aviesan
alliance nationale
pour les sciences de la vie et de la santé
ITMO Génétique, Génomique, Bioinformatique

FRANCE GENOMIQUE

International Symposium on Human Genomics

September 20-22, 2023

CNRS Headquarters, 3 rue Michel Ange, Paris, France

Book of Abstracts

PacBio
MGI
sptlabtech
SCIPIO BIOSCIENCE
Oxford NANOPORE Technologies
illumina
ACTIVE MOTIF Enabling Epigenetics Research
vizgen
Parse BIOSCIENCES
STANDARD BIOTOOLS
Olink
OZYME Des femmes et des hommes au service de vos recherches
10x GENOMICS

Table of contents

Oral Presentations

Posters

7000 years of settlement in France seen through the prism of paleogenomics: diffusion, migration and past societies

Mélanie Pruvost*†[1]

[1]De la Préhistoire à l'Actuel : Culture, Environnement et Anthropologie – Université de Bordeaux, Centre National de la Recherche Scientifique – France

Abstract

Thanks to consistent technological advancements, paleogenomics has gradually imposed itself over the last ten years as an essential discipline for studying past populations. Initially focused on problematics at the European scale, research in this field has confirmed the major population movements theorized by archaeologists through the study of material cultures. Consequently, it has become possible, for example, to validate the connection between the introduction of agriculture in Europe and the migration of Neolithic communities from Anatolia. Due to the growing strength of statistical techniques applied to ancient genomic data, it is currently feasible to delve into inquiries concerning intra-community matters like consanguinity, exogamy rates, biological kinship, and matrimonial systems. In this context, we will present the data collected during the ANR project "Ancestra", which encompasses information from more than 800 individuals coming from three regions of France: Hauts de France, Grand Est, and Occitanie. This dataset spans from the Mesolithic era to the High Middle Age. Its substantial size will enable us to discuss the diffusion and migration of populations in France in correlation with the cultural and technological transformations that occurred over a 7000-year period. We will also explore how paleogenomics can shed light on aspects of the structure of historical societies and the interplay between different groups in terms of both social and biological dynamics, by examining specific archaeological sites. Finally, we will talk about the potential and limitations of paleogenomics in addressing questions about gene selection and health.

Keywords: human paleogenomics, past populations

*Speaker

†Corresponding author: melanie.pruvost@u-bordeaux.fr

Paleogenomics of the peopling of western Europe

Eva-Maria Geigl *† [1], Oguzhan Parasayan , E. Andrew Bennett , Thierry Grange‡

[1] Institut Jacques Monod – Centre National de la Recherche Scientifique, Université Paris Cité – France

Major human migration waves in ancient times into Europe occurred from the southeast to the west. The territory of present-day France being at the most western end of the European continent, this was the area where they came to a natural end. The analysis of the DNA from individuals belonging to different time periods and areas enables the detailed reconstruction of these migration waves and their correlation with climate crises. We will present our large-scale paleogenomic studies with hundreds of genomes from the Upper Paleolithic to the Bronze Age. These genomes reveal the origins far in the East and the complex genesis of the population who in France produced the most stunning Gravettian art. They characterize genomically both the last autochthonous hunter-gatherers in France and the people who brought from the Middle East the Neolithic life style based on agriculture and husbandry. Finally, these ancient genomes unveil the population history of the Neolithic between 5000 and 2000 BCE detailing migration waves and admixture pulses which, beyond leading to the emergence of new cultures, led to the shaping of the present-day European genome.

Keywords: ancient DNA, paleogenomics, peopling, evolution, imputation, phasing

*Speaker

†Corresponding author: eva-maria.geigl@ijm.fr

‡Corresponding author: thierry.grange@ijm.fr

G-quadruplexes in ancient genomes

Jean-Louis Mergny * [1]

[1] Laboratoire d'Optique et Biosciences – Ecole Polytechnique – France

G-quadruplexes ("G4") are unusual DNA and RNA structures which can find applications in biology and medicine. We developped a new algorithm for the prediction of G4 propensity and apply it to the analysis of a number of species, current and extinct (*e.g.*, *H. neanderthalensis*), as well as viruses infecting *H. sapiens* (*e.g.*, Hepatitis B virus: HBV). Recent sequencing of ancient HBV viruses revealed that these viruses have accompanied humanity for several millenia. As G-quadruplexes are considered to be potential therapeutic targets in virology, we examined G-quadruplex-forming sequences (PQS) in modern and ancient HBV genomes. Our analyses showed the density of PQS motifs is lower in ancient HBV genomes than in their modern counterparts. This indicates that the PQS content in HBV increased over time to become closer to the PQS frequency in the human genome. These results constitute the first paleogenomics analysis of G4 propensity. We propose that, for viruses causing chronic infections, their PQS frequencies tend to converge evolutionarily with those of their hosts, as a kind of "genetic camouflage" to both hijack host cell transcriptional regulatory systems and to avoid recognition as foreign material

Keywords: G, quadruplex, paleogenomics, bioinformatics, Genome evolution

*Speaker

Single-cell analyses of the human airways in health and respiratory diseases

Laure-Emmanuelle Zaragosi *† [1]

[1] Institut de pharmacologie moléculaire et cellulaire – Université Nice Sophia Antipolis (1965 - 2019), Centre National de la Recherche Scientifique, Université Côte d'Azur – France

The defence function of the airway epithelium results from homeostasis of a complex cellular ecosystem which can be altered in chronic respiratory diseases. We have used single-cell RNA profiling to investigate the extent of these modifications in terms of cell compositions and differential gene expression at different levels of the airway tree.
Our initial single-cell atlas of the human healthy airways was established in young healthy adults and was then integrated in a larger Human Lung Cell Atlas (HLCA) that captures the cellular diversity of the human lung by describing the characteristics of 62 distinct cell types. We also directly compared patients suffering from early stages of Chronic Obstructive Pulmonary Disease (COPD) with healthy age-matched volunteers. This new dataset of 119 samples was assembled after collecting cells by brushings or biopsies from the nose to the 6th division of the airways and results in the first COPD airway cell atlas.
The benefits of integrating healthy and disease single-cell atlas in larger resources leads to a more robust characterization and better recovery of cell types and states. This establishes the global HLCA as a valuable resource for future investigations, and defines a general computational strategy transferrable to the analysis of other pathological samples.

Keywords: human airways, cell atlas, single cell

*Speaker
†Corresponding author: zaragosi@ipmc.cnrs.fr

Maturation exploration of kidney organoids from single-cell RNA-seq data analysis

Solène Brohard *† [1], Camille Lemercier [1], Alexandre Hubert [1], Jean-François Deleuze [1], Eric Bonnet [1]

[1] Centre National de Recherche en Génomique Humaine – Commissariat à l'énergie atomique et aux énergies alternatives – France

Organoids are complex 3D structures derived from stem cells. They display architectures and functionalities similar to *in vivo* organs. Organ-specific studies were able to confirm the structure and the composition of mature organoids by comparing cell populations found in organoids with the ones found in the target organ. This comparison is possible by analyzing singe-cell RNA-seq data of organoids to identify cell populations. In this context, we wanted to explore these sequencing data from organoids at different times of culture and observe, without *a priori*, their maturation by using descriptive and statistical methods.
We investigated single-cell RNA-seq public datasets of kidney organoids at different days of maturations from 2 different commercial cell lines (Subramanian et al, Nature communication, 2019). We integrated the datasets from the different maturation days to highlight cell specificities for each stage of maturation. We also studied cell trajectories to observe correlations between cell differentiation and the date of maturation. We finally used statistical methods to explore the variability of cell proportions between days of maturations.

From this preliminary study, computational analyses could defined 2 distinct early and late maturation levels in kidney organoids.

The perspectives of this project is to continue the maturation exploration on different datasets and possibly on organoids of other target organs. The final objective is to use machine-learning-based approaches to develop a computational method that can estimate the maturation level of organoids.

Keywords: Organoids, Maturation, Single, cell RNA, seq, Kidney

*Speaker
†Corresponding author: brohard@cng.fr

Tracking B cell immune responses through integrative single-cell sequencing of transcriptome and antibody genes

Pierre Milpied *† [1]

[1] Centre d'Immunologie de Marseille - Luminy – Aix Marseille Université, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique – France

B cells undergo rapid molecular and cellular evolution to produce high affinity antibodies in response to pathogens and vaccines. In my lab, we use genomics technologies for integrative analyses of transcriptome and antibody-coding gene sequences in single human B cells to decipher the dynamics of immune responses in complex tissues. I will highlight how such analyses enable to infer the dynamics of B cell responses occurring within human lung tumors.

Keywords: B cell, single cell, antibodies, immune response

*Speaker
†Corresponding author: milpied@ciml.univ-mrs.fr

The "dog" model for human genetic diseases: resources, tools and strategies applied to the genetic of longevity.

Thomas Derrien * [1], Benoit Hédan [1], Catherine André * † [1]

[1] Institut de Génétique et Développement de Rennes – Université de Rennes, Centre National de la Recherche Scientifique – France

With the longstanding artificial selection made by humans to shape modern dog breeds, the canine model represents a unique mammalian model for deciphering the genetic basis of various phenotypes (such as size, coat color, and behavior), diseases, and longevity. Our team works with dogs as a natural model to investigate the genetic and epigenetic basis of lifelong traits (https://igdr.univ-rennes.fr/equipe-genetique-du-chien). We have established the Cani-DNA Biological Resource Center, which houses a collection of 30,000 dog DNA samples and over 6,000 tissue samples available for biomedical research. Our approach aligns with the "One Health" concept and is conducted within the framework of ethical and participatory science. Our work has already facilitated the identification of genes and mutations in dogs, leading to the identification of orthologous genes in humans associated with rare cancers, dermatological and neurological diseases, among others. This research offers mutual benefits for both human and veterinary medicine. More recently, we have investigated the canine model's potential for studying longevity. In dogs, the average lifespan of a breed is inversely correlated with its average body weight, with smaller dogs living longer than larger dogs. Through the GOLDogs project (https://www.france-genomique.org/projet/goldogs/) funded by France Génomique, our objective is to create the most comprehensive catalog of genetic variations in 25 dog breeds. This will be achieved by combining low-pass sequencing of 500 aged dogs with long-read sequencing and genome assemblies of a diverse panel of dogs representing the main 25 breeds. This invaluable resource will not only contribute to the study of genomic diversity in dog populations but also provide a new strategy for mapping all forms of genetic variations to specific phenotypes of interest, using the fascinating example of longevity in dogs.

Keywords: Dog model, Longevity, Structural Variants, LongRead sequencing

*Speaker

†Corresponding author: catherine.andre@univ-rennes1.fr

Exploring the non-additive component of multifactorial traits through autozygosity in large Biobanks

Anne-Louise Leutenegger*†[1]

[1]Maladies neurodéveloppementales et neurovasculaires – Institut National de la Santé et de la Recherche Médicale, Université Paris Cité – France

Abstract

Genome-wide association studies (GWAS) have revealed numerous associations between genetic variants and multifactorial traits in population samples. While these studies have provided valuable insights, the genetic component of many diseases remains incompletely understood. GWAS primarily focus on common variants and assume an additive genetic model. Therefore, there is a growing interest in exploring the non-additive components of diseases, particularly the potential contributions of rare variants with recessive effects. We present a novel approach that relies on identifying an excess of autozygous (homozygous-by-descent) segments shared among cases compared to what would be expected in controls. We illustrate its performance on the UK Biobank cohort (500k participants aged 40-69 at recruitment) focusing on the individuals of non-European descent and type 2 diabetes phenotype. Our analysis reveals an overrepresentation of consanguineous individuals among those diagnosed with diabetes across all population ancestries and identified candidate regions for rare recessive variants associated with diabetes in the different population ancestries. We show the impact of the specific biobank data fields used for defining case and control individuals. The approach is implemented in the R package Fantasio (https://github.com/genostats/Fantasio).
Research work conducted using the UK Biobank biomedical database www.ukbiobank.ac.uk under application #59366 and funded by the Inserm cross-cutting program GOLD.

Keywords: multifactorial traits, autozygosity, Biobanks

*Speaker

†Corresponding author: anne-louise.leutenegger@inserm.fr

Transcriptome-wide analysis to better understand the molecular mechanism of pathogenesis in giant cell arteritis

Michal Zulcinski *† [1,2], Lubna Shafi [1,3], Arundhati Chakrabarty [4], Mark M Iles [2,5], Ann W Morgan [1,3]

[1] School of Medicine, University of Leeds – United Kingdom
[2] Leeds Institute of Data Analytics, University of Leeds – United Kingdom
[3] NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds – United Kingdom
[4] Department of Histopathology, St James Hospital, Leeds – United Kingdom
[5] Leeds Institute of Medical Research, University of Leeds – United Kingdom

Giant cell arteritis (GCA) is the most common form of vasculitis in people over 50 years old, characterised by inflammation in medium- and large-sized vessels. Better understanding of the underlying genetics and molecular mechanisms driving GCA is needed to discover alternative treatment options, currently limited to high-dose steroids on disease onset.
Our study seeks to identify genes and pathways associated with histological phenotypes of inflammatory infiltration in GCA and to investigate genetic correlation between GCA and different molecular traits such as gene expression and metabolites by computing genotypic risk scores for the intermediate traits.

We performed a comprehensive transcriptomic analysis of 82 temporal artery biopsies from UKGCA consortium and computed genotypic risk scores, using the GENOSCORES platform, constructed from genome-wide association study summary statistics in 4,866 genotyped individuals (1,936 GCA-cases; 2,930 unaffected controls). These scores were used in regression models to test for associations with different phenotypes. Confounding effects of clinical variables was assessed by logistic regression separately for each histological phenotype. No association with sex (%female_˜70) or the duration of steroid treatment (0-23 days) was detected, and age (59-92) was found to be positively correlated with the 'presence of infiltration around vasa vasorum' and the 'extent of inflammation in the intimal layer of artery' (r=0.067; P=0.045 and r=0.109; P=0.022, respectively). Differentially expressed genes were identified for each phenotype revealing that the histological phenotypes showing the greatest transcriptional changes between the groups were the presence of media destruction (5159 genes; FDR-corrected P< 0.01) and the extent of inflammation in the adventitia and media (3503 and 3333 genes respectively; all FDR-corrected P< 0.01).
This study provides new insights into molecular phenotypes of GCA and expands the list of genes currently known to be associated with the disease. Further analyses are underway to determine the functional and biological significance of identified genes and associations.

*Speaker
†Corresponding author: m.zulcinski@leeds.ac.uk

Current methods in statistical imputation: going beyond missing genotypes

Anthony F. Herzig * [1], Maël Guivarch [1], Aude Saint Pierre [1], Gaëlle Marenne [1], Gaëlle Le Folgoc [1], Frex Consortium [1], Francegenref Consortium [2], Popgen Consortium [1], Emmanuelle Génin [1,3]

[1] Univ. Brest, Inserm, EFS, UMR 1078, GGB, IBSAM, Brest, France – Centre de Recherche Inserm – France
[2] 2. LABEX GENMED, Centre National de Recherche en Génomique Humaine, Evry, France – CEA-DRF-IBFJ-CNRGH – France
[3] CHU Brest, Brest, France – CHU Brest – France

Statistical imputation of missing genotypes using population data has become exceedingly accurate. This is largely enabled by the increasing availability of enormous reference datasets and increased efforts to sequence population-specific panels. Highly accurate imputation enables for high-resolution sequencing data to be attained at a reduced cost, with resources such as the new UKBioBank imputation service carrying taglines along the lines of 'near-perfect genome imputation for less than £0.10'. Furthermore, the vogue for shallow whole-genome sequencing represents a new tendency for imputation to be integrated more directly into bioinformatics pipelines.
Here we discuss current efforts and future challenges for genotype imputation strategies in French populations; focusing on different large-scale sequencing projects. Notably, we demonstrate the impact of fine-scale population structure on imputation efficacy. Furthermore, the modelling of haplotype sharing that is the foundation of statistical imputation is being increasingly re-purposed to empower genetic association studies. We give an overview of such recent advances as well as introducing our recent method development of a surrogate-family based association test (SURFBAT).
Grant: French Ministry of Research PFMG2025, ANR IA-10-LABX-0013 FranceGenRef and ANR-11-INBS-0002 Constances.

Keywords: genotype imputation, population genetics, reference panel, haplotypes

*Speaker

Whole genome sequencing in cardiovascular diseases

Jean-Jacques Schott*†[1]

[1]unité de recherche de l'institut du thorax UMR1087 UMR6291 – Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Nantes Université - UFR de Médecine et des Techniques Médicales – France

Abstract

Sudden cardiac death (SCD) is a leading cause of total and cardiovascular mortality. Ventricular fibrillation (VF) is the most common arrhythmia causing SCD. In young individuals, SCD occurs primarily in the setting of rare cardiac disorders generally considered as Mendelian diseases. Among these are the primary electrical disorders associating lethal arrhythmias such as VF. These disorders are highly amenable to genetic studies since they manifest clinically with specific features on the electrocardiogram.
Whole genome sequencing allows to interrogate the non-coding portions of the genome and offers the possibility to provide a molecular diagnosis to patients with unsolved genetic disorders. We describe 7 families presenting with a complex heart disease associated electrical disorders to developmental anomalies that share overlapping deletions in a 1.5 Mbp gene desert on chromosome 4q25. The smallest deletion harbors two diverging CTCF binding sites and disrupts a TAD boundary. Among the genes located in the TADs, PITX2, a homeobox transcription factor, plays a key role in left-right asymmetrical development of organs including gut, stomach and heart, in cell fate determination, differentiation and organogenesis. Moreover, PITX2 is involved in the regulation of genes underlying electrophysiological properties of the left atrium and the *PITX2* locus has been implicated in atrial fibrillation by genome-wide association studies. Using human induced pluripotent stem cell (hiPSC) models and mouse models, we show that the deletions in the families cause chromatin conformation changes and dysregulation of PITX2 expression in the heart, explaining the observed phenotypes.

Keywords: Whole genome sequencing, cardiovascular diseases

*Speaker
†Corresponding author: Jean-Jacques.Schott@univ-nantes.fr

Unravelling the interplay between type 2 diabetes, genetics and metabolite levels in the UK Biobank cohort

Ozvan Bocher *† [1], Archit Singh [2], Ana Luiza Arruda [2], Peter Kreitmaier [2], Andrei Barysenka [2], William Rayner [2], Eleftheria Zeggini [2]

[1] Institute of Translational Genomics, Helmholtz Munich – Germany
[2] Institute of Translational Genomics, Helmholtz Munich – Germany

Type 2 diabetes (T2D) represents a major health burden for which genetics has been successfully investigated in large GWAS. The remaining challenge lies in fully understanding the role of these variants in biology giving rise to the disease, something which can be investigated through metabolomics. We sought to investigate the interplay between genetics, metabolomics and T2D risk in the UK Biobank cohort. We first conducted a bidirectional Mendelian Randomization study to assess causal relationships between metabolite levels and T2D risk. We found only a few of the 164 absolute metabolite levels tested to be causal of T2D, but half of them to be caused by T2D (with P value down to 10-61), including an increase in amino acids and glucose levels, and a decrease in cholesterol classes. Some of these metabolites are also seen to be associated with specific T2D complications such as HDL cholesteryl esters showing lower values in T2D individuals with kidney complications compared to those without complications (β=-0.55, p=3.66x10-8). Secondly, we assessed the interaction between T2D status and genetic variants through a differential metabolite QTL analysis. We find 22 metabolites that are differentially genetically regulated in individuals with and without T2D, including glycine (β=0.41, p=5x10-25 for the most significant SNP) and low-density lipoproteins (β=0.53, p=1.61x10-12) which show the most prominent signals. More than a third of these 22 metabolites were found to be caused by T2D. This work provides a better understanding of the metabolic changes induced by the occurrence of T2D. While further work is needed to confirm these results and better disentangle the underlying biology, they provide potential directions to investigate T2D consequences and subsequent complications.

Keywords: T2D, Metabolomics, UK Biobank, QTL, Mendelian Randomization

*Speaker
†Corresponding author: ozvan.bocher@helmholtz-munich.de

Tridimensional analysis of human head development

Alain Chedotal *† [1]

[1] Institut de la Vision – Institut National de la Santé et de la Recherche Médicale, Sorbonne Universite, Centre National de la Recherche Scientifique – France

The evolution and development of the head have long captivated researchers due to its crucial role as the gateway for sensory stimuli and its intricate structural complexity. While significant progress has been made in understanding head development in various vertebrate species, our knowledge of early human head ontogeny remains limited. We used advanced whole-mount immunostaining and 3D imaging techniques to generate the first comprehensive 3D cellular atlas of human head embryogenesis. We built detailed developmental series of diverse head tissues and cell types, including muscles, vasculature, bones and cartilage, peripheral nerves, and exocrine glands. These datasets, accessible through a dedicated web interface, provide unique insights into human embryogenesis. These insights into human embryology have important implications for understanding craniofacial defects, neurological disorders, and advancing diagnostic and therapeutic strategies.

Keywords: Genome 3D, human head, cell atlas

*Speaker

†Corresponding author: alain.chedotal@inserm.fr

Epigenetic regulation of genome architecture in development, cell differentiation and cancer

Giacomo Cavalli *† [1]

[1] Institut de génétique humaine – Centre National de la Recherche Scientifique, Université de Montpellier – France

The eukaryotic genome folds in 3D in a hierarchy of structures, including nucleosomes, chromatin fibers, loops, chromatin nanodomains, topologically associating domains (TADs), compartments and chromosome territories that are highly organized in order to allow for stable memory as well as for regulatory plasticity, depending on intrinsic cues, such as chromatin association of Polycomb proteins, CTCF and cohesin, and on environmental cues. We showed that TADs and chromatin loops can assist gene regulation, both in *Drosophila* and in mouse cells. Here, we focused on Polycomb components. Polycomb proteins are a set of highly conserved transcriptional repressor from flies to human that have the remarkable ability to induce stable epigenetic memory that can be inherited through cell division and across organismal generations. Surprisingly, we have shown that a transient alteration of Polycomb function can change cell fate and induce cancers that can not be rescued, even after returning the cells to normal levels of Polycomb. This is linked to the irreversible derepression of genes that can drive tumorigenesis, including JNK and JAK-STAT signalling pathways genes, which we show to be required for Polycomb-dependent tumorigenesis. These data show that a reversible perturbation of Polycomb-Group protein levels can induce cancer in the absence of driver mutations and suggest that this is achieved through epigenetic inheritance of altered cell fates.

Acknowledgements
This work was supported by the European Research Council Advanced Investigator grant (3DEpi), by the Horizon 2020 program (MuG), by the "Fondation ARC pour la recherche sur le cancer", by the "Fondation pour la recherche médicale" (FRM), by INSERM and the ITMO Cancer (MMTT project), by the INCa, the E-RARE IMPACT grant under the ERA-NET Cofund Horizon 2020 scheme, by the MSD-Avenir foundation and by the CNRS.

Keywords: Epigenetic regulation, development, cell differentiation, cancer

*Speaker
†Corresponding author: giacomo.cavalli@igh.cnrs.fr

Mechanisms of cell plasticity in breast cancer at single cell resolution

Céline Vallot*†[1]

[1]Centre de recherche de lÍnstitut Curie [Paris] – Institut Curie [Paris] – France

Abstract

The dynamic nature of chromatin and transcriptional features are expected to participate to tumor evolution. Our group focuses on the study of the dynamics of histone modifications in cancer cells upon cancer treatment as well as during the initial steps of tumorigenesis. We develop experimental and computational approaches to map histone marks at single-cell resolution, enabling the investigation of the dynamics of chromatin marks in tumor samples (Grosselin et al. Nat Genet 2019; Prompsy et al. Nat Comm 2020).

We have recently combined single-cell epigenomic and transcriptomic approaches to lineage tracing strategies to reveal the initial epigenomic events driving tolerance to chemotherapy in triple-negative breast cancer (Marsolier & Prompsy et al., Nat Genet 2022). We show that the repressive histone mark H3K27me3 is a lock to the activation of a drug-persistent expression program in breast cancers. Under chemotherapy, very few cells can survive the treatment, and these cells have a remodeled repressive epigenome, with targeted loss at key promoters. Using demethylase inhibitor in combination to chemotherapy, we improve the response rate and delay recurrence both in vitro and in vivo.

We also study mechanisms of cell plasticity in early breast tumorigenesis in vivo. We have recently mapped state transitions during Brca1-tumorigenesis in the mouse. We discovered that luminal progenitor cells undergo a partial epithelial to mesenchymal transition at the onset of tumorigenesis (Landragin &Saichi, unpublished 2022).

Keywords: cell plasticity, breast cancer, single cell

*Speaker

†Corresponding author: celine.vallot@curie.fr

Transcriptomic consequences of U4atac mutations, a snRNA component of the minor spliceosome

Audric Cologne , Deepak Khatri , Alicia Besson , Clara Benoit-Pilven , Audrey Putoux , Remy Bordonne , Patrick Edery , Anne-Louise Leutenegger , Vincent Lacroix , Marion Delous , Sylvie Mazoyer * [1]

[1] Equipe GENDEV, Centre de Recherche en Neurosciences de Lyon – CNRS UMR5292, INSERM U1028, Université Claude Bernard Lyon 1 – France

Intriguingly, most species have two distinct splicing mechanisms, one operated by the well-characterised major spliceosome that contains the U1, U2, U4, U5 and U6 snRNAs, and the other by the lesser-known minor spliceosome that contains U11, U12, U4atac, U5 and U6atac. Minor splicing, as its name suggests, concerns in most species a very small number of introns: in the human genome there are about 935 of them distributed in _˜750 genes.
Mutations in *RNU4ATAC*, the gene for U4atac, lead to three very rare recessive developmental diseases: Taybi-Linder syndrome (TALS, or microcephalic osteodysplastic primordial dwarfism type 1 (MOPD1)), Roifman syndrome (RFMN), and Lowry-Wood syndrome (LWS). They have common clinical features of varying severity: microcephaly, growth retardation, skeletal dysplasia, intellectual disability, TALS being the most severe (early death

Following the discovery in 2011 of *RNU4ATAC* as the TALS gene, we are conducting a research project that aims to understand the physiopathological and molecular bases of the *RNU4ATAC* syndromes, to identify which minor intron-containing genes are important for embryogenesis and brain development, and to decipher the "raison d'être" of minor splicing.
PolyA+ RNA-sequencing of patient cell lines showed that although most minor introns are retained when *RNU4ATAC* is mutated, the magnitude of these retentions differs depending on the cell type. The study of our zebrafish models showed that in KO fish, surprisingly, minor splicing can still take place, sometimes even more efficiently than in controls, but only in the early stage of development after which they stop developing and die. Concerning brain development, in KD fish, we saw in embryo head polyA+ RNAs a large number of genes with a high level of retention of their minor intron(s), and contrary to expectations, also changes in the splicing efficiency of some major introns, suggesting that minor and major splicing may rely on each other for the regulation of the expression of a subset of genes.

Keywords: Rare diseases, splicing, transcriptomic studies

*Speaker

Genomics of vascular brain disease across the lifespan and across ancestries

Stéphanie Debette*†[1,2,3]

[1]Université de Bordeaux – Université de Bordeaux – France
[2]Bordeaux Population Health Research Centre – Institut National de la Santé et de la Recherche Médicale - INSERM, Université de Bordeaux (Bordeaux, France) – France
[3]Institut Hospitalo-Universitaire VBHI – Université de Bordeaux (Bordeaux, France), CHU de Bordeaux; University of Bordeaux; Inserm U1219, Institut National de la Santé et de la Recherche Médicale - INSERM, L'Institut National de Recherche en Informatique et e n Automatique (INRIA), région Nouvelle-Aquitaine – France

Abstract

In recent years large collaborative genomic studies have led to substantial progress in the identification of genetic variants associated with vascular brain disease. In this presentation, we provide an overview of such collaborative studies on stroke and brain MRI markers of cerebral small vessel disease (cSVD), revealing over 150 independent loci associated with these conditions. We will expose how cross-ancestry efforts have enhanced our power to detect associations. In silico functional explorations of the observed genetic associations point to a major role of blood pressure-related pathways, but also mechanisms independent of vascular risk factors, such as extracellular matrix structure and function, membrane transport, vascular development, myelination, and blood-brain barrier. Leveraging next-generation sequencing data these studies also shed new light on the continuum between monogenic and multifactorial stroke and cSVD, with several genes harboring both rare mutations and common variants contributing to the disease. Intriguingly, we recently showed that MRI-cSVD risk loci identified in middle- and older age are associated with brain white matter microstructure and perivascular space burden already in young adulthood, suggesting that processes contributing to cSVD may find their root much earlier in life than previously thought. We provide preliminary insights on how these association patterns change across the adult lifespan and how this can inform clinical applications. We will further present how GWAS summary statistics have been leveraged, in combination with other omics resources, for genomics-driven drug discovery for stroke and cSVD. Finally, we will present how the combination of these new GWAS resources with innovative integrative polygenic score methodology has informed stroke risk prediction, for the first time across ancestries.

Keywords: vascular brain disease, lifespan, ancestries

*Speaker
†Corresponding author: stephanie.debette@u-bordeaux.fr

Adaptation to diet: human and microbial perspective

Laure Ségurel*†[1]

[1]Laboratoire de Biométrie et Biologie Evolutive - UMR 5558 – Université Claude Bernard Lyon 1, Institut National de Recherche en Informatique et en Automatique, VetAgro Sup - Institut national d'enseignement supérieur et de recherche en alimentation, santé animale, sciences agronomiques et de l'environnement, Centre National de la Recherche Scientifique, Centre National de la Recherche Scientifique : UMR5558 – France

Abstract

Understanding how species adapt to changes in their environment remains a central goal in evolutionary biology. The adaptive history of our species is especially interested to study, as human populations both faced contrasted environmental constraints when migrating all around the planet, but also strongly culturally modified their environment, which led to novel selective pressures. However, many of these cultural changes are too recent to impact the human genetic diversity. In contrast, our gut microbes evolve much faster than we do, and thus are informative about very recent timescales.
In this talk, I will discuss both how human populations have adapted to changes in diet linked to the neolithic revolution, 10 000 years ago, but also how our gut microbiota has adapted to to the growing urbanization and industrialization of human populations.
Key words : Adaptation to diet: human and microbial perspective

Keywords: diet, adaptation, microbes

*Speaker
†Corresponding author: laure.segurel@univ-lyon1.fr

The genetic legacy of past epidemics

Etienne Patin [*†] [1]

[1] 'Human Evolutionary Genetics' lab – Institut Pasteur de Paris – France

Infectious diseases have been one of the main causes of mortality in our species, *Homo sapiens*. Recent paleo-genetic studies have revealed the footprints that past epidemics have left on the human genome: genetic variants, which confer resistance against infection by pathogens, have spread through human populations by natural selection. But this genetic legacy also comes at a price...

Keywords: Past epidemics, genetic legacy

[*]Speaker
[†]Corresponding author: epatin@pasteur.fr

Compositional biases establish a link between spatial organization and expression regulation of genes and their products

Didier Auboeuf *† [1]

[1] Laboratoire de biologie et modélisation de la cellule – Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique – France

Since the discovery of operons, numerous studies have shown a link between genome organization, gene expression regulation, and gene product functions since genes in close physical proximity to each other tend to be co-expressed, co-regulated and their products tend to participate in the same biological functions. To date, however, there is no satisfactory explanation for this relationship. Inspired by our previous experiments on post-transcriptional co-regulation of gene products based on their nucleotide compositional biases, we show that genes that are in physical proximity to each other share the same nucleotide compositional biases. This at least partially explains their transcriptional co-regulation and spatial proximity, depending on histones and transcription factors. In addition, we found evidences that co-locating genes produce RNAs sharing the same compositional biases, contributing to their post-transcriptional co-regulation through splicing factors. Then, we show that co-locating genes having the same nucleotide composition biases generate proteins sharing the same amino acid composition biases and physicochemical properties. Finally, we show that proteins in proximity to each other in the cell and sharing the same biological functions tend to have the same amino acid composition biases and physicochemical properties. Altogether our data reveal that compositional biases of genes and consequently of their products provide an explanation of the relationship between the human genome spatial organization, transcriptional and post-transcriptional co-regulation of gene products, and their biological functions.

Keywords: Compositional biases, spatial organization, expression regulation of genes

*Speaker

†Corresponding author: didier.auboeuf@inserm.fr

Interplay between environmental exposure, genetic predisposition and somatic genomic alterations in liver oncogenesis

Jessica Zucman-Rossi*†[1]

[1]Centre de Recherche des Cordeliers – Ecole Pratique des Hautes Etudes, Institut National de la Santé et de la Recherche Médicale, Sorbonne Universite, Université Paris Cité – France

Abstract

Our research project aims to a better understanding of how different etiologies, including exposure to toxic agents, diet habits, genetic predisposition, chronic diseases, non-alcoholic fatty liver disease (NAFDL), and their cooperative actions, contribute to carcinogenesis in the liver and to tumor progression.
We explored this question in adults with hepatocellular carcinoma, using GWAS and WGS in tumors, and identified new genetic polymorphisms associated with the risk of HCC in the context of high alcohol intake and other exposures leading to specific mutational signatures in the tumors.

A similar strategy in benign liver tumors identified an interplay between genetic predisposition markers and different exposures to estrogen or androgen levels, leading to the development of different molecular subtypes of hepatocellular adenoma.

In children, integrated omic analyses identified chromosome 11p15 mosaic genetic alterations in the liver associated with early development of hepatoblastoma and specific mutational signatures related to cisplatin resistance.
These results led to new concepts in understanding the mechanisms of tumorigenesis that can be translated into clinical care of patients with liver tumors to improve early diagnosis and treatments.

Keywords: carcinogenesis, HCC, GWAS, WGS

*Speaker
†Corresponding author: jessica.zucman-rossi@inserm.fr

Bacteria mediated histone modifications in health and disease

Melanie Hamon *† [1]

[1] Institut Pasteur de Paris – Institut Pasteur de Paris – France

The team works on how bacteria modify the chromatin of host cells, a process that is actively induced to promote infection. Using the model bacterium *Streptococcus pneumoniae,* I will present evidence that it induces lasting histone modifications during infection, thereby leading to memory responses after bacterial clearance.

Keywords: bacteria, histone modifications

*Speaker
†Corresponding author: melanie.hamon@pasteur.fr

Large language models in genomics

Jean-Philippe Vert [*†] [1,2]

[1] Owkin France – United States
[2] Mines Paris - PSL (École nationale supérieure des mines de Paris) – Université Paris sciences et lettres – France

Large language models in genomics.

Keywords: Large language models

[*]Speaker
[†]Corresponding author: Jean-Philippe.Vert@mines.org

Leveraging genetic variant pleiotropy to improve fine mapping in human complex traits

Martin Tournaire * [1], Yves Rozenholc [1], Marie Verbanck [1]

[1] UR7537-BioSTM, Biostatistique, Traitement et Modélisation des données biologiques – Université Paris Cité – France

Background: Although pleiotropy, which occurs when a genetic variant affects at least two traits, is thought to play a central role in the genetic architecture of complex traits and diseases, it is a poorly understood mechanism. Our objective is to build a genome-wide map of pleiotropic variants of the human genome.
Methods: We have developed PleioVar, an algorithm to detect pleiotropic variants in the human genome. PleioVar derives variant pleiotropic labels from integrative Mendelian Randomization methods such as LHC-MR or MR-CUE, that take into account pleiotropy. The objective is to discern the biological mechanisms leading to observed pleiotropy. We propose 5 types of pleiotropy (direct, confounding-mediated, traits-mediated, linkage disequilibrium and serendipitous) and model them with Gaussian mixture models, at the level of variants. We selected highly heritable traits and diseases from UK Biobank and built a complex network GWAS simulation framework based on these traits to assess the performance of PleioVar. Then, using GWAS summary statistics from UK Biobank, we used PleioVar to detect pleiotropic variants of these same traits.

Results: On simulations, PleioVar successfully predicted variant pleiotropy with high accuracy. Simulated genetic variants were correctly clustered as displaying the different types of pleiotropy. Using GWAS summary statistics from highly heritable traits and diseases from the UK Biobank, pleiotropic genetic variants were identified by PleioVar, leading to a genome-wide variant-trait map.
Conclusion: This method, designed to map genetic variants to traits through pleiotropy, could help better understand the human genetic architecture and characteristics of complex traits and diseases.

Keywords: Pleiotropy, GWAS, Machine learning, Classification

*Speaker

Single cell transcriptome sequencing of stimulated and frozen human peripheral blood mononuclear cells

Céline Derbois , Marie-Ange Palomares , Jean-François Deleuze , Eric Cabannes , Eric Bonnet * [1,2]

[1] Centre National de Recherche en Génomique Humaine – Commissariat à l'énergie atomique et aux énergies alternatives – France
[2] Laboratoire de Bio-analyse – CEA-DRF-IBFJ-CNRGH – France

Peripheral blood mononuclear cells (PBMCs) are blood cells that are a critical part of the immune system used to fight off infection, defending our bodies from harmful pathogens. In biomedical research, PBMCs are commonly used to study global immune response to disease outbreak and progression, pathogen infections, for vaccine development and a multitude of other clinical applications. Over the past few years, the revolution in single-cell RNA sequencing (scRNA-seq) has enabled an unbiased quantification of gene expression in thousands of individual cells, which provides a more efficient tool to decipher the immune system in human diseases. In this work, we generate scRNA-seq data from human PBMCs at high sequencing depth ($>$ 100,000 reads/cell) for more than 30,000 cells, in resting, stimulated, fresh and frozen conditions. The data generated can be used for benchmarking batch correction and data integration methods, and to study the effect of freezing-thawing cycles on the quality of immune cell populations and their transcriptomic profiles.

Keywords: Human, PBMC, immune cells, single, cell, single cell RNAseq

*Speaker

PaintorPipe: a pipeline for genetic variant fine-mapping using functional annotations

Zoé Gerber [1], Michel Fisun [2], Hugues Aschard [2], Sarah Djebali * [1]

[1] Institut de Recherche en Santé Digestive – Université Toulouse III - Paul Sabatier, Ecole Nationale Vétérinaire de Toulouse, Institut National de la Santé et de la Recherche Médicale, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – France
[2] Génétique Statistique - Statistical Genetics – Institut Pasteur [Paris], Université Paris Cité – France

Genome Wide Association Studies (GWAS) have identified thousands of genetic variants associated with common diseases. However, pinpointing variants that are truly causal remains a challenge. Indeed, GWAS results likely include a mix of causal variants and variants in linkage disequilibrium (LD, i.e. highly correlated) with the causal variants. In order to identify actual causal variants, fine-mapping methods have been developed. These methods use GWAS results and LD information, to assign to each variant a probability of being causal. In this field, PAINTOR (1) has become a standard, and one of its advantages is its ability to take into account functional annotations. Since a PAINTOR run requires a lot of pre- and post-processing steps, we decided to wrap all these steps into a Nextflow pipeline called PaintorPipe (https://github.com/sdjebali/PaintorPipe). PaintorPipe uses three independent sources of information: GWAS summary statistics, LD information and functional annotations, to rank the variants according to their susceptibility to be involved in the development of the disease. The PAINTOR program is used to calculate the posterior probability of each SNP to be causal (a.k.a Bayesian fine-mapping). The resulting credible sets of variants are annotated with their biological functions and visualized using PAINTOR's visualization tool called CANVIS. This pipeline is implemented in the Nextflow pipeline specific language (DSL2) (2), can be run locally or on a slurm cluster and handles containerisation using Singularity (3). It is designed to be modular and customizable, allowing for an easy integration of diverse functional annotations. To validate PaintoPipe, we ran it (version 1.1.1) on GWAS results from the latest Coronary Artery Disease (CAD) meta-analysis involving 122,733 cases and 424,528 controls (4), and identified 149 loci with fine-mapping information. Out of the 161 CAD loci previously identified by the meta-analysis, and based on the presence of the lead SNP in our fine-mapped loci, 128 (78%) were common with our results. We additionally tested the impact of several input parameters of the pipeline, including the types of annotations used, on the pipeline's results on CAD, and also tested the pipeline on type 2 diabetes.
References

(1) Kichaev G, Yang WY, Lindstrom S, Hormozdiari F, Eskin E, et al. (2014) Integrating Functional Data to Prioritize Causal Variants in Statistical Fine-Mapping Studies. PLOS Genetics 10(10): e1004722.

(2) Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C.

*Speaker

(2017). Nextflow enables reproducible computational workflows. Nature Biotechnology, 35(4), 316–319. doi:10.1038/nbt.3820

(3) Kurtzer GM, Sochat V, Bauer MW (2017) Singularity: Scientific containers for mobility of compute. PLOS ONE 12(5): e0177459.

(4) van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. 2018 Feb 2;122(3):433-443. doi: 10.1161/CIRCRESAHA.117.312086. Epub 2017 Dec 6. PMID: 29212778; PMCID: PMC5805277.

Keywords: fine, mapping, causal variant, genetics, common disease, functional annotation, LD, coronary artery disease

Study of the impact of Formalin-Fixed Paraffin Embedded (FFPE) tissues compared to frozen (FF) tissues in PCR-free WGS

Alice Moussy[1], Mélanie Letexier*†[1], Jasmin Cevost[1], Edouard Turlotte[1], Margaux Gras[1], Marine Rouillon[1], Kévin Gorrichon[1], Paul Hofman[2], Alain Viari[1], Violette Turon*[1], and Jean-François Deleuze[3]

[1]Centre de Référence, d'Innovation, d'eXpertise et de Transfert (CRefIX) – CEA – France
[2]Laboratoire de Pathologie Clinique et Expérimentale. – Hôpital Pasteur [Nice] – France
[3]Centre National de Recherche en Génomique Humaine – CEA – France

Abstract

Background/Objectves: The France Genomic Medicine Plan 2025 (FMG2025) introduces WGS into healthcare pathway for rare diseases and cancers. Following a plan's request, The Reference Center for Technology, Innovation and Transfer (CRefIX) investigated the impact of formalin on WGS PCR Free data from FFPE *(Fresh Frozen Paraffin Embedding)* samples compared to FF *(fresh frozen)* samples. Chemical modifications made to DNA during tissue fixation (formalin, processing and storage) can lead to random fragmentation and sequence modification (loss or transition), which will interfere with variant detection and genome copy number. The aim of the study was to identify DNA molecular modifications caused by FFPE and their impact on the validity and reliability of results in the final interpretation.
Materials & Methods: Matched FF and FFPE samples (mirror block) and pbmc from three lung cancer tissues were sequenced on NovaSeq 6000, 2x150 bp. The depth of sequencing was 80X for FF and FFPE samples and 40 for pbmc. Somatic mutational were performed : SNV (by Mutect2 (GATK v4.1.7.0) on high-confidence regions; CNA (FACETS (bioconductor/3.13), SBS signature (R package, MutationalPatterns (V3.8.0)). The F1-score was also reported, FF samples used as gold standard.

Results & Conclusions: The results showed that fixation of tissues with formalin and paraffin embedded (FFPE) does not hinder with the final biological analysis of variants when appropriate filters are applied. CNA's analysis shows some discrepancies and further investigations are required. SNV artifacts can be corrected with an adequate read depth consistent with tumoral cellularity. The need to work from tissue sufficiently rich in tumor cells (> =40%) and a minimum depth of 80X is required. A variant frequency filter (VAF) is also effective and increases the F1-score.
Grant References: ANR-18-INBS-0001 (French National Research Agency

Keywords: French Medicine plan 2025, CRefIX, FFPE

*Speaker
†Corresponding author: melanie.letexier@cnrgh.fr

Novel statistical genetic tools to leverage WGS data to better understand the contribution of genetic variants in human diseases

Gaelle Marenne *† [1], Ozvan Bocher [2,3], Marie-Sophie C Ogloblinsky [3], Thomas E Ludwig [3,4], Chaker Aloui [5], Elisabeth Tournier-Lasserve [5,6], Emmanuelle Génin [3,4]

[1] Inserm, Univ Brest, EFS, UMR 1078, GGB, F-29200 Brest – Inserm, Univ Brest, EFS, UMR 1078, GGB, F-29200 Brest, France – France
[2] Institute of Translational Genomics, Helmholtz Zentrum München – Germany
[3] Inserm, Univ Brest, EFS, UMR 1078, GGB, F-29200 Brest – Inserm, Univ Brest, EFS, UMR 1078, GGB, F-29200 Brest, France – France
[4] CHU Brest, F-29200 Brest – CHU Brest, F-29200 Brest, France – France
[5] Université de Paris, NeuroDiderot, Inserm UMR 1141, F-75019 Paris – Université de Paris, NeuroDiderot, Inserm UMR 1141, F-75019 Paris, France – France
[6] AP-HP, Service de Génétique Moléculaire Neurovasculaire, Hôpital Saint-Louis, F-75010 Paris – AP-HP, Service de Génétique Moléculaire Neurovasculaire, Hôpital Saint-Louis, F-75010 Paris, France – France

Sequencing data from large case cohorts are generated to identify causal regions of the genome, either in research to understand the underlying mechanisms of diseases, or in clinics to resolve molecular diagnostics. The first challenge is the quality control (QC) of sequencing data, especially when external controls are needed for comparison. The second challenge is the need of interpretable regions to be used beyond genes to explore the non-coding genome but also to refine within-gene analyses.
Here we present a series of tools to analyse whole-genome sequencing data and ensure good quality results. First, we developed RAVAQ, a pipeline that performs QC to ensure high quality data and group comparability by limiting bias and batch effects. Second, we proposed to define functional genomic regions based on constrained regions of the genome with no highly pathogenic variants observed in large reference panels. These regions served as test units for rare variant association at the whole-genome scale in the RAVA-FIRST method. Third, we developed Easy-PSAP to prioritize predicted pathogenic variants in the functional genomic regions by leveraging allele frequencies from large reference panels and pathogenicity scores.

Through real data examples from moyamoya cases and the French exome Project, we show that the RAVAQ QC is accurate for rare-variant association testing with external controls. The functional regions strategy implemented in RAVA-FIRST accurately identifies the functional sub-unit of *RNF213* known to be associated with moyamoya disease, showing the interest of the strategy not only for the non-coding genome. Applying Easy-PSAP on two resolved moyamoya

*Speaker
†Corresponding author: gaelle.marenne@inserm.fr

cases accurately prioritize their recessive causal variant across their whole exome.
The functional regions we proposed along with the statistical methods we implemented are user-friendly tools for researchers to extend the analyses to the non-coding genome, but also to identify associations in specific regions of the exome.

Keywords: sequencing data, quality control, rare variants association test, variant prioritisation, non, coding genome, external controls

How reliable is the salivary microbiome information obtained from the whole genome sequencing of human saliva samples?

Lourdes Velo Suarez * [1,2], Anthony Herzig [1], Gaëlle Le Folgoc [1], Stephanie Gouriou [1], Marie Zins [3], Marcel Goldberg [3], Geneviève Hery-Arnaud [1,2], Emmanuelle Genin [1]

[1] Génétique, génomique fonctionnelle et biotechnologies (UMR 1078) – EFS, Université de Brest, Institut National de la Santé et de la Recherche Médicale – France
[2] CBAM – Centre Brestois d'Analyse du Microbiote (CBAM), CHU Brest, F-29200, Brest – France
[3] Cohortes épidémiologiques en population – Université de Versailles Saint-Quentin-en-Yvelines, Institut National de la Santé et de la Recherche Médicale, Université Paris-Saclay, Université Paris Cité – France

The past two decades have seen tremendous advances in understanding human genetic variation and its implication in disease. Similarly, the microbiome was also shown to play a significant role in human health and disease. However, the relationship between host genetic variation and microbiome composition is largely unknown, mainly because of the elevated cost of processing both types of samples with high throughput sequencing. This study explores how well an individual salivary microbiome can be characterized by using whole-genome sequencing (WGS) data obtained from DNA saliva kits designed to study human genome variation. WGS was performed on saliva samples from 35 healthy individuals who received saliva kits at home. The relative abundance distributions obtained from the analysis of non-human reads were compared against those obtained by specific 16S rRNA gene resequencing of the same DNA samples. The results showed that 16S sequencing detects only part of the salivary microbiome revealed by WGS analysis. Low-abundant taxa were identified on the WGS data that could not be captured by 16S sequencing. Interestingly, some of these taxa could be linked to oral diseases such as periodontitis. Our microbiome communities were very similar to those described in the Human Microbiome Project (HMP) and core salivary microbiome studies, showing that accurate microbiome profiles can be obtained from read data generated for human whole-genome sequencing. Furthermore, our metagenomic approach also provides additional insights into microbiome characterization such as the analysis of antimicrobial resistance in our samples.

Keywords: microbiome, saliva, human

*Speaker

Assessment of differential splicing of genes involved in motor neuron diseases in neuronal and non-neuronal cells, by long-read sequencing: proof of concept in the SPG11 gene

Marion Leblanc * [1,2], Thomas Gareau [3], Mathilde Bertrand [3], Ammara Mohammad [4], Elise Liu [5], Coline Jost Mousseau [5], Gaïzka Le Goff [5], Yannick Marie [4], Justine Guégan [3], Delphine Bohl [5], Thomas Carlile [6], Giovanni Stevanin [1,2,7]

[1] Equipe de Neurogénétique – EPHE, PSL University, Paris, France – France
[2] Team "Basic to Translational Neurogenetics" – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm U1127, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France
[3] Data Analysis Core – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm U1127, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France
[4] iGenSeq Genotyping-sequencing platform – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm U1127, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France
[5] Team "ALS : Causes and mechanisms of motor neuron degeneration" – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm U1127, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France
[6] Biogen, Cambridge (MA), USA – United States
[7] Équipe Recherche translationnelle sur les maladies neurogénetiques – INCIA, University of Bordeaux, CNRS UMR5287, EPHE, Bordeaux, France – France

SPG11 is a gene involved in various motor neuron diseases such as hereditary spastic paraplegia and amyotrophic lateral sclerosis type 5. However, no genotype-phenotype correlation has been established yet and the dynamic of expression of this gene is still not well known. Our aim is to identify all isoforms of SPG11 and establish their expression profiles, in non-pathological and in pathological models of the disease to determine if the presence of remaining isoforms could explain the variability in clinical presentation. We developed a long-read total RNA sequencing protocol on the MinION platform (Oxford Nanopore Technologies) on fibroblasts, peripheral blood mononuclear cells and motor neurons derived from induced pluripotent stem cells of control individuals and SPG11-mutated patients. The results showed that among the high-confidence transcripts detected on SPG11, the known and full-length transcript is the major transcript found in controls but not in patients, as expected. In addition, new undescribed SPG11 transcripts have also been found in both controls and patients, some with cell type specificities, indicating that the role of this gene and its isoforms could differ between cells. Long-term, we hope that this study will help unravel the pathological mechanisms of the disease heterogeneity at the clinical level.

*Speaker

Comparative Analysis of Computational Tools and Data for Predicting the Pathogenicity of Missense Variants

Ragousandirane Radjasandirane*[1], Alexandre De Brevern*†[1], and Jean-Christophe Gelly

[1]Université Paris Cité – Institut National de la Santé et de la Recherche Médicale - INSERM – France

Abstract

Amino acid substitutions on protein sequences are generally without consequences but a non negligible part of them can induce disease. Predicting the impact of a variant can be crucial for clinicians and should speed up the diagnosis of patients having missense variants suspected to cause disease. Nowadays, a multitude of computational tools have been developed with the aim of predicting the pathogenicity of variants using numerous methodologies. The most known tools are SIFT and PolyPhen accumulating more than 10k citations each. Recently, many tools have been developed using Artificial Intelligence and other new methods. Evaluating and categorizing the performance and usability of these tools is essential to guide future users and clinicians.
The goal of our work is to assess a significant number of these tools and evaluate them using quality data and metrics. We meticulously curated a high-quality, representative dataset while addressing potential issues of data circularity, where a tool is tested on data it was trained on. Performance evaluation employed diverse metrics such as MCC and AUC for each computational method.

Our findings demonstrate that the evaluated tools perform exceptionally well on data extracted from public databases like ClinVar. However, their performance diminishes significantly on clinical data, implying a potential bias within public database variant data that deep analysis before using them to train models.

Finer analysis has been done on the datasets of variants. Indeed, there are some variants that are always mispredicted by the majority of the tools. For example, a benign variant that is always predicted as pathogenic by 80% of tools. In this case, we analyzed these variants more deeply using 3D structure and interaction calculations to try explaining these mispredictions. Lastly, we examined tool behavior by conducting correlation analyses, identifying tools that exhibit similar prediction patterns.

Our work evaluates a broad range of computational tools, providing valuable insights into their performance and usability.

Key words : Bioinformatics ; ClinVar ; Prediction ; Missense variant ; Benchmark

*Speaker
†Corresponding author: alexandre.debrevern@gmail.com

Exploring Pleiotropy: Sparse Group Frequentist Meta-Analysis Models for Comprehensive Analysis

Pierre-Emmanuel Sugier * [1,2], Yazdan Asgari [3], Thérèse Truong [3], Benoit Liquet [4,5]

[1] Laboratoire de Mathématiques et de leurs Applications de Pau – Université de Pau et des Pays de l'Adour, UMR CNRS 5142, E2S-UPPA – France
[2] Equipe "Exposome et Hérédité" – Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP – France
[3] Equipe "Exposome et Hérédité" – Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP – France
[4] Laboratoire de Mathématiques et de leurs Applications de Pau – Université de Pau et des Pays de l'Adour, UMR CNRS 5142, E2S-UPPA – France
[5] School of Mathematical and Physical Sciences, Macquarie University, Sydney – Australia

An increasing number of genome-wide association studies (GWAS) summary statistics is made available to the scientific community. Exploiting these results from multiple phenotypes would permit identification of novel pleiotropic associations. In addition, incorporating prior biological information in GWAS such as group structure information (gene or pathway) has shown some success in classical GWAS approaches. Recent efforts have been made to develop methods allowing to explore pleiotropic associations at gene-level (1, 2). In particular, Baghfalaki, T. *et al.* (1) developed Bayesian meta-analysis approaches that can handle summary statistics generated from GWAS results by performing pleiotropy selection at variable (SNPs) and group-level (genes) with multiples diseases in each group independently.

We propose a novel penalized multivariate meta-analysis method adapted for pleiotropy that take into account the group structure information nested in the data by considering all genetic information at the same time. To consider all information in the same procedure can allow to gain statistical power by avoiding having to correct for multiple testing. We implement an alternating direction method of multipliers (ADMM) algorithm to perform both regularization at variable level and group level.

The performance of this method will be compared to GCPBayes methods and other benchmark meta-analysis approach on simulated data sets by considering different kind of summary data as inputs. Our method will be applied to the identification of potential pleiotropic genes between breast and thyroid cancer and the results will be presented.

References

Baghfalaki, T., Sugier, P. E., Truong, T., Pettitt, A. N., Mengersen, K., & Liquet, B. (2021). Bayesian meta-analysis models for cross cancer genomic investigation of pleiotropic effects using

*Speaker

group structure. *Statistics in Medicine*, 40(6), 1498-1518. Sutton M, Sugier PE, Truong T, Liquet B. (2022). Leveraging pleiotropic association using sparse group variable selection in genomics data. *BMC medical research methodology*, 22(1), 1-12.

Keywords: Pleiotropy, Meta, Analysis, Regularization, Group structure, Cancer

Exploring common genetic risk factors between breast cancer and thyroid cancer using GCPBayes Pipeline

Yazdan Asgari * [1], Pierre-Emmanuel Sugier [1,2], Taban Baghfalaki [3], Elise Lucotte [1], Mojgan Karimi [1], Mohammed Sedki [4], Amélie Ngo [1], Benoit Liquet [2,5], Thérèse Truong [1]

[1] Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP, Team Exposome and Heredity, Villejuif, France – Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP, Team Exposome and Heredity, Villejuif, France – France

[2] Laboratoire de Mathématiques et de leurs Applications de Pau, Université de Pau et des Pays de l'Adour, UMR CNRS 5142, E2S-UPPA, France – Laboratoire de Mathématiques et de leurs Applications de Pau, Université de Pau et des Pays de l'Adour, UMR CNRS 5142, E2S-UPPA, France – France

[3] Inserm U1219, Univ. Bordeaux, ISPED, Bordeaux, France – Inserm U1219, Univ. Bordeaux, ISPED, Bordeaux, France – France

[4] Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP, Team Psychiatrie du développement et trajectoires, Villejuif, France – Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP, Team Psychiatrie du développement et trajectoires, Villejuif, France – France

[5] School of Mathematical and Physical Sciences, Macquarie University, Sydney, Australia, Sydney, Australia – Australia

Exploring common genetic risk factors can help to better understand the underlying biological mechanisms between two traits. However, there is a lack of availability of user-friendly pipelines that let users easily perform pleiotropy analyses with genome-wide association studies (GWAS) summary statistic data as inputs. Here, we have developed a pipeline to handle GWAS summary statistics data and perform a feasible pleiotropy analysis using the GCPBayes package that was developed by our group. The pipeline harmonizes the alleles between the datasets, provides gene annotations for each variant, and explores pleiotropy at the gene level. A user could run the pipeline across three different environments (R, Bash, and a Shiny App). The pipeline offers user-friendly functionality, allowing users to easily adjust input parameters and execute analyses, and also contains a shiny App for visualization of results. A detailed tutorial is available on our GitHub page. We illustrate the use of our pipeline by exploring pleiotropic genes between breast cancer (BC) and thyroid cancer (TC). The link between BC and TC risks has been a subject of investigation through epidemiological studies, but there have been conflicting results thus far. We will present the results of this analysis at the conference.

Keywords: GWAS, Breast Cancer, Thyroid Cancer, Pleiotropy, GCPBayes Pipeline

*Speaker

Extension of a benchmark study of survival prediction methods using multi-omics data by inserting joint Dimension reduction methods

Gwendoline Mendes*†[1,2], Arnaud Gloaguen‡[1], and Edith Le Floch§[1]

[1]Centre National de Recherche en Génomique Humaine – CEA – France
[2]CentraleSupélec – CentraleSupélec, Université Paris-Saclay 3 rue Joliot Curie, 91190 Gif-sur-Yvette, France, CentraleSupélec, Université Paris-Saclay, 3 rue Joliot Curie, 91190 Gif-sur-Yvette, France – France

Abstract

Multi-omics studies aim at enhancing our understanding of complex diseases by providing insights into different molecular mechanisms. Yet, the high-dimensionality and the heterogeneity of such multi-omics datasets appears to be challenging. Furthermore, the systematic increase of prediction power when molecular and clinical data are joinlty analysed over exploiting clinical data only has not been consistently established. Though, in the context of survival analysis in 18 distinct cancer datasets, *(Herrmann et al. 2021)* showed that models considering the inherent group structure of multi-omics data can help leverage their full potential.
Our study aims to expand upon this benchmark by incorporating methods able to extract links between each block of omics data. For this purpose, joint Dimension Reduction (jDR) techniques, starting with RGCCA *(Tenenhaus et al. 2011)*, were selected, as they effectively capture shared information across blocks while highlighting the contributions of each omic modality.

Our approach involved sequential modelling, first estimating a reduced space from molecular data only, using unsupervised or supervised strategies following *(Bastien et al. 2008)*. Two values for the regularisation parameter and the number of extracted components were tested. Then, to assess the relative impact of molecular data, clinical data were included or not along the reduced space. Subsequently, a Cox model was trained for the survival prediction.

Our results show that RGCCA coupled with a Cox model, can outperform the methods investigated in *(Herrmann et al. 2021)*. Notably, RGCCA exhibits superior results over the reference model (Cox model on clinical data only) in half of the tested datasets, of which 5 where it outperforms every other method. Interestingly the unsupervised version could achieve performance levels similar to the supervised case when a greater number of components were extracted. Other methods are being added such as JIVE to further investigate the interest of jDR methods in multi-omics studies.
Key words : Multiomics; Joint Dimension Reduction methods; RGCCA; Survival Analysis; Benchmark

*Speaker
†Corresponding author: gwendoline.mendes@student-cs.fr
‡Corresponding author: agloague@cng.fr
§Corresponding author: edith.lefloch@cng.fr

Exploring Immune and Tumor Cells in Gliomas Highly Infiltrated by Lymphocytes through Single-Cell RNA-seq and CITE-seq Data Analysis

Jovana Brocic * [1], Julie Lerond [2], Emeline Mundwiller [3], Franck Bielle [2], Justine Guégan [1]

[1] Data Analyses Core – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France
[2] Genetic and Development of Brain Tumors team, M Sanson/E Huillar – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France
[3] iGenSeq Platform – Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France – France

Gliomas are one of the most lethal types of brain tumors, and immune cells such as lymphocytes have been shown to play a crucial role in their pathogenesis. Since all previous clinical trials and therapeutic approaches did not show any success in glioma patients, there is a need for further research and the development of new therapies. (1).

This study is conducted on four samples collected from patients with different histo-pathological glioma subtypes, all with unusually high lymphocytes infiltration. Fluorescence Activated Cell Sorting (FACS) and 10x Chromium experiment were followed by Illumina Next Generation Sequencing (NGS). We used single-cell RNA-seq (scRNA-seq) and cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) data to investigate the immune and tumor cells. T cell receptor sequencing (TCRseq) will be analyzed for further understanding of T cell clonotype expansion and T cell anti-tumor reactivity.

After quality control steps, filtering, and normalization, scRNA-seq and CITE-seq data were integrated into one assay with Weighted Nearest Neighbor (WNN) method and analyzed in R by using Seurat package, version 4.1.0 (2). Different public data sets - GBM (3) and PBMC (4) - were tested for cluster annotation, but final annotation was set manually according to gene expression levels. The following step was pseudotime analyses of CD8 T cells and CD4 T cells with partition-based graph abstraction (PAGA) (5).

We identified distinct subpopulations of immune and tumor cells based on gene expression profiles and cell surface marker expression, including CD8 T cells, CD4 T cells, Treg CD4 cells, natural killer cells, macrophages, and tumor cells. Beside this, PAGA results allowed us to distinguish between different functional stages of T cells. The detection of rare cell types may provide significant information for diagnosis and treatment.

*Speaker

Currently, tools benchmarking is ongoing for TCRseq data.

Keywords: Diffuse Gliomas, Single, cell RNA Sequencing, Cellular Indexing of Transcriptomes and Epitopes by Sequencing

G-quadruplex structures regulate promoter usage in cells

Rongxin Zhang [1,2], Jean-Louis Mergny *† [1]

[1] École polytechnique – Laboratoire d'Optique et Biosciences – France
[2] Southeast University – China

The tight regulation of gene transcription relies on promoters, and the selection of specific promoters for a particular gene significantly contributes to the diversity of transcripts. However, the precise regulatory mechanisms underlying the selection of promoters remain poorly understood. G-quadruplex (G4), a unique DNA non-canonical secondary structure, has been revealed to be closely associated with gene expression. Yet, whether the formation of G4 structures could potentially result in the expression and alternative use of promoters requires further investigation. In this study, we comprehensively characterized the relationship between G4 structures and promoters based on bioinformatics analysis methods, along with sequencing data including cap analysis of gene expression (CAGE) tag sequencing data. Our results show a close connection between G4 structures and promoter expression, particularly for those showing high expression levels. These G4-associated promoters often exhibit a higher occupancy of transcription factors, while this feature depends on the actual formation of G4 structures (experimentally confirmed G4) rather than the existence of linear sequences compatible with G4 formation. The differential formation of G4 structures between two cancer cell lines influences the use of alternative promoters. Finally, by analyzing RNA-seq data from G4 ligand-treated samples, we confirmed that the formation of G4 structures can induce the generation of alternative promoters. Altogether, our study provides new insights into the mechanisms that drive the formation of alternative promoters.

Keywords: G4 structure, promoter expression, alternative promoter, promoter usage, gene transcription

*Speaker
†Corresponding author: jean-louis.mergny@inserm.fr

Evolutionary signatures of uterine functions in human populations

Eulalie Liorzou *† [1,2], Hanna Julienne [3], Małgorzata Anna Gazda [2], Hugues Aschard [3], Camille Berthelot‡ [2]

[1] École Doctorale Bio Sorbonne Paris Cité [Paris] – Université Sorbonne Paris Cité, Université Paris Cité – France
[2] Institut Pasteur, Université de Paris, CNRS UMR 3525, INSERM UA12, Comparative Functional Genomics group, F-75015 Paris, France. – Institut Pasteur de Paris – France
[3] Institut Pasteur, Université Paris Cité, Department of Computational Biology, F-75015 Paris, France – Institut Pasteur de Paris – France

Understanding how the human genome has evolved since its divergence from chimpanzee and throughout demographic histories can provide critical insight into gene functions by documenting how genetic modifications fare through natural selection. This evolutionary lens is particularly relevant to study the uterus, as genetic variations affecting reproductive functions impact fertility, and therefore, evolutionary fitness. The uterus contains a dynamic mucosa undergoing cyclic differentiative and regenerative processes controlled by hormones. Among these processes, spontaneous decidualisation and menstruation are shared between humans and closely-related primates. However, some traits seem to have evolved specifically in humans, such as the degree of invasiveness of the placenta into the maternal mucosa. Moreover, uterine diseases are frequent across human populations with an average genetic heritability of 30% and affect women from different ancestries with inequal prevalence. The genetic pathways and mechanisms governing uterine physiology and its evolution are still poorly understood. Therefore, we propose to investigate how uterine physiology has impacted genome evolution in the human lineage and between human populations. To identify genomic regions influencing uterine physiology, we use a multitrait analysis combining Genome Wide Association Studies (GWAS) on diverse uterine phenotypes and diseases present in the UK Biobank and Finngen Biobank. We identify a set of single nucleotide polymorphisms (SNPs) associated with uterine conditions and we highlight high genetic correlations between several uterine traits and diseases, some of which confirm previous studies and some of which are novel. Next, we will investigate whether these regions have experienced changes in allele frequency in the last 100 000 years and compare these signatures across human populations. By leveraging recent evolutionary signature of uterine function, we expect to identify coding and non-coding elements crucial for uterine physiology and contributing to genome evolution.

Keywords: Human evolution, GWAS, Population genomics, Uterus, Female reproductive functions

*Speaker
†Corresponding author: eulalie.liorzou@pasteur.fr
‡Corresponding author: camille.berthelot@pasteur.fr

Exploring Positive Selection and Genetic Adaptations in Central and South-East Asian Populations

Johanne Adam * [1], Laurent Romain , Raphaelle Chaix , Laure Ségurel , Evelyne Heyer

[1] Muséum national d'Histoire naturelle – UMR7206 – France

Positive selection plays a crucial role in the dissemination of advantageous traits within populations, contributing to species diversity. However, research on human adaptation driven by positive selection has predominantly focused on Western European and Eastern Asian populations, primarily due to incomplete genomic datasets. Consequently, our understanding of human genetic adaptations to environmental pressures remains limited in other regions of Asia, including Central and South-East Asia.

To bridge this knowledge gap, we conducted a comprehensive analysis utilizing a large genomic dataset of 894 individuals from 27 populations across Central and South-East Asia. Our objective was to provide a broader understanding of adaptive findings encompassing the entire Asian continent, by integrating publicly available genomes from East Asia into our study. Employing the statistical methods iHS and nSL, we aimed to identify loci subject to positive selection within each population.

Through this investigation, we anticipate the identification of novel loci under positive selection that have yet to be reported. Additionally, we expect to observe population-specific signals associated with dietary adaptations, particularly in Central Asian populations that have transitioned from hunter-gatherer lifestyles to farming in recent times. Furthermore, we anticipate the presence of loci under strong positive selection shared across all human populations, such as HLA-B.

In summary, our findings will shed light on the genetic adaptations of Central and South-East Asian populations, offering valuable insights into the evolutionary constraints they have encountered throughout history. By examining the interplay between positive selection and human genetic variation, we enhance our understanding of the complex mechanisms driving human adaptation in diverse geographical regions.

Keywords: Positive selection, Human adaptation, Genomic datasets, Genetic adaptations

*Speaker

Biallelic mutation of OZF1 causes severe oligozoospermia and male infertility in human and mouse

Maha Al Dala Ali*†[1], yasmina Auguste[1], Eric Streichemberger[1], Longepied Guy[1], Cecile Mignon Ravix[1], Catherine Metzler Guillemain[1], Valérie Delague[1], Nicolas Lévy[1], André Megarbane[2], Laurent Daniel[3], and Michael Mitchell‡[1]

[1]U 1251 – Aix Marseille Univ, INSERM, MMG, Marseille Medical Genetics, Marseille, France – France
[2]Institut Jerome Lejeune – Institut Jerome Lejeune, Paris 75015, France – France
[3]Service Anatomopathologie, – Service Anatomopathologie, CHU la Timone, Marseille – France

Abstract

Clinical signs of male infertility are found in about half of the 15% of couples who experience conception difficulties. Oligozoospermia, low sperm count, is a common cause of male infertility. Although the underlying reasons for oligozoospermia remain unknown in most cases, assisted reproductive technology (ART) is frequently used to achieve conception. Nevertheless an oligozoospermia associated with genome instability could carry an increased risk of transmitting *de novo* secondary mutations to the conceptus.
Here, we present a homozygous nonsense mutation in the oligozoospermia factor 1 gene, *OZF1* (laboratory name), in five brothers with severe oligozoospermia. The mouse orthologue, *Ozf1*, has been found to be part of a piRNA-mediated transposon defense system. In mouse, OZF1, enhances the production of secondary PIWI-interacting RNAs (piRNAs). In perinatal prospermatogonia, secondary piRNAs bind to the PIWI protein, MIWI2, in the cytoplasm, and direct it into the nucleus where it targets active LINE-1 promoters for silencing by DNA methylation. C57BL/6 males lacking OZF1 show weakened LINE-1 repression but have normal spermatogenesis and fertility.

We have developed an infertile OZF1-KO mouse model on a mixed genetic background. Our infertile OZF1-KO males show severe oligozoospermia with partial meiotic arrest at zygonema. All the affected animals exhibit strong de-repression of LINE-1 retrotransposons. Interestingly, in newborn OZF1-KO mice, MIWI2 remains cytoplasmic in prospermatogonia, implying that MIWI2 fails to mediate DNA methylation at LINE-1 promoters in our OZF1-KO mice.

Our study shows that reduced genome stability during spermatogenesis can lead to severe oligozoospermia, in human and mouse. We also show that, in mouse, the requirement of OZF1 for male fertility is dependent on genetic background. Our findings have important implications for deciphering the genetic basis of male infertility. Our mouse model provides a means of studying LINE-1 repression and how its failure affects genome and epigenome quality in the gamete.
Key words :infertility ; oligozoospermia ; genome instability ; piARNs

*Speaker
†Corresponding author: maha.al-dala-ali@etu.univ-amu.fr
‡Corresponding author: michael.mitchell@univ-amu.fr

Expanding the phenome and variome of the ROBO-SLIT pathway in congenital heart defects: toward improving the genetic testing yield of CHD

Hager Jaouadi*[1], Chris Jopling[2], Fanny Bajolle , Alexis Théron , Adele Faucherre , Hilla Gerard , Sarab Al Dybiat , Caroline Ovaert , Damien Bonnet , Jean-François Avierinos , and Stéphane Zaffran†

[1]Marseille medical genetics - Centre de génétique médicale de Marseille – Aix Marseille Université, Institut National de la Santé et de la Recherche Médicale – France
[2]Institut de Génomique Fonctionnelle - Montpellier GenomiX – Institut de Génomique Fonctionnelle, BioCampus – France

Abstract

Background Recent studies have shown the implication of the ROBO-SLIT pathway in heart development. Within this study, we aimed to further assess the implication of the ROBO and SLIT genes mainly in bicuspid aortic valve (BAV) and other human congenital heart defects (CHD).
Methods We have analyzed a cohort of singleton exome sequencing data comprising 40 adult BAV patients, 20 pediatric BAV patients generated by the Pediatric Cardiac Genomics Consortium, 10 pediatric cases with tetralogy of Fallot (ToF), and one case with coarctation of the aorta. A gene-centered analysis of data was performed. To further advance the interpretation of the variants, we intended to combine more than 5 prediction tools comprising the assessment of protein structure and stability.
Results A total of 24 variants were identified. Only 4 adult BAV patients (10%) had missense variants in the ROBO and SLIT genes. In contrast, 19 pediatric cases carried variants in ROBO or SLIT genes (61%). Three BAV patients with a severe phenotype were digenic. Segregation analysis was possible for two BAV patients. For the homozygous ROBO4: p.(Arg776Cys) variant, family segregation was consistent with an autosomal recessive pattern of inheritance. The ROBO4: c.3001 + 3G > A variant segregates with the affected family members. Interestingly, these variants were also found in two unrelated patients with ToF highlighting that the same variant in the ROBO4 gene may underlie different cardiac phenotypes affecting the outflow tract development.
Conclusion Our results further reinforce the implication of the ROBO4 gene not only in BAV but also in ToF hence the importance of its inclusion in clinical genetic testing. The remaining ROBO and SLIT genes may be screened in patients with negative or inconclusive genetic tests.

Keywords: Congenital heart defects, Robo, Slit pathway, Genetics, Exome sequencing

*Speaker
†Corresponding author: stephane.zaffran@univ-amu.fr

A proteogenomic analysis of the adiposity colorectal cancer relationship identifies GREM1 as a probable mediator

Matthew Lee *† [1]

[1] Centre international de Recherche sur le Cancer – Centre international de Recherche sur le Cancer – France

Adiposity is an established risk factor for colorectal cancer (CRC). However, the pathways underlying this relationship, and specifically the role of the circulating proteome, is unclear.
Utilizing two-sample Mendelian randomization and colocalization, based on summary data from large sex-combined and sex-specific genetic studies, we estimated the univariable (UV) associations between: (I) adiposity measures (body mass index, BMI; waist hip ratio, WHR) and overall and site-specific (colon, proximal colon, distal colon, and rectal) CRC risk, (II) adiposity measures and plasma proteins, and (III) adiposity-associated plasma proteins and CRC risk. We used multivariable MR (MVMR) to investigate the potential mediating role of adiposity- and CRC-related proteins in the adiposity-CRC association.

BMI and WHR were positively associated with CRC risk, with similar associations by anatomical tumour site. 6,591 adiposity-protein (2,628 unique proteins) and 33 protein-CRC (8 unique proteins) associations were identified using UVMR and colocalization. 1 protein, GREM1 was associated with BMI only and CRC outcomes in a manner that was consistent with a potential mediating role in sex-combined and female-specific analyses. In MVMR, adjusting the BMI-CRC association for GREM1, effect estimates were attenuated - suggestive of a potential mediating role - most strongly for the BMI-overall CRC association in women.
These results highlight the impact of adiposity on the plasma proteome and of adiposity-associated circulating proteins on the risk of CRC. Supported by evidence from *cis*-SNP UVMR and colocalization analyses, GREM1 was identified as a potential mediator of the BMI-CRC association, particularly in women, and warrants further experimental investigation.

Keywords: mendelian randomization, cancer

*Speaker
†Corresponding author: leem@iarc.who.int

Real-time PCR genetic modification detection on the X9™ Real-Time PCR System

Amelie Barthelemy *† [1]

[1] Standard BioTools – Standard BioTools – France

Genetically modified crops are engineered to convey favorable traits such as disease resistance. Labeling and identification of genetic modifications (GMs) in commercial products have become a global concern. There are two standard methods for GM identification: quantitative real-time PCR (real-time PCR) and end-point digital PCR. In this study, we demonstrate the utility of the Standard BioTools X9 Real-Time PCR System with Dynamic Array™ Integrated Fluidic Circuits (IFCs) for real-time GM detection. Real-time analysis provides Ct values that identify false positives and provide greater data confidence. We use the Gossypium hirsutum genome (cotton; C=2.40; GM free) and GM event GHB 614 (614) detected by their respective TaqMan® assays, to determine the limit of GM detection on the 96.96 IFC and 192.24 IFC for real-time PCR without the use of any sort of pre-amplification.

Keywords: Genotyping, GM detection

*Speaker
†Corresponding author: amelie.barthelemy@standardbio.com

The input of patients in clinical research and care. NOTICEInfoBox : an innovative app for creating understandable regulatory-validated information notices.

Flavie Mathieu *† [1], And The " Notices D'information " Working Group

[1] Service sciences et société, département de l'information scientifique et de la communication – Institut National de la Santé et de la Recherche Médicale - INSERM – France

Introduction : Since the 1990s, public agencies have recognized the essential role of patients in the ethical management of research protocols and healthcare. However, progress remains to be made to respond to the societal demand to promote the autonomy of patients in their decision to participate in research or to be an actor in their own healthcare. In clinical research and care, this autonomy is difficult to archieve, as information notices are too often reduced to complicated and hard-to-understand mandatory documents. These considerations prompted the creation of a multidisciplinary working group " Notices d'information " in the fall of 2020, as part of the " College des relecteurs de l'Inserm ". The working group associates the different stakeholders involved in the development, evaluation and use of information notices: health and research professionals, representatives of patient' associations or research foundations, ethicists, jurists, scientific educators and communicators.
Objectives : (1) To describe this emancipation of patients and its consequences on the evolution of research and (2) To present NOTICEInfoBox, an innovative and free tool for clinicians and investigators, to easily create information notices tailored to the people concerned.

Method : Training of association members and establishment of collaborative and multidisciplinary methodologies and working groups.

Results : Associative or institutional review committees have gradually been recognized as one of the key actors in the implementation of research by promoters.

The " Notices d'information " working group has created a set of texts, pictograms and illustrations, adapted to the concerned people and accepted by all stakeholders. These contents will be easily used by professionals through the app NoticeInfoBox. A pilot phase was conducted to generate the notices of the France Genomic Medicine Plan 2025, used for genetic examinations.
Conclusion : Associative or institutional review committees establish patients' legitimacy in the evaluation of clinical research projects. The NoticeInfoBox app is a response to the societal demand to be an actor in its own healthcare and to adopt more ethical and responsible research.

*Speaker
†Corresponding author: flavie.mathieu@inserm.fr

DNA extraction kits and WGS PCR-free library preparation kits have an impact on whole-genome data quality from FFPE samples

Alice Moussy [1], Jasmin Cevost [1], Edouard Turlotte [*†] [1], Mélanie Letexier [1], Margaux Gras [1], Marine Rouillon [1], Kévin Gorrichon [1], Paul Hofman [2], Alain Viari [3], Violette Turon [*]

[1], Jean-François Deleuze [4]

[1] Centre de Référence, d'Innovation, d'eXpertise et de Transfert (CRefIX) – Commissariat à l'énergie atomique et aux énergies alternatives – France
[2] Laboratoire de Pathologie Clinique et Expérimentale. Hôpital Pasteur [Nice] – Hôpital Pasteur [Nice] – France
[3] Inria Lyon – Institut National de Recherche en Informatique et en Automatique – France
[4] Centre National de Recherche en Génomique Humaine – Commissariat à l'énergie atomique et aux énergies alternatives – France

Background/Objectives: The France Genomic Medicine Plan introduces WGS into healthcare pathway for rare diseases and cancers. The Reference Center for Technology, Innovation and Transfer (CRefIX) investigated the impact of preparation kits (extraction, library preparation and reparation) for WGS PCR free protocols on sequencing data accuracy and mutations analysis from FFPE samples.

Methods: Extraction kits (Qiagen, Covaris, 2 kits from Promega), WGS PCR Free library preparation kits (two kits from Illumina, two kits from NEB) with different DNA inputs (1000 - 100 ng) and a DNA repair kit (NEB) were tested on standard FFPE samples and on matched FF and FFPE (block mirrors) from lung cancer tissues. All samples were sequenced on NovaSeq 6000 (2x150 bp, 80X for tumors samples, 40X for gDNA or pbmc). Somatic mutational profiles (SNVs, CNVs, SVs, SBS signature ...) were performed. The F1-score was also reported, FF samples used as gold standard.

Results: FFPE extraction kits affected DNA quantities and quality. One extraction kit created more low frequency artefacts than the ones caused by FFPE. For one WGS PCR Free library preparation kit, insufficient library quantities to sequence were obtained, when artefacts specific to another kit were observed. The FFPE DNA repair kit tested had no effect on pre and post-sequencing results.

[*]Speaker
[†]Corresponding author: turlotte@cnrgh.fr

Conclusion: Introducing FFPE samples into clinical pathway still requires protocols adjustments, specifically for low input FFPE samples. The choice of kits combination is crucial and every kit needs testing as some increase the number of artefacts already high in FFPE samples. With the right kits combination, FFPE samples could be used with caution for clinical interpretation, although it will never reach the same quality as FF.

Grant References: ANR-18-INBS-0001 (French National Research Agency).

Keywords: French Medicine plan 2025, CRefIX, FFPE

The Reference, Innovation, Expertise and Transfer Center (CRefIX) of the French Genomic Medicine Plan 2025

Violette Turon * [1], Mélanie Letexier [1], Jasmin Cevost [1], Kévin Gorrichon [1], Margaux Gras [1], Marine Rouillon [1], Edouard Turlotte [1], Alain Viari [2], Jean-François Deleuze [3]

[1] Centre de Référence, d'Innovation, d'eXpertise et de Transfert (CRefIX) – Commissariat à l'énergie atomique et aux énergies alternatives – France
[2] Inria Lyon – Institut National de Recherche en Informatique et en Automatique – France
[3] Centre National de Recherche en Génomique Humaine – Commissariat à l'énergie atomique et aux énergies alternatives – France

Context & Missions: The Reference, Innovation, Expertise and Transfer Center (CRefIX) is one of the three key elements of the French genomic medicine plan 2025 (FMG2025), alongside the two sequencing platforms, SeqOIA and Auragen, and the Data Collector Analyzer (CAD). CRefIX is a strategic R&D center, building on the best expertise of the CEA, Inria, Inserm and the industrial world, to ensure maximum deployment of genomic medicine in the short, medium and long term. The mission of CRefIX is twofold: 1) to establish benchmarks and standards to be implemented in research projects and on diagnostic production sites: research pilots, genomic sequencing platforms, and data collection analyzers; 2) to innovate and accelerate transfers in cooperation with industrial players, in order to ensure the competitiveness of the system and to develop a national industrial sector. CRefIX has its own resources and is hosted by the National Center for Human Genomics Research (CNRGH) in order to create synergies and avoid duplications.

Ongoing and Future Projects: The multidisciplinary team of the CRefIX is currently or will soon work on: 1) the best practices to minimize bias caused by formalin when working with formalin-fixed paraffin embedded (FFPE) tumor tissue samples in a clinical context, 2) the new kits and automation systems specific to low input samples for WGS PCR Free procedures, 3) the benefit of long read sequencing for SV detection in the context of cancer and rare diseases in collaboration with the UK plan; 4) the evaluation of biological standards (cell lines, biopsies...) to optimize the detection of any type of mutation independently of the technology used and improve the comparability of the data between platforms, and 5) the best practices for high throughput ctDNA sequencing (large panels or WES).

Grant References: ANR-18-INBS-0001 (French National Research Agency

*Speaker

Optimization of low-coverage sequencing approach

Christian Daviaud*[1], Zuzana Gerber[1], Stéphane Meslages[1], Jeanne-Antide Perrier-Cornet[1], Johann Tassin[1], Delphine Bacq-Daian[1], Vincent Meyer[1], Anne Boland[1], Jean-François Deleuze†[1], and Robert Olaso‡[1]

[1]Centre National de Recherche en Génomique Humaine – CEA – France

Abstract

Low-coverage sequencing is a novel genotyping technique combining whole genome sequencing to an average depth of 1× with genotype imputation to identify genetic variants (1). Low-coverage sequencing is a cost-effective alternative to genotyping arrays to identify genetic variants. In this study, we compare different low-coverage sequencing methods. Sequencing libraries were prepared from a number of DNA reference standards, including an Ashkenazi Jewish trio, using miniaturized assays adapted from different commercial suppliers that have been optimized in our laboratory. Libraries were pooled to obtain low-coverage and sequenced on a NovaSeq 6000 (Illumina). Coverage metrics were assessed using our in-house bioinformatic pipeline.

We obtained homogeneous results for all libraries. Coverage was relatively homogeneous across all chromosomes. The "Illumina DNA Prep" process has shown the best results. We showed that the protocol significantly improved the homogeneity of depth of coverage. We examined the gaps in the genome coverage as they can potentially impact the accuracy of imputation. The Illumina process is now fully automated in our laboratory, with a seven-fold miniaturization.

This is a comparative study to evaluate our low-coverage sequencing approach. The new miniaturized process has been optimized for automation to be compatible with high-throughput sequencing platforms.

References: (1) Li et al. 2021. Genome Res 31(4):529-537.

Grants :LabEx GenMed (Medical Genomics)
Key words : Low coverage sequencing ; Genome sequencing

Keywords: Low coverage sequencing, Genome sequencing

*Speaker
†Corresponding author: deleuze@cng.fr
‡Corresponding author: olaso@cng.fr

Co-transcriptional cis-R-loop forming lncRNAs: a new lncRNA subclass?

Kevin Muret *† [1], Jean-François Deleuze , Eric Bonnet

[1] Centre National de Recherche en Génomique Humaine – Commissariat à l'énergie atomique et aux énergies alternatives – France

For more than a decade, lncRNAs have been the subject of many research fields. However, these non-protein-coding entities of more than 200 nucleotides represent a wide diversity of RNAs with very different roles and, despite efforts to subclassify these genes according to their genic environment, do not allow us to obtain subclasses of lncRNAs based on their function. LncRNAs are able to interact with other RNAs, DNA, peptides or proteins. Here, we focused on RNA:DNA interactions (R-loops) which can be studied by DRIP-seq based on the S9.6 antibody's high affinity for R-loops. Based on more than 160 DRIP-seq experiments and lncRNA annotation, we were able to show that 49% of lncRNAs are likely to form a cis-R-loop. We have also identified 1367 lncRNA/coding gene pairs for which we suspect a role for the lncRNA in regulating the expression of the nearby coding gene. The VIM/VIM-AS1 pair, a well-known case described by Boque-Sastre *et al.*, is also retrieved. These initial results are very promising; they will require experimental validations in the coming years. We hope, through this original approach, to annotate more precisely and subclassify lncRNAs in order to help researchers to envisage more adapted experimental methods for their functional studies.

Keywords: lncRNA, R, loop, DNA:RNA interaction, expression, regulation

*Speaker
†Corresponding author: kevin.muret@cnrgh.fr

Development of methylation and fragmentation workflows to identify potential biomarkers of tissue-of-origin in circulating cell-free DNA

Nouara Oussada * [1], Alexia Rabec [1], Mélanie Gou [1], Caroline Horgues [1], Jean-François Deleuze [1], Florence Mauger * † [1]

[1] Centre National de Recherche en Génomique Humaine – Commissariat à l'énergie atomique et aux énergies alternatives – France

In the context of precision medicine, the identification of novel non-invasive biomarkers such as circulating cell-free DNA (cfDNA) from plasma, urine and cerebrospinal fluid, is crucial for the diagnosis, prognosis and monitoring response to treatment of complex diseases such as cancer and neurodegenerative diseases.
In particular, the analysis of fragmentation and methylation of cfDNA could be promising biomarkers for identifying the tissue-of-origin (TOO). Indeed, the fragmentation of cfDNA reflects the chromatin structure of its TOO, especially at transcription start sites (TSS), allowing the inference of gene expression and the DNA methylation is specific of its TOO.

The analysis of cfDNA is challenging due to its low concentration and high fragmentation and it requires the development of specific methods to analyze sequencing data. Several approaches for fragmentation analysis have been developed, based on fragment size analysis, fragment end analysis and nucleosome profiling.

A methylation workflow was developed including: data treatment, quality-control step, differential methylation analysis in CpG sites or regions and deconvolution analysis to identify potential biomarkers.

Furthermore, two fragmentation workflows were developed: for TSS and for whole-genome analyses. The TSS workflow includes, data treatment, calculation of TSS features, quality-control step, selection of TSS features, and machine learning approaches to predict the expression status of genes in cfDNA. The whole-genome included: data treatment, scoring calculation (sizes profiling, end-motifs, fragment depths and size ratios), quality-control step and deconvolution and machine learning method to identify potential biomarkers.

Finally, we developed a complete workflow for the integration of fragmentomic and methylomic data from whole methylome sequencing data based on enzymatic conversion using an unsupervised multi-omics method to identify potential biomarkers.

*Speaker
†Corresponding author: florence.mauger@cng.fr

The combination of fragmentomic and methylomic analysis of cfDNA could improve the analysis of TOO in cfDNA to identify potential non-invasive biomarkers of complex diseases.

Keywords: circulating cell free DNA, methylomic, fragmentomic, sequencing data workflow, non invasive biomarkers, tissue of origin.

Comprehensive multiparametric analysis of cell-free circulating nucleic acids from a single extraction in relation to aging

Nicolas Tessier * [1], Lise M. Hardy [1], Antoine Daunay [1], Jean-François Deleuze [1,2], Alexandre How-Kit [1]

[1] Laboratory for Genomics, Foundation Jean Dausset – CEPH – Fondation Jean Dausset CEPH – France

[2] Centre National de Recherche en Génomique Humaine, CEA, Institut François Jacob – CEA Paris-Saclay, Institut François Jacob, Centre National de Recherche en Génomique Humaine – France

Circulating cell-free nucleic acids in plasma (ccfNAPs) are released into the bloodstream by cells from various tissues and organs through putative mechanisms such as apoptosis, secretion or necrosis. Circulating cell-free DNA (ccfDNA) and microRNAs (ccfmiRNAs) from plasma are the two components of ccfNAPs that have been most studied to date in many pathologies, while very few studies have focused on ccfNAPs in healthy individuals.

We here study the variations of a set of molecular parameters of circulating blood plasma nucleic acids (ccfNAPs) in a cohort of 140 healthy donors from the French Blood Bank (EFS) aged between 19 to 66 years old. We used a single extraction of all ccfNAPs, including DNA (nuclear and mitochondrial) and RNA (messenger, ribosomal and micro-RNA) combined to high resolution analysis methods. PCR assays targeting high copy number genomic elements (Line-1/Kpn, mtDNA and rRNA) were used to minimize the volume of ccfNAPs required for all analysis. They included real-time quantitative PCR, ultra-sensitive high-resolution capillary electrophoresis and pyrosequencing. These methods have been used to quantify and assess the integrity of ccfNAPs as well as quantification of global DNA methylation and of expression of a few candidate tissue-specific mRNAs and age-associated miRNAs.

Our study provides a workflow for multiparametric analysis of all types of ccfNAPs from a single extraction. These methods allowed to highlight the age-related molecular changes occurring in ccfNAPs as well as their inter-individual variability.

Keywords: Circulating cell free nucleic acids, plasma, aging, longevity

*Speaker

Inferring microsatellite mutation rates from allele frequencies obtained from (non-)isothermal DNA amplification approaches using thirty-two genetic models

Schayma Ben Marzougui[1], Sophie Jeanjean*[1], Antoine Daunay[1], Jean-François Deleuze[1,2], and Alexandre How-Kit†[1]

[1]Laboratory for Genomics, Foundation Jean Dausset – CEPH, Paris, France – Laboratory for Genomics, Foundation Jean Dausset – CEPH, Paris, France – France
[2]National Center of Human Genomics Research, CEA, François Jacob Institute of Biology, Evry, France – National Center of Human Genomics Research, CEA, François Jacob Institute of Biology, Evry, France – France

Abstract

Microsatellites, also known as Short Tandem Repeats (STRs), are sequences composed of 1 to 6 nucleotides repeated in tandem. Their high degree of polymorphism is due to the "Polymerase Slippage" mechanism, which allows the gain or loss of repetitions during DNA replication. In our study, we aimed to estimate microsatellite mutation rates by analyzing allele frequencies obtained from in vitro PCR and iso-thermal RPA (recombinase polymerase amplification) experiments and to identify the underlying genetic model for polymerase slippage.

We developed a method that infers microsatellite mutation rates using 32 genetic models. This model differs in the relationship between mutation rates and microsatellite allele length, as well as the number of steps allowed to be inserted or deleted during a slippage event. Our approach is based on the minimization criteria of the RMSE (Root Mean Square Error) between observed and simulated microsatellite allele frequencies using the Simulated Annealing algorithm (a stochastic optimization algorithm) combined to a Grid Search, to infer mutation rates and identify the genetic model best explaining the observed microsatellite mutation profiles.

Our method was successfully tested on experimental data generated with PCR and RPA experiments using 4 synthetic mono-nucleotide repeat microsatellites (A15, A19, A20 and A24) and 4 di-nucleotide repeat microsatellites (AC15, AC19, AC20 and AC24). This method should be a valuable tool for the study of microsatellite mutational mechanism.

Keywords: Microsatellites, Polymerase slippage, Short tandem repeats, microsatellite mutation rates, PCR, Recombinase Polymerase Amplification

*Speaker
†Corresponding author: alexandre.how-kit@fjd-ceph.org

List of participants

- Abriol Julien
- Adam Johanne
- Al Dala Ali Maha
- Aloui Chaker
- Amblard Elise
- Amoyal William
- Andre Catherine
- Asgari Yazdan
- Aubier Benjamin
- Auboeuf Didier
- Barbry Pascal
- Barthélemy Amélie
- Barthome Eli
- Batsche Eric
- Batut Aria
- Baulande Sylvain
- Berdou François
- Bertossi-Chazeirat Guillaume
- Besnier Nicolas
- Bihoreau Marie-Thérèse
- Blanché Hélène
- Blum Yuna
- Blunt Sierra
- Bocher Ozvan
- Bohec Mylène
- Boland Anne
- Bonnet Eric
- Bordes Constance
- Bouvet Delphine
- Brocic Jovana
- Brohard Solène
- Caro Ilana
- Carrier Maud
- Cavalli Giacomo
- C´evost Jasmin
- Champramary Simang
- Chantalat Sophie
- Charon Celine
- Chaudru Valérie
- Chauffert-Yvart Thibault
- Chedotal Alain
- Chettouh Hamza
- Chowdhury Tafsir
- Cliquet Freddy
- Costantino Félicie
- Coste Thibault
- Couvelard Linhda
- Cucchiarini Anne
- Dabbadie Anaïs
- Dandine-Roulland Claire
- Daviaud Christian
- De Massy Bernard
- Debette Stephanie
- Deleuze Jean François
- Djebali Sarah
- Domenighetti Cloé
- Duchesnay Edouard
- Dulary Cécile
- Egan Coailinn
- El Hage Perla

- Fakir Jamal
- Fauchereau Fabien
- Fleury Christophe
- Fleury Mathis
- Freby Cecile
- Geigl Eva-Maria
- Genin Emmanuelle
- Gerber Zuzana
- Gloaguen Arnaud
- Gorrichon Kévin
- Gourdon Genevi`eve
- Grange Thierry
- Gras Margaux
- Grellety Gautier
- Groussard Karine
- Hamon Mélanie
- Hanafi Latifa
- Hardy Lise
- Hedan Benoit
- Herzig Anthony
- Holt Sam
- How Kit Alexandre
- Jabot-Hanin Fabienne
- Jaouadi Hager
- Jeanjean Sophie
- Job Sylvie
- Journot Laurent
- Juiz Natalia
- Karimi Mojgan
- Kerbrat Stéphane
- Kora Hafid
- Lagarrigue Sandrine
- Langou-Quentin Karine
- Le Floch Edith
- Le Goff Vincent
- Le Nézet Louis
- Leap Katie
- Leblanc Marion
- Leblond Claire
- Lee Matthew
- Letexier Mélanie
- Leutenegger Anne-Louise
- Liorzou Eulalie
- Laure Segurel
- Maillard Valérie
- Marenne Gaëlle
- Martignetti Loredana
- Martin Caitlin
- Maî¾egot RaphaCel
- Mathieu Flavie
- Mauger Florence
- Mazoyer Sylvie
- Mendes Gwendoline
- Mergny Jean-Louis
- Meyer Vincent
- Millot Gaël
- Milpied Pierre
- Muchardt Christian
- Muret Kévin
- Nicklen Sukhvinder

- Nicolini Frederic
- Olaso Robert
- Oussada Nouara
- Parasayan Oguzhan
- Parisot Anne-Lyne
- Park Seehyun
- Patin Etienne
- Perdereau Aude
- Philippe Cathy
- Proux Clement
- Pruvost Mélanie
- Pulcini Françoise
- Rachez Christophe
- Radjasandirane Ragousandirane
- Raynal Virginie
- Richard Magali
- Rouillon Marine
- Roux Maguelonne
- Roy Robin
- Saint-Martin Cécile
- Schott Jean-Jacques
- Severac Dany
- Souaifan Hiba
- Stepanoff Manon
- Stevanin Giovanni
- Sugier Pierre-Emmanuel
- Teixeira Elisabeth
- Tessier Nicolas
- Thomas Derrien
- Tost Jorg
- Tournaire Martin
- Tournier-Lasserve Elisabeth
- Tragin Margot
- Truong Thérèse
- Turlotte Edouard
- Turon Violette
- Vallot Céline
- Vandiedonck Claire
- Velo Suarez Lourdes
- Vert Jean-Philippe
- Vinasco Sandoval Gloria Tatiana
- Wincker Patrick
- Zaragosi Laure-Emmanuelle
- Zhang Rongxin
- Zucman-Rossi Jessica
- Zulcinski Michal

Author Index