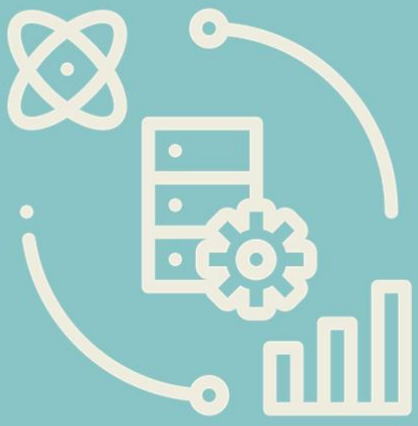
WORKSHOP

Computational analysis and multiomics in oncology

MONDAY 17TH OCTOBER (9AM - 5.30PM)

IUCT-ONCOPOLE (AMPHITHEATRE)

Keynote lectures :
Eduard Porta (IJC, Barcelona)
Nicolas Gaudenzio (Infinity, Toulouse)



Contact :
Agathe Redouté - agathe.redoute@inserm.fr
Axel Arthur - axel.arthur@inserm.fr
Vera Pancaldi - vera.pancaldi@inserm.fr

Programme

9h **Introduction**

9h15 ***Keynote Lecture*** **Eduard Porta**, IJC Barcelona
Understanding oncogenesis across biological scales

10h00 **David Simoncini**, IRIT
Artificial intelligence-based prediction models for acute myeloid leukemia using real-life data

10h30 *Coffee Break*

11h00 **Pauline Gravelle**, CRCT
Loss of synchrony during differentiation of sMZL B cells: demonstration through multimodal analysis including Spatial Transcriptomics

Frédéric Pont, CRCT
Single-cell Spatial Explorer: Easy exploration of spatial and multimodal transcriptomics

12h00 *Lunch Break*

14h00 ***Keynote Lecture*** **Nicolas Gaudenzio**, Infinity
Deconvolution of human skin immune architecture with Multiplex Annotated Tissue Imaging System (MANTIS)

14h45 **Nathaniel Polley**, CRCT
Follow the Trees: Novel XGBoost Forest Algorithm Elucidates Unforeseen Mechanisms of Action in Cancer Survival from Diagnostic Metabolic Flux Data

15h15 **Sébastien Bourdon**, CRCT
QUADRatlas: the RNA G-Quadruplex and RG4-binding proteins database

15h45 *Coffee Break*

16h15 **Matthieu Genais**, CRCT
Multi-omics prediction of clinical response to immunotherapy in advanced melanoma

16h45 **Vera Pancaldi**, CRCT
Characterisation and modelling of the tumour micro-environment

Keynote Lecture

Eduard Porta (IJC, Barcelona)

"Understanding oncogenesis across biological scales"

Abstract:

Tumors are complex ecosystems of different cell types, ranging from healthy cells from the original tissue to cancer cells that acquired malignant properties throughout somatic evolution. The final behavior of a tumor depends on the interplay between all these different cell types which, in turn, have different properties depending on both, the unique germline genome of each person as well as the distinctive set of somatic mutations acquired by cancer cells. In this talk I will present the recent work of our group, including analysis of proteogenomic data and spatial transcriptomic profiles, to understand how all these different phenomena interact with each other across omics layers and biological scales to determine the final properties of human tumors.

David Simoncini (IRIT, Toulouse)

"Artificial intelligence-based prediction models for acute myeloid leukemia using real-life data"

Abstract:

The DATAML registry compiles clinical and biological data of over 3000 Acute Myeloid Leukemia (AML) patients treated in Toulouse and Bordeaux University Hospitals with either intensive chemiotherapy or azacitidine. We designed artificial intelligence-based methods for overall survival prediction and treatment decision predictions. Using Boruta and SHAP, we were able to reduce the number of relevant features for predictions and show the impact of these features on the outcome of the predictive models. This talk will present the results we achieved for these prediction tasks, along with some recent preliminary work using diagnostic bone marrow smears scans from DATAML registry patients.

Co-presentation

Pauline Gravelle (CRCT, Toulouse)

"Loss of synchrony during differentiation of sMZLB cells: demonstration through multimodal analysis including Spatial Transcriptomics"

Abstract:

Splenic marginal zone lymphoma (sMZL) is a very rare small B cell malignancy which is still poorly described. Our study aimed at characterizing this disease through multimodal assays, realized on paired samples (blood/spleen) from 3 patients. First, single cell RNA sequencing (scRNAseq), CITE-seq 3' chemistry, was used to determine the gene expression profile (GEP) of tumor and microenvironment cells. Then, scRNAseq 5'-VDJ chemistry was performed to compute GEP data with B and T lymphocytes clonality characterization. Altogether, these scRNAseq data suggest a loss of synchrony of sMZL B cells during differentiation, but a homogeneous and conserved microenvironment between blood and spleen. Finally, spatial transcriptomics was performed in order to associate the previous scRNAseq data with topographical informations. Spatial data exploration, with the use of spatial explorer software (Pont F., & al, BioRxiv 2022), allowed us to visualize how B and T cells are organized and structured in sMZL tissues, and how these cells are characterized in terms of molecular functions.

Frédéric Pont (CRCT, Toulouse)

"Single-cell Spatial Explorer: Easy exploration of spatial and multimodal transcriptomics"

Abstract:
The development of single cell technologies yields large datasets of informations as diverse and multimodal as transcriptomes, immunophenotypes, and spatial position from tissue sections in the so-called "spatial transcriptomics". Currently however, user-friendly, powerful, and free algorithmic tools for straightforward analysis of spatial transcriptomic datasets are scarce. Here, we introduce Single-Cell Spatial Explorer, an open-source software for multimodal exploration of spatial transcriptomics (https://github.com/FredPont/spatial https://www.biorxiv.org/content/10.1101/2022.08.04.502890v1)

Keynote Lecture

Nicolas Gaudenzio (Infinity, Toulouse)

"Deconvolution of human skin immune architecture with Multiplex Annotated Tissue Imaging System (MANTIS)"

Abstract:

Routine clinical assays, such as conventional immunohistochemistry, often fail to resolve the regional heterogeneity of complex inflammatory skin conditions. Here we introduce MANTIS (Multiplexed Annotated Tissue Imaging System), a flexible analytic pipeline compatible with routine practice, specifically-designed for spatially-resolved immune phenotyping of the skin in experimental or clinical samples. Based on phenotype attribution matrices coupled to α-shape algorithms, MANTIS projects a representative digital immune landscape, while enabling automated detection of major inflammatory clusters and concomitant single-cell data quantification of biomarkers. We observed that severe acral lesions from systemic lupus erythematosus, Kawasaki syndrome or COVID-19-associated skin manifestations share common quantitative immune features, while displaying a non-random distribution of cells with the formation of disease-specific dermal immune structures. Given its accuracy and flexibility, MANTIS is designed to solve the spatial organization of complex immune environments to better apprehend the pathophysiology of skin manifestations.

Nathaniel Polley (CRCT, Toulouse)

"Follow the Trees: Novel XGBoost Forest Algorithm Elucidates Unforeseen Mechanisms of Action in Cancer Survival from Diagnostic Metabolic Flux Data"

Abstract:

Modern in vivo methods involving patient-derived xenograph models have facilitated countless advancements in experimental design due to the ability to apply multiple treatment combinations in parallel to existing human tissue. Nonetheless, the preliminary usage of in vivo screens in search of new biological targets in cancer research remains expensive, time-consuming, and rarely produces a sample size large enough for statistical significance to be inferred. As a consequence, the demand persists for an alternative method to accurately discover therapeutic mechanisms of action based upon genotypic and demographic variation directly from patient samples. Fortunately, the incorporation of metabolic flux data with modern machine learning methods has led to recent breakthroughs in survival prediction among AML patients. PolyMORPHOS (Metabolic-Oriented Regression Predicting Human Overall Survival) derives metabolism flux balances using bulk patient RNAseq data at diagnosis in order to calibrate a progressive XGBoost model, therefore providing a diverse metabolic landscape to predict the fate of patients at early, middle, and late stages of survival. Per each survival iteration, polyMORPHOS calculates the metabolic reactions most pertinent to death prediction, thus enabling researchers and clinicians to design individual therapies to best reverse any unfavorable prognosis with needlepoint precision. Additionally, we announce SOURIS (Simulation Of Universal Reactions In Silico), the most recent installment to our predictive analytics suite. Using the previous polyMORPHOS model, chronological metabolic survival conditions are simulated hundreds of times by random-sampling the initial patient cohort. Within each simulation, the cohort is stratified based upon high, baseline, or low expression of a variable chosen by the researcher e.g. gene, protein, age, mutation status, etc. The resulting metrics provide insight as to which metabolic reaction states result in morbidity or favorable outcome per each stratified patient characteristic. Hence, the precise elucidation of relevant mechanisms of action at early stages of the experimental process will empower researchers to strategically dedicate more resources to effective therapeutic design, thus further bridging the gap between theoretical and tangible results.

Sébastien Boudon (CRCT, Toulouse)

"QUADRatlas: the RNA G-Quadruplex and RG4-binding proteins database"

Abstract:

RNA G-Quadruplexes (RG4s) are non-canonical structures that have been increasingly recognized as fundamental posttranscriptional regulators of gene expression *(doi: 10.1016/j.tibs.2020.11.001)*. These elements are indeed able to affect cell physiology and pathology thanks to their dynamicity and wide array of interactors. In particular, they have been associated with the onset, progression, and therapy resistance in human cancers by us *(doi: 10.1038/s41467-020-16168-x, 10.1084/jem.20210571*) and several others. The current view is that RG4s are dynamic structures whose folding equilibrium and function are driven by RNA-binding proteins *(doi: 10.1126/science.aaf5371, 10.1038/s41467-018-07224- 8)*. Our previous work uncovered a set of such proteins as potential RG4 interactors, suggesting that their regulatory network is far wider than initially expected. Being able to explore interactions with RNA-binding proteins (RBPs) and the role of RG4 in post-transcriptional control of gene expression is thus paramount to understanding how RG4s are regulated and exploiting them as potential therapeutic targets. To empower this analysis, we thus built QUADRatlas, a database including experimentally-derived (using publicly available RG4-seq datasets) and computationally predicted RG4s (using three tools and their consensus as a "golden standard") in the human transcriptome. We enriched these datasets with annotations describing the biological function and disease phenotype associations for RG4s, as obtained from the literature. Users of our database can thus explore RG4s in specific transcripts of their interest, obtaining details on their function and pathological potential. The database currently includes only human data, but we plan to expand it to the mouse and potentially other organisms. To this end, the underlying software is already able to accommodate multiple organisms seamlessly. Recognizing that the interaction of RG4s with RBPs is key to their function, we then mined known interactions of RG4s with such proteins, complementing them with the extensive dataset of ENCODE eCLIP assays. We thus provide binding sites for many RBPs and allow our users to intersect the RG4s with their potential regulators and effectors. We thus enable the formulation of novel hypotheses on RG4 regulation, function,

and pathogenicity. This data is browsable through a genome-browser plugin and table views. Furthermore, search results can be downloaded easily from the same interface. Finally, to make the exploration of this vast amount of data easier, and empower hypothesis generation, we provide several analysis capabilities. In particular, the "overlap" function allows extracting the binding sites of user-specified RBPs with RG4s, computing a degree of overlap. This function thus can predict whether an RBP could also bind RG4s. From the opposite perspective, the "enrichment" function allows computing whether RG4s are enriched in a set of transcripts of interests (e.g., differentially expressed in some high-throughput experiment) and thus predicts their potential involvement with the observed phenotypes. Finally, we recognize the need to understand whether a specific sequence, not forcedly part of the human transcriptome, could contain an RG4. To meet this need, we thus offer the possibility of de novo prediction of RG4 on user-input sequences through three different algorithms, presenting the results in a sequence browser plugin and making them fully downloadable.

Matthieu Genais (CRCT, Toulouse)

"Multi-omics prediction of clinical response to immunotherapy in advanced melanoma"

Abstract:

Immune checkpoint inhibitors (ICI) such as anti-PD-1 act on T cells to restore their ability to kill cancer cells. Cutaneous melanoma is a poor-prognosis skin cancer that can be treated by ICI. Despite major advances in the field of immunotherapy, melanoma kills half of patients within 5 years of treatment induction, due to primary or acquired resistance. Over the last 10 years, our team has identified a mechanism of resistance to immunotherapy that depends on the production of TNF, a major inflammatory cytokine that acts as a brake on the immune response against tumours in mouse melanoma models and could ameliorate adverse events caused by immunotherapy treatment. Here, we present an analysis of bulk transcriptomics on a preclinical mice model and public patient cohorts using inferred TF and pathway activities to define characteristics of responders (both in mice and human).

Vera Pancaldi (CRCT, Toulouse)

"Characterisation and modelling of the tumour micro-environment"

Abstract:

Given the recent discoveries about the importance of specific interactions between immune and cancer cells, understanding the spatial properties of tumors at single-cell resolution becomes crucial. We have recently developed computational tools to describe spatial patterns and clustering of specific cell types in tissues using network theory. These approaches allow us to extract statistical properties from imaging or spatial-omics datasets that can be used as biomarkers. We will describe software to extract cellular networks from imaging or spatial-omics datasets and analyze spatial patterns by defining cellular neighborhood states across lung cancer samples. Some of the identified patterns can constitute biomarkers and aid in prediction of response to immunotherapy.

Finally, we will discuss our planned strategy to achieve an understanding of the dynamics of cellular interactions and phenotype transitions that can help us better target therapies in a personalized medicine approach. These models involve a combination of agent-based and logic modeling to represent processes at play in the tumor microenvironment at different levels. A major challenge remains in bridging the gap between literature based model construction and purely data-driven approaches, which are necessary to make the models patient-specific and clinically relevant.

ARe
Cancer, Ageing & Rejuvenation
Graduate school

CRCT
Centre de Recherches en Cancérologie de Toulouse
ONCOPOLE

UNIVERSITÉ TOULOUSE III PAUL SABATIER
Université Fédérale
Toulouse Midi-Pyrénées

Inserm
La science pour la santé
From science to health

cnrs

17th october 2022
IUCT-Oncopole, 1 avenue Joliot-Curie

Contact :
Agathe Redouté - agathe.redoute@inserm.fr
Axel Arthur - axel.arthur@inserm.fr
Vera Pancaldi - vera.pancaldi@inserm.fr