RESEARCH CENTRE

Inria Lyon Center

IN PARTNERSHIP WITH:

Institut national des sciences appliquées de Lyon, Université Claude Bernard (Lyon 1), CNRS

2022

ACTIVITY REPORT

Project-Team

BEAGLE

Artificial Evolution and Computational Biology

IN COLLABORATION WITH: Laboratoire d'InfoRmatique en Image et Systèmes d'information (LIRIS)

DOMAIN

Digital Health, Biology and Earth

THEME

Computational Biology

Inria

Contents

Project-Team BEAGLE

Creation of the Project-Team: 2013 January 01

Keywords

Computer sciences and digital sciences

A3.3.2. – Data mining

A6.1.1. – Continuous Modeling (PDE, ODE)

A6.1.3. – Discrete Modeling (multi-agent, people centered)

A6.1.4. – Multiscale modeling

A6.2.7. – High performance computing

A8.1. – Discrete mathematics, combinatorics

Other research topics and application domains

B1. – Life sciences

B1.1.2. – Molecular and cellular biology

B1.1.6. – Evolutionnary biology

B1.1.7. – Bioinformatics

B1.1.10. – Systems and synthetic biology

B1.1.11. – Plant Biology

B1.2.1. – Understanding and simulation of the brain and the nervous system

B3.5. – Agronomy

B3.6. – Ecology

B9.2.1. – Music, sound

B9.2.4. – Theater

B9.9. – Ethics

1 Team members, visitors, external collaborators

Research Scientists

- Hugues Berry [INRIA, Senior Researcher, HDR]
- Antonius Crombach [INRIA, Researcher]
- Thomas Guyet [Inria (détachement), Researcher, HDR]
- Eric Tannier [INRIA, Senior Researcher, HDR]
- Leonardo Trujillo Lugo [INRIA, Advanced Research Position]

Faculty Members

- Guillaume Beslon [Team leader, INSA LYON, Professor, HDR]
- Carole Knibbe [INSA LYON, Associate Professor, HDR]
- Christophe Rigotti [INSA LYON, Associate Professor, HDR]
- Jonathan Rouzaud-Cornabas [INSA LYON, Associate Professor]

Post-Doctoral Fellow

- Jean-Sebastien Beaulne [Inria, from Sep 2022]

PhD Students

- Paul Banse [INSA LYON]
- Lisa Blum Moyse [INSA LYON]
- Lisa Chabrier [INRIA]
- Julie Etienne [INSERM, and Inria from october 2022]
- Marco Foley [INRIA]
- Théotime Grohens [MESRT]
- Arnaud Hubert [Inria, from Oct 2022]
- Juliette Luiselli [Insa Lyon, from Sep 2022]
- Arsene Marzorati [Inria, from Oct 2022]
- Nathan Quiblier [INRIA]
- Hana Sebia [Inria, from Nov 2022]

Technical Staff

- David Parsons [INRIA]
- Arnaud Tilbian [Inria, until Jan 2022]

Interns and Apprentices

- Yanis Sindt Baret [Inria, from Jun 2022 until Jul 2022]
- Léo Dautun [Inria, from Feb 2022 until Aug 2022]
- Aoife Igoe [Inria, from Jun 2022]
- Victor Lezaud [Inria, from Mar 2022 until Jul 2022]
- Juliette Luiselli [ENS Lyon]
- Ambre Mychalski [Inria, from Jun 2022 until Aug 2022]
- Yoann Sabatier Montanaro [Inria, from Jun 2022 until Jul 2022]
- Hana Sebia [Inria, from Mar 2022 until Aug 2022]

Administrative Assistants

- Laetitia Lecot [Inria, until Aug 2022]
- Salwa Selmi [Inria, from Sep 2022]

2 Overall objectives

2.1 An interface between biology and computer science

The expanded name for the BEAGLE research group is "Artificial Evolution and Computational Biology". Our aim is to position our research at the interface between biology and computer science and to contribute new results in biology by modeling biological systems. In other words we are making artifacts – from the Latin *artis factum* (an entity made by human art rather than by Nature) – and we explore them in order to understand Nature. The team is an INRIA Project-Team since January, 2014. It gathers researchers from INRIA, INSA, who are members of three different labs, the LIRIS [1], the LBBE [2], and CARMEN [3]. It is led by Prof. Guillaume Beslon (INSA-Lyon, LIRIS, Computer Science Dept.).

Our research program requires the team members to have skills in computer science but also in life sciences: they must have or develop a strong knowledge in biosciences to interact efficiently with biologists or, ideally, to directly interpret the results given by the models they develop. A direct consequence of this claim is that it is mandatory to restrict the domain of expertise in life sciences. This is why we focus on a specific scale, central in biology: the cellular scale. Indeed, we restrict our investigations on the cell, viewed as a dynamical system made of molecular elements. This specific scale is rich in open questions that deserve modeling and simulation approaches. We also focus on two different kinds of constraints that structure the cellular level: biophysical constraints and historical constraints. The cell is a system composed of molecules that physically interact and the spatio-temporal nature of these interactions is likely to strongly influence its dynamics. But the cell is also the result of an evolutionary process that imposes its own limits on what can evolve (or is the most likely to evolve) and what cannot (or is the less likely to evolve). A better understanding of what kind of systems evolution is the most likely to lead to in a given context could give us important clues for the analysis of extant biological systems.

[1] Laboratoire d'Informatique en Image et Systèmes d'Information: UMR 5205 CNRS, INSA-Lyon, Univ. Claude Bernard Lyon 1, Univ. Louis Lumière Lyon 2, École Centrale de Lyon

[2] Laboratoire de Biometrie et Biologie Evolutive: UMR CNRS 5558, Univ. Claude Bernard Lyon 1.

[3] Laboratoire de Recherche en Cardiovasculaire, Métabolisme, Diabétologie et Nutrition: UMR U1060 INSERM, INSA-Lyon, INRA 1235, Univ. Claude Bernard Lyon 1.

2.2 An organization into two tools and four main axes

To study these two kinds of constraints we mainly rely on two specific tools: computational cellular biochemistry and evolution models. We use these tools to develop our “artifacts” and we compare their output with real data, either direct measurements collected by experimentalists or ancestral properties computationally inferred from their extant descendants. The team research is currently organized in four main research axes. The first two ones are methodologically-oriented: we develop general formalisms and tools for computational cellular biochemistry (research axis 1) and families of models to study the evolutionary process (research axis 2). The third “NeuroCell” axis (research axis 3) is the one in which biochemical models are specifically applied on brain cells (neurons and glia). Eventually the last axis aims at integrating the two tools, computational biochemistry and evolution, in what we call "Evolutionary Systems Biology" (research axis 4). The next four sections describe these four axes in more details. The biological questions described are not the sole topics tackled by the team. They are the ones that mobilize a substantial fraction of the researchers on the long run. Many other questions are tackled by individual researchers or even small groups. In the following these ones will be briefly described in their methodological context, *i.e.* in the two sections devoted to research axes 1 and 2.

2.3 A strategy

The scientific objective of the BEAGLE team is to develop a consistent set of concepts and tools – mainly based on computational science – to *in fine* contribute to knowledge discovery in systems biology. Our strategy is to develop strong interactions with life science researchers to become active partners of the biological discovery process. Thus, our aim as a team is not to be a computer science team interacting with biologists, nor to be a team of biologists using computer science tools, but rather to stay in the middle and to become a *trading zone* [46] between biology and computer science. Our very scientific identity is thus fuzzy, melting components from both sciences. Indeed, one of the central claims of the team is that interdisciplinarity involves permanent exchanges between the disciplines. Such exchanges can hardly be maintained between distant teams. That’s why the BEAGLE team tries to develop local collaborations with local scientists. That’s also why BEAGLE also tries to organize itself as an intrinsically interdisciplinary group, gathering different sensitivities between biology and computer science inside the group. Our ultimate objective is to develop interdisciplinarity at the individual level, all members of the team being able to interact efficiently with specialists from both fields.

3 Research program

3.1 Introduction

As stated above, the research topics of the BEAGLE Team are centered on the modeling and simulation of cellular processes. More specifically, we focus on two specific processes that govern cell dynamics and behavior: Biophysics and Evolution. We are strongly engaged into the integration of these level of biological understanding.

3.2 Research axis 1: Computational cellular biochemistry

Biochemical kinetics developed as an extension of chemical kinetics in the early 20th century and inherited the main hypotheses underlying Van’t Hoff’s law of mass action : a perfectly-stirred homogeneous medium with deterministic kinetics. This classical view is however challenged by recent experimental results regarding both the movement and the metabolic fate of biomolecules. First, it is now known that the diffusive motion of many proteins in cellular media exhibits deviations from the ideal case of Brownian motion, in the form of position-dependent diffusion or anomalous diffusion, a hallmark of poorly mixing media. Second, several lines of evidence indicate that the metabolic fate of molecules in the organism not only depends on their chemical nature, but also on their spatial organisation – for example, the fate of dietary lipids depends on whether they are organized into many small or a few large droplets (see e.g. [47]). In this modern-day framework, cellular media appear as heterogeneous collections of contiguous spatial domains with different characteristics, thus providing spatial organization

of the reactants. Moreover, the number of implicated reactants is often small enough that fluctuations cannot be ignored. To improve our understanding of intracellular biochemistry, we study spatiotemporal biochemical kinetics using computer simulations (particle-based spatially explicit stochastic simulations) and mathematical models (age-structured PDEs).

3.3 Research axis 2: Models for Molecular Evolution

We study the processes of genome evolution, with a focus on large-scale genomic events (rearrangements, duplications, transfers). We are interested in deciphering general laws which explain the organization of the genomes we observe today, as well as using the knowledge of these processes to reconstruct some aspects of the history of life. To do so, we construct mathematical models and apply them either in a "forward" way, *i.e.* observing the course of evolution from known ancestors and parameters, by simulation (*in silico experimental evolution*) or mathematical analysis (*theoretical biology*), or in a "backward" way, *i.e.* reconstructing ancestral states and parameters from known extant states (*phylogeny, comparative genomics*). Moreover we often mix the two approaches either by validating backwards reconstruction methods on forward simulations, or by using the forward method to test evolutionary hypotheses on biological data.

3.4 Research axis 3: Computational systems biology of neurons and astrocytes

Brain cells are rarely considered by computational systems biologists, though they are especially well suited for the field: their major signaling pathways are well characterized, the cellular properties they support are well identified (e.g. synaptic plasticity) and eventually give rise to well known functions at the organ scale (learning, memory). Moreover, electro-physiology measurements provide us with an experimental monitoring of signaling at the single cell level (sometimes at the sub-cellular scale) with unrivaled temporal resolution (milliseconds) over durations up to an hour. In this research axis, we develop modeling approaches for systems biology of both neuronal cells and glial cells, in particular astrocytes. We are mostly interested in understanding how the pathways implicated in the signaling between neurons, astrocytes and neurons-astrocytes interactions implement and regulate synaptic plasticity.

3.5 Research axis 4: Evolutionary Systems Biology

This axis, consisting in integrating the two main biological levels we study, is a long-standing and long-term objective in the team. We have started to see significant advances in this direction, mainly due to the evolution of the team staff and team projects. These novel developments allow us to give this axis back its central place. We have several short and middle term projects that integrate biochemical data and evolution. First results were reported in 2019 with respect to an evolutionary perspective on chromatin-associated proteins. Other, ongoing projects include reverse engineering the regulatory networks of 'old' and 'young' brain regions (i.e. neuro-evo-devo) and finding new therapeutic targets for lung tumours that evolve treatment resistance.

4 Application domains

4.1 Functional and Evolutionary Biology

We do not usually distinguish our research and its application domains. Our shared idea is that the research is oriented by a scientific question, which in the case of the Beagle team is a multidisciplinary one, most often of biological nature. We do not develop methodologies independently from this question and then look for applications. Instead we collectively work with other disciplines to solve a question, using our competencies.

In consequence the application domains are already listed in the description of our projects and goals. They concern functional and evolutionary biology, related to critical social questions as human or global health.

4.2 Implication domains

We still advocate for the "application domains" section of the activity report to be called "implication domains" to broaden its scope. Implication contains applications, but not conversely.

This could allow us and others to report for example on orientation activities of our research programs guided by a social demand rather than by an intrinsic dynamic of scientific evolution, a simple claim for “progress”, or a social demand coming only from industry.

This could allow a better awareness of social and environmental issues, and integrate them in this section.

5 Social and environmental responsibility

5.1 Footprint of research activities

The website we constructed two years ago ferme.yeswiki.net/Empreinte can still be used for simple carbon footprint calculations of a team, but is or will be supplanted by future internal tools from Inria or the released ones of Labo1p5.

5.2 Impact of research results

We organized several "Sciences-Environnements-Sociétés" workshops in collaboration with Sophie Quinton from Inria Grenoble. A dozen one-day workshops have been organised in 2022. It started in Lyon and Grenoble in 2021, and has now been deployed in Rennes, Paris, Marseille, Sophia, Nancy. We have requests to organize it in Montpellier, Avignon, Orleans in 2023.

Besides this, Eric Tannier regularly teaches research ethics at university of Lyon, at Inria and University of Lyon 1, wuth a significant environmental focus.

We also lead an "action exploratoire" related to environmental issues, on the development of agro-ecology, as it is recommended by the IPCC (GIEC) on climate change and IPBES on biodiversity.

6 Highlights of the year

This year we highlight several publications in rather good journals, which witness achievements in several long lasting team projects:

- One article in *PNAS* [11], a reputed generalist scientific journal, on lactate vs glucose supply in neuronal activity.
- Two articles on chromosome inversions, one about their influence on the evolvability of organisms [25] in *PLoS Computational Biology*, and one about their importance in the history of biology [21] in *History and Philosophy of the life Sciences*.
- One article in *PLoS Biology* [24], on the influence of ghost species, which has yielded reports in general public journals: *Pour la science*, *Epsiloon* and *L'Humanité magazine*.

Besides this, the team ends this year and a significant amount of work has been dedicated to the construction of future teams.

7 New software and platforms

7.1 New software

7.1.1 aevol

Name: Artificial Evolution

Keywords: Bioinformatics, Genomics, Evolution

Functional Description: Aevol is a digital genetics model: populations of digital organisms are subjected to a process of selection and variation, which creates a Darwinian dynamics. By modifying the characteristics of selection (e.g. population size, type of environment, environmental variations) or variation (e.g. mutation rates, chromosomal rearrangement rates, types of rearrangements, horizontal transfer), one can study experimentally the impact of these parameters on the structure of the evolved organisms. In particular, since Aevol integrates a precise and realistic model of the genome, it allows for the study of structural variations of the genome (e.g. number of genes, synteny, proportion of coding sequences).

The simulation platform comes along with a set of tools for analysing phylogenies and measuring many characteristics of the organisms and populations along evolution.

An extension of the model (R-Aevol), integrates an explicit model of the regulation of gene expression, thus allowing for the study of the evolution of gene regulation networks.

News of the Year: In the context of the ANR Project "Evoluthon" we developed a new version of the aevol software that extends the binary code used in aevol into a 4-bases genetic code that respects the universal genetic code. Using this new version of the software, we have been able to simulate evolution along a speciation tree. The outcome of these simulations consists in 40 final genomic sequences that diverged for different amount of time. These 40 sequences have then been aligned using on-the-shelf bioinformatic software and the aligned sequences have been used to reconstruct the speciation tree. Comparison between the original simulated and the final reconstructed one shows that both trees diverge only on few branches, specifically branches for which the divergence times are very short. To the best of our knowledge, this result is the first attempt to reconstruct a phylogenetic tree from data generated by an artificial-life simulation. It simultaneously constitutes an important cross-validation step, both for the simulation software and for the inference method and opens the way to the simulation of demanding test sets using the 4-basis version of the software. These results have been presented during the Jobim conference.

URL: http://www.aevol.fr/

Contact: Guillaume Beslon

Participants: Paul Banse, Guillaume Beslon, Marco Foley, Theotime Grohens, Juliette Luiselli, Jonathan Rouzaud-Cornabas, Laurent Turpin, Leonardo Trujillo Lugo

Partners: UCBL Lyon 1, INSERM, Université Paris-Descartes, Insa de Lyon, University of Sherbrooke (Canada)

7.1.2 bioindication

Name: Bioindication

Keywords: Environment perception, Agroecology

Functional Description: Bioindication is a web platform designed to facilitate the reading of the landscape by users: identification of species living in a space, calculation of biodiversity indices, location, indicator values, suggestions of species or varieties to cultivate.

Release Contributions: First version

News of the Year: Outside Inria Bioindication is used for a participative science project in the Lyon metropolis, and to survey the biodiversity in a small town of the Monts du Lyonnais.

Bioindication has also been forked to an educational version. We designed a sequence of computer science labs in which this educational version of Bioindication plays an important role. The aim of these labs is to analyze interactions in ecosystem using graph data structures. The total lab sequence is 6 hours long. By the end of 2022 we made a preliminary run of the sequence with 800 students at INSA of Lyon (involving about 15 teachers). First feedbacks are positive, and we plan to strengthen the approach and develop it further for the next year.

URL: http://bioindication.com

Contact: Eric Tannier

Participants: Arnaud Tilbian, David Parsons, Eric Tannier, Damien De Vienne, Jean-Sebastien Beaulne, Christophe Rigotti

7.2 New platforms

New platforms: RAS .

8 New results

8.1 Inversion mutations open access to fitter adaptive peaks in NK fitness landscapes

Participants: P. Banse, L. Trujillo, G. Beslon.

Molecular evolution is often conceptualised as adaptive walks on rugged fitness landscapes, driven by mutations and constrained by incremental fitness selection. It is well known that epistasis shapes the ruggedness of the landscape's surface, outlining their topography (with high-fitness peaks separated by valleys of lower fitness genotypes). However, within the strong selection weak mutation (SSWM) limit, once an adaptive walk reaches a local peak, natural selection restricts passage through downstream paths and hampers any possibility of reaching higher fitness values. In addition to the widely used point mutations, we introduced a minimal model of sequence inversions to simulate adaptive walks. We used the well known NK model to instantiate rugged landscapes and showed that adaptive walks can reach higher fitness values through inversion mutations, which, compared to point mutations, allows the evolutionary process to escape local fitness peaks. To elucidate the effects of this chromosomal rearrangement, we used a graph-theoretical representation of accessible mutants and show how new evolutionary paths are uncovered.

This result suggests a simple mechanistic rationale to analyse escapes from local fitness peaks in molecular evolution driven by (intragenic) structural inversions and reveals some consequences of the limits of point mutations for simulations of molecular evolution. It has been published in the international journal PLoS Computational Biology [25].

8.2 A historical and philosophical view on chromosome inversions

Participants: E. Tannier.

This other result on chromosome inversions published in *History and Philosophy of the life sciences*[21] is the story, told in the light of a new analysis of historical data, of a mathematical biology problem that was explored in the 1930s in Thomas Morgan's laboratory at the California Institute of Technology. It is one of the early developments of evolutionary genetics and quantitative phylogeny, and deals with the identification and counting of chromosomal inversions in Drosophila species from comparisons of genetic maps. A re-analysis of the data produced in the 1930s using current mathematics and computational technologies reveals how a team of biologists, with the help of a renowned mathematician and against their first intuition, came to an erroneous conclusion regarding the presence of phylogenetic signals in gene arrangements. This example illustrates two different aspects of a same piece: (1) the appearance of a mathematical in biology problem solved with the development of a combinatorial algorithm, which

was unusual at the time, and (2) the role of errors in scientific activity. Also underlying is the possible influence of computational complexity in understanding the directions of research in biology.

8.3 Modelling the modulation of cortical up-down state switching by astrocytes

Participants: L Blum, H Berry.

Up-Down synchronization in neuronal networks refers to spontaneous switches between periods of high collective firing activity (Up state) and periods of silence (Down state). Recent experimental reports have shown that astrocytes can control the emergence of such Up-Down regimes in neural networks, although the molecular or cellular mechanisms that are involved are still uncertain. We proposed in [10] neural network models made of three populations of cells: excitatory neurons, inhibitory neurons and astrocytes, interconnected by synaptic and gliotransmission events, to explore how astrocytes can control this phenomenon. The presence of astrocytes in the models is indeed observed to promote the emergence of Up-Down regimes with realistic characteristics. Our models show that the difference of signalling timescales between astrocytes and neurons (seconds versus milliseconds) can induce a regime where the frequency of gliotransmission events released by the astrocytes does not synchronize with the Up and Down phases of the neurons, but remains essentially stable. However, these gliotransmission events are found to change the localization of the bifurcations in the parameter space so that with the addition of astrocytes, the network enters a bistability region of the dynamics that corresponds to Up-Down synchronization. Taken together, our work provides a theoretical framework to test scenarios and hypotheses on the modulation of Up-Down dynamics by gliotransmission from astrocytes.

8.4 Control of ca2+ signals by astrocyte nanoscale morphology at tripartite synapses

Participants: H Berry.

Much of the Ca2+ activity in astrocytes is spatially restricted to microdomains and occurs in fine processes that form a complex anatomical meshwork, the so-called spongiform domain. A growing body of literature indicates that those astrocytic Ca2+ signals can influence the activity of neuronal synapses and thus tune the flow of information through neuronal circuits. Because of technical difficulties in accessing the small spatial scale involved, the role of astrocyte morphology on Ca2+ microdomain activity remains poorly understood. We used in [denizot:hal-03582629] computational tools and idealized 3D geometries of fine processes based on recent super-resolution microscopy data to investigate the mechanistic link between astrocytic nanoscale morphology and local Ca2+ activity. Simulations demonstrate that the nano-morphology of astrocytic processes powerfully shapes the spatio-temporal properties of Ca2+ signals and promotes local Ca2+ activity. The model predicts that this effect is attenuated upon astrocytic swelling, hallmark of brain diseases, which we confirm experimentally in hypo-osmotic conditions. Upon repeated neurotransmitter release events, the model predicts that swelling hinders astrocytic signal propagation. Overall, this study highlights the influence of the complex morphology of astrocytes at the nanoscale and its remodeling in pathological conditions on neuron-astrocyte communication at so-called tripartite synapses, where astrocytic processes come into close contact with pre- and postsynaptic structures.

8.5 Lactate supply overtakes glucose when neural computational and cognitive loads scale up

Participants: H Berry.

Neural computational power is determined by neuroenergetics, but how and which energy substrates are allocated to various forms of memory engram is unclear. To solve this question, we asked whether neuronal fueling by glucose or lactate scales differently upon increasing neural computation and cognitive loads. Using electrophysiology, two-photon imaging, cognitive tasks, and mathematical modeling, we show in [11] that both glucose and lactate are involved in engram formation, with lactate supporting longterm synaptic plasticity evoked by high-stimulation load activity patterns and high attentional load in cognitive tasks and glucose being sufficient for less demanding neural computation and learning tasks. Indeed, we show that lactate is mandatory for demanding neural computation, such as theta-burst stimulation, while glucose is sufficient for lighter forms of activity-dependent long-term potentiation (LTP), such as spike timing–dependent plasticity (STDP). We find that subtle variations of spike number or frequency in STDP are sufficient to shift the on-demand fueling from glucose to lactate. Finally, we demonstrate that lactate is necessary for a cognitive task requiring high attentional load, such as the object-in-place task, and for the corresponding in vivo hippocampal LTP expression but is not needed for a less demanding task, such as a simple novel object recognition. Overall, these results demonstrate that glucose and lactate metabolism are differentially engaged in neuronal fueling depending on the complexity of the activity-dependent plasticity and behavior.

8.6 Genome-wide simulation of the Transcription-Supercoiling Coupling (TSC)

Participants: Théotime Grohens, Guillaume Beslon.

DNA supercoiling (SC), the level of under- or overwinding of the DNA polymer around itself, is widely recognized as an ancestral regulation mechanism of gene expression in bacteria. Higher negative SC levels facilitate the opening of the DNA double helix at gene promoters, and increase the associated expression levels. Different levels of SC have been measured in bacteria exposed to different environments, leading to the hypothesis that SC variation can be an environmental response. Moreover, DNA transcription has been shown to generate local variations in the SC level, and therefore to impact the transcription of neighboring genes.

We studied the coupled dynamics of DNA supercoiling and transcription at the genome scale by implementing a genome-wide model of gene expression based on the transcription-supercoiling coupling (TSC). We show that, in this model, a simple change in global DNA SC is sufficient to trigger differentiated responses in gene expression levels via the TSC. Then, studying our model in the light of evolution, we demonstrate that this SC-mediated non-linear response to environmental change can serve as the basis for the evolution of specialized phenotypes. These results have been published in the Artificial Life journal [15]. Furthermore, a variant of the model has been used to study the impact of TSC on genome structure. We showed that regulation of gene activity through TSC leads to specific genomic organization at all levels (gene-pairs, motifs and whole genome). A preprint has been published on BiorXiv [40] that shall be submitted soon in an international journal.

8.7 The Danaïde genome

Participants: M. Foley, P. Panse, V. Lezaud, J. Rouzaud-Cornabas, G. Beslon.

Using the Aevol simulator we experimentally studied the dynamic of genome size in prokaryote-like organisms. To this aim we evolve five “Wild-Type” organisms with the simulator until the size of their genomes stabilizes (which occurs after 10 million generations). We then propagated 50 clones of each wild-type for 2 million generations and monitor the dispersal of their genome size and, more specifically of the size of non-coding compartment of their genome. Given that the non-coding compartment is not submitted to selection, its size should follow a random dispersal with a lower bound in zero. However, our experiments revealed that its dispersal is limited by two boundaries, a lower boundary that is much larger than zero and an upper boundary. To understand the origin of these boundaries, we developed a new

analysis tool called “Neutral Mutation Accumulation”. Neutral Mutation Accumulation revealed that the non-coding compartment size is driven by two forces. (*i.*) a neutral force due to a fixation bias between duplications and deletion. Indeed, neutral duplications appear to be more numerous (and longer) than neutral deletions. This neutral force create a permanent flux of genomic material from the coding to the non-coding compartment, hence explaining why the non-coding compartment never reaches the zero bound. (*ii.*) a selective force due to robustness constraints (the longer the genome, the less robust it is). This selective force limits the expansion of the genome, hence explaining its upper boundary. Both forces explain the observed dynamics of the genome in Aevol. Moreover, since only one of them is selective, we conjectured that the balance between these two forces is driven by the intensity of the selection, hence by the population size. Indeed, by changing the population size in our simulation, we observed that larger population sizes lead to shorter genomes and that, on the opposite, smaller population sizes lead to larger genomes. An empirical law that is well known in microbiology. A publication is in preparation.

8.8 X-Aevol: A massive parallelism model to support GPU and accelerators

Participant: J. Rouzaud-Cornabas, L. Turpin.

X-Aevol is a response to the need of more computational power. It was designed to leverage the massive parallelization capabilities of GPU. As Aevol exposes an irregular and dynamic computational pattern, it was not a straightforward process to adapt it for massively parallel architectures. We present in [GECCO 2021] how we have adapted the Aevol underlying algorithms to GPU architectures. We implement our new algorithms with CUDA programming language and test them on a representative benchmark of Aevol workloads. We show that, by using the power of a GPU, we managed to massively accelerate the evaluation process of Aevol. We do performance evaluation on NVIDIA Tesla V100 and A100. We show how we reach a speed-up of 1,000 over a sequential execution on a CPU and the speed-up gain up to 50% from using the newer Ampere micro-architecture in comparison with Volta one. However we have shown that this is not an easy task and that algorithms have to be re-designed to match this massive parallelism. Our work is then a successful GPU port of a program conveying irregular structures of data with variable size thanks to different parallel algorithms and their implementation using advanced hardware operations. Our experimental setup relies on populations built to control the heterogeneity of the genomes. The main interest is to let possible to generate worst and best scenarios to measure performance of X-Aevol.

Future work is to make real simulation in order to simulate the full evolution of an artificial organism. Another point of interest is the ability to execute our GPU port on other vendor GPUs than the ones from NVIDIA. As CUDA is a proprietary parallel computing platform, it cannot be used for AMD’s or Intel’s GPUs. Frameworks and languages have emerged recently to unify the development of parallel computing to use different kinds of accelerators with the same base code while maintaining a high performance portability. Last but not least, studies show the impact of the size of populations on the genome size and structures. Accordingly, Aevol can be required to simulate very large population exceeding million individuals. To do so, the computing power of a single GPU will not be enough. We would have to work on multi-GPUs implementation using partitioning algorithms that will take into account the micro architectural properties of GPU and our inner knowledge of the biological model of Aevol to cut the overall population into smaller ones assigned to different GPUs.

8.9 The influence of ghost lineages

Participant: E. Tannier.

Introgression, endosymbiosis, and gene transfer, i.e., horizontal gene flow (HGF), are primordial sources of innovation in all domains of life. Our knowledge on HGF relies on detection methods that exploit some of its signatures left on extant genomes. One of them is the effect of HGF on branch lengths

of constructed phylogenies. This signature has been formalized in statistical tests for HGF detection and used for example to detect massive adaptive gene flows in malaria vectors or to order evolutionary events involved in eukaryogenesis. However, these studies rely on the assumption that ghost lineages (all unsampled extant and extinct taxa) have little influence. We demonstrate here with simulations and data reanalysis that when considering the more realistic condition that unsampled taxa are legion compared to sampled ones, the conclusion of these studies become unfounded or even reversed. This illustrates the necessity to recognize the existence of ghosts in evolutionary studies. This result has been the subject of two published articles [23, 24], with releases in the generalist press (see highlights).

8.10 A digital twin of the enterocyte for the intestinal uptake of fatty acids

Participant: C. Knibbe, J. Etienne.

The absorption of dietary triglycerides has recently been revealed as a key step in cardio-metabolic health, but the underlying molecular mechanisms in the enterocyte remain incompletely understood and are still debated. While many studies focused primarily on the roles of membrane proteins, other have suggested that a critical force governing fatty acid uptake could be the intracellular metabolic demand for fatty acids, which would drive entry by passive diffusion. In 2021, we had tested the compatibility of these hypotheses with experimental uptake data by expressing each of them into a quantitative mathematical model and by fitting it to seven experimental datasets. This had led us to conclude that intracellular metabolism, more than active transport, was a major force driving fatty acid uptake. However, in 2022, we detected a calibration error in one of the parameter, the membrane permeability. It was a composite parameter, capturing several physico-chemical steps: desorption from the micelles, diffusion through the aqueous medium outside the cell, adsorption into the cell membrane, flip-flop from one membrane leaflet to the other, desorption from the membrane into the cytoplasm. However, the numerical value we had used was only based on the flip-flop step, and was therefore off by several orders of magnitude. This had important consequences on the model kinetics (stiffness) and on the conclusions. Moreover, a more detailed version of the model, explicitly modelling each physicochemical step, also revealed that the impact of pH was different depending on the detail level of the model. For these two reasons, we developed the more detailed version of the model, with a modular design allowing for simulated gene knock-outs. With this new version of the model, we observed that the active transport mechanism appeared important to correctly fit short-term uptake data, on the time scale of a few seconds. However, on the time scale of several hours, the most crucial mechanism remains intracellular metabolism, which is required to ensure total absorption of the dietary content. These updated results were about to be submitted for publication.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

Participants: Eric Tannier.

We are a partner in a project leaded by the company Greenshield that has been funded 2 million euros by BPI France following a PIA4 call on agro-ecology for environmental transition.

10 Partnerships and cooperations

10.1 National initiatives

10.1.1 ANR

- “SecNet” (Spatio-temporal dynamics of second messenger networks), 2023- ANR grant, Call AAPG ANR 2022. signals. We combine cell biology approaches and computational modeling to provide a description of compartmentalized networks of second messengers that specifically activate the cellular response to repellent molecules (Slits and ephrinAs) both in axons and endothelial cells. Our project will provide a generalizable model that will be useful as a starting point for other cell types and calcium and cyclic nucleotide-dependent signaling pathways. Supervisor: X. Nicol (Vision Institute, Paris). Total amount funded: 533 k€.

- ABC4M, 2020-, Approximate Bayesian computation-driven multimodal microscopy to explore the nuclear mobility of transcription factors, a project funded by the French National Agency for Research (ANR), Call “AAP génériques 2020". We combine computer simulations and Approximate Bayesian computation with simultaneous multiple microscopy methods (FCS and spt-PALM) to quantify the way transcription factors explore the nucleus to find their binding sites. The project is supervised by H. Berry. Other participants are Institut Langevin, ESPCI, Paris (I. Izeddin), Phlam laboratory, Lille (L. Héliot) and Univ. Berlkeley, CA, USA (X. Darzacq). Total amount funded: 565 keuro.

- EngFlea (Engram of fast learning in the striatum), 2021-, Call AAPG ANR 2021. Our goal is to study the link between endocannabinoid-dependent plasticity and fast learning of rodents thanks to a multidisciplinary approach combining in vitro and in vivo experimental neurophysiology with detailed subcellular biophysical models and large-scale neural network models. Supervisor: L. Venance (CIRB, Collège de France, Paris). Participant: H. Berry.

- Evoluthon (2019-2024): Artificial Life as a benchmark for evolutionary studies, a 4-year project leaded by E Tannier with 2 partners, Beagle Inria and Le Cocon, LBBE.

- ANR Equipex+ grant "Spatial Cell Id" (2021-) coordinated by Yad Ghavi-Helm (IGFL), Olivier Hamant (RDP), and Jonathan Enriquez (IGFL) - 4,2M€. Anton Crombach and Christophe Godin are contact persons between Inria teams (Beagle, Dracula, Mosaic) and Yad, Olivier, Jonathan.

- NeGA 2021-, Ne effect on Genetic Architecture. By studying several eukaryotic species as well as evolution models like Aevol, NeGA aims at a better understanding of the influence of the effective population size (Ne) on the Genetic Architecture of these species. The project is supervised by Tristan Lefebure (LEHNA, Lyon). Other participants are the Beagle team, the LBBE (Lyon) and the ISEM (Montpellier).

- PEPR Digital Agro-ecology: we are partner in two flagship projects "Coeditag" and "Cobreeding", starting in 2023.

10.1.2 Inria

- Naviscope (Inria Project Lab, 2018-2022): image-guided Navigation and VIsualization of large data sets in live cell imaging and microSCOPy. Nowadays, the detection and visualization of important localized events and process in multidimensional and multi-valued images, especially in cell and tissue imaging, is tedious and inefficient. Specialized scientists can miss key events due to complexity of the data and the lack of computer guidance. In Naviscope we develop original and cutting-edge visualization and navigation methods to assist scientists, enabling semi-automatic analysis, manipulation, and investigation of temporal series of multi-valued volumetric images, with a strong focus on live cell imaging and microscopy application domains. We build Naviscope upon the strength of scientific visualization and machine learning methods in order to provide systems capable to assist the scientist to obtain a better understanding of massive amounts of information. Such systems will be able to recognize and highlight the most informative regions of

the dataset by reducing the amount of information displayed and guiding the observer attention. Head: C. Kervrann (Serpico), other EPIs: Aviz, Beagle, Hybrid, Morpheme, Mosaic, Parietal, and MaIage (INRA unit).

- Action Exploratoire "Community Garden Book": IPBES's recent report on declining biodiversity calls for generalization of agroecological, productive, biodiversity and environmental friendly methods, oriented towards participatory action research. This exploratory action is a proposal to develop tools from open science, evolution science and algorithmics for the co-construction and use of an agroecological network of interactions between groups, species, varieties found in fields and gardens.

- Action Exploratoire ExODE: In biology, the vast majority of systems can be modeled as ordinary differential equations (ODEs). Modeling more finely biological objects leads to increase the number of equations. Simulating ever larger systems also leads to increasing the number of equations. Therefore, we observe a large increase in the size of the ODE systems to be solved. A major lock is the limitation of ODE numerical resolution so ware (ODE solver) to a few thousand equations due to prohibitive calculation time. The AEx ExODE tackles this lock via 1) the introduction of new numerical methods that will take advantage of the mixed precision that mixes several floating number precisions within numerical methods, 2) the adaptation of these new methods for next generation highly hierarchical and heterogeneous computers composed of a large number of CPUs and GPUs. For the past year, a new approach to Deep Learning has been proposed to replace the Recurrent Neural Network (RNN) with ODE systems. The numerical and parallel methods of ExODE will be evaluated and adapted in this framework in order to improve the performance and accuracy of these new approaches.

  but also verified by various numerical tests which show the compensation of the error with the increase of the system size.

  As we have already seen, this method is characterized by its simplicity, its efficiency and above all its vast field of application, especially in biology with large and complicated systems. By the way, following all these mentioned advantages we note that through this article the study of the precision was done by considering the rounding error, whereas we know well that this is not the only error involved in optimizing accuracy.

  This encourages us to deal with approximation errors, in order to obtain a solver and a numerical scheme compatible with our mixed precision method, so we can be able to offer an optimal precision for large scale systems in future works. In order to do so, we will use existing tools (PROMISE [16] and VerifTracer [6]) to evaluate the numerical quality of our code and quantify the magnitude of floating point related errors. Nonetheless, one of our goal is to improve performance (execution time) of ODE solver. Thus we will do a thorough performance evaluation of our method on the different proposed biologicalsystems. To conclude, we will assess how our method can benefit from next generation computing platform. Especially, we will work on porting our method to take into account silicon based mixed precision implementations that were tailored for IA/ML.

10.1.3 Other National Initiatives

- France is preparing a response to the next EuroHPC call for expressions of interest with a view to hosting and operating one of the European exaflopic machines planned for 2024 within a consortium in which GENCI is the "Hosting Entity" and the TGCC at the CEA the "Hosting Site". This report presents a vision of the current state of the applications of the organizations involved in the Exascale France Project and the sizing of the technical and human needs related to these applications that will allow them to run on exaflopic machines, in order to remain competitive in the global digital landscape and to better size the French response to the EuroHPC AMI. SP3 proposes four major recommendations, which are transversal to all the research communities concerned by exascale computing, with a particular focus on the human resources required for the applications identified [45]

Participant: Jonathan Rouzaud-Cornabas.

- Fondation ARC funds the project CEDRiC, a collaboration of Anton Crombach with Sandra Ortiz-Cuaran (head), Pierre Martinez, Karene Mahtouk, and Janice Kielbassa from the Cancer Research Center of Lyon (CRCL) / Centre Léon Bérard (CLB). This is a two year grant of 50k€ for experiments (2021-2023).

Participant: Anton Crombach.

- Institut National du Cancer funds the project CLAIRE, a collaboration of Anton Crombach with Sandra Ortiz-Cuaran (head), Virginie Marcel, and Gabriel Ichim from the Cancer Research Centre of Lyon (CRCL) / Centre Léon Bérard (CLB). This is a three-year grant of 526 k€, including a postdoc position for the Beagle team. Duration: November 2022 – November 2025.

Participant: Anton Crombach.

10.2 Regional initiatives

- IXXI project "Evolution de la Complexité des Génomes". We have been granted a research project by the Rhône-Alpes Complex Systems Institutes (IXXI). The aim of the project is to foster the emerging collaboration with Nicolas Lartillot (LBBE).

Participant: Guillaume Beslon, Juliette Luiselli.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

General chair, scientific chair.

- We organized a half day symposium at the Jobim conference in Rennes in July 2022 on biological sequence simulations.

Participants: E. Tannier, J. Rouzaud-Cornabas, G. Beslon, D. Parsons.

- We organized a half day workshop "Parlement des êtres vivants pour une recherche en Transition" at the "école de l'anthropocène" in january 2022 lien

Participants: E. Tannier, P. Jensen, J. Michel.

11.1.2 Scientific events: selection

Chair of conference program committees

- Thomas Guyet is chair of the steering committee of the French Platform of Artificial Intelligence.

Member of the conference program committees

- Christophe Rigotti was a member of the program committee of the 2022 ECML/PKDD Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases.
- Thomas Guyet is member of the program committee of the following conferences: IJCAI, AAAI, ECML, CIKM, IDA, TIME.

11.1.3 Journal

Member of the editorial boards.

- Guillaume Beslon served as an associate editor for the journal Frontiers in Ecology and Evolution
- Eric Tannier is a member of the editorial committee of Peer Community in Evolutionary and of Peer Community in Mathematical and Computational Biology
- Hugues Berry is an Associate Editor for PLoS Computational Biology
- Thomas Guyet parts of the editorial committee of the Revue Française d'Intelligence Artificielle

Reviewer - reviewing activities.

- Guillaume Beslon served as a reviewer for Genome Biology and Evolution, EMBO Review Commons and Molecular Systems Biology.
- Eric Tannier served as a reviewer for Systematic Biology, Journal of Mathematical Biology

11.1.4 Invited talks

- Carole Knibbe, "Intestinal uptake of fatty acids: Quantitative modeling validates intracellular metabolism as a major driving force", Biomathematics workgroup of the Camille Jordan Institute, January 2022.
- Christophe Rigotti, March 2022. Invited talk at the I-RISK seminar (natural gravity risks). University Savoie Mont Blanc, fully on-line.
- Christophe Rigotti, May 2022. Invited talk at the strategic seminar of INDURA (sustainable development of infrastructures in Auvergne Rhône-Alpes), Campus région du numérique, Charbonnières-les-Bains.
- Anton Crombach, May 2022. GDR Math-Bio seminar series, Lyon.
- Anton Crombach, September 2022. TIMC seminar series on Computational and Mathematical Biology, Grenoble.
- Eric Tannier, March 2022, EDISS doctoral school.
- Eric Tannier, "Intellectual property as a tool for environmental responsibility", Inria Sophia Antipolis.
- Carole Knibbe, "Intestinal uptake of fatty acids: Quantitative modeling validates intracellular metabolism as a major driving force", Biomathematics workgroup of the Camille Jordan Institute, January 2022.

- Hugues Berry December (2022). Les alternatives à l'expérimentation animale : innovation et fiabilité, Table ronde, Cité des sciences et de l'Industrie de La Villette, Paris.
- Hugues Berry December (2022). Spike-timing dependent potentiation: from experiments to computational model. Master Neuroscience of the University of Lyon 1.
- Hugues Berry October (2022). AI in experimental biomedical research. Master AI4OneHealth, University of Grenoble-Alpes.
- Hugues Berry July (2022). Modelling the modulation of cortical up-down state switching by astrocytes. CNS 2022, Workshop "Emerging Perspectives and Models for Neuron-Glial Interactions", Melbourne, Australia.
- Hugues Berry July (2022). Modelling astrocytic calcium dynamics in soma or thin branchlets. CNS 2022, "Tutorial Models of Neuron-Glia Interactions", Melbourne, Australia.
- Hugues Berry June (2022). AI and digital twins for cell biology and beyond. I-Stem Institute.
- Hugues Berry June (2022). A research project for AI approaches in neuro-glio-pharmacology. Centre de Recheche en Neurosciences de Lyon, CRNL.
- Hugues Berry May (2022). Les promesses et quelques réalisations de l'IA et des jumeaux numériques en santé. Ecole Nationale Superieure de la Police.
- Huhues Berry May (2022). L'apport de l'IA pour la recherche expérimentale biomédical. Groupement des Animaleries de Grenoble.
- Hugues Berry April (2022). AI in experimental research: digital twins and 3R perspectives. 6th scientific day on technological innovation.
- Thomas Guyet (March 2022, Paris). Presentation of the OPTISOINS projet to the kickoff of Bernoulli Lab.
- Thomas Guyet (March 2023, Paris). Combining temporal and ontological reasoning : application to epidemiological studies". AFIA Workshop on "time everywhere".

11.1.5 Leadership within the scientific community

- Hugues Berry INSERM committee for health technologies (2022-today): appointed member of the 7th specialized scientific board (CSS7) of Inserm, the French National Institute for biomedical research and human health.
- Hugues Berry Elaboration of the PEPR Digital Health (2021-today). The PEPR Digital Health is the national program of the French government to fund research in digital health for 2022-2030.
- Hugues Berry Comité d'Expertise et Scientifique pour les Recherches, les Études et les Évaluations dans le domaine de la Santé (CESREES)(2020-today) - appointed member.
- Thomas Guyet parts of the board of the French Association of Artificial Intelligence (link).
- Thomas Guyet parts of the steering committee of the TIME Symposium (link).

11.1.6 Scientific expertise

- Guillaume Beslon was president of the HCERES evaluation committee of the LITIS Lab.
- Guillaume Beslon served as an expert for the ANRT (Association Nationale de la Recherche et de la Technologie)
- Christophe Rigotti served as an expert for the ANRT (Association Nationale de la Recherche et de la Technologie).

- Hugues Berry Grant reviews for the NC3R (UK National center for the Replacement, refinement and reduction of animal research), the NSERC (Natural Sciences and Engineering Research Council of Canada) and the ANR (AAPG)

11.1.7 Research administration

- Christophe Rigotti, elected member of Insa Scientific board (Conseil Scientifique).
- Eric Tannier is a member of the administration council of Inria.
- Hugues Berry is Deputy Scientific Director of Inria for “Digital Health and Biology" (2018-today).

11.2 Teaching - Supervision - Juries

11.2.1 Teaching

- License: Jonathan Rouzaud-Cornabas, Computer Architecture, 100h, L3, Computer Science Department, INSA Lyon
- Master: Jonathan Rouzaud-Cornabas, High Performance Computing, 60h, M2, Computer Science Department, INSA Lyon
- Master: Jonathan Rouzaud-Cornabas, High Performance Computing, 40h, M2, Biosciences Department, INSA Lyon
- Master: C. Knibbe, "Why use modelling in nutrition research", 2h CM, M2, master "Cardiovascular, metabolic and nutritional regulations" of Lyon 1 University.
- Licence: C. Knibbe, Fundamentals of algorithmics and programming, 48 heqTD, L3, Bioinformatics and Modelling program of INSA-Lyon
- Licence: C.Knibbe, Introduction to automatic data processing, 16h eqTD, L3, Biosciences program of INSA-Lyon
- Licence: C. Knibbe, HTML/CSS, 4 heqTD, L3, Bisociences program of INSA-Lyon
- Master: C. Knibbe, Careers in bioinformatics and modelling, 20 heqTD, M1, Bioinformatics and Modelling program of INSA-Lyon
- Licence: Christophe Rigotti, Object-Oriented Programming and Graphical User Interfaces, 86h, L2, Department 1er cycle of INSA-Lyon.
- Licence: Christophe Rigotti, Simulation of Chemical Reactions, 26h, L2, Department 1er cycle of INSA-Lyon.
- Licence: Christophe Rigotti, Numerical Modelling for Engineering, 60h, L2, Department 1er cycle of INSA-Lyon.
- Master: Christophe Rigotti, Data Mining, 55h, M1, Bioinformatics and Modeling Department, and Civil Engineering Department of INSA-Lyon.
- Master: Eric Tannier, String algorithmics, 12h, M1, Bioinformatics UCBL.
- Master: Eric Tannier, Research Ethics, 6h, M2,Bioinformatics UCBL
- Doctorat: Eric Tannier, Research Ethics, 8h, Inria.
- Doctorat: Eric Tannier, "Sciences, Environnements, Sociétés" for a doctoral internship "Enjeux de la recherche", Insa Lyon, march 2022.
- Licence: Guillaume Beslon, Computer Architecture, 100h, L3, Computer Science Department, INSA-Lyon

- Master: Guillaume Beslon, Computational Science, 25h, M2, Computer Science Department, INSA-Lyon
- Licence: Guillaume Beslon, Stage Lighting, 25h, L2, Humanities Department, INSA-Lyon
- **E-learning**
  - MOOC: Eric Tannier, member of the pedagogical team of the Research Ethics MOOC, FUN, released 2018, still online, Ph-D candidates, 3000 registered participants at each session.
  - Online ethic courses: Eric Tannier, 2 videos on research ethics on vimeo, uploaded in 2020 to diversify distant courses.

11.2.2 Supervision

- PhD in progress (CORDI-S PhD grant): Lisa Chabrier, "Differential analysis of regulatory networks in multi-omics data", supervised by Anton Crombach, Christophe Rigotti and Sergio Peignier (BF2I UMR203 - Biologie Fonctionnelle Insectes et Interactions), started October 2021.
- Ph-D: Alexandre Laverré, co-supervised by E Tannier, defended in July 2022 "Relations entre l'évolution des paysages cis-régulateurs, l'évolution de l'expression des gènes et l'évolution phénotypique chez les vertébrés"
- Ph-D: Hugo Menet, co-supervised by E Tannier, defended in July 2022 "Multi-scale phylogenetic approaches for the evolution of the holobiont"
- M2: Maël Thomas, M2 "Stratégies et Design pour l'Anthropocène", from STRATE school in Lyon, on intellectual property as a tool for environmental responsibility of researchers. Supervised by Eric Tannier
- M2. Hana Sebia, co-supervised by T. Guyet, defensed in June 2022: Temporal phenotyping of patients from EHR data based on tensor decomposition - PhD in progress (ANR Grant) : Marco Foley, "Evoluthon, développement d'une version 4 bases du logiciel Aevol", supervised by Guillaume Beslon, Jonathan Rouzaud-Cornabas
- PhD in progress (INRIA/INSERM Grant) : Julie Etienne, "A mathematical model for the uptake of fatty acids by the enterocyte", supervised by Carole Knibbe
- PhD in progress (CDSN) : Paul Banse, "Approches formelles de l'impact des réarrangements chromosomiques", supervised by Guillaume Beslon
- PhD in progress (CDSN) : Juliette Luiselli, "Influence de Ne sur l'architecture des génomes eucaryotes", supervised by Guillaume Beslon and Nicolas Lartillot (LBBE, Lyon)
- PhD in progress (ANR Grant) : Arnaud Hubert, "Modelling endocannabinoid-mediated synaptic plasticity and it implication in fast learning", supervised by Hugues Berry
- PhD in progress (INRIA AEx) : Arsène Marzorati, "Introduction of new numerical methods to take advantage of the mixed precision", supervised by Jonathan Rouzaud-Cornabas and Samuel Bernard (INRIA Dracula)
- PhD in progress (ANR Grant) : Nathan Quiblier, "mathematical modeling of the dynamical behavior of transcription factors in the nucleus", supervised by Hugues Berry.
- PhD: Theotime Grohens, supervised by Guillaume Beslon, defended in December 2022 "Ride the Supercoiling: Evolution of Supercoiling-Mediated Gene Regulatory Networks through Genomic Inversions"
- M2. Victor Lezaud, supervised by Guillaume Beslon, defended in June 2022: "modeling linear genomes in Aevol"
- M2. Aoife Igoe, co-supervised by Guillaume Beslon and Jonathan Rouzaud-Cornabas, defended in September 2022: "estimating Ne in Aevol simulations"

11.2.3 Juries

- Thomas Guyet, PhD Salah Boukhetta, Université La Rochelle, France, August 2022 (reviewer)
- Thomas Guyet, PhD Lamia Djébour, Université Montpellier, France, Sept 2022 (reviewer)
- Thomas Guyet, PhD Esso-Ridah Bleza, Université Bretagne Sud, France, December 2022 (reviewer)
- Thomas Guyet, PhD Carole Favier, Université Paris Cité, France, December 2022 (reviewer)
- Thomas Guyet, Member of the hiring committee of assistant professor, MCF4782, Strasbourg, France
- Thomas Guyet, Member of the hiring committee of assistant professor, MCF4385, Lens, France
- Hugues Berry HDR N. Glade, Univ. Grenoble-Alpes, France, November 2022 (reviewer)
- Hugues Berry HDR J. Leroy-Dudal, Cergy-Pontoise University, France, February 2022 (examiner)
- Hugues Berry PhD S. Abbara, Univ. Paris-Saclay, Pasteur Institute, France, October 2022 (reviewer)
- Hugues Berry PhD H. Verdier, Univ. Paris-Cite', Pasteur Institute, France, October 2022 (reviewer)

11.3 Popularization

11.3.1 Articles and contents

Eric Tannier was interviewed for "Mediacité" and for "L'âge de faire" about the "Sciences Environnements Sociétés" workshops.

12 Scientific production

12.1 Major publications

[1] Y. Cui, Y. Yang, Z. Ni, Y. Dong, G. Cai, A. Foncelle, S. Ma, K. Sang, S. Tang, Y. Li, Y. Shen, H. Berry, S. Wu and H. Hu. 'Astroglial-Kir4.1 in Lateral Habenula Drives Neuronal Bursts to Mediate Depression'. In: *Nature* 554 (Feb. 2018), pp. 323–327. DOI: 10.1038/nature25752. URL: https://hal.science/hal-01683191.

[2] A. Davín, E. Tannier, T. A. Williams, B. Boussau, V. Daubin and G. Szöllősi. 'Gene transfers can date the tree of life'. In: *Nature Ecology & Evolution* 2.5 (May 2018), pp. 904–909. DOI: 10.1038/s41559-018-0525-3. URL: https://hal.science/hal-01913812.

[3] Y. Dembitskaya, C. Piette, S. Perez, H. Berry, P. J. Magistretti and L. Venance. 'Lactate supply overtakes glucose when neural computational and cognitive loads scale up'. In: *Proceedings of the National Academy of Sciences of the United States of America* 119.47 (14th Nov. 2022). DOI: 10.1073/pnas.2212004119. URL: https://hal.inria.fr/hal-03922367.

[4] J. Lehman, J. Clune, D. Misevic, C. Adami, J. Beaulieu, P. J. Bentley, S. Bernard, G. Beslon, D. M. Bryson, N. Cheney, A. Cully, S. Doncieux, F. C. Dyer, K. O. Ellefsen, R. Feldt, S. Fischer, S. Forrest, A. Frenoy, C. Gagneé, L. Le Goff, L. M. Grabowski, B. Hodjat, L. Keller, C. Knibbe, P. Krcah, R. E. Lenski, H. Lipson, R. MacCurdy, C. Maestre, R. Miikkulainen, S. Mitri, D. E. Moriarty, J.-B. Mouret, A. D. Nguyen, C. Ofria, M. Parizeau, D. Parsons, R. T. Pennock, W. F. Punch, T. S. Ray, M. Schoenauer, E. Shulte, K. Sims, K. O. Stanley, F. Taddei, D. Tarapore, S. Thibault, W. Weimer, R. Watson and J. Yosinksi. 'The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities'. In: *Artificial Life* 26.2 (June 2020), pp. 274–306. DOI: 10.1162/artl_a_00319. URL: https://hal.inria.fr/hal-01735473.

[5] Q. Li, C. Hagberg, H. Silva Cascales, S. Lang, M. Hyvönen, F. Salehzadeh, P. Chen, I. Alexandersson, E. Terezaki, M. Harms, M. Kutschke, N. Arifen, N. Krämer, M. Aouadi, C. Knibbe, J. Boucher, A. Thorell and K. Spalding. 'Obesity and hyperinsulinemia drive adipocytes to activate a cell cycle program and senesce'. In: *Nature Medicine* 27.11 (Nov. 2021), pp. 1941–1953. DOI: 10.1038/s41591-021-01501-8. URL: https://hal.inria.fr/hal-03479060.

[6] C. Rocabert, G. Beslon, C. Knibbe and S. Bernard. 'Phenotypic noise and the cost of complexity'. In: *Evolution - International Journal of Organic Evolution* (Aug. 2020). DOI: 10.1111/evo.14083. URL: https://hal.archives-ouvertes.fr/hal-02920356.

[7] T. Tricou, E. Tannier and D. M. de Vienne. 'Ghost lineages can invalidate or even reverse findings regarding gene flow'. In: *Plos Biology* 20.9 (14th Sept. 2022), e3001776. DOI: 10.1371/journal.pbio.3001776. URL: https://hal.science/hal-03781025.

[8] B. Verd, E. Clark, K. Wotton, H. Janssens, E. Jiménez-Guri, A. Crombach and J. Jaeger. 'A damped oscillator imposes temporal order on posterior gap gene expression in Drosophila'. In: *PLoS Biology* 16.2 (16th Feb. 2018), p. 24. DOI: 10.1371/journal.pbio.2003174. URL: https://hal.inria.fr/hal-01934923.

[9] S. Zeppilli, T. Ackels, R. Attey, N. Klimpert, K. Ritola, S. Boeing, A. Crombach, A. Schaefer and A. Fleischmann. 'Molecular characterization of projection neuron subtypes in the mouse olfactory bulb'. In: *eLife* 10 (22nd July 2021). DOI: 10.7554/eLife.65445. URL: https://hal.science/hal-03473445.

12.2 Publications of the year

International journals

[10] L. Blum Moyse and H. Berry. 'Modelling the modulation of cortical Up-Down state switching by astrocytes'. In: *PLoS Computational Biology* 18.7 (21st July 2022), e1010296. DOI: 10.1371/journal.pcbi.1010296. URL: https://hal.inria.fr/hal-03737962.

[11] Y. Dembitskaya, C. Piette, S. Perez, H. Berry, P. J. Magistretti and L. Venance. 'Lactate supply overtakes glucose when neural computational and cognitive loads scale up'. In: *Proceedings of the National Academy of Sciences of the United States of America* 119.47 (14th Nov. 2022). DOI: 10.1073/pnas.2212004119. URL: https://hal.inria.fr/hal-03922367.

[12] A. Denizot, M. Arizono, V. Nägerl, H. Berry and E. de Schutter. 'Control of Ca2+ signals by astrocyte nanoscale morphology at tripartite synapses'. In: *Glia* (13th Sept. 2022). DOI: 10.1002/glia.24258. URL: https://hal.inria.fr/hal-03582629.

[13] P. Dusseux, T. Guyet, P. Pattier, V. Barbier and H. Nicolas. 'Monitoring of grassland productivity using Sentinel-2 remote sensing data'. In: *International Journal of Applied Earth Observation and Geoinformation* 111 (July 2022), p. 102843. DOI: 10.1016/j.jag.2022.102843. URL: https://hal.archives-ouvertes.fr/hal-03833748.

[14] M. A. Fortuna, G. Beslon and C. Ofria. 'Digital evolution: Insights for biologists'. In: *Frontiers in Ecology and Evolution* (13th Oct. 2022). DOI: 10.3389/fevo.2022.1037040. URL: https://hal.inria.fr/hal-03912093.

[15] T. Grohens, S. Meyer and G. Beslon. 'A Genome-Wide Evolutionary Simulation of the Transcription-Supercoiling Coupling'. In: *Artificial Life* (1st Nov. 2022). DOI: 10.1162/artl_a_00373. URL: https://hal.inria.fr/hal-03912118.

[16] A. Laverré, E. Tannier and A. Necsulea. 'Long-range promoter–enhancer contacts are conserved during evolution and contribute to gene expression robustness'. In: *Genome Research* 32.2 (1st Feb. 2022), pp. 280–296. DOI: 10.1101/gr.275901.121. URL: https://hal.archives-ouvertes.fr/hal-03559410.

[17] H. Menet, V. Daubin and E. Tannier. 'Phylogenetic reconciliation'. In: *PLoS Computational Biology* 18.11 (2022), e1010621. DOI: 10.1371/journal.pcbi.1010621. URL: https://hal.archives-ouvertes.fr/hal-03258402.

[18] S. Pageaud, A. Eyraud-Loisel, J.-P. Bertoglio, A. Bienvenüe, N. Leboisne, C. Pothier, C. Rigotti, N. Ponthus, R. Gauchon, F. Gueyffier, P. Vanhems, J. Iwaz, S. Loisel and P. Roy. 'Predicted Impacts of Booster, Immunity Decline, Vaccination Strategies, and Non-Pharmaceutical Interventions on COVID-19 Outcomes in France'. In: *Vaccines* 10.12 (2022), p. 2033. DOI: 10.3390/vaccines10122033. URL: https://hal.archives-ouvertes.fr/hal-03883673.

[19] S. Penel, H. Menet, T. Tricou, V. Daubin and E. Tannier. 'Thirdkind: displaying phylogenetic encounters beyond 2-level reconciliation'. In: *Bioinformatics* (9th Feb. 2022), pp. 1–3. DOI: 10.1093/bioinformatics/btac062. URL: https://hal.archives-ouvertes.fr/hal-03563666.

[20] N. Sourbier, K. Desnos, T. Guyet, F. Majorczyk, O. Gesny and M. Pelcat. 'SECURE-GEGELATI Always-On Intrusion Detection through GEGELATI Lightweight Tangled Program Graphs'. In: *Journal of Signal Processing Systems* (2022). DOI: 10.1007/s11265-021-01728-1. URL: https://hal.archives-ouvertes.fr/hal-03593545.

[21] E. Tannier. 'A hapless mathematical contribution to biology: Chromosome inversions in Drosophila 1937-1941'. In: *History and Philosophy of the Life Sciences* (2022). URL: https://hal.archives-ouvertes.fr/hal-03153696.

[22] E. Tannier, V. Daubin and S. Quinton. 'The crisis of the scientific mind : an investigation, a tragedy and a collective redistribution of roles'. In: *Les Cahiers de Framespa : e-Storia* 40 (30th June 2022). DOI: 10.4000/framespa.13150. URL: https://hal.archives-ouvertes.fr/hal-03714886.

[23] T. Tricou, E. Tannier and D. M. de Vienne. 'Ghost lineages can invalidate or even reverse findings regarding gene flow'. In: *Plos Biology* 20.9 (14th Sept. 2022), e3001776. DOI: 10.1371/journal.pbio.3001776. URL: https://hal.archives-ouvertes.fr/hal-03781025.

[24] T. Tricou, E. Tannier and D. M. de Vienne. 'Ghost lineages highly influence the interpretation of introgression tests'. In: *Systematic Biology* 71.5 (2022), pp. 1147–1158. DOI: 10.1093/sysbio/syac011. URL: https://hal.archives-ouvertes.fr/hal-03455377.

[25] L. Trujillo, P. Banse and G. Beslon. 'Getting higher on rugged landscapes: Inversion mutations open access to fitter adaptive peaks in NK fitness landscapes: Inversion mutations open access to fitter adaptive peaks in NK fitness landscapes'. In: *PLoS Computational Biology* 18.10 (31st Oct. 2022), e1010647. DOI: 10.1371/journal.pcbi.1010647. URL: https://hal.inria.fr/hal-03866346.

International peer-reviewed conferences

[26] L. Beuchée, T. Guyet and S. Malinowski. 'Prédiction du niveau de nappes phréatiques : comparaison d'approches locale, globale et hybride'. In: EGC 2022 - Conférence francophone sur l'Extraction et la Gestion des Connaissances. Blois, France, 24th Jan. 2022. URL: https://hal.inria.fr/hal-03548071.

[27] T. Guyet and N. Markey. 'Logical forms of chronicles'. In: *Leibniz International Proceedings in Informatics*. TIME 2022 - 29th International Symposium on Temporal Representation and Reasoning. Virtual, Heard & McDonald Islands, 7th Nov. 2022, pp. 1–15. URL: https://hal.archives-ouvertes.fr/hal-03777471.

[28] T. Guyet, L. Spillemaecker, S. Malinowski and A.-I. Graux. 'Temporal Disaggregation of the Cumulative Grass Growth'. In: *Lecture Notes in Computer Science. LNCS 13364. Part II*. ICPRAI 2022 - 3rd International Conference on Pattern Recognition and Artificial Intelligence. Vol. 13364. Pattern Recognition and Artificial Intelligence. Paris, France: Springer International Publishing, 29th May 2022, pp. 383–394. DOI: 10.1007/978-3-031-09282-4_32. URL: https://hal.inria.fr/hal-03899264.

[29] T. Guyet, W. Zhang and A. Bifet. 'Incremental Mining of Frequent Serial Episodes Considering Multiple Occurrences'. In: ICCS 2022 - 22nd International Conference on Computational Science. Vol. LNCS -13350. Computational Science – ICCS 2022, 22nd International Conference London, UK, June 21–23, 2022 Proceedings, Part I. London, United Kingdom: Springer International Publishing, 15th June 2022, pp. 460–472. DOI: 10.1007/978-3-031-08751-6_33. URL: https://hal.inria.fr/hal-03898790.

[30] V. Guyomard, F. Fessant, T. Guyet, T. Bouadi and A. Termier. 'VCNet: A self-explaining model for realistic counterfactual generation'. In: ECML PKDD 2022 - European Conference on Machine Learning and Knowledge Discovery in Databases. Grenoble, France, 19th Sept. 2022, pp. 1–16. URL: https://hal.inria.fr/hal-03899151.

[31] M. F. Mbouopda, T. Guyet, N. Labroche and A. Henriot. 'Experimental study of time series forecasting methods for groundwater level prediction'. In: ECML/PKDD 2022 - ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data. Grenoble, France, 19th Sept. 2022, pp. 1–16. URL: https://hal.uca.fr/hal-03790275.

[32] H. Sebia, T. Guyet and E. Audureau. 'Une extension de la décomposition tensorielle au phénotypage temporel'. In: EGC 2023 - 23ème Conférence Francophone sur l'Extraction et Gestion des Connaissances. Lyon, France, 16th Jan. 2023. URL: https://hal.inria.fr/hal-03902300.

[33] N. Sourbier, J. Bonnot, F. Majorczyk, O. Gesny, T. Guyet and M. Pelcat. 'Imbalanced Classification with TPG Genetic Programming: Impact of Problem Imbalance and Selection Mechanisms'. In: GECCO 2022 - Genetic and Evolutionary Computation Conference. GECCO '22: Proceedings of the Genetic and Evolutionary Computation Conference Companion. Boston, United States, 9th July 2022, pp. 1–4. DOI: 10.1145/3520304.3529008. URL: https://hal.archives-ouvertes.fr/hal-03699228.

[34] L. Spillemaecker, T. Guyet, S. Malinowski and A.-I. Graux. 'Désagrégation temporelle du cumul annuel de croissance de l'herbe'. In: EGC 2022 - Conférence francophone sur l'Extraction et gestion des connaissances. Vol. RNTI-E-38. EGC 2022. Blois, France, 2022, pp. 27–38. URL: https://hal.inria.fr/hal-03548073.

Conferences without proceedings

[35] T. Guyet, P. Saurel, H. Muscat, C. Paillard, T. Allard, N. Le Meur, D. Bourcier, E. Oger and A. Happe. 'IA et données de santé : regards croisés sur les droits, devoirs et responsabilités des personnes publiques'. In: Enjeux scientifiques et sociaux de l'IA. Paris, France, 6th Jan. 2023. URL: https://hal.inria.fr/hal-03935333.

[36] V. Guyomard, F. Fessant, T. Bouadi and T. Guyet. 'Générer des explications contrefactuelles à l'aide d'un autoencodeur supervisé'. In: Extraction et Gestion des Connaissances, EGC'2022. Extraction et Gestion des Connaissances, EGC'2022. Blois, France, 1st Feb. 2022. URL: https://hal.archives-ouvertes.fr/hal-03551329.

[37] P. Sawadogo, T. Guyet and E. Audureau. 'Conceptions de phénotypes computationnels pour la recherche en santé publique'. In: Santé et IA 2022. Saint-Etienne, France, 28th June 2022. URL: https://hal.archives-ouvertes.fr/hal-03675618.

Scientific book chapters

[38] L. Foulon, C. Rigotti, S. Fenet and D. Jouvin. 'Anomaly Detection Based on Sequence Indexation and CFOF Score Approximation'. In: *Advances in Knowledge Discovery and Management*. Vol. 1004. Studies in Computational Intelligence. Springer International Publishing, 15th Mar. 2022, pp. 47–61. DOI: 10.1007/978-3-030-90287-2_3. URL: https://hal.archives-ouvertes.fr/hal-03881940.

Edition (books, proceedings, special issue of a journal)

[39] Z. Bouraoui, V. Camps, É. Gaussier, M. Guériau, C. Launois, N. Prcovic, C. Rouveirol, F. Saïs, F. Schwarzentruber, A. Wilczynski, Y. Demazeau, T. Guyet, D. Longin, D. Monticolo and A. Samet, eds. *Actes CNIA 2022 : Conférence Nationale d'Intelligence Artificielle*. 27th June 2022, pp. 1–239. URL: https://hal.science/hal-03777860.

Reports & preprints

[40] T. Grohens, S. Meyer and G. Beslon. *A Genome-Wide Evolutionary Simulation of the Transcription-Supercoiling Coupling: extended version*. 13th May 2022. URL: https://hal.archives-ouvertes.fr/hal-03667822.

[41] A. Laverré, E. Tannier, P. Veber and A. Necsulea. *GOntact: using chromatin contacts to infer Gene Ontology enrichments for cis -regulatory elements*. 20th Sept. 2022. DOI: 10.1101/2022.06.13.495495. URL: https://hal.archives-ouvertes.fr/hal-03781024.

[42] H. Menet, A. Nguyen Trung, V. Daubin and E. Tannier. *Host-symbiont-gene phylogenetic reconciliation*. 25th Nov. 2022. DOI: 10.1101/2022.07.01.498457. URL: https://hal.archives-ouvertes.fr/hal-03781023.

Other scientific publications

[43] L. Chabrier, A. Crombach, S. Peignier and C. Rigotti. 'Semi-supervised learning for tree-based regressors to improve the prediction of the interactions between genes'. In: Symposium MaDICS. Lyon, France, 11th July 2022. URL: https://hal.archives-ouvertes.fr/hal-03894314.

[44] J. Etienne, A. Penhoat, M.-C. Michalski and C. Knibbe. 'Uptake of fatty acids by the enterocyte: the role of intracellular metabolism'. In: 12th European Conference on Mathematical and Theoretical Biology. Heidelberg, France, 19th Sept. 2022. URL: https://hal.inria.fr/hal-03921859.

12.3 Cited publications

[45] F. Boillod-Cerneux, M. Daydé, C. Loscos and J. Roman. *Les applications françaises face à l'exascale*. Research Report. CEA - Commissariat à l'énergie atomique et aux énergies alternatives ; CNRS - Centre National de la Recherche Scientifique ; Université de Reims Champagne Ardenne (URCA) ; INRIA Bordeaux, July 2022. URL: https://hal.archives-ouvertes.fr/hal-03736805.

[46] P. Galison. *Image and Logic: A Material Culture of Microphysics*. University Of Chicago Press, 1997.

[47] C. Vors, G. Pineau, L. Gabert and al. 'Modulating absorption and postprandial handling of dietary fatty acids by structuring fat in the meal: a randomized crossover clinical trial'. In: *Am J Clin Nutr* 97 (2013), pp. 23–36.