Offer #2024-07092

Mathematical modeling and statistics for quantifying and predicting the evolution of tumor heterogeneity in chronic lymphocytic leukemia

Contract type: Internship agreement
Level of qualifications required: Graduate degree or equivalent
Fonction: Internship Research

Context

Internship context

The internship will take place in the Probability and Statistics team of the Institut Elie Cartan de Lorraine (IECL) and in the BIGS (Biology, Genetics and Statistics) team of Inria Nancy. The trainee will be involved in discussions with staff at the Strasbourg University Hospital on medical aspects of the project. During the internship, the trainee will have the opportunity to discover the world of mathematical research through the life of a dynamic mathematics laboratory, and to attend seminars and working groups in probability and statistics. The trainee will receive the standard stipend (around €600/month).

Follow up with a PhD thesis

At the end of the internship, the intern can apply for a thesis grant, which is part of a project funded by ITMO cancer (INSERM funding).

Supervision

The internship will be supervised by Nicolas Champagnat, Coralie Fritsch and Ulysse Herbach (IECL and INRIA Nancy - Grand Est).

Contacts

nicolas.champagnat@inria.fr, coralie.fritsch@inria.fr, ulysse.herbach@inria.fr

Assignment

Biological context

The development of targeted therapies has allowed considerable progress in the treatment of many cancers, but their efficacy is dependent on intra-tumor heterogeneity. In lymphomas and leukemias, the identification of gene alterations by high-throughput sequencing allows the characterization of this heterogeneity. In these hemopathies, the initial leukemic clone (i.e. population of tumor cells with the same genome) has a unique immune repertoire corresponding to specific gene sequences encoding the antigen receptor, called VDJ genes. The occurrence of additional mutations in VDJ genes may be responsible for the emergence of subclones with increased antigen receptor reactivity further complicating the clonal heterogeneity of these hemopathies. However, this second level of clonal heterogeneity and its evolution remain poorly characterized and is not considered in the management of these cancers.

Project description

We propose to develop a mathematical model for the evolution of the two levels of clonal heterogeneity in leukemia, allowing to characterize their evolution from temporal bulk sequencing data of VDJ and cancer genes mutations using a Bayesian approach. We will test the predictive performance of clonal evolution from the inferred model.

Main activities

Tasks

In this internship, we propose to tackle the problem of clonal reconstruction from data collected at a single time. The originality comes from the fact that data are heterogeneous: we will have the full profile of VDJ mutations of clones with frequencies and each cancer genes variants with allele frequencies. From the mathematical modeling perspective, VDJ data share common features with single-cell data. Existing packages for clonal heterogeneity analysis are B-SCITE (Malikic et al., 2017) and ddClone (Salehi et al., 2017). They are able to deal with both types of data (bulk and single-cell) and could in principle be used here. However, there are specificities of CLL that do not fit into these.

First, the trainee will construct a probabilistic model accounting for all the data. This model will contain the phylogenetic tree as latent variable, where each node in the tree corresponds either to a VDJ mutation, a mutation of cancer genes, or a chromosomic alteration, where each mutation occurs only once in the tree. The observations will then be obtained, following the classical rules of the infinitely many sites model, as linear combinations of the frequency of every clone in the sample (which are other latent variables), possibly with some noise. We will assume that the correspondence between frequencies of VDJ mutations and cancer genes mutations can be described using a single (unknown) proportionality parameter.

Treating latent variables as parameters, we could use the maximum likelihood method, but maximization is a difficult problem in practice due to the very large number of possible trees. We will test genetic algorithms (Metropolis-Hastings, MCMC...), but we expect better results using a Bayesian approach, combined with a variational method to maximize the a posteriori likelihood.

Second, the trainee will validate the method from data simulated from our model, then using the benchmarks simulation tools proposed by Foglierini et al. (2020), adapting them to our double heterogeneity context. Since nearly nothing is known about the
link between VDJ mutations and cancer genes mutations, it will be interesting to discuss the results with respect to the natural biological assumption that driver mutations occur before VDJ mutations, and then mutations of resistance to therapy. Our contacts with biologists from CHU Strasbourg will be crucial here. The case of patients with several VDJ clones of leukemic lymphocytes, which occurs sometimes, will also be particularly interesting.

Third, we will try to detect if groups of patients have similar mutational patterns (such as tree topology), which could correspond to a similar tumorigenesis, or a similar stage of progression, or a similar response to treatments. This is a clustering problem that can be addressed by model-free artificial intelligence tools (such as latent Dirichlet allocation: Pritchard et al., 2000; Falush et al., 2003), or using models developed by Beerenwinkel et al. (2004, 2005). The double heterogeneity brings a specific difficulty here. For example, will we find clusters sharing similar VDJ genes or only clusters sharing similar cancer genes mutations?

**Bibliography**


**Skills**

The candidate should have skills in applied mathematics, particularly statistics and/or stochastic modeling. R, Python or Matlab programming skills are also required. An affinity or experience with medical applications will be highly appreciated. The candidate should speak fluent French or English and have good skill in scientific English.

**Benefits package**

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

**Remuneration**

4.05 per hour of training

**General Information**

- **Theme/Domain**: Modeling and Control for Life Sciences
- **Town/city**: Villers-lès-Nancy
- **Inria Center**: Centre Inria de l'Université de Lorraine
- **Starting date**: 2024-04-01
- **Duration of contract**: 5 months
- **Deadline to apply**: 2024-02-11

**Contacts**

- **Inria Team**: SIMBA
- ** Recruiter**: Champagnat Nicolas / Nicolas.Champagnat@inria.fr

**About Inria**

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. Inria research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

**The keys to success**

Date limite de candidature : 15 February 2023

How to apply

Upload your file on jobs.inria.fr in a single pdf or zip file. Your file should contain the following documents:
- Your CV
- A cover/motivation letter describing your interest in this topic
- Your degree certificates and transcripts for Bachelor and Master (or the last 5 years)

In addition, one recommendation letter from the person who supervises(d) a research project or internship should be sent directly by his/her author to the supervisors.

**Warning**: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

**Instruction to apply**

**Defence Security:**
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

**Recruitment Policy:**
As part of its diversity policy, all Inria positions are accessible to people with disabilities.