5TH GENOMEPT SYMPOSIUM

ADVANCING GENOMICS
FOR A HEALTHIER FUTURE

ABSTRACT BOOK

- Invited lectures
- Oral Communications
- Technology Spotlights
- Poster abstracts

1 JULY 2026

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DE LISBOA
LISBON, PORTUGAL

Proceedings of the
5th GenomePT Symposium

www.genomept.pt | info@genomept.pt | genomept

HOSTING AND
ORGANISATIONAL SUPPORT

Ciências ULisboa
FCiências.ID
BioISI

5th GenomePT Symposium – Abstract Book

Faculty of Sciences, University of Lisbon
Lisbon, Portugal
1 July 2026

Published by **GenomePT – National Facility for Genome Sequencing and Analysis**

All abstracts are reproduced as submitted by the authors. The opinions expressed are those of the authors and do not necessarily reflect the views of GenomePT or the Scientific Committee.

CONTENTS

WELCOME

Welcome to the **5th GenomePT Symposium**

It is a great pleasure to welcome you to the **Faculty of Sciences of the University of Lisbon** for this annual meeting of the Portuguese genomics community. Whether you are joining us as a researcher, clinician, student, technology developer, industry partner or policymaker, your presence reflects the increasingly collaborative nature of genomics in Portugal.

GenomePT was established to strengthen the national capacity in genome sequencing and analysis through a distributed research infrastructure connecting complementary expertise across the country. Since its creation, the network has expanded its technological capabilities, fostered new collaborations and contributed to making advanced genomic technologies more accessible to the Portuguese scientific community.

This year's programme reflects the breadth of this mission. Alongside outstanding scientific contributions spanning human health, biodiversity, agriculture and biotechnology, we have dedicated a special session to one of the most exciting developments currently taking place in Portugal: the emergence of a coordinated national ecosystem for genomic medicine. Presentations on the i3S-HEALTH Centre of Excellence and the PT_MedGen National Strategy, followed by a multidisciplinary round table, provide an opportunity to discuss how research infrastructures, healthcare institutions, national initiatives and European partnerships can work together to accelerate the implementation of genomic medicine.

The symposium also highlights the importance of technological innovation, with presentations from our industrial partners and a keynote lecture on European Genomics Infrastructures, illustrating how national initiatives such as GenomePT connect with the broader European research landscape.

Finally, I would like to thank all speakers, participants, sponsors, members of the Scientific Committee, the GenomePT node representatives and the local organising team for their enthusiasm and commitment. Their collective efforts have made this symposium possible.

I hope this meeting will stimulate new ideas, foster collaborations and strengthen the community that continues to build genomics capacity in Portugal.

MARGARIDA GAMA-CARVALHO

Coordinator
GenomePT

About GenomePT

GenomePortugal – National Facility for Genome Sequencing and Analysis – is a distributed infrastructure integrating the Portuguese Roadmap of Research Infrastructures (RNIE).

Genome PT brings together national expertise and state of the art technologies and genomics to advance research, innovation and health for a healthier future.

MISSION

- Enable cutting-edge genomics research across multiple fields.
- Foster collaboration among institutions, researchers and industry.
- Support national and international scientific excellence.

NETWORK

- **National consortium** of academic and research institutions.
- **Distributed infrastructure** across Portugal.
- **Multidisciplinary teams** and complementary expertise.
- **Aligned with national** priorities and European strategies.

SERVICES

- **Sequencing** DNA, RNA, epigenomics, single-cell, spatial, long-read.
- **Bioinformatics** Data analysis, pipelines, interpretation and support.
- **Training & support** Capacity building and scientific dissemination.

www.genomept.pt

ORGANIZING COMMITTEE

SCIENTIFIC COMMITTEE

Margarida Gama-Carvalho	Ciências ULisboa	**Jorge Oliveira**	CGPP | IBMC | I3S
Conceição Egas	Biocant	**Luís Vieira**	INSA Ricardo Jorge
Gabriela Moura	Universidade de Aveiro	**Patricia Brito**	NOVA FCT
Henriqueta Silva	Universidade de Coimbra	**Vitor Sousa**	Ciências ULisboa

ORGANIZING COMMITTEE

Margarida Gama-Carvalho	**Sofia Correia**	**Cristiano Rocha**
Vitor Sousa	**Rúben Silva**	**Bruno Cardoso**
Lucas Monteiro	**Tomás Araújo**	
Raquel Luta	**Carolina Fernandes**	

LOCAL SUPPORT

Ana Martins	**Beatriz Rosa**
Maria Inês Santos	**Nuno Trindade**
Marta Silva	

We thank all committee members for their time, dedication and commitment to making this symposium possible

ACKNOWLEDGMENTS

The 5th GenomePT Symposium is made possible through the commitment and collaboration of many individuals and organizations.

We thank all who support our mission to advance genomics in Portugal.

SPONSORED BY

PLATINUM

soquímica
PacBio
Quilaban uma fórmula de vida
Grupo I.L.C.
Element Biosciences
BioPortugal Answers for life

GOLD

Oxford Nanopore Technologies
Palex Improving technologies Improving lives
STAB vida Your easy genetics laboratory

SILVER

BONSAILAB

SUPPORTING

Unilabs
specanalítica

WITH THE SUPPORT OF

Ciências ULisboa
FCiências.ID ASSOCIAÇÃO PARA A INVESTIGAÇÃO E DESENVOLVIMENTO DE CIÊNCIAS
BioISI

TOGETHER WE ADVANCE **GENOMICS**

PROGRAMME

MORNING SESSION

FROM GENOMES TO PRECISION HEALTHCARE

09:00 **REGISTRATION**

09:30 **WELCOME**
GenomePT & CIÊNCIAS ULisboa representatives

09:45 **GENOMEPT 2016-2026**
Building genomics capacity in Portugal
Margarida Gama Carvalho, GenomePT Coordinator

10:00 **i3S-HEALTH: Teaming to Drive the Future of Genomic Medicine** Cláudio Sunkel, Director, I3S

10:20 **PT_MedGen: National Strategy for Genomic Medicine**
Astrid Vicente, PT_MedGen Coordinator, INSA

10:40 **ROUND TABLE**
Building Genomic Medicine in Portugal: Strategies, Infrastructures and Implementation
Moderated by Manuel Santos (Director, MIA, UCoimbra)
Panel: Astrid Vicente (INSA), Cláudio Sunkel (i3S), Jorge Oliveira (BioData), Sérgio Sousa (ULS Coimbra), Margarida Gama Carvalho (GenomePT)

11:15 **COFFEE-BREAK AND POSTERS**

11:45 **Genetic test clinical utility and variant (re)classification: a long-read story and beyond**
Jorge Oliveira, CGPP - IBMC

12:05 **Learning hub in rare diseases and the Preventable Project** Carla Oliveira, I3S

SELECTED ORAL COMMUNICATIONS

12:25 **Transcriptional and Chromatin Determinants of Human Naïve Regulatory CD4 T Cell Identity** Alexandre Raposo, Gulbenkian Institute for Molecular Medicine

12:35 **Medieval Iasi Through the Lens of Ancient DNA: Population Structure, Ancestry and Cemetery-Level Variation**
João Brochado de Almeida, Universidade do Minho

12:45 **Detection of the BRCA2 c.156_157insAlu germline variant from FFPE tumor samples DNA using short-read targeted NGS** Manuela Pinheiro, IPO-Porto

12:55 **LUNCH AND POSTER SESSION**

AFTERNOON SESSION

FROM GENOMES TO SOCIETAL IMPACT

14:30 **EASIGEN-DS Consortium – Design Study for a Research Infrastructure on Advanced Genomics Technologies** Ivo Gut, Director, Centro Nacional de Análisis Genómico (CNAG), Spain

15:00 **Metabolic innovation through horizontal gene transfer in yeasts** Paula Gonçalves
UCIBIO, FCT-UNL

TECHNOLOGY SPOTLIGHTS

15:20 **Sequencing reimagined: introducing the Illumina Trupath Genome**
Stef Kupers, Illumina

15:40 **Beyond the Blind Spots: What Long-Read Sequencing is changing in Human Genomics**
Ximo Garcia, PacBio

16:00 **COFFEE-BREAK AND POSTERS**

16:30 **Past Genomes, Future Conservation: Museomics to the rescue**
Ricardo Lopes, Ce3C, CIÊNCIAS ULisboa

TECHNOLOGY SPOTLIGHTS

16:50 **New quality standard of sequencing with Aviti**
Agnieszka Ciesielska, Element Bioscience

17:10 **Rethinking WGS with MGI: Why Quality, Speed, and Cost No Longer Require Trade-offs**
Bernard Okere, MGI

17:30 **Setting new boundaries for single-cell and spatial multiomics analysis with 10X Genomics**
Ana Borges, BonsaiLab

SELECTED ORAL COMMUNICATIONS

17:45 **A genome language model for mapping DNA replication origins**
Francisco Berkemeier, University of Cambridge

17:55 **Transcriptomic signatures differentiating the successes and limitations of regeneration in a marine annelid**
Inês Moutinho Cabral, UCIBIO, FCT-UNL

18:05 **Single nucleus transcriptome atlas of the honey bee gut**
Waldan Kwong, Gulbenkian Institute for Molecular Medicine

18:15 **CLOSING**

PRACTICAL INFORMATION

Below you will find key information to help you navigate the campus and make the most of the symposium.

VENUE
Scientific Sessions

Building C8 – Auditorium 8.2.30

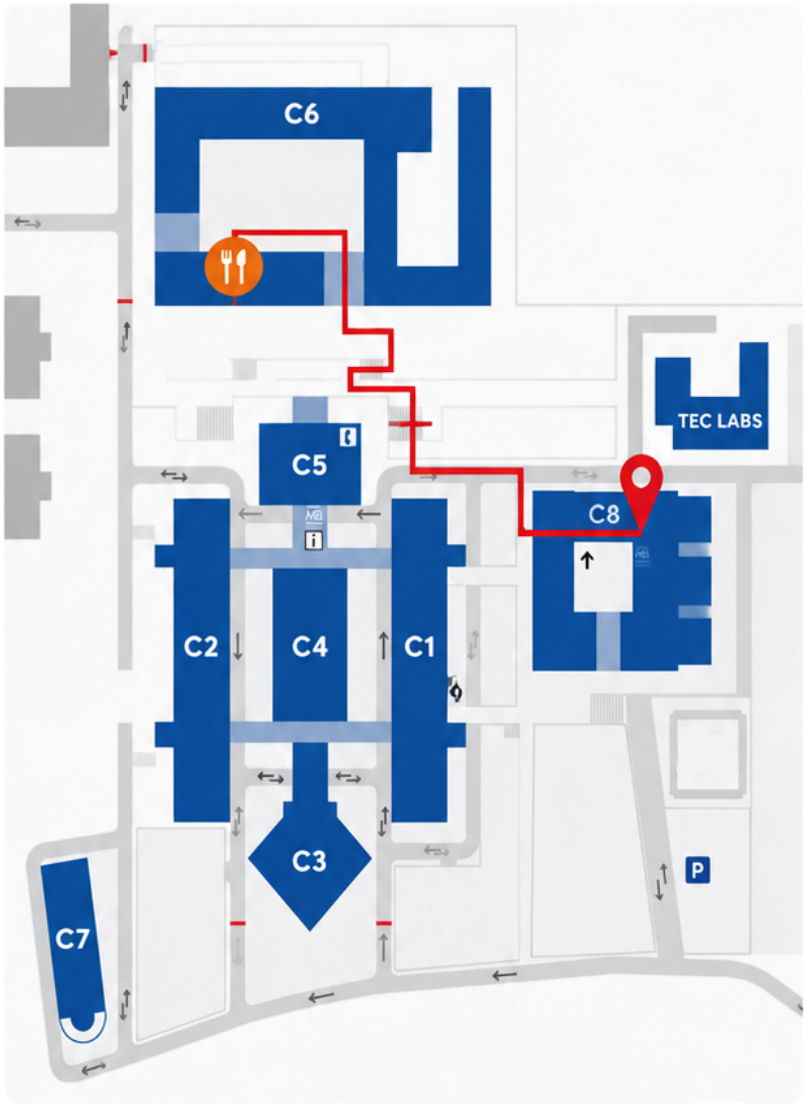
POSTERS, COFFEE BREAKS & SPONSOR EXHIBITION
Building C8 – Foyer and Courtyard

REGISTRATION
Building C8 –Courtyard

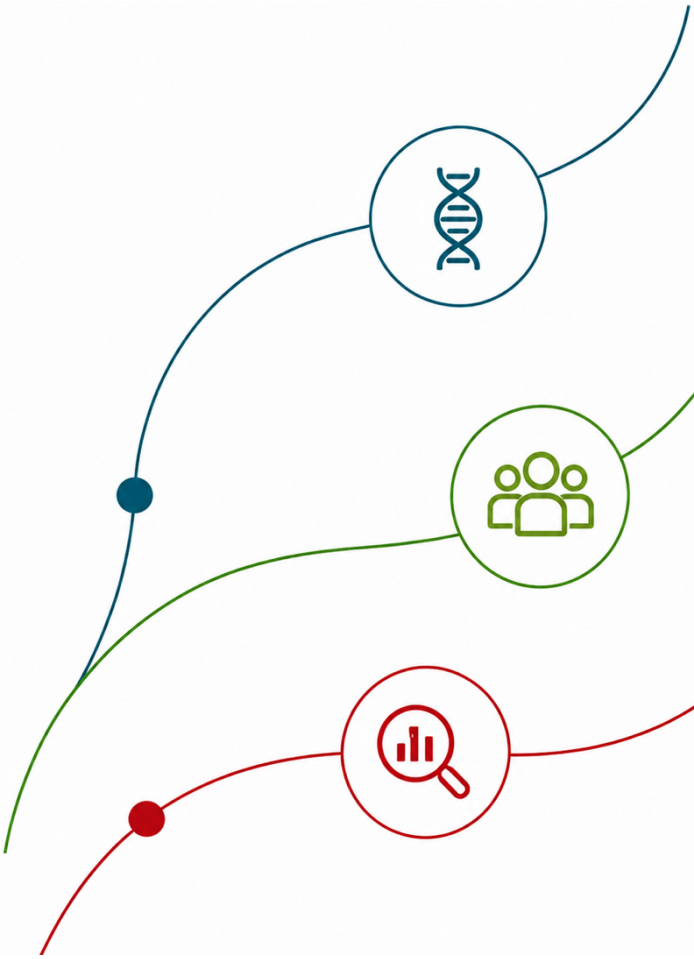
LUNCH
Building C6 –Cafeteria
(see map and directions)

HOW TO GET TO THE CAFETERIA
From Building C8, follow the route shown on the map (red line).
The cafeteria is located in Building C6.
Approximate walking time: 3–4 minutes.

LEGEND
- You are here (C8)
- Cafeteria (C6)
- Walking route
- Parking
- Main entrance



Wi-Fi
- **eduroam** is available throughout the campus.
- Visitors: connect to **guest@Ciencias** (password: **campanha**).



SPECIAL SESSION

Building Genomic Medicine in Portugal

SETTING THE SCENE

Towards Genomic Medicine in Portugal

i3S-HEALTH: Teaming to Drive the Future of Genomic Medicine

Genomic medicine is reshaping healthcare through more precise diagnosis, disease prevention and personalized therapeutic approaches. Achieving this transformation, however, requires more than advanced sequencing technologies: it depends on integrated infrastructures capable of connecting research, clinical practice, computational biology and innovation.

This keynote introduces i3S-HEALTH, Portugal's first Center of Excellence in Genomic Medicine, funded through the Horizon Europe Teaming for Excellence programme. Developed in partnership with EMBL and EMBL-EBI, the initiative will establish a state-of-the-art platform integrating next generation sequencing, artificial intelligence, bioinformatics and functional genomics. Beyond expanding national capacity for genomic research and diagnostics, this project aims to accelerate technology transfer, strengthen collaboration with healthcare providers and industry, and position Portugal as a European reference in genomic medicine

Claudio Sunkel

Speaker

Professor of molecular biology at ICBAS, University of Porto, and Director of i3S. An internationally recognized expert in cell division and genome stability, he has authored more than 100 scientific publications and has held several prominent international leadership roles, including President of the Council of the European Molecular Biology Laboratory (EMBL). He currently leads the implementation of the Horizon Europe i3S-HEALTH Centre of Excellence.

PT_MedGen: National Strategy for Genomic Medicine

Implementation of genomic medicine requires not only scientific and technological excellence but also a coordinated framework capable of integrating genomics into healthcare in an equitable, sustainable and clinically meaningful way, while promoting scientific progress.

This keynote introduces PT_MedGen, the Portuguese National Strategy for Genomic Medicine, outlining the priorities for implemening genomic medicine in healthcare and research within the European landscape. The strategy addresses governance, infrastructure, clinical integration, data management and access, public trust and workforce development, while fully aligning with major European and global initiatives, like the 1+ Million Genomes initiative, with the common goal of advancing genomic medicine through policy-driven, multinational coordination, and a strong emphasis on governance and trust. PT_MedGen provides the strategic framework to establish genomic medicine in Portugal while contributing to the European genomic medicine agenda, fostering healthcare transformation, research & innovation, and international collaboration.

Astrid Vicente

Speaker

Coordinator Researcher at the National Institute of Health Doutor Ricardo Jorge (INSA), Portugal, and Head of the Department of Health Promotion and Non-Communicable Disease Prevention, leading expert in human and public health genomics. She has extensive experience in European research and policy, representing the Portuguese Ministry of Health in the European 1+ Million Genomes (1+MG) Initiative, where she serves on the Coordination Group and leads the Working Group on Genomics Implementation in Healthcare. She is involved in major genomic European projects, and coordinated the development of Portugal's National Strategy for Genomic Medicine PT_MedGen.

> These two keynote presentations will set the context for the subsequent round table discussion, bringing together national leaders to reflect on strategies, infrastructures and implementation pathways for genomic medicine in Portugal.

ROUND TABLE

Building Genomic Medicine in Portugal: Strategies, Infrastructures and Implementation

Astrid Vicente
PT_MedGen

Cláudio Sunkel
i3S Health

Sérgio Sousa
ULS Coimbra

Manuel Santos
Moderator

Jorge Oliveira
BioData.pt

Margarida Gama-Carvalho
GenomePT Coordinator

DISCUSSION THEMES

- Building a national genomic medicine ecosystem
- Coordination between healthcare and research
- National and European infrastructures
- Data sharing and interoperability
- Future priorities for Portugal

An open discussion bringing together national leaders from research, healthcare, infrastructure and policy.

Questions from the audience are encouraged. The discussion will continue informally during the coffee break.

INVITED LECTURES

KEYNOTE SPEAKER

Ivo Gut

Centro Nacional de Análisis Genómico (CNAG), Spain

ABOUT THE SPEAKER

Ivo Gut is Director of the National Centre for Genomic Analysis (CNAG) in Barcelona, where he leads the Biomedical Genomics Group. An internationally recognised leader in genomics, his work has advanced large-scale sequencing technologies and their application to human disease, particularly through major European and international initiatives. He has coordinated several EU-funded infrastructure projects, including RD-Connect and EASI-Genomics, and currently coordinates **EASIGEN-DS**, supporting transnational access to cutting-edge genomic technologies. He has authored more than 500 scientific publications, holds numerous patents, and has played leading roles in international genomics consortia including ICGC, IHEC and IRDiRC.

KEYNOTE TALK

EASIGEN-DS Consortium – Design Study for a Research Infrastructure on Advanced Genomics Technologies

Genomics and related omics technologies are transforming research across health, life sciences, agriculture, and biotechnology. By integrating genomics with complementary approaches such as transcriptomics, proteomics, and metabolomics, researchers are increasingly able to address complex biological questions and societal challenges, including disease mechanisms, health emergency preparedness, and agricultural resilience. Despite this rapid progress, access to advanced genomics technologies, sequencing capacity, and specialized expertise remains fragmented and uneven across Europe. This creates disparities in research opportunities and limits Europe's ability to fully capitalize on its scientific and innovation potential.

The EASIGEN-DS project is a European design study aimed at establishing the foundations for a coordinated research infrastructure on advanced genomics technologies, with the long-term objective of inclusion in the ESFRI Roadmap. The infrastructure will provide integrated access to state-of-the-art sequencing technologies, data generation platforms, and specialized expertise across Europe.

The project is structured around three main objectives: (i) to develop a robust scientific and technological design for the future infrastructure; (ii) to define governance, access, and sustainability models ensuring long-term viability; and (iii) to engage key stakeholders, to ensure alignment with existing capacities and community needs.

In this context, collaboration with national infrastructures such as GenomePT is essential to build a coherent European genomics ecosystem, and support capacity building across countries.

In the long term, EASIGEN will enhance European capabilities in genomics and multi-omics, improve equitable access to advanced technologies, and reinforce Europe's competitiveness in research and innovation.

INVITED SPEAKER

Jorge Oliveira

Centro de Genética Preditiva e Preventiva
Instituto de Biologia Molecular e Celular, Porto, Portugal

ABOUT THE SPEAKER

Jorge Oliveira is Vice-Director, Laboratory Director and Senior Molecular Geneticist at the Centre for Predictive and Preventive Genetics (CGPP), IBMC–i3S, University of Porto, and Invited Assistant Professor at ICBAS. He is a specialist in Human Genetics and a certified European Clinical Laboratory Geneticist. His work focuses on the application of genomics, bioinformatics and multi-omic approaches to improve the diagnosis and discovery of rare neurological diseases. He has authored over 70 peer-reviewed publications, serves as Senior Editor of *Annals of Human Genetics*, and is an expert assessor for the European Molecular Quality Network (EMQN).

GENOMEPT LECTURE

Clinical utility of genetic testing and variant (re)classification: a longread story and beyond

Genetic testing often fails to fully resolve the initial clinical question. In the context of rare diseases, a substantial proportion of findings are variants of uncertain significance (VUS). These variants pose significant challenges across the clinical–laboratory continuum, reducing clinical actionability and, in some cases, restricting patients' access to a growing repertoire of disease-modifying therapies. An implicit expectation has emerged that these uncertainties can be resolved through high-throughput functional assays. In principle, such approaches may facilitate the reclassification of VUS and reduce diagnostic ambiguity. Nevertheless, they remain technically complex, may not be universally scalable, and are unlikely to resolve all – or even most - VUS, given the diversity of underlying genetic mechanisms and disease contexts. Importantly, uncertainty in medical genetics extends beyond the paradigm of VUS reduction through reclassification. This presentation will show that the identification of a pathogenic or likely pathogenic variant - even within a genotype consistent with the suspected disease - is not always sufficient to establish a definitive molecular diagnosis.

Additional layers of evidence are often required to assess pathogenicity and elucidate disease mechanisms. This is particularly evident in disorders caused by repeat expansions (e.g., Huntington disease, CANVAS, or SCA27B) or by epigenetic dysregulation (e.g., Prader–Willi/Angelman syndromes or FSHD2). In this context, long-read sequencing can provide critical insights that are not accessible through conventional diagnostic approaches.

Our current translational research targets these limitations by integrating multi-omics to improve diagnostic resolution, shorten time to diagnosis, and enhance clinical utility by systematically addressing genetic uncertainty.

INVITED SPEAKER

Carla Oliveira

i3S, FMUP & IPATIMUP, University of Porto, Portugal
ERN GENTURIS – European Reference Network on Genetic Tumour Risk Syndromes

ABOUT THE SPEAKER

Carla Oliveira is Principal Investigator of the Expression Regulation in Cancer Group at i3S and IPATIMUP, University of Porto, Affiliated Professor at the Faculty of Medicine of the University of Porto, and consultant in hereditary gastric cancer. An internationally recognised expert in hereditary cancer and E-cadherin (CDH1)-related diseases, her research focuses on genomics, bioinformatics and molecular oncology to improve cancer diagnosis, prevention and patient management. She coordinates the Horizon Europe **PREVENTABLE** project, serves as Secretary General of the European Society of Human Genetics, and is the National Coordinator of ERN GENTURIS.

GENOMEPT LECTURE

Learning hub in rare diseases and the Preventable Project

Rare Tumour Risk Syndromes (RTRS) are rare inherited disorders (≤5 per 10,000 people) caused by germline pathogenic variants that confer a high lifetime risk of multiple, early-onset cancers. Identifying at-risk individuals enables cancer prevention and early detection through surveillance and prophylactic interventions, but evidence on clinical and economic value remains limited. The PREVENTABLE project addresses this gap by evaluating prevention strategies across Europe.

We developed comprehensive care pathways from diagnosis to palliative care and created interoperable IT tools to collect real-world clinical data and link healthcare procedures to country-specific costs across nine European countries. Syndrome-specific models were generated for Birt-Hogg-Dubé syndrome (BHDS), Familial Malignant Melanoma (FMM), Hereditary Gastrointestinal Stromal Tumours (GIST), Hereditary Diffuse Gastric Cancer (HDGC), Hereditary Leiomyomatosis and Renal Cell Carcinoma (HLRCC), Li-Fraumeni syndrome (LFS), PTEN Hamartoma Tumour syndrome (PHTS), and Peutz-Jeghers syndrome (PJS).

Clinical data from 2,596 individuals included 1,516 carriers of pathogenic RTRS variants. Overall, 66% of carriers were identified before cancer development and entered preventive surveillance, ranging from 98.4% in HLRCC to 41.3% in LFS, where 58.7% of carriers already had cancer at genetic diagnosis. Survival was significantly improved for HDGC, LFS and HLRCC patients managed through prevention pathways. Across all syndromes, early detection during surveillance resulted in better clinical outcomes. Although prevention generated higher cumulative long-term costs, treatment pathways consistently incurred substantially greater per-patient expenditure, largely driven by a small number of high-cost interventions. Prevention pathways were associated with lower and more stable monthly costs across all syndromes.

These findings demonstrate that proactive surveillance and specialised care improve patient outcomes while reducing the economic burden of advanced disease, supporting a shift from reactive cancer treatment to preventive management of RTRS.

Learn more about RTRS and the PREVENTABLE project at: https://rtrs-hub.preventable.eu/

INVITED SPEAKER

Paula Gonçalves

UCIBIO
NOVA FCT, Caparica, Portugal

ABOUT THE SPEAKER

Paula Gonçalves is Associate Professor with Habilitation at NOVA FCT, where she teaches courses in the fields of Molecular Genetics and Microbiology and coordinates the NOVA Master's in Computational Biology and Bioinformatics. She obtained her PhD in Natural Sciences from Vrije Universiteit Amsterdam, working on yeast molecular genetics and transcriptional regulation, and joined NOVA FCT in 1999. In 2009, she co-founded the Yeast Genomics Lab @ NOVA. Her current research focuses on the impact of horizontal gene transfer on the evolution of metabolism in yeasts.

GENOMEPT LECTURE

Metabolic innovation through horizontal gene transfer in yeasts

Yeasts are excellent eukaryotic model organisms, and our laboratory uses comparative, functional, and population genomics to investigate their evolutionary biology. Among other topics, we have focused over the past decade on the *Wickerhamiella/Starmerella* clade (W/S clade), a lineage of approximately 100 yeast species predominantly associated with floral niches.

A striking feature of W/S-clade genomes is the unusually large number of genes acquired horizontally from bacteria and Pezizomycotina fungi. These xenologous genes have contributed to extensive metabolic remodelling, and are especially abundant in *Wickerhamiella versatilis*, which harbours the largest xenologous repertoire identified to date in the clade, representing roughly 6% of its total gene content.

Currently, we are interested in understanding why horizontally acquired genes have been retained at such unusually high frequencies in these yeasts. In *W. versatilis*, xenologous genes are preferentially located in large, terminal chromosomal regions characterized by low gene density, AT enrichment, and globally reduced transcription. We termed these regions "End domains". We propose that End domains act as genomic safe harbours, limiting the expression of potentially maladapted xenologous genes while allowing time for adaptive evolution.

INVITED SPEAKER

Ricardo Lopes

Ce3C & MUHNAC
CIÊNCIAS ULisboa, Portugal

ABOUT THE SPEAKER

Ricardo Lopes is Assistant Professor in the Department of Biology at Ciências ULisboa and researcher at cE3c. His research focuses on evolutionary ecology and conservation, using birds as a model system to investigate the evolution of phenotypic traits, trophic interactions and host–parasite relationships in isolated, insular, fragmented, migratory and domesticated populations. He is Curator of the Bird Collection at MUHNAC and Guest Curator of the bird collection at MHNC-UP, promoting the use of museum specimens in modern research. He teaches an advanced PhD course on museological techniques and actively contributes to the national scientific infrastructure PRISC through research that mobilizes its collections and data.

GENOMEPT LECTURE

Past Genomes, Future Conservation: Museomics to the rescue

I intend to show that museomics and collectomics are transforming how we view museum collections, leveraging natural history collections as invaluable repositories for genomic and multi-omic data. Museomics applies high-throughput sequencing to degraded DNA from historical to ancient museum specimens, enabling the retrieval of genetic information from extinct, rare, or difficult-to-sample taxa. Collectomics extends this framework holistically, encompassing the "extended specimen" concept, linking physical vouchers with digitized metadata, images, traits, genomic data, and associated environmental, ecological, or historical contexts.

These fields unlock unprecedented potential for genomic analysis by providing temporal depth and spatial breadth unattainable through contemporary sampling alone. I will provide several case studies that show that advances in DNA extraction from museum specimens, targeted enrichment and bioinformatics now yield high-quality genomes, phylogenomic reconstructions, population genomic insights, and microbiome profiles from samples spanning centuries. This facilitates studies of evolutionary history, genetic diversity loss, adaptation, hybridization, and host-pathogen dynamics, with direct applications in conservation, taxonomy, and ecosystem management.

SHORT ORAL COMMUNICATIONS

OC1

Transcriptional and Chromatin Determinants of Human Naïve Regulatory CD4 T Cell Identity

Alexandre A. S. F. Raposo[1]; Pedro Rosmaninho[1]; Beatriz Moleirinho[1,2]; Margarida Paulo-Pedro[1,2]; Susana L. Silva[1,2,3]; Afonso R. M. Almeida[1,2]; Ana E. Sousa[1,2*]

[1] GIMM - Gulbenkian Institute for Molecular Medicine, Lisbon, Portugal
[2] Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
[3] Hospital de Santa Maria, Unidade Local de Saúde de Santa Maria, Lisbon, Portugal

Human naïve regulatory CD4 T cells (Tregs) are a thymus-derived population essential for peripheral tolerance and tissue homeostasis, yet the mechanisms that preserve their steady-state identity remain poorly understood. Naïve Tregs co-express two superimposed programmes: Treg lineage and naïve T-cell identity. We resolved this entanglement into orthogonal programmes by confronting transcription and chromatin accessibility profiles across thymic, naïve, and memory Treg and conventional CD4 T-cell populations.

The core Treg programme, comprising 293 genes common to all three compartments, revealed a conserved regulatory architecture with AP1/ETS members organising gene regulatory modules. The lineage-independent naïve programme identified 175 genes peaking specifically in naïve cells, with overrepresentation analysis supporting translational preparedness as a naïve hallmark alongside cell trafficking, and POU2F1/OCT-1 emerging as key regulator. KLF/SP factors provided quantitative fine-tuning to both programmes.

These findings establish a two-programme model of human naïve Treg identity with implications for immunosenescence and Treg-based cell therapies.

THEME

Functional Genomics and Gene Regulation

OC2

Medieval Iasi Through the Lens of Ancient DNA: Population Structure, Ancestry and Cemetery-Level Variation

João Brochado de Almeida[1]; Ceiridwen Edwards[2]; Pedro Fernandes[1]; Luminita Bejenaru[3]; João Carneiro[1]; Maria Pala[2]; Teresa Rito[1]; Martin Richards[2]; Pedro Soares[1]

[1]University of Minho, Portugal
[2]University of Huddersfield, England
[3]University of Iasi, Romania

Romania occupies a key geographic position between Central Europe, the Pontic steppe, the Balkans and the eastern Mediterranean, yet its medieval population history remains comparatively underrepresented in ancient DNA (aDNA) research.

Historical and archaeological evidence points to repeated episodes of migration, political transition and cultural interaction across this region, involving local communities as well as influences connected to Steppe, Byzantine, Hungarian, Ottoman and broader eastern European.

Here, we present genome-wide ancient DNA data from medieval individuals excavated from three cemetery contexts in Iasi, northeastern Romania, from the 13th to 17th centuries. Petrous bone samples were processed using dedicated aDNA protocols, followed by high-throughput sequencing and a bioinformatic workflow designed for typical aDNA's degraded and low-coverage genomic data. Libraries were assessed through standard aDNA quality control, including endogenous DNA estimation, molecular sex determination, mitochondrial and Y-chromosomal haplogroup assignment, contamination assessment, fragment length distribution and post-mortem damage patterns. Population genetic analyses included principal component analysis (PCA), ADMIXTURE, outgroup-based f-statistics and comparative modelling against published ancient and present-day Eurasian datasets.

Our preliminary results reveal a structured but heterogeneous medieval population landscape within Iasi with different cemeteries having different amounts of eastern Mediterranean, Anatolian-related or steppe-related ancestry. The observed differentiation does not suggest a simple replacement model, instead, pointing towards a population structure shaped by local continuity, mobility, cultural affiliation and potentially social or religious organization. The genetic diversity observed across cemeteries mirrors the historically documented position of Moldavia as a frontier and contact zone between multiple political and cultural spheres during the medieval period.

This study contributes new genomic evidence from a region that remains crucial but under-sampled in European ancient DNA research. And by integrating archaeological context, laboratory authentication and population genetic analysis, we provide a first genomic window into the medieval communities of Iasi and highlight Romania as an important corridor for understanding population interaction between Europe, the steppe and the eastern Mediterranean. These results demonstrate the value of fine-scale cemetery-level sampling for reconstructing not only broad migration processes, but also the diversity of medieval communities.

THEME

Human and Clinical Genomics

OC3

Detection of the BRCA2 c.156_157insAlu germline variant from FFPE tumor samples DNA using short-read targeted NGS

Manuela Pinheiro[1]; Ana Peixoto[2]; Ana Barbosa[2]; Catarina Santos[2]; Carla Pinto[2]; Joana Guerra[2]; Carla Bartosch[3]; Gabriela Soares[4]; João Silva[4]; Manuel R. Teixeira[2]

[1]Cancer Genetics Group, IPO Porto Research Center (CI-IPOP), Portuguese Oncology Institute of Porto (IPO Porto)/Porto Comprehensive Cancer Center Raquel Seruca (P.CCC) & CI-IPOP@RISE (Health Research Network)
[2]Department of Laboratory Genetics, IPO-Porto, Portugal
[3]Department of Pathology, IPO-Porto, Portugal
[4]Department of Medical Genetics, IPO-Porto, Portugal

The BRCA2 c.156_157insAlu insertion is a well-established founder variant in the Portuguese population, accounting for approximately one-quarter of pathogenic BRCA1/2 variants identified in families with hereditary breast and ovarian cancer. Detection of this insertion using standard short-read next-generation sequencing (NGS) pipelines is challenging due to alignment difficulties and variant-calling limitations.

We developed a Python-based tool to detect the germline BRCA2 c.156_157insAlu founder variant from short-read targeted NGS data generated from formalin-fixed paraffin-embedded (FFPE) tumour samples. The tool automates the analysis of BAM files by interrogating a defined genomic region on chromosome 13 (chr13:32893278–32893320) encompassing the insertion site. Using embedded pattern-matching algorithms, the script identifies reads containing either the insertion junction or the wild-type sequence and employs Samtools to extract and quantify matching reads. A graphical user interface was implemented to facilitate file selection, result storage, and error logging.

The method was evaluated using BAM files from six tumour samples derived from patients carrying the germline BRCA2 c.156_157insAlu variant and 393 tumour samples from patients confirmed to be negative for this variant through germline testing performed on blood-derived DNA at the Genetics Department of IPO-Porto. The tool achieved 100% sensitivity and 100% specificity, correctly identifying all six positive samples and all 393 negative samples.

This variant-specific approach enables accurate detection of the BRCA2 c.156_157insAlu Portuguese founder variant using short-read targeted NGS data from FFPE tumour samples. The results highlight the value of dedicated bioinformatic strategies for identifying known insertion variants that may be missed by conventional NGS analysis pipelines, particularly in clinical settings where recurrent founder mutations are prevalent.

THEME

Technologies, Methods and Data Analysis in Genomics

OC4

A genome language model for mapping DNA replication origins

Francisco Berkemeier[1]; Pauline L. Pfuderer[1]; Michael A. Boemo[1]

[1]University of Cambridge, England

DNA replication in human cells initiates at tens of thousands of genomic loci, yet the sequence determinants that define replication origins remain poorly understood. Recent advances in genome language models have demonstrated that deep learning approaches can extract biologically meaningful information directly from DNA sequence, providing a powerful framework for investigating regulatory information encoded in the genome.

We present ORILINX, a genome language model trained to identify replication origins from sequence alone. The model accurately predicts origin locations across the human genome and generalizes to other vertebrate species, suggesting the presence of conserved sequence signals associated with origin activity. To interpret these predictions in a mechanistic framework, we integrate them with mathematical models of DNA replication derived from replication timing data. Strikingly, sequence-derived origin probabilities correlate strongly with origin efficiencies inferred from the model, linking intrinsic sequence features with the large-scale dynamics and kinetics of genome replication and opening new opportunities for connecting fundamental genome biology with the mechanisms of replication stress that underlie cancer and therapeutic response.

THEME

Technologies, Methods and Data Analysis in Genomics

Transcriptomic signatures differentiating the successes and limitations of regeneration in a marine annelid

Inês Moutinho Cabral[1]; Ana P. Rodrigo[1]; Mafalda Soares[1]; Ana R. Grosso[1]; Pedro M. Costa[1*]

[1]UCIBIO - Unidade de Ciências Biomoleculares Aplicadas, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal

OC5

Regeneration, the ability of healing and restoring body structures following amputation or autotomy, is not only widespread across Eumetazoa, but also highly variable. Regenerative biology is a frontline area of research with deep implications for therapeutics and biomedicine. The span of model organisms is expanding beyond the paradigmatic models (planarians and axolotls) toward marine annelids because they can regenerate posterior and albeit more rarely, anterior body sections. Integrating the cellular and molecular processes underlying these outcomes could help explaining why this ability is lost in some animals and identify potential targets for inducing the restorative capacity in non-regenerative tissues, addressing key questions in developmental biology and biomedicine.

We hypothesize that gene expression signatures can assist understanding the differences between anterior and posterior regeneration in these organisms. Therefore, comparative histopathology and transcriptomics of posterior and anterior regeneration in *Hediste diversicolor* were performed at 7 and 14 days post amputation. To circumvent hindrances arising from reduced genomic annotation, we leveraged from a pipeline that integrates RNA-Seq–based transcriptomics and annotation tools with differential gene expression and active-subnetwork-oriented pathway analyses with the goal of unravel the transcriptomic signatures driving the regenerative process.

We discovered that *H. diversicolor* successfully regenerating posterior body sections had a transcriptomic profile more similar to that of non-amputated controls and appeared to balance cell survival and proliferation with cell death. In addition, a controlled inflammation was evident in organisms regenerating posteriorly, after short time of recovery amputation, then presumably leading way to regeneration. On the other hand, marine annelids enduring anterior regeneration had more variable outcomes, ranging from limited to unsuccessful, having some individuals developed coelom protrusions. These organisms seemed to be presenting homeostatic disruption and cellular stress. mTOR (cell metabolism and growth), Wnt (cell proliferation) and NF-κβ (strongly associated to the inflammatory response and cell survival) signaling were differently dysregulated across all contrasts and can be pointed as key players in the regeneration of *H. diversicolor*. However, further characterization of these mechanisms is still needed, especially through the expansion of experimental endpoints as well as to overcome some challenges and restraints of computational tools related with non-model organisms.

The outcomes point towards the existence of specific biological triggers that may be investigated to induce restorative ability in non-regenerative tissue. This knowledge can bear great promise for the field of biomedicine, mainly through translation into new therapeutics at the level of healing and recovery.

THEME

Functional Genomics and Gene Regulation

OC6

Single nucleus transcriptome atlas of the honey bee gut

Waldan Kwong[1]; Ana Barradas[1]; Cláudia Campos[1]

[1] GIMM - Gulbenkian Institute for Molecular Medicine, Lisbon, Portugal

A highly specific and beneficial microbiome resides within the gut of honey bees (Apis mellifera) and other eusocial bees. While this microbial community has been studied extensively in the past decade, the organ in which it resides – the bee gut – remains poorly understood. It is unclear how many cells make up the gut, what types of cells there are, what genes they express, and how the gut cells change over time or upon colonization with the microbiota.

Here, we use single nucleus transcriptomics with the 10X Genomics platform to characterize the cell types in the honey bee gut. We combine this with high resolution imaging, flow cytometry, and in-situ hybridization to localize and quantify gut cells. We find that distinctive cell types present, which differ in expression from those described in Drosophila, thus requiring the use of different marker genes to identify. Cell types differed by gut compartment (midgut, ileum, rectum), and expression changed in younger versus older bees, and in the presence of the microbiota. We find evidence of cell renewal via stem cells in the midgut, but not in the ileum and rectum.

Overall, this work provides a foundational understanding of the cellular identity and physiology of the bee gut, which will be beneficial for unraveling the mechanistic basis of host-microbe symbiosis in this model.

THEME

Biodiversity, Evolution and Environmental Genomics

TECHNOLOGY SPOTLIGHTS

Brief presentations by industry experts highlighting emerging technologies and their impact on genomics research.

SESSION I

Sequencing reimagined: introducing the Illumina Trupath Genome

Stef Kupers
Strategy and Market Development
Illumina

Beyond the Blind Spots: What Long-Read Sequencing is changing in Human Genomics

Ximo Garcia
Territory Account Manager - Iberia
PacBio

Despite remarkable advances in genomic technologies, important regions and classes of variation remain challenging to resolve using conventional approaches. Long-read sequencing is helping to overcome many of these limitations, providing access to previously hidden genomic information and enabling a more complete characterization of human genetic variation.
This presentation will review recent scientific evidence highlighting the impact of long-read sequencing across a range of applications. It will also explore how the ability to capture multiple layers of genomic information within a single workflow is reshaping research strategies and expanding our understanding of genome biology.
By examining key studies and emerging trends, the session will discuss how long-read sequencing is contributing to a more comprehensive and accurate view of the human genome.

SESSION II

New quality standard of sequencing with Aviti

Agnieszka Ciesielska
Sr Technical Sales Specialist, Distributors EMEA
Element Bioscience

Element Biosciences' Avidite Base Chemistry™ (ABC™) on the AVITI™ system innovates each step of sequencing chemistry, separating base detection from strand extension to deliver unsurpassed accuracy (Q50) while reducing costly reagent consumption. This form of sequencing achieves high accuracy and improved sequencing performance through difficult homopolymer regions and enables a diversity of applications that include single-cell RNA-seq and whole-human-genome sequencing. Learn about first benchtop high throughput Vitari platform. And as we go beyond sequencing discover our 5D multiomics enabled on Aviti24 platform.

Rethinking WGS with MGI: Why Quality, Speed, and Cost No Longer Require Trade-offs

Bernard Okere
MGI-Tech Application Support Team, Europe&Africa
MGI

National genomics infrastructures require high-throughput sequencing that balances diverse workloads with financial sustainability. MGI's DNBSEQ™ platforms address these needs through an optimized triad of quality, speed, and cost-efficiency. Using DNA Nanoballs (DNB) and Rolling Circle Replication, MGI eliminates PCR bias, ensuring clinical-grade accuracy. Advanced chemistry accelerates turnaround times and independent multi-flow-cell flexibility drives down costs per gigabase. Crucially, this technology is backed by a highly robust customer support structure of dedicated local professionals across Europe. This complete ecosystem ensures operational continuity, minimizes downtime, and empowers research networks to maximize funding.

Setting new boundaries for single-cell and spatial multiomics analysis with 10X Genomics

Ana Borges
Field Application Specialist
BonsaiLab

POSTER ABSTRACTS

P01

Organ-Specific Metastasis and Survival Across Cancers

António Gomes[1]; Ana Rita Grosso[2]

[1]Hospital Prof. Dr. Fernando Fonseca, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal
[2]Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal

Metastatic disease remains the leading cause of cancer-related mortality, yet prognostic stratification continues to rely predominantly on the site of the primary tumor. Increasing evidence suggests that the metastatic organ itself may exert an independent biological and clinical influence on survival, potentially reflecting organ-specific interactions between tumor cells and the metastatic microenvironment driven by underlying genomic and transcriptomic programs.

This project aims to establish robust and reproducible clinical phenotypes by evaluating the impact of metastatic organ site on survival outcomes independent of primary tumor origin.

Using large independent pan-cancer cohorts, including population-based and institutional datasets such as SEER and Memorial Sloan Kettering (MSK), we analyzed overall survival (OS) and, where available, progression-free survival (PFS) according to metastatic site. Rather than pooling heterogeneous datasets, a parallel analytical framework was adopted to assess the consistency and reproducibility of organ-specific survival patterns across cohorts.

Survival analyses were conducted using Kaplan–Meier estimation and multivariable Cox proportional hazards models, complemented by machine learning approaches including Elastic Net Cox regression and Random Survival Forests. Internal validation included cross-validation, calibration assessment, concordance index evaluation, and SHAP-based interpretability analyses to explore complex and potentially non-linear interactions between variables.

Preliminary results demonstrate that metastatic organ site carries a consistent and clinically meaningful prognostic signal across cancers, independent of primary tumor origin and metastatic burden. Distinct survival gradients were reproducibly observed between organs such as liver, brain, lung, and bone metastases across datasets. Machine learning models further suggest the presence of non-linear and interaction effects linking tumor origin, metastatic distribution, and disease burden.

These findings support the hypothesis that metastatic organotropism reflects biologically relevant disease behavior with direct clinical implications.

THEME

Human and Clinical Genomics

P02

Uncovering the mechanisms underlying Pancreatic ductal adenocarcinoma development and heterogeneity through computational approaches

Frederico Água[1]; Rita Lourenço[1]; Ana Rita Grosso[1]

[1]UCIBIO - Unidade de Ciências Biomoleculares Aplicadas, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal

Pancreatic ductal adenocarcinoma (PDAC) is among the most lethal human cancers, with a 5-year survival rate below 10% and no reliable methods for early detection. Consequently, more than 80% of patients are diagnosed with advanced or metastatic disease, severely limiting therapeutic opportunities and contributing to the poor overall prognosis. Incidence continues to rise, and PDAC is projected to become the second leading cause of cancer-related mortality by 2030. Genomically, PDAC is marked by near-universal KRAS mutations and frequent inactivation of CDKN2A, SMAD4, and TP53, yet these canonical alterations do not fully account for the extensive clinical and biological heterogeneity observed across tumors. Transcriptomic profiling has revealed distinct PDAC subtypes associated with differences in patient outcome, therapeutic vulnerability, and tumor microenvironmental features. However, the coexistence of multiple transcriptional states within individual tumors, and even within individual malignant clones, suggests additional layers of regulation shape subtype identity and evolution. The mechanisms underlying this diversity remain poorly defined. Transposable elements (TE) activity has recently emerged as a potential contributor to cancer biology, with aberrant TE expression implicated in transcriptional instability, innate immune activation, and chromatin remodeling. Despite these insights, the role of TEs in PDAC heterogeneity has been minimally explored. At the same time, advances in computational methods to accurately quantify the expression of these elements to take new insights from previously published data. We leveraged RNA sequencing data to investigate TE expression patterns in PDAC and assess how these patterns relate to established transcriptomic subtypes and cellular phenotypes. By integrating TE-derived signals with malignant cell state classifications, we aimed to uncover previously unrecognized sources of heterogeneity and gain insight into the regulatory factors that influence subtype origin, maintenance, and plasticity.

THEME

Functional Genomics and Gene Regulation

Genomics of Listeria monocytogenes persistence in food production environments

Inês Augusta Casaca[1]; Teresa Nogueira[1]; Célia Domingues[1]; Diana Mortinho[1]

[1]INIAV - Instituto Nacional de Investigação Agrária e Veterinária, I.P, Portugal

P03

Foodborne diseases have profound health and socioeconomic impacts. Listeriosis, caused by Listeria monocytogenes (Lm) - a Gram-positive bacillus - is frequently associated with ready-to-eat (RTE) products contaminated in Food Processing Environments (FPEs). Persistent strains can colonize FPEs for years, surviving routine sanitation. Since the genetic basis of this persistence remains unknown, the identification of universal genetic markers, through Whole-genome sequencing (WGS), would enable genotype-based classification, replacing current inconsistent phenotype-based criteria that compromise cross-study comparisons.

For that being said, this study tested two hypotheses: Whether persistent Lm strains share exclusive genetic determinants (genes, clusters, or mobile elements) that distinguish them from sporadic strains, or whether persistence results from natural selection and evolutionary processes.

A comparative genomic analysis was performed on 30 Lm isolates from dairy and bakery FPEs, previously sequenced with WGS and classified as persistent (n=17) or sporadic (n=13) based on molecular typing aproches. Assembly was performed with SPAdes (v3.15.5) and annotation with Abricate (v1.4.0). A pangenome was constructed using Anvi'o (v7.1) to visualize accessory gene distribution. Statistical associations between gene presence/absence and persistence phenotype were tested using Chi-square tests.

Mobile Genetic Elements (MGE), including prophages, plasmids, and integrative elements, were found and compared with the method above. Some of the identified genes are associated with antibiotic resistance, biocides tolerance and virulence. However, no specific genes, clusters, or MGEs were significantly associated with the persistent phenotype. Persistent and sporadic strains shared largely overlapping accessory genomes; some genes were exclusively to one group, but none could justify the phenotype.

This study suggests that persistence of Lm in FPEs is not driven by universal genetic markers, as no evident pattern of exclusive genes was detected in persistent genomes. On the contrary, it suggests that persistence emerges from interactions between genetic background, environmental selective pressures (sanitization protocols, temperature, salinity), ecological factors (bacteria and prophages), and stochastic processes (genetic drift, dormancy, recolonization).

Future control strategies should prioritize environmental and ecological interventions over strain genotyping alone. This project aims to establish robust and reproducible clinical phenotypes by evaluating the impact of metastatic organ site on survival outcomes independent of primary tumor origin.

THEME

Pathogen and Public Health Genomics

Single-cell transcriptomic profiling of airway epithelial models for cystic fibrosis

P04

Immacolata Zollo[1]; Tomás Faria Araújo[1]; Margarida Gama-Carvalho[1]; Carlos M. Farinha[1]

[1]BioISI - Biosystems & Integrative Sciences Institute, Faculty of Sciences of the University of Lisbon, Portugal

Cystic fibrosis (CF) is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, leading to impaired ion transport and chronic respiratory disease. In the field of CF, the gold standard for human tissue-derived respiratory models have been primary cultures of human bronchial epithelial (HBE) cells obtained for lung explanted tissue during lung transplantation. Although a common procedure several years back, lung transplantation in people with cystic fibrosis (pwCF) has decreased substantially since the introduction of highly effective modulators that can address the molecular and cellular defect caused by CFTR variants in a large majority of pwCF. Thus, primary cultures of human nasal epithelial (HNE) cells have become a widely used respiratory model for CF studies in the evaluation of novel CFTR modulators, as they preserve patient-specific characteristics and native CFTR function. Their accessibility through minimally invasive nasal brushing also makes them particularly valuable for personalized medicine approaches and prediction of therapeutic responses.

Despite their physiological relevance, the use of HNE cultures in large-scale and long-term studies remains limited by donor variability, restricted lifespan, and labor-intensive culture procedures. Therefore, identifying complementary respiratory epithelial models that closely resemble HNE cells is essential for expanding preclinical CF research.

In this study, we compared primary HNE cultures with the immortalized airway epithelial cell line BCi-NS1.1 using single-cell RNA sequencing (scRNAseq).

Both HNE cells, collected from a healthy donor, and BCi-NS1.1 were seeded on transwell filters and grown under air-liquid interface (ALI) conditions to let them differentiate in a pseudostratified airway epithelium. After 30 days in ALI, epithelia were dissociated to produce a single-cell suspension for 10x Genomics scRNAseq. The analysis revealed a similar cellular composition between the two models, with the main cell clusters that include basal cells, ciliated cells, secretory cells and ionocytes. The gene expression profiles of HNE and BCi-NS1.1 were also comparable, allowing to identify specific gene markers for the different cell groups. Among others, we identified KRT5 and TP63 as markers for basal cells, FOXJ1 and TMEM212 for ciliated cells, XBP1 and CYP2F1 for secretory cells, FOXI1 and KCNMA1 for ionocytes.

Together, these findings support the relevance of BCi-NS1.1 cell line as a complementary in vitro model while reinforcing the value of HNE cells as a physiologically representative platform for the development and testing of novel CFTR modulators.

THEME

Technologies, Methods and Data Analysis in Genomics

Transcriptomic Evidence for Anticipatory Immunity in a Socially Transmitted Infection Context

P05

Patricia C. Lopes[1]; Bennett Hardy[1]; Arly Adame[1]; Amir Adams[1]; Madeleine Chang[1]; Sean Mccallum[1]; Luana Oliveira[1]; Morgan Rea[1]; Sophia Santangelo[1]

[1]Chapman University, United States of America

Parasites are well known to influence host physiology and behavior following infection. However, evidence suggests that hosts may also respond to the mere threat of parasitism, even prior to infection. Although this phenomenon has been examined through behavioral ecology, its physiological components remain largely unexplored. We investigated whether perceived infection risk alters immune responsiveness in healthy male and female domestic canaries (Serinus canaria). Birds were allowed to observe conspecifics either infected with Mycoplasma gallisepticum (MG) - a pathogen that induces behavioral and visible physical symptoms - or sham-infected, symptom-free controls. After six days of observation, we collected whole blood samples from the observers. An aliquot of the blood remained untreated (unstimulated blood), and a second aliquot was stimulated in vitro with lipopolysaccharides (LPS-stimulated blood) to mimic an immune challenge. RNA was obtained from both sample types and sent for RNA-sequencing. Using multivariate adaptive shrinkage, we assessed whether there were differences in the strength of the response to LPS between treatments and between sexes. We found that LPS elicited large transcriptional changes in thousands of genes across all groups but also triggered treatment- and sex-specific response patterns. We identified 855 genes that were responsive only in observers of sham-infected animals and 579 genes responsive only in observers of MG-infected animals. Additionally, 172 genes responded to LPS exclusively in female observers of MG-infected birds, indicating a sex-biased signature. When examining for enrichment in gene ontology terms in these gene sets, we found that pathways related to lymphocyte activation, homeostasis, and differentiation were downregulated in observers of MG-infected birds. These findings indicate that perceived parasitism risk can change how animals respond to infections, expanding our understanding of host-parasite dynamics beyond direct infection.

THEME

Functional Genomics and Gene Regulation

Cholesteryl Hemiesters Reprogram Macrophage Polarization Toward a Novel Immunometabolic Phenotype

Daniela Pinto[1]; Otília Vieira[2]; Ana Rita Grosso[2]

P06

[1]UCIBIO - Unidade de Ciências Biomoleculares Aplicadas, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal
[2]iNOVA4Health, NOVA Medical School, Faculdade de Ciências Médicas da Universidade NOVA de Lisboa, Portugal

Cardiovascular diseases (CVDs) continue to be the leading cause of death globally, with approximately 19.8 million deaths each year. Atherosclerosis is the central pathological process underlying most CVDs and is responsible for major vascular events such as heart attacks and strokes. It is marked by the accumulation of lipids, cholesterol, and other components within arterial walls, where they can undergo oxidative modification to form cholesteryl hemiesters (ChE). Macrophages play a key role in this process by taking up oxidized lipids and orchestrating inflammatory responses, which critically influence plaque stability and the likelihood of adverse cardiovascular outcomes. The main goal of our study was to identify the phenotypic changes that macrophages develop in response to ChA exposure, the most abundant ChE found in humans, using mouse models. To confirm the presence of ChA in well-established models of atherosclerosis, the lipid content of aortas of ApoE knockout (ApoE −/−) and Nur77-ApoE double knockout (Nur77−/−ApoE−/−) mice was quantified using MS-based shotgun lipidomics. Our results indicate that Nur77 deficiency may improve the relevance of mouse models for investigating human-like atherosclerosis, given its more similar lipid profile. Posteriorly, to uncover the molecular pathways and cellular processes affected by ChA exposure, the transcriptomic profile of ChA-treated BMDMs was analyzed using RNA-seq. These results show that ChA-treated macrophages adopt a distinct immunometabolic state when compared with the gene signature of macrophage phenotypes previously described in an atherosclerotic environment. ChA induced a novel phenotype in BMDMs of mice, with increased proliferation, reduced phagocytosis and metabolism changes, shaping macrophages' behaviour within atherosclerotic plaques. Thus, this lipid can modulate plaque progression and stability and may be relevant to identify novel therapeutic targets to modulate macrophage function in atherosclerosis.

THEME

Human and Clinical Genomics

Robust and Scalable Spatial Analysis of Pancreatic Cancer Dissemination

Marta Bica[1]; Jose Espejo Valle-Inclán[1]; Peter Bailey[1]

[1]Botton-Champalimaud Pancreatic Cancer Centre, Lisbon, Portugal

P07

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal malignancies worldwide. One of the major challenges in treating PDAC is its rapid acquisition of resistance to conventional therapies, which severely limits long-term treatment efficacy. PDAC is characterized by a highly desmoplastic and fibrotic tumour microenvironment (TME). This dense stromal architecture acts as a physical and biological barrier, impairing drug delivery, promoting tumour progression, and contributing to immune evasion. In addition, increasing evidence highlights the importance of TME-related processes such as neoneurogenesis, angiogenesis, and immune system remodeling in pancreatic cancer progression and metastasis.

We are leveraging access to hundreds of patient-derived tumour samples to perform comprehensive multi-omics characterization of PDAC, including genomic, transcriptomic, and spatial profiling approaches, and better understand tumour biology across multiple molecular layers and ultimately contribute to personalized medicine strategies. One of our major objectives is the development of a spatial transcriptomics atlas of pancreatic cancer using technologies such as Visium HD and Xenium. These technologies preserve the spatial architecture of tumour tissue sections, allowing us to investigate not only the composition of cellular populations and transcriptional states, but also how cells physically interact within the TME.

By integrating spatial transcriptomics data across multiple samples, we aim to identify spatial niches associated with tumour progression, therapeutic resistance, and clinical outcomes. However, the analysis and integration of large-scale spatial transcriptomics datasets are computationally demanding. To ensure scalability and reproducibility, we are developing robust bioinformatics pipelines using Nextflow. These pipelines are designed to run reproducibly across both local high-performance computing clusters and cloud-based environments, enabling efficient and standardized analysis of large patient cohorts.

Through the integration of advanced spatial technologies, large-scale multi-omics data, and reproducible computational frameworks, we aim to generate a comprehensive understanding of PDAC biology that can ultimately support the development of more effective and personalized therapeutic strategies for pancreatic cancer patients.

THEME

Technologies, Methods and Data Analysis in Genomics

Addressing Imprinting Defects in Induced Pluripotent Stem Cells to Enhance their Biomedical Potential

Mariana Luz[1]; Ana Rita Grosso[1]; Simão Teixeira da Rocha[2]; Maria Arez[2]

P08

[1]UCIBIO - Unidade de Ciências Biomoleculares Aplicadas, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal
[2]IBB - Instituto de Bioengenharia e Biociências, Universidade de Lisboa, Portugal

Induced pluripotent stem cells (iPSCs) have revolutionized biomedicine, but unfortunately their use is still limited due to the epigenetic errors (such as imprinting defects) that often arise during reprogramming, which inevitably restricts their translational applications.

Genomic imprinting is responsible for monoallelic expression due to the differing DNA methylation patterns between alleles inherited from the father and mother. When methylation in these regions is disrupted, the monoallelic expression is affected. This has been linked to developmental disorders, tumorigenesis and, in the case of iPSC-derived cell products, abnormal cellular behaviour and compromised functionality.

In mouse iPSCs, modulation of the expression levels of the Yamanaka factors has been shown to favourably impact imprinting maintenance during reprogramming. It's important to highlight that the Yamanaka factors have binding sites within the affected imprinted regions, which might explain this impact on imprinting defects as it's possible that they alter epigenetic profiles locally. The influence of the Yamanaka factors on imprinting has not, however, been modulated in human iPSCs yet.

In this work, we leverage previously published data to characterize transcription factor occupancy in early stages of human iPSC reprogramming and aim to decode the influence of the Yamanaka factors in imprinting maintenance during human iPSC reprogramming.

THEME

Human and Clinical Genomics

P09

Genomic Insights into Marine Brucella from Stranded Dolphins Along the Portuguese Coast

Ricardo Dias[1]; Sandra Cavaco[2]; Miguel L. Grilo[2]; Mónica Nunes[1]; Catarina Fogaça[3]; Ana Beatriz Costa[3]; Sofia Pardal[3]; Ana Cristina Ferreira[2]

[1]Faculdade de Ciências da Universidade de Lisboa, Portugal
[2]INIAV - Instituto Nacional de Investigação Agrária e Veterinária, I.P, Portugal
[3]RALVT - Rede de Arrojamentos de Lisboa e Vale do Tejo, Portugal

Marine mammals are recognized as important sentinels of ocean health and may host pathogens with implications for both ecosystem stability and public health. Among these pathogens, marine-associated Brucella species have been increasingly reported in cetaceans worldwide. Although primarily considered wildlife pathogens, Brucella spp. includes zoonotic agents capable of infecting humans and therefore represent a potential biosecurity concern. Monitoring their circulation in marine ecosystems is important for understanding pathogen dynamics at the wildlife–environment interface and for strengthening biosurveillance systems within a One Health and global health security framework. This study investigates the occurrence and genomic characteristics of Brucella spp. in marine mammals stranded along the Portuguese coast.

Between 2022 and 2024, stranded marine mammals recovered along the Portuguese coast were examined through necropsy. A total of 59 animals were sampled, with tissue samples and biological fluids collected for molecular screening. Detection of *Brucella spp.* was performed using PCR-based assays. Positive samples were further investigated through bacterial isolation, and recovered isolates were characterized using native real-time whole genome sequencing (rtWGS) to determine their genetic relationships with previously described strains.

Molecular screening detected Brucella DNA in 23.7% of the analyzed animals, suggesting that infection may be more frequent than indicated by culture-based detection alone. Bacterial isolation was achieved in three common dolphins (*Delphinus delphis*), representing 5.1% of the examined animals. The infected dolphins exhibited heterogeneous pathological findings, including neurological manifestations compatible with neurobrucellosis in one case, as well as other systemic lesions. Genomic characterization revealed genetic profiles closely related to previously reported Atlantic *Brucella ceti* strains, suggesting regional circulation of host-adapted lineages.

The detection and genomic characterization of *Brucella ceti* in stranded dolphins along the Portuguese coast highlights the importance of marine wildlife surveillance as part of broader biosafety and biosecurity monitoring systems. Marine ecosystems may act as reservoirs for pathogens with zoonotic potential, and environmental changes may influence their circulation and transmission dynamics. Integrating wildlife pathogen surveillance into biosurveillance frameworks can strengthen early detection of emerging biological risks and support global health security strategies under a One Health perspective.

THEME

Pathogen and Public Health Genomics

Unravelling the Signaling and Regulatory Programmes forming Tertiary Lymphoid Structures (TLS)

P10

Aminata Souleymane Sow[1]; Tomás Guerreiro[1]; Rui Do Amaral Vieira[1,2]; Beatriz Filipe[1,2]; Pedro Gaspar[1,2,3]; Luís Graça[1,2]; Tomás Gomes[1]

[1]Gulbenkian Institute for Molecular Medicine, Portugal
[2]Faculdade de Medicina, Universidade de Lisboa, Portugal
[3]Unidade Local de Saúde de Santa Maria, Lisbon Academic Medical Centre, Portugal

A healthy immune system defends the body against disease by distinguishing self from foreign antigens. However, dysregulation of this process can lead to autoimmune diseases, in which the immune system mistakenly attacks the body's own cells and tissues. These conditions can lead to the formation of tertiary lymphoid structures (TLSs), ectopic lymphoid aggregates that form in non-hematopoietic organs under pathological conditions. Similarly to secondary lymphoid organs, TLSs can support local immune responses by facilitating interactions between immune cells, promoting B cell maturation and T cell activation. TLSs are increasingly recognized as key immune structures in various diseases; however, their development and functional regulation remain incompletely characterized.

To address this, we analyzed publicly available single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics datasets collected from cells of different tissues across several autoimmune diseases, including multiple sclerosis, rheumatoid arthritis, Sjögren's syndrome, and systemic sclerosis. Data processing and comparative analysis were performed using Seurat in R and Python, to identify shared and unique cell types, as well as gene expression programmes associated with cells in specific tissues. Cell–cell communication events were inferred using inference methods implemented in the LIANA package.

Combined with the detailed cross-tissue cell type annotation performed, we aim to generate a first map of the network coordinating TLSs in autoimmune disease, including the identification of chemokine and chemokine receptor interactions involved in immune cell recruitment to TLSs. Furthermore, we expect to identify key gene regulatory programmes governing TLS organization and maintenance across tissues. This analysis will be further extended to cancer by incorporating published datasets of cancer-associated TLSs. Revealing conserved and disease-specific features of cellular interactions, chemokine-mediated recruitment, and gene regulation within TLSs will establish a reference to inform future therapeutic strategies targeting TLS formation and function.

THEME

Technologies, Methods and Data Analysis in Genomics

The use of transcriptomic data on a non-model species to understand the basis of chemical polymorphism

P11

Joel Laia[1]; Carlos Rodríguez Fernandes[2]; Ana Cristina Figueiredo[2]; Rui Elias[3]; Helena Trindade[1]

[1]PFE - Plant Functional Ecology, CE3C - Centre for Ecology, Evolution and Environmental Changes
[2]Faculdade de Ciências da Universidade de Lisboa, Portugal
[3]Universidade dos Açores, Portugal

Studies in transcriptomics are an intricate challenge especially for non-model organisms with no captive bred lines. This is the case of the *Lamiaceae* family, an extremely diverse taxon of aromatic plants for both species and chemical compounds. The most studied, diverse, and abundant group of specialised chemical compounds is terpenoids, which are being investigated in the medical areas with proven health benefits due to their antioxidant and antimicrobial properties. The interesting endemic thyme species *Thymus caespititius Brot.* shows a stable α-terpineol chemotype in the Portuguese mainland and the Madeira archipelago, whereas in the archipelago of Azores it exhibits a considerable chemical polymorphism, with the phenolic chemotypes thymol and carvacrol. Studying the origin of chemotype polymorphism in the Azores allows to better understand the biosynthetic pathways not yet addressed in the species and also expand the knowledge on genes coding for these enzymes.

A previous genomics study on population structure (Laia et al., submitted) showed that the Azorean *T. caespititius* populations are monophyletic and highly divergent from the populations of the mainland and Madeira. In the present study we make use of transcriptomic data from *T. caespititius* populations from the Azores and mainland, comprising three different chemotypes, to access the expression and transcript sequences coding for enzymes and transcription factors involved in the terpenoid biosynthesis. Our results demonstrate a highly separate biosynthetic pathway for α-terpineol, congruent with the literature, and fine downstream differences between the genetically similar populations showing the carvacrol and thymol chemotypes.

THEME

Biodiversity, Evolution and Environmental Genomics

Identification of Conserved Molecular Mechanisms in Endometriosis Through Transcriptomic Meta-Analysis

Matilde Baptista[1]; Pereira F.[1]

[1]UTAD - Universidade de Trás-os-Montes e Alto Douro, Portugal

P12

Endometriosis is a chronic pathology that affects more than 10% of women, representing a prevalent cause of infertility and pelvic pain that is frequently underdiagnosed due to its high clinical and biological heterogeneity. Given this complexity, this study aimed to identify a molecular signature transversal to different phenotypes and experimental models, seeking to find consistent biomarkers of tissue progression. To this end, an integrative meta-analysis of five gene expression studies obtained from the NCBI Gene Expression Omnibus (GEO) database was performed.

The results revealed a coordinated dysregulation across three fundamental axes: the intrinsic hormonal biosynthesis of the lesions (NR5A1 and STAR); the activation of innate immune response mediators (C1R, C1S, and C3); and the modulation of cellular architecture and adhesion (COL6A3, DMD, and AFDN). The identification of these common molecular targets validates the existence of conserved mechanisms in the pathogenesis of endometriosis, regardless of the diversity of tissues or experimental models analyzed.

These data suggest that the integration of transcriptomic analyses allows the isolation of essential biological processes above the clinical variability of the disease. In this manner, potential molecular targets that explain the survival of endometrial tissue in extrauterine locations are identified, paving the way for new diagnostic and therapeutic intervention strategies.

THEME

Functional Genomics and Gene Regulation

Indel-friendly Implementation of Quantum Algorithms to Next Generation Sequencing Alignment

Helena J. Maki[1]; André Souto[1]; Margarida Carvalho[2]

P13

[1]LASIGE Computer Science and Engineering Research Centre, Faculty of Sciences of the University of Lisbon, Portugal
[2]BioISI - Biosystems & Integrative Sciences Institute, Faculty of Sciences of the University of Lisbon, Portugal

Alignment is a computationally expensive and necessary step in Next Generation Sequencing (NGS) data analysis. Quantum search algorithms have the potential to provide a quadratic speedup in relative to classical computing-based methods. Alignment of indels has thus far been difficult for many quantum algorithms due to their causing different portions of the sequence to be out of phase with one another.

Here I implement and compare various types of quantum NGS alignment algorithms and address the problem of gaps in the context of a quantum pattern matching algorithm. By partitioning the query being searched against the reference in the case of unaligned reads, the gap is constrained to one half of the query allowing for the out of phase half to be aligned through extending out from the seed provided by the in-phase segment. Algorithms are tested using simulated data within the size allowed for by current hardware constraints. Implementations of quantum pattern matching showed both fidelity in the presence of an exact match as well as a high degree of tolerance in situations where the best match available had a large hamming distance.

Though the present state of quantum hardware prevents us from executing our algorithm at the scale necessary to significantly reap the benefits of their quadratic speedup, we show viability of quantum based algorithms in NGS alignment as well as the utility of quantum-classical hybrid algorithms regarding the problem of aligning indels.

THEME

Technologies, Methods and Data Analysis in Genomics

P14

Long-read Sequencing for Pharmacogenomics in the Treatment of Psychotic Disorders

Beatriz Reis[1]; Inês Silva[2]; Susana Valente[3]; Mariana Ribeiro[3]; Sara Morais[3]; Ana Pais[3]; Paulo Silva[3]; Jorge Oliveira[3]; Maria Augusta Vieira-Coelho[4]; Filipa Carvalho[1]; Cristina Joana Marques[1]

[1]RISE-Health, Genetics Unit - Department of Pathology, Faculty of Medicine, University of Porto, Portugal
[2]Faculdade de Ciências da Universidade do Porto, Portugal
[3]Centro de Genética Preditiva e Preventiva, Instituto de Biologia Molecular e Celular, Portugal
[4]Department of Psychiatry and Mental Health, São João University Hospital Center, Porto, Portugal

Psychotic disorders require prolonged treatments with antipsychotic medications. However, patient therapeutic response and the occurrence of adverse drug reactions (ADRs) are highly variable among individuals. Pharmacogenomics offers a proactive approach to predict individual response to treatments and optimize therapeutic outcomes. Traditional short-read sequencing often fails to accurately genotype structurally complex pharmacogenes, such as CYP2D6. Long-read sequencing methodologies have emerged as a critical tool to overcome these genomic blind spots. By enabling complete haplotype phasing, structural variant resolution and comprehensive profiling of these complex and challenging regions, long-read sequencing significantly advances personalized psychiatric care.

We are conducting a large-scale pharmacogenomic study with an expected cohort of approximately 800 psychiatric patients. To identify clinically actionable variants involved in antipsychotic metabolism and drug interactions, we are employing targeted long-read sequencing using Oxford Nanopore Technologies with adaptive sampling, a real-time enrichment approach that selectively sequences regions of interest based on rapid alignment to a reference genome. Our approach focuses on key pharmacogenes with established clinical guidelines, as well as additional genes of potential relevance in psychiatry, including COMT, and the serotonin and dopamine receptor genes. The comprehensive analysis pipeline is designed to detect simultaneously single nucleotide polymorphisms (SNPs), structural variants (SVs), copy number variants (CNVs), and potentially DNA methylation patterns, providing a multi-omic profile.

As a proof-of-concept, an initial test sample was successfully sequenced and processed. Diplotype analysis identified the participant as normal metabolizer for CYP2D6 (*1/*41), CYP2C19 (*1/*38), and CYP3A4 (*1/*1). Additionally, the sample was classified as an intermediate metabolizer for CYP3A5 (*1/*3) and possessed the CYP1A2 *1F/*1F diplotype, which indicated inducibility of enzyme activity in response to smoking. The CYP2C19 haplotype polymorphisms were successfully confirmed using Sanger sequencing, validating the accuracy of long-read variant calling. In addition, we identified variants in the COMT, and serotonin and dopamine receptor genes, with potential clinical relevance due to their previously reported associations with adverse drug reactions.

Our preliminary results demonstrate the strong potential of targeted long-read sequencing for accurate pharmacogenes genotyping. By successfully capturing critical structural and nucleotide-based variants in a single-assay, this methodology establishes a reliable analytical framework for our study. Furthermore, scaling this comprehensive sequencing approach will allow the translation of complex pharmacogenomic drivers into tailored therapeutic interventions.

THEME

Human and Clinical Genomics

Deep generative multi-omics mosaic integration from cancer cell lines, organoids, and tumors

P15

Ana Rita Baião[1]; Susana Vinga[1]; Emanuel Gonçalves[1]

[1]Hospital Prof. Dr. Fernando Fonseca, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal
[2]Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Portugal

Advances in high-throughput sequencing and other assay technologies have enabled the extensive molecular and phenotypic characterization of cancer cells, generating large and complex multi-omics datasets that provide a more holistic view of tumor biology. However, integrating these data across different cancer cell models remains a significant challenge due to their high dimensionality, heterogeneity, and pervasive missing omics. Bridging preclinical models with patient tumors is particularly critical, as it can unlock translational insights that directly inform therapeutic decision-making.

Here, we propose a conditional adversarial multiview variational autoencoder (VAE) that can integrate genomics, transcriptomics, DNA methylation, and CRISPR-Cas9 gene essentiality screens data from both preclinical and clinically relevant cancer models, when available. Specifically, we focus on the mosaic integration of Cancer Dependency Map data comprising cancer cell lines and organoids, with GDC TCGA tumor patient samples. The mosaic nature of these datasets allows the model to handle the inherent incompleteness of real-world multi-omics data, where not all modalities are observed for every sample.

This approach aligns heterogeneous datasets into a shared latent space, effectively bridging the gap between preclinical models and patient tumors. The learned latent space achieves strong cross-modality alignment, with reconstruction performance reaching an average Pearson correlation greater than 0.9 across observed omics layers. Notably, the framework accurately imputes missing molecular profiles, enabling in silico reconstruction of molecular and functional genomic dependencies across tumor samples.

Beyond technical performance, a key finding is that the latent representations learned by the model capture biologically meaningful cancer phenotypes. The framework recovers well-established transcriptional programs, including the epithelial-to-mesenchymal transition (EMT), a process central to tumor invasion, metastasis, and resistance to therapy. Furthermore, the model captures features associated with drug-tolerant persister (DTP) cell states, a clinically relevant population of cells that evade drug treatment and contribute to the emergence of acquired resistance. Taken together, these results demonstrate that the adversarial multiview VAE does not merely compress data but encodes interpretable and translationally relevant biological variation.

Overall, this framework provides a unified approach for data integration and augmentation across cancer models and establishes a foundation for the identification of patient-specific vulnerabilities, ultimately supporting the prioritization of effective therapeutic strategies for clinical validation.

THEME

Technologies, Methods and Data Analysis in Genomics

A bioinformatics framework for integrated small RNA analysis in grapevine-pathogen interactions using a telomere-to-telomere reference genome

P16

Lucas Monteiro[1]; João Proença Pereira[1]; Andreia Figueiredo[1]; Margarida Gama-Carvalho[1]

[1]BioISI - Biosystems & Integrative Sciences Institute, Faculty of Sciences of the University of Lisbon, Portugal

Small RNAs (sRNAs) are central regulators of plant gene expression, genome stability and immune responses. However, their characterization in grapevine remains challenging due to the complexity of repetitive genomic regions, incomplete annotation of sRNA loci, and the recent availability of telomere-to-telomere (T2T) genome assemblies.

Within a European collaborative project investigating grapevine responses to *Plasmopara viticola* and *Erysiphe necator*, we established a bioinformatics workflow for the analysis of small RNA sequencing datasets generated from multiple grapevine cultivars displaying distinct levels of disease tolerance.

The workflow combines re-annotation of known grapevine microRNAs against the latest T2T Vitis vinifera reference genome, standardized processing of small RNA sequencing data, genome-wide identification of small interfering RNA (siRNA) loci using ShortStack, and integration with transcriptomic and DNA methylation datasets. Particular emphasis was placed on improving the genomic localization of annotated miRNAs and on the detection of siRNA-producing regions associated with transposable elements, promoters and gene-rich regions.

Application of this framework enabled the generation of high-confidence count matrices for both miRNAs and siRNA clusters across pathogen infection experiments, providing the foundation for downstream differential expression, target prediction and epigenomic integration analyses.

This work illustrates how recent advances in grapevine genome assemblies, combined with dedicated bioinformatics workflows, enable a more comprehensive characterization of regulatory small RNAs and establish a framework for integrative multi-omics studies in grapevine.

THEME

Technologies, Methods and Data Analysis in Genomics

Transcriptomic Insights into Membrane Transport Dysfunction in Spinocerebellar Ataxia Type 3

Adriana Vaz[1,2,3]; Tiago Moreira-Gomes[1,2,3]; Clévio Nóbrega[1,2]; Carlos A. Matos[1,2]

P17

[1]Faculdade de Medicina e Ciências Biomédicas, Universidade do Algarve
[2]Algarve Biomedical Centre - Research Institute
[3]Doctoral program in Biomedical Sciences, Universidade do Algarve

Spinocerebellar ataxia type 3 (SCA3), also known as Machado-Joseph disease (MJD), is an autosomal dominant neurodegenerative disorder, which belong to the group of polyglutamine diseases. SCA3 is caused by an unstable CAG expansion in the ATXN3 gene, resulting in the translation of a polyglutamine-expanded ataxin-3 protein. This disorder is characterized by progressive motor impairments, which include cerebellar ataxia, impaired balance and coordination, among others. These motor deficits arise from neuronal degeneration, primarily affecting the cerebellum, brainstem and basal ganglia. Currently, the treatment of SCA3 is limited to symptom management, as no therapies are available to slow or halt disease progression. A major challenge in developing effective treatments is the lack of a complete understanding of the molecular mechanisms underlying this disease. Therefore, the identification of novel therapeutic targets becomes highly relevant.

Membrane transporters, mainly ion channels, are essential for maintaining neuronal function as they are responsible for upholding membrane potential, excitability, and overall cellular homeostasis. Disruption of these transport mechanisms has been implicated in the pathophysiology of a wide variety of neurodegenerative disorders, including polyQ disorders and spinocerebellar ataxias. Given the central role of ion channels in regulating both neuronal excitability and calcium dynamics, their dysfunction could contribute to impaired homeostasis in SCA3, exacerbating the disease pathology. Considering this evidence, transmembrane transport emerges as promising therapeutic targets for SCA3.

To identify candidate genes and pathways associated with transmembrane transport dysfunction in the disease, transcriptomic analyses of the cerebellum of a SCA3 mouse model were performed. Preliminary RNA sequencing data revealed changes in the expression of a set of transport-related genes, which suggests a dysregulation in pathways involved in membrane transport and ion homeostasis. Ongoing work focuses on characterizing these transcriptomic changes through additional bioinformatic analyses in order to identify key candidate genes for subsequent functional validation in cellular and animal models. By combining transcriptomic and functional approaches, this study highlights transmembrane transport pathways as potential contributors to SCA3 pathogenesis and therapeutic targets.

THEME

Human and Clinical Genomics

P18

Time to say goodbye: long-read sequencing is ready to replace array diagnostics in clinical routine

Ben Liesfeld[1]; Lena Hausdorf[1]; Aina Pi Roig[2]; Ana De Las Peñas[1]; Marco Graf[3]; Denny Popp[4]; Viktoria Bothe[4]

[1]Limbus Medical Technologies GmbH, Germany
[2]Nostos Genomics GmbH, Germany
[3]MVZ Dr. Eberhard & Partner Dortmund GbR, Germany
[4]Institute of Human Genetics, University of Leipzig, Germany

Structural variants (SVs) and copy-number variants (CNVs) are a major source of clinically relevant genomic variation, yet no single diagnostic approach robustly captures their full spectrum. Array-based diagnostics provide reliable detection of large CNVs but lack breakpoint resolution and cannot characterize complex rearrangements. Conversely, long-read sequencing enables precise SV detection but remains biased against very large CNVs when relying on breakpoint-based callers alone.

Here, we demonstrate how long-read sequencing can overcome these limitations by combining conventional SV calling with a complementary coverage-based CNV analysis. Using bin sizes approximately matching read length (5.25 kb), the CNV analysis is optimized for large copy-number events and naturally complements SV calling, which excels at breakpoint-resolved and complex variants.

Across multiple clinical cases, we show that integrating both approaches enables comprehensive variant interpretation including large complex rearrangements requiring combined CNV and SV signals, small deletions detectable only by SV calling, and concordant CNV/SV events affecting clinically relevant cancer genes. Together, these examples illustrate that integrated long-read analysis can consolidate the strengths of array and sequencing diagnostics into a single assay, supporting the routine replacement of array-based testing in clinical genomics.

THEME

Human and Clinical Genomics

P19

From VUS to Likely Pathogenic: Lipidomics as Functional Evidence for GM2A Variant Interpretation

Ana Margarida Gonçalves[1]; Carla Caseiro[1]; Helena Ribeiro[1]; Cátia Magro[1]; Célia Ferreira[1]; Filipe Godinho[2]; Tiago Oliveira[2]; Dulce Quelhas[1]

[1]CGMJM - Centro de Genética Médica Jacinto Magalhães, Porto, Portugal
[2]ULSSJ - Unidade Local de Saúde de São José, Portugal

The widespread implementation of next-generation sequencing (NGS) has transformed rare disease diagnostics but has also increased the identification of variants of uncertain significance (VUS). Although genomic data can reveal strong candidate variants, establishing clinical relevance often remains challenging, particularly in ultra-rare disorders lacking accessible functional assays. The interpretation of VUS is especially problematic when standard biochemical investigations are inconclusive. GM2 gangliosidosis AB variant is a rare lysosomal disorder caused by pathogenic variants in GM2A, which encodes the GM2 activator protein required for the degradation of GM2 ganglioside by β-hexosaminidase A. β-hexosaminidase assays are typically normal, despite impaired GM2 degradation.This study aimed to evaluate whether plasma glycosphingolipid profiling could provide functional evidence supporting the pathogenicity of GM2A variants initially classified as VUS.

Two siblings presenting with a slowly progressive neuromuscular phenotype underwent extensive clinical, neurophysiological, imaging and genetic investigations. β-hexosaminidase activities were measured in plasma and leukocytes. Mendeliome sequencing identified candidate variants in GM2A but none of the variants has been previously reported for GM2A deficiency. Given the discrepancy between clinical suspicion and routine biochemical findings, investigation continued by plasmatic GM2 gangliosidose analysis by LC-MS/MS, including total GM2 quantification and GM2/GM3 ratio assessment.

The two affected brothers developed progressive proximal lower-limb weakness during early adulthood, associated with gait instability and chronic motor neuronopathy. Muscle MRI demonstrated selective muscle involvement, while muscle biopsy supported a neurogenic process. Multiple neurogenetic investigations were inconclusive. Hexosaminidase activities were within the normal range, which excluded classical Tay-Sachs and Sandhoff disease. Mendeliome sequencing identified compound heterozygosity for two GM2A missense variants, c.428G>A (p.Gly143Glu) and c.496G>A (p.Gly166Arg), both initially classified as VUS according to ACMG criteria. Segregation was confirmed in the maternal allele Maternal, but not available in the father as it was already deceased. Evaluation of the functional impact of these variants required plasma lipidomic analysis. Both siblings showed markedly elevated total GM2 concentrations and increased GM2/GM3 ratios compared with controls. The lipidomic analysis demonstrated a biochemical profile revealing an impaired GM2 metabolism. Together, these data strengthened the pathogenicity assessment of the identified variants.

This work highlights an important limitation of NGS-based diagnostics: the frequent identification of VUS without sufficient evidence for clinical interpretation. In rare metabolic diseases, lipidomic biomarkers may provide critical functional evidence when conventional biochemical testing is non-informative. The integration of genomic, biochemical and clinical data contributed to solving diagnostic uncertainty in this family and supports the use of advanced lipidomics as a complementary strategy for VUS assessment and variant reclassification in lysosomal disorders. These findings illustrate how disease-specific lipidomic biomarkers can serve as biochemical evidence of disease, helping bridge the gap between genomic discovery and variant interpretation in ultra-rare lysosomal disorders.

THEME

Human and Clinical Genomics

Exploring the CFTR gene in publicly available cancer databases

Cori Mcfarlane[1]

[1]Instituto Superior Técnico - Universidade de Lisboa, Portugal

P20

Cystic fibrosis (CF) is a disorder caused by inheritance of pathogenic variants of the CFTR gene that lead to impaired function of the CFTR (Cystic Fibrosis Transmembrane Conductance Regulator) protein. CFTR is a chloride anion channel at epithelial surfaces that regulates salt and fluid homeostasis. Ion transport is impaired in CF patients, leading to build up of thick mucus in the airways, causing breathing difficulties and chronic infections.CF has historically been a devastating disease with patients in the early 1900s rarely surviving early childhood. Treatment advances, notably modulator therapies that restore function of the defective protein, are a clinical success story, with CF patients born in 2025 predicted to have a median survival age of >65 years. Whilst modulator therapies restore ion channel function, the disease state persists, including sustained inflammation. Therefore, this aging population of CF patients face novel treatment challenges that require modern solutions.At the forefront of this is the greater incidence of cancer seen in this cohort, with mounting evidence supporting CFTR as a tumour suppressor gene. The exact mechanistic link between CFTR and cancer remains unclear. However, there is increasing evidence implicating CFTR in fundamental cellular processes frequently dysregulated in cancer, including developmental processes, regeneration, differentiation and epithelial-mesenchymal transition (EMT).

This study aimed to apply bioinformatics approaches to interrogate publicly available cancer databases, to gain insights into the mechanisms implicating CFTR in these processes and how CFTR aberrations may be cancer-promoting. Using RNA-seq data from the DepMap Public 26Q1 release, CFTR expression was examined across colorectal cancer (CRC) cell lines and Gaussian mixture modelling identified two CRC subpopulations: a CFTR-low group and a CFTR-medium/high group. A CDH1/CDH2 ratio was calculated to capture bidirectional cadherin expression as a measure of epithelial-mesenchymal identity. CFTR-expressing colorectal lines exhibited a significantly higher CDH1/CDH2 ratio compared to CFTR-low lines ($p = 0.0079$), suggesting a more epithelial transcriptional state. Notably, a subset of CFTR-low lines showed concurrent CDH1 loss and marginal CDH2 gain, potentially suggesting a partial EMT phenotype. Median expression of EMT-associated genes was compared between CFTR-high and CFTR-low CRC cell lines, with genes such as HIF1A expressed at higher levels in CFTR-low lines. Differential expression analysis was performed on RNA-seq data from the TCGA Colon Adenocarcinoma dataset (cBioPortal, 2025), comparing CRC samples harbouring mutant versus wild-type CFTR to identify significantly differentially expressed genes. Mutant CFTR samples displayed a predominantly upregulated transcriptional profile. Pathway enrichment analysis (KEGG 2026) identified predominant enrichment in the PI3K-AKT, calcium, and neuroactive ligand-receptor interaction signaling pathways. Transcription factor enrichment analysis (ChEA 2022) identified 7 transcription factors as candidate regulators of this upregulated gene signature.

THEME

Human and Clinical Genomics

Comparative genomic and transcriptomic analysis of HSat1A, HSat1B, and HSat2/3 satellite DNA across Primates

Daniel Eleutério[1]; Carlos Vila-Verde[1]; Raquel Chaves[1]; Sandra Louzada[1]

[1]UTAD - Universidade de Trás-os-Montes e Alto Douro, Portugal

P21

Human satellite DNA families, including HSat1A, HSat1B, and HSat2/3, are major components of pericentromeric heterochromatic regions of the genome. Emerging evidence shows that HSat1A is transcriptionally active and overexpressed in some cancer cell lines, and HSat2/3 were shown to be actively transcribed under specific physiological, developmental, and pathological conditions. However, the evolutionary dynamics of these repeats, particularly how genomic sequence variation influences their transcription across closely related species, remains poorly characterized. This study presents a comparative sequence analysis of HSat1A, HSat1B, and HSat2/3 across genomic and transcriptomic datasets from human (*Homo sapiens*) and non-human primates, including chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), and orangutan (*Pongo abelii*).

Using long-read sequencing data alongside diverse RNA-sequencing (RNA-seq) datasets, we characterized the sequence conservation, copy number variation, and transcriptional diversity of these satellite families. Specialized bioinformatic pipelines for repetitive elements were utilized to quantify these satellite-derived reads and assess their transcription source.

This comparative analysis is designed to investigate potential lineage-specific patterns of amplification and sequence divergence. Specifically, this study will evaluate whether core monomeric motifs of HSat1A and HSat1B exhibit conservation across the Primates family and compare this with the hypothesized rapid evolutionary turnover in HSat2/3 families. Furthermore, the relationship between genomic abundance and transcriptional activity will be examined, addressing whether specific sequence variants and subfamilies of HSat1 and HSat2/3 are preferentially transcribed, or if their expression is a consequence of passive, pervasive read-through transcription. Also, species-specific transcript variants will also be examined, as that might reflect distinct evolutionary pathways since the divergence of the great ape lineages.

Ultimately, this comparative framework aims to open a discussion on the relationship between the genomic template and the transcriptomic output of major satellite families in Primates. By raising these key questions, this study seeks to contribute to a deeper understanding of repetitive DNA regulation and its potential functional roles in primate genome evolution.

THEME

Biodiversity, Evolution and Environmental Genomics

Transcriptomic analysis to uncover common mRNA signatures and regulators across smoking-related cancers

Rita Sousa[1]; Pedro G. Ferreira[1]; Alexandra Moreira[2]; Isabel Pereira-Castro[2]

P22

[1]Faculdade de Ciências da Universidade do Porto, Portugal
[2]i3S - i3S - Instituto de Investigação e Inovação em Saúde, Porto, Portugal

The tobacco epidemic is identified as a premier global health crisis, claiming over 7 million lives annually. It is considered a risk factor for various cancers, especially in head and neck, lung and bladder cancers. The co-transcriptional processes, 3' untranslated region alternative polyadenylation (3'UTR-APA) and intronic polyadenylation (IPA) produce mRNA isoforms of varying lengths due to the use of multiple polyadenylation sites throughout a gene. These variations in length are not merely structural, they also dictate regulatory traits, including the transcript's stability, subcellular localization, and translation efficiency. The relation between APA and disease is most evident in the described preference of rapidly proliferating and transformed cells for mRNA isoforms with short 3'UTRs. This shortening can be a functional driver of malignancy, which can eliminate miRNA and RNA-binding protein binding sites typically found in full-length 3'UTRs, and can actively promote an oncogenic phenotype.

In this work, we aim to use bioinformatics tools to obtain smoking-specific mRNAs produced by 3'UTR-APA and IPA in lung, bladder and head and neck cancers that may serve as potential risk biomarkers and therapeutic targets, and to identify common regulators of these signatures. For that, we previously identified 460 smoking-associated mRNA signatures exclusive to smokers with lung, bladder and head and neck cancers. Subsequently, a bioinformatic pipeline was developed to remove possible false positives emerging from overlapping genomic regions with other neighboring genes. The corrected dataset will then be used for motif discovery with Homer and MEME Suite. Of the 460 initial smoking-associated signatures, 384 were retained as high confidence genes with the pipeline built. From these, 366 were filtered for 3'UTR-APA (84.7%) and 18 for IPA (69,2%). With these results, we will analyse the 3'UTR of these signatures using Homer and MEME Suite and from the analysis of both tools, we expect to determine whether common cis-regulatory motifs are shared across most of these mRNAs or whether they can be classified into distinct co-regulated groups. It will also be possible to identify RNA-binding proteins that regulate the expression of these signatures.

Overall, this now filtered dataset of 384 smoking-specific mRNA signatures represents a promising resource for the identification of future risk biomarkers and therapeutic targets in smoking-related cancers. Moreover, this work will identify common cis-regulatory sequences and regulatory networks that can be used to modulate the mRNA signatures expression in these cancers.

THEME

Functional Genomics and Gene Regulation