



**Corso di dottorato in Scienze Biomolecolari**  
***PhD in Biomolecular Sciences***  
**Ciclo 41 / Cycle 41**  
**A.Y. 2025-2026**

**Reserved scholarship A**

Progetti finanziati nell'ambito dei Dipartimenti di Eccellenza 2023-2027.

**Curriculum Biologia Quantitativa**

Il vincitore sceglierà il progetto di ricerca dall'elenco sottostante.

*MUR-funded grants - Departments of Excellence 2023-2027.*

**Quantitative Biology Curriculum**

*The winner will choose from the list available below the research project.*

| Principal Investigator                       | Project title                                                                                                                                                                                                                                                             |
|----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>1 - Fulvio Chiacchiera</b>                | <i>Unveil differentiation trajectories and cell-cell communication during liver regeneration and tumour progression / Caratterizzare i processi differenziativi e le interazioni tra cellule durante la rigenerazione epatica e la progressione tumorale</i>              |
| <b>2 - Martin Hanczyc &amp; Toma Tebaldi</b> | <i>Noise as signal: coupling gene expression variance, metabolic fluxes and biological entropy in human disease / Il rumore come segnale: esplorare la connessione tra varianza dell'espressione genica, flussi metabolici ed entropia biologica nelle malattie umane</i> |
| <b>3 - Toma Tebaldi &amp; Luca Tiberi</b>    | <i>Building a Map of Human Specific Genes with Single-cell resolution (HuMap) / Costruire una mappa di geni "human specific" a risoluzione di singola cellula (HuMap)</i>                                                                                                 |
| <b>4 - Luca Marchetti</b>                    | <i>Exploring Cell-Type-Specific Metabolic Signatures in Breast Cancer through Transcriptome Deconvolution / Studio delle firme metaboliche a livello cellulare nel tumore al seno tramite deconvoluzione del trascrittoma</i>                                             |



## Project 1

### Unveil differentiation trajectories and cell-cell communication during liver regeneration and tumour progression

Caratterizzare i processi differenziativi e le interazioni tra cellule durante la rigenerazione epatica e la progressione tumorale

#### Laboratory:

Laboratory of Stem Cells and Cancer Genomics (<https://www.cibio.unitn.it/956/laboratory-of-stem-cells-and-cancer-genomics>)

Principal Investigator: Fulvio Chiacchiera ([fulvio.chiacchiera@unitn.it](mailto:fulvio.chiacchiera@unitn.it))

#### Keywords:

#### Synthetic description of the activity and expected research outcome

##### Introduction

Liver regeneration is a complex and still largely uncharacterized process. Under chronic stress conditions defined set of cells known as liver progenitor-like cells contribute to the regeneration of both hepatocytes and cholangiocytes. This process is activated and progresses through the interaction of many different cell types including immune cells, stellate cells, hepatocytes and cholangiocytes. Understanding how these cells interact, and which differentiation programs became activated is of crucial importance considering that proficient liver regeneration is of paramount importance to preserve liver integrity preventing liver tumour formation. Similar cell-cell interactions and differentiation trajectories have been suggested to have a role during liver tumour progression. Every year 13000 new hepatocellular carcinoma cases are diagnosed in Italy. Lungs are the most frequently observed distal site that can be colonized by liver tumour cells<sup>1</sup>. For most of the tumours, metastasis dissemination is an early event indicating that gradual accumulation of mutations during tumor growth is largely insufficient to explain cell dissemination and metastasis formation. These data suggest that metastatic potential are intrinsic features of the primary tumour and the result of an adaptive response to specific local or systemic environmental pressures<sup>2, 3</sup>. The possibility to identify potentially metastatic clones already present in the primary tumours and understand how they interact with all the other cells in the tumour microenvironment, would be a turning point for prognosis and therapy.

##### Aim

The goal of this project is to perform bioinformatics analyses to characterize cell-cell interaction and novel differentiation intermediates during liver regeneration and tumour progression. By analysing published and in-house produced single cell transcriptomics and spatial transcriptomics data, we aim at characterizing the crosstalk between immune cells and hepatocytes during liver regeneration, as well as between immune cells and cancer cells in both primary and metastatic lesions. We seek to understand how hepatocytes and immune cells reciprocally reshape the transcriptional and epigenetic landscape to promote liver progenitor cells expansion, cell dissemination or dormant lesions reactivation.

##### Experimental approach

We will take advantage of different models as well as already published datasets. Widely used mouse models for studying liver regeneration are already available in the lab. To investigate the mechanisms involved in metastasis formation, we developed a novel hepatocellular carcinoma mouse model in which the growth of sporadic primary tumours is followed by the formation of lung metastasis (proliferative macro- and dormant micrometastasis). This is similarly to what is observed in humans<sup>4</sup>. The student will perform bioinformatics analysis of spatial transcriptomics and proteomics data to characterize primary tumours and metastatic lesions in the mouse model. Moreover, single cell transcriptomics generated in the lab will be analysed to further characterize the immune cells associated with these lesions<sup>5</sup>. Spatial proteomics will be performed using laser-microdissection coupled to mass-spectrometry analysis to isolate and characterize specific regions of primary and metastatic lesions. For spatial transcriptomic we plan to use commercially available platforms and for single cell transcriptomic we will rely on the newly established single-cell facility at CIBIO. The single cell and spatial transcriptomics data will be further analysed to study cell-cell communications and cell trajectories in the pre-tumoral, primary and metastatic lesions. Indeed, the data obtained will be integrated with the analysis of the genetic and epigenetic alterations in the different cells and the role of immune infiltrate investigated using specific bioactive molecules (e.g. anti-CTLA4/Anti-PD1, clodronate loaded liposomes) and antibodies-based selective ablation. The translatability of the data obtained will be explored in collaboration with the pathology unit of S. Chiara Hospital.

##### Expected results

The computational analyses on the mouse models will shed light on the role of the liver and tumor microenvironment, in particular immune cells, in shaping cancer cell epigenome and transcription (and vice versa) during liver regeneration and tumour progression. The integration of spatially resolved transcriptional data, single cell-transcriptomic and genomics data will help to identify specific traits and key players involved in these processes with potential clinical applications.

##### References:

1. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med* 2019;380:1450-1462.
2. Vendramin R, Litchfield K, Swanton C. Cancer evolution: Darwin and beyond. *EMBO J* 2021;40:e108389.
3. Vegliante R, Pastushenko I, Blanpain C. Deciphering functional tumor states at single-cell resolution. *EMBO J* 2022;41:e109221.

4. D'Ambrosio A, Bressan D, Ferracci E, Carbone F, Mulè P, Rossi F, Barbieri C, Sorrenti E, Fiaccadori G, Detone T, Vezzoli E, Bianchi S, Sartori C, Corso S, Fukuda A, Bertalot G, Falqui A, Barbareschi M, Romanelli A, Pasini D, Chiacchiera F., *Increased genomic instability and reshaping of tissue microenvironment underlie oncogenic properties of Arid1a mutations. Science Advances* 15;10(11):eadh4435
5. Mulè P., Fernandez-Perez D., Brandini S., Diaferia G., Cuomo A., Manganaro D., Zanolini M., Rustichelli S., Ferrari K.J., Bisso A., Natoli G., Amati B., Chiacchiera F., Pasini D., *WNT oncogenic transcription requires MYC-dependent suppression of autophagy in gastric tumours. Gastroenterology* 167(5):903-918.

## **Descrizione sintetica dell'attività e dei risultati attesi**

### **Introduzione**

La rigenerazione epatica è un processo complesso e ancora in gran parte non caratterizzato. In condizioni di stress cronico, un insieme definito di cellule conosciute come cellule progenitriche epatiche contribuisce alla rigenerazione sia degli epatociti che dei colangiociti. Questo processo viene attivato e progredisce attraverso l'interazione di diversi tipi cellulari, tra cui cellule immunitarie, cellule stellate, epatociti e colangiociti. Comprendere come queste cellule interagiscono e quali programmi di differenziazione vengono attivati è di importanza cruciale, considerando che una rigenerazione epatica efficiente è fondamentale per preservare l'integrità del fegato, prevenendo la formazione di tumori epatici. Interazioni cellulari simili e traiettorie di differenziazione sono state suggerite come potenzialmente coinvolte nella progressione del tumore epatico. Ogni anno in Italia vengono diagnosticati 13.000 nuovi casi di carcinoma epatocellulare. I polmoni sono il sito di metastatizzazione più frequentemente osservato. Per la maggior parte dei tumori, la metastatizzazione è un evento precoce, indicando che l'accumulo graduale di mutazioni durante la crescita del tumore non è sufficiente a spiegare la disseminazione delle cellule e la formazione di metastasi. Questi dati suggeriscono che il potenziale metastatico è una caratteristica intrinseca del tumore primario e il risultato di una risposta adattativa a specifiche pressioni ambientali locali o sistemiche. La possibilità di identificare potenziali cloni metastatici già presenti nei tumori primari e comprendere come interagiscono con tutte le altre cellule nel microambiente tumorale potrebbe rappresentare una svolta per la prognosi e la terapia.

### **Obiettivo**

L'obiettivo di questo progetto è eseguire analisi bioinformatiche per caratterizzare l'interazione cellula-cellula e i nuovi intermedi di differenziamento durante la rigenerazione epatica e la progressione del tumore. Analizzando i dati di trascrittoma a singola cellula e trascrittoma spaziale pubblicati e prodotti internamente, miriamo a caratterizzare l'interazione tra cellule immunitarie ed epatociti durante la rigenerazione epatica, così come tra cellule immunitarie e cellule tumorali sia nelle lesioni primarie che metastatiche. Vogliamo comprendere come gli epatociti e le cellule immunitarie rimodellano reciprocamente il paesaggio trascrizionale ed epigenetico per promuovere l'espansione delle cellule progenitriche epatiche, la disseminazione cellulare o la riattivazione di lesioni dormienti.

### **Approccio sperimentale**

Sfrutteremo diversi modelli e set di dati già pubblicati. Per la rigenerazione epatica abbiamo già disponibili diversi modelli, ampiamente usati in letteratura, in grado di rispondere al danno in maniera coordinata e precisa. Per quanto riguarda la progressione tumorale abbiamo sviluppato un nuovo modello murino di carcinoma epatocellulare in cui la crescita di tumori primari sporadici è seguita dalla formazione di metastasi polmonari (macrometastasi proliferative e micrometastasi dormienti). Questo processo è simile a quanto osservato negli esseri umani. Lo studente eseguirà analisi bioinformatiche dei dati di trascrittoma spaziale e proteomica per caratterizzare i tumori primari e le lesioni metastatiche nel modello murino. Inoltre, la trascrittoma a singola cellula generata in laboratorio sarà analizzata per caratterizzare ulteriormente le cellule immunitarie associate a queste lesioni. La proteomica spaziale sarà eseguita utilizzando la microdissezione laser accoppiata all'analisi di spettrometria di massa per isolare e caratterizzare regioni specifiche di lesioni primarie e metastatiche. Per la trascrittoma spaziale, pianifichiamo di utilizzare piattaforme disponibili in commercio, mentre per la trascrittoma a singola cellula faremo affidamento sulla nuova struttura a singola cellula presso il CIBIO. I dati di trascrittoma spaziale e a singola cellula saranno ulteriormente analizzati per studiare la comunicazione cellula-cellula e le traiettorie cellulari nelle lesioni pre-tumorali, primarie e metastatiche. Infatti, i dati ottenuti saranno integrati con l'analisi delle alterazioni genetiche ed epigenetiche nelle diverse cellule e con l'indagine del ruolo dell'infiltrato immunitario utilizzando molecole bioattive specifiche (ad esempio, anti-CTLA4/Anti-PD1, liposomi caricati con clodronato) e ablazione selettiva basata su anticorpi. La traducibilità dei dati ottenuti sarà esplorata in collaborazione con l'unità di patologia dell'Ospedale S. Chiara.

### **Risultati attesi**

Le analisi computazionali sui modelli murini faranno luce sul ruolo del fegato e del microambiente tumorale, in particolare delle cellule immunitarie, nel plasmare l'epigenoma e la trascrizione delle cellule tumorali (e viceversa) durante la rigenerazione epatica e la progressione del tumore. L'integrazione di dati trascrizionali spazialmente risolti, trascrittoma a singola cellula e dati genomici aiuterà a identificare tratti specifici e attori chiave coinvolti in questi processi, con potenziali applicazioni cliniche.

### **References:**

1. Villanueva A. *Hepatocellular Carcinoma. N Engl J Med* 2019;380:1450-1462.
2. Vendramin R, Litchfield K, Swanton C. *Cancer evolution: Darwin and beyond. EMBO J* 2021;40:e108389.
3. Vegliante R, Pastushenko I, Blanpain C. *Deciphering functional tumor states at single-cell resolution. EMBO J* 2022;41:e109221.
4. D'Ambrosio A, Bressan D, Ferracci E, Carbone F, Mulè P, Rossi F, Barbieri C, Sorrenti E, Fiaccadori G, Detone T, Vezzoli E, Bianchi S, Sartori C, Corso S, Fukuda A, Bertalot G, Falqui A, Barbareschi M, Romanelli A, Pasini D, Chiacchiera F., *Increased genomic instability and reshaping of tissue microenvironment underlie oncogenic properties of Arid1a mutations. Science Advances* 15;10(11):eadh4435



5. Mulè P., Fernandez-Perez D., Brandini S., Diaferia G., Cuomo A., Manganaro D., Zanotti M., Rustichelli S., Ferrari K.J., Bisso A., Natoli G., Amati B., Chiacchiera F., Pasini D., WNT oncogenic transcription requires MYC-dependent suppression of autophagy in gastric tumours. *Gastroenterology* 167(5):903-918.

#### Candidate's profile (skills and competencies)

We are looking for passionate and curious open-minded candidates able to take risks and willing to fail. A mental attitude tuned towards problem solving and to collaborative work is required. He/she should hold a master's degree in informatics, bioinformatic, quantitative and computational biology or related fields. Experience in topics such as genomic and epigenomics, is preferred but is not mandatory.

#### Profilo del/la candidato/a

Cerchiamo candidati appassionati e curiosi, con una mentalità aperta, capaci di assumersi rischi e aperti al fallimento. È richiesta una mentalità orientata alla risoluzione dei problemi e al lavoro collaborativo. Il/la candidato/a dovrebbe possedere una laurea magistrale in informatica, bioinformatica, biologia quantitativa e computazionale o in campi correlati. È preferibile, ma non obbligatorio, avere esperienza in argomenti come la genomica e l'epigenomica.

## Project 2

### Noise as signal: coupling gene expression variance, metabolic fluxes and biological entropy in human disease

**Il rumore come segnale: esplorare la connessione tra varianza dell'espressione genica, flussi metabolici ed entropia biologica nelle malattie umane**

#### Laboratory:

Laboratory of Neural Development and Regeneration (<https://www.cibio.unitn.it/222/laboratory-of-neural-development-and-regeneration>)

Laboratory for Artificial Biology (<https://www.cibio.unitn.it/118/laboratory-for-artificial-biology>)

**Principal Investigator:** Martin M. Hanczyc ([martin.hanczyc@unitn.it](mailto:martin.hanczyc@unitn.it)) & Toma Tebaldi ([toma.tebaldi@unitn.it](mailto:toma.tebaldi@unitn.it))

#### Keywords:

#### Synthetic description of the activity and expected research outcome

Intrinsic gene expression noise is a pervasive feature of biological systems, significantly contributing to phenotypic variability and adaptive responses. Aberrant changes in expression variance are likely implicated in disease progression, ultimately resulting in the emergence of pathological phenotypes.

This project seeks to characterize the dynamic alterations in gene expression noise during disease development, leveraging the high resolution of single-cell sequencing data. We will utilize entropy-based methodologies to analyze transcriptional variability, aiming to determine the role of noise as a primary determinant of cellular heterogeneity in health and disease. A comprehensive understanding of the mechanisms governing stochastic fluctuations may provide critical insights into cellular variation and misregulation, potentially identifying targets for therapeutic modulation or reversal.

This research will encompass:

- (1) the analysis of single-cell RNA sequencing datasets;
- (2) the quantification of expression variability using entropy, mutual information, and related computational approaches; and
- (3) the application of flux balance analysis to demonstrate the functional consequences of gene expression variability on cellular metabolism.

#### Descrizione sintetica dell'attività e dei risultati attesi

Il rumore intrinseco dell'espressione genica è una caratteristica pervasiva dei sistemi biologici, che contribuisce significativamente alla variabilità fenotipica e alle risposte adattative. Alterazioni aberranti nella varianza dell'espressione genica sono probabilmente implicate nella progressione delle malattie, portando all'emergere di fenotipi patologici.

Questo progetto mira a caratterizzare le alterazioni dinamiche del rumore dell'espressione genica, sfruttando l'alta risoluzione dei dati di sequenziamento a singola cellula. Utilizzeremo metodologie basate sull'entropia per analizzare la variabilità trascrizionale, con l'obiettivo di determinare il ruolo del rumore come determinante primario dell'eterogeneità cellulare in salute e in malattia. Una comprensione approfondita dei meccanismi che regolano le fluttuazioni stocastiche potrebbe fornire informazioni critiche sulla variazione e sulla disregolazione cellulare, identificando potenzialmente bersagli per la modulazione terapeutica.

Questa ricerca comprenderà:



- (1) l'analisi di dataset di sequenziamento RNA a singola cellula;
- (2) la quantificazione della variabilità dell'espressione utilizzando entropia, informazione reciproca e approcci computazionali correlati; e
- (3) l'applicazione dell'analisi dei flussi metabolici per dimostrare le conseguenze funzionali della variabilità dell'espressione genica sul metabolismo cellulare."

#### Candidate's profile (skills and competencies)

The ideal candidate has prior experience in computational biology and demonstrates a proactive, interdisciplinary curiosity.

#### Profilo del/la candidato/a

Il candidato ideale ha precedente esperienza in biologia computazionale e dimostra una curiosità proattiva e multidisciplinare.

### Project 3

#### Building a Map of Human Specific Genes with Single-cell resolution (HuMap)

#### Costruire una mappa di geni "human specific" a risoluzione di singola cellula (HuMap)

##### Laboratory:

Laboratory of RNA and Disease Data Science (<https://www.cibio.unitn.it/1349/laboratory-of-rna-and-disease-data-science>)

Armenise-Harvard Laboratory of Brain Disorders and Cancer (<https://www.cibio.unitn.it/495/armenise-harvard-laboratory-of-brain-disorders-and-cancer>)

**Principal Investigator:** Toma Tebaldi ([toma.tebaldi@unitn.it](mailto:toma.tebaldi@unitn.it)) & Luca Tiberi ([luca.tiberi@unitn.it](mailto:luca.tiberi@unitn.it))

**Keywords:** Single-cell and spatial transcriptomics, cancer

#### Synthetic description of the activity and expected research outcome

Humans show unique features with respect to the closest living evolutionary species (chimpanzees and bonobos), such as brain size, height, gut, skeleton, immune system and metabolism. Molecular variations underlying these unique features have been evolving over the last 6 million years. Several hundreds of human specific genes, presenting human-specific genomic variations, have been reported in the last few years. Despite their huge relevance for understanding human-specific adaptation, most of human specific genes are currently poorly characterized in terms of their function, cell-type specific expression, regulatory relationships with other genes, and involvement in human diseases.

The HuMap project aims to characterize human specific genes in terms of their genomic features, cell-specific expression, causal relationship with other genes and association to human diseases.

The project will leverage innovative computational biology approaches such as causal network inference methods and integration of high quality up-to-date genomic, transcriptomic and proteomics data, with particular focus on high-resolution single-cell and spatially resolved omics atlases.

The information obtained by the HuMap project will be essential to identify disease-relevant candidates among human specific genes, ideally with a defined tissue or cell-specific expression and with strong causal relationships with disease-critical genes, in order to generate experimental follow-up using organoids and mouse models, where the function of the relevant human specific genes will be explored in-vivo. HuMap data will also be released as the first web server storing information on human specific genes and sharing all the results of the analyses for dissemination.

This proposal combines the complementary expertise of the Tiberi and Tebaldi labs, offers an original platform for studying human biology and has the strategic potential to open up multiple research lines answering the fundamental question of what makes humans unique at a molecular level, as well as addressing diseases occurring mainly in humans, such as brain, blood and bone tumors, or immune diseases.

#### **AIM 1: To obtain a high quality and comprehensive catalogue of human-specific genes and their single-cell expression patterns**

**Task 1a) Build the full set of human specific genes, including noncoding RNAs.**

The starting point of the analysis will be a collection of 856 human specific genes previously published (Bitar et al, 2019). A substantial update of this collection will be performed based on:

- a) recently released human genome assemblies, in particular the T2T genome assembly (Nurk et al, 2022) and the human PanGenome draft assembly (Liao et al, 2023)
- b) updated annotation of human genes and transcripts, using the latest Gencode reference. Importantly, the annotation will cover also noncoding RNAs, in order to include noncoding transcriptome variants in the final collection
- c) recent genome assemblies from the closest living evolutionary species (chimpanzees and bonobos)

The biological impact of human specific genes will be assessed by functional network and pathway analyses.

**Task 1b) Define the single-cell expression patterns of human specific genes**

Human single cell atlases will be used to explore the expression of human specific genes and identify:

- Human cell types showing the highest expression of human specific genes
  - Human specific genes with tissue or cell-type specific expression. Groups of co-expressed human specific genes will also be identified.
- The analysis will start from pan-tissue human atlases, such as the Tabula Sapiens dataset (Karkadinas et al, 2022) and the Human Cell Atlas (Regev et al, 2017) and will then focus on organ specific atlases, such as brain and immune atlases (Nieto et al, 2021; Siletti et al, 2022). The choice of specific datasets will be based on the results obtained in the analysis of pan-tissue atlases. Our preliminary results suggest that several human specific genes are selectively expressed in neuron and immune cell subtypes.

**Task 1c) Characterize human specific cells by integrating multi-species single-cell atlases**

Human single-cell atlases will be integrated with single-cell atlases from mouse (Tabula Muris Consortium, 2018; Han et al, 2018) and primates (Zhang et al, 2022; Han et al, 2022). After the integration, cells will be clustered according to transcriptional profiles, and specific clusters composed prevalently or exclusively by human cells will be identified. These human specific cells will be characterized, and the contribution of human specific genes in defining human specific cell types and states will be determined.

**AIM 2: Identify human specific genes potentially critical for human diseases**

**Task 2a Define the involvement of human specific genes in human diseases using spatially resolved expression data**

Based on cell type expression patterns collected in Task 1b, we will investigate the spatial expression patterns of the top relevant human specific genes. Given the human-specific morphology of brain regions such as the neocortex, we can anticipate the strategic relevance of collecting spatially resolved datasets available in literature and focusing on brain-related diseases (recently reviewed in Piwecka et al, 2023). Spatially resolved datasets will be analysed in order to identify human-specific genes with spatial expression pathways, possibly connected to alterations in the morphology induced by the pathology.

The most relevant candidates will be selected to generate experimental follow-up using organoids and mouse models, where the function of the relevant human specific genes will be explored in-vivo.

**References:**

- Amberger JS et al. OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.* 2019 Jan 8;47(D1):D1038-D1043. doi: 10.1093/nar/gky1151.
- Bitar M, et al. Genes with human-specific features are primarily involved with brain, immune and metabolic evolution. *BMC Bioinformatics.* 2019 Nov 22;20(Suppl 9):406. doi: 10.1186/s12859-019-2886-2.
- Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature.* 2005 Sep 1;437(7055):69-87. doi: 10.1038/nature04072.
- Dutrillaux B. Chromosomal evolution in primates: tentative phylogeny from *Microcebus murinus* (Prosimian) to man. *Hum Genet.* 1979 May 10;48(3):251-314. doi: 10.1007/BF00272830.
- Frankish A et al. GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res.* 2023 Jan 6;51(D1):D942-D949. doi: 10.1093/nar/gkac1071.
- Han L et al. Cell transcriptomic atlas of the non-human primate *Macaca fascicularis*. *Nature.* 2022 Apr;604(7907):723-731. doi: 10.1038/s41586-022-04587-3.
- Han X et al. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell.* 2018 Feb 22;172(5):1091-1107.e17. doi: 10.1016/j.cell.2018.02.001.
- Kundu I et al. GeDiPNet: Online resource of curated gene-disease associations for polypharmacological targets discovery. *Genes Dis.* 2022 Jun 13;10(3):647-649. doi: 10.1016/j.gendis.2022.05.034.
- Liao WW, et al. A draft human pangenome reference. *Nature.* 2023 May;617(7960):312-324. doi: 10.1038/s41586-023-05896-x.
- Nieto P et al. A single-cell tumor immune atlas for precision oncology. *Genome Res.* 2021 Oct;31(10):1913-1926. doi: 10.1101/gr.273300.120.
- Nurk S, et al. The complete sequence of a human genome. *Science.* 2022 Apr;376(6588):44-53. doi: 10.1126/science.abj6987.
- Piwecka M, Rajewsky N, Rybak-Wolf A. Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease. *Nat Rev Neurol.* 2023 Jun;19(6):346-362. doi: 10.1038/s41582-023-00809-y.
- Preuss TM. The human brain: rewired and running hot. *Ann N Y Acad Sci.* 2011 May;1225 Suppl 1(Suppl 1):E182-91. doi: 10.1111/j.1749-6632.2011.06001.x.
- Regev A et al. The Human Cell Atlas. *Elife.* 2017 Dec 5;6:e27041. doi: 10.7554/eLife.27041.
- Siletti K et al. Transcriptomic diversity of cell types across the adult human brain. *bioRxiv* 2022.10.12.511898; doi: <https://doi.org/10.1101/2022.10.12.511898>
- Tabula Sapiens Consortium et al. The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science.* 2022 May 13;376(6594):eabl4896. doi: 10.1126/science.abl4896.
- Tabula Muris Consortium et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature.* 2018 Oct;562(7727):367-372. doi: 10.1038/s41586-018-0590-4.
- Zhang X et al. Towards a primate single-cell atlas. *Zool Res.* 2022 Jul 18;43(4):691-694. doi: 10.24272/j.issn.2095-8137.2022.212.

**Descrizione sintetica dell'attività e dei risultati attesi**

Gli esseri umani mostrano caratteristiche uniche rispetto alle specie viventi più vicine dal punto di vista evolutivo (scimpanzé e bonobo), come le dimensioni del cervello, l'altezza, l'intestino, lo scheletro, il sistema immunitario e il metabolismo. Le variazioni molecolari alla base di queste caratteristiche uniche si sono evolute negli ultimi 6 milioni di anni. Negli ultimi anni sono state identificate diverse centinaia di geni specifici dell'uomo, "human specific", che presentano variazioni genomiche specifiche nel genoma umano. Nonostante la loro enorme rilevanza per la comprensione dell'adattamento specifico dell'uomo, la maggior parte dei geni specifici umani sono attualmente scarsamente caratterizzati in termini di funzione, espressione specifica del tipo cellulare, relazioni regolatorie con altri geni e coinvolgimento nelle malattie umane.

Il progetto HuMap mira a caratterizzare i geni specifici umani in termini di caratteristiche genomiche, espressione cellula-specifica, relazione causale con altri geni e associazione con malattie umane.

Il progetto sfrutterà approcci innovativi di biologia computazionale come metodi di inferenza di rete causale e integrazione di dati genomici, trascrittomici e proteomici aggiornati e di alta qualità, con particolare attenzione agli atlanti omici a singola cellula ad alta risoluzione e risolti spazialmente.

Le informazioni ottenute dal progetto HuMap saranno essenziali per identificare candidati rilevanti per malattie tra i geni specifici umani, idealmente con un'espressione specifica per tessuto o tipo cellulare e con forti relazioni causali con geni critici per malattie, al fine di generare un follow-up sperimentale utilizzando organoidi e modelli murini, in cui la funzione dei geni "human specific" rilevanti sarà esplorata in vivo. I dati HuMap verranno inoltre rilasciati come primo web server che contiene informazioni su geni specifici umani, per la condivisione di tutti i risultati delle analisi.

Questa proposta combina le competenze complementari dei laboratori Tiberi e Tebaldi, offre una piattaforma originale per lo studio della biologia umana e ha il potenziale strategico per aprire molteplici linee di ricerca che rispondono alla domanda fondamentale: cosa rende gli esseri umani unici a livello molecolare? I risultati saranno potenzialmente rilevanti per comprendere malattie che si verificano principalmente negli esseri umani, come tumori al cervello, al sangue e alle ossa o malattie immunitarie.

### **OBIETTIVO 1: Ottenere un catalogo completo e di alta qualità di geni specifici dell'uomo e dei loro profili di espressione a livello di singola cellula**

Obiettivo 1a) Costruire l'insieme completo di geni specifici umani, inclusi gli RNA non codificanti.

Il punto di partenza dell'analisi sarà una raccolta di 856 geni specifici umani precedentemente pubblicati (Bitar et al, 2019). Un aggiornamento sostanziale di questa raccolta verrà effettuato sulla base di:

a) assemblaggi di genoma umano rilasciati di recente, in particolare l'assemblaggio del genoma T2T (Nurk et al, 2022) e il progetto di assemblaggio PanGenome umano (Liao et al, 2023)

b) annotazione aggiornata dei geni e dei trascritti umani, utilizzando l'ultimo riferimento Gencode. È importante sottolineare che l'annotazione coprirà anche gli RNA non codificanti, al fine di includere varianti del trascrittoma non codificante nella raccolta finale

c) recenti assemblaggi di genomi delle specie viventi più vicine all'uomo evolutivamente (scimpanzé e bonobo)

L'impatto biologico dei geni specifici umani sarà valutato mediante analisi funzionali e reti di co-espressione.

Obiettivo 1b) Definire i pattern di espressione "single-cell" dei geni "human specific"

Gli atlanti "single-cell" umani verranno utilizzati per esplorare l'espressione di geni "human specific" e identificare:

a) Tipi cellulari umani che mostrano la massima espressione di geni specifici umani

b) Geni specifici umani con espressione specifica per tessuto o tipo cellulare. Verranno inoltre identificati gruppi di geni specifici umani co-espressi.

L'analisi inizierà da atlanti umani pan-tissutali, come il dataset Tabula Sapiens (Karkadinas et al, 2022) e lo "Human Cell Atlas" (Regev et al, 2017) e si concentrerà quindi su atlanti specifici per organi, come cervello e sistema immunitario (Nieto et al, 2021; Siletti et al, 2022).

La scelta dei dataset specifici sarà basata sui risultati ottenuti nell'analisi degli atlanti pan-tissutali. I nostri risultati preliminari suggeriscono che diversi geni umani specifici sono espressi selettivamente nei sottotipi di neuroni e cellule immunitarie.

Obiettivo 1c) Caratterizzare cellule umane specifiche integrando atlanti "single-cell" multi-specie

Gli atlanti "single-cell" umani saranno integrati con atlanti "single-cell" di topo (Tabula Muris Consortium, 2018; Han et al, 2018) e di primati (Zhang et al, 2022; Han et al, 2022). Dopo l'integrazione, le cellule verranno raggruppate secondo profili trascrizionali e verranno identificati cluster specifici composti prevalentemente o esclusivamente da cellule umane. Queste cellule umane specifiche verranno caratterizzate e verrà determinato il contributo dei geni specifici umani nella definizione dei tipi e degli stati cellulari specifici dell'uomo.

### **OBIETTIVO 2: Identificare geni "human specific" potenzialmente critici per le malattie umane**

Attività 2a Definire il coinvolgimento di geni "human specific" nelle malattie umane utilizzando dati di espressione a risoluzione spaziale.

Sulla base dei modelli di espressione dei tipi cellulari raccolti nell'Attività 1b, indagheremo i modelli di espressione spaziale dei principali geni specifici umani rilevanti. Data la morfologia specifica dell'uomo delle regioni del cervello come la neocorteccia, possiamo anticipare la rilevanza strategica della raccolta di set di dati risolti spazialmente disponibili in letteratura e concentrandoci sulle malattie legate al cervello (recentemente esaminate in Piwecka et al, 2023). Verranno analizzati set di dati risolti spazialmente al fine di identificare geni specifici dell'uomo con percorsi di espressione spaziale, possibilmente collegati ad alterazioni nella morfologia indotte dalla patologia.

I candidati più rilevanti verranno selezionati per generare un follow-up sperimentale utilizzando organoidi e modelli murini, in cui la funzione dei geni specifici umani rilevanti verrà esplorata in vivo.

### **Candidate's profile (skills and competencies)**

The ideal candidate has strong interest in exploring the biology and potential pathogenicity of human specific genes, and possibly has previous experience in computational biology, in particular single-cell and spatial omics data.

### **Profilo del/la candidato/a**

Il candidato ideale ha un forte interesse nell'esplorazione della biologia e della potenziale patogenicità di geni "human specific" e possibilmente ha precedenti esperienze in biologia computazionale, in particolare dati omici a risoluzione spaziale e di singola cellula.



## Project 4

### Exploring Cell-Type-Specific Metabolic Signatures in Breast Cancer through Transcriptome Deconvolution

#### Studio delle firme metaboliche a livello cellulare nel tumore al seno tramite deconvoluzione del trascrittoma

##### Laboratory:

Laboratory of Computational Modeling (<https://www.cibio.unitn.it/1321/laboratory-of-computational-modeling>)

**Principal Investigator:** Luca Marchetti ([luca.marchetti@unitn.it](mailto:luca.marchetti@unitn.it)) & Pier Paolo Di Fiore and Sara Sigismund (IEO)

##### Keywords:

##### Synthetic description of the activity and expected research outcome

Breast cancer (BC) is the most commonly diagnosed cancer worldwide and a leading cause of cancer-related mortality. Among its numerous hallmarks, the ability of cancer cells to reprogram their metabolism plays a pivotal role in tumor progression and therapeutic resistance. A prime example of this metabolic plasticity is aerobic glycolysis, which can be monitored in vivo through 18F-fluorodeoxyglucose positron emission tomography (FDG-PET).

Recent work by Confalonieri et al. (PMID: 38757578) analyzed RNA-Seq data from 120 treatment-naïve BC patients who underwent FDG-PET imaging prior to surgery, leading to the identification of a 54-gene signature, PETsign, capable of distinguishing tumors by metabolic phenotype. However, this analysis was based on bulk transcriptomic data, limiting insights into contributions from specific cell types within the tumor microenvironment.

This PhD project aims to address this gap by applying transcriptome deconvolution techniques to resolve cell-type-specific gene expression profiles in the same dataset, with a focus on metabolic signatures arising from immune and stromal compartments.

The work is part of a collaboration between the Laboratory of Computational Modelling at CIBIO and the research groups of Prof. Pier Paolo Di Fiore and Prof. Sara Sigismund at the European Institute of Oncology (IEO), Milan.

The candidate will develop expertise in cancer metabolism, computational transcriptomics, and systems biology. The project will progress from foundational training to the identification and functional analysis of metabolic signatures, concluding with their validation in public datasets and in a large retrospective cohort of 2,453 BC patients operated at the European Institute of Oncology between 1997 and 2000 (IEO BC 97-00 cohort).

##### Descrizione sintetica dell'attività e dei risultati attesi

Il carcinoma mammario è la neoplasia più frequentemente diagnosticata a livello globale e una delle principali cause di mortalità per cancro. Tra le sue caratteristiche distintive, la capacità delle cellule tumorali di riprogrammare il proprio metabolismo riveste un ruolo cruciale nella progressione della malattia e nella resistenza alle terapie. Un esempio emblematico di questa plasticità metabolica è la glicolisi aerobica, monitorabile in vivo attraverso la tomografia a emissione di positroni con 18F-fluorodesossiglucosio (FDG-PET).

Uno studio recente di Confalonieri et al. (PMID: 38757578) ha analizzato dati RNA-Seq provenienti da 120 pazienti con carcinoma mammario non trattati, sottoposti a FDG-PET prima dell'intervento chirurgico. Questo lavoro ha portato all'identificazione di una firma di 54 geni, denominata PETsign, in grado di distinguere i tumori in base al loro fenotipo metabolico. Tuttavia, l'analisi si è basata su dati trascrittomici bulk, limitando la possibilità di attribuire le variazioni di espressione genica a specifici tipi cellulari all'interno del microambiente tumorale.

Questo progetto di dottorato si propone di colmare questa lacuna applicando tecniche di deconvoluzione trascrittomica allo stesso dataset, con l'obiettivo di ricostruire profili di espressione genica specifici per tipo cellulare, focalizzandosi in particolare sulle firme metaboliche delle componenti immunitarie e stromali.

Il lavoro si svolgerà nell'ambito di una collaborazione tra il Laboratorio di Modellistica Computazionale del CIBIO e i gruppi di ricerca dei Professori Pier Paolo Di Fiore e Sara Sigismund presso l'Istituto Europeo di Oncologia (IEO) di Milano.

Durante il percorso di dottorato, il/la candidato/a acquisirà competenze avanzate in metabolismo tumorale, trascrittomica computazionale e biologia dei sistemi. Il progetto si articolerà in diverse fasi: dalla formazione iniziale allo sviluppo e caratterizzazione funzionale delle firme metaboliche, fino alla loro validazione su dataset pubblici e su una ampia coorte retrospettiva di 2.453 pazienti.

##### Candidate's profile (skills and competencies)

The ideal candidate will be a highly motivated student with a MSc in Quantitative Computational Biology, Bioinformatics, Biostatistics, or related disciplines, ideally with experience in transcriptomic data analysis and statistical or machine learning methods. Familiarity with R or Python programming and an interest in cancer biology and systems medicine will be considered a strong asset. The candidate should be enthusiastic about working at the interface of computational biology and clinical oncology and be eager to contribute to a multidisciplinary and collaborative research environment. The project will involve close interactions with teams at the University of Trento and the European Institute of Oncology in Milan and will require both independent problem-solving skills and a strong team-oriented mindset.



**Profilo del/la candidato/a**

Il/la candidato/a ideale è una persona fortemente motivata, con una laurea magistrale in Biologia Computazionale Quantitativa, Bioinformatica, Biostatistica o discipline affini, preferibilmente con esperienza nell'analisi di dati trascrittomici e nell'utilizzo di metodi statistici o di machine learning. La familiarità con linguaggi di programmazione come R o Python, unita a un interesse per la biologia del cancro e la medicina di sistema, costituiranno un valore aggiunto. Si richiede una forte attitudine al lavoro interdisciplinare e collaborativo, nonché la capacità di affrontare problemi complessi in modo autonomo. Il progetto prevede interazioni costanti con gruppi di ricerca dell'Università di Trento e dell'Istituto Europeo di Oncologia di Milano.