

# Funded PhD position in computational biology

Approximate Bayesian Computing for quantitative study of the evolution of bacterial pathogens

Antoine Frenoy and Olivier François, Université Grenoble Alpes

**Localization :** Université Grenoble Alpes, Laboratoire [TIMC](#), Team [MAGE](#)

**Supervision:** Antoine Frenoy (Asst. Prof. Grenoble INP) and Olivier François (Prof. Grenoble INP, HDR)

**Topic and keywords:** Computational biology; Models and simulations; Parameter inference (Markov chain Monte Carlo); Artificial intelligence for biology; Bacterial evolution

**Starting date:** October 1st 2023

**Salary:**  $\geq 2000\text{€}$  (monthly raw salary), in accordance with [national regulations](#) (already secured full funding by Grenoble INP)

## 1 Summary of the PhD project

The fluctuation test devised by Salvador Luria and Max Delbrück (publication in 1943, Nobel prize in 1969) was a landmark for our quantitative understanding of microbial evolution. It is one of the most common methods to quantify bacterial mutation rates from experimental data (reviewed by [1]). However this quantification relies on mathematical models making several restrictive implicit hypotheses on population growth dynamics. In particular, these models assume that the populations grow exponentially without death, that the studied mutations are selectively neutral, and that no sampling is performed between growth and mutants assay by selective plating.

These hypotheses are often problematic for modern research questions in microbial evolution, for example regarding within-host bacterial adaptation to antibiotic treatment. The field is thus stuck by the absence of a better model for the generation of genetic diversity in complex, realistic scenarios (bacteria within a host with a immune system and a complex microbiome, exposed to several biotic and abiotic stresses).

This project aims at developing a new simulation-based method for mutation rate inference in complex ecological scenarios using Approximate Bayesian Computing, and at using this method to study bacterial evolution under stress, and in particular adaptation to antibiotic treatments.

## 2 Objectives

1. Develop a simulation-based method to infer mutation rate from experimental data in arbitrary ecological scenario
2. Extend this method to estimate ecological parameters (such as those characterizing growth dynamics or fitness effect of the mutation) simultaneously.
3. Use these methods to study the development of antibiotic resistance *in vivo*, taking into account the intricate lifestyle of bacterial pathogens in their hosts (action of the immune system, competition with other microbes, potential antibiotic treatments).

## 3 Scientific context

The study of the process of bacterial mutagenesis is one of the oldest problems in computational biology. Fluctuation tests, devised by S. Luria and M. Delbrück in their seminal 1943 paper [2], permit to estimate the rate of a specific mutation from phenotypic data (number of mutants for a specific phenotype observed in several replicate experiments) in a simple overnight experiment. Despite the simplicity of the experiment, inferring the rate of the mutation of interest (probability of mutation per genome division) during growth from the observed number of mutants after growth requires a non-trivial model of growth with mutation.

This model makes a set of restrictive hypotheses (no fitness effect of the focal mutation, no death, no sampling), permitting an analytical formulation (proposed by Lea and Coulson in 1949 [3]) for the distribution of the number of observed mutants under a given mutation rate. This analytical formulation permits a maximum likelihood approach to infer mutation rate from experimental data [4], which remains the state of the art today.

However, deviation from this set of hypotheses makes it impossible to estimate the mutation rate based on this standard model. This is becoming an important limitation, as several current research questions lead to the consideration of experimental scenarios where these hypotheses are not met. It has for example been found that some antibiotic resistance mutations classically scored in fluctuation tests are not neutral [5]; the effect of bactericidal stresses on mutation rate has been subject to much scrutiny [6]; and dilution (sampling) before selective plating is sometimes required for experimental reasons.

Several authors thus proposed refined models which are able to circumvent one or several of these hypotheses. But no convincing analytical permits to take into account the more general case, characterized by:

**fitness effect** of the focal mutation (or set of mutations) conferring the phenotype of interest

**arbitrarily complex demographics** including death in both the wild-type and mutants populations (at a rate which may not be constant)

**sampling** of the population before plating (mostly resulting from purely experimental constraints)

To circumvent the absence of analytical model, the host team has recently developed a preliminary stochastic simulation model of the emergence of mutations in this complex ecological scenario. This particularly fast simulator will permit the development of a simulation-based inference

method of mutation rate, which would allow to bypass all mathematical hypotheses made by the classical analytical inference methods: any model that can be effectively simulated is suitable for simulation-based parameter inference.

## 4 Method

Approximate Bayesian Computing (ABC) permits to infer parameters of a model for which no expression of the likelihood is available, for example when (as in our case) the model has no analytical formulation but can be simulated [7, 8].

The first step of this work will consist in the development of an ABC method suited for the specificity of our problem. Our goal is to estimate the model parameters by confronting simulations with experimental data. In addition to mutation rate estimation, we will assess whether it is possible to simultaneously infer the ecological parameters of the model (those describing growth and death of the wild-type and the mutant) from the same dataset. We will explore the suitability of the standard candidate methods (including the rejection algorithm, pseudo-marginal Metropolis–Hastings algorithm, and more modern Markov chain Monte Carlo – MCMC – methods) and the technical choices to be made for the adaptation of these methods.

Beyond the initial development and use of this ABC method, the project encompasses several possible extensions which can be adapted to the background and interests of the candidate. This for example includes the adaptation of the simulator to a broader range of biological scenarios (such as mutation rate heterogeneity, spatial structure, phenotypes which can be conferred by several mutations), with applications to existing datasets.

More generally, beyond the initially envisaged topic, there is flexibility for developing the project in the directions which correspond to the PhD student’s interests and ambitions, and for tackling other projects related to computational and mathematical methods in ecology and evolutionary biology.

## 5 Supervision and environment

**Supervision:** The PhD student will be jointly supervised by Antoine Frenoy (assistant professor) and Olivier François (full professor, HDR), in the [MAGE](#) (Models and Algorithms for Genomics) team of the [TIMC](#) (Translational Innovation in Medicine and Complexity) institute. Antoine Frenoy has expertise in the computational study of mutagenesis and models of epidemiology of antibiotic resistance [6, 9, 10]. Olivier François is one of the pioneers in the development of Approximate Bayesian Computing and its application in population genetics and genomics [7, 11].

**Scientific environment:** TIMC gathers clinicians, biologists, mathematicians and computer scientists around interdisciplinary studies in biology and healthcare. The MAGE team has a longstanding scientific history in the field of computational biology, and was one of the early players in the use of artificial intelligence methods for biomedical research (in particular for the analysis of genetics, genomics, multi-omics and metagenomics datasets).

**Broader environment:** Knowledge of French is not a requirement: the scientific aspects of the PhD can be fully conducted in English, and the administrative staff of the lab can speak English. The social times in the lab are unfortunately not always conducted in English.

Grenoble is known for its student life and its proximity to the mountains. International students may be interested in the [practical guide for international PhD students](#) from the International Students and Scholars Office in Grenoble.

## 6 Application and desired profile of the PhD student

Two types of profiles could be relevant for this project: a student with a background in applied mathematics or computer science with a strong interest (but not necessarily previous knowledge) for biology or biomedical applications; or a student with a background in evolutionary biology or microbiology with strong quantitative skills (including the basics of programming and data analysis).

In both cases, the other required skills are:

- spoken and written english (at least B2, C1 preferred)
- interest in interdisciplinary research

The objectives of a PhD include training, and it is not expected that candidates be already experts in the specific subject.

Applications should be sent by email to Antoine Frenoy ([antoine.frenoy@univ-grenoble-alpes.fr](mailto:antoine.frenoy@univ-grenoble-alpes.fr)), and include a CV, a short application letter, the contact of at least one reference (or a reference letter), and a transcript from M1+M2 grades if available.

There is no strict deadline, review of applications will start immediately and stop when a strong candidate is found.

Prospective students who are not sure if they fit the profile or who would like more information about the project and the environment should feel free to email Antoine Frenoy for informal enquiries.

## References

- [1] P. L. Foster. Methods for determining spontaneous mutation rates. *Method Enzymol* (2006), 195–213. DOI: [10.1016/S0076-6879\(05\)09012-9](https://doi.org/10.1016/S0076-6879(05)09012-9).
- [2] S. E. Luria and M. Delbrück. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* (1943), 491. DOI: [10.1093/genetics/28.6.491](https://doi.org/10.1093/genetics/28.6.491).
- [3] D. E. Lea and C. A. Coulson. The distribution of the numbers of mutants in bacterial populations. *J Genet* (1949), 264–285. DOI: [10.1007/BF02986080](https://doi.org/10.1007/BF02986080).
- [4] W. T. Ma, G. vH. Sandri, and S. Sarkar. Analysis of the Luria–Delbrück distribution using discrete convolution powers. *J Appl Probab* (1992), 255–267. DOI: [10.2307/3214564](https://doi.org/10.2307/3214564).
- [5] S. Gagneux et al. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* (2006), 1944–1946. DOI: [10.1126/science.1124410](https://doi.org/10.1126/science.1124410).
- [6] A. Frenoy and S. Bonhoeffer. *Death and population dynamics affect mutation rate estimates and evolvability under stress in bacteria*. 2018. DOI: [10.1371/journal.pbio.2005056](https://doi.org/10.1371/journal.pbio.2005056).
- [7] K. Csilléry, M. G. Blum, O. E. Gaggiotti, and O. François. Approximate bayesian computation (ABC) in practice. *TREE* (2010), 410–418. DOI: [10.1016/j.j.tree.2010.04.001](https://doi.org/10.1016/j.j.tree.2010.04.001).

- [8] M. A. Beaumont. Approximate Bayesian Computation. *Annu Rev Stat Appl* (2019), 379–403. DOI: [10.1146/annurev-statistics-030718-105212](https://doi.org/10.1146/annurev-statistics-030718-105212).
- [9] B. Tepekule, H. Uecker, I. Derungs, A. Frenoy, and S. Bonhoeffer. *Modeling antibiotic treatment in hospitals: A systematic approach shows benefits of combination therapy over cycling, mixing, and mono-drug therapies*. 2017. DOI: [10.1371/journal.pcbi.1005745](https://doi.org/10.1371/journal.pcbi.1005745).
- [10] M. Vasse, S. Bonhoeffer, and A. Frenoy. *Ecological effects of stress drive bacterial evolvability under sub-inhibitory antibiotic treatments*. 2022. DOI: [10.1038/s43705-022-00157-w](https://doi.org/10.1038/s43705-022-00157-w).
- [11] K. Csilléry, O. François, and M. G. B. Blum. Abc: an R package for approximate Bayesian computation (ABC). *Method Ecol Evol* (2012), 475–479. DOI: [10.1111/j.2041-210X.2011.00179.x](https://doi.org/10.1111/j.2041-210X.2011.00179.x).