REPÚBLICA PORTUGUESA
SAÚDE

SNS SERVIÇO NACIONAL DE SAÚDE

Instituto_Nacional de Saúde Doutor Ricardo Jorge

GenomePT SYMPOSIUM

3rd Symposium of the National Research Infrastructure for Genome Sequencing and Analysis

Abstract Book

November 17th, 2023

Headquarters' Auditorium
National Institute of Health Doutor Ricardo Jorge
Av. Padre Cruz, 1649-016, Lisbon

Instituto_Nacional de Saúde Doutor Ricardo Jorge

biocant CENTRO DE INOVAÇÃO EM BIOTECNOLOGIA

UC GENOMICS

iBiMED Universidade de Aveiro

IBMC INSTITUTO DE BIOLOGIA MOLECULAR E CELULAR

cibio

ipatimup

CBMR CENTRE FOR BIOMEDICAL RESEARCH

MM Instituto de Medicina Molecular | João Lobo Antunes

cbma centre of molecular and environmental biology

INSTITUTO GULBENKIAN DE CIÊNCIA

IPOPORTO INSTITUTO PORTUGUÊS DE ONCOLOGIA DO PORTO FG, EPE
EPOP - ESCOLA PORTUGUESA DE ONCOLOGIA DO PORTO

UCIBIO Applied Molecular Biosciences Unit

UNIVERSIDADE BEIRA INTERIOR

BioISI Biosystems and Integrative Sciences Institute

This publication compiles the abstracts submitted to the 3rd Symposium of the National Research Infrastructure for Genome Sequencing and Analysis – Genome^{PT}. This event was organized by the consortium member Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA).

The information contained in the abstracts is the sole responsibility of the respective authors.

Instituto Nacional de Saúde Doutor Ricardo Jorge,
Headquarters' Auditorium,
Av. Padre Cruz, 1649-016, Lisbon.
Tel: +351 217 519 200
www.insa.pt

Organizing Committee

Luís Vieira, INSA
João Paulo Gomes, INSA
Luís Coelho, INSA
Rita Ferreira, INSA
Sara Rangel, INSA
Sílvia Duarte, INSA

Contacts:

genomept@insa.min-saude.pt
www.genomept.pt

Scientific Committee
Alexandre Raposo, IMM
Ana Rita Grosso, Univ. Nova
Conceição Egas, Biocant
Gabriela Moura, Univ. Aveiro
Henriqueta Silva, Univ. Coimbra
Isabel Gordo, IGC
Joana Xavier, Univ. Algarve
Jorge Oliveira, IBMC
José Melo-Ferreira, CIBIO
João Paulo Gomes, INSA
Manuel Lemos, UBI
Manuel Teixeira, IPO Porto
Margarida Gama Carvalho, BioISI
Pedro Santos, Univ. Minho

Communication Office
Diana Ferreira, INSA
Florisa Maria Mineiro, INSA
Francisco Branco, INSA
Joana Ribeiro, INSA
Nuno Crespo, INSA
Tiago Belim, INSA

Symposium host
Tiago Belim, INSA

Registration and payment management
Biocant

Graphic design
Miguel Pinto, INSA
Luís Coelho, INSA
Rita Ferreira, INSA
Sara Rangel, INSA
Sílvia Duarte, INSA

GenomePT website maintenance
Sérgio Cabaço

Acknowledgements
Conceição Egas, Biocant

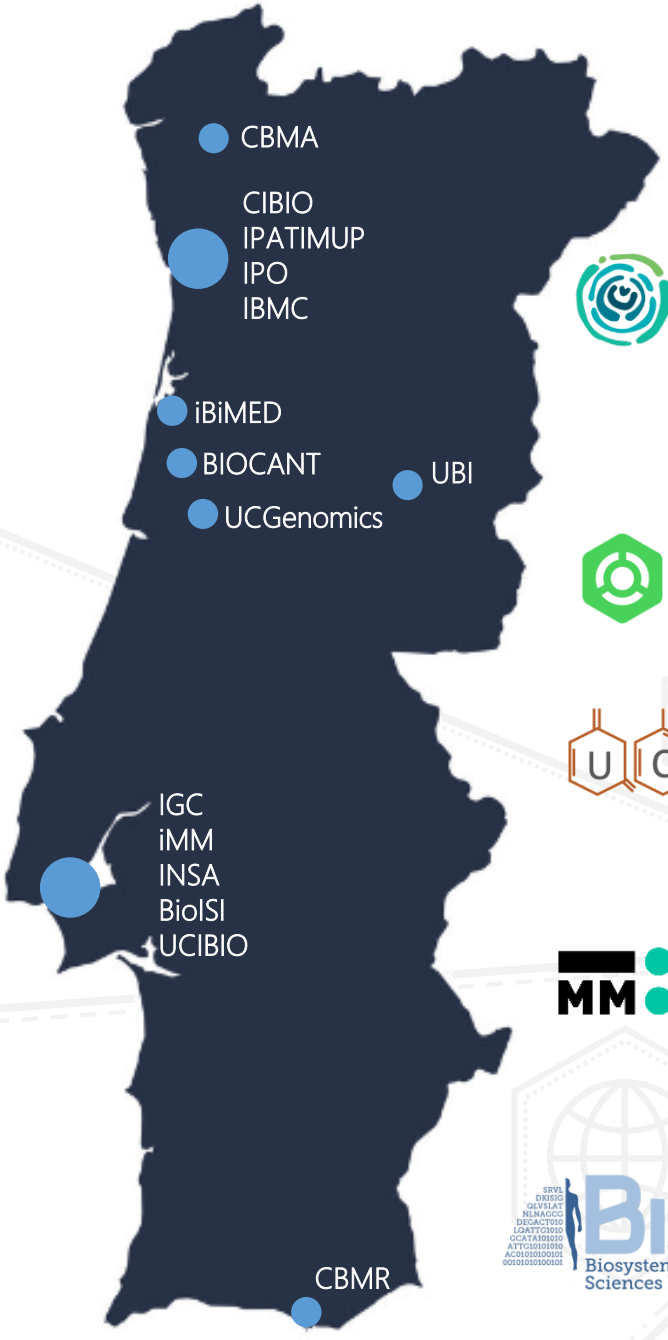
INSA's Training Office

Index

Index

Index

Consortium



centre of molecular and environmental biology

cibio

ipatimup

IBMC
INSTITUTO DE BIOLOGIA MOLECULAR E CELULAR
INSTITUTE FOR MOLECULAR AND CELL BIOLOGY

IPOPORTO
INSTITUTO PORTUGUÊS DE ONCOLOGIA DO PORTO FG, EPE
EPOP - ESCOLA PORTUGUESA DE ONCOLOGIA DO PORTO

iBiMED
Universidade de Aveiro

biocant
CENTRO DE INOVAÇÃO EM BIOTECNOLOGIA

UNIVERSIDADE BEIRA INTERIOR

UC GENOMICS

INSTITUTO GULBENKIAN DE CIÊNCIA

Instituto de Medicina Molecular | João Lobo Antunes

Instituto_Nacional de Saúde
Doutor Ricardo Jorge

BioISI
Biosystems and Integrative Sciences Institute

UCIBIO
Applied Molecular Biosciences Unit

CBMR
CENTRE FOR BIOMEDICAL RESEARCH

Venue

Instituto Nacional de Saúde Doutor Ricardo Jorge – Headquarters
Av. Padre Cruz, 1649-016, Lisbon

Symposium partners

Quilaban
uma fórmula de vida

Grupo I.L.C.

MGI

BIOPORTUGAL
Químico, Farmacêutica, Lda.

soquímica

VELSERA

avantor™

ThermoFisher
SCIENTIFIC

alfagene

3rd Symposium of the National Research Infrastructure for Genome Sequencing and Analysis - GenomePT

The **GenomePT Symposium 2023** is a one-day in-person event that centers around scientific and technological advancements in Genomics. This symposium places a special emphasis on the utilization of cutting-edge next-generation sequencing technologies and computational methodologies for data analysis, addressing a wide array of scientific inquiries.

The primary objective of the GenomePT Symposium 2023 is to serve as a pivotal platform for fostering connections among genome researchers, scholars, healthcare professionals, and graduate students. It aims to encourage collaboration and knowledge exchange within the field.

GenomePT is a distributed genome sequencing and analysis infrastructure embedded in the Portuguese Roadmap of Research Infrastructures. This consortium unites more than 50 dedicated researchers and technical experts, forming a robust infrastructure with a wealth of technological and bioinformatics expertise.

GenomePT provides comprehensive sequencing and bioinformatics services to a diverse range of genome projects, working in conjunction with both national and international partners. This includes collaborations with vital sectors such as the national healthcare system, as well as industries related to food, pharmaceuticals, biotechnology, winemaking, and the fishing sector.

Our overarching mission is to enhance research capabilities and align our strategic research objectives with regional and national developmental priorities. This approach enables us to make significant contributions to the advancement of regional development, bolstering the national economy, and ensuring the growing of highly skilled human resources within Portugal.

The organizing committee

Programme

9:00	**Reception and registration of participants**
9:35	**Welcome and opening of the symposium** Board of Directors – National Institute of Health Doutor Ricardo Jorge (INSA)
9:45	**Present and future of Genome^{PT}** Manuel Santos - Genome^{PT} Consortium Coordinator / Director - MIA (University of Coimbra)
10:10	Round table - **The future of genomics and its impact in science and society** Teresa Firmino – Moderator / Science Editor – Público Maria Gomes da Silva – Director of Hematology Service – IPO João Paulo Gomes – Head of the Genomics and Bioinformatics Unit - INSA José Matos – Head of the Molecular Biology Laboratory – INIAV
11:00	Poster walking: **Discover more works in genomics**
11:15	**Coffee-Break**
11:30	**Oral Communications I** **OC1** – **Microenvironment in breast cancer sentinel lymph nodes uncover two distinct immunoinflammatory profiles independent of metastasis status** Joana Ribeiro – FMUC **OC2** – **Effects of TiO2 nanoparticles on the genome-wide methylation of human epithelial intestinal cells** Célia Ventura – INSA **OC3** – **Differential transcription factor binding reveals the gene modules governing regulatory T cells and enables decoding mutational hotspots in human disease** Alexandre Raposo – IMM

Programme

12:10	**Clinical genomics: current applications, hurdles, and the road to future developments** João Freixo Clinical Director - Centre for Predictive and Preventive Genetics of the Institute of Molecular and Cellular Biology *(CGPP-IBMC)*
12:35	**The Portuguese strategy for genomic medicine and European context** Astrid Vicente Head of the Department of Health Promotion and Noncommunicable Diseases prevention – INSA
12:50	**Lunch**
14:10	**Machine learning as a tool to uncover the genotype-phenotype mapping: advances and current limitations** Miguel Rocha Director of the Masters Course in Bioinformatics – Minho University
14:35	**Forecasting species response to climate change using genomic prediction: advances and limitations** Octávio F. Paulo Head of Computational Biology and Population Genomics Group of cE3c & CHANGE – FCUL
15:00	**Oral Communications II** **OC4 – A new understanding of the transmission landscape of *Mycobacterium bovis* boosted by single cell analyses, genomics, and ecological modeling** Mónica Cunha – cE3c, FCUL **OC5 – Maternal HIV-1 genotypes and resistance mutations assessment in mother-to-child transmission with next-generation sequencing: perspectives of using this powerful tool** Diogo Ramos – INSA

Programme

OC6 – **Detection of the cell fusion agent virus in *Aedes aegypti* mosquitoes in Cape Verde using viral metagenomics**

Líbia Zé-Zé – INSA

OC7 – **Characterization of the microbiome of *Aedes albopictus* populations in Spain and São Tomé: implications for vector control**

Tiago Melo – IHMT, UNL

15:50	Exhibition: **Meet the technology partners for your goals**
16:05	**Coffee-Break**
16:20	**"One Genome… One Health" – How ATCG can help the nexus of animal, human and environmental health** (live online lecture) Carlos das Neves Chief Scientist – European Food Safety Authority (EFSA)
16:45	**What's new in the human (pan)genome?** (live online lecture) Adam M. Phillippy Head of the Genome Informatics Section – National Human Genome Research Institute (NHGRI)
17:10	**Final remarks and closing of the symposium** Luís Vieira – Head of the Technology and Innovation Unit – INSA

GenomePT
SYMPOSIUM

Round Table

The Future of Genomics and Its Impact in Science and Society

Teresa Firmino, JO

Moderator *– Science Editor*

Público newspaper

João Paulo Gomes, PhD

Head of Genomics and Bioinformatics Unit

National Health Institute Doutor Ricardo Jorge (INSA)

Maria Gomes da Silva, MD, PhD

Head of Hematology Service

Portuguese Institute of Oncology in Lisbon (IPO)

José António Matos, PhD

Head of Molecular Biology Laboratory

National Institute for Agricultural and Veterinary Research (INIAV)

GenomePT
SYMPOSIUM

Lectures

Present and Future of GenomePT

Manuel Santos, PhD

GenomePT Consortium Coordinator, Professor Faculty of Medicine, Director of MIA

Multidisciplinary Institute of Ageing (MIA) Coimbra University

"This presentation will center on the advancements in genomics in Portugal since 2008.

The GenomePT consortium was founded as part of the National Roadmap of Research Infrastructures (RNIE) initiated by Fundação para a Ciência e a Tecnologia (FCT) in 2014. Its primary objective was to unify the national laboratories engaged in genome and transcriptome sequencing into a cohesive distributed research infrastructure.

GenomePT's laboratories are strategically positioned throughout Portugal, ensuring widespread geographic coverage, and offer sequencing services to both their respective host institutions and the broader scientific community.

This presentation will provide an overview of GenomePT's activities since 2008, along with a glimpse into the future prospects for expanding and enhancing this research infrastructure in the years to come."

Clinical Genomics: Current applications, hurdles, and the road to future developments

João Freixo, MD

Clinical Diretor

Centre for Predictive and Preventive Genetics of the Institute of Molecular and Cellular Biology (CGPP-IBMC)

"Over the last decade, steered by next-generation sequencing (NGS) technology, medical genetics field engage on a substantial transformation. The focus has shifted from a gene-by-gene approach to a more comprehensive genomic perspective. NGS technology enables the simultaneous analysis of multiple genomic regions, even all gene coding regions (whole-exome sequencing, WES) or the entire genome. In this presentation, the current state of rare disease genetic testing and its expanded clinical applications will be explored. In addition, the potential for opportunistic screening of actionable genetic findings, assessment of carrier status for recessive diseases, and the use of genetic data for personalized treatment strategies will be updated given the experience generated in our genetics center.

Despite the maturity level reached by this technology, some limitations and difficulties should be considered for diagnostic purposes. These challenges comprise inherent technical limitations that may affect, specifically, the analysis of some *loci* and wide impact in terms of sensitivity and specificity.

Exciting new opportunities lie ahead for clinical genomics, given the availability of long-read NGS technology, advancements in transcriptomics and epigenomics, and the extended analytical capabilities of artificial intelligence-based software, especially for variant prioritization and interpretation. Global data sharing initiatives exponentiate the possibility of connecting specific genes to phenotypic traits, especially in patients with extremely rare genetic syndromes.

Considering the vast amount of genetic data generated in diagnostic laboratories, one should not neglect its potential utility towards a national precision medicine initiative. To corroborate this assumption, we will share our results and findings generated from data mining of >13,000 WES processed in CGPP over the last seven years."

The Portuguese strategy for genomic medicine and European context

Astrid Vicente, PhD

Head of the Department of Health Promotion and noncommunicable diseases (NCDs) prevention

National Institute of Health Doutor Ricardo Jorge (INSA)

"In 2018 the European Commission promoted the signature of a Declaration of Cooperation "Towards access to at least 1 million sequenced genomes in the European Union by 2022" by member states. This was considered a game changer for European health research and clinical practice: sharing more genomic data will improve understanding and prevention of disease, allowing for personalised medicine for European citizens. The declaration has now been signed by 26 European countries.

The 1 Million Genomes (1+MG) initiative established a network of organizations across Europe with a common roadmap for developing genomics for science and health. It is also promoting the development of national genomic initiatives in many countries, including Portugal. Portugal is developing the National Strategy for Genomic Medicine (PT_MedGen), addressing challenges and defining a way forward nationally. Portugal also actively participates in the 1+MG activities and projects, including the Beyond 1 Million Genomes, the Genomic Data Infrastructure and the upcoming Genome of Europe projects."

Machine Learning as a tool to uncover the genotype-phenotype mapping: advances and current limitations

Miguel Rocha, PhD

Director of the Masters Course in Bioinformatics

Minho University

"Artificial Intelligence (AI), and more specifically Machine Learning (ML) models and algorithms, are emerging as important tools to analyze biological sequences and predict distinct phenotypes. The development of different descriptors to represent DNA/RNA and protein sequences allowed the application of classical ML approaches, while more recently the use of deep learning methods has become dominant. Here, the surge of distinct vector representations, such as protein embedding, provided exciting new methods to map sequences to biological functions.

Here, we will look at the application of these approaches to predict enzymatic functions from protein sequences, as a representative example, showing both the potential of the method and current limitations."

Forecasting species response to climate change using genomic prediction: advances and limitations

Octávio F. Paulo, PhD

Head of Computational Biology and Population Genomics Group

Centre for Ecology, Evolution and Environmental Changes (cE3c) & CHANGE – Global Change and Sustainability Institute
Animal Biology Department, Faculty of Sciences of the University of Lisbon (FCUL)

"The capacity to forecast species response to climate change has become a hot topic in applied genomics to biodiversity since Bay et al. paper in Science in 2018.

The concept of "genomic vulnerability" and "genetic or genomic offset", become central in framework of genomic prediction to climate changes.

How good are the predictive models, their limitations and how validate the results are fundamental current concerns for this research area.

Our findings with oaks tress provide a preliminary assessment of the capacity of these species to respond to future climate change and contribute towards the understanding of the potential and limitations of these models."

“One Genome… One Health” – How ATCG can help the nexus of animal, human and environmental health

Carlos G. das Neves, PhD

Chief Scientist

European Food Safety Authority (EFSA)

“The recent COVID-19 pandemic has highlighted the importance of understanding that disease emergence is rooted in human interaction with the biodiversity of microbes and their reservoir host species. As much as 75% of emerging infectious diseases (EID) “jump species” from nonhuman animals to people, with wildlife playing a key role. Genomic sequencing and modern powerful computational and statistical methods used to analyse vast amounts of data have become crucial tools in predicting and preventing disease events. Genomics have the power to provide deeper understanding of the origin, spread and evolution of infectious diseases as these cross ecosystems compartments and evolve. Also in food-borne diseases, quick access to genomic sequencing has boosted authorities capacities to trace down outbreaks to their origins, track the spread of pathogens and help adjust surveillance and control measures.

Monitoring and surveillance is key to prevention, preparedness, response, and recovery. To ensure that prevention is effective, an integrated, multi-systems monitoring and surveillance system for infectious agents and their risk factors, based on genomic tools, is essential, and must be strongly built on synergies between disciplines and the use of innovative technologies and data sharing.

One Health is defined as the collaborative efforts of multiple disciplines working locally, nationally, and globally, to attain optimal health for people, animals and our environment. In this perspective, the OH approach can only succeed if we are to focus on these Integrated systems and connect data across wildlife, companion animals, livestock, food, the environment (e.g., soil and water), and humans.”

What's new in the human (pan)genome?

Adam M. Phillippy, PhD

Head of Genome Informatics Section

National Human Genome Research Institute (NHGRI)

"In 2022, the Telomere-to-Telomere consortium finished the first truly complete sequence of a human genome. This new reference sequence adds over 200 Mbp of novel sequence including recent segmental duplications, satellite repeats, and the short arms of all five acrocentric chromosomes.

I will provide an overview of what we have learned from these new regions of the genome and highlight ongoing studies that are generating many more reference genomes from a diverse collection of humans and non-human primates. These new references are unlocking the last uncovered regions of the genome to comparative, functional, and evolutionary studies for the first time, and revealing what we have been missing for the past 20 years."

Submitted Abstracts

OC1

Microenvironment in breast cancer sentinel lymph nodes uncover two distinct immunoinflammatory profiles independent of metastasis status

Joana Martins Ribeiro, João Mendes, Inês Gante, Margarida F Dias, Luís M Nogueira, Ana Gomes, Frederico Regateiro, Francisco Caramelo, Henriqueta Coimbra Silva

OC2

Effects of TiO2 nanoparticles on the genome-wide methylation of human epithelial intestinal cells

Ana Valente, Luís Vieira, Catarina Silva, Henriqueta Louro, Maria João Silva, Célia Ventura

OC3

Differential transcription factor binding reveals the gene modules governing regulatory T cells and enables decoding mutational hotspots in human disease

Alexandre A. S. F. Raposo, Pedro Rosmaninho, Susana L. Silva, Susana Paço, Maria E. Brazão, Ana Godinho-Santos, Yumie Tokunaga, Helena Nunes-Cabaço, Ana Serra-Caetano, Afonso R. M. Almeida, Ana E. Sousa

OC4

A new understanding of the transmission landscape of *Mycobacterium bovis* boosted by single cell analyses, genomics, and ecological modeling

André C. Pereira, Daniela Pinto, José Lourenço, Ana Botelho, Mónica V. Cunha

OC5

Maternal HIV-1 genotypes and resistance mutations assessment in mother-to-child transmission with next-generation sequencing: perspectives of using this powerful tool

Diogo Ramos, Inês João, Camila Fernandes, Vítor Borges, Miguel Pinto, Joana Isidro, Luís Vieira, João Paulo Gomes, Elizabeth Pádua

OC6

Detection of the cell fusion agent virus in *Aedes aegypti* mosquitoes in Cape Verde using viral metagenomics

Líbia Zé-Zé, Ingra M. Claro, Filipe R. R. Moreira, Celivianne Sousa, Helida Pires, Adéritow Gonçalves, Davidson Monteiro, Hugo Osório, Silvania Leal, Nuno R. Faria, Rita de Sousa

OC7

Characterization of the microbiome of *Aedes albopictus* populations in Spain and São Tomé: implications for vector control

Tiago Melo, Sarah Dellacour-Estrella, Daniel Bravo-Barriga, Carla A. Sousa, Gonçalo Seixas

Submitted Abstracts

P1

Detection of copy number variants in the human genome: is long-read sequencing an alternative to genomic microarrays?

Catarina Silva, José Ferrão, Bárbara Marques, Sónia Pedro, Hildeberto Correia, António S. Rodrigues, Luís Vieira

P2

Once upon athymia: the T-cell pool of a patient with *FOXN1* mutation, 17 years post-thymic transplant

Margarida Paulo-Pedro, Beatriz Moleirinho, Diana F. Santos, André M.C. Gomes, Susana L. Silva, Alexandre A.S.F. Raposo, Afonso R.M. Almeida, Ana E. Sousa

P3

The impact of outlying samples on case-control classification models for schizophrenia

Daniel Martins; Maryam Abassi; Joel P. Arrais; Conceição Egas

P4

Performance of metagenomics NGS probe-based pathogen detection in clinical samples

Rita Ferreira, Luís Coelho, Daniel Sobral, João Dourado, Joana Isidro, Verónica Mixão, Miguel Pinto, Alexandra Nunes, Sílvia Duarte, Luís Vieira, Vítor Borges, João Paulo Gomes

P5

How a control sample established a unique pre-natal diagnosis of Temple syndrome

Ana M. Capela, Isabel Marques, Céu R. Mota, Paula Jorge, Luís Guedes-Martins, Ana M. Fortuna, Rosário Santos, Cláudia Falcão-Reis

P6

Harnessing the power of gut metagenomics to trace seafood provenance: the Atlantic chub mackerel as case study

Eduardo Feijão, Irina A. Duarte, Marcelo Pereira, Pedro Pascoal, Rita. P. Vasconcelos, Carla Gameiro, Susanne E. Tanner, Ana Rita Matos, Ricardo Dias, Andreia Figueiredo, Vanessa F. Fonseca & Bernardo Duarte

P7

The complexity of identification of pathogenic variants

Joana Fino, Dezso David

P8

Living without a thymus: impact on CD4 T cell compartment

Beatriz Moleirinho, Margarida Paulo-Pedro, Diana F. Santos, André M.C. Gomes, Susana L. Silva, Afonso R.M. Almeida, Ana E. Sousa

Submitted Abstracts

P9

Human genetic susceptibility to infection by *Coxiella burnetii* preliminary results of massive parallel sequencing

Susana David, Liliana Castro, Elsa Duarte, Miguel Machado, Ulisses Gaspar, Maria Vanessa Cueto Rojo, Joana Mendonça, José Ferrão, Miguel Machado, João Lavinha, Luís Vieira, Ana Sofia Santos

P10

Evolution of the guanylate binding protein (GBP) multi-gene family in bats (Chrioptera)

J. Ricardo Borges, Ana Pinheiro, Pedro J. Esteves

P11

Addressing uncertainties of the species-level taxonomic classification of a *Pseudomonas sp.*

Pedro Soares-Castro, Filipa Soares, Pedro M. Santos

P12

Modern Lineages of the *Mycobacterium tuberculosis* Complex are associated with increased Cross-border Transmission of Multidrug-resistant Tuberculosis

Pedro Gomes, Jody Phelan, Susana Campino, Taane Clark, Isabel Portugal, João Perdigão

P13

Unravelling miRNA-3'UTR regulatory networks in primary human CD4 T cells through short and long-read sequencing approaches

Tomás Araujo, Cláudia Noronha-Estima, Pedro Pascoal, Ricardo Dias, Margarida Gama-Carvalho

P14

Innovative medical education: genomic medicine elective course enhancing communication and ethical understanding

Célia Azevedo Soares, Lúcia Lacerda, Paula Jorge

P15

A pilot initiative to sequence the genomes of Portuguese species

João P. Marques, Paulo C. Alves, Isabel R. Amorim, Ricardo J. Lopes, Mónica Moura, Manuela Sim-Sim, Carla Sousa-Santos, Vitor C. Sousa and José Melo-Ferreira

P16

Looking for Microbial Indicator Species by Long Read Nanopore Sequencing

Ana Cruz-Silva, Gonçalo Laureano, Marcelo Pereira, Ricardo Dias, Margarida Gama-Carvalho, Fiammetta Alagna, Bernardo Duarte, Andreia Figueiredo

GenomePT SYMPOSIUM

Oral Communications

OC1

MICROENVIRONMENT IN BREAST CANCER SENTINEL LYMPH NODES UNCOVER TWO DISTINCT IMMUNOINFLAMMATORY PROFILES INDEPENDENT OF METASTASIS STATUS

Joana Martins Ribeiro[1], João Mendes[1,2], Inês Gante[2,3], Margarida F Dias[2,3], Luís M Nogueira[1,2], Ana Gomes[4], Frederico Regateiro[2,5], Francisco Caramelo[2,6,7], Henriqueta Coimbra Silva[1,2,7]

[1] University of Coimbra, Faculty of Medicine, Laboratory of Sequencing and Functional Genomics of UCGenomics, 3000-548 Coimbra, Portugal

[2] University of Coimbra, Faculty of Medicine, Coimbra Institute for Clinical and Biomedical Research (iCBR) Area of Environment, Genetics and Oncobiology (CIMAGO), 3000-548 Coimbra, Portugal

[3] Coimbra Hospital and Universitary Centre (CHUC), Gynecology Department, Coimbra, Portugal

[4] Coimbra Hospital and Universitary Centre (CHUC), Department of Pathology, Coimbra, Portugal

[5] University of Coimbra, Faculty of Medicine, Institute of Immunology, Coimbra, Portugal

[6] University of Coimbra, Faculty of Medicine, Laboratory of Biostatistics and Medical Informatics (LBIM), 3000-548 Coimbra, Portugal

[7] University of Coimbra, Center for Innovative Biomedicine and Biotechnology (CIBB), 3000-548 Coimbra, Portugal

Introduction: The sentinel lymph node (SLN) is the first tumour-draining LN and is the natural environment to explore the immune response to breast cancer (BC). Many studies describe a complex immune profile evolution of LNs not always directly correlated with the presence of metastasis. Our aim was to analyse immune-related gene expression profile of SLNs to identify potential predictive and prognosis markers.

Methodology: SLNs samples from patients with Luminal A early stage BC, 16 with invaded and 16 with non-invaded SLNs, were used for targeted RNA-seq using an immune gene panel. Unsupervised clustering was performed using K-means. Up- and down differentially expressed genes (DEGs), identified using DESeq2, were separately analysed for functional and pathway enrichment with clusterProfiler and STRING. Hub genes were identified through a PPI network using STRING and Cytoscape. Normal LN RNA-seq data were retrieved from a public database. CIBERSORTx was used to estimate the abundances of immune populations.

Results: Unsupervised analysis revealed two clusters regardless the presence of metastasis. Cluster 1 (C1) upregulated DEGs were associated with "positive regulation of alpha-beta T cell activation" and "B cell proliferation", while cluster 2 (C2) upregulated DEGs were associated with "positive regulation of macrophage migration". Furthermore, five hub genes were identified: *CD80*, *CD40* and *TNF* were upregulated in C1, while *FCGR3A* and *CD163* were upregulated in C2. Data integration showed that the relative expression of *CD80*, *TNF*, and *CD163* is higher in BC SLNs compared to NLNs, while *CD40* shows a relatively lower expression in patients' SLNs. Furthermore, the LNN immune gene profile differs from that of cancer-free SLNs. Finally, C1 exhibited increased total B cells, follicular helper T cells and activated NK and mast cells compared to C2. The activated-to-resting DC ratio was elevated in C2.

OC1

MICROENVIRONMENT IN BREAST CANCER SENTINEL LYMPH NODES UNCOVER TWO DISTINCT IMMUNOINFLAMMATORY PROFILES INDEPENDENT OF METASTASIS STATUS

Joana Martins Ribeiro[1], João Mendes[1,2], Inês Gante[2,3], Margarida F Dias[2,3], Luís M Nogueira[1,2], Ana Gomes[4], Frederico Regateiro[2,5], Francisco Caramelo[2,6,7], Henriqueta Coimbra Silva[1,2,7]

Conclusions: Results suggest: 1) that immune profile of SLNs is not directly related to the presence of metastasis, and may be explained by interindividual variability of immune response; (2); the existence of two different profiles, one with evidence of adaptive anti-tumoral immune response (C1) and another (C2) with a more undifferentiated and pro-metastatic immune-environment (*CD163*); (3) both groups express immune related targets (CD80, CD40 and CD163).

Funding: GenomePT POCI-01-0145-FEDER-022184, CENTRO2020 [CENTRO-08-5864-FSE-000039 (PEP IN1194)07-2021].

OC2

EFFECTS OF TiO_2 NANOPARTICLES ON THE GENOME-WIDE METHYLATION OF HUMAN EPITHELIAL INTESTINAL CELLS

Ana Valente[1,2], Luís Vieira[1,3], Catarina Silva[1,3], Henriqueta Louro[1,3], Maria João Silva[1,3], Célia Ventura[1,3]

[1] Department of Human Genetics, National Institute of Health Doutor Ricardo Jorge, I.P. (INSA), Av. Padre Cruz, 1649-016 Lisbon, Portugal

[2] Department of Animal Biology, Faculty of Sciences, University of Lisbon, Campo Grande, 1749-016 Lisbon, Portugal

[3] Centre for Toxicogenomics and Human Health (ToxOmics), NOVA Medical School, NOVA University of Lisbon, 1169-056 Lisbon, Portugal

Introduction: Titanium dioxide nanoparticles (TiO_2NP) have multiple applications in industry (e.g., engineering, cosmetics, food additives), and biomedicine (e.g., targeted drug delivery and biosensing). Food-grade TiO_2 (E171) is applied as a food additive to whiten and improve the opacity of food products, while also having the ability to enhance its flavour. In 2021, EU member states banned E171 from all food products, since there is doubt about its genotoxicity. Nevertheless, the ingestion of TiO_2NPs may still occur through other sources, such as contaminated food or water, consumer products (e.g., toothpaste and lipstick) or pharmaceutics. To date, there is some *in vitro* evidence that TiO_2NP may induce changes in DNA methylation. However, very few studies were performed, and none used genome-wide approaches to identify possible differentially methylated genes induced by TiO_2NP exposure, and its impact on molecular pathways.

Methodology: Caco-2 epithelial intestinal cells were exposed to 14 µg/ml of anatase, rutile or brookite phase TiO_2NP for 24h. Genomic DNA was extracted from exposed and non-exposed cells. DNA libraries were generated using the Premium Reduced Representative Bisulfite Sequencing (RRBS) kit (Diagenode) and sequenced on the NextSeq 550 system (Illumina) using 100 bp paired reads. The Galaxy platform was used for read treatment and mapping, methylation calling and assessing of differentially methylated regions between exposed and non-exposed cells. Pathway analysis was performed using Reactome, and gene ontology analysis with the ClueGO plugin in Cytoscape.

Results: Significant differential methylation ($p \leq 0.05$) of 92 genes (21 hyper- and 71 hypo-methylated), 70 genes (12 hyper- and 58 hypo-methylated) and 88 genes (21 hyper- and 67 hypo-methylated) was observed for the anatase, rutile and brookite phase TiO_2NP, respectively. Functional pathway analysis of these methylation changes identified several relevant cellular pathways that may be altered by exposure, such as G alpha signalling events, being some associated to colon cancer.

Conclusions: All types of TiO_2NP induce changes in genome methylation of intestinal cells, which may affect cell proliferation, differentiation and survival. Moreover, although some dysfunctional pathways are shared between the three TiO_2NP, many are type-specific, suggesting different molecular mechanisms of action for each TiO_2NP.

Funding: FCT (PTDC/SAU-PUB/29481/2017), FCT UIDB/00009/2020 UIDP/00009/2020.

OC3

DIFFERENTIAL TRANSCRIPTION FACTOR BINDING REVEALS THE GENE MODULES GOVERNING REGULATORY T CELLS AND ENABLES DECODING MUTATIONAL HOTSPOTS IN HUMAN DISEASE

Alexandre A. S. F. Raposo[1], Pedro Rosmaninho[1], Susana L. Silva[1], Susana Paço[1], Maria E. Brazão[1], Ana Godinho-Santos[1], Yumie Tokunaga[1], Helena Nunes-Cabaço[1], Ana Serra-Caetano[1], Afonso R. M. Almeida[1], Ana E. Sousa[1]

[1] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Portugal

Introduction: Computational strategies to extract meaningful biological information from multiomics data are in great demand for effective clinical use. This is most relevant in immune-mediated disorders, where the combined impact of multiple variants is difficult to determine. Regulatory T cells (Tregs), particularly those lineage-committed in the thymus, are essential for immune homeostasis and self-tolerance, controlling inflammatory and autoimmune processes in many diseases with a multigenic basis.

Methodology: Here, we quantify the Transcription Factor (TF) differential occupancy landscape to uncover the Gene Regulatory Modules (GRM) governing human thymic Tregs and use these to prioritise variants in complex diseases. Combined RNA-seq and ATAC-seq generated a matrix of differential TF binding to genes differentially expressed in human Tregs, in contrast to their counterpart conventional CD4 single-positive thymocytes. We then mapped the GRM to the genomic variant landscape of patients with common variable immunodeficiency (CVID), here used as a model of polygenic-based disease with severe inflammatory and autoimmune manifestations.

Results: The gene *loci* of both established and novel genetic interactions uncovered by the GRM were significantly enriched in rare variants carried by CVID patients. GRM controlling the human thymic Treg signature can, therefore, be a valuable resource for variant classification in immune disorders and to reveal new therapeutic targets.

Conclusions: We propose differential binding as a reliable measurement of TF local activity, overcoming limitations in traditional approaches to define regulatory networks and variant-association studies. Overall, we provide a tool to decipher mutational hotspots in individual genomes, including those with non-coding variants, and unimpeded by size of cohort.

Acknowledgements: We thank Miguel Abecasis MD, Rui Anjos MD, and Hospital de Santa Cruz-Centro Hospitalar Lisboa Ocidental, and all parents and children for the access to thymic tissue; the iMM flow cytometry facility; Miguel Ângelo-Dias, Zoe Junginger, Adriana Motta-Raymundo, and Diana Santos for technical assistance.

Funding: PAC-PRECISE-LISBOA-01-0145-FEDER-016394 and LISBOA-01-0145-FEDER-007391, co-funded by FEDER through POR Lisboa 2020 - Programa Operacional Regional de Lisboa PORTUGAL 2020 and Fundação para a Ciência e a Tecnologia (FCT); PTDC/MED-IMU/0938/2020, funded by the Portuguese government through FCT; GenomePT PINFRA/22184/2016/POCI-01-0145-FEDER-022184; Marie Sklowdowska-Curie-COFUND Action MSCA-ITN-2015-675395 - H2020; and CEECIND/01474/2017 (FCT).

OC4

A NEW UNDERSTANDING OF THE TRANSMISSION LANDSCAPE OF *MYCOBACTERIUM BOVIS* BOOSTED BY SINGLE CELL ANALYSES, GENOMICS, AND ECOLOGICAL MODELING

André C. Pereira[1,2], Daniela Pinto[1,2], José Lourenço[2], Ana Botelho[3], Mónica V. Cunha[1,2]

[1] Centre for Ecology, Evolution and Environmental Changes (cE3c) & CHANGE - Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, Portugal.

[2] Biosystems & Integrative Sciences Institute (BioISI), Faculdade de Ciências, Universidade de Lisboa, Portugal. [3]Instituto Nacional de Investigação Agrária e Veterinária IP, Oeiras, Portugal.

Introduction: Animal tuberculosis (TB) is a communicable infectious disease, primarily caused by *Mycobacterium bovis*. Spreading across multi-host systems is mainly attributed to animal contacts, while exposure to contaminated substrates is assumed as a potential source of infection. This work aimed to improve global understanding of *M. bovis* transmission landscapes, via the development and integration of novel tools.

Methodology: To investigate the environmental contamination hypothesis, we developed a single-cell workflow based on flow cytometry, enabling detection, quantification and sorting of *M. bovis* cells in environmental matrices. Subsequently, we used this technology to analyse field samples collected within the official animal TB risk area. Coupling this workflow with a whole-genome enrichment strategy, we generated the first-ever whole genome sequences (WGS) of *M. bovis* from the environment, enabling phylogenomic comparisons at the environment-animal interface. In parallel, we generated WGS data at the livestock-wildlife interface and established eco-phylodynamic frameworks to reconstruct transmission processes across ecological gradients.

Results: Most environmental samples contain metabolically active and dormant *M. bovis* cells. The burden of viable *M. bovis* cells in sludge is compatible with the infectious dose required to infect new hosts. Environmental and animal *M. bovis* genomes are highly intertwined, with genomic data supporting several instances of environmental substrate contamination by infected animals. Evidence for the co-circulation of *M. bovis* Eu1, Eu2, and Eu3 clonal complexes in both animal and environmental samples is provided and the emergence and spread processes of the main lineages across animal populations are enlightened. Through ecological modeling, we additionally show that most host transitions occur toward climate, land use, and host density gradients, while phylogeographic reconstruction highlights ecological corridors of unrecognised importance in transmission.

Conclusions: Our work clarifies the different epidemiological roles exerted by different host species on TB maintenance and spread, indicates the environment as a source of animal infection, and highlights transmission routes in an era of global changes, supporting data-driven intervention.

Funding: Fundação para a Ciência e a Tecnologia, IP [2022.06014.PTDC, PTDC/CVT-CVT/29783/2017], PORLisboa [LISBOA-01-0145-FEDER-029783], COMPETE [POCI-01-0145-FEDER-029783].

OC5

MATERNAL HIV-1 GENOTYPES AND RESISTANCE MUTATIONS ASSESSMENT IN MOTHER-TO-CHILD TRANSMISSION WITH NEXT-GENERATION SEQUENCING: PERSPECTIVES OF USING THIS POWERFUL TOOL

Diogo Ramos[1], Inês João[2], Camila Fernandes[3], Vítor Borges[4], Miguel Pinto[4], Joana Isidro[4], Luís Vieira[4], João Paulo Gomes[4], Elizabeth Pádua[1]

[1] Reference Laboratory of HIV and Hepatitis B and C, Department of Infectious Diseases (DDI), National Institute of Health (INSA), Av. Padre Cruz, P-1649-016 Lisbon, Portugal

[2] Reference Unit of Sexually Transmitted Infections, Department of Infectious Diseases (DDI), National Institute of Health (INSA), Av. Padre Cruz, P-1649-016 Lisbon, Portugal

[3] Technology and Innovation Unit, Department of Human Genetics (DGH), National Institute of Health (INSA), Av. Padre Cruz, P-1649-016 Lisbon, Portugal

[4] Genomics and bioinformatics Unit, Department of Infectious Diseases (DDI), National Institute of Health (INSA), Av. Padre Cruz, P-1649-016 Lisbon, Portugal

Introduction: There is a broad consensus on antiretroviral drugs (ARD) use to prevent HIV-1 vertical transmission (VT). However, resistant variants can emerge during pregnancy, with risk of transmission to infants. Resistance may also emerge in infants after exposure to ARD prophylaxis. Through genetic markers identification, next-generation sequencing (NGS) may support a close surveillance of drug resistance, mandatory for a better management of the infection.

We aimed to assess if maternal HIV-1 resistance was associated with VT, and to describe the genetic diversity in mothers who transmitted (TM) or not transmitted (NTM) the virus to infants (IN).

Methodology: A cohort of 26 TM+IN pairs and 30 NTM (controls) was selected between 2001 and 2015. Although most mothers did not adhere to ARD regimen, infants predominantly fulfilled a prophylaxis regimen. Phylogenetic determination of HIV-1 genotypes and assessment of drug resistance by NGS were performed. An amplicon-based strategy was carried out, with proviral DNA amplification of six amplicons covering the whole HIV-1 genome. Samples were paired-end sequenced after library preparation with the Nextera XT protocol (Illumina). Results were submitted to the HIVdb-NGS tool by Stanford University and a primary analysis was based on the amplicon containing the PR and RT.

Results: We present preliminary results of 15 pairs TM+IN and 8 NTM. Several HIV-1 genotypes were detected: B, G, CRF02_AG, F1 in TM+IN and B, G, C, J in NTM. Amino acid substitutions of concern were detected in 8/15(53.3%) TM+IN pairs: A62V, D67N, T69N, K103N and T215A in RT, and L10V/I and G48R in PR. The same mutational profile was observed in 2/15(13.3%) TM+IN pairs; in 3/15(20.0%) the TM mutations were not transmitted to the IN; and for the remaining 3/15(20.0%) pairs T215A and D67N in RT and G48R in PR were only observed in IN. Regarding NTM, K103E and M184V in the RT region and L10I, K20R/I, G48R, A71V and G73R in the PR region were detected in 5/8(62.5%).

OC5

MATERNAL HIV-1 GENOTYPES AND RESISTANCE MUTATIONS ASSESSMENT IN MOTHER-TO-CHILD TRANSMISSION WITH NEXT-GENERATION SEQUENCING: PERSPECTIVES OF USING THIS POWERFUL TOOL

Diogo Ramos[1], Inês João[2], Camila Fernandes[3], Vítor Borges[4], Miguel Pinto[4], Joana Isidro[4], Luís Vieira[4], João Paulo Gomes[4], Elizabeth Pádua[1]

Conclusions: These preliminary results reveal a high genetic diversity of HIV-1. Although infected mothers were not using ARD, mutations of concern were identified, altogether with mutations only found in IN, probably selected due to ARD prophylaxis. Emergence of HIV-1 resistant variants directly affects ARD efficacy, emphasizing a mandatory monitoring of the infection with the beneficial brought by NGS.

Funding: Work supported by DDI INSA and DR by EU4H-2022-DGA-MS-IBA-01-02/Proj.101113460

OC6

DETECTION OF THE CELL FUSION AGENT VIRUS IN *AEDES AEGYPTI* MOSQUITOES IN CAPE VERDE USING VIRAL METAGENOMICS

Líbia Zé-Zé[1,2], Ingra M. Claro[3], Filipe R. R. Moreira[4], Celivianne Sousa[5], Helida Pires[5], Adéritow Gonçalves[5], Davidson Monteiro[5], Hugo Osório[1,6], Silvania Leal[5], Nuno R. Faria[3], Rita de Sousa[1]

[1] Centre for Vectors and Infectious Diseases Research (CEVDI), National Institute of Health Doutor Ricardo Jorge (INSA), Department of Infectious Diseases, Águas de Moura, Portugal

[2] Centro de Estudos de Ciência Animal (CECA), Instituto de Ciências, Tecnologias e Agroambiente (ICETA), Universidade do Porto, Porto, Portugal

[3] Imperial College, School of Public Health, Department of Infectious Disease Epidemiology, London, United Kingdom

[4] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[5] Medical Entomology Laboratory, National Institute of Public Health, Largo do Desastre da Assistência, Chã de Areia, CP Nº 719, Cidade da Praia, Cabo Verde

[6] Instituto de Saúde Ambiental (ISAMB), Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

Introduction: Mosquito-borne viruses are recognised as important pathogens and emerging global threats. Improving surveillance of mosquito vectors, their pathogens and potential host interactions is critical in mitigating the impact of vector-borne diseases on both human and animal populations. The One Health vector-borne project, ONESVEC, implemented in 2022, aims to improve the knowledge, research and surveillance of tick- and mosquito-borne pathogens in the Cape Verde Islands.

Methodology: Mosquitoes were collected from several islands in Cape Verde from April to December 2022. A total of 224 samples were morphologically identified and screened for viruses using pan-flavivirus RT-PCR. Individual positive samples were analysed using an untargeted metagenomic sequencing approach based on SMART technology, SMART-9N. Random priming was used for cDNA synthesis followed by PCR amplification. Libraries were generated by barcoding and equimolar pooling of samples, followed by loading onto FLO-MIN106 flow cells on the MinION instrument (ONT, UK). Sequencing was conducted following the standard 48-hour run script. Raw FAST5 files were basecalled, demultiplexed and trimmed using Guppy software. Taxonomic classification was performed with Kraken v2 viral database. The classified barcoded FASTQ files were then aligned and mapped to the reference genomes (NC_001564.2) using minimap2 and converted to a sorted BAM file using SAMtools. Genome statistics were obtained from SAMtools and the Tablet viewer. To recover consensus sequences, we called variants detected with Medaka for regions of the genome covered with at least 20 reads.

Results: In two female *Aedes aegypti* mosquitoes collected in Santiago Island we were able to characterize a viral RNA genome sequence of 10,682 nt identified as Cell Fusion Agent Virus (CFAV). The CFAV is an insect-specific flavivirus, first identified in an *Ae. aegypti* cell line, that shows the usual genome structure organisation of Flaviviruses.

OC6

DETECTION OF THE CELL FUSION AGENT VIRUS IN *AEDES AEGYPTI* MOSQUITOES IN CAPE VERDE USING VIRAL METAGENOMICS

Líbia Zé-Zé[1,2], Ingra M. Claro[3], Filipe R. R. Moreira[4], Celivianne Sousa[5], Helida Pires[5], Adéritow Gonçalves[5], Davidson Monteiro[5], Hugo Osório[1,6], Silvania Leal[5], Nuno R. Faria[3], Rita de Sousa[1]

Conclusions: We report for the first time the detection of CFAV infecting *Ae. aegypti* mosquito populations in Cape Verde. The presence of CFAV in natural mosquito populations may be important, as CFAV has been shown to reduce the spread of arboviruses (namely dengue virus serotype 1 and Zika viruses) *in vivo*. Genomic surveillance is essential to assess the prevalence and potential impact of arboviruses with a significant human disease.

Acknowledgments: We thank the delegates and anti-vector campaigners from the health centres in Cape Verde.

Funding: FCT/AGA-KHAN/541751551/2019: ONESVEC, "One Health action for enhancing surveillance on vector-borne diseases in Cape Verde" project.

OC7

CHARACTERIZATION OF THE MICROBIOME OF *Aedes albopictus* POPULATIONS IN SPAIN AND SÃO TOMÉ: IMPLICATIONS FOR VECTOR CONTROL

Tiago Melo[1], Sarah Dellacour-Estrella[2,3], Daniel Bravo-Barriga[4], Carla A. Sousa[1], Gonçalo Seixas[1]

[1] Global Health and Tropical Medicine, GHTM, Associate Laboratory in Translation and Innovation Towards Global Health, LA-REAL, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, IHMT /UNL, Portugal

[2] Animal Health Department, The AgriFood Institute of Aragon (IA2), School of Veterinary Medicine, University of Zaragoza, 50013 Zaragoza, Spain

[3] Departamento de Investigación y Desarrollo (I+D). Quimera. B.S. Calle Olivo, 14, 50016 La Puebla de Alfindén, Spain

[4] Department of Animal Health, Animal Health and Zoonosis Research Group (GISAZ), UIC Zoonosis and Emerging Diseases (ENZOEM), University of Córdoba, Córdoba, Spain

Introduction: *Aedes albopictus* is a well-known mosquito vector and nuisance agent for arboviruses. Over recent decades, this species has successfully spread to multiple continents, increasing the risk of pathogen transmission. Innovative vector control tools are urgently needed with recent focus on manipulating mosquito's microbiome.

Methodology: Mosquito eggs were collected in Spain and São Tomé in late 2021 and early 2022. Four mosquito colonies were established and maintained at VIASEF (In Vivo Arthropod Security Facility, IHMT). DNA extraction and Illumina 16S rRNA gene sequencing were performed on 19 female samples to study the natural microbiota. *Wolbachia* infection genotyping was done in 180 samples (males and females). Statistical analysis compared mixed and single infections by sex and population via chi-square tests.

Results: Illumina 16S rRNA sequencing yielded a total of 1,404,803 reads, which were categorized into 871 Operational Taxonomic Units (OTUs). Among the 19 samples, the proportion of reads attributed to the genus *Wolbachia* ranged from 92.4% to 98.8%, highlighting the dominance of *Wolbachia* in the *Ae. albopictus* microbiome composition. Other genera, including *Pelomonas*, *Asaia*, *Pseudomonas*, *Nevskia*, and *Sphingomonas*, were also detected, albeit with a lower number of assigned reads (0.1% to 1.8%). This study represents the first identification of the genera *Pelomonas* and *Nevskia* in *Aedes* mosquitoes. *Wolbachia* strains A and B were detected among the genotyped samples. Infection patterns varied, with mosquitoes exhibiting either double infection (strains A and B) or single infection (A or B). Dual infections were predominant, accounting for 71.11% of *Wolbachia*-infected mosquitoes, while single infections constituted 21.66%. Notably, female mosquitoes displayed a higher prevalence of double *Wolbachia* infections compared to males.

Conclusions: This study provides the first description of the *Ae. albopictus* microbiome from Spain and São Tomé populations. These findings may serve as valuable insights for the design of innovative vector control strategies, particularly tailored for *Ae. albopictus* populations. However, additional research is required to gain a deeper understanding of the impact of *Wolbachia* and other endosymbionts on this species.

Funding: This work was supported by the Fundação para a Ciência e Tecnologia funding to GHTM-UID/04413/2020 and LA-REAL-LA/P/0117/2020.

GenomePT
SYMPOSIUM

Posters

P1

DETECTION OF COPY NUMBER VARIANTS IN THE HUMAN GENOME: IS LONG-READ SEQUENCING AN ALTERNATIVE TO GENOMIC MICROARRAYS?

Catarina Silva[1,3,4], José Ferrão[1], Bárbara Marques[2], Sónia Pedro[2], Hildeberto Correia[2], António S. Rodrigues[3,4], Luís Vieira[1,3]

[1] Unidade de Tecnologia e Inovação, Departamento de Genética Humana, INSA, Lisboa.

[2] Unidade de Citogenética, Departamento de Genética Humana, INSA, Lisboa.

[3] Centro de Toxicogenómica e Saúde Humana (ToxOmics), NOVA Medical School|Faculdade de Ciências Médicas, Universidade Nova de Lisboa.

[4] NOVA Medical School | Faculdade de Ciências Médicas, Universidade Nova de Lisboa.

Introduction: Copy number variations (CNVs) represent ~13% of the human genome and can harbour important genes and regulatory elements. High-resolution whole genome microarray (MA) analysis is the gold standard tool for detection of CNVs associated with genetic disorders. While short-read sequencing (SRS) can address SV detection, the use of long-read sequencing as proven to overcome SRS mapping inaccuracy in highly repetitive DNA regions and improve genome contiguity. We applied whole genome nanopore sequencing (NS) to call CNVs and compared the results with those obtained by microarray.

Methodology: Genomic DNA from 2 cell lines (EOL-1 and 697) were processed using the CytoSan HD Array (Affymetrix) and ChAS software (ThermoFisher). A minimum CNV calling size threshold of 35 Kb was used. DNA was also sequenced on the MinION device (Oxford Nanopore Technologies) following a rapid library preparation method. Sequencing data were basecalled using Guppy, mapped with LRA, and SVs called using both CuteSV and Sniffles2. Sanger sequencing was performed to demonstrate breakpoint positions for 3 CNVs. R packages were used to perform comparisons between MA and NS data.

Results: A total of 49 CNVs were confirmed after curated MA analysis in both cell lines, ranging in size from 35 Kb to 79 Mb. From those, 43 CNVs (87.7%) were called in nanopore data by either one (4 CNVs) or both (39 CNVs) callers with a mean whole genome coverage of ~12X. Six of 43 CNVs were called as inversions instead. In 3 CNVs the size of the variant was found to be smaller (ranging from ~5 to 22 Kb) than the threshold of MA analysis. The correlation between CNV sizes obtained with MA and NS was of 0.71 with Sniffles2 and 0.74 with CuteSV, whereas the correlation between callers was of 0.99. The breakpoint precision obtained for NS was much higher (ranging for CuteSV from 2 to 42 bp; and for Sniffles2 from 0 to 87 bp) than the one obtained for MA (ranging from 774 to 7618 bp).

Conclusions: NS technology proved to be technically effective in the detection of CNVs of different types and sizes and thus posing itself as an alternative to MA in the detection of pathogenic SVs associated with genetic diseases. However, NS data analysis requires fine-tuning of the analysis conditions as well as the use of different methods, for greater reliability of results in a clinical context.

Acknowledgments: This work was supported by GenomePT (POCI-01-0145-FEDER-022184) and ToxOmics (UID/BIM/00009/2019).

P2

ONCE UPON ATHYMIA: THE T-CELL POOL OF A PATIENT WITH *FOXN1* MUTATION, 17 YEARS POST-THYMIC TRANSPLANT

Margarida Paulo-Pedro[1], Beatriz Moleirinho[1], Diana F. Santos[1], André M.C. Gomes[1], Susana L. Silva[1,2], Alexandre A.S.F. Raposo[1], Afonso R.M. Almeida[1], Ana E. Sousa[1]

[1] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Portugal.

[2] Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, Lisboa, Portugal.

Introduction: *FOXN1* deficiency, characterized by *alopecia totalis*, nail dystrophy, and severe T-cell immunodeficiency, leads to nude/severe combined immunodeficiency (SCID) phenotype. *FOXN1* gene is expressed by the thymic epithelium and mutations in this gene lead to developmental defects of thymic epithelial cells resulting in lack of T-cell development. Our lab has been following since birth an 18-year-old patient with *FOXN1* deficiency, that despite the lack of thymus presents circulating T-cells with a significant population with an aberrant phenotype (double negative αβT-cells (DN)). She was submitted to thymic transplantation at 14 months of age, with effective immunological reconstitution being achieved, despite HLA-mismatch. Thymic transplantation was preceded by immunosuppressive treatment, clearing all circulating T-cells, except the DNαβ T-cell population. Allogenic thymic graft functionality increased progressively reaching levels comparable to those found in healthy children, followed by a sharp decline at 4-year post-transplant. In this study we aim to characterize for the first time, to our knowledge, at a single-cell level the circulating T-cell compartment upon thymus transplantation. Specifically, we aim to investigate: 1) the mechanisms underlying the maintenance of naïve T-cells given the evidence of failure of the allogenic thymic graft, and 2) the aberrant DNαβ T-cells known to be present before thymic transplantation.

Methodology: Circulating T-cells were evaluated by high-dimensional spectral flow cytometry. In parallel, naïve CD4 and CD8 T-cells as well as DNαβ T-cells were enriched through FACS sorting and complemented with other CD3 T-cells. This enriched population was used to perform scRNAseq combined with TCR and cell surface protein sequencing (CITE-Seq). For the CITE-seq analysis, we selected a naïve marker (CD45RA) and CD31.

Results: From our flow cytometry analysis, we observed two major findings: 1) the persistence of the DNαβ T-cell population, and 2) a major contraction of the naïve T-cell compartment. Furthermore, the naïve CD4 T-cells display a unique aberrant phenotype with high levels of TCF1 expression. The scRNAseq dataset produced is currently being analyzed.

Conclusions: With our analysis we intend to uncover homeostatic mechanisms underlying the maintenance of these T-cell populations in this clinical setting. Ultimately, we expect to provide new insights on human CD4 T-cell biology and identify pathways for immune reconstitution.

P3

THE IMPACT OF OUTLYING SAMPLES ON CASE-CONTROL CLASSIFICATION MODELS FOR SCHIZOPHRENIA

Daniel Martins[1,2], Maryam Abassi[1,3], Joel P. Arrais[1], Conceição Egas[2,4]

[1] University of Coimbra, Centre for Informatics and Systems of the University of Coimbra, Department of Informatics Engineering, Coimbra, Portugal

[2] University of Coimbra, Centre for Innovative Biomedicine and Biotechnology., Coimbra, Portugal

[3] Polytechnic Institute of Coimbra, Applied Research Institute, Coimbra, Portugal

[4] Biocant – Transfer Technology Association, Cantanhede, Portugal

Introduction: Genetic factors play a significant role in the development of Schizophrenia (SCZ). However, its exact causes remain unclear, leading to a reliance on broad diagnostic frameworks and thus, hindering the assembly of reliable case-control groups. As an illustrative example of this issue, the Swedish Hospital Discharge Register, from where one of the largest SCZ cohorts was derived, has reported relatively high misdiagnosis rates for SCZ, ranging from 6% to 19%.

Methodology: To address this challenge, we used the referred large-scale case-control Whole-Exome Sequencing dataset from the Swedish population, containing more than 1.8 million variants and refined this dataset to the 18,970 variants that showed significant associations with the SCZ phenotype. Subsequently, we developed a Machine Learning model based on gene annotations and trained it on the entire dataset to improve our ability to distinguish between cases and controls. Outlying samples, more likely to be misclassified, were excluded from subsequent analysis. We conducted a second round of training to evaluate the performance of the model on a classification task on refined datasets.

Results: After excluding samples in proportion to the misdiagnosis rate associated with a stricter SCZ definition (19%), our classification model achieved an AUC value of approximately 0.81 for the test set. This level of differentiation between cases and controls aligns with existing estimates of Schizophrenia heritability in the literature. Additional analyses indicate a significant contribution of genes related to glucose metabolism pathways.

Conclusions: Our findings suggest that reported misdiagnosis rates for Schizophrenia may affect the composition of case-control cohorts, potentially undermining the reliability of genetic studies relying on such datasets. In comparison to the original data, the genetic profiles used to distinguish cases from controls in the refined dataset exhibited greater consistency across various iterations of the training process. This supports our approach of defining more informative subsets from large-scale cohorts and offers opportunities to apply adapted machine learning techniques for complex disease research using such data.

Acknowledgements: dbGaP phs000473.v2.p2.

Funding: FCT SFRH/BD/146094/2019; UIDB/00326/2020; UIDP/00326/2020; UIDB/04539/2020; UIDP/04539/2020; LA/P/0058/2020.

P4

PERFORMANCE OF METAGENOMICS NGS PROBE-BASED PATHOGEN DETECTION IN CLINICAL SAMPLES

Rita Ferreira[1], Luís Coelho[1], Daniel Sobral[1], João Dourado[1], Joana Isidro[1], Verónica Mixão[1], Miguel Pinto[1], Alexandra Nunes[1], Sílvia Duarte[2], Luís Vieira[2], Vítor Borges[1], João Paulo Gomes[1]

[1] Genomics and Bioinformatics Unit, Department of Infectious Diseases, National Institute of Health Doutor Ricardo Jorge (INSA), Lisbon, Portugal

[2] Technology and Innovation Unit, Department of Human Genetics, National Institute of Health Doutor Ricardo Jorge (INSA), Lisbon, Portugal

Introduction: In clinical settings, metagenomics Next-Generation Sequencing (mNGS) may detect accurately and rapidly unsuspected, unidentified and uncultivatable organisms, resulting in higher recovery outcomes due to timely and more specific treatments. However, mNGS presents several challenges, such as high costs per sample, a small microbial/human DNA ratio, and demanding laboratorial and computational infrastructures.

Methodology: We tested two recently developed Illumina panels for probe-based pathogen enrichment and detection, the Respiratory and the Urinary Pathogen ID/AMR panels (RPIP and UPIP), which together target 383 pathogenic agents (virus, bacteria, fungi and parasites). We selected 84 clinical samples of different nature (e.g., CSF, plasma, serum, urine, swabs, biopsies, etc.), for which at least one pathogen, mostly virus, had been previously identified by PCR in INSA's national reference laboratories. Data analysis was performed using both Illumina and in-house bioinformatics pipelines, with panels' performance being assessed in a combined fashion to mimic a real scenario where the panels would be used simultaneously to cover all 383 pathogens.

Results: Herein, the Illumina's pipelines detected 66.3% of the pathogens (highest value was 74.6%, for virus). Since we had previously developed an in-house bioinformatics pipeline for viral metagenomics detection (INSaFLU-TELEVIR, https://insaflu.insa.pt/), we assessed its usefulness to improve viral detection rate. This strategy increased Illumina accuracy for virus detection from 74.6% to 84.7%. Therefore, by improving viral identification we were able to increase the overall success rate of pathogen identification from 66.3% to about 72%. As such, we are currently developing a TELEVIR-like pipeline for bacteria to assess if a similar increment in detection is achieved.

Conclusion: Although these results rely on preliminary data, they provide strong evidence that these Illumina panels, coupled with complementary downstream pipelines, may constitute a powerful tool to help difficult diagnosis and support clinical decision, in which timely and specific treatment may be decisive for patients' clinical recovery.

P5

HOW A CONTROL SAMPLE ESTABLISHED A UNIQUE PRE-NATAL DIAGNOSIS OF TEMPLE SYNDROME

Ana M. Capela[1], Isabel Marques[2,3,4], Céu R. Mota[5], Paula Jorge[2,3,4], Luís Guedes-Martins[6,7], Ana M. Fortuna[1,3,4], Rosário Santos[2,3,4], Cláudia Falcão-Reis[1,8,9]

[1] Clinical Genetics Unit, Centro de Genética Médica Jacinto Magalhães (CGM), Centro Hospitalar Universitário de Santo António (CHUdSA), Porto, Portugal.

[2] Molecular Genetics Unit, Centro de Genética Médica Jacinto Magalhães (CGM), Centro Hospitalar Universitário de Santo António (CHUdSA), Porto, Portugal.

[3] UMIB – Unit for Multidisciplinary Research in Biomedicine, ICBAS – School of Medicine and Biomedical Sciences, University of Porto, Porto, Portugal.

[4] ITR – Laboratory for Integrative and Translational Research in Population Health, Porto, Portugal.

[5] Neonatology and Intensive Care Department, Neonatology Unit, Centro Materno-Infantil do Norte, Centro Hospitalar Universitário de Santo António (CHUdSA), Porto, Portugal.

[6] Fetal Medicine Centre (Medicina Fetal Porto) - Centro Materno-Infantil do Norte, Centro Hospitalar Universitário de Santo António (CHUdSA), Porto, Portugal.

[7] ICBAS – School of Medicine and Biomedical Sciences, University of Porto, Porto, Portugal.

[8] Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal;

[9] ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal.

Introduction: Temple Syndrome (TS, MIM #616222) is a genomic imprinting condition caused by disturbances of the imprinted region 14q32. TS is an ultra-rare condition with unknown prevalence and is likely underdiagnosed, since it is often not detected by first tier testing WES or array-CGH. We report a fortunate incidental prenatal diagnosis of TS and hope to highlight the role of methylation studies in the upcoming WGS era. We also aim to acknowledge that control samples, though very rarely, can harbor a diagnosis.

Methodology: A 29-year-old primigravida was evaluated at 21w+6d due to fetal growth restriction and ultrasound abnormalities in the 2nd-trimester evaluation. She underwent amniocentesis (at 22w+5d) for rapid aneuploidy test and karyotype. MS-MLPA for UPD7-UPD14 and array-CGH were also performed.

Results: Rapid aneuploidy test was negative and karyotyping 46,XX. In the context of a routine MS-MLPA for UPD7-UPD14, the molecular genetics laboratory used this fetal DNA sample as a control, detecting hypomethylation of the 14q32.3 region. After SNP-array in fetal and maternal samples revealed no uniparental disomy or deletions, epimutation was established as the most likely disease-causing mechanism for TS in our proband.

P5

HOW A CONTROL SAMPLE ESTABLISHED A UNIQUE PRE-NATAL DIAGNOSIS OF TEMPLE SYNDROME

Ana M. Capela[1], Isabel Marques[2,3,4], Céu R. Mota[5], Paula Jorge[2,3,4], Luís Guedes-Martins[6,7], Ana M. Fortuna[1,3,4], Rosário Santos[2,3,4], Cláudia Falcão-Reis[1,8,9]

Conclusion: A primigravida was referred to a fetal medicine consultation after findings of unilateral clubfoot and growth below the 5th centile on routine morphological ultrasound. Standard etiological investigations were negative and imaging follow-up resumed. A remaining fetal sample used as a control for a MS-MLPA detected hypomethylation of the 14q32.3 region. The relevance of this finding warranted un-pseudonymization and referral to Medical Genetics for genetic counseling. Although unspecific, the fetal phenotype was compatible with TS diagnosis and additional studies ensued. This diagnosis was well accepted by the couple, who feared a worse prognosis. The baby was delivered by C-section at 33w+6d due to fetal growth restriction. She progressed favorably and is currently 34 months old, with proportionate short stature and adequate psychomotor development after clubfoot repair. The early diagnosis of TS was invaluable in informing the parents and the medical team on prognosis, and allowed for specific preparations during pregnancy and labor, as well as early post-natal referrals for multidisciplinary care. The additional investigations performed established a low recurrence risk for future pregnancies.

P6

HARNESSING THE POWER OF GUT METAGENOMICS TO TRACE SEAFOOD PROVENANCE: THE ATLANTIC CHUB MACKEREL AS CASE STUDY

Eduardo Feijão[1,2], Irina A. Duarte[1], Marcelo Pereira[2], Pedro Pascoal[2], Rita. P. Vasconcelos[3], Carla Gameiro[3], Susanne E. Tanner[1,4], Ana Rita Matos[2,5], Ricardo Dias[2,6], Andreia Figueiredo[2,5], Vanessa F. Fonseca[1,4] & Bernardo Duarte[1,5]

[1] MARE - Marine and Environmental Sciences Centre & ARNET – Aquatic Research Network Associated Laboratory, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

[2] BioISI—Instituto de Biossistemas e Ciências Integrativas, Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[3] IPMA - Instituto Português do Mar e da Atmosfera, Av. Dr. Alfredo Magalhães Ramalho, 1495-165 Algés, Portugal.

[4] Departamento de Biologia Animal da Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016, Lisboa, Portugal.

[5] Departamento de Biologia Vegetal da Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016, Lisboa, Portugal.

[6] cE3c—Center for Ecology, Evolution and Environmental Changes & CHANGE—Global Change and Sustainability Institute, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

Introduction: This study addresses the challenges in ensuring seafood authenticity and provenance, given the escalating demand for seafood and the proliferation of illegal, unreported, and unregulated (IUU) fishing, which hampers sustainable fisheries. The primary goal is to leverage gut metagenomics together with machine learning models to establish robust markers for determining the geographical origin of commercially important seafood species. The study focuses specifically on the Atlantic chub mackerel (*Scomber colias*), a species of great economic and nutritional significance found along the Portuguese coast.

Methodology: To achieve this objective, fish samples were collected from five major fishing areas along the Portuguese coast, and their intestinal contents underwent metagenomic Nanopore 4th-generation sequencing. Data analysis was performed using a customized pipeline to assess bacterial diversity, both ecologically and functionally, among individuals from different regions. Abundance data was utilized to train random forest (RF) models, which were used to obtain reliable indicators of the specimens' geographical origins. The generation of geographical metagenomic biomarkers involved integrating RF model outcomes with Linear Discriminant Analysis Effect Size (LEfSe) and indicator species analysis (indicspecies package in R).

Results: Ecological differences in the gut bacterial communities of individuals collected from diverse sampling sites were highlighted. The differences were primarily attributed to variations in the abundance of taxonomic branches and units. During model validation, the trained RF models exhibited 85.0% accuracy in pinpointing the capture location of the specimens. Combining this approach with LEfSe and indicspecies analysis identified a condensed set of biomarkers characteristic of each fishing area. Additionally, a functional analysis of bacterial communities unveiled the presence of potentially pathogenic OTUs with varying abundance across fishing areas.

P6

HARNESSING THE POWER OF GUT METAGENOMICS TO TRACE SEAFOOD PROVENANCE: THE ATLANTIC CHUB MACKEREL AS CASE STUDY

Eduardo Feijão[1,2], Irina A. Duarte[1], Marcelo Pereira[2], Pedro Pascoal[2], Rita. P. Vasconcelos[3], Carla Gameiro[3], Susanne E. Tanner[1,4], Ana Rita Matos[2,5], Ricardo Dias[2,6], Andreia Figueiredo[2,5], Vanessa F. Fonseca[1,4] & Bernardo Duarte[1,5]

Conclusions: In summary, this study underscores the potential of gut metagenomic profiles to determine the origin of fish specimens accurately. Moreover, the developed analytical pipeline enables the extraction of biomarker OTUs, which can be employed to ascertain the species' provenance with high precision, reducing the number of OTUs required for traceability. Furthermore, the approach facilitates the assessment of human and fish pathogens in collected specimens, thereby addressing potential risks to human health.

P7

THE COMPLEXITY OF IDENTIFICATION OF PATHOGENIC VARIANTS

Joana Fino[1], Dezso David[1]

[1] Department of Human Genetics, National Health Institute Doutor Ricardo Jorge

Introduction: Natural occurring genomic variant, from single nucleotide to balanced, unbalanced and complex rearrangements, spanning large chromosomal regions, has been reported to cause human pathologies. As such, we present cases with neurodevelopmental disorder, infertility, and recurrent miscarriage, which reflect the complexity of the identification of pathogenic variants, considering the variation spectrum, the underlying pathogenic mechanisms, and the heterogeneous clinical presentations.

Methodology: Long and small insert genomic sequencing (GS) was applied to four cases. Variants were identified from GS data mapped against the reference human genome and confirmed through Sanger sequencing. Results were interpreted using SVInterpreter, Exomiser, genotype-phenotype correlation and convergent genomic data analysis.

Results: Although the first case is a carrier of a t(17;19)(p13.1;p13.3)mat, disrupting *GSG1L2,* and of a presumably paternally inherited dup(2)(q14.3q21.1), encompassing the autosomal dominant (AD) phenotype-associated *PROC* and *HS6ST1* genes, the identified novel frameshift c.4442del, p.(Gly1481Valfs*21) variant of *CHD4*, was considered the disease-causing variant, since the proband's phenotype fits the *CHD4*-associated Sifrim-Hitz-Weiss syndrome (Da Silva et al., 2022). Cases 2 and 3 were both reported with infertility, and carriers of t(5;9)(q31.3;p13) and t(4;21)(p14;q21.3), respectively. Our study revealed that the phenotype most plausibly resulted from a chromosomal position effect over *YIPF5* and *SPATC1L*. The last case, presented intellectual disability and recurrent miscarriage, associated to t(7;22)(p13;q13.1). The 7p13 breakpoint disrupts the brain specific *CAMK2B,* causing AD mental retardation 54 (OMIM #617799), whereas increased meiotic segregation of der(22), during gametogenesis, most likely explain the reported miscarriage (David et al., 2023).

Conclusions: These cases highlight the intricacy of pathogenic mechanisms leading to human disorders, the necessity for identification and evaluation of the "full" spectrum of genomic and genetic variants, of comparative reverse phenotyping, including patients with pathogenic variants affecting the same genes. Finally, highlight the need of introducing a more precise genomic medicine in clinical practice.

Funding: This research was supported by FCT—*Fundação para a Ciência e a Tecnologia*, Research Grant HMSP-ICT/0016/2013 of the Harvard Medical School—Portugal Program.

P8

LIVING WITHOUT A THYMUS IMPACT ON CD4 T CELL COMPARTMENT

Beatriz Moleirinho[1], Margarida Paulo-Pedro[1], Diana F. Santos[1], André M.C. Gomes[1], Susana L. Silva[1,2], Afonso R.M. Almeida[1], Ana E. Sousa[1]

[1] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa; [2]Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, Portugal.

Introduction: CD4 T cells play a pivotal role in the immune system by orchestrating and regulating immune responses, acting as conductors and coordinators of the activity of other immune cells. Peripheral CD4 T-cell homeostasis results from a dynamic equilibrium between the continuous replenishment by cells egressing the thymus, where T cells are produced, and peripheral mechanisms involving cell proliferation, survival and differentiation of naïve CD4 T cells into effector/regulatory populations. The thymic output is currently considered essential to maintain repertoire diversity throughout life. On the other hand, peripheral homeostatic mechanisms have been shown to be able to maintain naïve T cell counts. The latter is particularly important to counter-act the progressive decline of thymic activity with aging and the accelerated differentiation of naïve cells into memory/effector cells in the context of chronic infections.

Methodology: Thymectomy performed during corrective cardiac surgery in early infancy provides an extreme clinical setting to gain insights into T cell homeostasis. To investigate the long-term impact of thymectomy on the CD4 T-cell compartment we studied adults, 30 years after thymus removal in early childhood, alongside healthy age-matched controls. Circulating T-cells were evaluated by spectral flow cytometry, and, in parallel, sort-purified CD4 T cells were used to perform scRNAseq combined with TCR and cell surface protein sequencing (CITE-Seq). For the CITE-seq analysis, we selected naïve/memory markers and CD31, which we have previously shown to identify naïve cells with a higher proliferative capacity in response to IL-7.

Results: Our longitudinal analysis of these individuals revealed not only the resilience of the naïve compartment but also points to its expansion. We are currently focusing our analysis on the underlying mechanisms.

Conclusion: Ultimately, we expect to provide new insights on human CD4 T cell biology and identify pathways for immune reconstitution.

P9

HUMAN GENETIC SUSCEPTIBILITY TO INFECTION BY *Coxiella burnetii* PRELIMINARY RESULTS OF MASSIVE PARALLEL SEQUENCING

Susana David[1,2], Liliana Castro[1], Elsa Duarte[3], Miguel Machado[1], Ulisses Gaspar[1], Maria Vanessa Cueto Rojo[1], Joana Mendonça[1], José Ferrão[1], Miguel Machado[1], João Lavinha[1], Luís Vieira [1], Ana Sofia Santos[4,5]

[1] Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), Lisbon, Portugal.

[2] Host-Pathogen Interactions Unit, Research Institute for Medicines, iMed.ULisboa, Faculty of Pharmacy, Universidade de Lisboa, Lisbon, Portugal.

[3] MED – Instituto Mediterrâneo para a Agricultura, Ambiente e Desenvolvimento, Escola de Ciências e Tecnologia, Universidade de Evora, Portugal.

[4] Centro de Estudos de Vectores e Doenças Infecciosas Dr. Francisco Cambournac (CEVDI), Departamento de Doenças Infeciosas (INSA,IP), Lisboa, Portugal.

[5] Instituto de Saúde Ambiental, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, 1649-028 Lisboa, Portugal

Introduction: Host genetic factors are believed to influence the development of severe chronic Q fever upon infection by the etiological agent, *Coxiella burnetii*. Causal variants governing genetic susceptibility to infection have been particularly relevant to the management of other intracellular pathogens, sustained by cell-mediated immunity and the formation of granuloma. Candidate genes implicated in the immune response to these pathophysiological features were genotyped in a case-control genetic association study to identify genetic variants associated with the development of chronic Q fever. Variants identified were included in a replication study.

Methodology: NGS was used for genotyping 34 healthy individuals and 60 Q fever patients (43 acute and 17 chronic forms) using Nextera library construction and Illumina® sequencing. Exploratory data analysis and genotype-phenotype association were performed using PLINK. Sanger sequencing was used to validate the NGS findings and in an independent replication study on 27 Q fever patients of the same ethnic origin as in the original study (20 acute and 7 chronic forms).

Results: NGS provided 607 SNPs from 94 DNA samples. Of this 405 SNPs and 82 DNA samples passed quality control. 4 SNPs, in the genomic locus containing the *IFNGR2* and the *TMEM50B* genes, showed statistically significant variations between chronic and acute patients, passing the Bonferroni criteria. The NGS results were confirmed by Sanger sequencing. The trends in the association were confirmed in the replication study although without statistical significance due to the reduced sample size. The SNPs were identified as expression quantitative trait *loci* (eQTLs).

Conclusion: The NGS strategy used appeared promising for identifying genetic determinants of chronic Q fever. To test the genetic hypothesis an *in vitro* cell culture model will be generated.

Acknowledgements: This study was partially supported by the FCT grant PTDC/SAU-SAP/115266/2009 to A.S.

P10

EVOLUTION OF THE GUANYLATE BINDING PROTEIN (GBP) MULTI-GENE FAMILY IN BATS (Chrioptera)

J. Ricardo Borges[1,2,3], Ana Pinheiro[1,2], Pedro J. Esteves[1,2,3]

[1] CIBIO-UP, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus Agrário de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal

[2] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, 4485-661 Vairão, Portugal

[3] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal

Introduction: GBPs (guanylate binding proteins), an evolutionary ancient protein family, play a key role in the host innate immune response against bacterial, parasitic and viral infections. In Humans, seven GBP genes have been described (GBP1-7). Despite the interest that these proteins have received in the last years, evolutionary studies have only been performed in primates, Tupaia and rodents. These have shown a pattern of gain and loss of genes in each family, indicative for the birth-and-death evolution process. In this study, we analyzed the evolution of this gene cluster for several bat species, belonging to the Yangochiroptera and Yinpterochiroptera sub-orders.

Methodology: Coding sequences (CDSs) of GBP genes from various bat species were sourced from publicly available databases, including the NCBI and Ensembl. The inclusion and validation of bat species with sequenced genomes containing GBP genes were ensured using Basic Local Alignment Search Tool (BLAST). To investigate the genomic location and organization of GBP genes, synteny analysis was conducted using the Genome Data Viewer platform offered by NCBI. An examination of recombination events between GBP genes was undertaken using the online platform Genetic Algorithm for Recombination Detection (GARD). The CDSs of bat genes, as well as those of related outgroups, were subjected to alignment in the BioEdit software, utilizing the ClustalW algorithm. Following this alignment, meticulous manual corrections were implemented to guarantee the precision of the alignment, rectifying issues such as frameshifts and other alignment anomalies. To study the evolutionary relationships among GBP genes in bats, the aligned sequences were subjected to further analysis using the Molecular Evolutionary Genetics Analysis (MEGA) software.

Results: Detailed analysis show a conserved synteny and a history of gene expansion and loss. Phylogenetic analysis showed that bats have GBPs 1,2 and 4-6. GBP2 has been lost in several bat families being present only in Hipposideidae and Pteropodidae. GBPs1, 4 and 5 are present mostly as single copy genes in all families but have suffered duplication events, particularly in Myotis myotis and Eptesicus fuscus. Most interestingly, we demonstrate that GBP6 duplicated in a Chiroptera ancestral species into two genes, we called GBP6a and GBP6b, with different subsequent evolutionary histories. GBP6a has suffered several duplications in all families while GBP6b is present as a single copy gene and has been lost in Pteropodidae, Miniopteridae and Desmodus rotundus, a Phyllostomidae. With more than 14 GBP genes, Eptesicus fuscus and Myotis myotis stand out has having far more copies than all other studied bat species. Antagonistically, Pteropodidae have the lowest number of GBP genes in bats.

P10

EVOLUTION OF THE GUANYLATE BINDING PROTEIN (GBP) MULTI-GENE FAMILY IN BATS (Chrioptera)

J. Ricardo Borges[1,2,3], Ana Pinheiro[1,2], Pedro J. Esteves[1,2,3]

Conclusion: Bats are important reservoirs of viruses, many of which have become zoonotic diseases in the last decades. The study's findings unveil intriguing aspects of the evolutionary processes within the GBP gene family in bats. They indicate that these genes have undergone deletions, duplications, and differentiations over time, in line with the widely recognized birth-and-death evolutionary model typical of multigene families. Additionally, the research highlights a notable expansion of the GBP gene family in Myotis myotis and Eptesicus fuscus, while a reduction is observed in members of the Pteropodidae family. These trends are consistent with findings in similar studies focusing on other immune system-related genes.
A significant revelation in this study pertains to a gene duplication event, which led to the emergence of two new genes, GBP6a and GBP6b, each characterized by unique amino acids that set them apart from other GBP genes. Consequently, the study suggests that further dedicated research is needed to uncover the specific functions of these newly identified genes. Furthermore, the research raises questions about the nomenclature used for this gene family within the order. The results challenge the accuracy of the current nomenclature, particularly concerning the presence of GBP7 genes, which appear to be exclusive to primates.
To facilitate future investigations of a similar nature, providing more comprehensive insights across a wider range of species, it is crucial to expand the sequencing efforts for bat species and enhance the quality of existing genomic data. This approach will enable more thorough examinations of the evolutionary history and diversity of GBP genes in bats.

P11

ADDRESSING UNCERTAINTIES OF THE SPECIES-LEVEL TAXONOMIC CLASSIFICATION OF A *Pseudomonas sp.*

Pedro Soares-Castro[1], Filipa Soares [2], Pedro M. Santos [2]

[1] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
[2] Centre of Molecular and Environmental Biology, University of Minho, Braga, Portugal

Introduction: In the last decade, the taxonomy of the *Pseudomonas* genus (over 200 bacterial species) has greatly evolved due to the democratization of high-throughput sequencing. This has posed a challenge given the dubious taxonomic position of some strains already classified several years ago.

Despite being a very well-studied strain with its genome fully sequenced, the species-level classification of *Pseudomonas* sp. M1 is still unclear. Strain M1 holds a great biotech potential, being able to metabolize several toxic/recalcitrant hydrocarbons (*e.g.*, benzene, phenols), and producing value-added aroma compounds like myrcen-8-ol. Considering the current number of assigned *Pseudomonas* species, this work aimed to clarify whether strain M1 would be a member of a known species or represents a novel *Pseudomonas* species.

Methodology: *Pseudomonas* sp. M1 has been studied for over 20 years, covering genomics, proteomics, phenomics, and metabolite identification of its unique hydrocarbon catabolism. A bioprospection survey identified 9 *Pseudomonas* strains sharing over 99% of genome sequence identity, despite being isolated from different geographical locations and at different times. This work integrates the holistic characterization of strain M1, comparing it to closely related *Pseudomonas* spp. for taxonomical classification.

Results: Based on Average Nucleotide Identity (ANI) and *in silico* DNA-DNA hybridization (DH), strain M1 does not meet the currently defined cut-off to be designated as a new species, being very close to *Pseudomonas citronellolis* species (96-97% ANI, 95% cut-off; 70-85% DH identity, 70% cut-off). Physiological and Biolog plate data showed significant differences between strain M1 and *P. citronellolis* type strain DSM50332 in the primary metabolism of sugars, carboxylic acids, amino acids and alcohols.

Conclusion: Comparison of the M1-DSM50332^T genome sequence identity is close to the threshold values for species definition. However, the genetic and metabolic profile, and growth performance of both strains are significantly different, suggesting a divergent evolutionary track, probably leading to speciation.

Acknowledgment: sequencing services by Jörg Becker and João Sobral at Instituto Gulbenkian de Ciência.

Funding: PTDC/EAM-AMB/30488/2017, POCI01–0145-FEDER-022184, POCI-01–0145-FEDER-007569, by FCT I.P. and the European Regional Development Fund (ERDF), through COMPETE 2020—POCI; Ph.D. grants SFRH/BD/76894/2011 and PD/BD/146184/2019.

P12

MODERN LINEAGES OF THE *Mycobacterium tuberculosis* COMPLEX ARE ASSOCIATED WITH INCREASED CROSS-BORDER TRANSMISSION OF MULTIDRUG-RESISTANT TUBERCULOSIS

Pedro Gomes[1], Jody Phelan[2], Susana Campino[2], Taane Clark[2], Isabel Portugal[1], João Perdigão[1]

[1] Faculdade de Farmácia da Universidade de Lisboa

[2] London School of Hygiene and Tropical Medicine

Introduction: The "modern" lineages of *Mycobacterium tuberculosis* complex (MTBC) evolved through large successive genomic deletions, often associated with increased virulence. Higher mutation rates are also attributed to recently evolved lineages, which could substantiate their increased prevalence in outbreaks of multidrug-resistant tuberculosis (TB). This study intends to highlight the global dissemination of the major MTBC lineages, their associated drug resistance genotypes and uncover which lineages show an increased association with cross-border transmission of drug-resistant strains.

Methodology: A hierarchical clustering analysis was employed on a sample comprising 16,605 *M. tuberculosis* genomes, across 107 countries, based whole genome sequencing data. Clusters with epidemiological links were inferred though a distance of 5 single nucleotide polymorphisms (SNP) and those with strains isolated from different countries suggest cross-border transmission.

Results: Genomic characterization of lineage and drug resistance genotypes revealed eight MTBC lineages and four major drug resistance profiles. Globally, lineages 2 and 4 are the predominant genotypes (29.7–47.0%, respectively) across the majority of geographical regions, with emphasis on those within the western and eastern hemispheres, respectively. All lineages concentrate around their respective places of origin, although, for some regions a sizable proportion of TB cases come from atypical genetic backgrounds associated with immigrant populations. A significant 3,055 strains (54.0%) were classified as multidrug-resistant (MDR). Higher clustering rates were found amongst strains belonging to "modern" lineages (33.9–42.4%) as opposed to those belonging to "ancestral" lineages (7.1–29.4%). This implies 'modern' lineages have evolved to increase their virulence, allowing for a greater infectious potential that likely contributed the widespread proliferation of these strains in recent times. A majority of 75 cross-border clusters (96.2%) involve strains belonging to "modern" lineages, and 35 (77.8%) comprise of MDR strains, the majority of which are accredited to lineage 2 (77.1%).

Conclusion: The probability of novel *M. tuberculosis* infections deriving from "modern" MTBC lineages is high and likely to surge given their association with increased transmission. Multidrug-resistant TB rates are also likely to increase due to dissemination between countries driven by increased globalization.

P13

UNRAVELLING miRNA-3'UTR REGULATORY NETWORKS IN PRIMARY HUMAN CD4 T CELLS THROUGH SHORT AND LONG-READ SEQUENCING APPROACHES

Tomás Araujo[1], Cláudia Noronha-Estima[1], Pedro Pascoal[1], Ricardo Dias [1,2], Margarida Gama-Carvalho[1]

[1] BioISI—Instituto de Biossistemas e Ciências Integrativas, Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[2] cE3c—Center for Ecology, Evolution and Environmental Changes & CHANGE—Global Change and Sustainability Institute, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

Introduction: microRNAs (miRNAs) are key regulators of cellular programs that act through base-pairing interactions with mRNA sequences to inhibit their translation and/or promote degradation, predominantly through 3' UTR located miR recognition elements (MREs). Although several target-prediction algorithms have been developed to identify regulatory miRNA-mRNA interactions, their accuracy is very limited, with both high numbers of true and false positives. One of the reasons for this problem is the dynamic nature of 3' UTR sequences, that vary from cell type to cell type depending on the use of alternative poly-adenylation sites, while also being subject to both regulatory binding by RNA binding proteins and base-modification events.

Methodology: Our lab has been studying a novel miR regulatory network active in human primary CD4 T cells upon TCR stimulation. To obtain a comprehensive characterization of the relevant interactions, we are gathering data on the impact of miR antagonists on the cellular transcriptome using short read Illumina sequencing and generating a detailed map of expressed mRNA isoforms and 3'UTRs using nanopore cDNA sequencing.

Conclusion: Our preliminary results are starting to reveal novel insights into the complexity of the underlying regulatory networks and their role on primary T cell activation and function.

P14

INNOVATIVE MEDICAL EDUCATION: GENOMIC MEDICINE ELECTIVE COURSE ENHANCING COMMUNICATION AND ETHICAL UNDERSTANDING

Célia Azevedo Soares[1,2,3,4,5], Lúcia Lacerda[1,2,3], Paula Jorge[1,2,3]

[1] Departamento de Microscopia, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Portugal

[2] Centro de Genética Médica Dr. Jacinto Magalhães, Centro Hospitalar Universitário de Sto. António, Porto, Portugal

[3] Unidade Multidisciplinar de Investigação Biomédica, Instituto de Ciências Biomédicas Abel Salazar (UMIB/ICBAS) and Laboratory for Integrative and Translational Research in Population Health (ITR), Universidade do Porto, Porto, Portugal

[4] Departamento de Ciências Médicas, Universidade de Aveiro, Aveiro, Portugal.

[5] i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

Introduction: Genomic Medicine is at the forefront of modern healthcare, revolutionizing diagnostics and personalized treatment. The rapid advancements in genomics demand that medical students not only comprehend the science but also acquire skills in ethical considerations, effective communication, and navigating the complex patient diagnostic odyssey. To address this educational need, we developed an innovative elective course for medical students.

Methodology: Our Genomic Medicine elective at Abel Salazar Biomedical Sciences Institute - University of Porto integrates a multifaceted approach, offering comprehensive insights into the field including five main perspectives: 1) basic concepts in Genomic Medicine, 2) ethics in Genomic Medicine, 3) patient diagnostic odyssey, 4) national and international initiatives, and 5) communication skills. One unique aspect of our course is the inclusion of non-health professional guest lecturers: a patient association's representative/caregiver of a patient with a rare genetic disorder, as well as a health journalist.

Results: Our evaluation process focuses on the students' ability to convey complex Genomic Medicine concepts to three distinct audiences: a family, a patient organization member, and a health specialized journalist. Student reports reveal significant improvements in students' knowledge and communication skills. The caregiver's personal account provided valuable insight into the emotional and practical aspects of genetic/genomic conditions, enhancing students' empathetic understanding. Media-led sessions improved students' ability to translate intricate genomic information into layman's terms.

Conclusions: This innovative Genomic Medicine elective course enriches medical education by addressing the growing importance of genomic knowledge, ethics, and effective communication. It equips future healthcare professionals with the necessary tools to navigate the evolving landscape of genomic medicine, bridging the gap between scientific expertise and patient-centered care. By fostering empathy, ethical awareness, and communication skills, this course prepares students to become well-rounded and compassionate medical practitioners in the era of Genomic Medicine. We believe that our approach holds significant promise for advancing medical education and patient care in the genomic era.

P15

A PILOT INITIATIVE TO SEQUENCE THE GENOMES OF PORTUGUESE SPECIES

João P. Marques[1,2], Paulo C. Alves[1,2,3,4], Isabel R. Amorim[5], Ricardo J. Lopes[1,2,6], Mónica Moura[2,7,8], Manuela Sim-Sim[9,10,11], Carla Sousa-Santos[12], Vitor C. Sousa[a,9,13] and José Melo-Ferreira[a,1,2,3]

[1] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal

[2] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal

[3] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4099-002 Porto, Portugal

[4] EBM - Estação Biológica de Mértola,, 7750-329 Mértola, Portugal

[5] cE3c-Centre for Ecology, Evolution and Environmental Changes, Azorean Biodiversity Group, CHANGE – Global Change and Sustainability Institute, and U. of the Azores, Rua Capitão João d´Ávila, Pico da Urze, 9700-042, Angra do Heroísmo, Portugal

[6] MHNC-UP, Natural History and Science Museum of the University of Porto, 4099-002 Porto, Portugal

[7] CIBIO, Research Centre in Biodiversity and Genetic Resources, InBIO Associated Laboratory, Azores Unit, University of the Azores, Campus of Ponta Delgada, Rua da Mãe de Deus 13A, 9500-321 Ponta Delgada, Azores, Portugal

[8] UNESCO Chair – Land Within Sea: Biodiversity & Sustainability in Atlantic Islands

[9] cE3c - Centre for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[10] MUHNAC - Museu Nacional de História Natural e da Ciência, Universidade de Lisboa, Rua da Escola Politécnica, 58, 1250-102 Lisboa, Portugal

[11] Departamento de Biologia Vegetal, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[12] MARE-Marine and Environmental Sciences Centre/ARNET-Aquatic Research Network, ISPA-Instituto Superior de Ciências Psicológicas, Sociais e da Vida, Rua Jardim do Tabaco, 34, 1149-041 Lisbon, Portugal.

[13] Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

Introduction: Portugal's diverse landscape, stretching from continental Europe to the Atlantic archipelagos, has harboured a wealth of biodiversity shaped by Pleistocene glacial cycles. However, this biological richness faces imminent threats from anthropogenic factors, including climate change, invasive species, land use alterations, overexploitation, and the resurgence of pathogens. Genomics emerges as a powerful tool to generate knowledge about biodiversity that is essential to devise proper conservation strategies. To anchor such studies, high quality reference genomes representing the species are key resources, and several international initiatives are addressing this challenge. One of these initiatives is the European Reference Genome Atlas (ERGA), which is pioneering the building of a European knowledge and infrastructural network for reference genome generation.

P15

A PILOT INITIATIVE TO SEQUENCE THE GENOMES OF PORTUGUESE SPECIES

João P. Marques[1,2], Paulo C. Alves[1,2,3,4], Isabel R. Amorim[5], Ricardo J. Lopes[1,2,6], Mónica Moura[2,7,8], Manuela Sim-Sim[9,10,11], Carla Sousa-Santos[12], Vitor C. Sousa[a,9,13] and José Melo-Ferreira[a,1,2,3]

Methodology: As part of this effort, ERGA has developed a pilot project for reference genomes generation and assembly, which included six Portuguese species. These species, chosen for their endemic, iconic or endangered status, encompass plants, insects, and vertebrates (including fish, birds, and mammals) found in mainland Portugal and the Azores Islands.

This presentation will discuss the contribution and status of the Portuguese contribution to the ERGA pilot project.

P16

LOOKING FOR MICROBIAL INDICATOR SPECIES BY LONG READ NANOPORE SEQUENCING

Ana Cruz-Silva [1,2], Gonçalo Laureano [1,2], Marcelo Pereira [1], Ricardo Dias [1,3,4], Margarida Gama-Carvalho [1,5], Fiammetta Alagna [6], Bernardo Duarte [3,7], Andreia Figueiredo [1,2,3]

[1] Biosystems & Integrative Sciences Institute (BioISI), Faculdade de Ciências da Universidade de Lisboa,1749-016 Lisboa, Portugal

[2] Grapevine Pathogen Systems Lab, BioISI Faculdade de Ciências da Universidade de Lisboa,1749-016 Lisboa, Portugal

[3] Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[4] cE3c—Center for Ecology, Evolution and Environmental Changes & CHANGE—Global Change and Sustainability Institute, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[5] Departamento de Química e Bioquímica, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

[6] Energy Technologies and Renewable Sources Department, National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Trisaia Research Centre, 75026 Rotondella, MT, Italy

[7] MARE—Marine and Environmental Sciences Centre & ARNET—Aquatic Research Infrastructure NetworkAssociate Laboratory, Faculty of Sciences of the University of Lisbon,Campo Grande, 1749-016 Lisbon, Portugal

Introduction: Grapevine is one of the most important fruit crops worldwide and Portugal is one of the top wine producers. It is well established that sensory characteristics of wine from a particular region are defined by the physiological responses of the grapevine to its environment; thus, the concept of terroir in viticulture was established. Among all the factors that contribute to terroir definition, soil microorganisms play a major role from nutrient recycling to a drastic influence on plant fitness (growth and protection) and in wine production. In this study, we aimed to define microbiome signatures and indicator species of different terroirs in an organic vineyard located in the demarcated wine region of Douro Valley, the Quinta dos Murças. This vineyard presents a unique topography with six terroirs previously identified based on the edaphoclimatic characteristics.

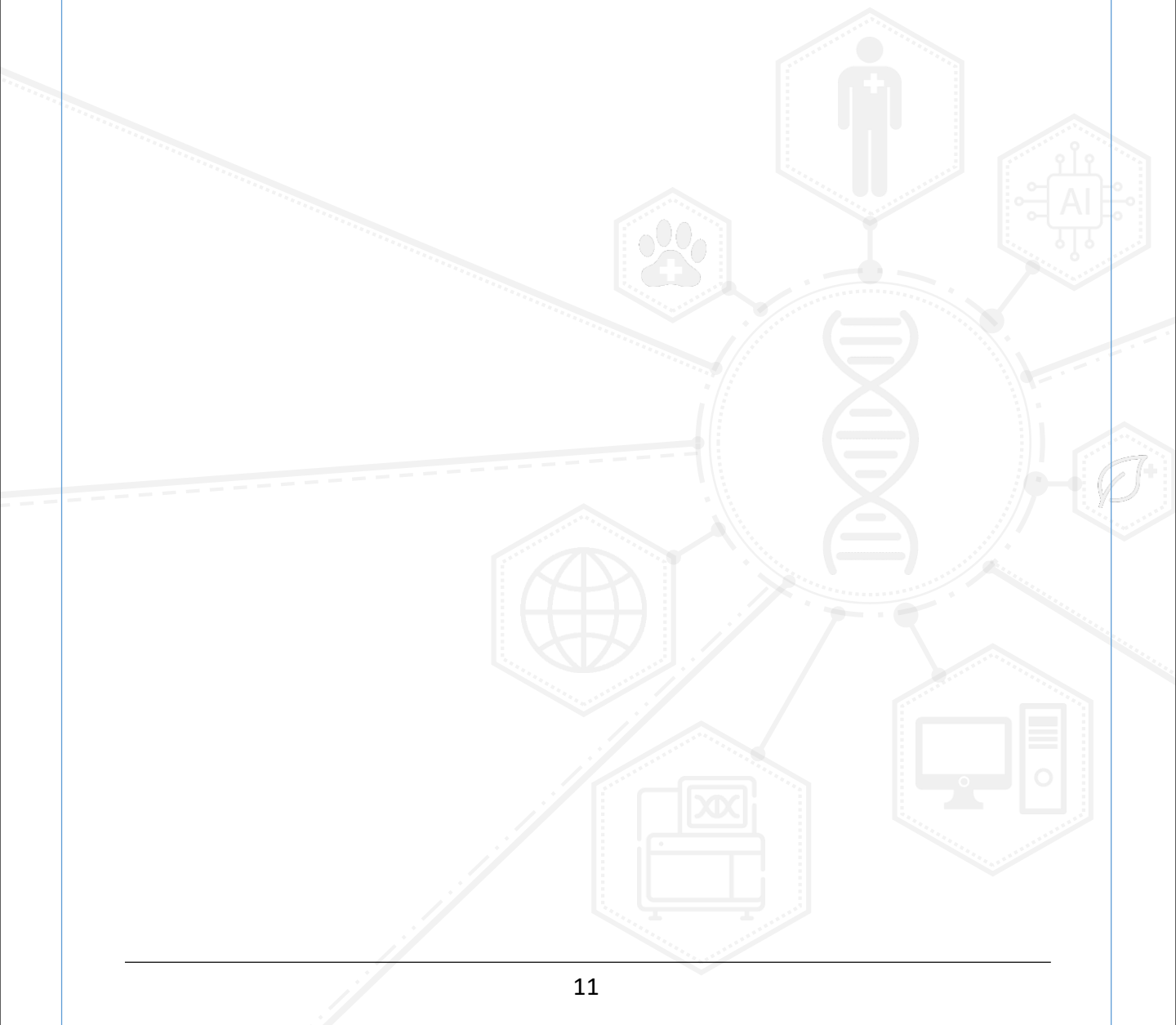
Methodology: Through long-read Oxford Nanopore sequencing we characterized the soil microbial diversity of four of the terroirs present at Quinta dos Murças.

Results: Our results demonstrate that the soil microbiome may constitute a terroir signature. Each terroir was also associated with signature functions, ecologies, and indicator species, suggesting that microbiome analysis is a viable method to distinguish terroirs. The indicator species identified are not only important to differentiate terroirs but potentially impact plant development and wine production.

Conclusion: The results pinpointed may aid the definition of adequate viticulture and oenological practices.

References: Cruz-Silva A, et al. A New Perspective for Vineyard Terroir Identity: Looking for Microbial Indicator Species by Long Read Nanopore Sequencing. Microorganisms. 2023; 11(3):672. https://doi.org/10.3390/microorganisms11030672

Genome^{PT} is funded by


AI