



Postdoctoral position (F/M) in computational biology/statistics/machine learning – Gustave Roussy, Paris-Saclay University, Villejuif

Gustave Roussy, the European leader in cancer research, is seeking an ambitious and motivated individual to contribute to exciting cutting-edge of machine learning/artificial intelligence at the intersection of imaging and computational biology. This project is dedicated to the development of a new methodology for the analysis of spatial transcriptomic data with the objective of improving our understanding of the tumor with its microenvironment. The applicant will benefit of the structure and the support of the experts of different disciplines of Gustave Roussy. Located close to Paris (Villejuif), Gustave Roussy regroups 1,000 people having research activities in fundamental, clinical and translational research, and places the innovation at the heart of fight against cancer, to face the current and future challenge of the development of the personalized medicine. He/she will also be affiliated to the INSERM Oncostat team (U1018, CESP), dedicated to the development of new methodologies for the precision medicine.

**Context of the project:**

The success of the targeted therapies for a precision medicine in oncology during the last decade relies on the identification of targetable molecular markers identified from genomic or transcriptomic sequencing of tumor samples. However, a third of children's tumor do not present any known targetable marker, driving clinicians and researchers to investigate tumor microenvironment (TME) and TME modulating agents, such as immunotherapies in children. Identifying the spatial pattern of the interactions between the different actors of the TME (such as the immune cells, osteoclasts, vascularization...) could be a key to better understand the TME dynamic, with the promise of the development of personalized intervention to regulate these populations. For this purpose, we have been generated the histopathological images and the related spatial expression using the in situ RNA Visium capture of 10x Genomics, resulting in a database of hundreds of thousands spots measured from dozens of tumor samples. The aim of this work is to develop a deep model for the spatial integration of the histological image and the spatial omics data, to define a high-resolution spatial characterization of the interactions between the cell populations of the TME. This project could be extended to the integration of other data sources (immunohistochemistry, single cell sequencing, other spatial transcriptomics technologies,...).

**Work environment and opportunities:**

This project is based on a collaboration of several teams of Gustave Roussy. Under the responsibility of Dr Damien Drubay, researcher in the Office of Biostatistics and Epidemiology, your contribution will be placed at the center of a stimulating teamwork between experts from several disciplines (bioinformaticians, anapathologists, biologists and clinicians). Your mission will be to build the mathematical consensus of these specialties by developing a new deep learning algorithm for TME characterization from histopathological image and spatial gene expression. In a second phase, the aim will be to develop a visualization tool for the interpretability of the algorithm and the interpretation of the results.

You will benefit from access to Gustave Roussy and CESP/INSERM IT infrastructures, including high performance computing clusters with several GPU nodes. This position offers the opportunity to publish in high quality scientific journals, as well as to build collaborative relationships with renowned researchers at Gustave Roussy and its partners.

**Expected profile:**

- PhD in Statistics, Computational Biology, Machine Learning or a closely aligned discipline is required before the time of appointment
- Strong mathematical background in computational science, and willingness to address the challenge of developing deep models based from the latest approaches such as flow-based models, geometric deep learning (especially graph modeling),... Bayesian theory knowledge would be a plus

- Good programming skills, including Python (or R) and at least one of the most popular machine learning platforms (Torch, TensorFlow, or JAX)
- In addition to your statistical/mathematical skills, you are known for your rigor, autonomy and interpersonal skills. You enjoy working in a multidisciplinary environment and are able to communicate complex results to a non-specialist audience.

**Type of contract:**

18 months full-time position

Starting date: September, 2023 or as soon as possible thereafter.

**How to apply?**

As part of our proactive policy to promote the integration of people with disabilities, all applications received are considered on an equal basis.

To submit an application for this vacancy, please send your *curriculum vitae* and a cover letter by email to [damien.drubay@gustaveroussy.fr](mailto:damien.drubay@gustaveroussy.fr), indicating the reference "RecrutCompBio2023" in the object of the email.

Post-doctorat (F/H) en biologie computationnelle/statistique/machine learning – Gustave Roussy, Université Paris-Saclay, Villejuif

Gustave Roussy, leader européen de la recherche contre le cancer, est à la recherche d'une personne ambitieuse et motivée pour contribuer au développement de méthodes innovantes dans le domaine de l'apprentissage automatique/intelligence artificielle à l'intersection de l'imagerie et de la biologie computationnelle. Ce projet s'inscrit dans le cadre de l'analyse de données de transcriptomique spatiale avec pour objectif d'améliorer notre compréhension de la tumeur avec son microenvironnement. Le candidat bénéficiera du soutien des experts de différentes disciplines de Gustave Roussy. Situé près de Paris (Villejuif), Gustave Roussy regroupe 1000 personnes ayant des activités de recherche fondamentale, clinique et translationnelle, et place l'innovation au cœur de la lutte contre le cancer, pour relever le défi actuel et futur du développement de la médecine personnalisée. Le candidat sera également affilié à l'équipe INSERM Oncostat (U1018, CESP), dédiée au développement de nouvelles méthodologies pour la médecine de précision.

### **Contexte du projet**

Le succès des thérapies ciblées pour une médecine de précision en oncologie au cours de la dernière décennie repose sur l'identification de marqueurs moléculaires ciblables identifiés à partir du séquençage génomique ou transcriptomique d'échantillons de tumeurs. Cependant, un tiers des tumeurs de l'enfant ne présentent aucun marqueur ciblé connu, ce qui pousse les cliniciens et les chercheurs à étudier le microenvironnement tumoral (MET) et les agents modulant le MET, tels que les immunothérapies chez l'enfant. Identifier le schéma spatial des interactions entre les différents acteurs du MET (tels que les cellules immunitaires, les ostéoclastes, la vascularisation...) pourrait être une clé pour mieux comprendre la dynamique du MET, avec la promesse du développement d'une intervention personnalisée pour réguler ces populations. Dans ce but, nous générons quotidiennement les images histopathologiques et l'expression spatiale associée en utilisant la capture in situ de l'ARN Visium de 10x Genomics, résultant en une base de données de centaines de milliers de spots mesurés à partir de dizaines d'échantillons de tumeurs. L'objectif de ce travail est de développer un modèle profond pour l'intégration spatiale de l'image histologique et des données omiques spatiales, afin de définir une caractérisation spatiale à haute résolution des interactions entre les populations cellulaires de la MET. Le cadre de ce projet pourrait être élargi à l'intégration d'autres types de données (immunohistochimie, séquençage single-cell, autre technologie de transcriptomique spatiale,...).

### **Missions :**

Ce projet s'articule dans une collaboration de plusieurs équipes de Gustave Roussy. Sous la responsabilité du Dr Damien Drubay, chercheur au sein du Bureau de Biostatistique et d'Epidémiologie (BBE), votre contribution se placera au centre d'un travail d'équipe stimulant entre les experts de plusieurs disciplines (bioinformaticiens, anapathologistes, biologistes et cliniciens). Votre mission consistera à construire mathématiquement le consensus entre ces spécialités, en développant un nouvel algorithme d'apprentissage profond pour la caractérisation du MET à partir de l'image histopathologique et de l'expression spatiale des gènes. Dans un second temps, il s'agira de développer un outil de visualisation pour l'interprétabilité de l'algorithme et l'interprétation des résultats.

Vous bénéficierez d'un accès aux infrastructures informatiques de Gustave Roussy et du CESP/INSERM, comprenant des clusters de calcul haute performance avec plusieurs nœuds GPU. Ce poste offre la possibilité de publier dans des revues scientifiques de haute qualité, ainsi que de construire des relations de collaboration avec des chercheurs renommés de Gustave Roussy et de ses partenaires.

### **Profile attendu :**

- Un doctorat en statistique, biologie computationnelle, apprentissage automatique ou dans une discipline proche est requis avant la prise de poste.

- Solides connaissances en science computationnelle, et volonté de relever le défi du développement de modèles profonds basés sur les dernières approches telles que les modèles de flux, l'apprentissage profond géométrique (notamment modélisation de graphes),... Une connaissance de la théorie bayésienne serait un plus
- Bonnes compétences en programmation, notamment en Python (ou R) et au moins une des plateformes d'apprentissage automatique les plus populaires (Torch, TensorFlow ou JAX)
- Au-delà de vos compétences en statistique/mathématique, vous êtes reconnu(e) pour votre rigueur, votre autonomie et vos qualités relationnelles. Vous aimez le travail dans un cadre pluridisciplinaire et savez vulgariser vos réflexions à un public non-spécialiste

### **Type de contrat**

Contrat à durée déterminée de 18 mois.

Date de prise de poste : Septembre 2023 ou dès que possible après cette date

### **Comment postuler ?**

Dans le cadre de notre politique volontariste en faveur de l'insertion des personnes en situation de handicap, toutes les candidatures reçues sont étudiées à compétences égales.

Merci d'adresser par email un CV et une lettre de motivation à l'adresse [damien.drubay@gustaveroussy.fr](mailto:damien.drubay@gustaveroussy.fr), en indiquant la référence « RecrutCompBio2023 » dans l'objet de l'email.