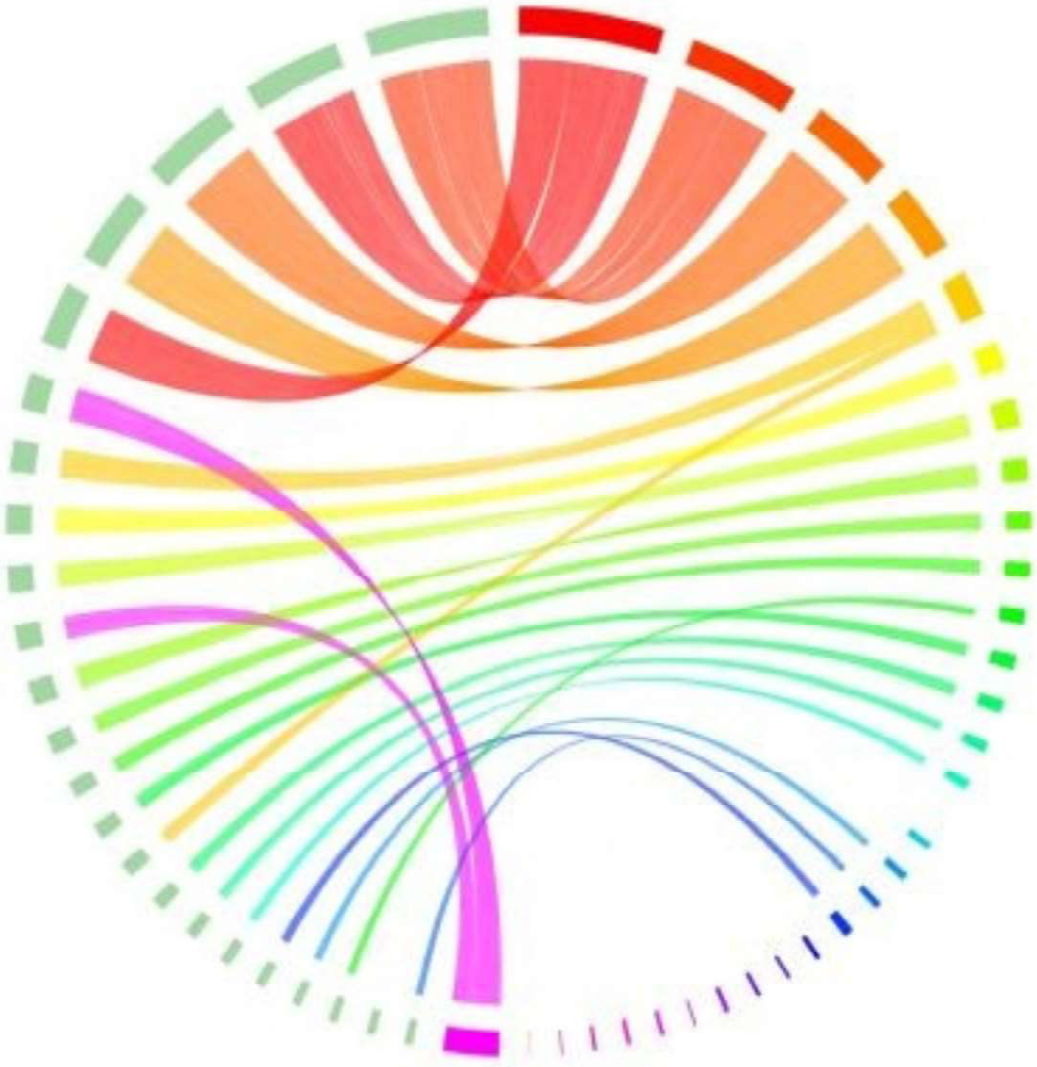
XII Jornada de Bioinformàtica i Genòmica

12-13 de desembre del 2024
Edifici del Centre de Cultures i Cooperació Transfronterera



Societat Catalana de BIOLOGIA

BIOINFORMATICS BARCELONA

Universitat de Lleida

XTS XARTEC SALUT

INSTITUT ESTUDIS ILERDENCS

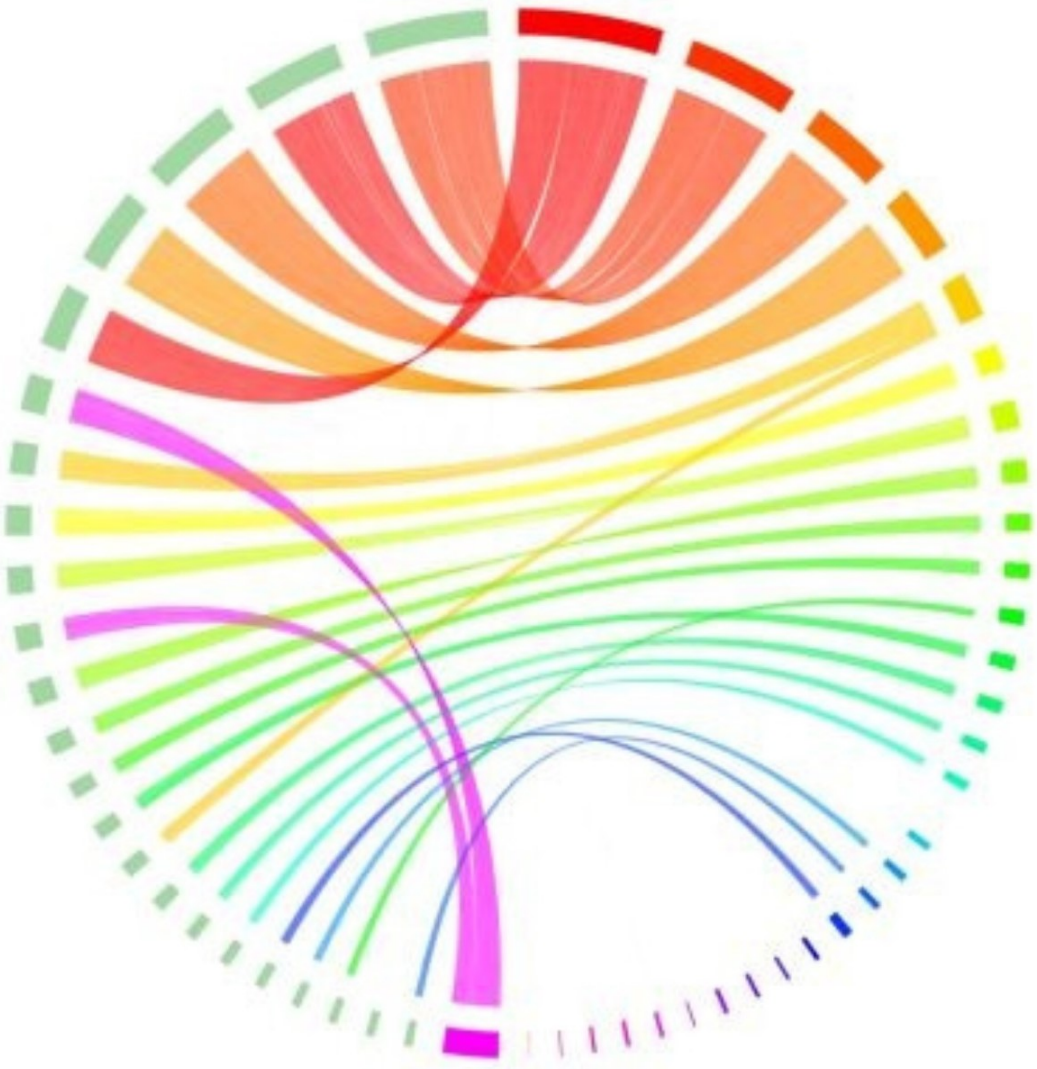
Diputació de Lleida
La força dels municipis

almirall

Alemany 1879

Comitè organitzador

Rui Alves
Ivan Benitez
Roger Ros
Ester Vilaprinyó
Marina Laplana
Alberto Marin
Jose Serrano



Suport a l'organització

Societat Catalana de Biologia

Roderic Guigó Mario Cáceres Mariàngels Gallego Paqui Lorite

Bioinformatics in Barcelona - BIB

Xavier Daura Rosa Bover

Xartec Salut

Alberto Rodríguez Cristina Páez

A Bit of History

The Bioinformatics and Genomics Annual Symposium of SCB began as a response to the growing importance of these fields in science. In its beginning, the Symposium aimed to connect local researchers in Catalonia. However, it quickly gained international participation, enriching knowledge exchange and collaboration. Over time, the conference has adapted to advancements like massive sequencing, omics data analysis, and AI in biology, fostering interdisciplinary dialogue across fields such as medicine, biochemistry, and computer science. Originally held in Barcelona, the event now rotates locations to promote inclusivity and regional development. In 2024, the 12th edition will take place on December 12–13 at the Edifici del Centre de Cultures i Cooperació Transfronterera, Universitat de Lleida.

Objectives of the XIIth Bioinformatics and Genomics Symposium

Share Scientific Advances: Present cutting-edge research in bioinformatics and genomics.
Promote Collaboration: Build connections across disciplines and institutions.
Showcase New Technologies: Highlight recent tools and innovations in biological data analysis.
Discuss Current Challenges: Create spaces to debate key trends and future directions.
Support Early-Career Researchers: Offer training opportunities through talks and workshops.
Drive Knowledge Transfer: Introduce dedicated sessions for applying research findings in society and industry.

These efforts aim to strengthen the scientific community and advance research in bioinformatics and genomics. We hope you will have a fruitful visit to Lleida and enjoy the meeting!

Una Mica d'Història

Les Jornades de Bioinformàtica i Genòmica de la SCB va començar com una resposta a la creixent importància d'aquests camps en la ciència. Inicialment, el seu objectiu era connectar investigadors locals de Catalunya. No obstant això, ràpidament va guanyar participació internacional, enriquint l'intercanvi de coneixements i la col·laboració. Amb el temps, el simposi s'ha adaptat als avenços com la seqüenciació massiva, l'anàlisi de dades òmiques i la intel·ligència artificial aplicada a la biologia, fomentant el diàleg interdisciplinari en àrees com la medicina, la bioquímica i la informàtica. Originalment celebrat a Barcelona, l'esdeveniment ara alterna ubicacions per promoure la inclusió i el desenvolupament regional.
El 2024, la 12a edició tindrà lloc els dies 12 i 13 de desembre a l'Edifici del Centre de Cultures i Cooperació Transfronterera de la Universitat de Lleida.

Objectius de la XII Jornada de Bioinformàtica i Genòmica

Compartir avenços científics: Presentar investigacions innovadores en bioinformàtica i genòmica.
Fomentar la col·laboració: Crear connexions entre disciplines i institucions.
Mostrar noves tecnologies: Destacar eines i innovacions recents en l'anàlisi de dades biològiques.
Debatre reptes actuals: Obrir espais per discutir tendències i direccions futures.
Recolzar joves investigadors: Oferir oportunitats de formació mitjançant conferències i tallers.
Impulsar la transferència: Incloure sessions específiques per aplicar els resultats de la recerca a la societat i la indústria.

Aquests esforços tenen com a objectiu enfortir la comunitat científica i impulsar la recerca en bioinformàtica i genòmica. Us desitgem una visita profitosa a Lleida i que gaudiu del simposi!

Programa General

XII JORNADES DE BIOINFORMÀTICA I GENÒMICA

12 I 13 DE DESEMBRE

Edifici del Centre de Cultures i Cooperació Transfronterera, Universitat de Lleida

Conferències a l'auditori

Pòsters a la planta baixa

PROGRAMA

12 DE DESEMBRE

8:30 - 9:30	***Inscripcions***
9:30 - 9:50	Comitè de Benvinguda
9:50 - 10:40	Conferència Plenària (James Sharpe - EMBL)
10:40 - 11:30	Conferències curtes
11:30 - 11:50	***Pausa per cafè amb pòsters***
11:50 - 13:00	Conferències curtes
13:00 - 14:30	***Pausa per dinar***
14:30 - 16:00	Conferències curtes
16:00 - 17:30	***Pausa per cafè amb pòsters****
17:30 - 18:30	Conferències curtes
18:30 - 21:00	***Esdeveniment Social***

***Trobada amb l'Empresa: Almirall**

13 DE DESEMBRE

9:00 - 9:50	Conferència Plenària (Jose Pereira-Leal/Ophiomics)
9:50 - 10:20	Conferència (Joao Curado - Flomics)
10:20 - 10:50	Taula Rodona: Què busca la indústria en un bioinformàtic?
10:50 - 11:30	***Pausa per cafè amb pòsters i networking***
11:30 - 12:00	Masterclass: com comunicar la teva idea (Brenda Serrano Casasola - Arkansas State University)
12:00 - 12:50	Conferència Plenària (Ana Conesa - CSIC)
12:50 - 13:05	Recerca en Català (Toni Hermoso i Roderic Guigó)
13:05 - 13:30	***Premis i Cloenda***

Organitzen: Secció de Bioinformàtica i Genòmica de la SCB i Bioinformatics Barcelona

Program at a glance

December 12th

08:30 – 09:30	Late Registration
09:30 – 09:50	Welcome committee
09:50 – 10:40	Plenary Talk: James Sharpe
10:40 – 11:30	Short Talks: Session 1
11:30 – 11:50	Coffee break and Poster setup
11:50 – 13:00	Short Talks: Session 2
13:00 – 14:30	Lunch break
14:30 – 16:00	Short Talks: Session 3
16:00 – 17:30	Poster Session with networking, coffee break and Meet the company (Almirall)
17:30 – 18:30	Short Talks: Session 4
18:30 – 21:00	Social event: Cocktail dinner at La Llotja

December 13th

09:00 – 09:50	Plenary Entrepreneurship Talk: José Pereira Leal
09:50 – 10:20	Guest Entrepreneurship Talk: João Curado
10:20 – 11:00	Round Table: What does the industry look for in a Bioinformatician?
11:00 – 11:30	Poster Session with coffee break
11:30 – 12:00	Invited Master Class: How to present your idea
12:00 – 12:50	Plenary Talk: Ana Conesa
12:50 – 13:05	Research in Catalan
13:05 – 13:30	Awards Session and Symposium Closing

Detailed Program

Invited Talks

12th of December

09:50: Turing patterns on Turing machines. Understanding the how organ morphology is encoded by dynamical mechanisms, presented by **James Sharpe**, Head of EMBL Barcelona.

Dr. Sharpe's research focuses on understanding the intricate processes of organ development and tissue regeneration. His work has shed light on the complex interplay between genetic and environmental factors that shape our organs and tissues.

13th of December

09:00: Entrepreneurship Talk – Diagnostics, big data and challenges thereof, presented by **Jose Pereira Leal**, CEO of Ophiomics.

Dr. Pereira-Leal has been a driving force behind tools like HepatoPredict, revolutionizing clinical decision-making in liver cancer.

09:50: Startups: tools to accelerate the impact of health research, presented by **João Curado**, CEO of Flomics Biotech.

Dr. Curado is the co-founder of Flomics Biotech, a Barcelona-based company pioneering RNA-based liquid biopsy technology for the early detection of cancer.

11:30: Master Class on how to present your idea, presented by **Brenda Serrano Casasola**, Professor, Arkansas State University at Queretáro.

Strong media and communication professional, graduated from Tecnológico de Monterrey (Mexico) and Universidad Rafael Landivar (Guatemala).

12:00: Single-molecule sequencing of the transcriptome: the TODOs, presented by **Ana Conesa**, CSIC Research Professor at Institute for Integrative Systems Biology.

Prof. Conesa is a renowned bioinformatics researcher specializing in functional genomics and multiomics analysis. Her work advanced understanding of gene expression dynamics and their applications in both basic and clinical research.

Round Table: What Does the Industry Look for in a Bioinformatics Professional?

December 13th, at 10:20

Introduction

With the entrepreneurial invited talks, we explore how to transform intellectual property into a startup and learn about the steps involved in starting your own company. This workshop aims at shifting the focus to the other side of the equation: What do companies seek when hiring Bioinformatics professionals?

To address this question, we have invited a diverse panel of experts who will share insights from both industry and academia. Representing the industry, we have José Pereira-Leal and João Curado, who have previously shared their entrepreneurial journeys and insights into running successful companies. This time, they will provide their perspective on what they look for when hiring Bioinformatics professionals.

On the academic and research side, we have Alex Pereira and Xavier Daura.

- **Alex Perera** is a Full Professor at UPC, Director of CREB, and Director of Xartec Salut, a network connecting over 100 entities to drive HealthTech innovation, foster collaboration, and support technology transfer.
- **Xavier Daura**, an ICREA Research Professor, has also held management roles such as Director of IBB and Academic Director of Bioinformatics Barcelona (BIB). BIB is a non-profit association promoting education, research, and industry collaboration in Bioinformatics.

Our panel includes individuals who are both consumers (industry employers) and producers (academic trainers) of Bioinformatics professionals, providing a comprehensive view of the skills, traits, and qualifications valued in the field.

Structure

We aim to have two main rounds of questions, with a potential third round if time permits. Each round will begin with questions directed at industry representatives ("consumers") followed by questions for academic experts ("producers").

Award Session and Symposium Closing program

13:05 – 13:10	Presentation by BIB and SCB representatives
13:10 – 13:15	Presentation of Best Poster and Best Short Talk Awards
13:15 – 13:20	Presentation of Best Bioinformatics Thesis award
13:20 – 13:30	Short Talk: Best Bioinformatics Thesis Awardee
13:30	Symposium Closing

Short Talks Program and Abstracts

Short talks should be prepared to be 7 minutes leaving time for questions

Session 1 (10:40 – 11:30): Single-Cell and Multi-Omics Approaches in Biology

Talk 1.1: Realistic scRNA-read simulation at isoform resolution using scr4eam, Marcel Schilling.

Talk 1.2: High content of nuclei-free low-quality cells in reference single-cell atlases: a call for more stringent quality control using nuclear fraction, Tomàs Montserrat-Ayuso.

Talk 1.3: Single-cell atlas of the human immune system reveals sex-specific dynamics of immunosenescence, Maria Sopena.

Talk 1.4: Single-cell characterization of an in vitro model of Patient-Derived Follicular Lymphoma Tumoroids, Víctor Jiménez-Martínez.

Talk 1.5: From Blood to Behavior: Identifying Biomarkers of Treatment-Resistant Depression via Multi-Omics, Anna Sirés.

Session 2 (11:50 – 13:00): RNA Biology and Gene Regulation

Talk 2.1: Long-read transcriptomics of a diverse human cohort reveals widespread ancestry bias in gene annotation, Pau Clavell.

Talk 2.2: Scalpel, a new tool for isoform quantification at single cell resolution, Franz Ake.

Talk 2.3: Ribonet identifies cell state-dependent patterns in mRNA translation, Xavier Hernandez.

Talk 2.4: Chromatin dynamics during cellular transdifferentiation, Silvia Gonzalez.

Talk 2.5: Slow transdifferentiation of the human B cells is the by-product of the Alu repeats expansion, Ramil Nurtdinov.

Talk 2.6: Targeting the "untargetable": inhibition of the transcription factor foxc1 with crispr/cas9 ribonucleoproteins reduces triple negative breast cancer aggressiveness, Marta Marqués.

Talk 2.7: Predicting response to cisplatin-based neoadjuvant chemotherapy for muscle-invasive bladder cancer: transcriptomic features outrank genomic biomarkers, Ariadna Acedo.

Session 3 (14:30 – 16:00): Evolutionary Genomics and Adaptation

Talk 3.1: The Faculty of Language and the Regulatability of Gene Expression, Julio Collado-Vides.

Talk 3.2: Understanding evolutionary novelties: a single-cell perspective on the turbanate eye in mayflies, presened by Maria Rosselló.

Talk 3.3: Hybrid assembly and comparative genomics unveil insights into the evolution and biology of the red-legged partridge, Abderrahmane Eleiwa.

Talk 3.4: Exploring adaptation across Europe through a polygenic perspective, Xavier Martí.

Talk 3.5: Exploring the functional consequences of sppl2c variants within a positively selected inversion in human populations, Illya Yakymenko.

Talk 3.6: The impact of population susbtructure in demographic inference Alba Nieto.

Talk 3.7: Using pangenomes to characterize complex transposon insertions in plants, Noemia Morales-Díaz.

Talk 3.8: Exploring functional conservation in silico: a new machine learning approach to RNA-editing, Michal Zawisza.

Session 4 (17:30 – 18:30): Model Systems and Translational Research

Talk 4.1: Modeling the effect of daytime duration on the biosynthesis of terpenoid precursors, Oriol Basallo.

Talk 4.2: Early environmental challenges may remodel murine phenotypes across multiple generations via small noncoding RNA, Flavio Palmieri.

Talk 4.3: Genetic analysis of mitochondrial dna copy number as an endophenotype, Eduard Molinero.

Talk 4.4: Multiomic integration analysis identifies atherogenic metabolites mediating between novel immune genes and cardiovascular risk, Robert Carreras.

Talk 4.5: Single-neuron deep generative model uncovers underlying physics of neural firing and connectivity, Jordi Abante.

Realistic scRNA-read simulation at isoform resolution using sc^{4}ream

Marcel Schilling[1,2], Franz Arnold AKE[1,2], Mireya Plass[1,2,3]

1. Gene Regulation of Cell Identity, Regenerative Medicine Program, Bellvitge Institute for Biomedical Research (IDIBELL), 08908, L'Hospitalet del Llobregat, Spain
2. Program for Advancing Clinical Translation of Regenerative Medicine of Catalonia, P-CMR[C], 08908, L'Hospitalet del Llobregat, Spain
3. Center for Networked Biomedical Research on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), 28029, Madrid, Spain

Presenter e-mail: mschilling@idibell.cat

In recent years, single-cell RNA-sequencing (scRNA-seq) techniques have transformed the field of transcriptomic studies. This has led to the development of not only a multitude of new specialized experimental methods but also of many new computational approaches to analyze scRNA-seq data, creating a strong demand for benchmarking approaches.

In the case of scRNA-seq data, existing tools to generate artificial 'ground truth' data are often limited to simulating gene counts per cell, *i.e.* digital gene expression matrices (DGEs). While this may suffice to benchmark workflows for gene quantification or differential gene expression analysis, they are insufficient to benchmark other types of tools, such as isoform quantification tools, which require the simulation of synthetic reads. Currently, most of the existing scRNA-seq read generators do not use a realistic distribution of reads along gene loci. Therefore, they cannot be used to benchmark software that considers the sequence and location of reads along a genomic region.

Here we present scr^{4}eam (GitHub: plasslab/scr4eam), a versatile user friendly and efficient tool to simulate realistic scRNA-seq reads based on a reference data set. In contrast to other tools, scr^{4}eam considers the relative expression of different transcripts from a gene to generate synthetic reads. These reads thus recapitulate real read distributions along a gene region and reflect existing biases in scRNA-seq data.

Together, we consider that scr^{4}eam will be a useful tool for the evaluation of new software working with scRNA-seq data.

High content of nuclei-free low-quality cells in reference single-cell atlases: a call for more stringent quality control using nuclear fraction

Tomàs Montserrat-Ayuso[1,2], Anna Esteve-Codina[1,2]

1. Centre Nacional d'Anàlisi Genòmica (CNAG), Baldiri Reixac 4, 08028 Barcelona, Spain
2. Universitat de Barcelona (UB), Barcelona, Spain

Presenter email: tomas.montserrat@cnag.eu

The advent of droplet-based single-cell RNA-sequencing (scRNA-seq) has dramatically increased data throughput, enabling the release of a diverse array of tissue cell atlases to the public. However, we will show that prominent initiatives such as the Human Cell Atlas, the Tabula Sapiens and the Tabula Muris contain a significant amount of contamination products (frequently affecting the whole organ) in their data portals due to suboptimal quality filtering. Our work addresses a critical gap by advocating for more stringent quality filtering, highlighting the imperative for a shift from existing standards, which currently lean towards greater permissiveness. We will show the importance of incorporating cell intronic fraction in quality control -or MALAT1 expression otherwise- showcasing its informative nature and potential to elevate cell atlas data reliability. In summary, here, we unveil the hidden intronic landscape of every tissue and highlight the importance of more rigorous single-cell RNA-sequencing quality assessment in cell atlases to enhance their applicability in diverse downstream analyses.

Single-cell atlas of the human immune system reveals sex-specific dynamics of immunosenescence

Maria Sopena-Rios[1,2*], Aida Ripoll-Cladellas[1,2*], Fatemeh Omidi[3], Sara Ballouz[4], Jose Alquicira-Hernandez[5,6,7,8,9], Roy Oelen[10], Lude Franke[10], Monique G.P. van der Wijst[10], Joseph E. Powell[9,11] & Marta Melé[1] *Contributed equally

1. Life Sciences Department, Barcelona Supercomputing Center (BSC), Barcelona, Spain
2. Universitat de Barcelona, Barcelona, Spain
3. Research and Development Department, qGenomics Laboratory, Barcelona, Spain.
4. School of Computer Science and Engineering, University of New South Wales, Sydney, Australia
5. Center for Data Sciences, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA.
6. Divisions of Rheumatology and Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
7. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA.
8. Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA
9. Translational Genomics, Garvan Institute of Medical Research, Sydney, NSW, Australia.
10. Department of Genetics, University of Groningen, Medical Center Groningen, Groningen, the Netherlands
11. UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, Australia

Immunosenescence, the gradual deterioration of the immune system with age, leads to an increased susceptibility to a range of diseases associated with immune dysfunction. Notably, sex is an important variable underlying how immune aging unfolds, as, for instance, autoimmunity develops with aging differently between males and females. Even though some clinical and molecular differences have been identified between male and female immunosenescence, it is not known to what extent sex affects the dynamic composition of immune cells over time. Here, we analyze a large single-cell RNA-sequencing dataset of peripheral blood mononuclear cells from a sex-balanced cohort of 982 human donors providing novel transcriptional and cellular insights into immune aging at an unprecedented resolution. We uncover that aging induces cell type-dependent and sex-specific transcriptional shifts that translate into a differential abundance of distinct immune cell subpopulations. These shifts predominantly involve translation-related genes, indicating a strong link between transcriptional and translational throughput with cell function and consequent immune cell composition. This sexual dimorphism overlaps known autoimmune disease-related genetic variants and results in the differential enrichment of functionally distinct immune populations. Specifically, we uncover that a cytotoxic CD8+ T effector memory subpopulation with an NK- like phenotype accumulates with age only in females and identify a distinct B cell subpopulation that expands with age exclusively in males, representing novel sex-specific hallmarks of immune aging. Our findings underscore the hidden complexity of immune aging and demonstrate the value of high-resolution, single-cell analyses in large population cohorts. This research paves the way for future sex-specific interventions targeting immunosenescence to ultimately promote a personalized approach to foster healthy aging.

Single-cell characterization of an *in vitro* model of Patient-Derived Follicular Lymphoma Tumoroids

Jiménez-MartínezV.[1,2], GiménezR.[1], SerratN.[1], FernándezPérezR.[1,3], Martínez-FarranA.[4], RoblesJA.[5], López-Guillermo A.[1,2,4,7], Campo E.[1,2,4,6], Villanueva J.[4,5], Martin-Subero JI.[1], Hodson DJ.[7], Colomer [1,2,4,6], Pérez-Galán P.[1,6]

1. Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), Barcelona, Spain.
2. University of Barcelona, Medical School, Barcelona, Spain.
3. Departamento de Bioquímica y Biología Molecular, Facultad de Veterinaria, Universidad Complutense de Madrid, Madrid, Spain.
4. Secció Hematopatologia, Servei d'Anatomia Patològica, Hospital Clínic, Barcelona, Spain.
5. Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain.
6. Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Instituto de Salud Carlos III, Madrid, Spain.
7. Department of Haematology, Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge Cambridge, United Kingdom.

Presentere-mail: VIJIMENEZ@recerca.clinic.cat

Follicular lymphoma (FL) is among the most common types of non-Hodgkin lymphoma, tipically manifesting as an indolent disease. Originating in mature B cells, the role of the immune tumor microenvironment (TME) is capital to understand disease onset, progression, and therapy resistance. This makes it an excellent paradigm to study tumor cells and TME crosstalk in lymphoma. In this work, we have developed an FL *in vitro* model based on patient-derived lymphoma tumoroids (PDLT). The system is composed of tumor B cells and autologous T cells cohabiting with supportive follicular dendritic cells (FDC; YK6 cell line). The B cells can be stimulated by inducing expression of specific cytokines and factors typical of germinal centers (GC) –*i.e* CD40L and IL-21– and thus simulate tumor progression and expansion. Here, we present a single-cell multiomic characterization of the model using 10X Chromium Singel Cell Immune Profiling data (RNA-seq and B cell receptor [BCR] sequencing). We have characterized a diagnosis biopsy, PDLT and stimulated PDLT (FDC-GC) belonging to the same case. Our results showed that both tumor cells and TME maintained a FL phenotype when cultured in the PDLT. While coculture with FDCs yielded a profile more closely related to an indolent FL, the FDC-GC stimulation recapitulated features of a high-grade FL in terms of proliferation, metabolism or cell signaling. In addition, B cell clonality did not change throughout PDLT culture and stimulation, with cells maintaining the same major clonotype. In conclusion, this PDLT system stands as a promising tool to study FL progression, transformation, and therapies *in vitro*.

From Blood to Behavior: Identifying Biomarkers of Treatment-Resistant Depression via Multi-Omics

Anna Sirés[1], Nicola Lorenzon[2,3], María Martínez de Lagrán[2,3], Massimo Gennerelli[4], Bernardo Carpiniello[5], Juan Navarro[6], Carmen Pedraza[6], Ewa Ferensztajn-Rochowiak[7], Filip Rybakowski[7], Ferran Sanz[1], Alessio Squassina[8], Mirko Manchia[9], PROMPT Study group, Marie Claude Potier[10], Alessandra Minelli[11,12], Bernhard T. Baune[13], Fernando Rodriguez de Fonseca[6*], Mara Dierssen[2,3*], Júlia Perera-Bel[1*]

Presenter e-mail: amartinez3@researchmar.net

1. Research Program on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Barcelona, Spain
2. Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain
3. Universitat Pompeu Fabra (UPF), Barcelona, Spain
4. Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy; Genetics Unit, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy.
5. Section of Psychiatry, Department of Medical Science and Public Health, University of Cagliari & Psychiatric Unit, University Hospital, 09127 Cagliari, Italy.
6. Grupo de Neuropsicofarmacología, Instituto de Investigación Biomédica de Málaga y Plataforma en Nanomedicina-IBIMA Plataforma BIONAND, Unidades Clínicas de Neurología y Salud Mental, Málaga, Spain
7. Department of Adult Psychiatry, Poznan University of Medical Sciences, Poznan, Poland.
8. Department of Biomedical Sciences, Section of Neuroscience and Clinical Pharmacology, University of Cagliari, Cagliari, Italy; Department of Psychiatry, Dalhousie University, Halifax, NS, Canada.
9. Section of Psychiatry, Department of Medical Sciences and Public Health, University of Cagliari, Italy; Unit of Clinical Psychiatry, University Hospital Agency of Cagliari, Italy
10. Institut du Cerveau-Paris Brain Institute-ICM, Institut national de la santé et de la recherche médicale, Centre National de la Recherche Scientifique, Assistance Publique-Hôpitaux de Paris, Hôpital de la Pitié Salpêtrière, Sorbonne Université, Paris, France.
11. Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy
12. Genetics Unit, San Giovanni di Dio Fatebenefratelli Center (IRCCS), Brescia, Italy
13. Department of Psychiatry and Psychotherapy, University of Münster, Münster, Germany

Treatment-resistant depression (TRD) poses a significant clinical challenge within major depressive disorder (MDD) due to its limited response to conventional therapies, underscoring the need for biomarkers that could better predict treatment outcomes. To address this, we applied a multi-omics approach, analyzing both whole-blood RNAseq profiles and small RNAseq data from MDD patients categorized as TRD and non-TRD, complemented by in vivo validation. RNAseq analysis revealed unique TRD patterns, including elevated expression of non-coding RNAs—especially lncRNAs—and ribosomal activity alongside a dysregulation of immune functions, suggesting immune and cellular stress deficiencies in TRD. Additionally, small RNAseq analysis showed significant downregulation of multiple miRNAs, which combined with down-regulation of genes involved in miRNA regulation suggest a coordinated regulatory imbalance of miRNA biogenesis in TRD. Using intracerebroventricular injections of antagomirs in mice, we induced depressive-like behaviors, validating these miRNAs' role in mood regulation. Together, these findings lay the groundwork for improved diagnostic and open new avenues for targeted therapies in TRD, aiming to shift the treatment paradigm in MDD toward more personalized, molecularly-informed approaches.

Long-read transcriptomics of a diverse human cohort reveals widespread ancestry bias in gene annotation

Pau Clavell-Revelles[1,2], Fairlie Reese[1], Sílvia Carbonell-Sala[3], Carme Arnan[3], Fabien Degalez[3], Roderic Guigó[3,4*] & Marta Melé[1 *]

* Corresponding author and equal contribution

1. Life Sciences Department, Barcelona Supercomputing Center (BSC), Barcelona, Catalonia.
2. Universitat de Barcelona (UB), Barcelona, Catalonia.
3. Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Catalonia.
4. Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra (UPF), Barcelona, Catalonia.

Gene and transcript annotations are essential for interpreting biological findings in genetics and genomics. Yet, gene annotations and most available transcriptomic datasets either derive exclusively from European individuals or rely on short-read technologies which can not unambiguously elucidate complete transcript structures. Therefore, whether these annotations are European biased at the full-length transcript level and the consequences of such a bias are currently unknown. Here, we sequence over 600 million full-length RNA reads from 8 distinct human populations and demonstrate that existing reference gene annotations systematically better represent transcripts from European ancestry. From our data, we build a novel gene annotation which represents human transcriptomic diversity and identify over 2300 population-specific transcripts which are more frequently already annotated in Europeans. Finally, we prove that the addition of these novel transcripts to GENCODE significantly increases discovery potential of allele-specific transcript usage, especially in non-Europeans. Overall, our work highlights the pressing need of developing a gene annotation that better represents the transcriptome diversity in all present-day human individuals worldwide.

SCALPEL, a new tool for isoform quantification at single cell resolution

Franz Ake[1,2], Ana Gutierrez-Franco[1,2], Sandra Maria Fernandez-Moya[1,2], Mireya Plass-Portulas[1,2,3]

1 Gene Regulation of Cell Identity, Regenerative Medicine Program, Bellvitge Institute for Biomedical Research (IDIBELL), 08908, L'Hospitalet del Llobregat, Barcelona, Spain
2 Program for Advancing Clinical Translation of Regenerative Medicine of Catalonia, P-CMR[C], 08908, L'Hospitalet del Llobregat, Barcelona, Spain
3 Center for Networked Biomedical Research on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), 28029, Madrid, Spain
Presenter e-mail: fake@idibell.cat

Alternative polyadenylation (APA) is a widespread mechanism of gene regulation that that affects more than 70% of human genes. mRNA isoforms generated though APA have distinct 3' ends, which can affect mRNA regulation as well as the resulting proteins. APA is well known to be regulated during cell differentiation and it is a major source of gene regulation in the brain. Yet, it is not known till which extend APA contributes to the transcriptomic variability across individual cell populations.

The rapid development of 3'tag-based single-cell RNA sequencing (scRNAseq) has enabled the study of gene expression and the implementation of methods for describing isoform usage at single cell resolution. Yet, most of these methods have low sensitivity when quantifying isoforms at the single-cell level. Here we present SCALPEL, a Nextflow based tool to quantify and characterize isoform differential usage at the single-cell level. Benchmark analysis shows that SCALPEL outperforms current published tools in the identification of genes expressing multiple isoforms and detecting changes in isoform usage. We used SCALPEL to study the changes in isoform usage during the differentiation of human induced pluripotent stem cells (iPSCs) to neural progenitor cells (NPCs). The results from our analysis show clear changes in 3'end usage between iPSCs and NPCs. We aim to use SCALPEL to investigate the role of APA during neural differentiation and how these changes are altered in neurodegenerative diseases.

RIBONET identifies cell state-dependent patterns in mrna translation

Xavier Hernandez-Alias[1,2], Selay Kaya[1], Julien Gagneur[2], Danny Nedialkova[1,3]

1. Mechanisms of Protein Biogenesis, Max Planck Institute of Biochemistry, Martinsried, Germany
2. School of Computation, Information and Technology, Technical University of Munich, Garching, Germany
3. School of Natural Sciences, Technical University of Munich, Garching, Germany

Presenter e-mail: hernandezalias@biochem.mpg.de

Cells react to environmental stimuli by regulating gene expression at many different levels and time scales. While factors and motifs that mediate changes in transcription have been extensively studied, the regulation of posttranscriptional mechanisms is much less understood. Ribosome profiling enables the mapping and quantitation of ribosome positions along mRNAs with codon resolution, yielding rich informative signal to quantitatively model translation processes from sequence. To uncover the underpinning regulatory code in both untranslated regions (UTRs) and coding sequences, we developed RiboNet, a deep learning model that predicts absolute ribosome counts and their distribution along endogenous mRNA transcripts. We developed RiboNet as a multi-task framework across measured conditions, which allows for flexible model interpretation approaches by in silico mutation of regulatory sequences and comparisons between differential conditions. We benchmarked RiboNet performance with ribosome profiling datasets from diverse eukaryotes, uncovering known and novel predictive motifs and codons upon tRNA pool perturbations. RiboNet thus provides an end-to-end modeling and interpretation framework that identifies the intricate regulatory patterns of both mRNA translation initiation and elongation between contrasting cell states.

Chromatin dynamics during cellular transdifferentiation

Silvia González-López[1,2], Marina Ruiz-Romero[1], Sílvia Pérez-Lluch[1], Roderic Guigó[1,2]

1. Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Barcelona (BIST), Catalonia, Spain
2. Universitat Pompeu Fabra (UPF), Barcelona, Catalonia, Spain Presenter e-mail:
silvia.gonzalez@crg.eu

Candidate *cis*-regulatory elements (cCREs) have been one of the focuses of the ENCODE consortium since the release of their first registry of these regions, in 2016; however, their mechanism of action is yet to be fully understood. Analyzing the logic behind the regulatory functions that these elements are involved in requires a wide array of high-resolution temporal data, hence, in this study, we have used publicly available datasets describing the chromatin landscape –ChIP-seq (H3K4me1/2/3, H3K9ac/me3, H3K27ac/me3, H3K36me3, H4K20me1) and ATAC-seq– during the transdifferentiation of human B-cell precursors to macrophages in order to explore the dynamics of a set of system-specific cCREs.

Taking advantage of this information, we have observed that histone post-translational modifications (hPTMs) are deposited in an order that is conserved across cCRE categories: first H3K4me1/2, followed by H3K27ac, H3K9ac and, if applicable, H3K4me3, and vice versa for their removal. Furthermore, we have found that chromatin can be defined in few states –i.e. combinations of hPTMs marking– which, in turn, lead to limited profiles –or trajectories– over time; particularly, regions proximal to transcription start sites mostly remain in the same state during transdifferentiation, while distal ones are much more dynamic. Besides, we have identified groups of genomic regions that display coordinated chromatin dynamics during transdifferentiation.

Overall, thanks to the broad description of chromatin attributes in this dynamic system, we have been able to explore the patterns of chromatin decorations involving regulatory regions and their evolution. This enables a better comprehension of the logic behind some epigenetic processes happening during transdifferentiation and their possible biological implications.

Slow transdifferentiation of the human B cells is the by-product of the Alu repeats expansion

Ramil Nurtdinov, Carme Arnan, María Sanz, Amaya Abad, Alexandre Esteban, Sebastian Ullrich, Rory Johnson, Sílvia Pérez-Lluch and Roderic Guigó

Centre de Regulació Genòmica, CRG-UPF, Barcelona, Catalonia, Spain

Many developmental and differentiation processes take substantially longer in human than in mouse. To investigate the molecular mechanisms underlying this phenomenon, we have specifically focused on the transdifferentiation from B cells to macrophages. The process is triggered by exactly the same molecular mechanism, the induction by the transcription factor CEBPA, but takes three days in mouse and seven in human.

We collected a time-course of densely-spaced, high-depth RNA-Seq and CEBPA ChIP-Seq data during human and mouse transdifferentiation. We found quite good concordance of expression changes for the bulk of orthologous genes, however, both up- and, in particular, down-regulation were substantially delayed in human cells. Although the expression levels of CEBPA transgene were comparable, the CEBPA genomic binding was up to two-fold stronger in mouse cells. Analyzing this phenomenon, we found a large number of human- specific Alu repeats carry strong instances of the CEBPA binding motif. Overall, Alu repeats attract approximately twice more ChIP-Seq reads than expected using the random input controls. As a consequence, we hypothesize that Alu repeats work as sponges for CEBPA molecules, driving to the overall attenuation of CEBPA binding in the human genome.

To directly test this hypothesis, we attempted to overcome Alu competition using in-house CRISPR-dCas9 system. The system uses specific gRNAs to attract inactive dCas9 protein to target regions. We designed gRNAs to target as many Alu repeats showing strong CEBPA instances as possible. Further on, CEBPA and dCas9 binding and gene expression dynamics were monitored by ChIP-Seq and RNA-Seq, respectively in Alu-targeted and control cells. At three hours after induction of transdifferentiation, we found substantial increase in CEBPA binding to its canonical binding sites and increase in dCas9 binding at targeted Alu regions. We also found acceleration of gene expression up-regulation in Alu-targeted cells compared to control ones.

Overall, our work highlights the highly complex mode in which biological information is encoded in genome sequences, evolutionarily connecting, in an unexpected way, lineage-specific transposable element expansions to species-specific changes in developmental tempos.

Targeting the "untargetable": inhibition of the transcription factor *foxc1* with crispr/cas9 ribonucleoproteins reduces triple negative breast cancer aggressiveness

Marta Marqués[1], Izaskun Urdanibia[1], Iván Hidalgo[1], Joel Gabàs[3], Mariona Pont[1], Arnau Oronich[1], M. Alba Sorolla[1], Eva Parisi[1], Robert Montal[2], Rui Alves[3], Antonieta Salud[2], Anabel Sorolla[1#]

1. Research Group of Cancer Biomarkers, Lleida Institute for Biomedical Research Dr. Pifarré Foundation (IRBLleida), Av Alcalde Rovira Roure, 80, 25198, Lleida, Spain.
2. Oncology Department, Arnau de Vilanova University Hospital, Av Alcalde Rovira Roure, 80, 25198 Lleida, Spain.
3. Scientific service of Bioinformatics and Computational Biology, University of Lleida, Av Alcalde Rovira Roure, 80, 25198, Lleida, Spain.
#asorolla@irblleida.cat
Presenter e-mail: mmarques@irblleida.cat

Triple-negative breast cancer (TNBC) is an aggressive subtype of breast cancer (BC) that lacks oestrogen, progesterone, and human epidermal growth factor 2 (HER2) receptors, rendering unresponsive to standard hormonal and anti-HER2 therapies, which contributes to limited treatment options and poor prognosis. The aggressiveness of TNBC is largely determined by the overexpression of transcription factors (TFs) which drive tumour growth, survival, metastasis, and treatment resistance. Particularly, the TF *Forkhead Box C1* (*FOXC1*) was found to be the most overexpressed TF in TNBC, which regulates gene networks that drive malignancy, including cell proliferation, migration, and metastasis, representing a promising therapeutic target. Indeed, our bioinformatic analysis demonstrated that high expression of *FOXC1* correlates with lower patient survival.

In this study, we have employed CRISPR/Cas9 technology to knock out *FOXC1* in MDA-MB-231 and MDA-MB-468 TNBC cell lines, delivered via liposome- encapsulated single guide RNAs (sgRNAs) and Cas9 ribonucleoprotein (RNP) complexes. This resulted in an effective inhibition of *FOXC1* which caused a significant decrease in cell viability, proliferation, migration, invasiveness, and tumorigenic potential *in vitro and* in an *in vivo* MDA-MB-231 xenograft model. Additionally, the inhibition of *FOXC1* altered the protein expression of key epithelial-to-mesenchymal transition (EMT) markers, highlighting its role in TNBC progression. Further analyses, including RNA sequencing and coupled bioinformatics identified the introduction of a premature stop codon on the protein sequence, leading to a truncated inactive protein affecting the expression pattern of a subset of genes.

Altogether, our findings highlight the role of *FOXC1* in maintaining a more malignant state in TNBC. At the same time, it points out *FOXC1* as an attractive therapeutic target which could effectively be suppressed with the utilization of CRISPR/Cas9 RNPs. Further exploration will allow the delivery of FOXC1 RNPs *in vivo*, in form of therapy, to reduce the aggressive phenotype of TNBC, constituting an interesting targeted therapy for TNBC.

Predicting response to cisplatin-based neoadjuvant chemotherapy for muscle-invasive bladder cancer: transcriptomic features outrank genomic biomarkers

Ariadna Acedo-Terrades[1], Alejo Rodriguez-Vida[1,2], Oscar Buisan[3], Marta Bódalo-Torruella[1], Maria Gabarrós[1], Miquel Clarós[1], Nuria Juanpere[4], Marta Lorenzo[4], Sergio Vázquez Montes de Oca[1], Alejandro Rios-Hoyo[2], Cristina Carrato Moñino[5], Tamara Sanhueza[5], Eduardo Eyras[1,6], Eulàlia Puigdecanet[7], Gottfrid Sjödahl[8], Júlia Perera-Bel[1#], Lara Nonell[9#] & Joaquim Bellmunt[1,10#] *#co-corresponding authors*

[1]Hospital del Mar Research Institute (HMRI), Barcelona, Spain.

[2]Medical Oncology Department, Hospital del Mar, Barcelona, Spain.

[3]Urology Department, Hospital de Bellvitge, Barcelona, Spain.

[4]Department of Pathology, Hospital del Mar, Barcelona, Spain.

[5]Urology Department, University Hospital Germans Trias i Pujol, Badalona, Spain.

[6]EMBL Australia Partner Laboratory Network at the Australian National University, Canberra Australian Capital Territory, Australia.

[7]IRIS-CC & MoD Lab Research Group. Faculty of Medicine, UVic-UCC, Vic, Spain.

[8]Division of Urological Research, Department of Translational Medicine, Lund University and Department of Urology, Malmö, Sweden.

[9]Bioinformatics Unit, Vall d'Hebron Institute of Oncology, Barcelona, Spain.

[10]Dana Farber Cancer Institute, Harvard Medical School Boston Massachusetts, USA.

Muscle-invasive bladder cancer (MIBC) is associated with poor predictability of response to cisplatin- based neoadjuvant chemotherapy (NAC). Consequently, the benefit of NAC remains unclear for many patients due to the lack of reliable biomarkers predicting treatment response. In order to identify biomarkers and build an integrated and highly accurate model to predict NAC response, we performed a comprehensive transcriptomic and genomic profiling on tumors from 100 MIBC patients. Our results showed that the expression of the top genes associated with response, as well as the expression of growth factor genes and cell cycle regulators are highly correlated with NAC response. Most importantly, we found a novel signature related to the WNT signaling pathway that alone was highly correlated with NAC response and showed high accuracy in predicting NAC response (AUC=0.76). Additionally, mutations in the DNAH family genes (DNAH8, DNAH6 and DNAH10) and deletion in KDM6A were also highly correlated with NAC response. Using our comprehensive molecular analysis as a backbone, we developed two machine learning (ML) models, one incorporating both transcriptomic and genomic features (RF-RW), and the other using only transcriptomic data (RF-R). Both models demonstrated promising performance (AUC=0.82) as predictive models of response to NAC in MIBC. RF-RW and RF-R, after external validation, could potentially change the management of MIBC patients by selecting ideal candidates for NAC.

The Faculty of Language and the Regulatability of Gene Expression

Julio Collado-Vides
Center for Genomic Sciences
National Autonomous University of Mexico
Center for Genomic Regulation, UPF, Barcelona

In spite of the considerable progress in the understanding of the faculty of language, its nature is frequently misunderstood. In this talk I will first revise different concepts of what language is, as an introduction to the notion of generative grammars. I will compare so- called, the mostly descriptive "enhancer grammars" with a generative grammar of bacterial regulation. This introduces the question of how to study a biological capacity such as the regulatability of gene expression. As a consequence, I propose that the genetics of gene regulation is the theory species have naturally built in their evolution.

Understanding evolutionary novelties: a single-cell perspective on the turbanate eye in mayflies

Maria Rosselló[1,2], Helena García-Castro[3], Sophie Tandonnet[1,2], Rafath Chowdhury[1,2], Tòt Senar-Serra[1,2], Noelle Anderson[4], Scott Roy[4], Fernando Casares[5], Jordi Solana[3], Isabel Almudí[1,2]

1. Dept Genètica, Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Spain
2. Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain
3. Living Systems Institute, University of Exeter, Exeter, UK
4. Department of Biology, San Francisco State University, San Francisco, California, USA
5. Centro Andaluz de Biología del Desarrollo, CSIC-UPO-JA, Sevilla, Spain

Presenter e-mail: mariarossello@ub.edu

New morphological traits frequently play pivotal roles in evolutionary processes, yet the genetic mechanisms underlying such innovations remain poorly understood. The turbanate eye, a male-specific pair of eyes with an associated optic lobe found in *Cloeon dipterum*, offers a unique opportunity to study the genetic basis of this evolutionary novelty. These eyes, absent in females, represent a new and distinct trait within the species, located above the lateral compound eyes that are shared by both sexes.

To investigate the genetic basis of turbanate eye formation, we employed a multi-omics approach that integrates single-cell RNA sequencing to profile gene expression, along with bulk ATAC sequencing to assess chromatin accessibility. By integrating these datasets, we were able to uncover the gene regulatory network involved in the development of the turbanate eye.

Our analysis led to the identification of *dsxl*, a male-specific gene expressed during turbanate eye development. We hypothesize that *dsxl*, a putative de novo gene, plays a central role in the formation of this sexually dimorphic trait by regulating male-specific visual system components, with its function being further tested through functional experiments.

Ultimately, our findings provide new insights into how gene regulatory networks, such as those involving dsxl, contribute to the evolution of evolutionary innovations. By focusing on Clo*eon dipterum*, this study expands the application of state-of-the-art bioinformatics tools to non-model organisms, offering a novel framework for investigating the genetic basis of complex traits and their role in evolutionary adaptations in species with unique morphological innovations.

Hybrid assembly and comparative genomics unveil insights into the evolution and biology of the red-legged partridge

Abderrahmane Eleiwa, Jesus Nadal, Ester Vilaprinyo, Alberto Marin-Sanguino, Albert Sorribas, Oriol Basallo, Abel Lucido, Cristobal Richart, Ramona N. Pena, Roger Ros-Freixedes, Anabel Usie & Rui Alves

[1]Systems Biology Group, Department Ciències Mèdiques Bàsiques, Faculty of Medicine, Universitat de Lleida, Lleida, Spain
[2]IRBLleida, Lleida, Spain
[3]Grup de Recerca Consolidat de la Generalitat de Catalunya 2021 SGR 1353 MathSy2Bio
[4]Departament de Ciència Animal, Universitat de Lleida, Lleida, Catalonia, Spain.
[5]Agrotecnio-CERCA Center, Lleida, Catalonia, Spain.

Presenter e-mail:Abderrahmane Eleiwa,

The red-legged partridge *Alectoris rufa* plays a crucial role in the ecosystem of southwestern Europe, and understanding its genetics is vital for conservation and management. Here we sequence, assemble, and annotate a highly contiguous and nearly complete version of its genome. This assembly encompasses 96.9% of the avian genes flagged as essential in the BUSCO aves_odb10 dataset. Moreover, we pinpointed RNA and protein-coding genes, 95% of which had functional annotations. Notably, we observed significant chromosome rearrangements in comparison to quail (Coturnix japonica) and chicken (Gallus gallus). In addition, a comparative phylogenetic analysis of these genomes suggests that A. rufa and C. japonica diverged roughly 20 million years ago and that their common ancestor diverged from G. gallus 35 million years ago. Our assembly represents a significant advancement towards a complete reference genome for A. rufa, facilitating comparative avian genomics, and providing a valuable resource for future research and conservation efforts for the red-legged partridge.

Exploring adaptation across europe through a polygenic perspective

Xavier Martí-Pérez[1], Jorge García-Calleja[1], Rafael de Cid[2], Francesc Calafell[1], Elena Bosch[1]

1 Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona 08003, Catalonia, Spain
2 Genomes for Life-GCAT lab Group, Institute for Health Science Research Germans Trias i Pujol (IGTP), Badalona 08916, Spain
Presenter e-mail: xavier.martii@upf.edu

Most of the phenotypical spectrum is constituted by complex traits, which are influenced by many loci scattered across the genome. Though many minor allele frequency changes, polygenic selection acting on such multiple redundant loci of small effect can produce remarkably rapid phenotypic shifts. As a result of pleiotropy, adaptative processes happening on complex traits can lead to an increase in prevalence of complex diseases as a direct evolutionary trade-off. Here, we infer whether selection has acted in a recent (up to ~60,000 ya) and very recent (up to ~3,000 ya) time scales in 5 immune response conditions to different bacterial and viral stimuli and 90 polygenic traits from the UK Biobank (UKBB) including immune-related phenotypes, as well as metabolic, mental and physical constitution traits. To this aim, we use three polygenic selection statistics: the trait integrated haplotype score (tiHS), the trait singleton density score (tSDS), and the polygenic adaptation likelihood model (PALM) test. Additionally, we perform the joint PALM (J-PALM) analysis to detect signals of correlated selection and antagonistic selection. We applied these tests on almost 800 whole-genome sequenced (WGS) individuals of the GCAT cohort, which serves as a proxy for the Iberian population, and on the WGS European populations of the 1000 Genomes Project. Alongside the well-known selection towards lighter skin pigmentation, our results show selection signals in most European populations towards a generalized increase of prevalence of light hair colors and nervousness, as well as towards a decrease in the pulse rate and the prevalence of black hair color. We also appreciate highly complex correlated selection signals between pigmentation and hair color, which suggest that there might have been synergic selection processes in these traits, rather than one actually being the driver of the others.

Exploring the functional consequences of sppl2c variants within a positively selected inversion in human populations

Illya Yakymenko[1,2], David Aranda-García[1], Konstantinos Karakostis[1,2], Jana Selent[1], Mario Cáceres[1,2,3]

1. Research Program on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Barcelona, Spain
2. Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain
3. ICREA, Barcelona, Spain

Presenter e-mail: illya.yakymenko@uab.cat

Genomic inversions are structural variants that contribute to evolutionary adaptations. The 17q21.31 inversion is a well-studied ~1 Mb-long human inversion that is associated to multiple phenotypic traits and has been proposed to be positively selected in European populations. Due to recombination suppression, the two inversion haplotypes (H1 and H2) accumulate around 3000 variants in perfect linkage disequilibrium, affecting the regulation of multiple coding and non-coding transcripts. However, linking individual variants to specific traits remains very challenging. In this study, we focus on the potential functional effect of changes in the single-exon gene *SPPL2C*, which is located within the inverted region and encodes a 684 amino acid aspartyl protease. Notably, *SPPL2C* harbors seven missense mutations between the two orientations that may alter the function of the protein. Particularly, the substitution of a highly conserved arginine at position 461 with a proline (R461P) has been classified as potentially damaging by several tools (CADD score: 23.3; SIFT score: 0.01; AlphaMissense score: 0.851). We generated molecular dynamics simulations of both SPPL2c variants with one of its substrates, VAPB, embedded in the endoplasmic reticulum membrane. The R461P change apparently induces an allosteric conformational modification affecting the interaction of crucial residues of the catalytic center (D386 and D448) with VAPB. Specifically, VAPB forms a hydrogen bond with D386 during 20.8% of the simulated time in the R461 variant and with D448 for 29.0% of the time in the P461 variant. Interestingly, the other catalytic residue loses any interaction with VAPB in both cases. Thus, we hypothesize that the variants associated to the 17q21.31 inversion alter the catalytic activity of SPPL2c and, consequently, play a role in the regulation of the SNARE complex. This could be related to the protective effect against neurodegeneration of the H2 haplotype and experimental validation of the findings is currently underway.

The impact of population susbtructure in demographic inference

Alba Nieto Heredia
ISYEB-EPHE-PSL and IBE-CSIC

Gaining insights into the historical demographic dynamics helps to identify potential threats to species survival and to implement targeted conservation measures. Typically, the main metrics to assess the conservation status from genetic data is the variation in effective population size through time. This is related to the inverse of the coalescence rate (CR), and ultimately depends on the true demographic history of a species. SMC-based algorithms predict the variation of coalescent rate (CR) through time employing various population genetics summary statistics obtained from genomic data. However, their performance when analyzing demes issued from structured meta-populations remain poorly investigated.

Moreover, under panmixia, variations in CR correspond to changes in population size over time, which has a simple biological interpretation. Conversely, the interpretation of CR variation in structured population is challenging as it depends on the joint effect of all evolutionary forces acting on the meta-population. Fully understanding of the CR variation is therefore of paramount importance, claiming for the need for novel approaches to accurately interpret the CR and integrate it with other population genetic statistics. Here we evaluate the accuracy of SMC-based algorithms by testing their performance in complex structured demographic scenarios for which the true CR was generated either by the structured coalescent or by simulations.

We demonstrate how overlooking population substructure is introducing systematic biases in SMC-based algorithms inferences. These biases are prone to be included in further analyse involving the test of conservation policies, the evolutionary history of species and complex links in the ecosystem.

Using pangenomes to characterize complex transposon insertions in plants

Morales-Díaz Noemia[1], Svitlana Sushko[2], Lucía Campos-Dominguez[1], Venkataramana Kopalli[3], Agnieszka A. Golicz[3], Raúl Castanera[1,4], Josep M. Casacuberta[1]

1. Centre for Research in Agricultural Genomics, CRAG (CSIC- IRTA-UAB-UB), Campus UAB, Cerdanyola del Vallès, Barcelona, Spain
2. Department of Molecular Biology, Max Planck Institute for Biology Tübingen, Tübingen, Germany
3. Department of Plant Breeding, Justus Liebig University Giessen, Giessen, Germany
4. IRTA, Genomics and Biotechnology, Edifici CRAG, Campus UAB, 08193 Bellaterra, Catalonia, Spain

Transposable elements (TEs) are DNA sequences capable of mobilizing within host genomes, contributing to genomic diversity. Traditionally, TE insertions have been studied by aligning short-read sequences to a single reference genome and detecting discordant reads. However, this approach can introduce a bias, as reads containing non-reference alleles may be misaligned. With the recent availability of hundreds of long-read sequences, it is now possible to address this limitation by constructing a pangenome. In a pangenome, all DNA sequences from the accessions or varieties under study are used to build a variation graph containing nodes (DNA segments), edges (that connect nodes) and paths (routes through the different nodes of the graph), that represent all possible variants, including TE insertions or deletions.

Here we show how this approach is particularly useful to analyze transposon insertion polymorphisms in rice. Using multiple long-read assemblies, we analyzed Long Terminal Repeat Retrotransposon (LTR-RT) Tandems, which are complex structures composed of arrays of LTR-RT sequences sharing internal LTRs. We then explored two different graph pangenome construction approaches, minimap2+SVIM-asm+Truvari+vg and minigraph- cactus+vg, to assess their accuracy representing the intricated LTR-RT Tandem sequences. Minigraph-cactus showed better precision representing the structural variation associated with LTR-RT Tandems. Our analysis suggests that pangenome approaches are extremely useful to characterize complex regions in multiple genomes. Nevertheless, the preliminary results from genotyping these structures using short-read data and the pangenome reference indicate that further refinements in the computational tools are needed.

Exploring functional conservation in silico: a new machine learning approach to RNA-editing

Michal Zawisza-Alvarez[1], Jesus Peñuela-Melero[1], Esteban Vegas[1], Ferran Reverter[1], Jordi Garcia Fernandez[1], Carlos Herrera-Ubeda[1]

1. Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain
Presenter e-mail: m.zawisza@ub.edu

Around 50 years ago, molecular biology opened the path to understand changes in forms, adaptations, complexity or the basis of human diseases, trough myriads of reports on gene birth, gene duplication, gene expression regulation and splicing regulation, among other relevant mechanisms behind gene function. Here, with the advent of big data and artificial intelligence (AI), we focus on an elusive and intriguing mechanism of gene function regulation, RNA editing, in which a single nucleotide from an RNA molecule is changed with a remarkable impact in the increase of the complexity of transcriptome and proteome. We present a new generation approach to assess the functional conservation of the RNA-editing targeting mechanism using two AI learning algorithms, random forest (RF) and bidirectional long short-term memory (biLSTM) neural networks with attention layer. These algorithms combined with RNA- editing data coming from databases and variant calling from same-individual RNA and DNA-seq experiments from different species, allowed us to predict RNA-editing events using both primary sequence and secondary structure. Then, we devised a method for assessing conservation or divergence in the molecular mechanisms of editing completely in silico: the cross-training analysis. This novel method could set the basis to understand the conservation of the editing mechanism through evolution.

Modeling the effect of daytime duration on the biosynthesis of terpenoid precursors

Oriol Basallo[1,2,3], Abel Lucido[1,2,3], Albert Sorribas[1,2,3], Alberto Marin-Sanguino[1,2,3], Ester Vilaprinyo[1,2,3], Emilce Martinez[1,2,4], Abderrahmane Eleiwa[1,2,] Rui Alves[1,2,3*]

[1]Systems Biology Group, Department Ciències Mèdiques Bàsiques, Faculty of Medicine, Universitat de Lleida, Lleida, Spain
[2]IRBLleida, Lleida, Spain
[3]Grup de Recerca Consolidat de la Generalitat de Catalunya 2021 SGR 1353 MathSy2Bio
[4]Instituto Nacional de Tecnología Agropecuaria, Argentina

Terpenoids are valued chemicals in the pharmaceutical, biotechnological, cosmetic, and biomedical industries. Biosynthesis of these chemicals relies on polymerization of Isopentenyl di-phosphate (IPP) and/or dimethylallyl diphosphate (DMAPP) monomers, which plants synthesize using a cytosolic mevalonic acid (MVA) pathway and a plastidic methyleritritol-4-phosphate (MEP) pathway. Circadian regulation affects MVA and MEP pathway activity at three levels: substrate availability, gene expression of pathway enzymes, and utilization of IPP and DMAPP for synthesizing complex terpenoids. There is a gap in understanding the interplay between the circadian rhythm and the dynamics and regulation of the two pathways. In this paper we create a mathematical model of the MVA and MEP pathways in plants that incorporates the effects of circadian rhythms. We then used the model to investigate how annual and latitudinal variations in circadian rhythm affect IPP and DMAPP biosynthesis. We found that, despite significant fluctuations in daylight hours, the amplitude of oscillations in IPP and DMAPP concentrations remains stable, highlighting the robustness of the system. We also examined the impact of removing circadian regulation from different parts of the model on its dynamic behavior. We found that regulation of pathway substrate availability alone results in higher sensitivity to daylight changes, while gene expression regulation alone leads to less robust IPP/DMAPP concentration oscillations. Our results suggest that the combined circadian regulation of substrate availability, gene expression, and product utilization, along with MVA- and MEP-specific regulatory loops, create an optimal operating regime. This regime maintains pathway flux closely coupled to demand and stable across a wide range of daylight hours, balancing the dynamic behavior of the pathways and ensuring robustness in response to cellular demand for IPP/DMAPP.

Early environmental challenges may remodel murine phenotypes across multiple generations via small noncoding RNA

Flavio Palmieri [1], Marta Mourin-Fernandez [1,2], Signe Skog [3], Violeta de Anca Prado[4], Marcela Parra-Vargas [1,2], Ivonne Palacios-Marin [2], Carles Lerin [2], Carlos Guerrero Bosagna [4], Anita Öst[3], Josep C. Jiménez-Chillarón [1,2]

1. Department of Physiological Sciences, University of Barcelona, c/ Feixa Larga, s/n 08907 L'Hospitalet de Llobregat, Spain
2. Institut de Recerca Sant Joan de Déu, c/ de Sta. Rosa, 39, 08950, Barcelona, Spain
3 Department of Biomedical and Clinical Sciences, Linköping University, SE-581 83 Linköping, Sweden
4. Department of Organismal Biology, Uppsala University, SE-751 05 Uppsala, Sweden Presenter e-mail: flavio.palmieri@ub.edu

Besides random mutations, developmental plasticity can be considered a driving factor of phenotypic adjustments, helping organisms cope with environmental change (Kenkel *et al.*, 2016). However, the mechanisms involved in its generation remain largely uncharacterised in mammals. Here, we explored specific physiological adaptations in mice to an environmental challenge: nutrition.

For that purpose, three consecutive generations of C57BL6 mice were exposed to two nutritional conditions: appropriate and excessive caloric intake during the suckling period. Three-generation follow-up and six-generation long control groups were included at chow-diet (Parra-Vargas *et al.*, 2020).

Samples of sperm and white adipose tissue (WAT) were collected and RNA was extracted and sequenced. Sperm analysis focused on microRNA (miRNA) and transfer RNA fragments (TRF), small non-coding RNAs described as potential epigenetic information carriers (Lee *et al.,* 2022).

Body weight and adiposity exhibited plastic responses to the calorie intake and the metabolic phenotypes were stably inherited by the follow-up group. TRF showed very small variation between groups and the significance did not survive after adjustments, suggesting that their impact seems to be relatively limited. We found several up- and down-regulated miRNAs across generations. Interestingly, some of the target genes of these miRNAs are mainly involved in the regulation of metabolic and lipo-related processes. We confirmed that these targets appeared altered in the offspring's WAT samples (RNA-seq) denoting a possible modification of the germline, in response to environmental challenge.

We found that an obesogenic environment causes both phenotypic changes and, more interestingly, it could also affect the germline. These latter appear to promote phenotypic variations in the following generations and need further validation.

Genetic analysis of mitochondrial dna copy number as an endophenotype

Eduard Molinero[1,2], Ramona N. Pena[1,2], Joan Estany[1,2], Roger Ros-Freixedes[1,2]

1. Departament de Ciència Animal, Universitat de Lleida, Lleida, Catalonia, Spain.
2. Agrotecnio-CERCA Center, Lleida, Catalonia, Spain.
Presenter e-mail: eduard.molinero@udl.cat

Mitochondria play a crucial role in regulating cellular energy metabolism. Mitochondrial DNA copy number (mtDNA_CN) is an endophenotype that serves as an indicator of mitochondrial activity and that can be easily derived from whole-genome sequence data. Overall, the existing evidence suggests that a reduction of mitochondrial activity might be associated with greater feed efficiency and, in a Duroc pig population, we found that mtDNA_CN is correlated with carcass traits, positively with backfat thickness and negatively with loin thickness. The objectives of our study were to uncover the genetic basis of mtDNA_CN, as well as to assess the impact of mitochondrial haplotypes on production traits in pigs. We gathered performance data from 234 pigs and extracted DNA from their skeletal muscle tissue. Whole-genome sequencing was employed to determine mtDNA_CN based on the coverage depths of nuclear and mitochondrial genomes. Sequencing reads were processed using a standard bioinformatics pipeline and aligned to the Sscrofa11.1 reference genome. A genome-wide association study was conducted for mtDNA_CN using GEMMA, applying a linear mixed model that accounted for batch effects and the kinship matrix. Variants with a p-value below 10-5 were considered significant, and we investigated genes located within ± 50 kb of these variants. The GWAS identified 29 significant variants across 12 regions on chromosomes 3, 4, 5, 7, 10, 12, 13, 14, and 17. These regions included candidate genes such as MTHFD2 (rs342882286 ; p-value = 4.7×10^{-6}), which plays a role in the mitochondrial folate pathway, the KRT gene family (rs333838194; p-value = 4.7×10^{-6}), related to mitochondrial homeostasis and function, and HLX (rs340491579; p-value = 4.2×10^{-6}) and NEBL (rs699992004; p-value = 3.8×10^{-6}), both related to mitochondrial accumulation. No association was found between mitochondrial haploblocks and the studied production traits. In conclusion, our findings provide novel insights into the genetic basis of the endophenotype mtDNA_CN, although the genetic variants associated with mtDNA_CN did not translate into noticeable effects on the production traits.

Multiomic integration analysis identifies atherogenic metabolites mediating between novel immune genes and cardiovascular risk

Robert Carreras-Torres,[1,2†] Iván Galván-Femenía,[3,4†] Xavier Farré,[4,5] Beatriz Cortés,[4] Virginia Díez-Obrero,[2,6] Anna Carreras,[4] Ferran Moratalla-Navarro,[2,6,7,8] Susana Iraola-Guzmán,[4,5] Natalia Blay,[4,5] Mireia Obón-Santacana,[2,6,7] Víctor Moreno,[2,6,7,8*] Rafael de Cid [4,5*]

[1]Digestive Diseases and Microbiota Group, Girona Biomedical Research Institute (IDIBGI), Salt, 17190, Girona, Spain.

[2]ONCOBELL Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, 08908, Barcelona, Spain.

[3]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute for Science and Technology, Barcelona, Spain.

[4]Genomes for Life-GCAT lab. CORE Program. Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain.

[5]Grup de Recerca en Impacte de les Malalties Cròniques i les seves Trajectòries (GRIMTra) (IGTP), Badalona, Spain.

[6]Unit of Biomarkers and Susceptibility (UBS), Oncology Data Analytics Program (ODAP), Catalan Institute of Oncology (ICO), L'Hospitalet del Llobregat, 08908, Barcelona, Spain.

[7]Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), 28029, Madrid, Spain.

[8]Department of Clinical Sciences, University of Barcelona, Barcelona, Spain.

* Corresponding authors: rdecid@igtp.cat; v.moreno@iconcologia.net

Understanding genetic-metabolite associations has translational implications for informing cardiovascular risk assessment. Interrogating functional genetic variants enhances our understanding of disease pathogenesis and the development and optimization of targeted interventions.

In this study, a total of 187 plasma metabolite levels were profiled in 4,974 individuals of European ancestry of the GCAT| Genomes for Life cohort. Results of genetic analyses were meta-analyzed with additional datasets, resulting in up to approximately 40,000 European individuals. Results of meta- analyses were integrated with reference gene expression panels from 58 tissues and cell types to identify predicted gene expression associated with metabolite levels. This approach was also performed for cardiovascular outcomes in three independent large European studies (N=700,000) to identify predicted gene expression additionally associated with cardiovascular risk. Finally, genetically informed mediation analysis was performed to infer causal mediation in the relationship between gene expression, metabolite levels and cardiovascular risk.

A total of 44 genetic loci were associated with 124 metabolites. Lead genetic variants included 11 non-synonymous variants. Predicted expression of 53 fine-mapped genes was associated with 108 metabolite levels; while predicted expression of 6 of these genes was also associated with cardiovascular outcomes, highlighting a new role for regulatory gene *HCG27*. Additionally, we found that atherogenic metabolite levels mediate the associations between gene expression and cardiovascular risk. Some of these genes showed stronger associations in immune tissues, providing further evidence of the role of immune cells in increasing cardiovascular risk.

These findings propose new gene targets that could be potential candidates for drug development aimed at lowering the risk of cardiovascular events through the modulation of blood atherogenic metabolite levels.

Single-neuron deep generative model uncovers underlying physics of neural firing and connectivity

Jordi Abante[1-6], Angelo Piga[7,8], Clara F. López-León[7,8],Josep M. Canals[1,3-6], Jordi Soriano[7,8]

1 Dept. of Biomedical Sciences, Universitat de Barcelona, Barcelona, Spain 2 Centre for Genomic Regulation, Barcelona, Spain
3 Institute of Neurosciences, Universitat de Barcelona, Barcelona, Spain
4 Creatio, Production and Validation Center of Advanced Therapies, Universitat de Barcelona, Barcelona, Spain
5 Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS)
6 Institut de Neurociències UB (UBNeuro), Universitat de Barcelona, Barcelona, Spain 7 Dept. of Condensed Matter Physics, Universitat de Barcelona, Barcelona, Spain
7Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona, Spain

Presenter e-mail: jordi.abante@ub.edu

Calcium imaging has become a powerful alternative to electrophysiology for studying neuronal activity, offering spatial resolution and the ability to measure large populations of neurons in a minimally invasive manner. This technique has broad applications in neuroscience, neuroengineering, and medicine, allowing researchers to explore the relationship between neuron location and activity. Recent advancements in deep generative models (DGMs) have facilitated the modeling of neural population dynamics, uncovering latent representations that offer insights into behavior prediction and neural variance. However, these models typically rely on spike inference algorithms and are focused on population-level dynamics. To address this gap, we propose a novel framework for single-neuron representation learning using autoregressive variational autoencoders (AVAEs). This approach embeds individual neurons' spatiotemporal signals into a reduced-dimensional space without the need for spike inference algorithms. Our model enables tasks such as visualization, clustering, and data compression, while also capturing underlying connectivity patterns between neurons. Using both simulated and real data, we demonstrate that the proposed AVAE approach achieves competitive reconstruction performance, offers significant data compression capabilities, and produces a superior representation to standard linear approaches. Our model lays out the foundation for future work directed at integrating RNA-seq and Ca imaging single-cell data of neural populations.

Poster Sessions Program and Abstracts

Venue: Hall one level below the main conference room

12th of December

Poster set-up: 11:30 – 11:50 downstairs from the main conference hall

Poster Session 1: 16:00 – 17:30

Posters with coffee break and networking
Meet the company: Almirall

13th of December

Poster Session 2: 10:50 – 11:30

Posters with coffee break and networking

WHEN TO STAND BY YOUR POSTER	
ODD-NUMBERED POSTERS	FIRST HALF OF EACH SESSION
EVEN-NUMBERED POSTERS	SECOND HALF OF EACH SESSION

POSTER 1: Integrative Modeling of Synthetic Biology Interventions in Maize and Rice: From Gene Expression to Whole-Plant Phenotypes, **Rui Alves**

POSTER 2: Understanding the role of the RNA-binding protein Staufen 2 during neurogenesis using single cell transcriptomics, **Akshay Jaya Ganesh**

POSTER 3: Large-scale meta-analysis of plasma cell-free RNA-Seq methods, **Cristina Tuñí i Domínguez**

POSTER 4: Lipidomics meets Bioinformatics: Overcoming Barriers in Lipid Analysis, **Elia Obis**

POSTER 5: Genetic polymorphisms lead to major, locus-specific, variation in piRNA production in mouse, **Adrià Mitjavila-Ventura**

POSTER 6: Biological properties and pharmacological implications of Cysteine Rich Peptides from Mediterranean Plants, **Julia Lisa-Molina**

POSTER 7: Three-dimensional simulations of host-microbiota and microbiota-microbiota interactions, **Nikolai S. Bykov**

POSTER 8: Biofunctional and GANGO: Advancing Omics Data Interpretation Through Functional and Network Analysis, **Xavi Tarragó**

POSTER 9: BPGA: a Shiny app to perform Basic Population Genetic Analysis, **Joan Fibla**

POSTER 10: Genomic risk prediction for common disease in rural Pyrenean populations, **Marina Laplana**

POSTER 11: Brain EV Proteomes provide novel molecular clues for the major prevalence of dementias in subjects diagnosed with schizophrenia, **Jose Antonio Sánchez Milán**

POSTER 12: Climate acts as an environmental filter to plant pathogen community assembly in Europe, **Maria Caballol**

POSTER 13: Consumption of tomatoes with distinct (poly)phenolic hallmarks promotes significant changes in aminoacidic and lipidic metabolisms in a photoperiod-dependent manner, **Eduardo Gallardo-Baena**

POSTER 14: Implementation of a scalable cloud server for bioinformatics in hospital settings, **Maria Berman-Riu**

POSTER 15: DNA sequence variation in the porcine glp1r gene and adiposity, **Gerard Otin**

POSTER 16: Development of an automated pipeline oriented to capture-based viral metagenomics, **Maria Tarradas-Alemany**

POSTER 17: Complete characterization of human polymorphic inversions and other complex variants from long read data, **Ricardo Moreira Pinhal**

POSTER 18: Exploring Structural Variability and Repeat Dynamics in the *Physcomitrium patens* Pangenome, **Marc Pulido**

POSTER 19: Systems Biology and Proteomics to Unveil Common Pathways Associated with Oxidative Stress and Amyloidosis in Cardiovascular Diseases and Cognitive Decline, **Patricia M. Bota**

POSTER 20: Forestforward: a platform to understand plant biodiversity changes, **Eva Tejada**

POSTER 21: Exploring functional conservation in silico: a new machine learning approach to RNA-editing, **Michal Zawisza-Alvarez**

POSTER 22: Development of image alignment methods for spatial transcriptomics data of spinal cord, **Víctor A. Gaya-Martín**

POSTER 23: Automating data management and analyses in sequencing facilities using a django-based lims, a custom api and nextflow pipelines, **Toni Hermoso**

POSTER 24: Genomic Insights into the Evolutionary Nature of Autism Spectrum Disorder, **Ariadna**

Bada-Navarro

POSTER 25: SNPeBoT: A method for predicting Transcription Factor Allele Specific Binding, Patrick Gohl

POSTER 26: Long-term inversion recurrence and segmental duplication conservation in mammal genomes, **Maria Diaz-Ros**

POSTER 27: Infering phylodinamics of bacterial epidemics using a bayesian multi-type birth- death model, **Jordi Cabrera-Gumbau**

POSTER 28: DNA markers associated with the host response to porcine reproductive and respiratory syndrome, **Houda Laghouaouta**

POSTER 29: Effect of nasal microbiota at weaning on lung lesions caused by an experimental inoculation with *Mycoplasma hyopneumoniae,* **Fernando Moreira Petri**

POSTER 30: Systems Biology and Proteomics to Unveil Common Pathways Associated with Oxidative Stress and Amyloidosis in Cardiovascular Diseases and Cognitive Decline, Patricia M. Bota

POSTER 31: GSVA 2.0: Pathway-centric analysis at single-cell and spatial resolution, Robert Castelo

POSTER 32: Analysis of the imputability of human inversions using different phasing and imputation software, Adrià Mompart

POSTER 33: Mathematical analysis on a cam skeleton model: exploring the underlying causes for cam rhythmicity, Dario Herrera

POSTER 1: Integrative Modeling of Synthetic Biology Interventions in Maize and Rice: From Gene Expression to Whole-Plant Phenotypes

Rui Alves[1,2,3,*], Jorge Comas[3], Oriol Basallo[1,2,3], Abel Lucido[1,2,3], Ester Vilaprinyo[1,2,3], Alberto Marin-Sanguino[1,2,3] & Albert Sorribas[1,2,3]

[1]Systems Biology Group, Department Ciències Mèdiques Bàsiques, Faculty of Medicine, Universitat de Lleida, Lleida, Spain
[2]IRBLleida, Lleida, Spain
[3]Grup de Recerca Consolidat de la Generalitat de Catalunya 2021 SGR 1353 MathSy2Bio

Presenter e-mail: rui.alves@udl.cat

Synthetic biology offers powerful means to enhance the nutritional and agricultural value of staple crops like maize and rice by engineering metabolic pathways. Our work leverages mathematical modeling to bridge the gap between molecular interventions and broader phenotypic outcomes, unraveling the complexity of these modifications.

In maize, targeted transcriptomic and metabolomic profiling, combined with modeling, identified temporal bottlenecks in the carotenoid biosynthetic pathway [1,2]. This approach pinpointed genetic intervention points to optimize carotenoid production in the maize endosperm, underscoring the need for multigene engineering. Data-driven models further refined these strategies, predicting both general and line-specific engineering tactics to maximize carotenoid yields, highlighting the importance of tailored approaches for each genetic background. Additionally, we modeled the biosynthesis of strigolactones [3]—key hormones in root development—integrating these models with 3D simulations of maize root architecture [4]. This integration allowed us to predict how altering strigolactone levels through synthetic biology could improve root system architecture, enhancing nutrient uptake and resilience to environmental stressors and parasites.

In rice, our focus was on manipulating isoprenoid biosynthesis to increase the production of terpenoid precursors, crucial for various high-value compounds. By introducing an ectopic MVA pathway, we aimed to boost IPP/DMAPP production while managing interactions with native pathways [5–7]. Mathematical models were developed to integrate molecular and macroscopic data, predicting how these synthetic pathways affect metabolite concentrations and plant traits such as seed morphology and hormone balance. These models provided critical insights into the impact of these interventions, offering a predictive framework for optimizing metabolic engineering strategies.

Together, these studies demonstrate the power of mathematical modeling in synthetic biology, enabling a comprehensive understanding of how genetic modifications translate into desired phenotypic outcomes in maize and rice. The models become important tools for predicting the effects of synthetic biology interventions across different scales, paving the way for more efficient and predictable crop engineering.

POSTER 2: Understanding the role of the RNA-binding protein Staufen 2 during neurogenesis using single cell transcriptomics

Akshay Jaya Ganesh[1,2†], Sandra María Fernández-Moya[1†‡], Ana Gutiérrez-Franco[1], Natalie Chaves Ferreira[1], Damià Romero-Moya[1], Loris Mularoni[1], Monika Piwecka[3], Alessandra Giorgetti[1,4], Agnieszka Rybak-Wolf[5] and Mireya Plass[1,4]

[1] Program for Advancing Clinical Translation of Regenerative Medicine of Catalonia, P-CMR[C], Regenerative Medicine Program, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, Barcelona, Spain

[2] Doctoral Program in Biomedicine, Universitat de Barcelona, Spain

[3] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

[4] Center for Networked Biomedical Research on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain.

[5] Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine, Berlin, Germany

† These authors contributed equally to this work

‡ Present address: Department of Basic Sciences, Iron metabolism: Regulation and Diseases, Universitat Internacional de Catalunya (UIC), Sant Cugat del Vallès, Barcelona, Spain

Neurogenesis is a crucial process involving the formation of new neurons in the developing cortex of embryos, regulated by multiple factors including RNA-binding proteins (RBPs). Staufen 2 (STAU2) is an RBP implicated in the asymmetric distribution of mRNAs in radial glial cells, thereby regulating the balance between neural stem cell maintenance and differentiation. However, the molecular mechanisms of STAU2-mediated regulation in human neurogenesis remain largely unknown. To investigate the role of STAU2, we first generated STAU2 KO iPSC lines using CRISPR-Cas9 technology. We differentiated the iPSCs to neural cell types and profiled the impact of STAU2 KO at multiple timepoints during the differentiation (day 0, 11, 25, 55 and 70) using single-cell transcriptomics (scRNA-seq). In our data, we identified cell types characteristic of the different timepoints including iPSCs, progenitor cells, and mature neurons and astrocytes. Using differential expression (DE) analysis, we discovered that most DE genes between STAU2 KO and control cells were found in the neuroepithelial cell cluster. Gene Set Enrichment Analysis of neuroepithelial cells, primarily found at day 11, showed significant upregulation of oxidative phosphorylation and glycolysis pathways in STAU2 KO. Such metabolic shifts have been previously linked to the transition between progenitor and differentiated cells, suggesting a faster differentiation of neuroepithelial cells towards neurons upon STAU2 KO. We confirmed the accelerated differentiation in STAU2 KO using 2D cell cultures, which have a larger proportion of cells with mature neuronal marker expression and morphology, and cortical organoids, which present alterations in the neuroepithelial cell organisation, faster neural differentiation and smaller size. To identify potential drivers of this phenotype, we performed gene regulatory network (GRN) analysis. Our results showed significant downregulation of the GRNs controlled by CHD2 and ARID3A transcription factors in STAU2 KO neuroepithelial cells, potentially driving the accelerated differentiation observed. Considering that CHD2 and ARID3A are known targets of STAU2 and have been previously shown to regulate the differentiation of stem cells, we propose a model where STAU2 controls human neurogenesis at the neuroepithelial stage by modulating the expression of these two transcription factors.

POSTER 3: Large-scale meta-analysis of plasma cell-free RNA-Seq methods

Cristina Tuñí i Domínguez[1], Lluc Cabús[1], Phil Sanders[1], Marc Weber[1], Julien Lagarde[1]

[1] Flomics Biotech SL, Carrer de Roc Boronat, 31, 08005 Barcelona, Spain

Presenter e-mail: cristina.tuni@flomics.com

The use of plasma cell-free RNA (cfRNA) as a diagnostic biomarker in liquid biopsies, while promising, presents important technical challenges. These include the acquisition of high-quality, diverse cfRNA molecules, mitigating genomic DNA contamination in sequencing libraries, and the optimisation of bioinformatics pipelines. In addition, there is currently no gold-standard workflow for the profiling of plasma cfRNA. A systematic large-scale comparison of cfRNA-Seq experimental workflows is thus needed to address these issues. In this study, we present a first-of-its-kind comprehensive meta-analysis of all publicly available plasma cfRNA-Seq datasets, which we re-analyzed using a uniform bioinformatics pipeline. We systematically compare various analytical metrics across 2,700 samples from a comprehensive, annotated list of 19 published peer-reviewed cfRNA-Seq studies. Our results show how details of the experimental protocol influence cfRNA-Seq data quality in often drastic ways, leading to considerable variability and batch effects across studies. Such a benchmark contributes to the advancement of cfRNA-based diagnostic methods and sets the stage for future improvements and standardizations of cfRNA profiling workflows.

POSTER 4: Lipidomics meets Bioinformatics: Overcoming Barriers in Lipid Analysis

Obis, Elia*; Roig, Vicent; Oronich, Arnau; Ramírez, Omar; Martín, Meritxell

Plataforma Lipidòmica de Catalunya (PLICAT) – Lipidomics Unit – Institut de Recerca Biomèdica de Lleida (IRBLLeida)
*Presenting and corresponding author

The rapid expansion of lipidomics, driven by advancements in high-resolution mass spectrometry, has unveiled the immense complexity of lipidomes and their roles in cellular processes. However, translating this complexity into meaningful biological insights requires addressing critical challenges in data processing, analysis, and integration.

Key hurdles include correcting signal bias introduced by sample injection order and mitigating batch effects across diverse analytical runs, both of which are essential to ensure data reliability. Furthermore, the accurate identification of lipid species is hampered by their structural diversity and isobaric overlaps, requiring advanced lipid-centric search tools and comprehensive databases. Efficient management of metadata, coupled with scalable computational solutions, is vital to process and analyze increasingly large datasets.

In addition, the field demands improved methodologies for the analysis and visualization of results, enabling researchers to interpret complex lipidomics data effectively. Lastly, integrating lipid knowledge into biochemical pathways and interactive maps is essential to connect lipidomics findings with broader biological contexts.

Here, we present bioinformatics solutions that address these challenges, combining data correction models with analytical and visualization tools. This comprehensive approach not only enhances the precision and usability of lipidomics data but also supports its integration into systems biology, paving the way for new discoveries and applications in health and disease research.

POSTER 5: Genetic polymorphisms lead to major, locus-specific, variation in piRNA production in mouse

Adrià Mitjavila-Ventura[1,2,3], Tanya Vavouri[2]

1. Germans Trias i Pujol Research Institute (IGTP), Can Ruti Campus, 08916 Badalona, Spain
2. Regulatory Genomics group, Josep Carreras Leukaemia Research Institute (IJC), Campus ICO-Germans Trias i Pujol, 08916 Badalona, Spain.
3. PhD Program in Bioinformatics, Universitat Autònoma de Barcelona (UAB), 08193 Cerdanyola del Vallès, Spain.
Presenter e-mail: amitjavila@carrerasresearch.org

PIWI-interacting RNAs (piRNAs) are small noncoding RNAs that silence transposons in the animal germline. They are produced from long single-stranded transcripts, as well as transposons. Despite some sites producing piRNAs are conserved syntenic regions, piRNA sequences and the loci producing them evolve rapidly and vary a lot across and within species, much more than other functional regions in the genome. To unravel the sequence changes that contribute to the fast evolution of piRNA, we analysed transcription and piRNA expression of piRNA-producing loci between the germline genetically distinct male mice, five inbred and one outbred strain. We found that genetic variation underlies most of the piRNA expression variation between genetically different mice, and report that polymorphic insertions of endogenous retroviruses (ERVs) are associated with significant differences in piRNA production in several loci including protein-coding genes. Our findings provide evidence that transposable elements contribute to inter-individual differences in expression, and potentially to the fast evolution of piRNA-producing loci in mammals.

POSTER 6: Biological properties and pharmacological implications of Cysteine Rich Peptides from Mediterranean Plants

Julia Lisa-Molina[1,2,3 *], María Mulet[1,2], Jose Antonio Sánchez-Milán[1,2], Itziar Molet i Arias[1,2], Aida Serra[1+], Xavier Gallart-Palau[2+]

1. Department of Basic Medical Sciences, University of Lleida (UdL) – Biomedical Research Institute of Lleida (IRB Lleida) - +Pec Proteomics Research Group (+PPRG) - Neuroscience Area, 25198 Lleida, Spain.
2. Biomedical Research Institute of Lleida Dr. Pifarré Foundation (IRBLLEIDA) - +Pec Proteomics Research Group (+PPRG) - Neuroscience Area – University Hospital Arnau de Vilanova (HUAV) – Health Campus, Psychology Dept., University of Lleida, Spain.
3. Algèmica Barcelona S.L. Camí Sant Bartomeu 5, 08600 Berga (Barcelona), Spain.

+ Correspondence: Xavier Gallart-Palau, PhD. xgallart @irblleida.cat ; Aida Serra, PhD. aida.s erra@udl.cat .

*Presenter: Julia Lisa-Molina jlm 1 3 @alumnes .ub.edu

The Mediterranean basin is among the richest plant biodiversity hotspots in the World, with a vast history of traditional uses of medicinal plants (1,2). Nonetheless, countless bioactive principles of said plants (especially peptidic ones) remain to be comprehended (3). This study focuses on a subset of Bioactive Peptides, the Cysteine-Rich Peptides (CRPs), that hold promising applications in the food and pharmaceutical industries because of their general specificity, safety and stability (3; 4). The Plant Kingdom constitutes a riveting source for their isolation (5), and even though a handful of species are known to produce CRPs (6,7), many remain to be explored and delved into. In order to broaden the knowledge of CRPs synthesized by Mediterranean plants and prospect for possible new local medicinal products, 3921 species belonging to 27 different Mediterranean plant families that could potentially produce CRPs were identified. Afterwards, a bioinformatics sequence-pattern-based approach was used to screen transcriptomic, genomic, and proteomic data available in public repositories and to predict potential CRP sequences associated with the mentioned plant species (53 plants have been currently screened, 31 have the potential of producing CRPs, and 118392 cysteine-rich sequences have been found). The predicted CRP sequences of 11 species have been empirically evaluated using chromatographic techniques along with Mass Spectrometry-driven proteomics leading to identifying at least 6 CRPs. Furthermore, metabolomic and proteomic analyses are currently being performed to evaluate their bioactivity through *in vitro* assays, and stability as well as cytotoxicity studies are being conducted to assess their potential scalability and translation to the market. The data that is being generated has therefore the potential to enhance existing compound libraries and identify novel nutraceutical and therapeutic candidates for comprehensive screening in a wide range of biotechnological and biomedical applications.

References

1. Gonzalez-Tejero, M.R., et al. J Ethnopharmacol 2008;116(2):341-357.
2. Ramos-Gutiérrez, I., et al. Global Ecol Biogeogr 2021;30(10):1951-1957.
3. Meena, S., et al. Biol Futur 2020;71(3):195-208.
4. Lavergne, V., Taft, R.J. and Alewood, P.F. Curr Top Med Chem 2012;12(14):1514-1533.
5. Fan, H., et al. Journal of Future Foods 2022;2(2):143 -159.
6. Zorin, E.A., et al. Front Plant Sci 2022;13:884726
7. Tam, J.P., et al. Sci Rep 2018;8(1):16201.

POSTER 7: Three-dimensional simulations of host-microbiota and microbiota-microbiota interactions

Nikolai S. Bykov[1], Bryan Wang[2], Zoe Horisberger[2], Marta Contreras[3], Carlotta Pietroni[3], Urvish Trivedi[2], Antton Alberdi[3], and Marc A. Marti-Renom[1,4,5]

1. CNAG, National Center for Genome Analysis, Barcelona, Spain
2. Department of Biology, Section of Microbiology, University of Copenhagen, Denmark
3. Center for Evolutionary Hologenomics, Globe Institute, University of Copenhagen, Denmark
4. Center for Genomic Regulation, The Barcelona Institute for Science and Technology, Barcelona, Spain
5. ICREA, Barcelona, Spain

Understanding how host tissues and microbes interact is crucial for improving animal production efficiency. Conventional multi-omic approaches have limitations in revealing these interactions due to their inability to capture the spatial arrangement of relevant biological and chemical components. In the 3D'omics consortium, we aim to address this challenge by developing a new technology to generate three-dimensional data of the intestinal ecosystem and a computational framework that will enable the reconstruction of host-microbiota interactions based on this data with unprecedented spatial resolution. As a part of this framework, we introduce *Gorgona*, a tool for simulating the three-dimensional architecture of host tissues and bacteria based on the generated 3D'omics data. By running simulations from a random state of the environment, *Gorgona* reveals experimentally testable interaction patterns within the microbiota. In summary, *Gorgona* provides a visual and computational perspective to study the spatial distribution of microbiota in host tissues.

POSTER 8: Biofunctional and GANGO: Advancing Omics Data Interpretation Through Functional and Network Analysis

Xavi Tarragó[1], Alejandro Rodríguez[1], Antonio Monleón[1]

1. BIOST3 Section of Statistics. Departament of Genetics, Microbiology and Statistics. Edifici Aulari (2n Floor) Faculty of Biology. University of Barcelona Avda Diagonal 645, 08028 Barcelona (Spain)

Biofunctional is a Shiny-based bioinformatics tool developed to enhance the funcional analysis of omics data, focusing on gene ontology (GO) terms and KEGG pathways. It addresses limitations in existing tools by enabling multi-factor, multi-level analysis across distinct experimental groups, such as various disease stages (e.g., onset, treatment phases). Biofunctional integrates visualization features, showing dynamic relationships and regulation states (up/down/neutral) between ontologies and pathways, while enrichment analysis (EA-value) signifies the strength of associations for different disease stages (Rodriguez & Monleon Getino, 2024). The tool automatically generates barplots and heatmaps, facilitating visual interpretation of enrichment patterns, particularly useful in understanding disease mechanisms like intestinal inflammation related to the microbiota. Biofunctional's intuitive interface, combined with AI for result interpretation, supports comprehensive biomedical research applications.

In addition, we have also developed GANGO, an innovative algorithm designed to interpret metagenomic and metatranscriptomic data (Monleon-Getino et al., 2020). It is based on the ecological concept of consortia, which refers to biologically interconnected groups. By utilizing clustering network analysis and genetic ontologies, GANGO allows the identification and interpretation of complex ecological networks. It helps reveal the relationship between taxa/genes, the number of groups, their relations and their functionalities using the annotated genes of an organism in a database (e.g. UniProt or Ensembl). This tool significantly improves our understanding of the functional implications of metagenomic data, facilitating the interpretation of intricate biological networks.

References:

Rodríguez, A. & Monleón Getino, A. 2024. bioRxiv. https://doi.org/10.1101/2024.10.08.616405

Monleón-Getino, A., *et al.* 2020. bioRxiv. https://doi.org/10.1101/2020.06.10.140103

POSTER 9: BPGA: a Shiny app to perform Basic Population Genetic Analysis

Joan Fibla

Secció Biomedicina, Instiut d'Estudis Ilerdencs (IEI). Plaça de la Catedral, s/n, 25002
Lleida
Presenter e-mail: joan.fibla@gmail.com

This Shiny app provides an interactive platform for visualizing population genetic structure. Users can upload their own datasets or use example files to explore population relationships, genetic diversity, and ancestral components. The app dynamically reads PLINK binary files and integrates them with worldwide reference populations from the 1000 Genomes Project (1000G) and the Human Genome Diversity Project (HGDP), enabling basic population analyses such as Principal Component Analysis (PCA), ADMIXTURE, and FST analyses.

The app performs PCA for dimensionality reduction, producing scaterplots that reveal population clusters. ADMIXTURE analysis is employed to assess population ancestry components, with a geographic visualization feature that displays these components as pie charts on a world map—each slice representing an ancestral proportion, and chart sizes scaled according to population size. Finally, FST analysis enables users to quantify genetic differentiation between pairs of populations.

Designed for both researchers and students in population genetics, the app offers a user-friendly interface for exploring genetic structure through multiple methods. It facilitates interactive data exploration and produces publication-ready plots. Overall, the app integrates PCA, ADMIXTURE, and FST analyses, providing a comprehensive view of population dynamics and evolutionary relationships.

Code available at : https://github.com/jfibla/BPGA-a-Shiny-app-to-perform-Basic-Population-Genetic-Analysis.git

POSTER 10: Genomic risk prediction for common disease in rural Pyrenean populations

Marina Laplana [1,2], Joan Fibla [2]

[1] Departament CMB, Universitat de Lleida - IRBLleida

[2] Secció Biomedicina, Institit d'Estudis Ilerdencs

Presenter e-mail: marina.laplana@udl.cat

Polygenic risk scores (PRS) provide valuable insights into the genetic predisposition to complex diseases by aggregating the effects of multiple genetic variants. The Genetic Project of the Pyrenees (GenPIR) focuses on characterizing the genetic variability of rural populations in the Catalan Pyrenees, a unique model for studying how isolation, environmental pressures, and historical events shape genetic structure. These populations, shaped by geographic isolation due to geo-historical barriers, exhibit distinct genetic substructures that may influence disease susceptibility.

To investigate these influences, we analyzed genotype data from 395 individuals across rural Pyrenean communities. The dataset was expanded through statistical imputation to over 10 million single nucleotide polymorphisms (SNPs), allowing for a comprehensive examination of genetic variation. By utilizing curated risk markers from the PGS Catalog and genotype data from the Iberian Peninsula population as a reference, we evaluated the distribution of PRS for several diseases.

Our analysis revealed significant differences in the mean PRS z-scores between Pyrenean and Iberian Peninsula populations for common diseases and cancers, including breast, prostate, and thyroid cancers, as well as Alzheimer's disease. Moreover, we identified geographic patterns in genetic risk, with statistically significant variations in PRS by latitude and longitude for thyroid cancer and Alzheimer's disease.

These findings are made accessible through an interactive Shiny web service, enabling users to visualize PRS distributions and facilitating comparisons between populations. The study of rural populations like those in the Pyrenees provides key insights into the genetic factors that influence health, with implications for targeted healthcare interventions.

Available at: https://pyrprs.shinyapps.io/Shiny_GENPIR_prs_interpolENG/

POSTER 11: Brain EV Proteomes provide novel molecular clues for the major prevalence of dementias in subjects diagnosed with schizophrenia

Jose Antonio Sánchez Milán [1,2+], Itziar Molet Arias [1,2], María Mulet [1,2], Julia Lisa [1], Aida Serra [2,*], Xavier Gallart-Palau [1,*]

[1] Biomedical Research Institute of Lleida Dr. Pifarré Foundation (IRBLLEIDA) - +Pec Proteomics Research Group (+PPRG) - Neuroscience Area – University Hospital Arnau de Vilanova (HUAV), 80 Av. Rovira Roure, 25198, Lleida, Spain

[2] Department of Medical Basic Sciences, University of Lleida (UdL), +Pec Proteomics Research Group (+PPRG), 25198, Lleida, Spain

+ Presenter: Jose Antonio Sánchez Milán *Correspondence: aida.serra@udl.cat & xgallart@irblleida.cat

The high prevalence of age-associated dementias affecting individuals with psychotic disorders has been epidemiologically observed; however, a molecular mechanism linking psychoticism and age-associated dementias remains to be elucidated. Extracellular Vesicles (EVs) are small membranous structures that facilitate intercellular communication. While their roles have been extensively studied in neurological diseases, their potential implications in the neuropathology underlying human psychiatric disorders and their connection to dementias remain mostly unexplored. In this study, using advanced bioinformatics and discovery-driven proteomics, we have analyzed the brain EVs proteome of patients with psychotic disorders and patients with early stages of Alzheimer's Disease (AD), Vascular Dementia and Controls. Our results indicate that a total of 32 brain EVs proteins share a common expression pattern between patients with psychiatric disorders and patients with Age-related Dementias. Functional characterization of these proteins revealed that many of these proteins are involved in functions related to brain development, axonogenesis, regulation of synaptic signaling and immune response. These interesting results could explain the molecular mechanisms altered in psychiatric diseases that lead to the development of dementia, with the aim of identifying new therapeutic targets and biological markers in these neurodegenerative diseases.

POSTER 12: CLIMATE ACTS AS AN ENVIRONMENTAL FILTER TO PLANT PATHOGEN COMMUNITY ASSEMBLY IN EUROPE

Maria Caballol[1,2], Miguel Ángel Redondo[3], Núria Catalán[4], Tamara Corcobado[5,6], Thomas Jung[5], Benoît Marçais[7], Ivan Milenković[5,8], Miguel Nemesio-Gorriz[9], Jan Stenlid[3], Jonàs Oliva[1,2]

1. Department of Agricultural and Forest Sciences and Engineering, University of Lleida, 25198, Lleida, Spain
2. Joint Research Unit CTFC – AGROTECNIO-CERCA, 25198, Lleida, Spain
3. Department of Forest Mycology and Plant Pathology, Swedish University of Agricultural Sciences, Box 7026, 750 07, Uppsala, Sweden
4. Institute of Environmental Assessment and Water Research, IDAEA-CSIC, 08034, Barcelona, Spain
5. Mendel University in Brno, Faculty of Forestry and Wood Technology, Department of Forest Protection and Wildlife Management, Phytophthora Research Centre, 613 00, Brno, Czech Republic
6. Austrian Research Centre for Forests (BFW), Seckendorff-Gudent-Weg 8, 1131, Vienna, Austria
7. Université de Lorraine - Inrae, UMR Interactions Arbres/Microorganismes, 54000, Nancy, France
8. University of Belgrade-Faculty of Forestry, 11030, Belgrade, Serbia
9. Forestry Development Department, Teagasc, D15DY05, Dublin, Ireland

Presenter e-mail: maria.caballol@udl.cat

Understanding how climate influences the distribution of plant pathogens is paramount given the upraising effects of climate change. While this has been studied for other plant-associated microbes, such as bacteria and mycorrhizal and endophytic fungi, the role of climate on oomycetes has been rarely explored from a community ecology perspective. Oomycetes include the genus *Phytophthora* which comprises some of the most damaging plant pathogens of agricultural, forestry and natural ecosystems. We explored the role of climate in the assembly of *Phytophthora* species at >250 river sites across two gradients, a latitudinal gradient spanning from Mediterranean to Arctic conditions, and an altitudinal gradient including the Spanish Pyrenees. *Phytophthora* communities were obtained by sequencing the internal transcribed spacer (ITS) region of river filtrates. The diversity and functional diversity of *Phytophthora* communities were correlated with environmental variables. Signs of environmental filtering were obtained by performing randomisation analysis. Climate, rather than other environmental factors such as geography or tree diversity, determined *Phytophthora* biogeography. Two key processes determined species assembly. In southern latitudes, a hot dry climate posed an environmental filter for *Phytophthora* communities resulting in communities dominated by drought-tolerant species with thick oospores and high optimum temperature for growth. In northern regions, winter temperatures acted as an environmental filter on *Phytophthora* community assembly, selecting species adapted to survive low minimum temperatures. This study showed how functional traits related to morphological and physiological characteristics of *Phytophthora* species explained community assembly processes. Future research could also include genomic data as functional traits to study whether certain gene ontology families could explain pathogen community assembly.

POSTER 13: Consumption of tomatoes with distinct (poly)phenolic hallmarks promotes significant changes in aminoacidic and lipidic metabolisms in a photoperiod-dependent manner

Eduardo Gallardo-Baena[1], Antonio J. Cortés-Espinar[1,2,3], Manuel Suarez[1,2,3], Begoña Muguerza[1,2,3]

[1]*Universitat Rovira i Virgili, Departament de Bioquímica i Biotecnología, Grup de Recerca en Nutrigenòmica, Tarragona, Spain*

[2]*Institut d'Investigació Sanitària Pere Virgili (IISPV), Reus, Spain*

[3]*Center of Environmental, Food and Toxicological Technology (TecnATox), Tarragona, Spain*

Presenter e-mail: eduardo.gallardo@estudiants.urv.cat

Seasonal rhythms, driven primarily by the photoperiod as key environmental signals, are cyclic physiological and behavioural changes which help organisms adapt to fluctuating conditions and maintain homeostasis. Increasing evidence shows that (poly)phenols contained in fruits may also regulate them. In this regard, (poly)phenol-rich diets are associated with an improvement in health due to their anti-inflammatory and antioxidant properties. However, it remains unclear whether fruits with distinctive (poly)phenolic compositions can differentially affect the serum metabolome in a photoperiod-dependent manner. Therefore, 72 male Fischer 344 rats were fed for 11 weeks a standard chow diet supplemented with either a vehicle (VH), or with two tomatoes with the same amount of (poly)phenols but distinct profile (namely, Tomato 1 and Tomato 2). In addition, animals were exposed to three distinct photoperiods: L6 (6 h light/day), L12 (12 h light/day) and L18 (18 h light/day) to mimic the different seasons of the year. Serum metabolomic analysis was conducted under untargeted approach using NMR, and the resulting data were processed using MetaboAnalyst. Results showed that consumption of both tomatoes induced alterations in the amino acid metabolism, as evidenced by increased serum levels of certain amino acids, such as phenylalanine and serine, predominantly in L6. On the other hand, only the consumption of Tomato 2 reduced the levels of lipidic species, including triglycerides and cholesterol, primarily in L18. Our results demonstrate that the consumption of tomatoes with distinctive (poly)phenol profiles induces metabolic changes in a photoperiod-dependent manner. Nevertheless, further studies are still needed to unravel the molecular mechanisms associated with these effects.

Acknowledgements: This study was funded by European Regional Development Fund of the European Commission though the Operative Program ERDF of Catalonia 2014–2020 (PECT-NUTRISALT) and by Ministerio de Ciencia e Innovación: PID2021-128813OB-I00, MCIN/AEI/10.13039/501100011033/.

A.J.C-E is recipient of the grant 2023 INV-2 00033 funded by the European Union, NextGenerationEU, and the Generalitat de Catalunya through the Martí i Franquès Postdoctoral Fellowship from the Universitat Rovira i Virgili.

POSTER 14: IMPLEMENTATION OF A SCALABLE CLOUD SERVER FOR BIOINFORMATICS IN HOSPITAL SETTINGS

Maria Berman-Riu[1], Pep Mut[2], Josep Muncunill[1]

[1]Plataforma de Genómica i Bioinformàtica, Institut d'Investigació Sanitària de les Illes Balears (IdISBa), Palma

[2]Plataforma de Recerca amb Informació Sanitària de les Illes Balears (PRISIB), Institut d'Investigació Sanitària de les Illes Balears (IdISBa), Palma

Presenter e-mail: maria.berman@ssib.es

In response to the growing demand for computational resources in bioinformatics, particularly in hospital environments with limited IT infrastructure, we implemented a scalable server using Amazon Web Services (AWS).Traditional physical servers entail high costs and maintenance burdens, making them unsustainable in the long term. Cloud computing, by contrast, offers a flexible, scalable, pay-as-you-go model, enabling efficient adaptation to evolving bioinformatics needs without infrastructure constraints. AWS was selected for its versatility, scalability, and robust security protocols. We deployed an EC2 (Elastic Compute Cloud) virtual server, which supports three predefined instances: one for off-peak (low-demand usage), another for moderate computational tasks (e.g., microarrays analysis), and a third for high-performance requirements (e.g., RNA-seq). Amazon S3 (Simple Storage Service) was used for scalable and secure data storage and retrieval, ideal for backups and data analysis without overloading the instances.

Currently, the server supports 27 active users, enabling programming in R and Python, dynamic report visualization, and the execution of Shiny applications. Usage monitoring allows us to assess the need for instance adjustments while providing detailed tracking of individual user activity.

AWS provides an efficient solution to the limitations of physical servers, offering real-time scalability, cost control, and a secure environment that enhances productivity in bioinformatics research.

POSTER 15: DNA SEQUENCE VARIATION IN THE PORCINE *GLP1R* GENE AND ADIPOSITY

Gerard Otín[1], Ramona N. Pena[1], Roger Ros-Freixedes[1], Joan Estany[1]
1. Departament de Ciència Animal, Universitat de Lleida – Agrotecnio-CERCA Center, 191 Rovira Roure, 25198 Lleida, Catalunya, Spain
Presenter e-mail: gerard.otin@udl.cat

The glucagon like peptide 1 receptor (GLP1R) has recently become notorious in human medicine for its relationship with satiety regulation and obesity treatment. Using the pig as a model, in the present work we explored the impact of sequence variation of the GLP1R gene on adiposity. With this purpose, we (1) looked into sequence variation in the GLP1R gene (±0.5 Mb) and (2) investigated the association of identified variants with fat-related traits. A total of 436 whole-genome sequenced pigs (311 purebred Duroc and 125 of other genetic types) were used. We found 969 variants in the GLP1R region, three of which were missense. An association analysis of the identified variants was conducted for body weight, backfat and loin thickness, and intramuscular fat content and fatty acid composition collected in 276 Duroc individuals from the same line. No association was found between the three missense mutations and the target traits, but associations involving 26 other variants were detected ($p < 0.01$). Most of them (24) were with intramuscular fat content (1 upstream, 11 intronic, 1 synonymous and 4 downstream) and loin thickness (3 upstream and 4 intronic). Unlike missense mutations, which segregated at a very low frequency (< 3 %), the allele frequency of the associated candidate variants was relatively high (> 15 %). These results suggest that sequence variation in the GLP1R gene may potentially be associated with fatness in pigs. The most promising variants are currently being validated with all the pigs registered (~3,000) in our (UdLGIM) biobank.

POSTER 16: DEVELOPMENT OF AN AUTOMATED PIPELINE ORIENTED TO CAPTURE-BASED VIRAL METAGENOMICS

Maria Tarradas-Alemany[1,2], Sandra Martínez-Puchol[2], Cristina Mejías-Molina[2,3], Marta Itarte[2,3], Marta Russiñol[2,3] Sílvia Bofill-Mas[2,3], Josep F Abril[1]

1. Computational Genomics Lab., Department of Genetics, Microbiology and Statistics, Universitat de Barcelona (UB); Institut de Biomedicina UB (IBUB).
2. Laboratory of Viruses Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, Universitat de Barcelona (UB).
3. The Water Research Institute (IdRA), Universitat de Barcelona (UB).

Presenter e-mail: mtarradasi@ub.edu

Target Enrichment Sequencing or Capture-based metagenomics has emerged as useful approach to solve some of the major challenges of viral metagenomics in complex samples, such as the low concentration of viral sequences or the sample biases that may appear in non-clinical environments. These datasets are usually analyzed with standard downstream Bioinformatics protocols, which leads to unnecessary resources consumption and huge undirected results reports. CAPTVRED (*Capture-based metagenomics Analysis Pipeline for tracking ViRal species from Environmental Datasets*), has been designed to analyze the virome in complex samples, focused on datasets obtained by Target Enrichment Sequencing approach. This tool provides a user-friendly resource that complements this semi-directed sequencing approach for the total or partial virome description, especially from environmental matrices. The pipeline resulted in a user-friendly, customizable tool which achieves time and computational cost reduction. It provides a setup module, a pre- filtering stage to discard non-viral sequences, data cleaning, assembly, taxonomic classification and optional contamination discarding. Unlike other available protocols, CAPTVRED offers the flexibility to adjust almost any parameter at each step, making it adaptable to the unique characteristics of viral metagenomic datasets. Results are integrated and gathered into an HTML report providing comprehensive, reproducible, and accessible results. The virome from a set of samples retrieved from sewage and bat guano have been already analyzed with this pipeline; moreover, sequences obtained by whole-genome sequencing and target enrichment sequencing approaches have been also considered, in order to assess the performance of the capture panel, as well as for the pipeline. The results show an increased number of assigned viral contigs in the capture approach, which also recalls higher coverage and similarity with respect to reference sequences of potentially zoonotic viruses.

POSTER 17: Complete characterization of human polymorphic inversions and other complex variants from long read data

Ricardo Moreira Pinhal[1,2], Konstantinos Karakostis[1,2], Illya Yakymenko[1,2], Maria Diaz Ros[1,2], Andrés Santos[3], Jaime Martinez-Urtaza[3], Marta Puig[1,3], Mario Cáceres[1,2,4]

1. Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain.
2. Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Barcelona, Spain.
3. Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain.
4. ICREA, Barcelona, Spain.

Inversions are a special type of structural variants whose study has lagged behind due to their balanced nature and the presence of large inverted repeats (IR) at their breakpoints. New techniques are finally making possible to identify the full spectrum of human inversions. However, in most cases, just a limited number of individuals has been studied, which precludes the analysis of the effects of the detected variants. Here, we take advantage of long reads to generate and characterize an exhaustive catalogue of IR-mediated inversions.We have developed a computational pipeline that uses Oxford Nanopore Technologies (ONT) long read data to genotype inversions. Using this method, we have interrogated 614 candidate inversions from different studies, ranging from 265 bp to 4.4 Mb and flanked by up to 190 kb long IRs, in a diverse set of 108 individuals.We detected both orientations in 180 inversions, validating 126 novel inversions, with frequencies ranging between 1-50%. We also showed that ONT genotypes were highly accurate, matching perfectly previous experimental genotypes of 54 inversions. Importantly, 282 additional SVs were identified on the predictions, revealing the complexity of repeat-rich regions. Finally, some of these regions show discrepancies with the new Pangenome reference, which indicates that they might not be well resolved.Long reads therefore have a great potential for the characterization of currently missed inversions and other complex genomic regions in multiple individuals, opening the door to determine their real functional impact.

POSTER 18: Exploring Structural Variability and Repeat Dynamics in the *Physcomitrium patens* Pangenome

Marc Pulido[1], Pol Vendrell-Mir[2], Pierre-François Perroud[3], Fabien Nogué[3], Leandro Quadrana[2], Josep M Casacuberta[1]

1. Centre for Research in Agricultural Genomics, CRAG (CSIC- IRTA-UAB-UB), Campus UAB, Cerdanyola del Vallès, Barcelona, Spain
2. Institute of Plant Sciences Paris-Saclay, Centre Nationale de la Recherche Scientifique, Institute National de la Recherche Agronomique, Université Evry, Université Paris-Saclay, Orsay, France
3. Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay, 78000 Versailles, France

The moss *Physcomitrium patens* is a model plant species widely used in evolutionary studies due to its phylogenetic position as a basal clade among land plants. Its unique genome organisation is characterized by evenly distributed repetitive sequences and retrotransposon-rich centromeres. In this project, we aim to construct the first pangenome for *P. patens* using long-read assemblies from more than 20 natural accessions. This will allow us to investigate intraspecific variation, focusing specifically on the evolutionary dynamics of LTR-retrotransposons, which likely play a pivotal role in shaping chromosome structure and centromeres.

POSTER 19: Systems Biology and Proteomics to Unveil Common Pathways Associated with Oxidative Stress and Amyloidosis in Cardiovascular Diseases and Cognitive Decline.

Patricia M. Bota, Melisa E. Zeylan, Saja Almabhouh, Simge Senyuz, Ruben Molina- Fernandez, Narcis Fernandez-Fuentes, Pol Picon-Pagès, Anna Garcia-Elias, Marta Tajes, Eduard Barbu, Raul Vicente, Stanley Nattel, Angel Ois , Albert Puig-Pijoan, Jordi Garcia- Ojalvo, Ozlem Keskin, Attila Gursoy, Francisco J. Muñoz and Baldomero Oliva.

Department of Medicine and Life Sciences, (SBI-GRIB), Universitat Pompeu Fabra, Barcelona 08003, Catalonia, Spain

This study investigates the molecular landscape of cellular aging and its implications for cognitive impairment, Alzheimer's disease (AD), and type 2 diabetes through a systems biology approach. We employed an in vitro model of astrocytes subjected to Aβ1-42 oligomers and H_2O_2 to simulate neurodegenerative environments. Proteomic profiling via mass spectrometry identified key proteins significantly affected by these treatments. We constructed protein interaction networks and compared them with established networks related to cognitive impairment, AD, type 2 diabetes, cardiovascular diseases, oxidative stress, and inflammation. Gene regulatory networks influenced by single nucleotide polymorphisms (SNPs) associated with AD were analyzed to pinpoint candidate disease genes. Our findings elucidate the altered protein interactions and regulatory mechanisms underlying aging and disease, highlighting potential therapeutic targets and biomarkers for age-related cognitive disorders.

POSTER 20: Forestforward: a platform to understand plant biodiversity changes

Emilce Soledad Martinez[1,2], Eva Luz Tejada-Gutiérrez[1], Rui Alves[1]

1. University of Lleida & Biomedical Research Institute (IRB), Avda Alcalde Rovira Roure nº80 · 25198 Lleida, Catalonia, Spain
2. National Institute of Agricultural Technology (INTA), Agricultural Experimental Station (EEA) RP32 km 4,5, Pergamino, Buenos Aires, Argentina.
Presenter e-mail: martinez.emilce@inta.gob.ar

Climate change is increasingly affecting global plant biodiversity, with consequences for ecosystem services and human well-being. Global Strategy for Plant Conservation promotes the development of online platforms with plant information in open access at a global scale (Convention on Biological Diversity, 2012). In response to these challenges, the ForestForward platform was developed by Tejada-Gutiérrez *et al.* (2022). The platform integrates over 4,000 historical forest census datasets, offering a comprehensive tool for analyzing global plant biodiversity dynamics. Here we illustrate how ForestForward enables large scale plant biodiversity studies. By using the platform and its data, we estimated Species Richness and Shannon index and analyze biodiversity changes over the years (from 1950 to 2020), and across the world. Our analysis finds an initial increase in species richness, followed by a decline post-1990, suggesting climate-driven biodiversity loss. Furthermore, using the Shannon diversity index we identify priority areas for the development and implementation of strategies for plant biodiversity conservation. ForestForward provides quality data to understand changes in plant biodiversity over time and geography. The platform is useful for plant biodiversity research and ecology modeling studies.

POSTER 21: Exploring functional conservation in silico: a new machine learning approach to RNA-editing

Michal Zawisza-Alvarez[1], Jesus Peñuela-Melero[1], Esteban Vegas[1], Ferran Reverter[1], Jordi Garcia Fernandez[1], Carlos Herrera-Ubeda[1]

1. Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain
Presenter e-mail: m.zawisza@ub.edu

Around 50 years ago, molecular biology opened the path to understand changes in forms, adaptations, complexity or the basis of human diseases, trough myriads of reports on gene birth, gene duplication, gene expression regulation and splicing regulation, among other relevant mechanisms behind gene function. Here, with the advent of big data and artificial intelligence (AI), we focus on an elusive and intriguing mechanism of gene function regulation, RNA editing, in which a single nucleotide from an RNA molecule is changed with a remarkable impact in the increase of the complexity of transcriptome and proteome. We present a new generation approach to assess the functional conservation of the RNA-editing targeting mechanism using two AI learning algorithms, random forest (RF) and bidirectional long short-term memory (biLSTM) neural networks with attention layer. These algorithms combined with RNA- editing data coming from databases and variant calling from same-individual RNA and DNA-seq experiments from different species, allowed us to predict RNA-editing events using both primary sequence and secondary structure. Then, we devised a method for assessing conservation or divergence in the molecular mechanisms of editing completely in silico: the cross-training analysis. This novel method could set the basis to understand the conservation of the editing mechanism through evolution.

POSTER 22: DEVELOPMENT OF IMAGE ALIGNMENT METHODS FOR SPATIAL TRANSCRIPTOMICS DATA OF SPINAL CORD

Víctor A. Gaya-Martín[1], Quique Vidal-Beneyto[1], Sonia Tarazona[2], Ana Conesa[1]

1. Genomics of Gene Expression Lab, Institute for Integrative Systems Biology, Spanish National Research Council (I2SysBio, CSIC-UV), Carrer del Catedràtic Agustín Escardino Benlloch, 46980 Paterna, Valencia, Spain
2. BiostatOmics Group, Departamento de Estadística e Investigación Operativa Aplicadas Y Calidad, Universitat Politècnica de València (UPV), Camí de Vera, s/n, 46022 Valencia, Spain
Presenter e-mail: vagayam11@gmail.com

Spatial transcriptomics (ST) based on array mRNA capture is an innovative technique that enables the detection and quantification of gene expression in spatially resolved tissue samples. By using spatial location barcoding and Next Generation Sequencing (NGS), ST allows researchers to map gene expression to specific tissue regions, providing insights into tissue organization and cell-to-cell communication in biological contexts such as embryo development, tumours, and injuries.

Spinal cord injury (SCI) is a serious condition that affects mobility; although recovery is challenging, it is sometimes possible under certain conditions. ST methods can be used to study SCI by helping to understand the changes that occur during injury and regeneration, potentially offering insights for therapeutic targets. However, in ST experiments, tissue sections may differ in size, shape, or spatial layout, making comparison difficult since the ST coordinates on two slides may not capture equivalent tissue regions. While various alignment methods and algorithms have been developed to address this issue, many combine alignment to a common coordinate system with gene expression data integration into a single dataset.

In our study, we proposed a pairwise alignment strategy that uses only the image data, without gene expression, to enable direct comparison of specific regions across different tissue slices and samples. We tested three different approaches to align four similar slices from the same patient and four distinct tissue samples: 1) Geometric Transformation Estimation Model: This approach involves manually selecting landmarks and estimating transformation parameters manually. 2) Procrustes Transformation: This method also uses manual landmark selection but calculates transformation parameters automatically with published R packages. 3) ImageJ plugin Register Virtual Stack Slices: This tool automatically selects landmarks and calculates transformations.

The alignment results varied between methods and images, but the Geometric Transformation Estimation Model generally produced good outcomes. This approach successfully enabled the selection of regions of interest, facilitating future comparative analyses between samples and slices.

POSTER 23: Automating data management and analyses in sequencing facilities using a django-based lims, a custom api and nextflow pipelines

Anna Delgado[1], Luca Cozzuto[1], Clemente Borges[1], Toni Hermoso[1], Kathrin Meindl[1], Rodny Hernandez[1], Emyr James[1], Berta Fuste[1,2], Julia Ponomarenko[1,2]

1. Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology, Doctor Aiguader 88, 08003 Barcelona, Catalonia, Spain
2. Universitat Pompeu Fabra (UPF), 08003 Barcelona, Catalonia, Spain
Presenter e-mail: toni.hermoso@crg.eu

The increasing complexity and volume of data generated in sequencing facilities necessitate robust solutions for data management and analysis. We present an integrated approach to automating data processing, analysis, and management by combining a Django-based Laboratory Information Management System (LIMS) with a Nextflow pipeline. The LIMS facilitates efficient tracking and management of samples, projects, and instrument runs, while the Nextflow pipeline automates the execution of demultiplexing, quality control, and data transfer. By integrating these systems, we achieved seamless data flow, ensuring reproducibility and reducing manual errors. Furthermore, leveraging containerization and HPC computing with Nextflow enhances the scalability and portability of the procedures. This combined approach provides a comprehensive solution for sequencing facilities, streamlining both data management and computational processes, ultimately accelerating biomedical research and discoveries.

POSTER 24: Genomic Insights into the Evolutionary Nature of Autism Spectrum Disorder

Ariadna Bada-Navarro[1]*, Carla Casanova[2]*, Selena Aranda[1], Dora Koller[1], Ditte Demontis[3], Anders Borglum[3,4], Bru Cormand[1,5]&, Marina Mitjans[1,6]&, Oscar Lao[2]&

[1]Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Catalonia, Spain; Institut de Recerca Sant Joan de Déu (IRSJD), Esplugues de Llobregat, Catalonia, Spain; Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain

[2]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Catalonia, Spain

[3]iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Aarhus, Denmark; Department of Biomedicine, Aarhus University, Aarhus, Denmark; Center for Genomics and Personalized Medicine, Aarhus, Denmark

[4]Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark

[5]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBER-ER), Madrid, Spain

[6]Centro de Investigación Biomédica en Red de Salud Mental (CIBER-SAM), Madrid, Spain

Presenter e-mail: carla.casanova@ibe.upf-csic.es

*Equally contributed &Equally supervised

Given the impairing nature and reproductive drawbacks of Autism Spectrum Disorder (ASD), its high prevalence in the population poses an evolutionary paradox. In principle, natural selection acts on removing alleles that decrease fitness of the carriers – i.e. risk alleles with large effects on predisposition to mental illnesses. However, a substantial fraction of ASD genetic risk is attributable to a polygenic predisposition driven by common variation of small to moderate effect size. Thus, we aimed to investigate the evolutionary nature of ASD from a genomic perspective using different state-of-the-art evolutionary genomic tools and taking advantage of the most recent ASD-GWAS meta-analysis including 38,717 cases and 232,735 controls (unpublished).

We used MAGMA to test whether ASD-associated variants are enriched in LoF-intolerant genes (Leeuw et al. 2015). We tested whether ASD-risk alleles (stratified by MAF bins) tend to be enriched for the derived or ancestral state (allele present in the chimpanzee (*Pan Troglodytes*)). Moreover, we investigated the load of ASD-risk alleles in ancient genomes from the Allen Ancient DNA Resource (AADR) (Mallick et al. 2024).

MAGMA analyses suggest that LoF-intolerant genes (ExAC pLI≥0.9) are enriched with ASD-associated variants (p=6.64e-08). This enrichment persists in LoF-intolerant brain-expressed (FPKM≥1, p=2.99e-07) and highly-brain-expressed (FPKM≥5, p=5.44e-07) genes. Our results indicate that the ancestral state generally confers protection against autism susceptibility compared to the derived state. Furthermore, trend analysis of ancient DNA samples revealed a statistically significant increase in the proportion of ASD-risk alleles towards the present (R=-0.16, p=2.4e-06).

These results, together with new genomic approaches that we are currently conducting using archaic genomes will give us some clues about the evolutionary nature of ASD, providing a new dimension for understanding this disorder. Further analyses to study the pleiotropy of ASD-associated variants with other phenotypes will help us to unravel the selective forces shaping ASD.

POSTER 25: SNPeBoT: A method for predicting Transcription Factor Allele Specific Binding

Patrick Gohl[1], Baldomero Oliva[1]

1. Department of Medicine and Life Sciences, (SBI-GRIB), Universitat Pompeu Fabra, Barcelona 08003, Catalonia, Spain.
Presenter e-mail: patrick.gohl@upf.edu

We created SNPeBoT (Single Nucleotide Polymorphism effect on Binding of Transcription Factors) a tool to predict the effects of Single Nucleotide Polymorphisms (SNP) on transcription factor binding. Allele specific binding (ASB) data from ChIP-seq experiments were paired with high sequence-identity motifs assessed in Protein Binding Microarray experiments. For each transcription factor a paired motif was selected from which we derived E-score profiles for reference and alternate DNA sequences of ASB events. A Convolutional Neural Network was trained to predict whether these profiles were indicative of ASB gain/loss or no change in binding. More than 23,000 E-score profiles from various transcription factors were split into train, validation and test data. Upon benchmarking we demonstrate an improvement of predictive performance when compared to state of the art tools. SNPeBoT was made available on a web server and as a standalone package. Next we show how SNPeBoT may be applied to Genome Wide Association Studies (GWAS) to discern which SNP disease associations are relevant in target disease etiology through the disturbance of transcription factor regulation. For this a pipeline was created which takes as input a disease term and retrieves GWAS associations with information on their regulators. Using SNPeBoT, the pipeline is then able to label the SNPs that affect transcription factor binding and predict disease associated genes affected by this change in regulation. Finally, we applied a data mining approach to provide the user with potential explanations for the role of the affected gene in target disease etiology.

POSTER 26: Long-term inversion recurrence and segmental duplication conservation in mammal genomes

Maria Diaz-Ros[1,2], Mario Cáceres[1,2,3]

1 Research Program on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Barcelona, Spain

2 Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

3 ICREA, Barcelona, Spain

Segmental duplications (SDs) are highly identical duplicated DNA fragments longer than 1 kb, which contribute to recurring rearrangements through non-allelic homologous recombination (NAHR) and are a potential source of phenotypic diversity and disease. However, they are difficult to sequence and assemble reliably and, therefore, they are one of the least known regions of the genome. Recently, thanks to new long-read sequencing techniques and ongoing projects focused on building high-quality reference genomes, it is possible to investigate SD conservation and evolution across the mammalian phylogeny for the first time. Here, by analyzing a set of validated human NAHR inversions in 13 diverse mammals, we found conservation of the flanking repeats in 18 out of 44 inversions (41%) and independent inversion recurrent events for eight of them. Next, we extended the analysis to 1164 human SD pairs throughout the genome, and assessed their evolution in 41 species with a wide range of divergence times. This allowed us to estimate the origin of each pair of SDs and the long-term rate of gains and losses. SD conservation beyond primates was 6.6% in autosomes and 29.1% in chromosome X. Thus, these results help to identify the most conserved regions and their potential functional implications.

POSTER 27: Infering phylodinamics of bacterial epidemics using a bayesian multi-type birth- death model

Jordi Cabrera-Gumbau[1,2], Jaime Martínez-Urtaza[1]

1. Department of Genetics and Microbiology, Autonomous University of Barcelona, Barcelona, Spain.
2. Institute of Biotechnology and Biomedicine, Autonomous University of Barcelona, Barcelona, Spain.

Presenter e-mail: jordi.cabrera@uab.cat

The structured coalescent model has been employed in several occasions to model migration rates and events between the different epidemic locations. However, the primary assumptions of the coalescent model are not well-suited to epidemic dynamics. For instance, epidemical populations are known to have overlapping generations but Coalescent models assumes there is no overlap at all. Another important assumption is the constant population size that the coalescent models rely on. Here we will try the multi-type birth-death model (BDMM), a forward-in-time model mainly thought to be used with infectious diseases, that has already been used to infer the phylogeography of viral epidemics but not widely in bacteria.

To be able to compare the results of the structured coalescent model and the BDMM we will use the Bayesian framework BEAST2 with the core alignment of our bacterial genomes. We will set our research in the cholera epidemic that hit Latin America during the 90's decade. Vibrio cholera is a bacterium with a high recombination/mutation rate known to live in water-like sources like rivers or the coastline, with a rare spread from person to person which leads these bacteria to a perpetual cycle between the host and the environment. These characteristics among others make it extremely difficult to study. Several studies have tried to answer the unknowns surrounding this epidemic with greater or lesser success. With this new approach we aim to understand how these bacteria arrived to the Peruvian coast, how it expanded throughout the continent, and which are the main elements that in one hand let them conquer this diverse territory in just two years and in the other led them to disappearance in the following six.

POSTER 28: DNA markers associated with the host response to porcine reproductive and respiratory syndrome

Houda Laghouaouta[1,2], Lorenzo J Fraile[1,2], Romi Pena[1,2]

1. Departament de Ciència Animal, Universitat de Lleida, Lleida, Catalonia, Spain.
2. Agrotecnio-CERCA Center, Lleida, Catalonia, Spain.
Presenter e-mail: houda.laghouaouta@udl.cat

Porcine Reproductive and Respiratory Syndrome (PRRS) is a highly challenging viral disease that causes major losses in the porcine sector. The genetic component of the host response to PRRS is well-evidenced, indicating that depending on their genetics, pigs may be more resilient or susceptible to the disease. Thus, the objective of this work is to investigate the molecular basis of the host response to the PRRS virus by identifying associated DNA markers and candidate genes. Previous research in our group revealed that response of young females to an attenuated vaccine can be used as a proxy for the host response to PRRS. Females with no viremia 7 and 21 days after vaccination are considered resilient, with an improved host response to PRRS. In contrast, females positive at 7 and 21 days after vaccination are considered susceptible. Therefore, viremia at 7 and 21 days was used as a proxy to assess the host response to PRRS. A whole genome-wide sequencing-based association study was performed using 128 sows (66 resilient and 62 susceptible) and 7,795,669 variants to identify DNA markers associated with the classification. A total of 35 variants and 11 genomic regions were found associated in pig chromosomes 1, 4, 10, 13 and 16. Candidate genes within the associated genomic regions are mainly related to the immune response, such as Fc receptors (*FCRL1*, *FCRL3*, *FCRL4*, and *FCRL5*), CD1 molecules (*CD1A*, *CD1B*, *CD1D*, and *CD1E*), and the Eomesodermin gene (*EOMES*). These findings corroborate the critical role of the immune system in the modulation of the host response to virus infection. We are currently exploring bioinformatic pipelines for effective prioritization of underlying variants in the most relevant candidate genes.

POSTER 29: Effect of nasal microbiota at weaning on lung lesions caused by an experimental inoculation with *Mycoplasma hyopneumoniae*

Fernando Moreira Petri[1], Gerard Ale[2,3], Virginia Aragón[2,3], Luís Guilherme de Oliveira[1], Marina Sibila[2,,3]*, Florencia Correa-Fiz[2,,3]*

1. São Paulo State University (Unesp), School of Agricultural and Veterinary Sciences, Jaboticabal, Brazil.
2. IRTA, Animal Health, Centre de Recerca en Sanitat Animal (CReSA), Campus de la Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Catalonia, Spain
3. Unitat mixta d'investigació IRTA-UAB en Sanitat Animal, Centre de Recerca en Sanitat Animal (CReSA), Campus de la Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Catalonia, Spain.
*Equally contributing last authors
Presenter e-mail: fernando.moreira@irta.cat

Mycoplasma hyopneumoniae causes enzootic pneumonia in pigs worldwide, primarily affecting cranio-ventral lung areas. The variable severity seen in these lung lesions might be related to different factors, being the nasal microbiota one of them. This study explores the impact of nasal microbiota on lung lesion development and extension in *M. hyopneumoniae* experimental infection. Twelve piglets were challenged with 5mL (10e6 CCU/ml) of *M. hyopneumoniae* via endotracheal route; nasal swabs were collected at weaning (D0), at challenge (D21), and two weeks post-challenge (D35). Samples were submitted to Illumina sequencing of *16S* rRNA gene and analyzed using Quantitative insights into microbial ecology (Qiime2, version 2023.9). Animals were classified based on the percentage of affected lung tissue observed at necropsy on D56 as good (above median) or bad (below median) responders. At D0, the nasal microbiota of bad responders showed lower richness (Chao1, observed features indexes, $p<0.05$) and different diversity (Bray Curtis, $p<0.05$) when compared to the good ones. At D35, the differential analyses (ANCOM-BC) showed that *Mycoplasma* genus was associated with bad responders while *Lactobacillaceae* family was enriched in the good ones. These findings suggest that the nasal microbiota at weaning may have an important role in *M. hyopneumoniae* experimental infection outcome.

POSTER 30: Systems Biology and Proteomics to Unveil Common Pathways Associated with Oxidative Stress and Amyloidosis in Cardiovascular Diseases and Cognitive Decline

Patricia M. Bota[1], Melisa E. Zeylan [2], Saja Almabhouh[1], Simge Senyuz[2], Ruben Molina- Fernandez[1], Narcis Fernandez-Fuentes[1], Pol Picon-Pagès[3], Anna Garcia-Elias[3], Marta Tajes[3], Eduard Barbu[4], Raul Vicente[4], Stanley Nattel[6], Angel Ois[7] , Albert Puig-Pijoan[7], Jordi Garcia- Ojalvo[8], Ozlem Keskin[2], Attila Gursoy[2], Francisco J. Muñoz[5] and Baldomero Oliva[1]

1 - Structural Bioinformatics Group, Research Programme on Biomedical Informatics, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Catalonia 08003, Spain
2- Computational Sciences and Engineering, Graduate School of Science and Engineering, Koç University, Istanbul, Türkiye
3. Laboratory of Molecular Physiology, Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona 08002, Spain
4.Institute of Computer Science, University of Tartu, Tartu, 50090, Estonia
5. Departamento de Medicina y Ciencias de la Vida Laboratorio de Fisiología Molecular, Universitat Pompeu Fabra, Barcelona, Catalonia 08003, Spain
6. Department of Medicine and Research Center, Montreal Heart Institute and Université de Montréal; Institute of Pharmacology, West German Heart and Vascular Center, University Duisburg-Essen, Germany; IHU LIRYC and Fondation Bordeaux Université, Bordeaux 33000, France
7. Department of Neurology, Hospital Del Mar. Hospital Del Mar - Medical Research Institute and Universitat Pompeu Fabra, Barcelona 08003, Spain
8. Laboratory of Dynamical Systems Biology, Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona 08002, Spain
Presenter e-mail: petrislp@gmail.com

This study investigates the molecular landscape of cellular aging and its implications for cognitive impairment, Alzheimer's disease (AD), and type 2 diabetes through a systems biology approach. We employed an in vitro model of astrocytes subjected to Aβ1-42 oligomers and H_2O_2 to simulate neurodegenerative environments. Proteomic profiling via mass spectrometry identified key proteins significantly affected by these treatments. We constructed protein interaction networks and compared them with established networks related to cognitive impairment, AD, type 2 diabetes, cardiovascular diseases, oxidative stress, and inflammation. Gene regulatory networks influenced by single nucleotide polymorphisms (SNPs) associated with AD were analyzed to pinpoint candidate disease genes. Our findings elucidate the altered protein interactions and regulatory mechanisms underlying aging and disease, highlighting potential therapeutic targets and biomarkers for age-related cognitive disorders.

POSTER 31: GSVA 2.0: Pathway-centric analysis at single-cell and spatial resolution

Axel Klenk[1], Pere Moles[1], Andrew J. Sedgewick[2], Justin Guinney[2], Robert Castelo[1]

1. Department of Medicine and Life Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain
2. Tempus Labs, Inc., Chicago, IL.
Presenter e-mail: robert.castelo@upf.edu

GSVA (https://bioconductor.org/packages/GSVA) is an R/Bioconductor package that enables pathway-centric analyses of data produced by high-throughput molecular profiling technologies. The interpretation of biological findings from such data is one of the cornerstones of biomedical research, and GSVA facilitates that goal by performing a conceptually simple but powerful change in the functional unit of analysis, from genes to gene sets. Here we describe our efforts to adapt GSVA to data produced at single-cell and spatial resolution, increasing its robustness and scalability, and improving the user interface and documentation. By using pathways instead of gene-centric features, this new version of GSVA contributes to improve exploratory data analysis, and to develop lower-dimensional statistical and machine learning models, on data from single-cell and spatial transcriptomics experiments.

POSTER 32: Analysis of the imputability of human inversions using different phasing and imputation software

Adrià Mompart[1], Illya Yakymenko[1,2], Mario Cáceres[1,2,3]

1. Research Program on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Barcelona, Spain
2. Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain
3. ICREA, Barcelona, Spain
Presenter e-mail: amompart@researchmar.net

Abstract

Inversions are genomic structural variants (SVs) where a DNA segment gets reversed, which have been linked to many phenotypic traits and adaptations in both humans and other species. Depending on their formation mechanisms, they can be divided into homologous and non-homologous inversions. However, the study of inversions is hindered by the difficulties in their genotyping, especially for homologous inversions due to the presence of large inverted repeats at the breakpoints and their recurrence. Imputation has been widely used as an alternative to infer genotypes of missing variants with great success, but it has been mainly limited to simple variants. Therefore, the extent to which current imputation software can impute inversions accurately remains to be determined. Here, we benchmark four widely-used imputation tools – Beagle, Impute2, Impute5 and Minimac4 – and ScoreInvHap to assess their ability to impute a subset of human polymorphic inversions, most of which have been generated by homologous mechanisms. By taking advantage of the 1000 Genomes Project high-coverage data and the accurate experimental genotypes from the InvFEST project, we have examined the overall and per inversion imputation accuracy of each tool using whole genome sequences and microarrays with different SNP density. The results suggest an overall good inversion imputation performance for all methods, with up to ~65% of inversions being imputable in different human populations from sequence data and slightly less from SNP arrays. In particular, Minimac4 and Impute5 show higher imputation accuracy and less individual loss with respect to the other methods. Also, we have found that sample size per inversion and the application of posterior genotype probability filtering are important factors for inversion imputation accuracy. Thus, our findings provide insights into the optimal conditions for imputation software application and demonstrate the potential of these methods to improve inversion genotype prediction, contributing to determine the functional impact of these little characterized variants.

POSTER 33: Mathematical analysis on a cam skeleton model: exploring the underlying causes for cam rhythmicity

Darío Herrera-García[1], Asdrúbal Burgos[1], Rui Alves[2]

1. Laboratorio de Biotecnología, CUCBA, Universidad de Guadalajara, Mexico
2. Institute of Biomedical Research of Lleida, IRBLleida, Lleida, Spain

Presenter e-mail: braulio.herrera@alumnos.udg.mx

Although the biochemistry of CAM metabolism has been studied since the mid-twentieth century, several aspects of its regulation remain ill understood. One of these aspects is the regulation of circadian intracellular malate oscillations (IMO). In CAM plants, malate accumulates in the vacuole at night when stomata are open, allowing external CO_2 uptake. During the day, malate is exported to the cytoplasm. Remarkably, malate oscillations still occur in the absence of a day-night cycle, when keeping plants under constant conditions. The control of autonomous CAM malate oscillations has been primarily attributed to two mutually exclusive mechanisms: an oscillator at the tonoplast, which changes the vacuolar state from accumulating to expelling malate, and PEPC phosphorylation, which increases PEPC activity during the night (Nungesser *et al.*, 1984; Nimmo, 2000). However, both hypotheses have extensive experimental and theoretical support. Here we use mathematical modeling and analysis to understand the influence of both mechanisms on autonomous IMO. We expanded and integrated previously developed CAM mathematical models into a single model that considers tonoplast oscillations, PEPC phosphorylation, circadian regulation, and regulation of malate transport by pH. We used mathematically controlled comparisons to identify the intrinsic effects of each mode of regulation on the dynamic behavior of IMO. Preliminary results indicate that PEPC phosphorylation contributes to a minor extent to the oscillations compared to the vacuole transporter. The influence on pH on the vacuolar oscillator is minimal too. With this work we attempt to shed light on the extent of how these mechanisms contribute to CAM circadian rhythms, and, in a broader perspective, on how CAM evolved from a C3 pathway.

References

Nimmo, G. A., Wilkins, M. B., Fewson, C. A., and Nimmo, H. G. (1987). Persistent circadian rhythms in the phosphorylation state of phosphoenolpyruvate carboxylase from *Bryophyllum fedtschenkoi* leaves and in its sensitivity to inhibition by malate. *Planta*, 170(3), 408–415. doi:10.1007/bf00395034

Nungesser, D., Kluge, M., Tolle, H. and Oppelt, E. (1984). A dynamic computer model of the metabolic and regulatory processes in Crassulacean acid metabolism.*Planta*,(162):204-214