiotic genes included MCE family protein, MMPL family transporter, TetR/AcrR family transcriptional regulator protein, and different types of restriction endonuclease proteins. In silico screening identified a suitable, antigenic, non-virulent, and non-toxic hypothetical protein (HP). A linear B-cell epitope from this HP showed promise for vaccine design.

DISCUSSION: The result underlined the fact that pangenome is still open while the core genome has stabilised. The accessory genes contributed to the diversity. The COGs functional categories revealed that the core, accessory and unique genes are associated with amino acid, lipid, metabolism and biosynthesis of different metabolites. This enhanced the adaptation and sustainability of *M. abscessus* within the host. A number of COG genes are associated linked to virulence of *M. abscessus* provided insights into pathogenic lifestyle.

CONCLUSION: The study emphasized the role of accessory genes in influencing evolutionary dynamics. Pangenome analysis identified key genes in transcription, metabolism, and pathogenesis,

showcasing diverse metabolic capabilities for host ecosystem survival. Screening of hypothetical proteins revealed potential vaccine candidates. The outcome will aid in better comprehension of host-pathogen relationship.

A19

TRANSCRIPTOMIC PROFILING REVEALS EPIGENETIC REGULATION OF ENDOTHELIAL PROTHROMBOTIC PHENOTYPE UNDER HYPOXIA

Mohd Yasir Khan, Kashika Singh, Alia Hashmi, Syed Mohd & Mohammad Zahid Ashraf*

Cardiovascular Genomics Lab, Department of Biotechnology, Jamia Millia Islamia, New Delhi, India.
* Corresponding author: zashraf@jmi.ac.in

BACKGROUND: The endothelium is recognized as a crucial regulator of vascular homeostasis, with endothelial cell activation being linked to the onset of cardiovascular complications. Hypoxia induces an inflammatory response and is implicated as a contributing factor to endothelial activation. The involvement of hypoxia in the development of cardiovascular complications has been extensively studied. Increasing evidence supports the significance of long non-coding RNAs in cardiovascular diseases.

OBJECTIVES: In this study, we explored the potential role of lncRNAs as regulators of the endothelial pro-thrombotic phenotype in the context of hypoxic conditions.

METHODS: Transcriptomic datasets from hypoxia-exposed HUVECs (0.1 – 1% O2) were analyzed to identify differentially expressed lncRNAs and genes. Subsequent analyses included Gene Ontology, Pathway, and Disease Ontology enrichment, as well as Gene Set Enrichment Analysis. Co-expression analysis of differentially expressed lncRNA-mRNA pairs, along with ChIP-seq, ATAC-seq, and TF binding site analyses, were performed. Bioinformatic findings were validated *in-vitro* using Ea.hy926 cells subjected to 24 hours of hypoxia (1% O_2), DMOG treatment (1mM, 12 hrs.), DIM (10µM, 24 hrs.) and silencing of LINC00607 through LNA gapmers.

RESULTS: Transcriptomic analysis indicated a marked upregulation of LINC00607 in endothelial cells subjected to hypoxia. LINC00607 was identified as a dual attribute—being both HIF1α-regulated and specific to endothelial cells. Analysis of lncRNA and mRNA co-expression revealed a significant correlation between LINC00607 and the prothrombotic gene TSPAN18. Transcriptomic data from LINC00607 knock-out experiment revealed a downregulation of TSPAN18. ATAC-seq analysis following LINC00607 knock-out revealed decreased accessibility at the TSPAN18 locus. LINC00607 regulates endothelial expression by modulating ERG transcription factor binding. H3K27Ac ChIP-Seq analysis with siERG demonstrated a significant decrease in acetylation levels at the TSPAN18 locus, mirroring the alterations observed under hypoxic conditions. *In-vitro* validation through RT-PCR and immunofluorescence confirmed the HIF1α-mediated regulation of LINC00607, resulting in the upregulation of TSPAN18 and suggesting enhanced vWF release under hypoxic conditions.

CONCLUSION: In summary, we identify LINC00607 as a potential regulator of the TSPAN18 gene in hypoxic conditions, leading to a prothrombotic phenotype. This long non-coding RNA may serve as a biomarker and therapeutic target for hypoxia-induced thrombosis. This study is the first to report an elevation in TSPAN18 expression under hypoxic conditions.

A20

IDENTIFICATION OF BASELINE AND RESISTANCE-ASSOCIATED SUBSTITUTIONS IN NS3 AND NS5 REGION BY NEXT-GENERATION SEQUENCING AMONG HCV GENOTYPE 3

Sagnik Bakshi[#], Raina Das, Supradip Dutta, Shreyashi Nath, Shanta Dutta & Provash Chandra Sadhukhan*

ICMR-National Institute of Cholera and Enteric Diseases, P-33 CIT Road, Scheme XM, Beliaghata, Kolkata; India
*Corresponding author: sadhukhan.pc@icmr.gov.in
#Presenting author: bakshi445@gmail.com; 7044716201

BACKGROUND: Hepatitis C virus (HCV) is a blood-borne positive-stranded RNA virus associated with morbidity and mortality in humans. This virus is linked with liver-related complications like chronic liver diseases or hepatocellular carcinoma (HCC). To combat against HCV, before 2011, a combination of Pegylated-interferon (Peg-IFN) and Ribavirin (RIB) was the gold standard for HCV treatment. However, lower sustainable viral response (SVR) and adverse side effects to the patient lead to the introduction of direct-acting antivirals (DAAs). The advent of DAAs has changed HCV treatment scenarios around the globe. However, an increasing number of treatment failure is reported in different parts of the world which may cause a major concern for public health.

OBJECTIVES: To conduct a detailed analysis of DAAs efficacy against HCV and to find baseline and resistance-associated substitutions (RAS) among HCV infected DAAs treatment-naive and non-responders' patients.

METHODS: In this study, 198 HCV sero-reactive patients were enrolled over three years for DAAs treatment. Blood samples were collected before and after the DAAs treatment and proceeded for viral RNA isolation. HCV RNA was quantified by using the qRT-PCR. HCV genotype was determined by nested RT-PCR of partial HCV core (405bp) gene amplification followed by Sanger sequencing and NCBI genotyping tool. GT-3 specific baseline substitutions and RAS were detected by amplifying whole NS3 and NS5 regions followed by next-generation sequencing (NGS) and HCV Geno2pheno online tool.

RESULTS: The efficiency of DAAs varied between ~ 94% and 100%. The maximum numbers of DAAs non-responders were observed against SOF/DAC combination (n=7; 63.63%). DAAs non-responders were high among decompensated cirrhosis of liver patients (90.9%), and HCV GT-3b infected patients (n=5, 45.45%). We performed baseline substitutions and RAS analysis of GT-3-specific NS3 and NS5 regions. Two important RAS in NS5A and NS5B regions such as V31M and S282T were circulating among GT-3b drug-resistance patients, whereas the prevalence of baseline mutations like A98T, D172E and R114K was observed in NS3, NS5A and NS5B region of GT-3a infected naïve patients.

DISCUSSION AND CONCLUSIONS: Extensive genomic variability was observed due to high error-prone RNA-dependent RNA polymerase (RdRP) activity in the HCV. In India, GT3 is the most prevalent genotype, followed by GT-1. GT3 is a comparatively difficult genotype to treat. The increasing number of DAAs treatment failure cases among GT-3b infected patients indicates a serious concern for public health because maximum decompensated cirrhosis of liver cases is connected with GT-3 infection. Moreover, the circulation of drug-resistance mutations like V31M and S282T is a serious concern in terms of the treatment of HCV. Thus, frequent screening of baseline mutation and RAS is required to reach our goal of eradicating HCV by 2030 and to find new DAAs treatment combinations in the future.

A21

TUNING THE SYMPHONY: DYNAMICS OF TRNA-DERIVED SMALL NON-CODING RNAS DURING HUMAN MONOCYTE-TO-MACROPHAGE MATURATION

Anahita Jayaram[#] & Kamlesh Pawar*

Centre of Excellence in Epigenetics, Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence Deemed to be University, Delhi-NCR, India
*Corresponding author: kamlesh.pawar@snu.edu.in
#Presenting author: aj758@snu.edu.in; +91 7259200667

Monocytes, originating from the bone marrow, represent a type of white blood cell that plays a pivotal role as innate immune cells. In response to infection or tissue damage, monocytes are swiftly mobilized to the affected tissue, where they undergo differentiation into tissue macrophages. Recent studies have highlighted the involvement of short non-coding RNAs (sncRNAs) during the process of monocyte-to-macrophage differentiation. In this context, a major subset of sncRNAs, such as tRNA-derived sncRNAs, has not been studied. It is important to study the expression of tRNA-derived sncRNAs as it has gained attention because of their role in immunity and immunological pathways. tRNA-derived sncRNAs have been difficult to study because of the presence of cyclic phosphate (cP) at their 3'-end, which cannot be detected by standard small RNAseq platforms.

This study investigates the dynamic regulation of tRNA-derived sncRNAs during human monocyte-to-macrophage differentiation and their potential impact on the functional characteristics of macrophages. We used THP-1, a human leukemia monocytic cell line, as a model to study human monocyte-to-macrophage differentiation. Monocyte-to-macrophage differentiation was carried out by exposing the THP-1 cells to a potent differentiating compound, phorbol 12-myristate-13-acetate (PMA). RNA isolated during differentiation was subjected to T4 polynucleotide kinase (T4 PNK) treatment to convert the RNA termini to 5'-P/3'-OH-ends (thus rendering them available for 5'-/3'-AD ligation), which is required to capture whole sncRNAs. Illumina's TruSeq small RNA sequencing was used to sequence T4 PNK treated RNAs.

Our results reveal significant alterations in the abundance and diversity of tRNA-derived sncRNAs during this differentiation process. Furthermore, we identified specific tRNA-derived sncRNAs that exhibit stage-specific expression patterns, suggesting a potential role in orchestrating the molecular events associated with macrophage maturation.

A22

VIRAL GENOMIC SURVEILLANCE IN *AEDES* MOSQUITO POPULATIONS

Chitra Patankar[1], Umay Kulsum[1], Dipesh Kale[1], Shruti Mahajan[2], Shaswati Nema[1], Vineet Sharma[2] & Debasis Biswas[1]*

[1]Department of Microbiology, All India Institute of Medical Sciences, Bhopal, Madhya Pradesh, India.
[2]Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, Madhya Pradesh, India.
*Corresponding author: debasis.microbiology@aiimsbhopal.edu.in

BACKGROUND: Mosquito-borne viral diseases pose a significant impact on human health, with the potential to cause widespread outbreaks of diseases. Monitoring viral genomes in mosquito populations can lead to informed risk assessment and promote early diagnosis. However, there is lack of standardized methodology aimed at deciphering circulating viral sequences in mosquito populations collected from human habitats.

OBJECTIVE: To establish and evaluate a system of viral metagenomic analysis in Aedes mosquito population.

METHODS: Mosquitoes were collected using CDC approved BG sentinel trap and battery-operated vacuum aspirators from different locations of Bhopal region. They were sorted on the basis of genus, gender, location and date of collection. RNA was extracted from the homogenised mosquito pools and reverse transcribed. cDNA was amplified using Sequence Independent Single Prime Amplification (SISPA) approach to increase the quantity of viral nucleic acid. Further, PCR products were sequenced using NovaSeq 6000 platform. Bioinformatic analysis of the reads was performed using Trimmomatic for trimming of low-quality raw reads. Later Kraken and Bracken were used for identification of viral sequences.

RESULTS AND DISCUSSION: From a total of 31mosquito pools 428,276,613 paired end reads were obtained and after quality filtration 375,010,668 reads remained. The metagenomic analysis using bioinformatic software revealed that viral abundance was significantly higher during monsoon months in mosquito population at family, genus and species level. In addition to this temporal variation, significant geographical variation was also observed in the viral reads collected across the locations. We failed to find any viral sequence belonging to pathogenic arboviruses, which corresponded with the low number of cases of diseases like Dengue, Chikungunya in the same year in the national portal of vector-borne diseases.

CONCLUSION: Our finding suggests that early detection of viruses with human pathogenic potential can be made through viral genomic surveillance in vector population enabling us to employ targeted vector control measures and avert potential outbreaks of these diseases in the future.

A23

EXPLORATION OF THE COVID-19 GENOME USING OXFORD NANOPORE TECHNOLOGY WITH SALIVA AS AN ALTERNATIVE SPECIMEN SOURCE

Anvita Gupta Malhotra[1#], Anindita Datta[1], Dipesh Kale[1], Divya Namdeo[1], Ashvini Kumar Yadav[1], Satya Prakash Vishwakarma[1], Shashwati Nema[1], Rajnish Josh[2] & Debasis Biswas[1]*

[1]Department of Microbiology, All India Institute of Medical Sciences (AIIMS), Bhopal
[2]Department of General Medicine, All India Institute of Medical Sciences (AIIMS), Bhopal
* Corresponding Author: debasis.microbiology@aiimsbhopal.edu.in ; +91-7747010312
#Presenting Author: anvitagupta16@gmail.com; +91-9669536156

BACKGROUND: The conventional method for diagnosing COVID-19 with RT-PCR requires collection of Nasopharyngeal/Oropharyngeal (NP/OP) swabs which presents challenges related to biosafety concerns and reliance on healthcare infrastructure for sampling. To address these issues, this study aimed to assess the viability of using saliva samples as an alternative sample for diagnosis of COVID-19.

METHODS: The study involved collecting paired samples of NP/OP swabs and saliva from 515 COVID-19 suspects. These samples were processed in a blinded manner, with NP/OP swabs following the standard RNA extraction protocol and saliva samples undergoing RT-PCR either directly (SD) or after RNA extraction (SR).

RESULTS: Among the 515 COVID-19 suspects, 114 were diagnosed as positive using NP/OP swabs, 117 with SR, and 147 with SD protocols. The sensitivity and specificity of SR compared to NP/OP swabs were 61.4% and 88.28%, respectively, while for SD, they were 70.17% and 83.29%. Notably, 16.69% of the samples that tested positive in either saliva sample methods were negative in NP/OP swabs. To investigate this discrepancy, Nanopore sequencing was performed on 69 of these 86 saliva samples with Ct values below 32, detecting SARS-CoV-2 RNA in 78.26% of them, reinforcing the effectiveness of saliva samples in enhancing case detection.

CONCLUSION: This study underscores the significance of improving COVID-19 diagnosis through the use of saliva samples, addressing biosafety concerns. Additionally, it marks the inaugural submission of complete Indian SARS-CoV-2 genome sequences obtained from saliva samples through the Oxford Nanopore platform.

Keywords: Saliva, COVID-19, SARS-CoV-2, NP/OP swabs, RT-PCR, Next Generation Sequencing, Oxford Nanopore MinION.

A24

THE PANGENOME ARCHITECTURE OF *MYCOBACTERIUM AVIM* COMPLEX

Anindita Banerjee* & Saubashya Sur

Department of Botany, Ramananda College, Bankura University, Bishnupur, West Bengal, India
*Corresponding author: aninditabanerjee960@gmail.com

SUMMARY: Mycobacterium avium complex (MAC) comprises a diverse assemblage of non-tuberculosis mycobacteria linked to various diseases. The pangenomic analysis exhibited an open trend allowing it to accumulate more genes. Phylogenetic analysis divulged diversity within the subspecies. Mapping of pangenomic components in COGs and KEGG emphasized the impact of strong metabolic capacity and virulence in maintaining it's lifestyle.

BACKGROUND: *Mycobacterium avium* complex (MAC) comprises a diverse assemblage of non-tuberculosis mycobacteria causing diseases across different hosts. A necessity was felt to make a comprehensive investigation of the pangenome, for a deeper understanding of the factors contributing to genomic diversity, adaptations, virulence, and lifestyle.

OBJECTIVES: Deciphering the evolutionary dynamics and pathogenic lifestyle

METHODS: A dataset consisting of 90 high quality MAC genomes were retrieved from the PATRIC database. BPGA pangenome analysis pipeline was utilized to characterize the pangenome. USEARCH was used for clustering. Phylogenetic analysis module and concatenated alignment of core genes were used for estimating evolutionary dynamics. They were visualized using tvBOT software. Functional analysis were performed using COG and KEGG mapping.

RESULTS: The pangenome comprised of 46.1% core genome, 52.43% accessory genome, and 1.46% unique genome. It exhibited an open trend without reaching a plateau, indicating that it can accumulate more genes. The Pan-Core plot revealed increasing trajectory for the pan and a decreasing core. Phylogenetic trees revealed number of interspersed organisms and some close to each other. Notable sequevar differences were observed for *Mycobacterium avium* subsp.*avium*, *Mycobacterium avium* subsp. *hominissuis* and *Mycobacterium avium* subsp. *paratuberculosis*, alongside variations among bovine and ovine *Mycobacterium avium* subsp. *paratuberculosis* strains. Phylogenetic analysis revealed that *Mycobacterium avium* subsp. *hominissuis* strains had highest sequence diversity and *Mycobacterium avium* subsp. *paratuberculosis* had the least. The outcome revealed that the genes are associated with different COG categories. Metabolism-related COGs were most abundant, while cellular processes and signaling COGs were the least represented. Virulent, pathogenic, toxin-related genes included in core, accessory, and unique genomes are ESX. Mmpl, TetR, PPE family protein, ABC transporter protein, etc. KEGG analysis identified the genes associated with different categories involved in pathways related to human diseases, metabolism, cellular processes, environmental information, genetic information processing, and organismal systems. Metabolism had the highest gene count, while organismal systems had the lowest.

DISCUSSION: Accessory genomes catalyzed diversity in MAC. This is a strategy for maintaining its pathogenic lifestyle. The diverse metabolic capability allowed *M. abscessus* to survive within the host ecosystem and sustain its lifestyle. It assisted in adaptation and survival under varied environments and hosts.

CONCLUSION: The outcomes provided valuable insights, into evolutionary dynamics, pathogenic mechanisms, and lifestyle adaptations. This enhances comprehension of the roles played by functional genes in pathogenesis, virulence, and antibiotic resistance. These outcomes could be significant for drug development, genomic surveillance, genome plasticity, and drug development.

A25

AFFORDABLE DNA SEQUENCERS IN MICROBIAL METAGENOMICS, A BOON TO BUDDING LABS: EXPERIENCES FROM INVESTIGATION ON GUT-MICROFLORA-BRAIN AXIS UPON XENOESTROGEN(S) EXPOSURE

Dhirendra Pratap Singh* & Santasabuj Das

Neurotoxicology and Immunotoxicology Laboratory, Division of Biological Sciences, ICMR-National Institute of Occupational Health, Ahmedabad, Gujrat, India
*Corresponding author: singh.dpratap@icmr.gov.in

Endocrine-disrupting chemicals (EDCs) are potential xenoestrogenic chemicals which affect hormonal balances. Studies assessing their effects on the reproductive, endocrine, and hepato- renal neurological system have shown their ability to cause toxicological alteration. However, due to the non-monotonic dose response many of them have shown in varied dose ranges. Especially, a substance such as bisphenol A (BPA) has shown such toxic effects in the doss well below their No Observed Adverse Effects Levels (NOAEL). Due to regulatory constraints over BPA use, its alternatives are being used in consumer goods which still hold the promise to affects similar to it. We have studied the effects of two analogs of bisphenols at doses lower than the NOAEL in adult mice to assess the impact of this exposure on brain functions, specifically to depression and anxiety like behavior. Further, considering the role of gut and resident microbes in host physiology and in particular to mood disorders we have done gut microbial metagenomics analysis covering 16S rRNA gene amplicons sequencing using Oxford Nanopore Sequencing (MinIOn Mk1B). Results from the study suggested that the effects of these bisphenol analogs are sex-specific as male are more pronounced toward the negative effect on behavioral features related to anxiety and depression. Further, these exposures increased oxido-nitrosative stress and inflammation in brain tissues an altered neurotransmitter turn-overs linked to their biosynthesis, metabolism and reuptakes. For the gut specific markers, it increases inflammation in ileum along with alteration in gut barrier functions. Cecal metagenomics analysis revealed distinct microbial signature in male and female mice. This is also linked with the alteration in diversity, and homogeneity of microbial population dynamics. Use of affordable DNA sequencing technologies helped us a lot in executing such work in greater number of samples covering varied experimental conditions. As we are in an establishing phase, such technologies can not only assist in deciphering the intricacies of gut-microbial-brain function related studies but also provide other avenues exploring the potential of genomics technology in the area of behavioral pharmacology.

A26

GENETIC APPROACHES TO UNRAVEL COMPLEX MOLECULAR MECHANISMS OF ION TRANSPORT IN *PLASMODIUM* FOR DEVELOPING NOVEL ANTIMALARIALS

Ankit Gupta*

Department of Biochemistry, All India Institute of Medical Sciences (AIIMS), Uttar Pradesh, India
*Corresponding author: guptaankitbiochem07@gmail.com; +91-9839082545

Rising drug resistance to currently available antimalarials is one of the greatest threats to global malaria eradication which results into increased morbidity and mortality. Therefore, there is an urgent need of developing novel drug targets to combat malaria infection and support malaria elimination program. The intracellular parasite exploits the host machinery to create a hospitable environment for its optimum growth through acquisition of essential nutrients and ions from human plasma. The ion transport in *Plasmodium* is a complex process which is attributed to epigenetic regulation of multigene family. The gene locus responsible for the nutrient/anion transport was identified by the linkage analysis and genetic mapping. The functional role of key parasite proteins in ion transport was deciphered by making transgenic parasite lines using a number of genetic manipulation tools e.g. CRISPR-Cas9. The inhibitors targeting key ion transport were screened on *in vitro* asexual blood culture and novel molecular insight into mechanism was revealed. The findings provide a novel example of gene family expansion to allow ion transport and insight into pharmacology relevant to drug development against malaria parasite.

A27

FUNCTIONAL CHARACTERIZATION OF PUTATIVE SERINE/THREONINE PROTEIN KINASE IN *PLASMODIUM FALCIPARUM*

Vikas Kumar, Manoj Yadav & Ankit Gupta*

Department of Biochemistry, All India Institute of Medical Sciences (AIIMS), Uttar Pradesh, India
*Corresponding author: vikaskamalvanshi@gmail.com; +91-8090986083

The intracellular parasite modulates the host machinery to create a hospitable environment for its optimum growth by trafficking a number of proteins in host compartment. One such protein known as serine/threonine protein kinase (STPK) was found in host erythrocyte membrane fraction based on unique peptides through LC-MS/MS analysis. Later on, STPK was reported as an orthologue of 3-phosphoinositide-dependent protein kinase-1 (PDK-1) based on multiple sequence alignment however, it lacks a characteristic c-terminal phospholipid-binding pleckstrin homology (PH) domain required for membrane localization by recognizing phosphatidylinositol phosphate (PIP) lipids. Here, we successfully expressed and purified recombinant STPK protein with c-terminal 6xHis tag at ~65kDa which was further used to raise polyclonal antibodies for localization. In addition, the kinase activity was checked on purified recombinant STPK using kinase-Glo kit. A selection-linked integration (SLI) approach was utilized to generate conditional knockout of endogenous STPK gene using dimerazable Cre recombinase (diCre) method. Furthermore, mutant STPK (M51R) protein was recombinantly expressed using site-directed mutagenesis and functionally characterized by checking its kinase activity. The findings provide a key insight into signalling cascade involved in merozoite invasion and imply that STPK is a potential new drug target.

A28

PHENOTYPIC AND GENOMIC CHARACTERIZATION OF PLANT GROWTH PROMOTING BACTERIAL SYMBIONT OF HIGH ALTITUDE ENDANGERED MEDICINAL PLANT *ACONITUM HETEROPHYLLUM*

Refad Ahmed[1,2], Diganta Kumar Saikia[1,2], Natarajan Velmurugan[1,2]*

[1]CSIR-North East Institute of Science and Technology (CSIR-NEIST), Branch Laboratory-Itanagar, Arunachal Pradesh, India
[2]Academy of Scientific and Innovative Research (AcSIR), Uttar Pradesh, India
*Corresponding author: velmmk@gmail.com, natarajan@neist.res.in

Due to lack of accessibility to endangered medicinal plants of high-altitude Himalayan Mountains, the biotechnological potential of microorganisms associated with these plants has not been largely explored. Here, we demonstrate the beneficial phenotypic and genomic characteristics of culturable bacterial symbionts associated with an endangered medicinal plant *Aconitum heterophyllum* from high-altitude Himalayan Mountain ranges. We have successfully isolated and characterized a bifunctional and metabolically stable *Pantoea* sp. AHE68 from *A. heterophyllum*. *Pantoea* sp. AHE68 was demonstrated to simultaneously convert inorganic phosphate into organic phosphate (29.8% increase in available phosphorus content) and synthesis growth-promoting IAA (45.34±0.82 μg mL^{-1}) thereby enhances photosynthetic microalgal productivity in synthetic co-culture system. AHE68 was found to simultaneously secrete organic acid (10.42 mM mL^{-1}) to convert the inorganic phosphate into organic phosphate as well as produces high level IAA under stress conditions. Interactions of AHE68 and long chain fatty acids (LCFA) producing photosynthetic microalga *Micractinium* sp. GA001 were characterised in different combinations of synthetic co-culture systems. The endophyte AHE68 was found to significantly enhance 22.71% and 79.27% increase in microalgal cell numbers and lipid contents in co-culture systems, respectively. In co-culture studies, AHE68 was found to be metabolically capable to simultaneously synthesise active forms of IAA and organic acids under salt-stress conditions. The findings of the phenotypic and genotypic analysis give us the proof that the bacterial symbionts were metabolically stable for extended period (up to 30 days) to improve microalgal productivity. The findings were further complemented with genomics studies of symbiont which was subjected to multiple sequencing approaches to assemble and annotate their genome, resulting in key genes involved in PGP activities, metabolites production and transportation being identified. The bacterial symbiont was fully characterized by the whole genome sequencing with two different sequencing approaches. Hybrid genome of the bacterial symbiont had been obtained with SPAdes using a hybrid method, combining Illumina paired end reads with Nanopore single-end reads. Hybrid genome assembly is a strategy to overcome the limitations of individual sequencing techniques including long-read and short-read techniques. Illumina and PacBio/Nanopore were well-known next generation sequencing (NGS) techniques to produce short and long length reads, respectively. However, several limitations were associated with these techniques such as large genome fragmentations (in Illumina), identification of transposons, mobile genetic elements, unambiguous reads, horizontal gene transfers events, and large-scale errors (in PacBio/Nanopore). The approach helped us to assemble and annotate symbiotic bacterial genome yielded highly accurate, fully finished hybrid genome along with plasmids and core genes. This study expands the benefits and bioprocessing potential of bacterial symbiont of Himalayan medicinal plant.

A29

IRC20 MODULATES LOH FREQUENCY AND DISTRIBUTION IN *S. CEREVISIAE*

Sameer Joshi[1†]*, Suman Dash[1†], Nikilesh Vijayan[1] & Nishant K.T.[1,2].

[1]School of Biology, Indian Institute of Science Education and Research Thiruvananthapuram, Kerala, India
[2]Center for High-Performance Computing, Indian Institute of Science Education and Research Thiruvananthapuram, Kerala, India
[†]Equal contribution
*Corresponding author: sameerjoshi21@iisertvm.ac.in, 9673967265

Loss of Heterozygosity (LOH) due to mitotic recombination is frequently associated with the development of various cancers (e.g. retinoblastoma). LOH is also an important source of genetic diversity, especially in organisms where meiosis is infrequent. *Irc*20 is a putative helicase, and E3 ubiquitin ligase involved in DNA double-strand break repair pathway. We analyzed genome-wide LOH events, gross chromosomal changes, small insertion-deletions and single nucleotide mutations in eleven *S. cerevisiae* mutation accumulation lines of *irc*20Δ, which underwent 50 mitotic bottlenecks. LOH was enhanced significantly in *irc*20Δ as compared to the wild type. Interstitial and terminal LOH events were similar in *irc*20Δ, unlike in wild type, where terminal LOH events are reduced. LOH events in *irc*20Δ were more telomere proximal compared to the wild type. Further, we observed that terminal LOH tracts in *irc*20Δ were longer than the wild type. Gross chromosomal changes, single nucleotide mutations and in-dels were not enhanced in the *irc*20Δ mutant. Locus based and genome-wide analysis of meiotic recombination showed that meiotic crossover and non- crossover frequencies are not altered in *irc*20Δ. These results suggest IRC20 primarily regulates mitotic recombination and does not affect meiotic recombination. Our results suggest that the IRC20 gene is important for regulating LOH frequency and distribution.

A30

MUTATIONS IN SARS-COV-2 LINKED TO IMMUNE RESISTANCE RESULTING IN BREAKTHROUGH INFECTIONS

Chetan Sahni*

Department of Anatomy, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India
*Corresponding author: chetansahni@bhu.ac.in; +91 8802528235

BACKGROUND: India had a catastrophic second COVID-19 epidemic, with most virus sequences identified as belonging to the B.1.617.2 lineage (Delta-variant). As the globe and India concentrated on immunization, news of the virus evading vaccination—referred to as "breakthrough cases"—arose everywhere.

OBJECTIVES: Finding the alterations linked to SARS-CoV-2 breakthrough infections was the main goal of our investigation.

METHODS: In this investigation, we extracted the ribonucleic acid, or SARS-CoV-2 RNA, from patients who tested positive for COVID-19 using reverse transcription polymerase chain reaction (RT-PCR). We then randomly selected 150 samples and sent them to the Centre for Cellular & Molecular Biology in Hyderabad for sequencing. 150 SARS-CoV-2 virus samples had their whole genome sequences examined in detail. The findings of the genome sequencing mostly revealed B.1.617 and its sub-lineages.

RESULTS AND DISCUSSION: It was discovered via additional examination of the patient data that nine individuals had previously received a SARS-CoV-2 vaccination. The B.1.617/B.1 or A strains that these nine individuals carried had comparable genetic alterations in spike proteins and non-structural proteins (NSPs). The Spike (S), NSPs, and open reading frame (ORF) sections of these sequences were found to have undergone mutations that would have changed the amino acids in a way that would have improved viral replication, conferred drug resistance, influenced host-cell contact, and caused antigenic drift.

CONCLUSIONS: These unique changes may indicate increased virulence, which leads to vaccination immunity evasion. Our findings add to the mounting evidence that mutations in the S (Spike) and ORF genes of SARS-CoV-2 lead to immune evasion.

A31

UNVEILING THE MICROBIAL COMPLEXITY OF NECROTISING SOFT-TISSUE INFECTIONS: GENOMIC EXPLORATION THROUGH 16S RRNA SEQUENCING IN AN INDIAN COHORT

Tanvi Vashist[1]*, Archana Angrup[1], Vikas Sharma[1], Divya Nair[1], Cherring Tandup[2], Pallab Ray[1] & Amit Arora[1]

[1]Department of Medical Microbiology, Postgraduate Institute of Medical Education and Research, Chandigarh, India
[2]Department of General Surgery, Postgraduate Institute of Medical Education and Research, Chandigarh, India
*Corresponding author: tanvipgi@gmail.com

SUMMARY: Necrotising Soft-tissue infections (NSTIs) pose a significant health challenge, marked by high mortality rates, especially in polymicrobial cases. This study delves into the microbial diversity within NSTIs using 16S rRNA gene sequencing, with the goal of advancing diagnostic precision and refining treatment strategies.

BACKGROUND: NSTIs, characterized by the rapid destruction of skin, muscles, and fascia, present diagnostic intricacies, particularly in severe cases. The underestimated microbial diversity in NSTIs, especially in polymicrobial scenarios, necessitates advanced diagnostic methodologies. Aims and Objectives: This study aims to assess the microbial diversity directly in patient tissues diagnosed with NSTIs, utilizing prokaryotic 16S ribosomal genes and intergenic regions. Our objectives include categorizing patient tissue biopsies and uncovering key taxa through Illumina MiSeq sequencing, targeting the V3-V4 region of 16S rRNA gene.

METHODS: Patient tissue biopsies, encompassing cases of Necrotizing fasciitis, Gas gangrene, Fournier's Gangrene, and others, underwent 16S metagenomic sequencing for pathogen identification. This involved 16S rRNA NGS Library preparation, followed by analysis using QIIME2-2023.2, including taxonomic profiling, OTU clustering, and core microbiome analysis. Visualization was achieved through Cytoscape 3.10.1.

RESULTS: Key taxa, such as *Streptococcus pyogenes*, *Acinetobacter baumannii*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Proteus mirabilis*, were identified. Culture-based analysis categorized samples into polymicrobial, monomicrobial, and sterile cases.

DISCUSSION: Elucidating NSTI dynamics revealed Acinetobacter baumannii significant prevalence (>80%) alongside Streptococcus pyogenes and Escherichia species. Noteworthy findings include the prominence of Acinetobacter baumannii in NSTIs and the identification of unculturable taxa, enriching our understanding of NSTI microbiomes.

CONCLUSIONS: This study represents the initial exploration of NSTI microbiomes in India, employing metagenome sequencing. The complex microbial landscape uncovered, dominated by Acinetobacter baumannii, underscores the necessity for routinely incorporating next-generation sequencing in clinical setups. This research establishes a precedent for future NSTI investigations, offering insights into microbial diversity critical for informing diagnostic and treatment strategies.

A32

COMPUTATIONAL EXPLORATION OF LIGNIN PEROXIDASE ISOZYMES FROM *PHANEROCHAETE CHRYSOSPORIUM*: UNVEILING STRUCTURAL DIVERSITY AND FUNCTIONAL INSIGHTS

A. Abdul Basith[#] & R. Sathishkumar*

Plant Genetic Engineering Laboratory, Department of Biotechnology, Bharathiar University, Coimbatore, India.
*Corresponding author: rsathish@buc.edu.in
#Presenting author: abdulbasithphd@gmail.com

BACKGROUND: Lignin peroxidases (LiPs) are pivotal enzymes involved in lignin degradation, showing immense potential in a wide range of industrial and environmental applications. The presence of high lignin content in woodchips, sawdust, forestry residues, and agro-waste. As a result, it causes numerous environmental problems due to its slow decay. Understanding the structural and functional attributes of LiP isozymes from *Phanerochaete chrysosporium* (PC) is crucial for harnessing their enzymatic capabilities. However, the pairwise identity, evolutionary relationships, motifs, and domains of all ten PC-LiPs remain unknown.

OBJECTIVES: This study employs an *in silico* approach to investigate the structural and functional characteristics of LiP isozymes from the fungus *P. chrysosporium* Through comprehensive sequence analysis, phylogenetic studies, and N-Glycosylation site prediction, we elucidated the structural motifs and functional domains within the PC-LiP isozymes.

METHODS: Studied the evolutionary links, comprehensive sequencing analysis, and structural motifs within the PC-LiP isozymes. Various bioinformatics tools were used to identify conserved domains, active sites, and substrate-binding pockets. The N-Glycosylation sites and solubility of the isozymes in *E. coli* were analyzed using online tools NetNGlyc v1.0 and SoluProt v1.0.

RESULTS: The study revealed the evolutionary relationships, sequence identities, and structural features important for understanding the functional aspects of PC-LiP isozymes. The findings identified the conserved domains, active sites, and glycosylation sites and possible implications in industrial applications for the degradation of the lignin.

DISCUSSION AND CONCLUSION: Our *in silico* analysis contributed to a comprehensive understanding of the structural and functional characteristics of *P. chrysosporium* LiP isozymes. This study forms the foundation for our future approaches in production of recombinant lignin peroxidase candidates in *E. coli* and also to transiently express in *Nicotiana benthamiana* plant host system. This technology will be very useful for the degradation of high lignin content in the agricultural wastes that is prevalent in many parts of our country.

A33

A NOVEL MACHINE LEARNING FRAMEWORK FOR CLASSIFICATION OF ADENOCARCINOMA AND SQUAMOUS CELL CARCINOMA USING LUNG MICROBIOME

Pragya Kashyap[1], Kalbhavi Vadhi Raj[2], Naveen Dutt[3] & Pankaj Yadav[1,4]

[1]Department of Bioscience & Bioengineering, Indian Institute of Technology, Jodhpur, Rajasthan, India
[2]Department of Electrical Engineering, Indian Institute of Technology, Jodhpur, Rajasthan, India
[3]Department of Pulmonary Medicine, All India Institute of Medical Sciences, Jodhpur, Rajasthan, India
[4]School of Artificial Intelligence and Data Science, Indian Institute of Technology, Jodhpur, Rajasthan, India
*Corresponding author: kashyap.3@iitj.ac.in

BACKGROUND: Non-small-cell lung cancer (NSCLC) accounts for 80-85% of lung cancer (LC) cases, mainly classified into Adenocarcinoma (AC) and Squamous cell carcinoma (SCC). A late diagnosis at an advanced stage, high metastasic rate, and the development of therapy resistance are probable reasons for approximately 95% of mortality. Owing to high heterogeneity and variances in subtypes, it is important to precisely classify them for treatment. The lower airways harbour a dynamic microbial population sustained by the immigration, elimination, and local growth conditions of the lung. However, the disruption in the homeostasis of microbiome compositions was found to be correlated with the increased risk of LC. Although artificial intelligence techniques are used extensively in the early screening, and treatment of LC. Recently, the use of CT/MRI scan images in prediction models results in false-positive rates and requires subsequent tests for further exploration which delays the prognostication. Therefore, early LC diagnosis, and prognosis are critical to enhance survival and reduce death.

OBJECTIVES: We aim to classify AC and SCC subtypes of NSCLC using lung microbiome with machine learning (ML) and deep learning (DL) based algorithms for early diagnosis and prognosis to enhance survival.

METHODS: We have obtained 16S rRNA sequencing data from the NCBI database. In total, 149 AC and 145 SCC patients' samples for their lung microbiome and metadata were analyzed. The differential microbial features of the analyzed data were extracted for classifying the NSCLC subtypes. Further, various ML and DL algorithms were implemented to select the best microbiome features for subtype classification. The prediction performance of the models was evaluated for their classification potential of the selected features using 5-fold validation.

RESULTS: ML and DL-based models were developed to discriminate NSCLC subtypes based on their microbial information. Consequently, 17 features were extracted as a biomarker, and they showed good performance in distinguishing AC from SCC with an accuracy of 81% in the K-nearest neighbour and 71% in Deep neural network when demonstrated on the validation dataset.

DISCUSSION AND CONCLUSIONS: This study proposed a supervised ML framework where we can rely on taxonomic features along with ML and DL techniques to classify overlapped AC and SCC metagenomic data. This framework provides lung microbiome as a predictive and diagnostic biomarker in LC. Moreover, this framework will also be helpful to obtain further biomarkers and perform analysis of overlapped subtypes in different diseases.

A34

A COMPARATIVE ANALYSIS OF THE GENOMIC DATA GENERATED BY DIFFERENT METHODS FROM THE ARCHAEOLOGICAL SAMPLES OF TAMIL NADU, INDIA

Dheerkha Darsha Kannan[1#], Anisha Marina[1], Kaviya Rajamani[2], Baskar Ponnusamy[2], Prabakaran Murugan[2], Victor Gnanaraj[2], Thangadurai T[2], Ranjith J[2], Suresh Kaliyan[2], Paranthaman Seetharaman[2], Sakthivel K[2], Ramesh Masethung[2], Ajay Kumar Ramamoorthy[2], Kumaresan Ratha[1], Arun Raj[4], Yatheesh Kumar[4], Muthukumar V[4], Sivanantham N Ramalingam[2], Rajan K[3] & Kumaresan Ganesan[1]*

[1]Ancient Genomics Laboratory, Dept. of Genetics, Centre for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University, Madurai, India.
[2]Department of Archaeology, Government of Tamil Nadu, India.
[3]Institute of Archaeology, Department of Archaeology, Government of Tamil Nadu, India.
[4] Trichy Circle, Archaeological Survey of India, India
*Corresponding author: kumaresan@mkuniversity.ac.in
#Presenting author: kddarsha483@gmail.com; 9080511635

Investigation of ancient DNA from archaeological remnants offer the opportunity to unravel the historical mysteries related to food, agricultural, and cultural practices, as well as the complex patterns of human migration. About 10,000 ancient human genomes have been sequenced worldwide, most sequenced ancient human DNA samples have originated from temperate regions, where preservation conditions are optimal. A few genome sequences have been established from Indus Valley and Roopkund, India. From the unearthed archaeological human remains of Department of Archaeology, Government of Tamil Nadu, and ASI of Government of India, we are working to explore the genomics and to trace the migration, and cultural patterns. The excavation sites Sivagalai, Adichanallur, Kondagai, Kodumanal, and Kilanamandi of Tamil Nadu date from 300 BC to 2000 BC. From the human skeletal remains of the above sites, we have isolated DNA, constructed library and sequencing some of the libraries. For the construction of library, we have adopted various methods including ssDNA library and dsDNA library. In all the samples, out of 30-50 million reads sequenced, 1000-10000 DNA reads being human with damage. This shows the need for screening many samples as well as the need for enriching and amplifying human DNA. Also, the GC content of the constructed aDNA libraries being ~60%. The GC content of the successfully sequenced ancient human DNA from Indus valley is in the range of ~50%. The elevated GC content, minimal number of human reads, relatively lengthier insert in library all indicating the less endogenous DNA from the screened samples. Working further to enrich the ancient human DNA to maximize the possibility of reconstructing the ancient genome from Tamil Nadu. This investigation expected to decode the ancestral linkages and the early migratory patterns in ancient Tamil Nadu.

A35

COMPARATIVE ANALYSIS OF THE COMPLETE MITOCHONDRIAL GENOMES OF TWO SYMPATRIC FRESHWATER FISH *GARRA* SPP. AND ITS PHYLOGENETIC IMPLICATIONS

Sumouli Bhadra* & Saurabh Chakraborti

Department of Zoology, Bidhannagar College, Kolkata, West Bengal, India
*Corresponding author: mailto:sumoulibhadra@gmail.com; +918250761596

SUMMARY: The complete mitochondrial genomic structures of two hill-stream, cyprinid species of *Garra,* namely *G. gotyla* and *G. lamta*, currently under the threat habitat loss, have been characterized using the next-generation sequencing method (using Illumina NovaSeq 6000 Platform). Total length of the mitochondrial genomes, GC%, protein-coding genes, r-RNA genes, and t-RNA genes and relative skewness of codon usage (RSCU) have been analyzed with bioinformatics methods. Data showed an accelerated nucleotide substitution rate in these species, possibly owing to the high selective constraints to sustain in a fragile ecosystem.

BACKGROUND: This study examines and compares the mitochondrial genomic structures of two benthopelagic, sympatric, cyprinid species of *Garra* spp. – *G. gotyla* and *G. lamta* (IUCN Red List Status, LC), found in the hill streams of Eastern Himalayas. Their population is rapidly dwindling, primarily because of habitat loss, threats from human activity and climate change.

OBJECTIVES: Although studies on the morphology, ecology and taxonomy of these fish species belonging to the same genus have been conducted, extensive mitogenomic studies to understand their phylogenetic relationships are unavailable. Considering the rapid rate of habitat loss, such genomic studies are of great scientific importance.

METHODS: The specimens were collected from the Rangeet River in Jorethang Town (27.13° N, 88.27° E), Namchi District, Sikkim, India. After collection the specimens were fixed and stored in 99.9% ethanol. The steps involving mitochondrial DNA isolation, library preparation for next generation sequencing (Illumina NovaSeq 6000 Platform), and assembly were carried out at the Genotypic Technology Pvt. Ltd. Bangalore, India. Subsequently various bioinformatic analyses were performed using free software. The complete mitochondrial sequences were deposited in NCBI GenBank.

RESULTS: The lengths of the mitogenome sequences of *G. gotyla*, and *G. lamta* are 17074bp and 16854bp respectively. Both the genome contains 13 protein-coding genes, two ribosomal RNAs, and 22 transfer RNA genes. The GC% of *G. gotyla* and *G. lamta* are respectively 40.9% and 41.2%. Nucleotide frequency of *G. gotyla* and *G. lamta* is G (15.1%), C (25.8%), A (32.7%), T (26.4%) and G (15.3%), C (25.9%), A (32.5%), T (26.3%).

DISCUSSION AND CONCLUSION: Due to the high selective constraints to thrive in a fragile ecosystem, these specimens showed an accelerated nucleotide substitution rate. Further investigations may contribute to our understandings of the evolutionary history, phylogeography, and conservation of these species.

A36

COMPARATIVE MITOGENOMICS OF SUBTERRANEAN LOACH, *PANGIO BHUJIA*, PROVIDE INSIGHTS INTO ITS PHYLOGENETIC POSITION AND MOLECULAR SELECTION

Tania Ann Varghese[1], Remya L. Sundar[2], Arya Sidharthan[3], Rajeev Raghavan[2] & Neelesh Dahanukar[1]*

[1]Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India.
[2]Faculty of Ocean Science and Technology, Kerala University of Fisheries and Ocean Studies (KUFOS), Kochi, India
[3]Department of Fisheries Resource Management, Kerala University of Fisheries and Ocean Studies (KUFOS), Kochi, India
*For correspondence: neelesh.dahanukar@snu.edu.in

BACKGROUND: The subterranean eel-loach, *Pangio bhujia*, inhabiting the lateritic aquifers in Kerala, India, shows multiple morphological adaptations such as extremely reduced eyes, absence of pigmentation, absence of dorsal and pelvic fins, and well-developed sensory system, in response to its underground mode of life. Understanding the genomic basis of such evolutionary adaptations can provide new insights into evolution of forms and functions in vertebrates. As a preliminary analysis, we focus on the evolution of mitochondrial genome to understand the possible adaptations with respect to oxidative phosphorylation in subterranean fishes.

OBJECTIVES: Construction of complete mitogenome of *Pangio bhujia* using Next Generation Sequencing and its comparative genomics to understand phylogenetic position of the species and molecular selection on protein coding genes.

METHODS: Mitogenome of *Pangio bhujia* was sequenced using Illumina short read sequencing chemistry. Mitogenome was constructed using MitoZ 3.6. Mitogenome was annotated using MitoAnnotator and circular map of the mitogenome was constructed using Chloroplot. Additional mitogenome sequences were downloaded from NCBI GenBank. Protein coding sequences were extracted using EZmito, concatenated and a maximum likelihood phylogeny was performed using IQ-TREE2. Time of divergence was estimated using three calibration points, employing RelTime implemented in MEGA 11. Molecular selection analysis was done in Datamonkey using adaptive Branch-Site Random Effects Likelihood (aBSREL), RELAX and Mixed Effects Model of Evolution (MEME) methods.

RESULTS: In the molecular phylogeny, *Pangio bhujia* was recovered as a sister to *Lepidocephalichthys*, questioning its taxonomic status as *Pangio*. Time to the most common recent ancestor of *P. bhujia* and *Lepidocephalichthys* was 21.3 million years ago in Aquitanian age of lower Miocene. Molecular selection analysis revealed that the branch of *P. bhujia* has undergone episodic diversifying selection ($\omega1 = 0.00654$, $\omega2 = 0.281$, $\omega3 = 520$, $P = 0.0011$). Detailed site wise selection analysis showed that COXI, ND2, ND4 each had four codons under episodic positive diversifying selection, COXII and ND6 had three, Cytb and ND1 had two, while ATPase6, ATPase8 and ND4 had one codon each. In addition, we found that there was also selection for relaxation ($K = 0.38$, $P < 0.0001$, $LR = 79.12$), which could likely explain why there were diversifying selection.

DISCUSSION: The sister group relationship of *P. bhujia* with *Lepidocephalichthys* suggests that the classification of this taxon may require revision. Our analysis indicates that *P. bhujia* probably became adapted to subterranean habitats in the Miocene, which correlates with the major droughts that took place in southern India. Molecular selection analysis suggests that after entering subterranean habitats, the mitogenome of *P. bhujia* might have experienced significant selection relaxation as a response to environmental conditions, which might have led to episodic diversification in ten out of its thirteen

protein coding genes which are responsible for oxidative phosphorylation. Further understanding of the whole genome sequencing might provide better information about the genomic basis of evolutionary adaptation in this unique subterranean fish.

CONCLUSIONS: We provide the first complete mitogenome of subterranean eel-loach, *Pangio bhujia* and provide its phylogenetic position within family Cobitidae. Our analysis of protein coding genes suggests that there is episodic diversifying selection in its oxidative phosphorylation genes.

A37

DECODING THE MITOGENOME OF *TRAVANCORIA ELONGATA*, AN ENDEMIC AND THREATENED MOUNTAIN LOACH FROM THE WESTERN GHATS

Arya Sidharthan[1], Rajeev Raghavan[2] and Neelesh Dahanukar[3]

[1]Faculty of Ocean Science and Technology, Kerala University of Fisheries and Ocean Studies (KUFOS), Kochi, India
[2]Department of Fisheries Resource Management, Kerala University of Fisheries and Ocean Studies (KUFOS), Kochi, India
[3]Department of Life Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
For correspondence: [1]aryaanandhu06@gmail.com; [2]rajeevraq@hotmail.com

BACKGROUND: *Travancoria elongata* is an endangered mountain loach (Family Balitoridae), endemic to the Western Ghats, where it is restricted in distribution to two small-west flowing rivers (Chalakudy and Periyar) in the state of Kerala. *Travancoria elongata* is one of the two species within the genus, the other being *T. jonesi*. The species has a unique morphology, and use their whole body as a hydrodynamic adhesion device for clinging on to rocks and pebbles in swift-flowing streams. This study aims to unravel the mitochondrial genome of *Travancoria elongata* to better understand its evolutionary and ecological adaptations to torrential, mountain streams.

OBJECTIVES: To decode the complete mitogenome of *Travancoria elongata* using NGS, and use this information to determine its phylogenetic relationships, and molecular selection pressures on the protein coding genes.

METHODS: Fin clips were extracted from a sample of *Travancoria elongata* collected from the Chalakkudy River. Genomic DNA was isolated using the Qiagen DNAeasy Blood and tissue kit and the mitogenome was sequenced using NovaSeq 600 (Illumina). The genome was subsequently constructed using MitoZ 3.6, annotated using MitoAnnotator, and circular map of the mitogenome prepared using Chloroplot. Additional mitogenome sequences representing members of the family balitoridae were downloaded from NCBI GenBank. Protein-coding sequences were extracted using EZmito, and they were concatenated to generate a maximum likelihood phylogenetic tree in IQ-TREE2. Selection analysis was performed in Datamonkey using adaptive Branch-Site Random Effects Likelihood (aBSREL) and RELAX methods. Codons under diversifying selection were identified using FEL (**F**ixed **E**ffects **L**ikelihood) method which uses a maximum-likelihood (ML) approach to infer non-synonymous (dN) and synonymous (dS) substitution rates on a per-site basis for a given coding alignment and corresponding phylogeny

RESULTS: The phylogeny based on mitogenomes recovered *T. elongata* into a well-resolved clade of balitoridae with another Western Ghats endemic species, *Bhavania australis* as its sister taxon. The branch of the two Western Ghats endemic balitorids, *T. elongata* and *B. australis* were under episodic diversifying selection (optimized branch length = 0.0157, Likelihood ratio test = 6.8117, P = 0.0118, $\omega1 = 0.0297$ (99%), $\omega2 = 100000$ (0.64%)). Further, the test for selection relaxation on the branch of balitorids endemic to Western Ghats ($K = 0.61$) was significant ($p = 0.007$, LR = 7.38). Site wise selection analysis using FEL analysis revealed that on the branch of Western Ghats endemic lineages, 136 sites were under purifying selection ($p \leq 0.1$), while 22 sites were under diversifying positive selection ($p \leq 0.1$), and included nine codons in ND5, three codons in ND2, two codons in ATPase6, ND1 and ND4, and one codon each in cox3, cytb, ND3 and ND6.

DISCUSSION: Members of family Balitoridae are rheophilic and can withstand extreme water turbulence in the stream rapids through their adhesion devices. As a result, we hypothesize that the genes involved in energy production will be under selection based on the unique habitat of the species. Since the mitochondrial protein coding genes are responsible for oxidative phosphorylation, it is not

surprising that our analysis indicates various selection pressures acting on them. While the purifying selection is important for purging deleterious and potentially harmful mutations, the overall relaxation on the purifying selection on the branch leading to the Western Ghats endemic lineages could be a result of unique mountainous habitat of the species. The relaxation in selection can possibly explain the episodic diversifying selection in mitochondrial genes, which might have helped the species in occupying high elevation, cold-water habitats in the Western Ghats of India. Molecular selection acting on the mitogenome, however, provides only limited understanding of how the energy metabolism of the species might have been shaped by the environmental clues, and finer resolution understanding will require the deciphering of the whole genome of the species.

A38

META-BARCODING OF PELLETS UNRAVELS THE FORAGING PREFERENCES OF SPOTTED DEER (*AXIS AXIS* ERXLEBEN); A CASE STUDY FROM TAMIL NADU

Rajendran Harani, Gnanasekar Swetha, James J. Amuldoss & Umapathy Senthilkumar*

Division of DNA Barcoding Laboratory, Department of Botany, Madras Christian College, East Tambaram, Chennai, India
*Corresponding author: senthilumapathy@mcc.edu.in

DNA barcoding, a tool that obtains species-specific DNA signatures based on the simple premise within small stretches of the organism's genome. It can provide a ''biological barcode'' to identify any organism at the species level. It is believed that DNA barcoding, will provide a "universal key" that will allow the identification of a species by running unknown DNA sequences through a DNA barcode database such as BOLD, NCBI etc... Further, "barcoding", has helped clarify taxonomic position of an 'Apparent species complex' by revealing several cryptic species within a 'single' species described through conventional taxonomy. This study aims to understand the grazing preferences of Spotted Deer within the Madras Christian College campus by analysing the pellets using a Meta-barcoding approach. DNA markers have greater fidelity to unravel the diets even if the plant materials have been processed through the gut. This study has been identified the niches (foraging grounds) for spotted deer, hence the demarcation can be derived to avoid niche overlaps and animal conflicts within the campus. Thus, this study can be established further as a model to implement the methods and protocols for the first time in India to understand such patterns amongst other herbivores such as endemic-Nilgiri Tahr, Asian Elephants, Blackbuck deer, etc. towards the conservation and restoration of the foraging grounds.

Keywords: Diet analysis, eDNA, Metabarcoding, Spotted Deer, Tropical Dry Evergreen Forests (TDEF)

A39

COMPLETE MITOCHONDRIAL GENOMES OF 5 REEF-FORMING ACROPORA CORALS FROM OMAN COAST

S. Senthil Kumar[1] & Sreenath Chandran[2]

Department of Applied Biotechnology, (University of Technology and Applied Sciences, Sur) Sultanate of Oman.
*For correspondence: [1]senthil.kumar@utas.edu.om; [2]sreenath.chandran@utas.edu.om

SUMMARY: This study investigates the complete mitochondrial genomes of five reef-forming Acropora corals found along the coast of Oman. The research aims to understand the genetic makeup of these corals and its implications for their conservation and management. DNA extraction, sequencing, and bioinformatics analysis were conducted to assemble and annotate the mitochondrial genomes. Preliminary analysis revealed a high degree of conservation in gene content and order among the studied species, with variations observed in non-coding regions. Phylogenetic analysis provided insights into evolutionary relationships and genetic diversity. The findings contribute to the field of coral genomics and offer valuable resources for studying population genetics, adaptation, and responses to climate change in these corals. Conservation strategies can be informed by these findings to protect the ecologically important Acropora corals and their associated reef ecosystems along the Oman coast.

BACKGROUND: Coral reefs are highly diverse and ecologically important ecosystems, providing habitat for numerous marine species. The Acropora genus is a key reef-building coral group, known for its significant contribution to coral reef structure and function. Understanding the genetic makeup of Acropora corals is crucial for their conservation and management, particularly in the face of increasing threats such as climate change and anthropogenic activities. This study focuses on investigating the complete mitochondrial genomes of five reef-forming Acropora corals found along the coast of Oman.

OBJECTIVES: The aim of this research is to analyze the complete mitochondrial genomes of five reef-forming Acropora corals from the Oman coast. The objectives include: (1) Perform DNA extraction and sequencing on coral samples collected from various sites along the Oman coast. (2) Assemble and annotate the mitochondrial genomes using bioinformatics tools and databases. (3) Compare the assembled genomes with known mitochondrial genomes of other Acropora species to identify similarities and differences. (4) Conduct phylogenetic analysis to understand the evolutionary relationships and genetic diversity among the studied Acropora corals and related species. (5) Discuss the implications of the findings for the conservation and management of these ecologically important corals along the Oman coast.

METHODS: Coral samples were collected from different sites along the Oman coast. The collected samples were preserved immediately with 70% ethanol and carefully transported to the laboratory for further process. Raw data quality assessment was performed using FastQC v.0.11.9 (default parameters). The data was preprocessed using Fastp v.0.20.1(parameters: -f 5 -q 30 -l 50 -c). Post filtering data was further assessed using FastQC. Post and pre-filtered cleaned data were summarised using MultiQC. Acropora digitifera (NC_022830.1) and Acropora millepora (MT593341.1) mitochondrion, complete genomes were used as a seed sequence to assemble the preprocessed reads using NOVOPlasty v.4.3.1 (default parameters) Annotations and circular mitogenome plots were done using Prokka implemented in Proksee online server. Comparative analysis was conducted with known mitochondrial genomes of other Acropora species. Phylogenetic analysis was performed to understand the evolutionary relationships and genetic diversity.

RESULTS: Preliminary analysis of the complete mitochondrial genomes revealed a high degree of conservation in gene content and order among the five studied Acropora coral species. Variations were observed in the non-coding regions, indicating potential differences in regulatory elements and adaptive

traits. Phylogenetic analysis provided insights into the evolutionary relationships and genetic diversity among the studied corals and related species.

DISCUSSION: The high conservation in gene content and order among the studied Acropora corals suggests a strong evolutionary relationship and functional importance of these genes. The variations observed in the non-coding regions may be indicative of potential differences in adaptive traits and responses to environmental changes. The phylogenetic analysis provides a deeper understanding of the evolutionary history and genetic diversity of these corals.

CONCLUSIONS: The complete mitochondrial genomes of the five reef-forming Acropora corals from the Oman coast contribute to the growing field of coral genomics. These genomes serve as valuable resources for studying population genetics, adaptation to environmental changes, and potential responses to climate change in these corals. The findings have important implications for the conservation and management of these ecologically significant corals and their associated reef ecosystems along the Oman coast. Conservation strategies can be informed by the genetic information obtained from this study to protect these corals and their habitats.

A40

GENOMIC INSIGHTS INTO EVOLUTIONARY ADAPTATIONS TO BROOD PARASITISM IN KOEL: PRELIMINARY OBSERVATIONS FROM COMPARATIVE MITOGENOMICS

Devina Ghose[1#], Shagnik Saha[2], Ekta Gupta[2], Manaswini Parija[1], Sanjeev Galande[2] & Neelesh Dahanukar[1]*

[1]Molecular Ecology and Evolution Laboratory, Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
[2]Center of Excellence in Epigenetics, Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
*Corresponding authors: neelesh.dahanukar@snu.edu.in
#Presenting author: dg551@snu.edu.in

BACKGROUND: Numerous cuckoo species within the Cuculiformes order are well-known for their brood parasitism, a strategy where they deposit eggs in the nests of other birds, leaving the host raise their offspring. The evolutionary success of such reproductive strategy hinges on the brood parasite acquiring specific traits that enable their young to mimic the physical and chemical clues for seamlessly blending into the host's brood. Additionally, the parasitic offspring must exhibit higher metabolic activity to effectively exploit host resources while soliciting food and attention. Understanding the genomic basis of evolutionary adaptation to brood parasitism can be gained through a comprehensive genome-wide analysis of molecular selection acting on genes in both parasitic and non-parasitic members of the Cuculiformes order. This study focuses on Koel as an illustrative example to unravel the genomic basis of brood parasitism. The Koel is particularly intriguing due to its diverse array of host species and its unique adult diet exclusively consisting of fruits. Although the complete assembly of the whole genome of Koel, sequenced using both short and long read chemistries, is currently underway, this work presents a fully annotated mitogenome of the species. The exploration of molecular selection is approached through comparative mitogenomics, raising questions about the genomic adaptations associated with brood parasitism.

OBJECTIVES: (1) Whole genome sequencing of *Eudynamys scolopaceus* (common name Koel) using short and long read high throughput sequencing chemistries. (2) Assembly and annotation of complete mitogenome of Asian Koel. (3) Molecular phylogeny and evolutionary selection analysis based on mitochondrial coding genes of Cuculiformes.

METHODS: High molecular weight DNA was extracted using MagAttract HMW DNA Kit and sequenced using long read (Nanopore) and short read (MGI) high throughput sequencing chemistries. Initial assembly of Nanopore reads was performed using Flye, whereas initial assembly of MGI reads was performed using MEGAHIT. Whole mitogenome was culled out from the MEGAHIT assembly, annotated using MitoAnnotator and visualized using Chloroplot. Phylogenetic analysis was performed in IQTREE2 and ancestral traits in phylogeny were determined using Bayestraits. Molecular section analysis was performed in Datamonkey.

RESULTS AND DISCUSSION: Preliminary assembly of the whole genome of Koel suggested that the genome of the species is about 1.1 Gb with a 41.90% GC content, which is consistent with other cuckoo genomes available in NCBI database. The whole mitochondrial genome is 17,084 bp with 43% GC content and encodes 13 protein coding genes (PCGs), two rRNAs and 22 tRNAs. Examining the 13 protein coding genes (PCGs) in mitochondria, we observed that obligate brood parasitic cuckoos formed a monophyletic clade, indicating a likely single evolution of obligate parasitism within this group of organisms. Ancestral state reconstruction suggested that the basal clade of cuckoos was non-brood parasitic, but the parasitic strategy likely emerged early in the group's evolution and was subsequently lost in some branches while retained in others. Selection analysis revealed a more

pronounced selection pressure (K = 1.13, P = 0.04, LR = 4.21) on the PCGs in brood parasitic species compared to their non-parasitic counterparts, indicating strong purifying selection that has preserved the mitogenomes of parasitic cuckoos. However, concerning other parasitic cuckoos, the Koel exhibited significant relaxation (K = 0.71, P < 0.001, LR = 61.31), as reflected in 22 codons under diversifying positive selection, including one each in the *cytb*, *ATPase6* and *ATPase8* genes and 19 codons in various subunits of ND genes. We hypothesize that the extensive diversity of host species may have influenced the molecular selection in the Koel.

CONCLUSIONS: Our analysis indicates that the mitogenome, while initially thought to have a narrow role in the evolutionary adaptation to brood parasitism, exhibits molecular selection on its genes and codons. This implies that the brood parasitic strategy has played a role in shaping the selection pressures on the mitogenome, especially in the case of the cuckoos and, notably, the Koel. Consequently, a comprehensive whole-genome analysis of molecular selection is expected to offer intriguing insights into evolutionary adaptations associated with brood parasitism in the Koel.

A41

EXPLORING THE GENETIC BASIS OF SEED LONGEVITY IN WHEAT (*TRITICUM AESTIVUM* L.) – A GENEBANK PERSPECTIVE

Sampa Saha[1]*, Sherry Rachel Jacob[1], Abhijith KP[2], Arun Kumar MB[2], Aradhana Mishra[1], J. Aravind[1], Sundeep Kumar[1], Jyoti Kumari[1], Sudhir Navathe[3] & Yashvantha Kumar[3]

[1]ICAR-National Bureau of Plant Genetic Resources, New Delhi, India
[2]ICAR-Indian Agricultural Research Institute, New Delhi, India
3Agharkar Research Institute, Pune, Maharashtra, India
*Corresponding author: sampasaha777@gmail.com; 8617566567

A Genome-Wide Association Analysis was conducted to understand the mechanism behind wheat seed longevity and to identify the genomic regions associated with component traits of seed longevity using 104 wheat Germplasm accessions conserved in National Genebank, New Delhi and multiplied at ARI, Pune. Parameters for phenotypic evaluation included total antioxidant capacity, GSH-GSSG redox potential and mean germination time (MGT) after artificial ageing following two procedures-Accelerated Aging (AA) and Controlled Deterioration (CD). Genotyping was done with Axiom® bread wheat 35K array and 19090 SNPs were used for GWAS. The study identified ten significant MTAs with major candidate genes, for seed longevity. In K- net miner analysis, ATRAESCS5A03G0963300 gene was found on chr 5 which is an orthologue of *Arabidopsis Ats*1T1 gene that leads to accumulation of ROS, growth inhibition and cell death. On Chr 4B, TRAESCS4B03G0930800 gene was found, which is an orthologue of *Arabidopsis* AtRBOHB and rice OsRBOHB genes that modulates ROS accumulation and PCD. Candidate gene identified on chr 7B belongs to co-expression cluster of biological ageing process and that on chr 3A, has contributory role in ROS accumulation. This study complements the reports of other genome association studies that identified multiple loci and confirms the complex polygenic nature of the trait. The results also showed positive correlation with actual genebank conservation data. This information generated in this study could be of potential value for streamlining genebank conservation protocols and also can be utilized for improving the genetic base of seed longevity traits, using marker assisted selection.

Keywords: Wheat, gene bank, seed longevity, GWAS

A42

EXPLORING THE CRRLK1L GENE FAMILY IN WHEAT: INSIGHTS INTO DEVELOPMENT AND STRESS RESPONSE

Nilesh D. Gawande* & Subramanian Sankaranarayanan

Department of Biological Science and Engineering, Indian Institute of Technology Gandhinagar, Gujarat, India
*Corresponding author: nilesh.g@iitgn.ac.in

Catharanthus roseus receptor-like kinase 1-like (CrRLK1L) genes encode a subfamily of receptor-like kinases (RLK) that regulate diverse processes during plant growth, development, and stress responses. The first CrRLK1L was identified from the *Catharanthus roseus* Madagascar periwinkle, commonly known as Madagascar periwinkle. Subsequently, CrRLK1L gene families have been characterized in many plants. The genome of *T. aestivum* encodes 15 *CrRLK1L* genes with 43 paralogous copies with three homeologs each, except for -*2*-D and -*7*-A, which are absent. Chromosomal localization analysis revealed a markedly uneven distribution of *CrRLK1L* genes across seven different chromosomes, with chromosome 4 housing the highest number of genes, while chromosome 6 lacked any *CrRLK1L* genes. Tissue-specific gene expression analysis revealed distinct expression patterns among the gene family members, with certain members exhibiting increased expression in reproductive tissues. Gene expression analysis in response to various abiotic and biotic stress conditions unveiled differential regulation of gene family members. The analysis of *cis*-acting elements in the promoter regions identified specific elements crucial for plant growth and developmental processes. This comprehensive genome wide analysis and expression study provides valuable insights into the essential functions of CrRLK1L members in wheat.

A43

DE NOVO ASSEMBLY AND ANNOTATION OF *CAESALPINIA BONDUCELLA* L. SEED TRANSCRIPTOME IDENTIFIES KEY GENES INVOLVED IN THE BIOSYNTHESIS OF BONDUCELLIN, A HOMOISOFLAVONOID

Balasundaram Usha* & Kandasamy Veerapandiyan

[1]Department of Genetic Engineering, SRM Institute of Science and Technology, Tamil Nadu, India.
*Corresponding author: sundaram.usha@gmail.com; ushab@srmist.edu.in

Caesalpinia bonducella L. is a traditional medicinal plant containing a potential homoisoflavonoid, bonducellin, with therapeutic values against polycystic ovary syndrome, oxidative damage, pathogenic bacteria, irregular menstrual cycle, ovarian cancer and diabetes. Owing to the multi-therapeutic properties of bonducellin, knowledge of its biosynthetic pathway genes will help understand its regulatory mechanism and thus improve the yield. This study sequenced *C. bonducella* seed mRNA transcriptome to identify the genes in bonducellin biosynthesis. Before this, the presence of bonducellin in the seed samples was analysed by HPLC using the chemically synthesized bonducellin as the standard. Seven key genes encoding enzymes involved in the synthesis of bonducellin via the phenylpropanoid pathway were identified. The expression of selective genes from the bonducellin biosynthetic pathway was validated using qRT-PCR and was comparable with RNA-Seq data. Here, we put forth the sequences of 67,560 genes from *C. bonducella* and highlight the bonducellin biosynthetic pathway genes.

A44

COMPARATIVE ASSESSMENTS OF GENE FAMILIES AMONG FIBER-PRODUCING SPECIES

Hari Shankar Gadri[#], Sarbani Roy[#] & Pankaj Bhardwaj*

Molecular Genetics Lab, Department of Botany, Central University of Punjab, Bathinda, India
*Corresponding author: pankajbhardwaj@cup.edu.in
#Presenting authors: haribotany1994@gmail.com, 9929889204;
sarbani.rumu@gmail.com, 7978738207

In plant taxa, several gene families underlying significant traits exhibit structural variation, thus resulting in species diversification. Here, it has been discussed that the fiber-producing species belong to different families. Comparative gene family evaluation permits us to reveal the evolutionary history of species by genomic information. High-quality transcriptomics of twenty-one fiber-producing species were retrieved from a public repository and were examined. We annotated the protein-coding genes to assess the comparative gene families across the fiber-producing species. The study unveiled essential regulatory gene families, Cellulose synthase A, Expansin, Beta-glucosidase, etc., aid in fiber development. Further, a thorough gene family analysis established the phylogenetic relationships among the fiber-producing species.

Keywords: Gene family, Transcriptomics, fiber-producing species, phylogenetic relationship.

A45

B-BOX 20 GENE EXPRESSION AS AN EARLY ADAPTATION TO UV-B TOLERANCE IN *DUNALIELLA* SP.

Vasantha G, Priyanka V & Dronamraju VL Sarada*

Department of Biotechnology, School of Bioengineering, SRM Institute of Science and Technology, Tamil Nadu, India
*Corresponding author: saradad@srmist.edu.in; +91-9884825471

There has been a growing demand for nutraceuticals, food supplements, and skin care products of natural origin over synthetic ones. Carotenoids are synthesized by several algae as accessory pigments that perform various functions, including photosynthesis, pigmentation and maintenance of redox homeostasis. The synthetic carotenoid market value is estimated about $1.84 billion by 2027 and $2.26 billion is the estimated market value of natural carotenoids by 2030. This research aims to increase the carotenoid content by elicitation of a carotenoid biosynthetic pathway by UV-B signaling transduction pathway. *Dunaliellas sp.* was isolated from water samples collected from various sources and was exposed to UV-B radiation for a definite time intervals. The strain was analyzed for by gene expression of Box X-20 key gene being involved in activating the rate determining gene of carotenoid biosynthetic pathway gene phytoene synthase (PSY). BBX20 is known to bind to the G-Box promoter of the PSY gene which would be involved in increasing the carotenoid content in cell. The study would improve our understanding of the UV-B signaling responses and might pave way for enhancement of carotenoids in microalgae.

Keywords: Microalgae, UV-B radiation, BBX-20, Carotenoids.

A46

EXPRESSION ANALYSIS OF RALF FAMILY IN DROUGHT-STRESSED *ORYZA SATIVA*

Durvasula Sumana Pratyusha, Vidyasagar G & Dronamraju VL Sarada*

Department of Biotechnology, School of Bioengineering, SRM Institute of Science and Technology, Tamil Nadu, India
*Corresponding author: saradad@srmist.edu.in; +91-9884825471

Oryza sativa is one of the prominent cereals with high nutritional value and is consumed by around 60% of the world's population. The yield and productivity of the crop are affected by various abiotic stresses specifically drought. In response to stress signals, plants secrete small peptides involved in abiotic stress responses and rapid alkalinization factor (RALF) is one of these peptide families. RALF, a cysteine-rich peptide is processed in response to stress and regulates phytohormone-based growth regulation. The nucleotide sequences of the RALF family were retrieved from RAP-DB (Rice Annotation Protein – Database) and Phytozome. Further, the gene ontology analysis for functional enrichment, prediction of subcellular location, determination of cis-regulatory elements, and transcription factor binding analysis were performed to shortlist 20 RALFs that could be involved in drought. The standardization of the concentrations of PEG6000 to mimic drought was determined by analyzing several morphological, physiological, and biochemical parameters and the gene expression of RALFs was analyzed by performing RT-PCR with a PEG-based drought model. The effect of the exogenous application of the phytohormones (ABA, IAA, BR) on the expression of RALFs and their receptors will be analyzed by performing qRT-PCR. We could classify the RALFs that can be upregulated and downregulated during drought and their role in hormonal regulation during water-deficit conditions.

***Keywords**:* Peptides, Drought, RALF, Phytohormones, Gene expression.

A47

CRISPR/CAS9 MEDIATED RESISTANCE AGAINST COTTON LEAF CURL MULTAN VIRUS

Rahmat Khan[#], Ankit Singh & Jawaid A. Khan*

Jamia Millia Islamia, New Delhi, India
*Corresponding author: jkhan1@jmi.ac.in
[#]Presenting author: rahmatkhan4mail@gmail.com; 9520572486

In recent decades begomovirus complex limited the production of cotton in the Indian subcontinent by causing CLCuD. These Begomovirus belong to family *Geminiviridae*, having circular, single-stranded DNA in association with symptom modulating betasatellite molecule. A number of short term (mainly management practices) and long-term strategies (developing resistant cotton varieties) have been devised to control the disease. In long term strategies such as breeding, considerable efforts have been made to develop cotton varieties resistant to CLCuD. However, there has been no comprehensive assessment of the resistance in cotton varieties. Development of transgenic cotton using both pathogen and non- pathogen derived approaches are in progress. Various *in vitro* and *in vivo* studies find the potential of different site-specific nucleases to target viral genome and generate resistant in plant system. Among these CRISPR/Cas9 system has been proved very cost effective and convenient. In prokaryotes it evolved as molecular immunity against invading phages and plasmids. By time natural CRISPR/Cas9 system engineered and successfully used in eukaryotic system for instant providing molecular immunity to plants against invading viral nucleic acid. In our study we designed a construct targeting the genome of CLCuMuV by over lapping extension PCR method, then cloned and mobilized it into *Agrobacterium tumefaciens*. *Gossypium hirsutum* explants were transformed with agrobacterium harbouring the desired construct. Transformed plants confirmed by PCR, challenged with viruliferous whitefly and further subjected to various biological and molecular studies.

Keywords: CLCuMuV, Begomovirus, CRISPR/Cas9, transgenic, PCR

A48

ENGINEERING CRISPR/CAS SYSTEM-MEDIATED RESISTANCE AGAINST BEGOMOVIRUSES IN *GOSSYPIUM HIRSUTUM*: HARNESSING MOLECULAR SCISSORS FOR CROP PROTECTION

Divya Sharma[#], Rashmi Rishishwar and Jawaid A. Khan*

Molecular Virology Laboratory, Department of Biosciences, Jamia Millia Islamia, New Delhi, India
*Corresponding author: jkhan1@jmi.ac.in
[#]Presenting author: divyas750@gmail.com; 7827454917

Gossypium hirsutum (cotton) has been a crucial cash crop for many countries, and its cultivation and trade have played a pivotal role in shaping the Indian economy. The incessant threat of Cotton leaf curl disease (CLCuD) in cotton poses a significant challenge to global cotton production. CLCuD is induced by begomoviruses (family *Geminiviridae*) in association with satellite molecules and are transmitted by the whitefly vector (*Bemisia tabaci*). Various conventional and non-conventional methods have been applied to control these viruses. CRISPR/Cas genome editing is based on innate immunity against viruses found in bacteria and archaea. Recent studies have shown that CRISPR/Cas system can be utilized to develop virus resistance and reduce the probability of off-targets in genome editing. This study explores the potential of the CRISPR/Cas system as a targeted and efficient tool for conferring resistance against begomoviruses infecting cotton. Here, a CRISPR/Cas9 system equipped with dual guide RNAs that strategically target the single-stranded DNA genome of *Cotton Leaf Curl Multan Virus* was employed. *G. hirsutum* cv. HS6 plants were transformed following *Agrobacterium*-mediated transformation. Cotton seeds were surface sterilized and germinated on seed germination media. Shoot tip explants were isolated from 10-12 days old cotton seedling, precultured and soaked in *A. tumefaciens* strain LBA4404 culture with MS, followed by co-cultivation for 48 hrs. After co-cultivation, the explants were washed and transferred to selection medium. Regular sub-culturing of explants on new selection media was carried out weekly. Once rooting was achieved, the plantlets were transferred to hoagland medium for acclimatization. Following successful acclimatization, the plantlets were shifted to pots containing a mixture of peat moss and soil, and covered with polybags to create an optimal humidity. Finally, the plants were transferred to a greenhouse setting, providing a controlled environment for further growth and development. Transgene detection was done using PCR. The evaluation of the transformed plants is currently underway, involving a combination of biological and molecular methodologies, mutation detection, etc.

***Keywords**:* CLCuD, begomovirus, whitefly, CRISPR/Cas, virus resistance.

A49

UNRAVELLING THE CROSSTALK BETWEEN DIFFERENTIAL METHYLATION UNDER SALINITY STRESS AFFECTING GENE EXPRESSION IN CHICKPEA

Namra Ali[#] & Rohini Garg*

Plant Gene Regulatory Lab, Centre of Excellence in Epigenetics, Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
*Corresponding author: rohini.garg@snu.edu.in
[#]Presenting author: na815@snu.edu.in

Chickpea is one of the important crop plants with high amount of protein making it an alternate source of vegan protein along with vitamins, good sterols, and dietary fibres. This food crop is highly affected under abiotic stress. Salinity is one of the most affecting abiotic stresses contributing to yield loss. Plants combat detrimental effects of salt stress by changing plant physiological processes which are regulated by altered gene expression under stress conditions. Dynamics of gene regulatory machinery can be regulated by multiple regulatory pathways involving signalling and epigenetic mechanisms. There are many types of epigenetic changes that occur in plants, DNA methylation is one such epigenetic modification that could regulate gene expression under changing environmental cues to modulate plant responses. DNA methylation occurs at the cytosine base of the DNA strand in three different sequence contexts (CG, CHH and CHG) in plants, each one having a specific effect on gene expression based on its localization with respect to different genic features, thereby making plants to have a broader paradigm of methylation regulated genes under salt stress. Our lab has previously identified differentially methylated and differentially expressed genes under salinity stress in stress sensitive and tolerant genotypes. I will present our results on differential methylation and differential gene expression of selected candidate genes (involved in root development) in both the genotypes. I will further highlight the effect of demethylating agent on DNA methylation and gene expression of these genes in both sensitive and tolerant genotypes to study the crosstalk between DNA methylation and salinity stress.
Keywords: DNA methylation, chickpea, gene expression, salinity stress

A50

DEVELOPMENT OF CLCUD RESISTANT *G. HIRSUTUM* PLANTS USING ARTIFICIAL MIRNA

Masnad Khan[#] & Jawaid A. Khan*

Molecular Virology Laboratory, Department of Biosciences, Jamia Millia Islamia, New Delhi, India
*Corresponding author: jkhan1@jmi.ac.in
[#]Presenting author: khanmasnad1994@gmail.com; 9452120093

Cotton leaf curl disease (CLCuD) is a serious threat to cotton production worldwide and causes severe losses in north-western India. It is caused by *Begomovirus* (family Geminiviridae) in association with satellite molecules namely betasatellite and alphasatellite. They are exclusively transmitted by the insect vector, *Bemisia tabaci* (whitefly). The genome of monopartite begomoviruses consists of a circular, single-stranded DNA molecule. It is organized into 7 ORFs with C1-C5 on the complementary sense strand and V1-V2 on the virion sense strand. The conventional strategies to control viral infection have certain limitations. Currently, the amiRNA-based approach has been employed to suppress virus infections in crops. This approach uses a natural precursor miRNA backbone in which mature miRNA sequences are replaced with designed artificial miRNA (amiRNA) sequences that can target specific genes utilizing the same processing mechanisms as natural miRNAs. Present study aims at generating amiRNA-mediated resistance in *G. hirsutum* plants against genome of CLCuMV. Potential amiRNAs were designed using a computational approach and cloned into a plant transformation vector. The developed plant transformation vector carrying amiRNA construct was mobilized into *Agrobacterium tumefaciens* and *G. hirsutum* explants were transformed. The presence of transgenes in cotton transformants was confirmed by PCR analysis. amiRNA expression level and their antiviral resistance in transformed plants were studied. These results suggest that amiRNA effectively downregulate viral DNA accumulation in cotton plants.

A51

REGULATION OF GENE EXPRESSION ACROSS ECOTYPES OF WOOD TRANSCRIPTOME IN *TECTONA GRANDIS*

Maheswari Patturaj, Nithishkumar Kannan, Balasubramanian Aiyar & Ramasamy Yasodha*

Institute of Forest Genetics and Tree Breeding, Coimbatore, India
*Corresponding author: yasodha@icfre.org

In woody species, the adaptive response of trees with changing environment conditions will influence the regulation of wood formation. The shift in gene expression can be studied using transcriptome which will unveil about the variation in expression among different ecotypes. In the present study, the transcriptomic changes of the wood core samples across two ecotypes in *Tectona grandis* was investigated using the RNA-seq method. The pair-end transcriptome profiling generated 6GB data containing 48,503,292 raw reads, after removing contaminants, the clean reads were subjected to reference-based transcriptome assembly generation. A total of 51,956 transcripts were annotated and their functions were identified using databases of UniProtKB, GO, KEGG, PlantTFDB and iTAK. The goal of this work is to discover the differentially expressed transcripts and the regulation of the wood related genes across ecotypes. From the 31,680 annotated transcripts, a total of 25,230 transcripts were brought forward for differential expression analysis and co-expression network construction. Two ecotype DE analyses, resulted in 1530 differentially expressed genes (DEGs) of which almost 730 genes were up-regulated and 800 were down-regulated. GO and KEGG enrichment analysis of DEGs revealed that the metabolic processes of lignin, phenylpropanoid were enriched, and many biosynthetic pathways such as biosynthesis of secondary metabolites, sesquiterpenoid and triterpenoid biosynthesis and stilbenoid, diarylheptanoid and gingerol biosynthesis pathways, were also involved. The co-expression networks were constructed from gene expression data and unveiled the candidate genes responsible for variation. Overall, this study will enrich the molecular mechanisms of wood related genes and provide insights with respect to environmental constrains for further improvement in conservation efforts and future breeding strategies of Teak.

A52

HETEROLOGOUS EXPRESSION OF ATIPK IN SALVIA OFFICINALIS FOR THE PRODUCTION OF COMMERCIAL HIGH-VALUE TERPENOIDS

Sreelekshmi V A, Bharadwaj Revuru & Sathishkumar R*

Plant Genetic Engineering Laboratory, Department of Biotechnology, Bharathiar University, Coimbatore, Tamil Nadu, India
*Corresponding author: rsathish@buc.edu.in

Salvia officinalis (garden sage), belongs to the family Lamiaceae, a fascinating medicinal herb known for its essential oil and having medicinally potent constituents like carnosol, carnosic acid (terpenes) and rosmarinic acid (poly phenols). Due to the huge demand for the metabolites, there is a prerequisite to engineer *S. officinalis* with homologous/heterologous genes to improve the inplanta content of carnosol and carnosic acid. The two 5-carbon moieties Isopentenyl diphosphate (IPP) and its interconvertible dimethylallyl diphosphate (DMAPP) are the central precursors for the terpene diversity in higher eukaryotes. Scientists recently discovered a novel enzyme termed as Isopentenyl phosphate kinase (IPK) in plants which is known to convert Isopentenyl phosphate (IP) to IPP and dimethylallylmonophosphate (DMAP) to DMAPP through ATP dependent phosphorylation. IPK acts a key player in improving the precursor pool (IPP and DMAPP), thereby serving a predominant role in terpenoid biosynthesis. In the current study, we heterologously overexpressed Arabidopsis thaliana IPK (AtIPK) in S. officinalis to improve the precursor pool thereby diverting the flux towards biogenesis of diterpenes like carnosol and carnosic acid. Accordingly, AtIPK was cloned in pBI121 and mobilized to Agrobacterium tumefacians GV3101 for transient expression studies by vacuum infiltration. Post infiltration, infiltrated garden sage leaves improved the expression of terpene biosynthetic pathway genes. Our current study paves the way for future metabolic engineering of the biosynthesis of commercially important terpenoids.

A53

PRODUCTION, CHARACTERIZATION AND EFFICACY TESTING OF RECOMBINANT ANTIMICROBIAL PEPTIDE MEUCIN-18 FROM *NICOTIANA BENTHAMIANA* PLANT SYSTEM FOR THERAPEUTIC APPLICATIONS

Sakshi Chavhan, Parthiban Subramanian, Rudranil Saha & Sathishkumar Ramalingam*

Plant Genetic Engineering Laboratory, Department of Biotechnology, Bharathiar University, Coimbatore, Tamil Nadu, India
*Corresponding author: rsathish@buc.edu.in

A mixture of peptides and proteins are found in scorpion venoms, which serve two basic biological functions, predation and defence. Venom peptides as antimicrobial peptide (AMP) have high bactericidal activity against pathogenic Gram-positive and Gram-negative bacteria. Meucin-18 is an antimicrobial peptide from the venom of *Mesobuthus eupeus* (lesser Asian scorpion). The sequence of Meucin-18 is FFGHLFKLATKIIPSLFQ, molecular weight is about 2.11 kDa, with an isoelectric point of 10. Our research aims to produce recombinant antimicrobial peptide Meucin-18 in potential alternative plant expression system *Nicothiama benthamiana*.

Meucin-18 peptide gene sequence was retrieved from NCBI (Accession no.: ADT89761), codon-optimized and *de novo* synthesized for plant expression and custom cloned in pENTR-D-TOPO vector. Plant high expression construct pEAQ-HT-DEST3-Meucin-18 was generated using Gateway cloning technology with a C-terminal (6X) his-tag and an ER localization signal KDEL. The construct was transformed to *Agrobacterium tumefaciens* GV3101 by electroporation. *N. benthamiana* plants were infiltrated with pEAQ-HT-DEST3-Meucin18 to produce His-tagged recombinant antimicrobial Meucin-18 peptide and purified using Ni-NTA metal affinity chromatography under native conditions.

Gateway-based LR cloning reaction was performed to generate pEAQ-HT-DEST3-Meucin-18 expression constructs; maintained in cloning host *E. coli* DH5α and transformed into *A. tumefaciens* GV3101. The *Agrobacterium* culture was induced with 100 mM acetosyringone and syringe infiltrated into *N. benthamiana* leaves. The crude peptide was extracted from plant leaves; purified using Ni-NTA metal affinity chromatography under native conditions; analyzed on Tricine SDS-PAGE and confirmed using ant-His_6 antibody by Western blot. Antimicrobial activity of plant expressed recombinant Meucin-18 will be evaluated by broth microdilution assay and the plant-produced Meucin 18 will be utilized for therapeutic research purpose.

A54

DE NOVO GENOME ASSEMBLY AND CONTENT ANALYSIS OF *LANTANA CAMARA*

Abhishek Prasad* & Gitanjali Yadav

Computational Biology Laboratory, National Institute of Plant Genome Research, New Delhi, India
*Corresponding author: abhishek1992@nipgr.ac.in; 9572513313

Lantana camara is one of the most noxious invasive plants in the world, and has been listed in the IUCN's 100 most invasive species globally. To decipher the genomic basis of its invasion, two datasets were investigated from two distinct geographical locations, namely India (Joshi et al, 2022) and Australia (Shah et al, 2022). Both these reports contain merely preliminary analysis of the assembled genome, and do not explore the genome or its biological features in detail. Therefore, we re-analysed raw data of both genomes, and were able to assemble the genome of *Lantana camara* from Australia to a greater contiguity than the authors. While the expected 1C size of *Lantana camara* is 2.5 Gb, our assembled genome was 1.99 Gb (79.6 %). N50 was 4,243, BUSCO completeness was 92.5 % and GC content was 38.56. In all, we assembled 28,193 contigs of length 10,000 bp or higher. The repeat content of the genome revealed 80.47 % of the genome were repeat elements, out of which 27.3 % were retroelements, 1.5 % were DNA transposons, 0.72 % were simple repeats, and about 50 % repeat sequences could not be classified. Prediction of gene models in this assembled genome using the nr viridiplantae database and *Lantana camara* RNA seq raw reads generated 57,027 gene models. To functionally characterise the gene models, we performed the KEGG analysis on the predicted genes, which resulted in 6,073 genes being involved in various pathways. 603 of these genes were involved in secondary metabolte pathways. Of these, 153 were involved in different terpenoid biosynthetic pathways, 41 in carotenoid biosynthesis, 139 in phenylpropanoid biosynthesis and 44 in flavonoid biosynthesis. Our current analysis gives a glimpse of the genomic characters of this invasive plant. Further pathway enrichment analysis and functional characterisation is currently underway to understand the genomic basis of invasion of this extremely inoxious species.

A55

DEEP LEARNING-BASED IDENTIFICATION OF DISEASE-RESISTANT PROTEINS IN *ORYZA SATIVA* AND RELATED SPECIES

Vedikaa Dhiman[1], Soham Biswas[1], Asish Kumar Swain[1], Ayan Sadhukhan[1] & Pankaj Yadav[1,2]

[1]Department of Bioscience and Bioengineering, Indian Institute of Technology, Jodhpur, Rajasthan, India
[2]School of Artificial Intelligence and Data Science, Indian Institute of Technology, Jodhpur, Rajasthan, India
*Corresponding author: pyadav@iitj.ac.in

BACKGROUND: Rice (*Oryza sativa*) is the major cereal crop consumed by more than half of the global population. It is a staple crop affected by biotic stress, leading to pathogenic diseases. Developing disease-resistant cultivars by incorporating resistance proteins is required to reduce pathogenicity and yield losses. Machine Learning (ML) models are used to identify resistance proteins, but due to low prediction power, Deep Learning (DL) models are extensively used.

OBJECTIVES: A comparative analysis between ML and DL algorithms predicting disease-resistance proteins in *Oryza sativa* and related species.

METHODS: Protein sequence data, including known disease-resistance proteins as positive and whole proteome as negative datasets from 13 species related to rice, were retrieved. The negative dataset was optimized to reduce the data imbalance. 80/20 (training/testing) partition, feature extraction, and selection were implemented on both datasets. Further, five classifiers of ML and DL were executed, and their performance was evaluated with 10-fold cross-validation. Various model parameters were assessed, namely, accuracy, Area Under Receiving Operating Characteristic curve (AU-ROC), F1-score, precision, and recall. Moreover, the model was tested on a new dataset including proteins with less information on disease resistance mechanism. The developed approach is publicly available for disease-resistance protein identification on GitHub.

RESULTS: In total, 21,289,622 and 19,503 features on the protein sequences were extracted and selected, respectively. DL-based MLP model outperformed applied ML models with the highest accuracy, AU-ROC F1-score, precision, and recall of 91.59%, 97.84%, 93.77%, 95.65%, and 92.5%, respectively. Furthermore, the DL-based MLP model predicted the new dataset with the highest accuracy of 90.1%.

DISCUSSION AND CONCLUSIONS: The high performance and robustness of the DL-based MLP model provide valuable information on the underlying properties of the disease-resistance proteins. The model can rapidly identify these proteins to develop novel disease-resistant varieties. This will assist in transforming traditional rice farming practices into a new age of smart rice farming.

A56

DO ENDOPHYTES MAKE FINGER MILLET STRESS TOLERANT AND AN OLIGOTROPH?

Anita Pande*

College of Biotechnology, Birsa Agricultural University, Ranchi, Jharkhand, India.
*Corresponding author: anitapande66@gmail.com; 9771002247

Finger millet (*Eleusine coracana* L. Gaertn) is a cereal widely consumed in Asia and Africa and is popularly called ragi or madua. It is inherently tolerant to various abiotic stresses like drought, acidity and salinity. It is rich in iron, calcium and the sulphur containing amino acid methionine which is absent in the popular cereals like rice and wheat. Work undertaken for finger printing of the genotypes led to certain SCAR markers. The two SCAR sequences, surprisingly, matched to bacterial regions of the Asp_Glu_race superfamily domain of the gene for hydantoin racemase (Genbank accession no. KC020190 and KC020192) and the HATPase_c domain of the NAR Q receptor protein (Genbank accession no. KC020191), respectively.

Hydantoin racemase has not been reported in any eukaryote. It is found in the bacteria *Ralstonia pickettii* which has been identified in biofilm formation in industrial plastic water piping/ high-purity water. It can survive in low nutrient (oligotrophic) conditions and is capable of degrading contaminates as aromatic hydrocarbons like cresol, phenol and toluene and can survive in environments highly contaminated with metals like Copper, Nickel, Iron and Zinc. Hydantoin racemase, forms part of the reaction cascade known as the "hydantoinase process", which allows the total conversion of hydantoins (byproduct of cytosine and thymine bases of DNA following cell death) into optically pure D- or L-amino acids. Thus, finger millet and *R. Pickettii* exhibit similar behavior vis a vis survival on poor and harsh soils. Is *Ralstonia pickettii,* as an endophyte, helping finger millet to survive in harsh environments?

The second SCAR mached with the HATPase_c superfamily domain found in the NarQ receptor protein. There was 96% homology with the NarQ protein of *Pantoea vagans* which is a bacterial epiphyte of a broad range of plants and is used for biological control of fire blight in pear and apple. This receptor protein is also part of a system that controls the expression of many anaerobic genes in response to nitrate and nitrite in E*schericia coli*. The nitrate assimilation pathway used by higher plants converts ammonium to nitrate in a series of reactions. It is possible that the conversion of ammonium to nitrate in finger millet is due to the combined processes of the eukaryotic and prokaryotic (in the form of an endophyte) pathways. It has been shown that the endophytic fungus *Piriformospora indica* stimulates the expression of nitrate reductase in tobacco and Arabidopsis through a transcription factor upregulated by the fungus. Subsequently, different genotypes of finger millet were analyzed for the presence of endophytes. The DNA isolated from the culturable endophytes was then subjected to partial 16s / 18s DNA sequencing. The comparison with reported sequences suggest that the endophytes have varied properties like drought tolerance, halotolerance, nitrogen fixation and anti-pathogen activities.. Once the role of the endophytes or the metagenomics / meta transcriptomics of the consortium is available it will be possible to transfer some or all of the microbes to other crop species so that specific properties like abiotic stress tolerance can be transferred to them.

A57

ASSOCIATION OF G-QUADRUPLEX FORMING REGIONS WITH DIFFERENTIAL GENE EXPRESSION UNDER SALINITY STRESS IN CHICKPEA

Shubhangi Singh[#] & Rohini Garg*

Plant Gene Regulatory Lab, Centre of Excellence in Epigenetics, Department of Life Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
*Corresponding author: rohini.garg@snu.edu.in
[#]Presenting author: ss496@snu.edu.in

Chickpea is one of the most cultivated legume crops in India that holds high economic and nutritional value. However, the extreme environmental conditions such as drought, salinity, high light intensity, cold, floods, etc. can hamper the plant growth and development resulting in yield losses. Among abiotic stresses, salinity stress conditions have been shown to highly affect the crop production and yield. Upon exposure to salt stress a plant can witness changes in its phenotype and associated gene regulatory network to combat the stressed conditions. Other than the coding regions, non-coding regions including elements such as long non-coding RNAs, miRNAs, silencers, etc. can also be involved in the regulation of gene expression. Enhancers are one such regulatory element that can enhance the expression of their target gene via mediator complex and chromatin loop formation. Identification of enhancer regions associated with gene expression under normal versus stressed conditions can, therefore, give more insight into the molecular mechanisms underlying the stress responses in chickpea. In this study we aimed to identify the putative DNA secondary structure forming sequences across chickpea genome and studied its distribution suggesting its preferential presence in intergenic regions. The overlap between the differentially expressed genes identified previously in our lab with the identified GQS forming regions suggested the possible involvement of these regions in biological processes such as response to salt stress, ion transport, response to osmotic stress, etc. Our work will shed light onto the possible mechanisms by which the presence/absence of DNA secondary structures can modulate the stress responses via differential gene expression under salinity stress and activity of the associated enhancer regions and providing tolerance to the plant.

Keywords: chickpea, salinity, enhancers, G-Quadruplex

A58

FIRST TELOMERE TO TELOMERE INDIAN RICE GENOME

S. Srivignesh[1], Poornima Reddy[2], Nihar Bachan Das[3], Raja Mugasimangalam[2,3] & Rajashekhar Ballari[2*]

[1]Central University of Tamil Nadu, Thiruvarur, Tamil Nadu, India
[2]Qtlomics Technologies, Bengaluru, Karnataka, India
[3]Genotypic Technology, Bengaluru, Karnataka, India
*Corresponding author: rajashekhar@qtlomics.com

Rice (*Oryza sativa* L.) is an important cereal crop and represents the staple food of more than half of the global population. Basmati rice is considered a unique varietal group because of its aroma and superior grain quality with unique quality traits such as extra-longslender grain, lengthwise excessive kernel elongation upon cooking, soft and fluffy texture after cooking, and aroma. These unique characteristics makes Basmati rice economically important crop and its market demand.

Panjab Basmati 3 is a bacterial blight (BB) resistant and dwarf version of tall and highly acclaimed variety Basmati 386 (J Res Punjab agric Univ 51 (2): 206-207, June 2014). In this study we have Sequenced the Punjab Basmati 3 genome by Nanopore PromethaION sequencing. Total 110 Gb data was generated which is around 300X coverage with 17 million reads of 12.5 kb N50. Assembly generated 571 contigs with total contigs legnth of 373.38 Mb. Reference guided mapping of draft Basmati Rice assembly genome indicated ~ 99% of genome coverage at chromosome level.

A59

PATHWAY AND NETWORK ANALYSIS OF GENES LINKED TO TYPE 1 DIABETES - A COMPUTATIONAL APPROACH

Saubashya Sur*

Department of Botany, Ramananda College, Bishnupur, West Bengal, India
*Corresponding author: saubashya@gmail.com

SUMMARY: Type 1 diabetes (T1D) is a complex autoimmune disease damaging pancreatic islet β cells. A comprehensive analysis of T1D-associated genes using bioinformatics tools revealed enriched pathways like cytokine-mediated signaling and immune response activation. Meta-analyses identified common genes and pathways. The findings highlight the role of immune-linked processes in T1D, offering insights for precision medicine development.

BACKGROUND: Type 1 diabetes (T1D) is an autoimmune disorder destroying insulin-producing pancreatic cells, affecting both adults and children. Genetic susceptibility, particularly high-risk genes, plays a major role, but current studies often overlook a comprehensive analysis of the collective role and connections among all associated genes.

OBJECTIVES: This study aimed meta-analysis of T1D-associated genes using bioinformatics tools, focusing on identifying enriched pathways, understanding protein complex roles, clustering enrichments, and mapping pathway networks, to reveal relationships between genes.

METHODS: PubMed and NCBI were utilized to extract Type 1 Diabetes (T1D)-associated genes from human studies. After comprehensive swotting, a dataset comprising 505 T1D-associated genes was curated. Additionally, a separate dataset of 247 T1D-related genes from Single Nucleotide Polymorphisms (SNPs) was prepared using published data. Functional enrichment and pathway analysis were conducted using Metascape. Protein-protein interactions and correlations among T1D genes, were investigated through BioGrid, InWeb_IM, and OmniPath interactome databases using Metascape.

RESULTS: Notable pathways showing enrichment included cytokine-mediated signaling, cytokine production, interferon gamma production, myeloid leukocyte activation, immune response activation, lymphocyte activation, adaptive immune response, and Th17 cell differentiation. The ontology cluster network from T1D-associated genes enhanced our understanding of both inter- and intracluster similarities. Protein-protein interaction analysis uncovered enrichment in toll-like receptors, Vitamin D, interleukins, chemokines, *HLA, VDR, PTPN22, NOTCH 1, NOTCH 3*, and *TNF*. Meta-enrichment analysis highlighted increased proportions of GO biological processes: interferon gamma production, cytokine production, angiogenesis, myeloid leukocyte activation, activation of immune response, lymphocyte activation, and adaptive immune response. It portrayed their involvement in inflammation and immune response pathways to T1D.

DISCUSSION: Enrichment analysis of genes associated with Type 1 Diabetes (T1D) emphasized the crucial role of immune-related mechanisms in metabolism, disease progression, and the etiology of T1D. Overrepresentation and greater degree of significance of the identified GO biological processes, KEGG pathway and canonical pathway from the list of T1D-associated genes compared with T1D-related genes from SNPs, demonstrated that the former may be a better target for clinical investigations. MCODE components underscored the significance of pathways related to vitamin D metabolism, interleukin signaling, toll-like receptors, chemokines, *PD-1, NOTCH*, and antigen processes.

CONCLUSION: This study uncovered vital T1D pathways and genes, revealing 153 significant genes central to its physiology. These offer potential therapeutic targets, notably *HLA*s, *VDR*, and *PTPN22*.

As precision medicine advances, this information guides therapeutic design based on genetic backgrounds. The study also highlighted key SNPs, suggesting targets for microbiome analysis to develop novel T1D strategies.

A60

UNRAVELLING PROGNOSTIC SIGNATURES IN GASTRIC CANCER THROUGH MACHINE LEARNING ANALYSIS OF GENE EXPRESSION DATASETS

Rohit Kumar Verma[1], Prashant Kumar Srivastava[2]** & Ashutosh Singh[1]*

[1]Department of Life Sciences, School of Natural Sciences (SONS), Shiv Nadar Institution of Eminence, Delhi-NCR, India
[2] National Health and Lung Institute, Imperial College London, London, United Kingdom
*Corresponding author: ashutosh.singh@snu.edu.in

BACKGROUND: Gastric cancer is frequently diagnosed at advanced stages, leading to poor prognoses, underscoring the critical need for novel biomarkers in its identification. The discovery of prognostic biomarkers has the potential to enhance prognosis and personalised treatment regimens.

METHODS: To find the prognostic signatures in gastric cancer, we acquired and consolidated several datasets from the NCBI-GEO to build a complete dataset. TCGA stomach cancer (STAD) was used for survival and other bioinformatics analysis. A subset focused on known cancer driver genes was also constructed from databases such as COSMIC, IntOGen, and Bailey. Recursive feature elimination (RFE) was employed to extract top 10, 20, 30, 40 and 50 gene signatures from both datasets. Furthermore, eleven different machine learning classifiers were employed on these selected features to classify cancer and normal samples. Alongside, differential expressed gene analysis, survival and enrichment analyses were also conducted.

RESULT: RF and SVC classifiers outperformed on the top 30 and 40 gene features of the complete and driver datasets, respectively, with test dataset. On the external validation datasets QDA and KNN classifiers outperformed in complete and driver datasets with higher sensitivity and specificity. Among the identified 70 gene features, 28 exhibited differentially regulated. Lasso-penalized Cox regression further refined the gene selection to 8. The median risk score derived from these 8 genes significantly stratified patients, revealing that low-risk individuals experienced significantly better overall survival in GHR, LOX, MAGEA6, CBX7, VCAN, ATAP4, MTTP and EZH2. Pathway enrichment analysis highlighted associations with cancer-related pathways, PI3K-Akt signalling, and gastric cancer.

CONCLUSION: This study underscores the effectiveness of an integrated approach, combining machine learning and bioinformatics analysis, in the discovery of novel gastric cancer biomarkers. Finally, we believe that understanding the role of these prognostic genes in gastric cancer may further provide potential targets for future intervention and therapy.

Keywords: Gastric cancer, Machine Learning, Differentially expressed genes, Prognostic biomarker, Survival Analysis

A61

SCREENING FOR HIF1A INHIBITORS REVEALS THE PATHWAY INTERCONNECTIONS

Yoga Soundarya Mohan, Balaji T Sekar & Kumaresan Ganesan*

Unit of Excellence in Cancer Genetics, Dept. of Genetics, Centre for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University, Madurai, India.
*Corresponding author: kumar@oncocellomics.org

The current study aims to understand the HIF1α associated pathways, to identify the optimal drugs and the targetable subgroup of gastric tumors. From the HIF1α inhibitor drug screening, top hits were analyzed for the possible targets of the compounds. The analysis revealed the association among HIF1α and the identified four key cellular Pathways. Further, these pathways were found to be co-activated in a subset of gastric tumors across multiple cohorts of gastric tumors. Notably, these pathways are enriched in a subset of gastric tumors which are associated with worse overall survival and a higher risk of recurrence. The current investigation also presents a list of potential drugs for a distinct subset of gastric cancer patients.

A62

STATISTICAL PATTERNS IN THE DISTRIBUTION OF TRANSPOSONS IN THE FISH GENOMES

Spandan Patra, Manaswini Parija & Neelesh Dahanukar*

Molecular Ecology and Evolution Lab, Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Delhi-NCR, India
*Corresponding author: neelesh.dahanukar@snu.edu.in

BACKGROUND: Transposons, or the jumping genes, move within a host genome from one genetic locus to another either through copy-and-paste or cut-and-paste mechanism. Since they exploit cellular resources for their movement and copy number increase, they are also called as selfish genetic elements. In addition to the metabolic burden incurred on the host, if transposons get integrated into vital genes, they can disrupt essential functions that can affect the host's survivability. Therefore, understanding how different types of transposons are distributed in the genome with respect to each other and with respect to the host genes is a fundamental question to understanding the ecology and evolution of these selfish genetic elements. We study the patterns of transposon distribution using teleost fish genomes as a model system since they represent lower vertebrates and have relatively smaller genomes.

OBJECTIVES: (1) Annotation of whole genomes of teleost fishes for transposons using EDTA pipeline. (2) Understanding distribution of transposons in the genic and intragenic regions of host genome. (3) Statistical analysis of distribution patterns of different transposable elements with respect to each other and with respect to host genes.

METHODS: Whole genomes of teleost fishes with assembled and annotated chromosomes were downloaded from NCBI RefSeq database. Chromosomes were annotated for transposons using EDTA pipeline. A script was developed to compare the transposon and gene annotations and filtering our transposons present in intergenic and genic regions (PCGs, snRNA, snoRNA, rRNA, tRNA, lncRNA, tRNA, pseudogenes). Univariate and multivariate statistical analysis was performed using opensource software package PAST 3.15.

RESULTS: Cut-and-paste DNA transposons were highly abundant in teleost fishes and represented more than 75% of the total transposons. However, they were smaller in size ranging from less than 200bp to not more than 6000bp. The only exception was helitron, which ranged from less than 200bp to up to 20000bp. On the other hand, copy-and-paste RNA transposons were relatively low in numbers and ranged in size from 1000 to 18000bp. Principle Componenet Analysis (PCA) showed that number of different types of transposons as well as their cumulative size varied in genomes of different species. Further, there was a significant difference in the number of transposons in the chromosomes of different species (PERMANOVA, 9999 permutations, $F = 8.172$, $P = 0.0002$). Even after correcting the number of transposons for chromosome size, the sizes showed positive correlation to the 1st and 2nd PCA axis indicating that larger chromosomes have a disproportionately larger number of transposons than expected under linearity. A significant number of transposons were found inside PCGs, while very few were present in tRNAs and pseudogenes. No transposable elements were found inside rRNAs, snRNAs, snoRNAs or IncRNAs.

DISCUSSION: We provide the first in-depth analysis of distribution of transposons along the chromosomes of teleost fishes with respect to each other and the host genes. Our analysis shows that a large proportion of DNA transposons are present in the PCGs, which could be because these regions are transcriptionally more active. However, it is possible that most are present in the introns and non-coding regions of PCGs as their presence in vital genes will be selected against. On the contrary to common belief, pseudogenes had very few transposons, suggesting that some of these genes might have vital functions and their disruption by transposons could be selected against.

A63

DECODING PANCREATIC DUCTAL ADENOCARCINOMA: INSIGHTS FROM MICROARRAY META-ANALYSIS

Akhil Dinesan[1#], Deepak Krishna Vyas[2] & Ashutosh Singh[1]

[1]TBCGRL, Department of Life Science, Shiv Nadar Institute of Eminence, Uttar Pradesh, India
[2]Department of Biotechnology, Lachoo Memorial College of Science and Technology, Jodhpur, Rajasthan, India
*Corresponding author: ashutosh.singh@snu.edu.in
[#]Presenting author: ad305@snu.edu.in

SUMMARY: PDAC is the most prevalent type of adenocarcinoma, with an alarming 5-year survival rate of ~10%. Due to this, there is a dire need to identify diagnostic and prognostic markers to enable early detection and enhance patient outcomes. Our study employs microarray meta-analysis, integrating datasets from different platforms to unravel the genomic landscape of PDAC. Through the study, we identified 9078 differentially expressed genes & The KEGG enrichment analysis depicted that upregulated DEGs were enriched in protein digestion and absorption, ECM-receptor interaction, and focal adhesion. In contrast, downregulated DEGs were enriched in pancreatic secretion, fat digestion and absorption, and protein digestion and absorption.

BACKGROUND: Pancreatic ductal adenocarcinoma is the most common type of adenocarcinoma originating from the pancreatic duct. The 5-year overall survival rate of pancreatic ductal adenocarcinoma is ~10% since most of the patients are diagnosed at advanced stages.

OBJECTIVES: (1) Comprehensive Data Integration: Integration of microarray data for a nuanced understanding of PDAC gene expression. (2) Identification of Robust Biomarkers: Use advanced statistical methods to pinpoint consistent differentially expressed genes associated with PDAC. (3) Functional Annotation: Uncover the biological relevance of identified DEGs through functional annotation and pathway analysis.

METHODS: (1) Data Acquisition: 6 PDAC microarray datasets, GSE15471, GSE16515, GSE28735, GSE62165, GSE71729, and GSE101448, were downloaded from the GEO database in SOFT file format. Only cancer and normal pancreas tissue samples were selected from these datasets for further analysis. (2) Data Preprocessing: Data was normalized according to platform-specific normalization methods, and the values were converted into log2 transformed values. (3) Data Integration: Data integration was performed by using COMBAT from the sva package, removing batch effects arising due to different platforms. (4) Differential Expression Analysis: Limma package was utilized to identify differentially expressed genes in PDAC, forming the foundation for potential biomarkers. (5) Functional Annotation and Validation: GO and KEGG enrichment analysis was performed to identify the functions and pathways disrupted by the differentially expressed genes, with further validation through a literature review.

RESULT: A total of 9078 differentially expressed genes were identified, with 78 upregulated and 39 downregulated genes selected for further analysis.

DISCUSSION: From the meta-analysis of the PDAC microarray dataset, we observed that S100P, CEACAM5, and COL10A1 were the top three upregulated genes, while PNLIPRP2, SERPINI2, and PNLIPRP1 were the top 3 genes downregulated genes. The KEGG enrichment analysis depicted that upregulated DEGs were enriched in protein digestion and absorption, ECM-receptor interaction, and focal adhesion. In contrast, downregulated DEGs were enriched in pancreatic secretion, fat digestion and absorption, and protein digestion and absorption.

CONCLUSION: In conclusion, our comprehensive microarray meta-analysis of pancreatic ductal adenocarcinoma (PDAC) datasets has yielded valuable insights into the genomic landscape of this challenging malignancy. The alarming 5-year overall survival rate of ~10% necessitates the urgent need to identify diagnostic and prognostic markers to facilitate early detection and improve patient outcomes.

A64

RENEGENE-GI: IRREGULAR COMPUTING WITH GENOMIC BIG DATA

Cherisha*

Shiv Nadar University Chennai, Tamil Nadu, India
*Corresponding author: cherisha23610016@snuchennai.edu.in

BACKGROUND: We present ReneGENE-GI, an innovatively engineered Genome Informatics pipeline that performs genome mapping with high precision and accuracy, without unduly penalizing biological fidelity of the results. The pipeline hosts a unique blend of highly dynamic multi-dimensional data structures and parallel algorithms designed for executing the irregular genomic computing across accelerator platforms.

OBJECTIVES: ReneGENE-GI aims to exploit the inherent parallelism and scalability of the underlying hardware at the level of micro and system architecture. It engages fine-grain synchronization and allows the application to scale up with dynamic load balancing techniques. The principal novelty is engineering the solutions using existing algorithms using a data streaming approach that minimizes heap memory footprint and input/output bottlenecks.

METHODS: The ReneGENE-GI pipeline (Figure 1) performs Short Read Mapping (SRM). The small fragments of genome from the NGS platforms, generally known as short reads, are mapped or aligned against a reference genome string through SRM. SRM works on the massive input data set of short reads, typically of the order of petabytes, and aims to find the region of origin of each short read string with respect to the reference, and hence find regions of similarity or dissimilarity.

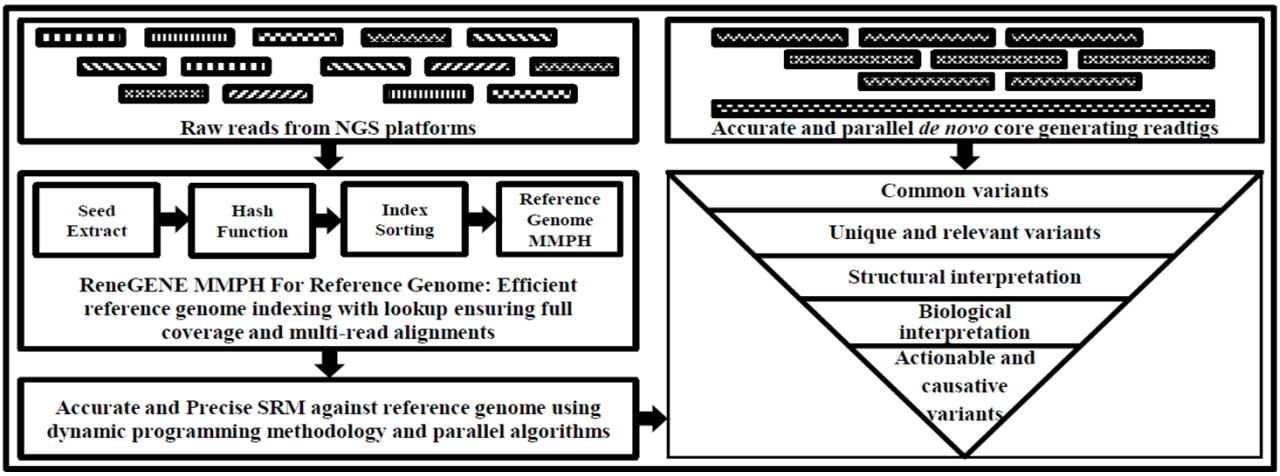

Figure 1| ReneGENE-GI pipeline.

RESULTS: The novelty of the ReneGENE-GI pipeline is the unique blend of comparative genomics and de novo sequence assembly, offering the most precise and accurate SRM. The Comparative Genomics Module (CGM) exploits parallel dynamic programming methodology to accurately map the short reads against the reference genome. The alignment is backed by an accurate indexing and lookup of reads against the reference using the parallel implementation of dynamic Monotonic Minimal Perfect Hashing (MMPH) method, allowing a heuristic free multi-read alignment across repeat regions of the reference.

The de novo module is implemented as a parallel map-reduce contig generation technique. The contigs, or larger genomic segments generated by grouping short reads based on the parallel de novo assembly algorithm, are further mapped on to the reference genome to encompass the possible

insertions and deletions of genetic alphabets at certain locations, thereby widening the map space and coverage.

Performance of ReneGENE-GI with state-of-the-art aligners for small organism genomes is provided in Table 1.

Table 1| Comparative table for performance of ReneGENE-GI.

Genome	Bowtie2	BWA-MEM	HISAT2	GMAP-GSNAP	BISMARK	STAR	MINIMAP2	ReneGENE-SRM Unoptimized
E.coli (e.c)	131.847	36.156	82.349	11.681	176.683	101.443	23.932	**43.28**
P.aeruginosa (p.a)	128.41	29.941	76.704	18.729	39.387	101.301	29.09	**17.12**
S.cerevisiae (s.c)	278.1	79.6	203.22	28.851	1217	305.18	73.531	**37.08**
S.pombe (s.p)	307.6	59.778	81.02	23.38	163.418	3076.064	17.856	**41.01**

DISCUSSION: We have two flavours of ReneGENE-GI's CGM module, ReneGENE-AccuRA for FPGAs and ReneGENE-GMAccS for GPUs. ReneGENE-GI uses the k-mer based dynamic Monotonic Minimum Perfect Hashing (MMPH) algorithm for reference genome indexing provides an accurate hash table, Supplemented with a multi-threaded firmware architecture, the CGM in ReneGENE-GI precisely aligns short reads, at a fine-grained single nucleotide resolution, and offers full alignment coverage of the genome including repeat regions. ReneGENE-GI is a fully streaming solution that eliminates memory bottleneck and storage issues, thus reducing the computing and I/O burden on the host significantly.

CONCLUSIONS: The precise secondary analysis offered by ReneGENE-GI, associated with an efficient tertiary analysis down- stream serves to be a promising target to derive more meaningful inferences from NGS data with biological and clinical significance.

A65

ACCURA RENEGENE-GI: ACCURATE ALIGNMENT OF SHORT READS ON RECONFIGURABLE ACCELERATORS

Shyam Sundar*

Shiv Nadar University Chennai, Tamil Nadu, India
*Corresponding author: shyam23610029@snuchennai.edu.in

BACKGROUND: AccuRA is a massively parallel, scalable, high performance reconfigurable accelerator for accurate alignment of short reads. AccuRA precisely aligns at a fine-grained single nucleotide resolution. The AccuRA prototype, hosting eight kernel units on a single reconfigurable device, aligns short reads with an alignment performance of 20.48 Giga Cell Updates Per Second (GCUPs). AccuRA also scales and aligns genomes of various sizes.

OBJECTIVES: AccuRA aims to generate precise, affordable, reliable and actionable results from SRM, to support any application with uncompromised accuracy and performance. AccuRA is a massively parallel, scalable, high performance reconfigurable accelerator for accurate alignment of short reads.

METHODS: The AccuRA architecture (Figure 1) represents a heterogeneous Single Instruction Multiple Data (SIMD) system, aimed at accelerating the SRM process in NGS pipeline. The architecture consists of a multicore host platform running AccuRA's firmware, tightly coupled to a Reconfigurable Hardware Platform (RHP) configured with AccuRA's parallel hardware units.

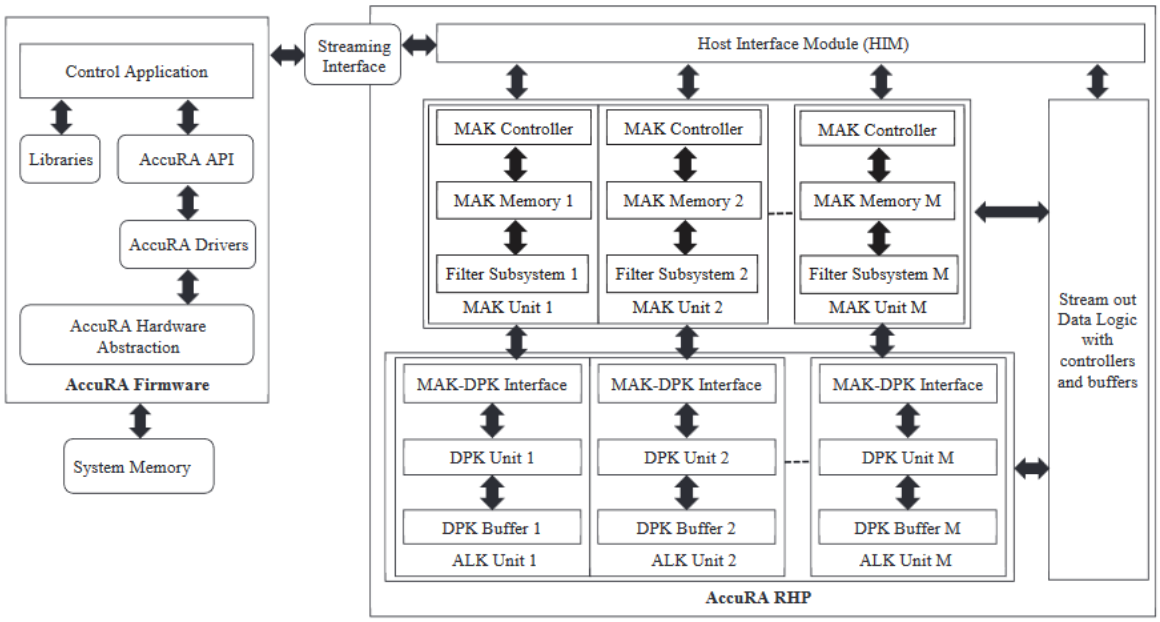

Figure 1| AccuRA Architecture.

RESULTS: The AccuRA firmware preprocesses the reference genome to generate the Reference Index Table (RIT) elements. The Host Interface Module consists of upstream and downstream data paths to and from the streaming interface. It acts as a gatekeeper, and orchestrates the flow of control, data and configuration information between host and RHP.

The RHP is made up of MApper Kernel (MAK) units and ALigner Kernel (ALK) units. The MAK unit filters out those reads which share a certain range of similarity with the RIT elements. MAK units achieve filtering in a deterministic minimal possible time. Each ALK unit has a Dynamic

Programming Kernel (DPK). The DPK Kernel architecture has a parallel linear array of CFUs, acting on an input sequence pair of a short read. The DPK Kernel constructs an alignment matrix with one index for each sequence in a recursive manner. Each CFU is responsible for populating one column of the alignment matrix. The calculation of current CFU depends on the results of the previous CFU or parent. Thus, each CFU is one cycle behind the predecessor in the linear chain. As the length of the sequences increases, the DPK Kernel also has to grow in size. Multiple instances of the DPK kernel within an RHP will not help, as the number of instances is limited by the RHP.

Comparison of the performance of AccuRA with BFAST is provided in Table 1 and human genome alignment results using AccuRA are shown in Table 2.

Table 1| Comparison of performance of AccuRA with BFAST.

ReadSet E. coli	Num: Reads	BFAST Time in seconds	AccuRA Time in seconds, 8 units	AccuRA Time in seconds, 48 units
SRR1948068	596100	35	0.66	0.11
SRR072097	14764330	207	17.81	2.97
SRR072103	23239433	302	24.23	4.04
SRR071759	49202822	1084	80.02	13.33

Table 2| Human genome alignment results using AccuRA.

Read Data Sets	SRR1559281	SRR1559282	SRR1559283
No. of Reads	142992687	146386600	144082500
No. of Pairs	5067156377	4898853334	5061571327
Alignment Time(s)	6214.239978	5962.010015	6066.540026
Alignment Time (min)	103.5706663	99.36683358	101.1090004

DISCUSSION AND CONCLUSIONS: Have presented AccuRA, an accurate, reliable and massively parallel solution for the SRM problem. The DPK units in AccuRA's hardware host a highly efficient and parallel DPK kernel to achieve traceback in hardware in the shortest deterministic time and agnostic to short read length based on a DP alignment algorithm. Traceback in hardware, AccuRA eliminates the memory bottleneck issues and reduces the compute intensive tasks on the host. The results from AccuRA are available in the Sequence Alignment/Map (SAM) format, making it compatible with the downstream applications in the NGS pipeline.

A66

QGENE: QUANTUM ACCELERATOR FOR SHORT READ MAPPING

Madhan Kumar*

Shiv Nadar University Chennai, Tamil Nadu, India
*Corresponding author: madhankumarp@snuchennai.edu.in

BACKGROUND: We present QGENE, an innovatively engineered short read mapping tool that performs genome mapping with high precision and accuracy, without unduly penalizing biological fidelity of the results. QGENE leverages quantum computing for efficient and accurate sequence comparison using Smith Waterman algorithm for local alignment.

OBJECTIVES: QGENE is part of ReneGENE-GI, a larger genome informatics pipeline, that hosts a unique blend of highly dynamic multi-dimensional data structures and parallel algorithms designed for executing the irregular genomic computing across accelerator platforms. ReneGENE-GI aims to exploit inherent parallelism and scalability of underlying hardware at the level of micro and system architectureOur novelty is engineering the solutions using existing algorithms using a data streaming approach that minimizes heap memory footprint and I/O bottlenecks.

METHODS: QGENE Quantum Circuit (Figure 1) is a quantized version of Smith Waterman algorithm for short read mapping (SRM). QGENE uses 5 qubits in total, for one pair of sequences of length N. The number of qubits remain constant, as the sequence length increases. The circuit is initialised based on the nucleotides in both the sequences. This is done by adding X gate appropriately to encode the 4 nucleotides (A – 00, C – 10, G – 01, T – 11). The first two qubits represent the nucleotide value of sequence A and the next two represent sequence B. CX gates are then applied to entangle qubits 0 and 3, 1 and 4. Qubit 2 and 3 are applied a X gate additionally. Finally, a CCX gate is applied to qubits 2, 3 and 4 and the output of qubit 4 is measured. The circuit is measured to determine if a particular nucleotide matches with another, with the following logic formulations.

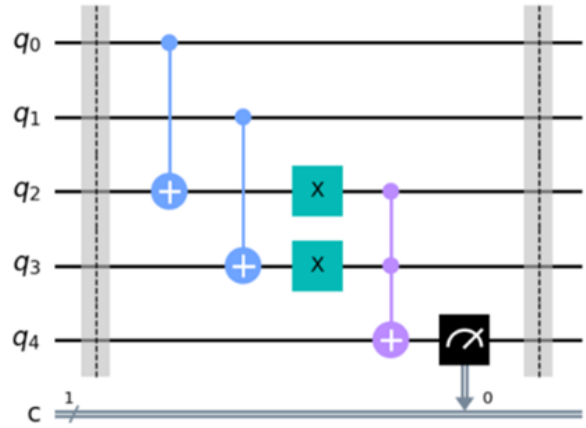

Figure 1| QGENE Quantum Circuit.

QGENE is implemented in Python using IBM's Qiskit framework and executed using Statevector_Simulator, in a noise-free environment, with probability amplitudes of all possible states in the system. The simulator type used is Schrödinger wavefunction, Version 0.1.547, allowing a maximum of 32 qubits and 1000 shots, and provides perfect execution of quantum gates without any imperfections.

$$
\begin{aligned}
\text{Initialization:} \quad & |\psi_{\text{initial}}\rangle = |0\rangle^{\otimes 4} \\
\text{If } seq[i] =' C' : \quad & |\psi_{\text{final}}\rangle = X_1|\psi_{\text{initial}}\rangle \\
\text{If } seq[i] =' T' : \quad & |\psi_{\text{final}}\rangle = (X_0 \otimes X_1)|\psi_{\text{initial}}\rangle \\
\text{If } seq[i] =' G' : \quad & |\psi_{\text{final}}\rangle = X_0|\psi_{\text{initial}}\rangle \\
\text{Application of CX Gates:} \quad & |\psi_{\text{final}}\rangle = CX_{0,2}CX_{1,3}|\psi_{\text{final}}\rangle \\
\text{Application of X-gates:} \quad & |\psi_{\text{final}}\rangle = (X_2 \otimes X_3)|\psi_{\text{final}}\rangle \\
\text{Application of CCX Gate:} \quad & |\psi_{\text{final}}\rangle = CCX_{2,3,4}|\psi_{\text{final}}\rangle \\
\text{Measurement:} \quad & |\psi_{\text{final}}\rangle = M_4|\psi_{\text{final}}\rangle
\end{aligned}
$$

Parallel SRM using QGENE: The alignment scores are calculated for every result from the circuit based on Smith Waterman local alignment formulation. The whole process is parallelized by dividing the scoring matrix into anti-diagonals of the alignment matrix, where each anti-diagonal represents a set of cells that can be computed independently. This is done by assigning each anti-diagonal computation to a separate processor or thread. The circuits of the corresponding nucleotides in the anti-diagonal are computed, and measured in parallel along with the score matrix computation.

RESUKTS AND DISCUSSION: As per the Table 1 below, QGENE (Sequential), a non-parallelized version takes a maximum time of 21.18 seconds. The QGENE parallel version works on the Anti Diagonals of the matrix in parallel thus clocking 15.92 seconds. The third version runs in 15.4 seconds, that utilizes anti-diagonals in parallel, but the computation in each antidiagonal is performed sequentially.

Table 1| Comparison of Quantum version of Smith Waterman for a sequence of length 40.

Algorithm	**Time (s)**
Quantum Smith Waterman (Sequential)	21.18
Quantum Smith Waterman (Using Anti Diagonal and Parallelisation)	15.92
Quantum Smith Waterman (Using Anti Diagonal)	15.40

A67

DEEP SCIENCE & DEEP TECH FOR BIOPHARMA R&D

Manasvi Sharma* & Rammohan Pratapa

Aganitha Cognitive Solutions Pvt Ltd., Hyderabad, Telangana, India
*Correspondence email: manasvi@aganitha.ai

BACKGROUND: Advances in computation biology and AI tools are revolutionizing nearly every stage of the drug discovery process by establishing a virtuous cycle of innovation; combining deep science with deep tech. High throughput datasets such as biobanks globally have amassed millions of human genomes alongside detailed medical histories, helping create broader coverage of human biology. The ability to read epigenomic and transcriptomic state at single-cell resolution has revolutionized the study of biological systems, etc. Technology like AlphaFold, Generative AI are allowing us to overcome challenges of limited data, enabling researchers to go even further than mere target identification and accelerating the design of appropriate drugs that bind to them

OBJECTIVE: To accelerate BioPharma R&D with AI powered in silico methods in the field of:

- ***Target Identification and Validation:*** AI algorithms analyze biological data to identify potential drug targets associated with specific diseases, reducing the time and resources required for the early stages of drug discovery.
- ***Drug Design and Optimization:*** Generative models can propose novel biological structures, speeding up the drug development process.
- ***Virtual Screening:*** AI enables virtual screening of large compound libraries to identify potential drug candidates thereby saving time to generate leads from hits

METHODS:

- Variant, gene-burden and eQTL association analyses using linear, logistic, and linear mixed regression; to identify putative risk loci for a given phenotype.
- Analyzing large volumes of Omics data with proprietary & open source platforms. Identify the mechanism of action(MoA) in both the disease and treatment options. Data pertinent prediction for responder versus non responders to treatment in patient populations
- Generative AI to de novo design and optimize bio molecules as therapeutic candidates by combining sequence and structure property modeling.
- Large Language Models to gather more information around target selection for an indication
- *in silico* solutions in computational biology to efficiently navigate ever evolving complex multi-dimensional data and interconnections, at scale, thereby accelerating drug discovery and development

RESULTS AND DISCUSSIONS: We will discuss how by combining the power of deep science and deep tech we have been able to:

- For an autoimmune disorder, identified key biomarkers at a single cell level in PBMC samples that separated responders from non responders to a plasma derived therapy
- For a tropical infectious disease identified a cell cycle related target that is now being further investigated for therapeutic design purposes
- Optimized chemical and biosynthesis process reaction for a blockbuster drug in animal cell lines. Able to predict yield of the biomolecule with 85% accuracy at a seed train level.
- For a cancer target, we have been able to de novo design and optimize a mAb which shows higher in silico binding than the approved monoclonal antibody for the target in focus

A68

'CELLS-TO-CDNA ON CHIP': GENE EXPRESSION ANALYSIS IN NANOLITER DROPLETS ON A DROPLET MICROARRAY

Shraddha Chakraborty[1,5]*, Charlotte Luchena[1], Jonathan J. Elton [1], Marcel P. Schilling [2], Markus Reischl [2], Margaux Roux [3], Pavel A. Levkin [1,4] & Anna A. Popova [1,4]

[1] Karlsruhe Institute of Technology (KIT), Institute of Biological and Chemical Systems–Functional Molecular Systems (IBCS–FMS), Hermann–von–Helmholtz-Platz 1, Germany
[2] Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann–von–Helmholtz-Platz 1, Germany
[3] Cellenion SASU, Bioserra 2, Lyon, France
[4] Institute of Organic Chemistry, Karlsruhe Institute of Technology, Fritz-Haber Weg 6, Germany
[5] Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, India (current address)
*Corresponding author: shraddha.chakraborty@gmail.com

SUMMARY: The outcome of cellular screening experiments can be best understood by employing an integrative approach that combines cellular phenotypic observations with gene expression analysis. Here we describe the first established experimental strategy enabling the entire course of cell screening, phenotypic microscopy-based assessments along with mRNA isolation and its conversion to cDNA for gene expression analysis on an open Droplet Microarray (DMA) platform with down to single cell per droplet sensitivity. The principle demonstrated in the current study sets a beginning for myriad of possible applications to obtain detailed information about molecular changes in cells after high-throughput screenings.

BACKGROUND: Droplet-Microarray (DMA) platform is based on hydrophilic-superhydrophobic patterning, enabling researchers to grow cells in the nanoliter droplets formed on hydrophilic spots and perform wide range of cellular screenings in hundreds of cell culture nanoreservoirs [1]. Here we performed phenotypic and transcriptomic analysis of cells cultured on DMA in miniaturized volumes, within the same droplet.

OBJECTIVES: In this study, we aimed to develop a 'sample in-answer out' protocol, where DMA was used as a wall-free, miniaturized platform for cell culture, nucleic acid isolation and manipulation of the same in micro- or nanoliter volumes within a single droplet. Using our novel protocol, we demonstrated successful quantitative gene expression analysis down to single cells in 200 nanoliter total reaction volume per sample [2,3].

METHODS: A 'humidity chamber' was built to fit inside a thermocycler, which would prevent sample evaporation from the nanoliter droplets on DMA. For 'Cells-to-cDNA on Chip', differential numbers (10,100,500 and 1000) of HeLa-CCL2 cells were seeded in 100 nanoliter culture media volumes onto the consecutive hydrophilic spots of DMA. Cells were thereafter lysed within the droplet and mRNA was isolated with poly-T magnetic beads. cDNA synthesis was carried out onto the same nanoliter droplets containing resultant mRNA on DMA. After cDNA synthesis, real-time PCR (qPCR) was performed to analyse particular gene expressions in different samples. Furthermore, targeted gene knockdown was performed by anti-GFP siRNA transfection to HeLa-GFP cells cultured on DMA, followed by 'Cells-to-cDNA on Chip' and qPCR. Quantitative gene expression studies were also performed with single cells per droplet on DMA.

RESULTS: Using our novel 'Cells-to-cDNA on Chip' protocol on DMA, we could successfully perform cell screening, isolate mRNA from cells, convert it into cDNA and collect the cDNA from each sample for qualitative and quantitative gene expression studies, all within a total reaction volume of 200 nanoliters and down to single cell per sample. No cross contamination between the droplets were

observed as the qPCR Ct values indicated a distinct inverse relationship with cell numbers on DMA. 90% reduction in GFP transcripts was observed upon performing qPCR after siRNA transfection to HeLa-GFP cells on DMA, which was also in perfect agreement with microscopy-based fluorescence readouts. Furthermore, this protocol can be used in a high-throughput setup to analyze gene expression in single cells per nanoliter droplet on DMA with qPCR-based readouts.

DISCUSSION & CONCLUSIONS: 'Cells-to-cDNA on Chip' is the first established protocol that enables phenotypic cell screening along with targeted gene expression analysis on an open DMA platform with a mere input of single cell per sample. This methodology sets a multitude of possibilities to obtain detailed insight onto the molecular changes in cells upon incubation with various stimuli for the variety of high throughput applications.

[1]. Chakraborty S, Gourain V, Benz M, et al. (2021) Droplet microarrays for cell culture: effect of surface properties and nanoliter culture volume on global transcriptomic landscape. *Mater Today Bio.* 11:100112.
[2]. Chakraborty S, Luchena C, Elton JJ, et al. (2022) "Cells-to-cDNA on Chip": Phenotypic Assessment and Gene Expression Analysis from Live Cells in Nanoliter Volumes Using Droplet Microarrays. Adv Healthc Mater 11: e2102493
[3]. Levkin PA, Chakraborty S, Popova AA (2023) Parallel direct isolation and manipulation of mRNA from cultured cells in nanoliter droplets. EP4186981A1, European Patent application.

A69

GENETIC PERSPECTIVES ON AUTISM: A CGH ARRAY STUDY

Nargis Khanam*

Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India
*Corresponding author: knargis@bhu.ac.in

BACKGROUND: ASD typically manifests early in childhood, often before the age of three, and its prevalence has been steadily increasing globally. The disorder can significantly impact various aspects of an individual's life, including their relationships, academic performance, and daily functioning. Autism exhibits a notable degree of heritability and features an intricate genetic mutational pattern associated with limited social behavior and impaired social communication. Despite numerous studies aiming to uncover mechanisms related to ASD, the underlying pathology remains inadequately comprehended. In this study we focus to evaluate genes associated with autism spectrum disorder. And performed the CGH array of DNA samples of Autistic patients.

OBJECTIVES: To perform CGH array of DNA samples of patients having autism spectrum disorder

METHODS: We obtained samples in EDTA vials from the Outpatient Department (OPD) of Co-supervisors at the Department of Psychiatry and the Department of Pediatrics, IMS, BHU after seeing proper inclusion and exclusion criteria. Subsequently, we extracted DNA from these samples, conducted a CGH microarray analysis, and thoroughly analyzed the obtained results.

RESULTS: Upon scrutinizing the results of the CGH array, we identified amplifications in the CNTNAP family genes among certain individuals diagnosed with autism spectrum disorder.

DISCUSSION: Despite numerous studies aiming to uncover mechanisms related to ASD, the underlying pathology remains inadequately comprehended. Notably, genetic variants within CNTNAP family gene have been identified as contributing factors to the genetic risk associated with ASD. This CNTNAP family gene has been linked to various neurodevelopmental disorders such as autism, intellectual disability, schizophrenia, epilepsy, and ADHD. Typically, the absence or mutation of just one copy of the gene is adequate to trigger the onset of the disease.

CONCLUSION: We can conclude that the CNTNAP family gene encodes a member of the neurexin family, playing a crucial role in the vertebrate nervous system as both cell adhesion molecules and receptors. Disruptions in this gene family may potentially contribute to the onset of autism spectrum disorder.

A70

BIOCLAY BASED DELIVERY OF DSRNA AGAINST CLCUD

Shabnam[#], Meryam Sardar & Jawaid A. Khan*

Department of Biosciences, Jamia Millia Islamia, New Delhi, India
*Corresponding author: jkhan1@jmi.ac.in
[#]Presenting author: shabnam.s2694@gmail.com; 8802355134

There is approximately 30-40% per year reduction in global crop yield due to plant pathogens leading to huge economic and food loss. Current genetic transformation approaches to provide resistance against plant pathogens is a major environmental concern. So innovative approaches are urgently requiring to overcome genetic transformation issues. Among plant pathogens, viruses cause significant losses to agricultural and horticultural crops. During the last couple of decades, Cotton leaf curl disease (CLCuD) caused by *Begomovirus*, has emerged as a serious threat to plants causing enormous losses in the Indian sub-continent. As a safe alternative to transgenic crops, clay nanosheet based delivery of dsRNA against viral infection can be a sustainable and ecofriendly approach. We intend to investigate potential application of sheet-like clay nanoparticles as a dsRNA carrier, for prolonged and effective protection against cotton-infecting begomovirus. In this study, silencing suppressor gene of begomovirus was cloned in vector to express dsRNA in *E.coli* HT115. Synthesized clay nanosheets were loaded by the expressed dsRNA, and nanosheet-dsRNA complex (Bioclay) topically sprayed on leaves of *Nicotiana tabaccum* plants. Bioclay stabilize and provides a continuous supply of dsRNA on the leaf surface of plants. After delivery, dsRNA taken up by plant, and potentially silence the viral genes by its RNA interference-mediated defensive mechanism.

A71

MOLECULAR ENGINEERING OF PHENYLALANINE AMMONIA LYASE TO ENHANCE THERMOSTABILITY

Sudipt Kumar Dalei[†#], Siddhi Pavale[†], Preeti Sokhal, Kunal Meena, Biswambher Biswas & Nidhi Adlakha*

Regional Centre for Biotechnology, Faridabad, Haryana, India
*Corresponding author: nidhi.adlakha@rcb.res.in
#Presenting author: sudipt.kumar@rcb.res.in; 7735886105
[†]Joint first authors

BACKGROUND: Phenylalanine ammonia-lyase (PAL) is an enzyme with significant implications in biotechnology, particularly in the synthesis of phenols, antioxidants, and nutraceuticals. However, the enzyme catalysis capability is impeded by lack of thermotolerance property.

METHODS: To tackle this issue, enzyme engineering strategies using mutagenesis techniques and recombinant DNA technology have been employed. Two different approaches were considered under the spectrum of recombinant DNA technology-based mutagenesis technique. One was the rational approach and the other was the random approach. Error prone PCR was employed in Random mutagenesis approach and Site directed mutagenesis (SDM) was utilized to incorporate site specific mutations. The mutagenized DNA fragments were cloned using restriction digestion and ligation in expression vectors for protein expression and further purification of mutant proteins.

RESULTS AND DISCUSSION: Through a systematic process involving high throughput screening and subsequent biochemical characterization, a thermotolerant mutant was identified. Among eight meticulously selected mutations, the T102E mutation stood out as the mutation demonstrating best thermotolerance. T102E from mutagenesis through rational approach demonstrated a substantial six-fold reduction in PAL's affinity for cinnamic acid and a two-fold increase in the half-life compared to the native PAL enzyme. The mutant also demonstrated twice the specific activity compared to the wild type.

By immobilizing the enzyme on carbon nanotubes, its robustness and reusability were significantly improved. The immobilized mutant PAL was found to be significantly more effective in deaminating phenylalanine present in protein hydrolysate mixture than its non-immobilized counterpart.

Molecular dynamic simulations confirmed the increased tolerance to cinnamic acid, offering further insights into the underlying mechanisms. The findings not only broaden the understanding of PAL's sequence-function relationship but also pave the way for future advancements in enzyme engineering. The ultimate goal is the development of an efficient version of PAL thereby expanding its potential applications in biotechnology.

A72

BEYOND SHORT AND LONG READS: A COMPREHENSIVE COMPARISON OF SINGLE-CELL RNASEQ WITH ILLUMINA AND NANOPORE TECHNOLOGIES

Subhashree Devasenapathy[1]*, Garima Sanoria[2], Aditi Mehta[2] & Sudha Narayana Rao[3]*

[1]NGS Lab, Genotypic Technology Pvt Ltd., Bengaluru, India
[2]Genome Informatics, Genotypic Technology Pvt Ltd., Bengaluru, India
[3]Genotypic Technology Pvt Ltd., Bengaluru, India
*For correspondence: [1]subhashree.devasenapathy@genotypic.co.in,
[3]sudha.rao@genotypic.co.in

Single cell transcriptomics (scRNAseq) has revolutionised our understanding of cellular heterogeneity within tissues and populations. Unlike bulk RNAseq, scRNAseq allows quantification of gene expression at single cell resolution, providing insights into the diversity of gene expression among individual cells within a population, aiding to advance our understanding of diseases such as cancer and immunopathologies, by dissecting the heterogeneity and identifying cell types underlying the pathology, in turn providing insight into designing interventional strategies. Illumina sequencing, with its short-read technology, has been the cornerstone of scRNA-seq workflows, enabling comprehensive transcriptomic profiling across thousands of individual cells. Yet, the inherent limitations of short reads, such as difficulties in resolving complex isoforms and repetitive regions, necessitate the need for alternative approaches to fully capture the intricacies of cellular transcriptomes. Long-read sequencing platforms, such as Oxford Nanopore, offer compelling solutions to these challenges by enabling sequencing of full-length transcripts. Long read data deliver more accurate representation of gene isoforms and facilitate the detection of alternative splicing events with greater precision. We at Genotypic Technology are collaborating with 10X Genomics, to enable researchers adopt this cutting-edge technology as a tool to explore complex transcriptomic pathways and processes not just at single cell resolution but also with significantly increased ability to identify alternative splice isoforms that underlie differential gene expression. The current poster aims to address and showcase our offerings and solutions in the single cell genomics space, using both short read and long read technologies, each with its unique advantages and challenges.

A73

COMPREHENSIVE GENOME AND TRANSGENE CHARACTERIZATION OF TRANSGENIC CELL LINES FOR BIOPRODUCTION PROCESSES

Meenal Vyas[1]*, Nihar Bachchan Das[2], Garima Sanoria[2], Manjunatha B L[3] & Sudha Narayana Rao[1]*

[1]Genotypic Technology Pvt Ltd., Bengaluru, India
[2]Genome Informatics, Genotypic Technology Pvt Ltd., Bengaluru, India
[3]Nanopore lab, Genotypic Technology Pvt Ltd., Bengaluru, India
*For correspondence: meenal.v@genotypic.co.in (MV); sudha.rao@genotypic.co.in (SNR)

In the realm of bioproduction, development and optimization of transgenic cell lines play a pivotal role in enhancing the efficiency of biopharmaceutical manufacturing. Comprehensive characterization of a transgenic cell line engineered for bioproduction purposes is one of the most unique and customized offerings at Genotypic Technologies Pvt Ltd. The package amalgamates Nanopore sequencing technology and advanced bioinformatics tools to analyze the sequencing data allowing for accurate mapping and annotation of transgene insertions and genomic modifications. Nanopore sequencing offers a unique advantage over traditional sequencing methods by enabling real-time, long-read, and single-molecule DNA and RNA sequencing. The approach enables a holistic understanding of the transgenic cell line's genomic landscape, including the integration sites of the introduced genetic elements, integrity of the transgene, and identification of any unintended genomic alterations. RNA seq data provides a detailed examination of gene expression patterns, alternative splicing events, and the identification of any chimeric transcripts. The findings from this study demonstrate and showcase Genotypic Technologies` potential as an expert collaborator to facilitate bioproduction optimization, regulatory compliance, and ensure reliability and reproducibility of biopharmaceutical manufacturing processes.

A74

PERFORMANCE OF BACTFAST® & FUNGIFAST® CE-IVD ASSAYS USING NANOPORE SEQUENCING

Lalanika Manahari Abayasekara, Aishwariya Ramanathan & Vaz Gnanam

Microgenomics India Pvt Ltd., Tamil Nadu, India
*For correspondence: manahari@credencegenomics.com (LMA);
aishwariya@microgenomics.in (AR)

SUMMARY: Bactfast® & Fungifast® are pan-infectious diagnostics test that were originially developed using semi-conductor based NGS. The results here are a proof of principle for the same pan-infectious ID using nanopore sequencing. Selected samples were run in parallel with the legacy workflow & results coroborated with culture. The results from this limited sample size show that Nanopore sequencing and be adapted to the Bactfast® & Fungifast® workflow.

BACKGROUND: Bactfast® & Fungifast® are two pan-infectious diagnostic tests that were clinically validated in 2017. These tests using 16s rRNA and ITS1 respectively have been available for clinical diagnostic use in India since 2019. The legacy product used semi-conductor based NGS to identify microbial organisms within a sample. As the Bactfast® & Fungifast® workflows are direct from sample, both products have functional application in clinical treatment specifically in ICU setting. Considering the versatility of Nanopore technology and cost-effective workflows, validating Bactfast® & Fungifast® for Nanopore sequencing would provide a significant benefit to the user. Therefore as a proof of concept, these tests were re-engineered and tested on selected samples using the Minion MK1C.

METHODS: Legacy Bactfast® & Fungifast® Assays were redesigned to allow for the long read sequencing on the Mk1C. Redesigned assays were designated as Nano_BF & Nano_FF. A total of 7 samples were tested (4 clinical samples prescribed for Bactfast®/bacterial culture; 3 fungal culture plates prescribed for Fungifast®). Sample pre-processing and DNA extraction was carried out according to validated protocols using the QIAmp DNA mini kit. Amplification of targetted fragments was carried out using Nano_BF & Nano_FF; and size selection carried out using nuSeek® Size Select. Final purified libraries were sequenced using the Native Barcoding V14 Sequencing kit. Barcoded samples were multiplexed and run on the MK1C. 40MB of data was generated per sample. FAST5 files were analyzed using the Credence Genomics proprietary pipeline for Bactfast® & Fungifast®.

RESULTS: Results comparison (Table 1) with legacy Bactfast®, Fungifast® and culture was conducted to evaluate the validity of the Nanopore workflows.

DISCUSSION: All samples tested identified the same organisms as in the legacy workflows and were corroborated by culture results. A variation in relative abundance was observed in 3 samples, however as this is less than 1% it was not considered significant.

CONCLUSIONS: Based on these results, it is clear that Nanopore sequencing can also be applied for Bactfast® & Fungifast® product workflows. A full scale validation study with a greater sample number would provide more information for validation of Bactfast® & Fungifast® product workflows using Nanopore Sequencing.

Table 1| Comparison of results with Bactfast®, Fungifast® and culture analysis.

Sample ID	Sample Type	Legacy Results	Nanopore workflow results	Culture results
679	Bronchoalveolar Lavage	*Mycobacterium tuberculosis* (41%)	*Mycobacterium tuberculosis* (40.6%)	*Mycobacterium tuberculosis*
697	Fungal Culture plate	*Flaciformispora tompkinsi* (100%)	*Flaciformispora tompkinsi* (100%)	-
713	Fungal Culture plate	*Magnusimyces capitatus* (97%)	*Magnusimyces capitatus* (97%)	-
752	Pus	*Burkholderia pseudomallei* (100%)	*Burkholderia pseudomallei* (100%)	*Burkholderia pseudomallei*
792	Fungal Culture plate	*Medicopsis romeroi* (100%)	*Medicopsis romeroi* (100%)	-
799	Pus & tissue	*Staphylococcus aureus* (99%)	*Staphylococcus aureus* (99.1%)	*Staphylococcus aureus*
809	Pus	*Mycobaterocides abscessus* (98%)	*Mycobaterocides abscessus* (97.1%)	Acid -fast Bacill

Further Reading

1. Abayasekara LM, Perera J, Chandrasekharan V *et al.* (2017) Detection of bacterial pathogens from clinical specimens using conventional microbial culture and 16S metagenomics: a comparative study. BMC Infect Dis 17, 631. https://doi.org/10.1186/s12879-017-2727-8
2. Ramanathan A, Marimuthu A, Abayasekara LM *et al.* (2023) Enhancing ICU care through NGS – based identification of infectious agents : A comparative study. Research Square, preprint. https://doi.org/10.21203/rs.3.rs-3498989/v1

A75

To explore miRNAs targeting Stem Cell Transcription Factor for The Reversal of Therapy Resistance in Breast Cancer

Komal Barhela[#], S. K. Mishra & Binayak Kumar [*]

Department of Life Sciences, Sharda University, Uttar Pradesh, India
* Corresponding author: binayak.kumar@sharda.ac.in
#Presenting author: komalbarhela110029@gmail.com

BACKGROUND: Breast cancer is the most frequent incidence among women with an estimated 1,300,000 new cases and 465,000 fatalities each year. It remains the top cause of cancer death among women globally. It is recognised as a heterogeneous illness involving numerous carcinogenic biological pathways and/or genetic/epigenetic changes. Acquired therapy resistance is a major challenge in the patient recovery along with post therapy recurrence of breast cancer. There are many reports suggest that stem cells are the major factor in acquired therapy resistance as well as the recurrence of breast cancer after post- therapy. The Stem cell transcription factor plays a major role in maintaining the characteristic features of cancer stem cells. miRNA mediates a major role in post- transcriptional regulation gene expression along with SCTFs as well that involved in maintaining stem cell phenotype.

METHODOLOGY: miRNAs and corresponding target SCTFs were screened from the review and literature by using PubMed search engines. Integral analysis of both is done by using various bioinformatics tools such as miR-Net and miR-DB etc. SCTFs are further analyzed for their interacting partners by using String V8.0 and established PPI-network as shown in result.

RESULT: We have screen 300 miRNAs associated with therapy resistance as well as target SCTFs and 100 major SCTFs associated with acquired therapy resistance in Breast cancer. Further integral analysis of both miRNA and SCTFs using as mentioned tools we obtained many miRNAs target SCTFs such as miRNA-200a, miRNA-125a, miRNA-15b and so on and the transcription factors which are targeted by miRNAs are p53, TCF, ATF2 and so on showed strong interaction at in-silico level further validation of these candidate need to be done.

CONCLUSION: At the level of in-silico study we found hsa-miRNA-200a, hsa-miRNA-125a, and hsa-miRNA-15b etc. and the transcription factors which are targeted by miRNAs such as p53, TCF and ATF2 etc as potential candidates for further validation at invitro as well as in-vivo level.

Keywords: miRNA, SCTFs, Breast Cancer, Drug Resistance and Genomics.

GIC2024 SPONSORS

Participating Nations

USA
UK
Austria
UAE
Switzerland
Spain
Germany
Singapore
Bangladesh
Nepal
Turkey
Srilanka
Tanzania
Uganda

Contact us at

+91-9900035744, +91- 9611714948
conference@genotypic.co.in
www.genomicsindia.co.in

Follow us at