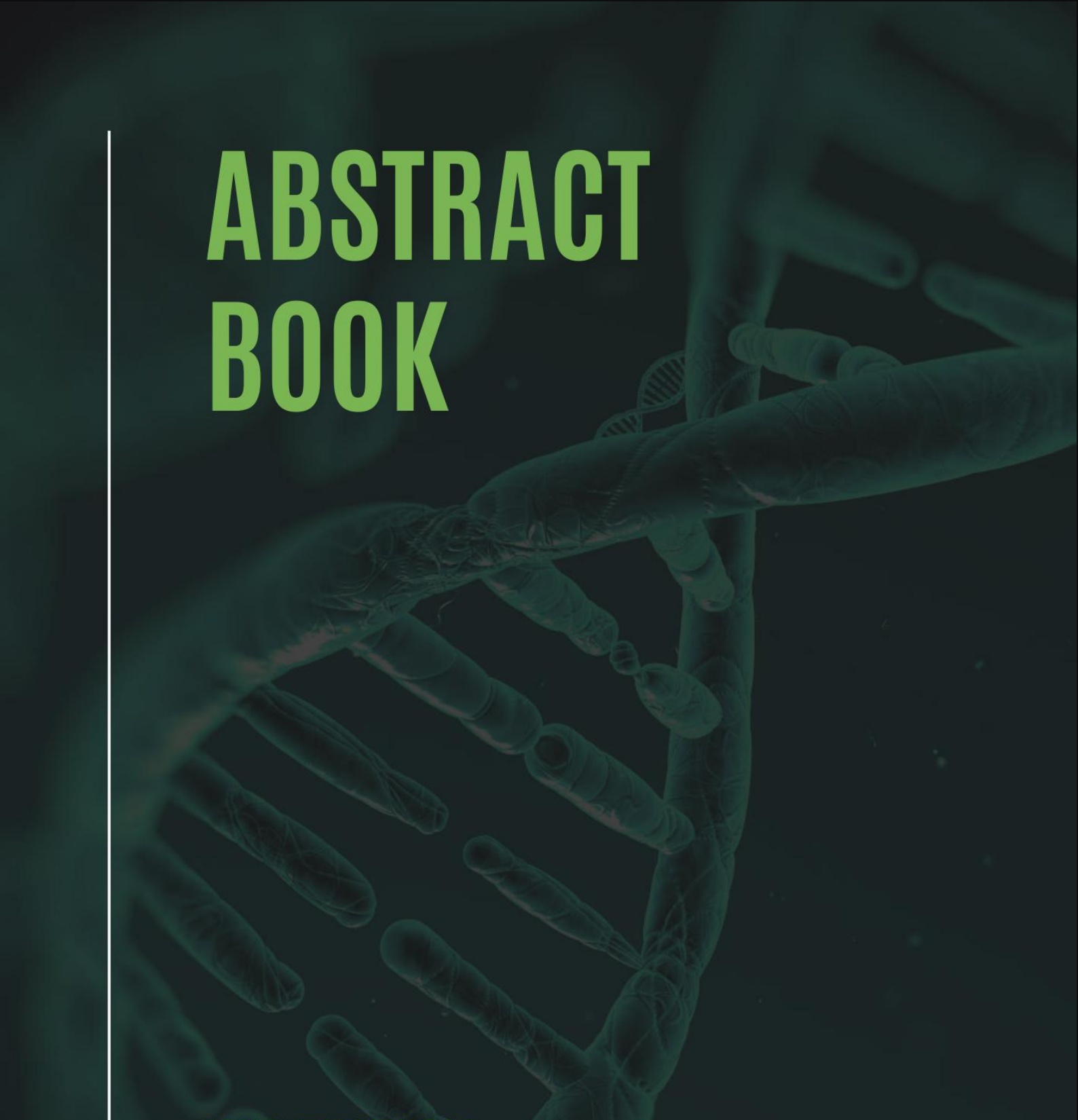
ABSTRACT BOOK



XIV EDITION
26 - 28 MARCH

University of Minho
Gualtar Campus

Bioinformatics
Open Days 2025

Bioinformatics Open Days 2025

SCIENTIFIC SUBMISSIONS

ORAL PRESENTATIONS

Nº	Authors	Title
O1	Isabel Duarte	Mitochondrial Gene Expression Is Independent of Organ Metabolic Rates: Do Cancer Selective Pressures Override the Tumor Microenvironment?
O2	João Miranda	The Neighbour Effect: How Interacting Proteins Influence Driver Mutations and Cancer Patient Outcomes
O3	Bruno Sá	Implementing Metabolic Transformation Algorithms and their application in Ageing-related research
O4	Rogério Ribeiro	Integration of multi-modal datasets to estimate human aging
O5	Ricardo Pinto	DepMap and GDSC data mining to inform ancestry-driven mechanisms in triple-negative breast cancer
O6	Rita Teixeira	Guiding computational design for SARS-CoV-2 Spike protein using Molecular Dynamics simulations
O7	Pedro Fernandes	Tracing the prevalence and pathogenicity of artifact mutations in ancient Human mitochondrial DNA
O8	Gonçalo Apolinário	DeepTransyt, annotation of transporter proteins using deep learning
O9	João Correia	Improving Docking Predictions with a ML-Based Scoring Ensemble
O10	Alexandra Balola	Computational Engineering of PETase for Improved PET plastics Substrate Affinity and Degradation Efficiency
O11	Herlander Azevedo	Application of bioinformatics in the characterization of multiple aspects of grapevine diversity
O12	Ricardo Franco-Duarte	Phylogenomics and functional annotation of 530 non-Saccharomyces yeasts from winemaking environments reveals their fermentome and flavorome
O13	Luis Gonçalves	Genome-wide characterization of plant resistance genes in cork oak (Quercus suber)
O14	Raquel Ríos-Castro	Eukaryotic Communities and Potential Pathogens in Wastewater Effluents in Ría de Vigo
O15	José Morim	Deciphering the venoms of Conus species with transcriptomics

O1 - Mitochondrial Gene Expression Is Independent of Organ Metabolic Rates: Do Cancer Selective Pressures Override the Tumor Microenvironment?

Catarina Gomes Ferreira[1], Miguel Rocha[1*], Isabel Duarte[2,3*]

[1]Centro de Engenharia Biológica, Universidade do Minho, Braga, Portugal; [2]RISE Health, Universidade do Algarve, Faro, Portugal; [3]Pattern Institute, Faro, Portugal; [*]*Equal contribution*

Metabolic reprogramming is a recognized hallmark of cancer, playing a crucial role in its initiation and progression. Recent evidence highlights the impact of non-cancer cell metabolism within the tumor microenvironment in modulating tumor development, emphasizing the importance of metabolic factors in understanding cancer biology.

To identify differentially expressed genes (DEGs) in cancer, with a specific focus on mitochondria-located genes, and to compare gene expression profiles between organs with high metabolic-rates (brain, liver, kidneys) and low metabolic-rates (bladder, colon, skin).

RNA-seq data for normal and cancer tissues from the six organs under study were sourced from the GTEx and TCGA databases. The dataset included 2,675 normal tissue samples from GTEx, and 3,633 cancer tissue samples from TCGA. Differential expression analysis was performed using General Linear Models (GLMs), while hierarchical and soft fuzzy clustering were applied to identify distinct gene expression patterns. All analyses were performed in R, and the code is openly accessible in GitHub [https://github.com/orgs/MitoProfiles/repositories].

The analysis identified mitochondrial DEGs such as ACSM1, ACSM5, and PRODH, likely reflecting cancer cell adaptations to metabolic and microenvironmental stress. Significant differences in FDX2 (iron-sulfur protein biogenesis) and ACSM2B (free fatty acid activation) suggest the involvement of these pathways in oncogenesis. However, the expression differences were largest between normal and cancer tissues, overshadowing the variation between high and low metabolic-rate organs.Mitochondrial gene expression profiles differ more significantly between cancer and normal tissues than between cancers originating from organs with varying metabolic-rates. This highlights the dominant role of cancer-induced cellular reprogramming over organ-specific microenvironment signals. This raises the hypothesis of a functional convergence of mitochondrial pathways in cancer, particularly those involved in processes such as iron-sulfur cluster biogenesis and fatty acid metabolism, which may be selectively enhanced/suppressed to meet the metabolic demands of cancer cells. This suggests the possibility of a general cancer mitochondrial expression-signature, warranting further investigation.

O2 - The Neighbour Effect: How Interacting Proteins Influence Driver Mutations and Cancer Patient Outcomes

Márcia F. S. Vital[1,§], João A. I. Miranda[1,2,§], Margarida Carrolo[3,§], António Quintela[3], Francisco R. Pinto[1]

[1]BioISI—Biosystems and Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
[2]INESC-ID—Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa, Portugal
[3]CUF Oncology—Hospital CUF Descobertas, Lisboa, Portugal
[§]These authors have contributed equally to this work

Cancer is driven by the accumulation of somatic mutations, including driver mutations that confer selective advantages to cancer cells. These driver proteins are central players in several cellular processes and their function is performed in conjugation with neighbouring interacting partners. Here, we evaluated the association between driver mutations and neighbouring genes expression. We hypothesize that neighbour abundance can modulate the phenotypic effects of driver mutations and influence patient survival.

We used correlation analysis and linear regression models on a “The Cancer Genome Atlas” pan-cancer cohort to identify associations between driver mutations and neighbour gene expression. Proportional hazard models were applied to evaluate the impact of neighbour expression on overall survival and to selected neighbours simultaneously enriched in driver and survival associations with coherent effect signs. Experimental evidence supporting the causal roles of selected neighbours was manually curated.

We found a significant correlation between the number of driver associations for each neighbour and sign-coherent survival associations. Neighbours enriched in positive driver-association were more likely to be enriched in survival associations, particularly where high neighbour expression correlates with increased driver mutations and poorer survival outcome. 87% of the 247 positively associated neighbours identified have causal roles supported by literature, compared to 52% of the 39 negatively associated neighbours.

Our study suggests a strong link between driver mutations, neighbour gene expression, and patient survival in cancer. The identified neighbours represent promising candidates for experimental validation of their functional roles and for exploring their therapeutic potential as drug targets.

O3 - Implementing Metabolic Transformation Algorithms and their application in Ageing-related research

Bruno Sá[1], Alexandre Oliveira[1] and Miguel Rocha[12]

[1]Centre of Biological Engineering, University of Minho, 4710-057, Braga, Portugal
[2]LABBELS – Associate Laboratory, Braga/Guimarães, Portugal

This work presents the Python-based implementation and validation of the Metabolic Transformation Algorithm (MTA) and its robust counterpart (rMTA), originally developed in MATLAB. MTA/rMTA facilitates the identification of therapeutic metabolic interventions by simulating gene knockouts and evaluating their potential to redirect metabolic fluxes toward healthier phenotypes. The transition to Python capitalizes on its open-source ecosystem, strong scientific computing libraries, and integrated frameworks for constraint-based modeling and optimization.

The new Python toolkit utilizes the Gurobi Python API for addressing core Mixed-Integer Quadratic Programming (MIQP) tasks and offers a modular, extensible pipeline for efficiently simulating metabolic perturbations. Validation against the original MATLAB implementations, including studies on RRM1 and RRM2 gene knockouts, confirmed strong consistency in outputs, ensuring methodological reliability and reproducibility.

Beyond replication, the pipeline was applied to an ageing-related case study in *Caenorhabditis elegans*, focusing on metabolic responses to the knockdown of the unc-62 gene. This analysis uncovered key candidate genes and pathways of potential therapeutic relevance, illustrating rMTA's capacity to navigate complex solution spaces and refine intervention strategies.

A key innovation of this work is the integration of Evolutionary Algorithms (EAs) into rMTA, enabling the optimization of multi-gene knockout strategies. By exploring synergistic gene deletions, the enhanced approach identifies interventions that more effectively shift metabolic states toward healthier phenotypes.

Overall, the Python-based MTA/rMTA software, augmented by evolutionary optimization techniques, provides a more accessible, scalable, and versatile resource for metabolic engineering and therapeutic target discovery. Its successful application to ageing-related metabolic research underlines its value for tackling intricate biological questions and advancing computational approaches in systems biology.

O4 - Integration of multi-modal datasets to estimate human aging

Rogério Ribeiro[1,2], Athos Moraes[1,2], Marta Moreno[1,2], Pedro G. Ferreira[1,2]

[1]Department of Computer Science, Faculty of Sciences, University of Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal

[2]Laboratory of Artificial Intelligence and Decision Support, Institute for Systems and Computer Engineering, Technology and Science, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Aging involves complex biological processes leading to the decline of living organisms. As population lifespan increases worldwide, the importance of identifying factors underlying healthy aging has become critical. Integration of multi-modal datasets is a powerful approach for the analysis of complex biological systems, with the potential to uncover novel aging biomarkers. In this study, we leveraged publicly available epigenomic, transcriptomic and telomere length data along with histological images from the Genotype-Tissue Expression project to build tissue-specific regression models for age prediction. Using data from two tissues, lung and ovary, we aimed to compare model performance across data modalities, as well as to assess the improvement resulting from integrating multiple data types. Methylation outperformed the other data modalities, with a mean absolute error of 3.36 and 4.36 in the test sets for lung and ovary, respectively. These models achieved lower error rates when compared with established state-of-the-art tissue-agnostic methylation models, emphasizing the importance of a tissue-specific approach. Additionally, this work has shown how the application of Hierarchical Image Pyramid Transformers for feature extraction significantly enhances age modeling using histological images. Finally, we evaluated the benefits of integrating multiple data modalities into a single model. Combining methylation data with other data modalities only marginally improved performance likely due to the limited number of available samples. Combining gene expression with histological features yielded more accurate age predictions compared with the individual performance of these data types. Given these results, this study shows how machine learning applications can be extended to/in multi-modal aging research. Code used is available at https://github.com/zroger49/multi_modal_age_prediction.

O5 - DepMap and GDSC data mining to inform ancestry-driven mechanisms in triple-negative breast cancer

Ricardo J. Pinto[1,2], Ana C. Magalhães[1], Fernanda Rosário[1], Lúcio Lara Santos[3], Emanuel Gonçalves[4], Luísa Pereira[1]

[1]i3S and Ipatimup
[2]ICBAS, Universidade do Porto
[3]IPO-Porto and University Fernando Pessoa
[4]Instituto Superior Técnico - IST and INESC-ID

DepMap and the GDSC are largely contributing to the identification of cancer vulnerabilities by combining genome-scale CRISPR and drug screens with molecular and genetic profiling of large panels of cell lines.

We selected Sub-Saharan African (SSA) and European (EUR) triple-negative breast cancer (TNBC) cell lines from DepMap to perform a differential gene essentiality analysis between both ancestries using the limma and ssGSEA algorithms; correlations between mRNA and mutation levels were also conducted. We identified BLM has an essential gene in SSA TNBC, a result further supported by the SSA signature of enrichment in DNA damage related pathways for the genome-wide essentiality scores. BLM is an important helicase as evidenced by carriers of germline mutations in this gene displaying the Bloom syndrome, with increased risk of tumour development leaving to early death. We confirmed that BLM expression was positively correlated with structural mutations in tumour cells.

Mining (limma linear model) GDSC data to evaluate ancestry influence on TNBC drug sensitivity, we found that SSA TNBC cell lines are also significantly more sensitive to Pevonedistat than the EUR ones. This drug interferes with the neddylation post-translation transformation of proteins, controlling their degradation or action in the cells. The use of this drug is being evaluated in several clinical trials for blood-related tumours (rich in structural mutations) as well as solid tumours (including breast cancer). Our transcriptome analyses of in vitro Pevonedistat assays in TNBC cell lines from both ancestries allowed to establish a link with our DepMap results, as we confirmed that this drug leads to a significant decrease in BLM expression. This link was not previously established and needs further investigation.

The results of this work will be relevant to explore novel mechanisms underlying TNBC development and treatment response for the SSA context. At least in vitro, Pevonedistat is a promising treatment for TNBC in the SSA ancestry, in significant lower doses than the EUR counterparts.

O6 - Guiding computational design for SARS-CoV-2 Spike protein using Molecular Dynamics simulations

Rita I. Teixeira[1], Carolina C. Buga[1-3], Diogo Silva[1], Pedro Moreira[1,4], Ana S. Veiga[2,3], Miguel A. R. B. Castanho[2,3], João B. Vicente[1], Manuel N. Melo[1], Diana Lousa[1], Cláudio M. Soares[1]

[1]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal
[2]Gulbenkian Institute for Molecular Medicine, Lisboa, Portugal
[3]Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
[4]Centro de Engenharia Biológica, Escola de Engenharia da Universidade do Minho, Braga, Portugal

Enveloped viruses pose a serious threat to human health, as recently witnessed with the COVID-19 pandemic. This highlights the urgent need for effective pandemic preparedness strategies and rapid therapeutic development to combat emerging viral outbreaks. One promising approach involves targeting viral entry mechanisms, specifically viral fusion proteins (VFP), to disrupt infection. In previous work, we implemented a computational framework to design de novo tailor-made miniproteins (MP) targeting VFPs, leveraging AI and physics-based methods, with a focus on SARS-CoV-2 as proof-of-concept. Thousands of MPs were designed to target the Spike protein receptor-binding domain (RBD). After rounds of design and computational screening, the promising candidates were experimentally tested, and a subset of these MPs successfully bound to the RBD and blocked SARS-CoV-2 infection. Here, we emphasize the critical role of molecular dynamics (MD) simulations in guiding computational design, by evaluating the efficacy of these designed proteins. We selected the best candidate to perform large-scale MD simulations to investigate its interaction with the fully glycosylated Spike protein, providing a realistic representation of the conditions encountered during viral infection. Due to the complexity of the system, we leveraged access to the Deucalion supercomputer, enabling the simulation setup, performance benchmarking and ultimately, the execution of simulations to capture detailed insights into the stability and conformational adaptability of the MP-Spike complex. By analyzing key interaction metrics from these simulations, we identified the molecular mechanisms underlying binding affinity and structural stability. This work underscores the value of MD simulations in guiding and refining antiviral protein design, offering a deeper understanding of MP behavior in realistic viral environments. Our combined approach of computational protein design and MD simulations provides a robust, adaptable strategy for the rapid development of effective therapeutics, enhancing our preparedness for future viral pandemics.

O7 - Tracing the prevalence and pathogenicity of artifact mutations in ancient Human mitochondrial DNA

Pedro Fernandes[1,2], Bernardo Pinho[1], Bárbara Miguéis[1,2], Ceiridwen J. Edwards[3], João Brochado Almeida[1,2], Martin B. Richards[3], Teresa Rito[1,2], Pedro Soares[1,2]

[1]Centre of Molecular and Environmental Biology (CBMA), University of Minho, Braga, Portugal
[2]Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Braga, Portugal
[3]School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK

Ancient DNA (aDNA) research explores genetic material recovered from ancient specimens, such as archaeological remains. Therefore, aDNA provides a unique record for past demographic phenomena, offering insights on human prehistoric migrations that shaped modern genetic variation. However, its preservation is heavily influenced by taphonomic and diagenetic processes, resulting in fragmentation, scarcity, contamination and miscoding lesions, which pose difficulties in its retrieval. This explains the previous focus of human aDNA research on mitochondrial DNA (mtDNA), a genetic marker with high copy number that offers insights on sex-specific demographic events.

Phylogeography uses this marker to explore the geographical and temporal distribution of human mtDNA lineages. A mutation accumulation rate can be applied to date clades and demographic events, normally using modern mtDNA data. This mutational clock can be estimated by comparing the genetic changes between humans and other species, considering their divergence time based on fossil evidence. Another approach explores variation between ancient and modern samples without the external split time, resulting in higher mutation rate estimations and earlier inferred demographic events, possibly related to artifact mutations accumulated by aDNA.

We reconstructed a mtDNA phylogeny through a newly developed software using 3757 ancient and 63965 modern mtDNA sequences, focusing on major European haplogroups. We defined three groups of mutations: ancient and modern terminal modifications, and pre-terminal variations, aiming to assess if aDNA samples accumulate more artifact mutations in the form of non-synonymous, tRNAs, rRNAs or stop-inducing variations. To predict the impact of protein sequence variants, pathogenicity scores were determined using three software (SIFT, MutPred2 and PolyPhen-2), considering the amino acid changes relatively to the reference. Moreover, we determined the proportion of these mutation types to synonymous ones.

We found higher pathogenicity scores and ratios associated with ancient variations, suggesting a mutation load connected to damage-related artifacts. These results alert for caution when incorporating aDNA for estimations of mutation rates and demographic events of human past.

O8 - DeepTransyt, annotation of transporter proteins using deep learning

Gonçalo Apolinário[1,3], Filipe Liu[3], Christopher S Henry[3] and Oscar Dias[2]

[1]Centre of Biological Engineering, University of Minho, Braga, Portugal
[2]LABBELS – Associate Laboratory, Braga/Guimarães, Portugal
[3]Argonne National Laborary, Chicago, Illinois, USA

Genome-Scale Metabolic (GSM) models are essential tools for studying and engineering cellular metabolism, with transporter annotation representing a critical step in their reconstruction. However, traditional annotation methods, such as homology-based approaches like BLAST, are very dependent on the existence of available data, struggling when it comes to novel examples. To address these challenges, DeepTransyt is introduced as an alternative tool for the identification and annotation of transporters using deep learning.

DeepTransyt leverages state-of-the-art transformer-based models, like ESM2, to generate highly informative protein embeddings. A cornerstone of DeepTransyt is the Transporters Classification DataBase (TCDB), which serves as the primary source of data and annotations for this work. The workflow integrates binary classification to distinguish transporters from non-transporters and multi-class classification to assign transporters to their respective Families and Subfamilies.

The primary challenge in transporter annotation is the accurate identification of each protein's substrate. A major challenge is that many substrate annotations are overly generic, using broad terms like "sugar" or "lipid", making it difficult to determine true substrate specificity. Various approaches were explored to address this issue; however, the most significant limitation remains the availability and quality of data.

The tool was validated on two case studies, *Escherichia coli* strain K-12 MG1655 and *Chlorella* vulgaris strain 221/11P, for the metabolic models: *i*ML1515 and *i*GA1312, respectively. Additionally, a comparison was established against the previous iteration of the transporter annotation tool, TranSyT. For the *i*ML1515 model, DeepTransyt identified 280 additional transporters, 183 of which overlapped with TranSyT's predictions. Among these, 81 had annotations on EcoCyc related to transport functions. For the *C. vulgaris*, DeepTransyt predicted 37 additional transporters, with 22 linked to transport-related annotations on Uniprot.

Beyond matching previous predictions, DeepTransyt expands on existing knowledge by identifying potential transport proteins, particularly in cases where homology-based methods are limited by available data.

DeepTransyt is available as an open-source package at https://github.com/Apolinario8/deeptransyt and on PyPI, enabling seamless integration into broader computational workflows. Ongoing development of a module for the KBase platform aims to increase accessibility for researchers with limited computational expertise.

O9 - Improving Docking Predictions with a ML-Based Scoring Ensemble

João Correia[1], Sofia Ferreira[1], Isabel Rocha[1], Caio S. Souza[1], Diana Lousa[1], Cláudio M. Soares[1]

[1]ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras. www.itqb.unl.pt/pm

Molecular docking has emerged as a critical computational approach for addressing a variety of scientific challenges, such as identifying protein-ligand interactions and discovering and optimizing drug lead compounds. By modelling the favourable conformations of a ligand, typically a small organic molecule, as it binds to a protein's active site, docking offers detailed structural and binding information, including thermodynamic binding affinities. This is achieved through optimization algorithms and scoring functions. While current algorithms manage ligand pose generation effectively, the accuracy of docking predictions remains limited by the imprecision of conventional scoring methods.

To address this limitation, machine learning (ML)-based scoring functions have been developed, offering a promising alternative to traditional scoring approaches. These non-parametric methods bypass the need for predefined functional forms, instead leveraging experimental data to directly infer the contributions of intermolecular interactions to binding affinities. By capturing complex binding interactions that are difficult to model explicitly, ML-based scoring functions provide improved generality and accuracy. Several studies have already demonstrated their ability to enhance docking performance.

In this study, we assess seven ML-based scoring functions aimed at enhancing the accuracy of binding free energy predictions. Our analysis specifically targets the ranking step within a docking-based platform developed for Gene Discovery and Enzyme Engineering (GDEE) (to be published). By focusing on this critical aspect of the workflow, we aim to refine the ranking methodology commonly used in molecular docking platforms. The results demonstrate the substantial potential of these ML-based approaches to improve the precision of binding affinity predictions for protein-ligand complexes. These findings highlight the impact of integrating machine learning into docking workflows, offering a pathway to more reliable and accurate predictions in gene discovery and enzyme engineering applications.

O10 - Computational Engineering of PETase for Improved PET plastics Substrate Affinity and Degradation Efficiency

Alexandra Balola[1], Sofia Ferreira[1], Diana Lousa[1], Caio Souza[1], Cláudio M. Soares[1], Isabel Rocha[1]

[1]Instituto de Tecnologia Química e Biológica António Xavier, Oeiras, Portugal

Polyethylene terephthalate (PET) is a major contributor to global plastic waste pollution, demanding innovative solutions for sustainable recycling. PETase, a PET hydrolase from *Ideonella sakaiensis*, offers a promising enzymatic approach to PET degradation at moderate temperatures. However, its industrial application remains constrained by suboptimal catalytic efficiency and lack of structural robustness.

In this study, we used an in-house protein engineering computational platform to enhance the binding affinity of PETase for the PET plastic polymer. The engineering process was preceded by a detailed analysis of the enzyme's catalytic cavity using a PET analog molecule as a ligand for docking simulations. Subsequently, thousands of mutant variants with the docked model substrate were computationally generated, targeting amino acid residues within the catalytic cavity. These mutants were then evaluated and ranked based on binding free energy calculations and relevant pre-defined metrics, with the goal of identifying and selecting mutations that optimize the binding pose of the PET substrate. Stable and more thermally robust PETase mutants, previously reported in the literature, were used as scaffolds in these engineering rounds, ensuring that the resulting variants would have enhanced structural and functional properties.

The selected top-ranking PETase mutants were expressed and analysed *in vitro* for their ability to degrade PET films. Our results demonstrate that several engineered mutants exhibit enhanced degradation efficiency compared to the benchmark FastPETase mutant at different tested conditions. These findings validate our rational high-throughput engineering approach and pave the way for subsequent engineering rounds.

By improving the performance of PETase, this work contributes to the development of enzyme-based solutions for PET recycling. More specifically, we anticipate that these improved mutants will aid in our goal of developing an *Escherichia coli* strain capable of secreting PETase for *in vivo* PET biodegradation and converting the resulting monomers into biomass and high-value molecules, supporting a circular economy.

O11 - Application of bioinformatics in the characterization of multiple aspects of grapevine diversity

Sara Freitas[1,2,3]; João Nunes[1,2]; David Azevedo-Silva[1,2,3]; Pedro Humberto Castro[1,2], João Tereso[1,3,4,5]; Miguel Carneiro[1,2,3]; Herlander Azevedo[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]MHNC-UP - Museum of Natural History and Science of the University of Porto - PO Herbarium, University of Porto, Praça Gomes Teixeira, 4099-002, Porto, Portugal
[5]Centre for Archaeology, UNIARQ, School of Arts and Humanities, University of Lisbon, Portugal
*E-mail: hazevedo@cibio.up.pt

Grapevine (Vitis vinifera L.) diversity richness results from a complex domestication history over multiple historical periods. Unfortunately, the recent favoring of specific varieties/clones, climate change and the globalization-driven exposure to pathogens, has led to extensive genetic erosion in this widely cultivated and economically significant crop. Fighting this genetic erosion whilst addressing issues of resilience to climate change, yield and other traits, requires a crucial understanding of the genetic basis of grapevine variation. Such studies have been significantly enabled by the use of genomics approaches based on Next Generation Sequencing (NGS). We present an overview of bioinformatic approaches that have been successfully used by the PlantBio group at BIOPOLIS-CIBIO to tackle multiple aspects of grapevine diversity. Research interests include: 1) a clarification of grapevine domestication history and the contribution of Iberian germplasm to the worldwide genetic pool of this crop, using hundreds of modern grapevine genomes and the application of ancient DNA approaches; 2) the characterization of modern clonal diversity on Portuguese varieties and its use on new phenotype-genotype inferences associated with different traits of interest.

Funding: Fundação para a Ciência e Tecnologia (FCT/MCTES) for project GrapeVision (PTDC/BIA-FBT/2389/2020; DOI 10.54499/CEECIND/00399/2017/CP1423/CT0004); FCT/MCTES and POCH/NORTE2020/FSE for support to S.F. (SFRH/BD/120020/2016) and DAS (2022.11910.BD); FCT/MCTES and POPH-QREN/FSE for support to M.C. (CEECINST/00014/2018/CP1512/CT0002).

O12 - Phylogenomics and functional annotation of 530 non-Saccharomyces yeasts from winemaking environments reveals their fermentome and flavorome

Ricardo Franco-Duarte[1,2*], Ticiana Fernandes[1,2], Maria João Sousa[1,2], Paula Sampaio[1,2], Teresa Rito[1,2], Pedro Soares[1,2]

[1]CBMA (Centre of Molecular and Environmental Biology), Department of Biology, University of Minho, Braga, Portugal
[2]Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, 4710-057 Braga, Portugal
*Presenting author e-mail: ricardofilipeduarte@bio.uminho.pt,ricardofrancoduarte@gmail.com

The winemaking industry is currently facing significant challenges driven by climate change and evolving market demands, with far-reaching economic and societal impacts. Non-Saccharomyces yeasts have emerged as promising candidates in addressing these issues while simultaneously enriching the sensory complexity of wine. Building on our previous identification of 293 non-Saccharomyces species linked to winemaking environments, this study expands the scope by analyzing 661 publicly available genomes. By employing a bioinformatics pipeline with stringent quality control checkpoints, we obtained a high-quality dataset of 530 annotated non-Saccharomyces proteomes, belonging to 134 species, and accessible to the research community. After, we performed a comprehensive phylogenomic analysis to explore metabolic networks relevant to fermentation and flavor/aroma enhancement. Our functional annotations revealed distinctive metabolic traits unique to non-Saccharomyces yeasts, underscoring their potential role in enology. In particular, this work pioneers the identification of a non-Saccharomyces 'fermentome', a specific set of six genes uniquely present in fermentative species and absent in non-fermentative ones, and an expanded set of 35 genes constituting the complete fermentome. Moreover, we delineated a 'flavorome' by examining 96 genes across 19 metabolic categories implicated in wine aroma and flavor enrichment. These discoveries provide valuable genomic insights, offering new avenues for innovative winemaking practices and research.

O13 - Genome-wide characterization of plant resistance genes in cork oak (*Quercus suber*)

Luís Miguel Gonçalves[1], Pedro Miguel Barros[1], Maria Margarida Oliveira[1]

[1]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

The Nucleotide-Binding Leucine-Rich Repeat (NLR) genes are essential for plant immune responses, canonically divided in three main sub-families: TNL, CNL and RNL. In this study, we aimed to identify NLR and NLR-like genes and characterize the ones that act in the cross-talk between biotic and abiotic stress. We developed InterNLR, a tool to identify members of the NLR gene family, based on sequence similarity, and used it to characterize the NLR-ome from *Quercus suber* (cork oak). We identified 918 genes, encoding both canonical NLRs and non-canonical NLR-like proteins. Phylogenetic analysis of the NLR-conserved NB-ARC domain revealed the ancestral nature of TNLs and the specialized role of RNLs as "helper" NLRs, with gene duplications occurring predominantly within the RNL clade. Tissue-specific gene expression patterns showed that RNLs had significantly higher expression in xylem in comparison other stem layers, indicating a specialized function in this critical tissue. Under drought stress, seven NLRs were differentially expressed in either phellem, inner bark or xylem. These NLRs shared similarity to genes from other species known to respond to abiotic stress. An example is the orthologue of ADR1 (RNL clade) which responds to both biotic and abiotic stresses in xylem, suggesting a potential crosstalk between these stress responses. Population-level diversity analyses also indicate that RNLs tend to be more evolutionarily conserved than other NLR subfamilies in cork oak. This research enhances understanding of the NLR gene family's evolution, expression, and roles in cork oak, providing insights for future functional studies.

O14 - Eukaryotic Communities and Potential Pathogens in Wastewater Effluents in Ría de Vigo

Ríos-Castro, R1; Ramilo, A1; Rodríguez, H1; Pascual, S1; Abollo, E1

[1]Ecobiomar, Instituto de Investigaciones Marinas, Consejo Superior de Investigaciones Científicas, IIM-CSIC, 36208 Vigo, Spain

Coastal marine ecosystems support diverse eukaryotic microorganisms that contribute to ecological processes and ecosystem services. However, anthropogenic activities, particularly wastewater discharges, introduce potential pathogens and pollutants. This study assessed eukaryotic diversity and pathogen presence in wastewater effluents from three municipalities in the Ría de Vigo (NW Spain), a region of high aquaculture significance.

Water and sediment samples were collected near wastewater treatment plant (WWTP) outfalls in Vigo, Cangas, and Redondela. Three genetic markers were used:the V9 and V4 hypervariable regions of the 18S rRNA gene for eukaryotic analysis and ITS-2 for fungal identification. Bioinformatic analyses were performed using the DADA2 algorithm, and taxonomic classification of ASVs was conducted with SILVA 18S [5] and UNITE databases.

Eukaryotic diversity varied across locations and sequencing regions. Redondela exhibited the highest diversity, particularly in sediments, while Cangas had the lowest across all genetic markers. The V9 region detected greater taxonomic richness, including rare taxa (>1% of total reads), whereas the V4 region exhibited a higher total read count. The V9 region provided better representation of Rhizaria, Discoba (Excavata), and Stramenopila, while V4 exhibited amplification biases favoring Alveolata. ITS-2 analysis revealed a higher prevalence of Ascomycota and Basidiomycota and provided more precise taxonomic resolution than 18S rRNA sequencing. Pathogen abundance was highest in Redondela.

Genetic material from human and zoonotic pathogens was detected, including fungal species causing opportunistic infections (*Rhodotorula mucilaginosa*) and protozoan parasites (*Acanthamoeba* sp., *Cryptosporidium* sp.), underscoring potential public health risks. Aquaculture pathogens were notably present, such as Scuticociliata, responsible for scuticociliatosis in fish, and *Haplosporidium edule*, a bivalve parasite affecting *Cerastoderma edule*. Harmful algal species associated with shellfish poisoning were detected in Vigo sediments, including *Alexandrium minutum* (Paralytic Shellfish Poisoning, PSP) and *Pseudo-nitzschia* spp. (Amnesic Shellfish Poisoning, ASP).

This study highlights the effectiveness of a multi-marker eDNA metabarcoding approach for assessing eukaryotic diversity and pathogen distribution in wastewater-impacted coastal ecosystems. The findings emphasize the importance of sediment analysis and comprehensive monitoring strategies under a "One Health" framework to mitigate environmental and public health risks in anthropogenic influenced marine environments.

O15 - Deciphering the venoms of *Conus* species with transcriptomics

José Morim[1,2], Yihe Zhao[1,2] and Agostinho Antunes[1,2]

[1]CIIMAR/CIMAR—Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos, s/n, 4450-208 Porto, Portugal

[2]Department of Biology, Faculty of Sciences, University of Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

Conus species are a group of marine gastropods known for their beautiful coloured shells and highly powerful and complex venoms, deadly even to humans, which they use for predation and defence. Each *Conus* sea snail's arsenal boasts roughly a couple hundred bioactive well-structured polypeptides known as conotoxins. These toxins are the main component of the venom cocktail, and have been linked with biomedical potential. Despite being studied for decades, the remarkable variability of *Conus* venoms is intriguing and its causes remain mysterious to the present day. The reason behind this state of affairs could be connected with the lack of in-depth studies targeting the broad transcriptome diversity. Such big-scope studies could potentially allow investigations to unveil the trends behind such erratic variabilities, finally deciphering the puzzling enigma.

Following this aim, a simple, straight-forward bioinformatic pipeline was specifically designed to gather all suitable, publicly available *Conus* venom gland transcriptomes and polish, assemble and annotate them. In total, 76 transcriptomes from 20 different *Conus* species were analysed. The transcriptomic repertoire analyses provided several important insights on various aspects such as venom variability and evolution, and supporting signs of ubiquitous symbiotic relationships. Furthermore, this analysis revealed that several gene variations are caused by predatory diets, indicating the feeding habit to be the main force driving venom evolution. Lastly, an extension from the original objective was undertaken with the evaluation of the antimicrobial potential of *Conus* venom peptides, showcasing the reach of broad transcriptome studies. Since this work relied entirely on previously published neglected transcriptomic data, this study highlights the importance and the inherent strength of dedicated bioinformatics.

SCIENTIFIC SUBMISSIONS

POSTER AND SOFTWARE PRESENTATIONS

Nº	Authors	Title
P1	Gonçalo Sousa	WGCNA as a tool to identify key gene expression networks associated with mistranslation and azole resistance in the human fungal pathogen Candida albicans
P2	Maria Inês Gomes	Benchmarking Causal Reasoning Algorithms for enhanced Drug Discovery: Insights from Clarivate's pre-competitive Algorithm Benchmarking Consortium
P3	Tiago Miranda	Towards a Machine Learning Framework for Predicting MicroRNAs involved in Candidiasis
P4	Diana Silva	Comprehensive Molecular Characterization of Anti-EFG1 2'OMe Effects on C. albicans: A Bioinformatics Study
P5	Inês Carvalho	A machine learning model based on geometric morphometric data for Skeletal Malocclusion diagnosis
P6	Filipa Ferreira	The Diagnosis Network App (DiNA): a computational system for the diagnosis of depressive major disorder
P7	Grigore Platon	Expanding the Kinetic-Regulatory Model of Escherichia coli for High-Value Compound Production from Aromatic Amino Acids
P8	Elias Barreira	Predictability of Genomic Evolution at the Molecular Level
P9	Constança Ilunga	Design and Characterization of MHETase Mutants for Improved PET Degradation: A Combined In Silico and In Vitro Study
P10	Raquel Ríos-Castro	High Throughput sequencing for Monitoring Eukaryotic Pathogens in Castellana WWTP: Insights from the AWARE Project
P11	Rafael Vleira	Introducing AptaCom - A centralized aptamer database
P12	Mahmoud Abdallah	Exploring the Genomic Potential of Macroalgae: A Bioinformatics Pipeline for Protein and Gene Discovery
P13	João Guimarães	Evaluating the Viability of BEAST Software for Accurate Phylogenetic Dating
P14	David Henriques	Enhancing the simulation of anaerobic flux distributions in Saccharomyces cerevisiae using genome-scale metabolic models
P15	Sirine Gaieb	Structural Dynamics of H5N1 Hemagglutinin Globular Head: Insights from Molecular Dynamics Simulations
P16	Sofia Torres	Characterization of tissue immunity in metastatic breast cancer by single-cell RNA-sequencing
P17	Raquel Romão	Akna as a novel RNA-binding protein: uncovering its role in immune cell function through iCLIP-seq analysis

SCIENTIFIC SUBMISSIONS

POSTER AND SOFTWARE PRESENTATIONS

Nº	Authors	Title
P18	Sofia Ferreira	Understanding lactoferrin multifunctionality by analyzing the full spectrum of its interacting partners and its evolutionary history
P19	Daniela Holdych	Antiphospholipid Syndrome: Insights into the cellular basis of autoimmunity
P20	Mariana Vasques	Characterizing the Heterogeneity and Differentiation of Murine T Follicular Cells using Single-Cell
P21	Bruna Cruz	Decoding Early Neuronal Responses to Injury: A Bioinformatics Approach to Identify Regeneration- Associated Genes
P22	Ana Paulino	The molecular regulatory mechanisms behind the vegetative-to-reproductive transition in Cynara cardunculus
P23	Diana Lourenço	Plastic Biodegradation by Microalgae: Database Development and Sequence Retrieval
P24	Ana Lima	Comprehensive multi-omics database for highly infectious viruses: a focus on HIV, Ebola and SARS-CoV-2
P25	Raquel Domingues	Structural Bioinformatics insights into the stability of a promising RSV-targeting nanobody
P26	André Bagão	Molecular Dynamics Simulations of Glycan Shielding on Prefusion RSV F Protein: Implications for Epitope Accessibility and Vaccine Design
P27	Benedita Pereira	Developing a Repository for the Storage and Analysis of Novel Computationally Designed Proteins
P28	Raquel Ríos-Castro	Behaviour of Infective Stages (L3) of Anisakis simplex in Water Mass and Fish-waste: Transcriptomic Contribution to understanding the Parasite Life Cycle
P29	Elias Barreira	Developing Genomic Models to Predict Breast Cancer Metastatic Progression

P1 - WGCNA as a tool to identify key gene expression networks associated with mistranslation in the human fungal pathogen *Candida albicans*

Gonçalo Sousa[1], Ana Rita Guimarães[1], Sofia Marques[1], Gabriela Moura[1], Manuel Santos[2], Ana Rita Bezerra[1]

[1]Department of Medical Sciences, Institute of Biomedicine – iBiMED, University of Aveiro, Portugal

[2]Multidisciplinary Institute of Aging, MIA-Portugal, Faculty of Medicine, University of Coimbra, Portugal

Candida albicans is the leading cause of life-threatening invasive infections, with mortality rates approaching 40% despite intensive treatment efforts. *C. albicans* possesses a highly heterogeneous proteome due to its unconventional translation of the CUG codon, which is predominantly decoded as serine (97%) and partially as leucine (3%). A previous study demonstrated that *C. albicans* can tolerate high levels of mistranslation by increasing CUG-Leu misincorporation from 3% (control strain T0) to 20% (strain T1) and 67% (strain T2). This atypical translation mechanism fosters a proteome with exceptional adaptability, driving both phenotypic and genomic diversity and enhancing resistance to various stress conditions. Additionally, experimental evolution of these hypermistranslating strains under non-stress conditions revealed distinct genomic alterations, potentially linked to survival under diverse stress scenarios. In this study, to further explore the relationship between mistranslation and gene expression changes, we employed Weighted Correlation Network Analysis (WGCNA). This approach enabled the identification of clusters (modules) of highly correlated genes in strains evolved under non-stress conditions with varying levels of mistranslation. Our objective was to uncover the evolutionary mechanisms by which leucine misincorporation influences adaptive responses.

This work is supported by the Portuguese Foundation for Science and Technology (FCT) through project FunResist (https://doi.org/10.54499/PTDC/BIA-MIC/1141/2021) and grant (2022.11692.BD). The iBiMED research unit is supported by FCT funds under UIDB/04501/2020 https://doi.org/10.54499/UIDB/04501/2020), and UIDP/04501/2020 (https://doi.org/10.54499/UIDP/04501/2020).

P2 - Benchmarking Causal Reasoning Algorithms for enhanced Drug Discovery: Insights from Clarivate's pre-competitive Algorithm Benchmarking Consortium

Maria Inês Gomes[1, 2], Alexandr Ishkin[1], Cecilia Klein[1], Filippo Ciceri[1]

[1]Clarivate Analytics, Barcelona, Spain
[2]NOVA School of Science and Technology, Lisbon, Portugal

Drug development is a complex process, as it requires integrating information from many different knowledge areas. During the preclinical phase, OMICs measurements are often used to elucidate the phenotypical changes induced by exposure to a therapeutic agent thereby allowing to reconstruct the cascade of biochemical changes responsible for the pharmacologic effect (often termed Mechanism of Action reconstruction). This is a crucial step in the characterization of a therapeutic agent and researchers rely on numerous computational tools for this process. However, the abundance of available solutions and the lack of standardized assessment frameworks create challenges in identifying reliable and efficient tools. To address this challenge, Clarivate leads the pre-competitive subscription-based Algorithm Benchmarking Consortium (ABC), which evaluates computational tools for a variety of use cases.

Within the ABC framework, several algorithms for for identifying key regulators and inferring dysregulated signaling proteins from transcriptomics data (also termed "causal reasoning") were benchmarked, to provide an unbiased assessment of the available methods. The benchmarking initiative evaluated 11 state-of-the-art algorithms using curated datasets, and a benchmark network derived from curated resources such as OmniPath and MetabaseTM. Whilespecific results remain confidential, performance metrics included accuracy in predicting causal regulators, computational scalability, and robustness to input data variations, providing an advanced understanding of algorithm strengths and limitations in this domain.

ABC's framework enables not only evidence-based algorithm selection for diverse tasks, ensuring reliable and reproducible results, but also promotes the development of more effective computational workflows in drug discovery by empowering researchers to select the most optimal tool for their needs. Through this use case, ABC highlights the importance of causal regulation as a tool to provide a deeper understanding of the pharmacological effects of therapeutic agents. Ultimately, this is essential to characterize the safety profile, to identify new indications (drug repurposing), potential combinatorial strategies with synergistic effects and clinically relevant patient subpopulations.

P3 - Towards a Machine Learning Framework for Predicting MicroRNAs involved in Candidiasis

Tiago Miranda[1], Filipa Ferraz[1], Mariana Henriques[2,3], Sónia Silva[2,3], Bruna Gonçalves[2,3]

[1]ALGORITMI Research Centre/LASI, University of Minho, Braga, Portugal
[2]Centre of Biological Engineering, University of Minho, Gualtar Campus, Portugal
[3]LABBELS—Associate Laboratory, Braga/Guimarães, Portugal

Infections caused by *Candida* species present significant clinical challenges, leading to high morbidity and mortality rates. Emerging research suggests that microRNAs (miRNAs) may serve as valuable biomarkers for diagnosing and treating candidiasis, although the field is still developing. The scarcity of reported miRNAs associated with candidiasis stems from the complexities and costs of screening numerous potential biomarkers, making *in silico* prediction essential for identifying candidates for experimental validation. This study aims to enhance the understanding of miRNAs related to candidiasis using bioinformatics approaches, focusing on two primary objectives: analyzing existing experimental data and creating a novel bioinformatics tool using Machine Learning (ML) to predict candidiasis-related miRNAs. For this end, a dataset with 80 experimental miRNAs related to candidiasis was constructed through a systematic literature search and using bioinformatics databases such as *miRBase*, *miRPathDB* and *miRDB*. Notably, 63 of these miRNAs exhibited high cellular activity, with five (*miR-17-3p*, *miR-222-3p*, *miR-133a*, *miR-132-5p*, and *miR-100*) showing ideal characteristics for diagnostic biomarkers, highlighting their potential application. To support further identification of biomarkers, a ML-based tool is being developed, able of predicting miRNAs related to different types of candidiasis, also providing valuable insights as cellular localization and specificity. Our prediction hypothesis posits that if a gene is a target of candidiasis and also a target of a miRNA, that miRNA may be associated to candidiasis, suggesting that a higher number of shared targets correlates with a higher prediction score. As so, ML models, including ensemble and hybrid models, have being tested, and the results showed about 70% of accuracy correlating the predicted candidiasis-associated miRNAs with the ones obtained through literature. Further developments include optimizing the models, as well as experiment others able to classify the candidiasis type, obtaining a comprehensive analysis of these relations. Upon successful completion, this work is poised to make significant contributions to the fields of bioinformatics and biomedicine, ultimately contributing to mitigate the devast consequences of candidiasis.

P4 - Comprehensive Molecular Characterization of Anti-*EFG1* 2'OMe Effects on *C. albicans*: A Bioinformatics Study

Diana Silva[1], Sónia Silva[2,3], Mariana Henriques[2,3], Daniela Araújo[2,4], Bruna Gonçalves[2,3]

[1]Department of Informatics, University of Minho, Braga, Portugal
[2]Centre of Biological Engineering, University of Minho, Braga, Portugal
[3]LABBELS – Associate Laboratory, Braga/Guimarães, Portugal
[4]INIAV, IP - National Institute for Agrarian and Veterinary Research, Vairão, Portugal

Candidiasis is a fungal infection primarily caused by *Candida albicans*. It affects millions of people worldwide and poses significant clinical challenges, resulting in high morbidity and mortality rates, largely due to increasing resistance to conventional antifungal therapies resulting from their widespread use [1]. Innovative therapies, such as antisense oligonucleotides (ASOs), are being explored to address this issue. One promising candidate is anti-*EFG1* 2'OMe, which control *C. albicans* filamentation, through Efg1 inhibition [2,3]. This study aimed to perform a comprehensive characterization of the molecular effects of anti-*EFG1* 2'OMe on *C. albicans* using a bioinformatics approach. For this, various bioinformatics tools were employed to process and analyze proteomic datasets of *C. albicans* cells exposed to the ASO, obtained by LC/MS-MS. Initial data processing involved decoding protein identifiers via Uniprot database and identifying ASO-targeted proteins based on abundance ratios and utilizing the BioinfoGP Venny tool. The results indicated that the presence of the ASO led to the repression and induction of 282 proteins, highlighting significant modulation of the proteome. Further analysis using the FungiFun tool revealed that the most common functions of both the induced and repressed proteins were related to protein binding and metabolism, including carbohydrate metabolism. This suggests that while the ASO may impair critical functions, *C. albicans* cells potentially induce others to adapt to stress. Additionally, using PathoYeastract and *Candida* Genome Database tools, 44 Efg1 targets and 26 proteins with filamentation-related phenotype were identified among proteins repressed by the ASO, demonstrating a direct relationship with Efg1 inhibition. However, 63 Efg1 targets and 31 proteins with filamentation-related phenotypes were induced by the ASO, indicating a potential compensatory response to the ASO's effects. Among these, 9 proteins represent potential targets for the design of new ASOs to combine with anti-*EFG1* 2'OMe, creating an optimal therapeutic cocktail. Overall, these bioinformatics analyses enhanced the understanding of the molecular mechanisms underlying anti-*EFG1* 2'OMe effects, paving the way for more effective therapies against candidiasis.

P5 - A machine learning model based on geometric morphometric data for Skeletal Malocclusion diagnosis

Maria Cristina Faria-Teixeira[1,2], Inês M. N. Carvalho[2], Alexandra Dehesa-Santos[1], Francisco Salvado e Silva[2], Helena Afonso Agostinho[3], Francisco Do Vale[4], António Vaz-Carneiro[2,5], Leixuri De Frutos-Valle[1], Joao C. Guimaraes[2], and Alejandro Iglesias-Linares[1,6]

[1]Complutense University of Madrid, School of Dentistry, 28040 Madrid, Spain
[2]Faculdade de Medicina, Universidade de Lisboa, Portugal 1649-028, Lisboa, Portugal
[3]São João University Hospital of OPorto, 4200 Porto, Portugal
[4]University of Coimbra, 3004 Coimbra, Portugal
[5]Institute for Evidence Based Healthcare, Faculdade de Medicina, Universidade de Lisboa, 1200 Lisbon, Portugal
[6]BIOCRAN (Craniofacial Biology) Research Group, Complutense University, 28040 Madrid,Spain

Skeletal Class III malocclusion (SCIII) is a heterogeneous condition characterized by maxillary retrusion, excessive mandibular growth (prognathism), or a combination of both. This variability hinders accurate diagnosis and classification of orthodontic conditions. To address this, we propose a data-driven approach using twelve annotated cephalometric landmarks from cephalic x-rays to stratify patients and identify clinical subphenotypes associated with SCIII.

Our workflow begins by eliminating biases unrelated to the morphology of interest through generalized Procrustes analysis, which aligns cephalometric coordinates to generate a mean SCIII shape. The resulting Procrustes residuals, which capture deviations in craniofacial morphology relevant to the mean SCIII shape, are then used as features for unsupervised clustering via the K-means algorithm. We have selected six final clusters based on their ability to capture clinically significant patient characteristics.

A key aspect of our approach is evaluating the stability of the identified subgroups. We employed a cross-validation method to assess the robustness of the clusters, framing the problem within a supervised learning framework. Specifically, we quantified cluster stability by examining how well the clustering assignments are preserved under dataset perturbations, yielding cluster membership accuracy. To assess the model's ability to generalize, we evaluated its capacity to predict cluster membership for previously unseen data through the accuracy of cluster assignments from a fitted K-Nearest Neighbors classifier.

Finally, we found that our data-driven clusters were significantly correlated with clinically relevant outcomes, validating the effectiveness of our pipeline in capturing meaningful patient characteristics. This work demonstrates the potential of data-driven methodologies to address the complexity of heterogeneous conditions like SCIII, offering a standardized approach for diagnosis and subphenotype identification. This pipeline serves as a versatile framework, applicable to other challenges involving spatial data and geometric morphology beyond SCIII.

P6 - The Diagnosis Network App (DiNA): a computational system for the diagnosis of depressive major disorder

Ferreira, F[1]., Mendes, P[1]. Nascimento, A[1]. & Gysi, D[2].

[1]University of Maia, Maia, Portugal
[2]Department of Statistics, Federal University of Paraná, Curitiba, Brazil

Major depressive disorder (MDD) is a highly prevalent and complex mental disorder that significantly impacts an individual's health, functioning, and quality of life, and imposes a high level of social and economic burden. Also, MDD etiopathogenesis remains poorly understood. This knowledge gap partly emanates from dependence on different nosographic systems, which rest on latent variable theories. In these theories, symptom covariance is due to specific latent underlying causes, such as genetic or psychological mechanisms. However, classic latent variable models have important shortcomings, because rely on psychometric instruments that are demonstrably inadequate to capture the true complexity of MDD.

Current diagnostic tools for MDD have low content validity, unidimensionality, and poor temporal invariance. Such shortcomings preclude measuring symptom changes during treatment, accounting for the clinical heterogeneity of the disorder, or elucidating patterns of comorbidity.

The network theory of mental disorders provides another promising alternative framework in which symptoms are directly interacting and mental disorders arise from networks of these symptoms. Drawing on this framework, it would enable symptom networks of individuals to be mapped out, with further possibilities for an adaptive, patient-centered orientation of diagnosis and treatment.

Theoretically promising and to give a start in this direction, we presented the Diagnosis Network App (DiNA), a computational system integrating network theory into clinical assessment and diagnosis. DiNA embodies network-based methodologies to analyze symptom interactions and generate diagnostic profiles. To examine its potential, we start by mapping out all disorders and its symptoms, and simulate 5000 patients with MDD, ranging from 3 to 22 symptoms – described by the DSM-5 to characterize the disorder, and several combinations of those symptoms. We measure patients' proximity from the set of simulated symptoms to all other mental disorders and access our ability to correctly identify MDD as the patient's disorder. We find that with 50% of the symptoms (10 symptoms) we reach an AUC of 0.91. Such findings point toward the possibility that DiNA can act as a new diagnostic tool in the treatment of mental health disorders. The method DiNA addresses, most of the deficiencies of the current diagnostic tools in a quick, accurate, detailed manner.

P7 - Expanding the Kinetic-Regulatory Model of *Escherichia coli* for High-Value Compound Production from Aromatic Amino Acids

Grigore Platon[1], Sofia Ferreira[1], Isabel Rocha[1]

[1]ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier

This project aims to integrate the production of various high-value compounds such as L-DOPA, dopamine, and 5-HTP, derived from Aromatic Amino Acids (AAAs) into the in-house kinetic-regulatory model. The initial focus was on incorporating a novel dopamine production pathway into the kinetic model. This was achieved by integrating kinetic parameters of the two key enzymes: tyrosine decarboxylase from *Levilactobacillus brevis* (BRENDA:EC4.1.1.25) and polyphenol oxidase from *Mucuna pruriens var. utilis* (sabiorkR:11664). Using COPASI BASICO Python package, the new dopamine biosynthetic pathway was integrated into the model, Time-course analysis confirmed that the model's topology remained intact, validating the plug-and-play model extension approach. Since the existing kinetic model represents a qualitative continuous culture framework, an additional dynamic model from the literature was selected for batch culture simulations. This alternative model includes improved reactions for biomass production, glucose uptake and TCA cycle regulation compared to previously existing models, but lacks the shikimate and AAA pathways. These pathways are highly regulated, presenting a significant challenge for model integration. To address this, rate laws and parameters from the initial model will be utilized to extend the batch model, followed by performance evaluation Once the two models are reconstructed, optimization algorithms will be employed to maximize dopamine production. Additionally, protein availability will be assessed as a potential limiting factor using *E. coli* GECKO constraints. Upon successfully demonstrating this workflow for dopamine production, this approach will be extended to other AAA-derived compounds, such as 5-HTP and serotonin, to create tailored, product-specific models. This project lays the foundation for scalable and efficient bioproduction of multiple AAA-derived high-value compounds.

P8 - Predictability of Genomic Evolution at the Molecular Level

Elias Barreira[2], Margarida Matos[1,2] & Pedro Simões[1,2]

[1]CE3C - Centre for Ecology, Evolution and Environmental Changes and CHANGE -Global Change and Sustainability Institute, Lisbon, Portugal

[2]Departamento de Biologia Animal, Faculdade de Ciências, University of Lisbon, Lisbon, Portugal; Universidade de Lisboa, Faculdade de Medicina, Lisboa, Portugal

Understanding the predictability of evolution at the genomic level remains a critical challenge in evolutionary biology. While phenotypic evolution has demonstrated some degree of predictability, the underlying genomic mechanisms driving such changes are less understood. This study investigates whether the predictability observed at higher biological organizational levels is mirrored at the genomic scale. To address this, we analyzed populations of *Drosophila subobscura* with contrasting initial histories, collected from distinct latitudes and years, and subjected them to independent laboratory adaptation.

Using the latest genomic tools and a reference genome, we reconstructed existing analysis pipelines and developed custom algorithms—CAR (Coverage Adjusted Rho) and CDD (Convergence Divergence Detection)—to enhance our ability to detect evolutionary patterns. These tools enabled us to differentiate between parallel, convergent, and divergent selection.

The analyses revealed that large genomic regions associated with chromosomal inversions underwent convergent evolution across populations, driven by shared selective pressures. Despite these patterns, no common single nucleotide polymorphisms (SNPs) were observed between populations, suggesting limited predictability at finer genomic scales. Our methods also uncovered peaks of selection and a significant component of parallel evolution, which may arise from codon-level convergence or hitchhiking effects associated with inversions. Functional annotation of candidate regions failed to identify consistent motifs across populations, implying that predictability may be constrained to larger structural features like chromosomal inversions.

This work highlights the complex interplay of drift, selection, and historical contingency in shaping genomic evolution. While predictability appears limited at the SNP level, our findings underscore the importance of structural genomic changes in driving convergent evolution. Future research with higher-resolution genomic data and advanced annotation techniques will be crucial to elucidate the mechanisms underlying these patterns, paving the way for more accurate predictions of evolutionary trajectories at different biological levels of organization.

P9 - Design and Characterization of MHETase Mutants for Improved PET Degradation: A Combined *In Silico* and *In Vitro* Study

Constança Ilunga[1], Alexandra Balola[1], Sofia Ferreira[1], Cláudio Soares[1], Isabel Rocha[1]

[1]Instituto de Tecnologia Química e Biológica António Xavier (ITQB NOVA), Oeiras, Portugal

Polyethylene terephthalate (PET) is a petroleum-based plastic known for its chemical inertness, making it exceptionally resistant and durable. Its widespread use in consumer products has resulted in the production of vast quantities annually. However, improper disposal and inefficient recycling methods have led to its accumulation in ecosystems, posing significant environmental challenges.

Recently, a bacterium capable of biodegrading PET was discovered — *Ideonella sakaiensis* produces two enzymes, a PET hydrolase (PETase) and an MHET hydrolase (MHETase), to break down PET plastic. PETase initiates the process by breaking down PET into smaller intermediates, including mono-(2-hydroxyethyl) terephthalate (MHET), bis-(2-hydroxyethyl) terephthalate (BHET), terephthalic acid (TPA), and ethylene glycol (EG). MHETase then hydrolyzes MHET into the two final monomers TPA and EG, completing the degradation process.

While PETase has been extensively studied and engineered to improve its thermal stability and activity, MHETase remains largely underexplored. This study addresses this gap by designing and characterizing MHETase mutants to improve its efficiency in degrading MHET and potentially other larger PET intermediates generated by PETase. Additionally, it tackles MHETase's tendency to aggregate during expression, which lowers yields and hinders industrial scalability.

An in-house computational pipeline was employed to design thousands of MHETase candidate mutants by modifying amino acid residues on the enzyme's active site. This pipeline ranks mutants based on Autodock Vina metrics, such as binding free energy. The most promising candidates were selected for *in vitro* validation to assess their enzymatic activity. Molecular dynamics simulations will be performed to understand how these mutations influence substrate interactions, providing insights for further optimization.

Beyond activity enhancement, machine learning tools such as AggreProt and MutCompute were used to identify mutations outside the active site that potentially improve MHETase's solubility. These mutations will be soon tested at the laboratory.

Ultimately, optimized MHETase variants can be combined with efficient PETase variants to develop a more efficient PET biodegradation process. In the future, these enzymes could be integrated into PET-degrading *Escherichia coli* strains that secrete PETase and MHETase, enabling the upcycle of the resulting monomers into new plastics or valuable compounds.

P10 - High Throughput sequencing for Monitoring Eukaryotic Pathogens in Castellana WWTP: Insights from the AWARE Project

Ríos-Castro, R.[1]; Di Sansebastiano G.P.[2]; Pascual, S.[1]; Abollo, E.[1]

[1]Ecobiomar, Instituto de Investigaciones Marinas, Consejo Superior de Investigaciones Científicas, IIM-CSIC, 36208 Vigo, Spain
[2]DiSTeBA (Department of Biological and Environmental Sciences and Technologies), University of Salento, Via Lecce-Monteroni, Campus ECOTEKNE, 73100 Lecce, Italy

Environmental DNA (eDNA) metabarcoding is a powerful tool for assessing microbial diversity in wastewater treatment plants (WWTPs), providing valuable insights into treatment efficiency and potential public health risks [1]. This study is part of the AWARE project, which aims to establish Europe's first Recirculating Aquaponic System (RAS) utilizing reclaimed water free from residual contaminants, including eukaryotic pathogens, to safely produce fish and vegetables for human consumption.

To evaluate eukaryote community dynamics and assess the efficiency of the RAS system, high-throughput sequencing of the V9 region of the 18S rRNA gene [2] was conducted on wastewater samples collected from various treatment stages at the Castellana WWTP (Italy) between July 2023 and October 2024. The analysis focused on detecting zoonotic, aquaculture and agricultural pathogens.

During the construction phase of the Castellana WWTP (July 2023–May 2024), diverse pathogenic taxa were identified at different stages of the treatment process, highlighting fluctuations in microbial communities as the system was established. However, as the advanced treatment processes became fully operational, a marked improvement in water quality was observed. By September 2024, the proportion of potential pathogens in the RAS had dropped to 0.5% of the total eukaryotic community, further declining to 0.1% by October 2024. These results underscore the high efficiency of the wastewater purification system in removing pathogenic eukaryotes.

The near-total elimination of pathogens in the final treatment stages validates the effectiveness of the RAS in ensuring water safety. However, as eDNA metabarcoding detects genetic material but does not confirm organism viability, complementary methodologies are recommended for a more comprehensive risk assessment.

P11 - Introducing AptaCom - A centralized aptamer database

Rafael Vieira[3], Diogo Pratas[1], João Carneiro[2], Sérgio F. Sousa[3]

[1]IEETA/LASI, Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro
[2]CBMA, Center of Molecular and Environmental Biology, University of Minho
[3]LAQV/REQUIMTE - BioSIM, Department of Biomedicine, Faculty of Medicine, University of Porto

Aptamers are small single stranded nucleotides, capable of binding to an extensive array of targets, with high affinity and specificity1 - similarly to antibodies. Unlike antibodies however, aptamers trigger a lower immune response, and display higher degree of chemical malleability. Since the inception of the field of aptamer research, there has been a large gap in how aptamer information is shared and structured. Information is often incomplete and scattered throughout various different platforms2. AptaCom was developed to tackle this issue by combining more than 3500 aptamer entries from 6 of the largest aptamer databases into one single platform. The end goal is to facilitate access and cross- reference of aptamer information in preliminary stages of research.

AptaCom is now online (https://jc-biotechaiteam.com/AptaCom/) and supports submission of new aptamer entries by external users, which promotes further expansion of the aptamer library.

This work received financial support from the PT national funds (FCT/MECI, Fundação para a Ciência e Tecnologia and Ministério da Educação, Ciência e Inovação) through the project UID/50006 -Laboratório Associado para a Química Verde - Tecnologias e Processos Limpos.
This work received financial support from the European Union (FEDER funds POCI/01/0145/FEDER/007265) and national funds from Fundação para a Ciência e a Tecnologia and Ministério da Educação para a Ciência (FCT/MEC - SFRH/BD/137844/2018, IF01310/2013, IF/00052/2014, and PTDC/QUI-QFl/31689/2017), under the partnership agreement PT2020-UID/QUI/50006/2013, and co-financed by the ERD under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007728).

P12 - Exploring the Genomic Potential of Macroalgae: A Bioinformatics Pipeline for Protein and Gene Discovery

Mahmoud Bassyouni[1,2],Mena Khalaf[1,2],Mohamed Emam[1,2], Yihe Zhao[1,2] Agostinho Antunes[1,2]

[1]CIIMAR/CIMAR, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos, s/n, 4450–208 Porto, Portugal

[2]Department of Biology, Faculty of Sciences, University of Porto, Rua do Campo Alegre, 4169-007, Porto, Portugal

In the last two decades, scientists have collected over 80 terabytes of data on 10,000 species of macroalgae, highlighting their crucial role in marine ecosystems. Seaweeds support biodiversity, mitigate climate change by absorbing CO_2, and offer high-yield, sustainable biomass. Despite their potential, many species remain underutilized due to limited genetic and functional insights. Traditional resource management lacks comprehensive data, risking over-exploitation and supply chain instability. To address this, we developed a bioinformatics pipeline for protein and gene discovery in algae, integrating MMseqs2 [1] for clustering and HMMScan [2] for functional annotation. Our approach enables comparative genomics, protein domain identification, and evolutionary insights, aiding in the sustainable utilization of macroalgae. Our analysis, conducted as part of the iCulture project, revealed that Ankyrin repeat (Ank) proteins and kinase-related genes are highly prevalent, emphasizing their importance in cellular regulation and adaptation. These findings provide a foundation for biotechnological and aquaculture applications. This pipeline advances algal genomics, unlocking new opportunities for sustainable bioresource management and industrial applications. The pipeline is available on GitHub: https://github.com/Mohamedema/ICulture.

P13 - Evaluating the Viability of BEAST Software for Accurate Phylogenetic Dating

João Guimarães[1,2],Teresa Rito[1,2], Pedro Soares[1,2]

[1]CBMA-Centre of Molecular and Environmental Biology, Department of Biology, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal
[2]IB-S-Institute of Science and Innovation for Bio-Sustainability, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal

Phylogenetic dating is essential for reconstructing evolutionary timelines, providing critical insights into the origins and diversification of species. Accurate estimation of divergence times is key to understanding evolutionary processes and correlating biological events with geological history. BEAST (Bayesian Evolutionary Analysis by Sampling Trees) is one of the most widely used tools for Bayesian phylogenetic analysis, offering a robust framework for integrating molecular data, fossil records, and evolutionary models. In this study, we evaluate the performance of BEAST as a tool for dating phylogenetic trees by analysing a dataset comprising 471 individuals. To enhance the precision of divergence time estimates, we incorporate calibration points derived from ancient samples. The study focuses on assessing BEAST's accuracy, reliability, and computational efficiency when applied to this dataset. By testing BEAST with carefully calibrated temporal data, we aim to elucidate its strengths and limitations, contributing to its application in evolutionary biology and phylogenetic research.

P14 - Enhancing the simulation of anaerobic flux distributions in *Saccharomyces cerevisiae* using genome-scale metabolic models

Gustav Sjöberg[1], Eduard J Kerkhoven[2,*], Melek Dhemaied[1], Antonius J.A. van Maris[1], David Henriques[3,*]

[1]Department of Industrial Biotechnology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm, Sweden
[2]Department of Life Sciences, SciLifeLabChalmers University of Technology Gothenburg SE412 96 Sweden
[3]Aquatic Biotechnology group, IIM-CSIC, Vigo, Spain

Genome-scale metabolic models are widely employed to predict microbial phenotypes and estimate intracellular fluxes in a variety of biotechnological applications. *Saccharomyces cerevisiae*, one of the most extensively studied organisms of biotechnological relevance, benefits from a community-based consensus metabolic reconstruction. Historically, the validation and refinement of yeast metabolic networks have predominantly relied on data obtained under aerobic conditions. However, the application of metabolic networks to yeast fermentation processes has grown significantly in recent years.

In this study, we compiled physiological, omics, and fluxomics data obtained under anaerobic growth conditions and compared them with simulations performed using the pFBA algorithm applied to the Yeast-GEM 9 model. Based on this effort, we propose several modifications to reaction directionality, stoichiometry, and metabolite charges, which enhance the overall agreement of simulations with the known physiology of *S. cerevisiae*. Decisions regarding model curation were guided by experimental evidence. As a result, we significantly improved the accuracy of flux predictions in the genome-scale metabolic model when compared to three independent fluxomic studies. Additionally, growth predictions across various conditions were also improved. Notably, we identified the importance of an oxaloacetate-malate exchange reaction for describing anaplerosis in a manner consistent with published data. While this reaction is not directly compatible with kinetic data from purified transporter studies in liposomes, introducing an intermediate bidirectional transport of oxaloacetate and malate exchanged by sulfate resolved this inconsistency. These results will be incorporated into the upcoming Yeast-GEM 9.2.0 release.

P15 - Structural Dynamics of H5N1 Hemagglutinin Globular Head: Insights from Molecular Dynamics Simulations

Sirine Gaieb[1], Raquel Domingues[1], André Bagão[1], Diana Lousa[1], Cláudio M.Soares[1]

[1]ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal

Highly pathogenic avian influenza virus (HPAIV) H5N1 poses a significant threat due to its increasing ability to infect mammals. The latest strain has caused an outbreak in dairy cattle across the United States, that has also spread to poultry and humans working in close contact with infected animals. While no human-to-human transmission has been confirmed, the virus' continued spread and adaptation to new mammalian hosts raise concerns about its potential to acquire this capability, potentially leading to human-to-human transmission.

This study focuses on the hemagglutinin (HA) protein, which is crucial for host cell interaction, since it binds to the host cell sialic acid receptors and is also responsible for fusing the viral envelope with the host cell membrane, making it a key target for viral inhibition. To advance our understanding and inform potential inhibitor designs, μs-long molecular dynamics simulations were conducted to assess the structural characteristics and dynamic behavior of the HA globular head, which contains the receptor binding site (RBS) and is, thus, a very relevant target.

The simulations were performed on both the predicted structure of the A/cattle/Texas/2024 strain and the X-ray structure of the whole protein truncated to obtain the globular head (PDB ID: 9DIP). Three replicates were run for each structure to ensure robustness in the findings. The data extracted allows for a detailed examination of conformational changes and stability in regions critical for receptor binding and may provide valuable insights into HA-receptor interactions and specificity.

By enhancing our understanding of H5N1 HA dynamics, a foundation is laid for future targeted interventions to contain the pandemic potential of this evolving virus.

P16 - Characterization of tissue immunity in metastatic breast cancer by single-cell RNA-sequencing

Sofia Torres[1], Ana Luísa Correia[2], João C. Guimarães[1]

[1]Faculdade de Medicina, Universidade de Lisboa, Portugal (FMUL)
[2]Champalimaud Research, Champalimaud Foundation, Lisbon, Portugal

In metastatic breast cancer, the timing and mechanisms of metastasis remain poorly understood. Disseminated tumor cells (DTCs) can establish themselves in distant organs and remain dormant for years before progressing to overt metastases. Previous research has demonstrated that the microenvironment significantly influences this progression: tumor- free livers exhibited increased abundances of NK cells, whereas metastatic livers showed a higher prevalence of myeloid-derived cells. Additionally, it was also shown that this balance is disrupted, and metastases are initiated, by fibrotic injury driven by activated hepatic stellate cells.

Building on this foundation, our study utilizes single-cell RNA-seq data to characterize the tumor microenvironment of various organs targeted by disseminated breast cancer cells, alongside the molecular mechanisms of the different cell types found there. Our dataset includes single-cell data of liver and lung tissues across four metastasis development stages: normal, dormant, micrometastatic, and metastatic.

Cells annotations were obtained with CellTypist, leveraging a pre-annotated dataset of healthy liver samples belonging to the Liver Cell Atlas. A preliminary analysis using scCoda, a tool for compositional analysis of cell types across conditions, reveals a significant increase in monocytes and monocyte-derived cells, alongside a confirmed decrease in NK cells, during the micrometastatic and metastatic stages. These insights shed light on the role of the immune microenvironment in metastasis development and may inform future immunotherapy strategies aimed at preventing metastatic progression.

P17 - Akna as a novel RNA-binding protein: uncovering its role in immune cell function through iCLIP-seq analysis

Raquel A. Romão[1], Lisa Kifinger[2], Vigo Heissmeyer[2], João C. Guimarães[1]

[1]Universidade de Lisboa, Faculdade de Medicina, Lisboa, Portugal
[2]Ludwig Maximilian University of Munich, Germany

RNA-binding proteins (RBPs) are emerging as key regulators of immune cell function. Here, we reveal that Akna, a centrosomal protein known for regulating microtubule organization, is a novel RBP potentially involved in immune regulation.

Experimental observations revealed that Akna knockout in T and B cells leads to reduced T follicular helper (Tfh) and germinal center B cell populations. Additionally, Akna was found to localize at the interface of B cell-T cell synapses during immune interactions, suggesting its potential role in enabling a full germinal center response.

We sought to elucidate Akna's molecular mechanisms in immune cell function, specifically focusing on its role as an RBP in B cells.

We developed a comprehensive data analysis workflow for reliable detection of mRNA target binding sites from Akna iCLIP-seq data, covering all steps from quality control of sequencing reads to peak calling and quantification of binding enrichment.

Our results revealed that Akna binds to mRNAs encoding key genes involved in cytoplasmic translation, synapse function, and immune system processes. Targets involved in cytoplasmic translation were predominantly ribosomal protein genes, suggesting that Akna may play a role in regulating cellular translation. Additionally, Akna showed binding to mRNAs encoding genes crucial for B cell activation and immune responses, including CD20 and CD79a. This integrated bioinformatics and biological approach highlights the putative role of Akna in modulating immune cell response through its RNA binding function, opening new avenues for understanding the complex regulation of these processes.

P18 - Understanding lactoferrin multifunctionality by analyzing the full spectrum of its interacting partners and its evolutionary history

Sofia Ferreira[1], Joana P. Guedes[1], Juliana F. Rocha[2], Sérgio F. Sousa[2], Lígia R. Rodrigues[3,4], Manuela Côrte-Real[1], Pedro Soares[1,5], Cátia Santos- Pereira[3,4]

[1] CBMA - Centre of Molecular and Environmental Biology, University of Minho, Braga, Portugal
[2] LAQV/REQUIMTE, BioSIM, Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal
[3] CEB - Centre of Biological Engineering, Universidade do Minho, Braga, Portugal
[4] LABBELS - Associate Laboratory, Braga/Guimarães, Portugal
[5] IB-S - Institute of Science and Innovation for Bio Sustainability, University of Minho, Braga, Portugal

Lactoferrin (LF) is a multifunctional protein involved in critical biological processes such as immune response, antimicrobial activity, and iron regulation [1]. Its diverse roles include inhibiting inflammation, infections, and carcinogenesis [2], as well as modulating cellular processes by targeting proton-pumping ATPases such as V-ATPases in cancer cells [3], F-ATPases in bacteria [4], and P-ATPases in yeast [5]. Despite its inclusion in commercial products like toothpastes, infant formulas, and cosmetics, and its broad therapeutic potential [1;6], the mechanisms underlying LF's interactions with binding partners remain underexplored. Moreover, its evolutionary trajectory across species is poorly understood, leaving significant gaps in our knowledge of its biological diversity and potential applications. This study aims to bridge these gaps by conducting a comprehensive review of LF binding partners and their functional annotations, alongside an evolutionary analysis of LF among diverse species. By integrating bioinformatics and computational tools, we seek to elucidate LF's interactions with binding partners, uncover evolutionary relationships with homologous proteins and identify conserved motifs critical for LF-ATPase interactions. Our methodology includes the compilation and annotation of LF binding partners, functional characterization and interaction network analysis, and phylogenetic reconstruction of LF homologs using advanced computational approaches. This integrative approach leverages databases like STRING, EggNOG, and KEGG, alongside phylogenetic tools such as MEGA and BEAST. Overall, this study provides a comprehensive view of LF's functional landscape, highlighting its role as a critical regulator of essential biological processes. Preliminary data in LF evolutionary history also contributes to enhance our understanding of LF's multifaceted biological roles.

P19 - Understanding lactoferrin multifunctionality by analyzing the full spectrum of its interacting partners and its evolutionary history

Daniela Holdych[1,2], Miguel M. Santos[1], Pedro Gaspar[1,3], Filipa Ribeiro[1], Rui do Amaral Vieira[1], Beatriz Filipe[1], Válter R. Fonseca[1], Vasco C. Romão[1,3], Luís Graca[1]

[1]Gulbenkian Institute for Molecular Medicine, Centro Académico de Medicina de Lisboa, Universidade de Lisboa, Lisbon, Portugal
[2]NOVA School of Science and Technology, Universidade NOVA de Lisboa, Largo da Torre, Caparica, Portugal
[3]Hospital Santa Maria, ULS Santa Maria, Lisbon, Portugal

The study of autoimmunity, characterized by dysregulated B and T cells culminating in the production of autoantibodies, remains challenging due to the diverse underlying mechanisms. Antiphospholipid Syndrome (APS) is a rare autoimmune disease characterized by the presence of antiphospholipid autoantibodies (aPL), leading to thrombotic events and pregnancy morbidity. While aPL levels may fluctuate or even become undetectable over time, the mechanisms of immune dysregulation during aPL changes remain poorly understood. Since APS mainly targets the coagulation system, patients are treated with anticoagulants, providing the ideal opportunity to explore this immune dysregulation.

APS is mediated by autoantibodies produced by B cells under T cell regulation. Preliminary data revealed quantitative differences in CD4+ T cell subpopulations: aPL-positive patients exhibited an expansion of defined subpopulations of T cells, while aPL-negativized patients have an immune phenotype similar to healthy controls. B cells showed no differences across patients. Using scRNA-seq we aim to characterize the molecular changes in these cells during disease progression and aPL fluctuations.

Using computational biology approaches, we identified well defined sub-populations of B cells, such as naïve, switched and unswitched memory B cells. Gene ontology analysis of the differentially expressed genes of naïve and unswitched memory B cells from patients revealed enrichment in pathways associated with cell activation, differentiation and antigen reception/processing. This result suggests a more activated profile in APS patients compared to healthy controls, while switched memory B cells had a similar transcription profile between groups.

Regarding CD4+ T cells, we could confidently identify regulatory and cytotoxic T cells. However, their homogeneous transcriptional profile in circulation combined with low RNA levels of well characterised marker genes make it hard to identify other subtypes relevant for disease pathogenesis. To overcome this, we are currently performing CITE-seq of T cells to refine our classification and enable deeper sub-populations analysis, including trajectory analysis.

This study uses a multiomics approach to uncover the immune cell dysregulation during disease progression in APS. We expect to provide new insights on the onset of autoimmunity and potentially opening new therapeutic options for these patients.

P20 - Characterizing the Heterogeneity and Differentiation of Murine T Follicular Cells using Single-Cell

Mariana Vasques[1,2], Filipa Ribeiro[1], Tomás Gomes[1], Saumya Kumar[1,3], Afonso Basto[1,4], Jimena Tosello[5], Eliane Piaggio[5], Válter Fonseca[1,6], Luís Graça[1,6]

[1]Gulbenkian Institute for Molecular Medicine, Lisboa, Portugal
[2]Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa, Lisboa, Portugal
[3]Centre for Individualised Infection Medicine (CiiM), a joint venture between the Helmholtz Centre for Infection Research (HZI) and Hannover Medical School (MHH), Hannover, Germany
[4]CIISA Centro de Investigação Interdisciplinar em Sanidade Animal, Faculdade de Medicina Veterinária, Universidade de Lisboa, Lisboa, Portugal
[5]Institut Curie, PSL Research University, Paris, France. [6]Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal

The formation of germinal centres (GC), within secondary lymph organs, is key to ensure that B cells differentiate and proliferate to mount effective responses against pathogens and generate lasting immunity through processes like vaccination. Tight regulation of GCs is critical to ensure the selection of B cells with high affinity antibodies, while avoiding the emergence of autoreactive B cell clones. Central to this regulation are T follicular helper (Tfh) and T follicular regulatory (Tfr) cells, therefore understanding the biology of these specialized T follicular cells is paramount. However, studying them, particularly Tfr cells, remains challenging due to their close resemblance, the absence of distinct markers for their identification and the lack of insight into their distinct developmental trajectories.

To address this problem, we isolated three cell populations ($Foxp3^{-}Cxcr5^{+}$ Tfh cells, $Foxp3^{+}Cxcr5^{+}$ Tfr cells and $Foxp3^{+}Cxcr5^{-}$ Treg cells), collected from mice popliteal lymph nodes, at day 0, 6 and 11 after immunization, and performed single-cell transcriptomics, using scRNA-seq. Additionally, we also generated data on TCR diversity, using scVDJ-seq. We then used state of the art computational approaches, to compare gene expression levels at a single-cell resolution and delve into the heterogeneity of Tfh and Tfr cells. This analysis revealed distinct localization and activation states, which allowed us to differentiate and characterize the transcriptomic profile of these specialized cell populations along their developmental trajectories during a humoral immune response.

Ultimately, through this research we can advance our understanding of the biology of Tfh and Tfr cells in the GCs and may unveil key regulators and signalling pathways involved in their development and function. In turn, this knowledge can potentially lead to the identification of new therapeutic targets for effectively modulating immune responses, notably contributing to the development of improved vaccines and treatments for immune-related disorders, such as autoimmunity.

P21 - Characterizing the Heterogeneity and Differentiation of Murine T Follicular Cells using Single-Cell

Bruna C. da Cruz[1], Patrícia D. Correia[1], Ana Rita Guimarães[1], Gabriela Moura[1], Frank Bosse[2], Sandra I. Vieira[1]

[1]Institute of Biomedicine (iBiMED), Department of Medical Sciences, University of Aveiro, Portugal

[2]University-Hospital Düsseldorf, Department of Neurology, Heinrich-Heine-University, Germany

Spinal Cord Injury (SCI) is a debilitating condition with no effective treatments. Future therapies may depend on regeneration-associated genes (RAGs), but the core upstream RAGs capable of driving full regenerative responses remain unknown. To address this, we performed a comparative transcriptomic analysis of sensory and motor neurons following nonregenerative SCI and regenerative peripheral nerve injury (PNI).

Injured neurons were collected via laser capture microdissection 24 hours post-injury, and their RNA profiles subjected to high-throughput sequencing. Differentially expressed genes (DEGs), relatively to sham controls, were identified using an adjusted p-value threshold of <0.05. Principal Component Analysis (PCA) revealed distinct clustering of samples by both injury type and neuronal subtype, with stronger transcriptional responses in regenerating neurons (post-PNI) compared to non-regenerating neurons (post-SCI).

Bioinformatics analyses focused on DEGs, particularly the 104 genes uniquely regulated in the regenerating condition. Protein-protein interaction (PPI) networks were constructed using STRING and visualized in Cytoscape, identifying highly connected hubs potentially driving regeneration. Gene Ontology (GO) enrichment and KEGG pathway analysis were performed in ShinyGO, revealing significant overrepresentation of processes such as steroid biosynthesis and immune responses among PNI-exclusive DEGs. Conversely, genes shared between PNI and SCI were enriched for stress response, developmental processes, and signalling regulation, indicating a conserved injury response across conditions.

These findings underscore the utility of integrative bioinformatics approaches in unravelling the molecular basis of neuronal regeneration. By identifying candidate RAGs, this study lays the groundwork for future functional experiments to validate their therapeutic relevance in SCI.

P23 - Plastic Biodegradation by Microalgae: Database Development and Sequence Retrieval

Lourenço, D[1,2], Pratas, D[3,4,5], Sousa, SF[6], Carneiro, J[7,8]

[1]FCUP - Department of Computer Science, Faculty of Sciences, University of Porto, Porto, Portugal
[2]CIIMAR - Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Porto, Portugal
[3]IEETA/LASI - Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Aveiro, Portugal
[4]DETI - Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal
[5]DV - Department of Virology, University of Helsinki, Helsinki, Finland
[6]LAQV/REQUIMTE BioSIM - Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal
[7]CBMA - Centre of Molecular and Environmental Biology, Department of Biology, University of Minho, Campus Gualtar, 4710-057 Braga, Portugal
[8]IB-S - Institute of Science and Innovation for Bio-Sustainability, University of Minho, 4710-057 Braga, Portugal

Plastic pollution is a global environmental challenge, and biodegradation offers a promising solution to fix this problem. While bacterial degradation of plastics has been widely studied, research into the role of microalgae in plastic biodegradation is still very underdeveloped. Microalgae have demonstrated potential in the biodegradation process through the production of plastic-degrading enzymes, including ligninolytic enzymes and exopolysaccharides (EPS). Bioinformatic tools provide a powerful approach to better understand these biodegradation mechanisms.

This study utilizes bioinformatic tools to investigate plastic-degrading enzymes from microalgae, focusing on their interaction with plastic polymers and the mechanisms involved in plastic biodegradation.

A comprehensive literary review identified microalgae species that have shown plastic-degrading potential. Information on species, degraded plastic types, enzymes, and degradation efficiency was compiled into a database to facilitate future analysis. A python script for sequence retrieval was developed using the NCBI API to automate the process of obtaining relevant sequences.

In the literary analysis several microalgae species were identified with plastic-degrading properties, this included more than 40 microalgae, cyanobacteria and diatoms species, 9 naturally produced enzymes and 2 genetically engineered enzymes capable of degrading common plastics such as, polyethylene (PE), low density polyethylene (LDPE) and polyethylene terephthalate (PET). In the NCBI database, only one sequence was available (esterase from *Dunaliella salina)*, while two other sequences (dbph-1 and dbph-2) were found in the respective article.

This bioinformatic approach provides valuable insights into the potential of microalgae for plastic remediation. The developed database and ongoing homology modelling will aid in identifying key enzymes and their interaction with microplastics. Future work will focus on refining enzyme models, conducting molecular docking to examine enzyme-polymer interactions and exploring additional enzymes with enhanced plastic degradation capabilities.

P24 - Comprehensive multi-omics database for highly infectious viruses: a focus on HIV, Ebola and SARS-CoV-2

Lima, AS[1,2], Carneiro, J[,3,4], Sousa, SF[5], Sá, VJ[1,9], Pratas, D.[6,7,8]

[1]LabRP/CIR, School of Health, Polytechnic of Porto, Porto, Portugal
[2]CIIMAR – Interdisciplinary Centre of Marine and Environmental Research of the University of Porto, University of Porto, Matosinhos, Portugal
[3]Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus Gualtar, Braga, Portugal
[4]Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Braga, Portugal
[5]LAQV/REQUIMTE - BioSIM, Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal
[6]IEETA/LASI, Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro
[7]DETI, Department of Electronics, Telecommunications and Informatics, University of Aveiro
[8]DV, Department of Virology, University of Helsinki; (9) ALGORITMI/LASI, University of Minho, Braga, Portugal

Highly infectious viruses, such as HIV, Ebola, and SARS-CoV-2, continue to pose significant threats to global health, underlining the urgent need for new therapeutic approaches. Recent advancements in genomic and proteomic databases, along with 3D homology modelling, have enabled detailed simulations of virus-host interactions, providing insights into infection mechanisms and helping identify potential therapeutic targets.

This study aimed to create a unified database of highly infectious viruses and to conduct structural analyses of key viral proteins to explore potential therapeutic strategies. Structural information for proteins involved in the infection process was sourced from the Protein Data Bank and UniProt, while 3D homology models for significant viral variants were generated using AlphaFold. The quality of these models was assessed using AlphaFold-specific metrics, including pLDDT (per-residue confidence scores) and PAE (predicted aligned error), ensuring the structural reliability for further analyses. Identification of the most relevant structural changes was done through alanine scanning in Schrödinger's Biologic Suite, with posterior studies on how those changes affected the infection process. Simulations of virus-host interactions were conducted using docking algorithms, namely HADDOCK, with visualizations performed using PyMOL.

This integrative approach highlights high-confidence therapeutic targets and provides a foundation for developing novel effective treatments for highly infectious diseases.

This research was supported by Portuguese national funds through the Foundation for Science and Technology (FCT) within the scope of UIDB/04423/2020 (CIIMAR) and UIDP/04423/2020 (CIIMAR), 10.54499/LA/P/0008/2020,10.54499/UIDP/50006/2020, UIDB/04050/2020 (UMinho CBMA - https://doi.org/10.54499/UIDB/04050/2020), 10.54499/UIDB/50006/2020 (LAVQ), UIDB/00319/2020 (ALGORITMI/LASI) and UIDB/00127/2020 (IEETA/LASI - doi.org/10.54499/UIDB/00127/2020). JC acknowledges the FCT funding for his research contract established under the transitional rule of Decree Law 57/2016, amended by Law 57/2017. D.P. is funded by national funds through FCT–Fundação para a Ciência e a Tecnologia, I.P., under the Scientific Employment Stimulus—Institutional Call—reference CEECINST/00026/2018.

P25 - Structural Bioinformatics insights into the stability of a promising RSV-targeting nanobody

Raquel Domingues[1], André Bagão[1], Sirine Gaieb[1], Iebe Rossey[2], Xavier Saelens[2], Cláudio M. Soares[1], Diana Lousa[1]

[1]ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal
[2]Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium Center for Medical Biotechnology, VIB, Ghent, Belgium

Respiratory syncytial virus (RSV) is a major cause of respiratory infections, particularly in children under five and older adults. In 2019, RSV caused 33 million infections and 101,400 deaths in children. Treatment of patients primarily focuses on symptom relief, with approaches varying by age and severity. The fusion protein (F), essential for RSV's ability to infect host cells, is a key target for therapeutic and preventive measures. In 2023, two vaccines targeting the F protein were approved for older adults: Arexvy and Abrysvo. Later in May of 2024, mRESVIA, an mRNA vaccine with similar efficacy to Abrysvo, was also approved. For infants, two monoclonal antibody therapies targeting the fusion (F) protein are currently approved for RSV prevention (palivizumab and nirsevimab).

Nanobodies (VHHs), with their unique antigen-binding properties and ease of production in microbial expression systems, are emerging as promising tools for combating RSV. For example, nanobody ALX-0171 has been evaluated in a phase II clinical trial in children that were hospitalized with RSV lower respiratory tract infection. Another VHH, named F-VHH-4, potently neutralizes RSV A and B strains by targeting a prefusion F-specific quaternary epitope.
However, the biopharmaceutical potential of F-VHH-4 remains to be explored.

In this work, we focus on studying the structural stability of F-VHH-4 through molecular dynamics (MD) simulations. We analyzed its stability through RMSD, the flexibility using RMSF and the secondary structure using DSSP. Our simulations revealed that F-VHH-4 is a highly stable structure. Our studies will be important for the improvement of nanobody's biopharmaceutical potential.

P26 - Molecular Dynamics Simulations of Glycan Shielding on Prefusion RSV F Protein: Implications for Epitope Accessibility and Vaccine Design

André Bagão[1], Raquel Domingues[1], Rita I. Teixeira[1], Cláudio M. Soares[1], João B. Vicente[1], Diana Lousa[1]

[1]ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal

Respiratory Syncytial Virus (RSV) is a primary contributor to acute respiratory infections, with intense severity in infants and the elderly. The virus's Fusion (F) glycoprotein plays a crucial role in viral infection and serves as a key target for neutralizing antibodies. As a class I fusion protein, RSV F facilitates the merging of viral and host cell membranes, enabling viral entry. The prefusion form of RSV F exists in a metastable state, undergoing conformational changes during the fusion process. This investigation explores the glycan shielding of established epitopes on the prefusion F protein to identify prospective targets for protein design and discover of new potential epitopes.

For the first time, molecular dynamics (MD) simulations of the prefusion trimeric F protein structure incorporating glycans were performed, offering innovative perspectives into the protein's surface accessibility. The findings indicate that while glycans provide substantial shielding to much of the protein surface, they afford reduced protection to two specific epitopes unique to the prefusion state. These results could significantly impact RSV vaccine and therapeutic development. The lower glycan shielding of these two sites implies enhanced antibody accessibility, making them promising candidates for protein design strategies. Moreover, the inclusion of glycans in MD simulations of the RSV F protein in this study yields a more precise representation of the protein's dynamics in its natural environment.

P27 - Developing a Repository for the Storage and Analysis of Novel Computationally Designed Proteins

Benedita Pereira[1]Pedro Moreira[1,2], João B. Vicente[1], Cláudio M. Soares[1], Diana Lousa[1], Manuel N. Melo[1]

[1]ITQB NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal
[2]Centro de Engenharia Biológica, Escola de Engenharia da Universidade do Minho, Braga, Portugal

The design of novel proteins, either through artificial intelligence (AI)-driven methods or physical-based methods, holds great promise for advancing various scientific fields. A critical challenge in protein design lies in efficiently storing, managing, and analysing the vast array of data generated throughout the design process. In this work, we present the development of a specialised database designed to store, categorise, and analyse data related to designed proteins. While existing databases such as ProtaBank, UniProt, and the Protein Data Bank (PDB)3 offer valuable resources for storing and analysing protein-related data, they often lack integration between computational design strategies and experimental validation workflows. This database serves as a comprehensive repository that integrates information on both computational design strategies and experimental validation, facilitating access to key information across several stages of the development of novel proteins: from the initial AI-driven or physical-based design phase to the experimental production, purification, and characterisation of the designed proteins. It includes structural data, design parameters, production protocols, yields, and performance metrics. By providing a centralised platform for storing and querying this data, the database facilitates the identification of successful design strategies, the optimisation of production protocols, and the reproducibility in protein engineering, providing a valuable resource for researchers and practitioners in the field of protein design. We will discuss the underlying data models, the challenges in data integration, and the potential applications of this tool in advancing the understanding of designed proteins. Ultimately, this database aims to aid the process of computationally designing novel proteins, improve reproducibility, and bridge the gap between computational prediction and experimental reality, offering a tool for accelerating the discovery of novel proteins with desired properties while fostering collaboration across different research groups.

P28 - Behaviour of Infective Stages (L3) of *Anisakis simplex* in Water Mass and Fish-waste: Transcriptomic Contribution to understanding the Parasite Life Cycle

Ríos-Castro, R[1]; Ramilo, A[1]; Rodríguez, H[1]; Abollo, E[1]; Pascual, S[1]

[1]Ecobiomar, Instituto de Investigaciones Marinas, Consejo Superior de Investigaciones Científicas, IIM-CSIC, 36208 Vigo, Spain

Anisakis simplex has a complex life cycle involving marine mammals as definitive hosts, where adults release eggs into the water. The larvae hatch and infect crustaceans, developing into the L3 stage. Fish and squid serve as paratenic hosts, accumulating larvae that are transferred up the food chain. When consumed by marine mammals, the larvae mature into adults, completing the cycle. Humans can be accidental hosts through raw seafood consumption, but larvae do not mature in humans.

The aim of this study is to analyse the plasticity of gene expression in *A. simplex* L3 larvae under different environmental and host conditions to understand their adaptability and survival. It addresses whether the spilled infected viscera by *Anisakis* by the fishing fleet during gutting operations reinforces the parasite's biological cycle, increasing its spread in fish and cetaceans.

Using RNA sequencing (RNA-seq), transcriptomic profiles of larvae were analysed. the survival of L3 larvae in the water column at different temperatures (15°C and 28°C) and under in vitro conditions simulating the intermediate and definitive hosts. A bioinformatic pipeline involving sequence alignment (STAR) [1] with *A. simplex* genome, differential gene expression analysis (DEseq2) [2], and Gene Ontology (GO) enrichment was applied.

Transcriptomic profiles showed significant gene regulation differences under environmental stress. At 15°C, larvae experienced progressive energy depletion and cytotoxicity by day 15, surviving up to 60 days. At 28°C, short-term resistance mechanisms were activated, limiting survival to seven days. In agar penetration assays, gene activation related to moulting and L4 development was observed under homeothermic conditions, while in poikilothermic hosts, gene expression focused on survival without developmental progression.

Anisakis simplex larvae exhibit remarkable adaptability, adjusting gene expression to environmental conditions. Key genes related to metabolism and cuticle structure play vital roles in survival. These findings highlight the importance of molecular studies in assessing the ecological impact and developing strategies to control *A. simplex* spread as a panzootic animal disease, reducing risks to public health and marine ecosystem stability.

P29 - Developing Genomic Models to Predict Breast Cancer Metastatic Progression

Elias Barreira[1], Marta Martins[1,2,4], Luis Costa[1,2,3], Joao C. Guimaraes[1]

[1]Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
[2]Gulbenkian Institute for Molecular Medicine, Lisbon, Portugal
[3]Oncology Division, Hospital de Santa Maria, Centro Hospitalar Lisboa Norte, Lisbon, Portugal
[4]iPLuso, I-MVET, Faculdade de Medicina Veterinária da Universidade Lusófona, Lisbon, Portugal

Breast cancer metastasis is a complex and heterogeneous process with significant implications for patient outcomes. Patients are typically classified as oligometastatic (≤5 lesions in ≤2 organs) or polymetastatic (>5 lesions or >2 organs), with the former generally associated with better prognosis. While some oligometastatic patients remain stable for years, the transition to polymetastatic disease remains poorly understood. Elucidating the mechanisms underlying these divergent trajectories is critical for advancing prognostic models, refining treatment strategies, and developing predictive genomic tools.

This study presents a novel, iterative genomic analysis pipeline integrating whole-exome sequencing (WES) data from a cohort of 55 breast cancer patients. Variant calling was conducted using DeepVariant, with annotation and functional insights enriched via Ensembl databases and advanced plugins. The analysis spanned multiple biological organizational levels, including genes and proteins, employing supervised machine learning (ML) models enhanced by explainability techniques and ontological frameworks to balance predictive accuracy with interpretability.

Key challenges addressed include distinguishing somatic mutations from germline variants, as possible driving factors, using a limited panel of normals.

Preliminary findings reveal distinct genomic clusters correlated with metastatic phenotypes, though definitive biological motifs have not yet been identified. A predictive ML model with very good accuracy was developed, demonstrating promise for future applications despite the inherent complexities of the genomic landscape.

This work aims to create a robust genomic test capable of stratifying patients by metastatic progression risk while providing actionable biological insights to guide personalized treatment strategies. These findings highlight the promise of integrated genomic analyses in advancing precision oncology.

Bioinformatics Open Days 2025

SPONSORS AND COLLABORATORS

Universidade do Minho

BRAGA
SOA A FUTURO.

Segundino Coelho Lda
CAIXILHARIAS EM PVC

BioData.pt

Clarivate™

GERMANO DE SOUSA
CENTRO DE MEDICINA LABORATORIAL

TRANSPORTES URBANOS DE BRAGA

aauminho

Saborosa

Thank you!!!

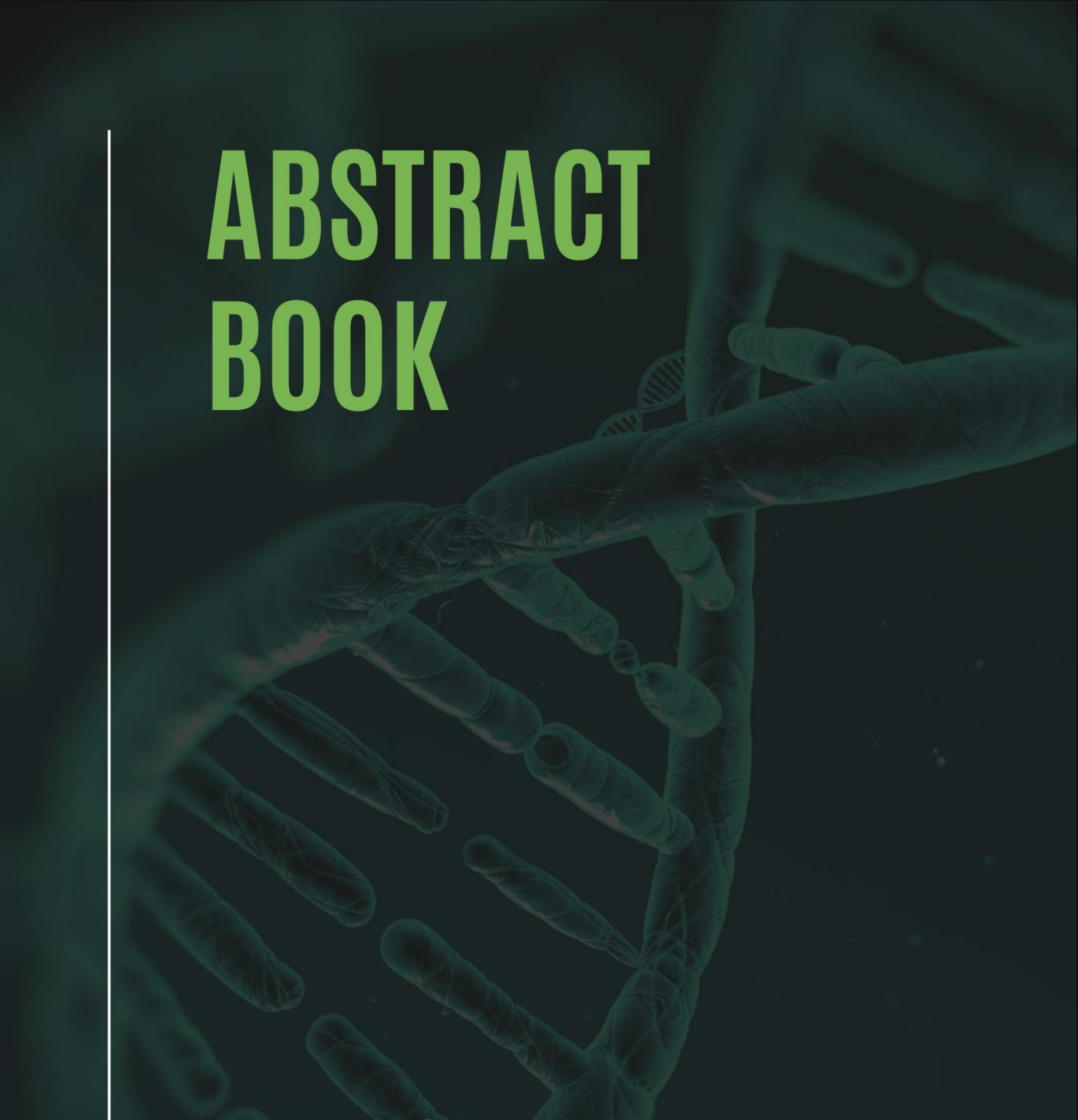
XIV EDITION

Bioinformatics Open Days 2025

Thank you for joining us at XIV BOD!

Your participation and support are greatly appreciated. We look forward to seeing you at future event editions and continuing to grow the Bioinformatics community.

ABSTRACT BOOK



XIV EDITION

26 - 28 MARCH

University of Minho
Gualtar Campus

Bioinformatics
Open Days 2025