2023
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

B1.1.2. – 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

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

- Antonius Crombach [Inria, Researcher]

- Eric Tannier [Inria, Senior Researcher, HDR]

- Leonardo Trujillo Lugo [Inria, Advanced Research Position, until Oct 2023]

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 Fellows

- Jean-Sebastien Beaulne [Inria, Post-Doctoral Fellow]

- Hamza Chegraoui [Inria, Post-Doctoral Fellow, from Nov 2023]

PhD Students

- Paul Banse [INSA LYON, until Aug 2023]

- Lisa Blum Moyse [INSA LYON, until Aug 2023]

- Lisa Chabrier [Inria]

- Julie Etienne [Inria, until Mar 2023]

- Marco Foley [Inria, until Apr 2023]

- Romain Galle [Inria, from Oct 2023]

- Juliette Luiselli [INSA LYON]

- Arsene Marzorati [Inria]

- Thibaut Peyric [Inria, from Nov 2023]

Technical Staff

- Mouhamad Al-Sayed Ali [Inria, Engineer, from Jul 2023]

- David Parsons [Inria, Engineer, Research Engineer]

Administrative Assistant

- Claire Sauer [Inria]

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.

2.2   An organization into two tools and three 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). Eventually the last axis aims at integrating the two tools, computational biochemistry and evolution, in what we call "Evolutionary Systems Biology" (research axis 3). The next three sections describe these three 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* 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

---

[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.

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. 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. To improve our understanding of intracellular biochemistry, we study spatiotemporal biochemical kinetics using mathematical models and numerical simulations.

Specifically, with our biomedical partner lab (CarMeN) and our collaborators (Karolinska University Hospital, Stockholm), we investigate the molecular and cellular mechanisms that drive the fate of lipids as they are digested, absorbed, stored and mobilised in the body. Our goal is to build quantitative models of lipid fate based on data from in vitro digestion studies (enzymatic kinetics), from cellular cultures (gene expression data for genes involved in lipid metabolism or inflammation), from animal studies and from clinical studies (physiological data at the level of tissues and organs, like subcutaneous fat mass and visceral fat mass). We aim at taking into account the influence of the spatial organisation of dietary lipids (e.g. whether and how they are emulsified) on their fate in the body.

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: 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. Over the last years, we have accomplished 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 a central place in the next (future) team. We have several middle term projects that integrate molecular data and evolution. First results were reported in 2019 with respect to an evolutionary perspective on chromatin-associated proteins. In this report, we discuss results derived from data analysis of single-cell multi-omics RNA expression data of different mouse brain regions. Other, ongoing projects include reverse engineering the regulatory networks of 'old' and 'young' brain regions through image bioinformatics 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

We gave several conferences on the environemental footprint of research, in Rennes, Marseille, Lyon, and one online at Ecopolien.

5.2 Impact of research results

We organized several "Sciences-Environnements-Sociétés" workshops in collaboration with Sophie Quinton from Inria Grenoble. About 50 one-day workshops have been organised since 2021. It started in Lyon and Grenoble in 2021, and has now been deployed in Rennes, Paris, Marseille, Sophia, Nancy, Dijon, Bordeaux, and several others are planned in 2024.

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

We also handle this question as a research theme. One of the future teams arising from Beagle, hopefully in 2024, will have a focus on social, politics and environmental impacts of science and technologies.

6   Highlights of the year

This is possibly the last activity report of the Beagle team, as it is supposed to end before 2024. So remarkable events in the life of the team in 2023 have concerned the transformation of the team in two future teams, called BioTiC (Biologie Théorique et Computationnelle, Team Leader: Guillaume Beslon) and Semis (A multidisciplinary research on science and anthropocene, Team Leader: Eric Tannier).

7   New software, platforms, open data

7.1   New software

7.1.1   aevol

Name: Artificial Evolution

Keywords: Evolution, Simulation

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-bases version of the software. This dataset has been proposed to the community during the Alphy conference (Grenoble, February 2023) and several groups tested their reconstruction algorithms with it. Moreover, In the context of Juliette Luiselli's PhD, an extension of the aevol model is under development to extend the platform to eukaryotic genomes. This includes the possibility to simulate linear diploid chromosomes and sexual reproduction.

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

Contact: Guillaume Beslon

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

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: Bioindication was used, in an educational version, in two lincence modules at Insa-Lyon. This involved around 20 teachers and several hundred students.

URL: <http://bioindication.com>

Contact: Eric Tannier

Participants: Arnaud Tilbian, David Parsons, Eric Tannier, Damien De Vienne, Jean-Sebastien Beaulne, Christophe Rigotti, Hugo Daudey, Julien Barnier, Simon Penel

7.1.3 TopShapLite

Keyword: Explainable Artificial Intelligence

Functional Description: TopShapLite is an efficient implementation of our TopShap algorithm for scikit-learn regression models. TopShap is an algorithm that computes the top-K absolute SHAP values, including possible ties, and their confidence intervals. TopShap is agnostic in the sense that it can be applied to any kind of models, and only used the model as a black box. TopShap performs an iterative refinement of the set of top-K candidates by interleaving sampling operations, to improve SHAP value estimates, and pruning steps of the remaining candidates. The pruning is effective and leads to an important reduction of the number of calls to the model prediction function. TopShapLite is an implementation of TopShap based on NumPy vectorization and on batch-based sampling. It takes as input any scikit-learn regression model, but can also be used on any non-scikit-learn regressor object that has a method "predict" with the same signature.

URL: <https://gitlab.inria.fr/topshap/topshaplite_int>

Contact: Christophe Rigotti

Participants: Christophe Rigotti, Sergio Peignier, Lisa Chabrier, Antonius Crombach

Partners: LIRIS, BF2I, Insa de Lyon

8 New results

8.1 Mammalian olfactory cortex retains molecular signatures of ancestral cell types

Participants: A. Crombach.

The repertoire of behavioural and cognitive abilities of mammals is thought to arise from the vast diversity of neuronal cell types and circuits of the cerebral cortex. To understand cortical circuit functions, it is thus essential to reconstruct the molecular logic driving the diversification of cell types across cortical areas. We performed single-nucleus transcriptome and chromatin accessibility analyses to compare the molecular identities of neurons across three- to six-layered cortical areas of adult mice and across tetrapod species. We found that, in contrast to the six-layered mouse neocortex, glutamatergic neurons of

the three-layered mouse olfactory (piriform) cortex displayed a continuous rather than discrete variation in transcriptomic profiles. Surprisingly, subsets of glutamatergic cells with conserved transcriptomic profiles were distinguished by distinct, area-specific epigenetic states. Furthermore, we identified a prominent population of immature neurons in piriform cortex and observed that, in contrast to the neocortex, piriform cortex exhibited divergence between pyramidal cells in lab versus wild-derived mice. These results suggest a critical role for adult immature neurons in enhancing the adaptability of olfactory circuits. Finally, we showed that, unexpectedly, piriform neurons displayed marked transcriptomic similarities to cortical neurons in turtles, lizards, and salamanders. In summary, despite over 200 million years of co-evolution alongside the mammalian neocortex, olfactory cortex neurons seem to retain molecular signatures of ancestral cortical identity.

8.2 The automatic construction of plant identification keys

Participants: E. Tannier.

We have investigated a way to build a convivial plant identification tool halfway between the complex determination keys of botanists and the more recent but poorly explainable approaches based on AI image recognition. Our approach consists of a formal language to organize morphological traits and a Bayesian technique to describe plants with possible polymorphisms at all taxonomic levels, and to handle errors and uncertainties. From these structured data, automatic approaches can be designed to generate versatile determination keys, ie decision trees, which are otherwise tedious to design by hand. We published an article in the proceedings of the conference "Computing within limits". [18]

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

Participants: C. Knibbe, J. Etienne.

Intestinal absorption of dietary fatty acids is a key step in cardio-metabolic health. However, the molecular mechanisms underlying fatty acids uptake by the absorptive cells of the intestine, the enterocytes, remain incompletely understood. In 2023, we have completed the development and calibration of a quantitative and mechanistic model of the intestinal uptake of long-chain fatty acids, taking into account their hydrophobicity and their sensitivity to pH. This system of ordinary differential equations is composed of passive diffusion and of different modules (active transport, fatty acid-binding proteins (FABPs), intracellular metabolism), that we removed in turn to simulate gene knockouts. To allow for the quantitative comparison of uptake rates, we built a standardized dataset of long-chain fatty acid uptake rates based on nine published experimental datasets, originally expressed in various units. Our simulations show that intracellular metabolism, acting as a sink for passive diffusion through the membrane, is critical to ensure total absorption of the dietary content, at the time scale of several hours. Removing FABP does not prevent total absorption, but delays the process by more than a hundred hours. The presence of active transport does not impact the long-term uptake dynamics, but is required to properly fit experimental data at the time scale of a few seconds or minutes. Beyond goodness-of-fit, dissecting the quantitative contribution of each subflux shows that two kinds of dynamics can underly a good fit. In the "physiological" dynamics, fatty acids enter the cell via passive diffusion and active transport and end up metabolized. In the "anomalous" dynamics, fatty acids enter via active transport, but most of them leave via outward passive diffusion instead of being metabolized. Finally, we warn the community against a risk of misinterpreting experimental uptake response data: A saturated dose-response curve is not necessarily a hint that an active membrane transporter is primarily responsible for the uptake. We have submitted these results for publication and are working on a revised version of the manuscript. Julie Etienne, the PhD student who worked on this project with an Inria-Inserm fellowship, defended her thesis in June 2023.

8.4   Characterizing the fate of duplicated genes

Participants: Guillaume Beslon, Juliette Luiselli.

We initiated a collaboration with Sherbrooke University (Manuel Lafond, Reza Kalhor, Sherbrooke University, Canada) and ISEM (Céline Scornavacca, Montpellier, France). The objective of this collaboration is to use the aevol platform to simulate gene duplication in order to help characterizing the evolutionary fate of the duplicates. Indeed, although gene duplication has a central role in evolution, little is known on the fates of the duplicated copies, their relative frequency, and on how environmental conditions affect them. Moreover, the lack of rigorous definitions concerning the fate of duplicated genes hinders the development of a global vision of this process. We proposed a new theoretical framework aiming at characterizing and formally differentiating the fate of duplicated genes. This new framework has been tested via aevol simulations. Our results show several patterns to confirm previous studies and exhibit new tendencies; this opens up new avenues to better understand the role of duplications as driver of evolution. This research has been presented during the Recomb-CG conference [16] and an extended version is in revision for the Journal of Computational Biology. A long stay of Juliette Luiselli in Sherbrooke is planning for April-July 2024.

8.5   Origin of genome streamlining

Participants: Juliette Luiselli, Jonathan Rouzaud-Cornabas, Guillaume Beslon.

Genome streamlining, i.e. genome size reduction, is observed in bacteria with very different life traits —including cyanobacteria and endosymbiotic bacteria— raising the question of its evolutionary origin. None of the hypotheses proposed in the literature is firmly established, mainly due to the many confounding factors related to the diverse habitats of streamlined species. Computational models may help overcome these difficulties and rigorously test hypotheses. We used Aevol to test two main hypotheses: increase in either population size or mutation rate. Preevolved individuals were transferred into new conditions, characterized by either a population size increase, or a mutation rate increase. Both conditions lead to streamlining. However, the increased population size and mutation rate resulted in very different genome structures. Under increased population size, genomes have lost a significant fraction of non-coding sequences, but keep their coding genome size, resulting in densely packed genomes akin to cyanobacteria genomes. On the opposite, under increased mutation rates, genomes have lost both coding and non-coding sequences, akin to endosymbiotic bacteria genomes. In both cases, genome streamlining is largely driven by structural genomic variations and is due to an increased selection for robustness to structural genomic variants. However, under increased population size, selection for robustness is secondary to selection for fitness, hence the maintenance of coding sequences, while under increased mutation rate, selection for robustness outweighs selection for fitness, resulting in a loss of both coding and non-coding sequences.

8.6   Origin of evolutionary bursts in viruses

Participants: Paul Banse, Guillaume Beslon.

Viruses are known to evolve by bursts, which are often triggered by exogenous factors such as environmental changes, antiviral therapies or spill-overs from reservoirs into novel host species. However, other types of events have been suggested to be able to trigger evolutionary burst: either fitness valley crossing or a neutral exploration of a fitness plateau until an escape mutant is found on a neutral ridge. In order to investigate the importance of these different causes of evolutionary burst, we used aevol

to perform massive evolution experiments of viral-like genomes. We tested two conditions: after an "environmental" change or in constant conditions, this latter situation guaranteeing the absence of an exogenous triggering factor. As expected, an environmental change is almost systematically followed by an evolutionary burst. However, we show that bursts also occur, although much less frequently, in constant conditions. We analyze how many of these latter bursts are triggered by deleterious, neutral or beneficial mutations and we show that while bursts can occasionally be triggered by valley crossing or traveling along neutral ridge walking, many of them were triggered by chromosomal rearrangements, and in particular segmental duplications. Our results suggest that the difference in combinatorics between the different mutation types leads to punctuated evolutionary dynamics, with long periods of stasis occasionally interrupted by short periods of rapid evolution, akin to what is observed in virus evolution. This research is currently submitted. This work has been done in collaboration with Santiago Elena (professor at the Integrative Systems Biology, Valencia, Spain)

8.7 Modeling the evolution of non-coding sequences

Participants: Juliette Luiselli, Paul Banse, Marco Foley, Jonathan Rouzaud-Cornabas, Guillaume Beslon.

The various levels of non-coding sequences in genomes along the tree of life puzzles the evolutionary theory for decades. Starting from aevol simulations in which we observed that non-coding sequences are finely regulated, we proposed a new theory to explain their maintenance and regulation despite the absence of direct selective forces. We first conducted a very large scale simulation campaign in order to characterize the forces that act on non-coding sequences and showed that two opposite forces are at work: a neutrality bias that leads to non-coding sequences accumulation and indirect selection for robustness that limits genome expansion, hence the accumulation of non-coding sequences. Then, to generalize these results, we developed a probabilistic model that allows computing the equilibrium size of non-coding sequence. The model reproduce reasonably the amount of non-coding sequences observed in several organisms. This new model has been presented in several conferences and seminar and two articles are in preparation

8.8 Forward-in-time simulation of chromosomal rearrangements: The invisible backbone that sustains long-term adaptation

Participants: Paul Banse, Juliette Luiselli, David Parsons, Théotime Grohens, Marco Foley, Leonardo Trujillo Lugo, Jonathan Rouzaud-Cornabas, Carole Knibbe, Guillaume Beslon.

While chromosomal rearrangements are ubiquitous in all domains of life, very little is known about their evolutionary significance, mostly because very few models take them into account. As a consequence, we lack a general theory to account for their direct and indirect contributions to evolution. The platform Aevol is one of the few simulation platforms specifically dedicated to unraveling the evolutionary significance of chromosomal rearrangements (CR) compared to local mutations (LM). Using the platform, we evolve populations of organisms in four conditions characterized by an increasing diversity of mutational operators, from substitutions alone to a mix of substitutions, InDels and CR, but with a constant global mutational rate. Despite being almost invisible in the phylogeny owing to the scarcity of their fixation in the lineages, we show that CR make a decisive contribution to the evolutionary dynamics by comparing the outcome in these four conditions. As expected, chromosomal rearrangements allow fast expansion of the gene repertoire through gene duplication, but they also reduce the effect of diminishing-returns epistasis, hence sustaining adaptation on the long-run. Last, we show that chromosomal rearrangements tightly regulate the size of the genome through indirect selection for reproductive robustness. Overall, these results confirm the need to improve our theoretical understanding of the contribution of chromosomal rearrangements to evolution and show that dedicated platforms like Aevol can efficiently contribute to this agenda [10].

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 International initiatives

10.1.1 Participation in other International Programs

As participant on the NIH R01 grant "Gene regulatory network control of olfactory cortex cell type specification", Anton Crombach was awarded funding for a postdoctoral researcher and an engineer. This is a 5 year grant led by Alexander Fleischmann (Brown University, USA) with 2 partners, Ritambhara Singh (Brown University) and Anton Crombach. Total amount funded: $500k$.

Participants: Anton Crombach.

10.2 International research visitors

10.2.1 Visits of international scientists

Ruggero Pensa, Associate Professor from the University of Torino (Italy) We have benefited from support of Inria Lyon Centre through the program *Invited Researchers*, which allowed us to host Dr Ruggero Pensa for one month. Dr. Pensa is an expert in co-clustering methods and we started working together on the development of a new method to analyse gene expression data. In summary, in the analysis of gene expression data, we use so-called Shapley values to assess the importance of transcription factors (the features) to predict the expression of genes (the target variables) in different biological contexts. This leads to large tensors (transcription factor × target genes × biological contexts) of Shapley values that are very difficult to interpret. We are exploring co-clustering techniques as a highly effective approach to find patterns in these tensors by revealing 3-dimensional "blocks" composed of a group of transcription factors, a group of genes, and a group of biological contexts, that share a coherent set of large Shapley values. This collaboration will continue in 2024, thanks to an additional one-month visit of Dr. Pensa planned for springtime.

10.2.2 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€.

- 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.

- Flores (2023-2027): Participatory plant identification, 4 year project leaded by Simon Castellan (Inria Rennes), an answer to the call "Science with and for society", with academic and non academic partners. Participant: Eric Tannier (17keuros for Inria Lyon)

- Flowers (2023-2027): Plant trait evolution, 4 year project leaded by Sylvain Glemin (Rennes). Participant: Eric Tannier (92 keuros for Inria Lyon) -

- 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, flagship "Coeditag". A social science project around digital agroecology, on the economical and societal aspects of the digital technologies in agro-ecological farms. Participant: Eric Tannier

- PEPR Digital Agro-ecology, flagship "Cobreeding", a agronomic project the diversity of breeding animals and crops. Participant: Eric Tannier

- PEPR Santé Numerique, Flagship "AI4scMed" or "MultiScale AI for SingleCell-Based Precision Medicine", a project to develop methodology on the intersection of biology and medicine. Participants: Anton Crombach and Thibaut Peyric

10.2.3   Inria

- 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.

10.2.4   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 [21]. 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).

- 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, Hamza Chegraoui

10.3 Regional initiatives

Eric Tannier is one of the two leaders of a project "common health" with Anne-Laure Fougère, from university of Lyon 1, funded by "Shapemed" (university of Lyon 1). 150keuros.

Participants: E. Tannier.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: selection

Member of the conference program committees

- Christophe Rigotti was a member of the program committee of the IEEE International Conference on Data Mining (ICDM 2023).

- Eric Tannier was a member of the program committee of ISMB/ECCB 2023 conference on computational biology

- Guillaume Beslon and Jonathan Rouzaud-Cornabas were members of the program committee of the international conference on artificial life (ALife 2023)

11.1.2 Journal

Member of the editorial boards Eric Tannier is an editor for PCI evolutionary biology and PCI mathematical and computational biology

Reviewer - reviewing activities

- Eric Tannier reviewed for Journal of Mathematical Biology, and Systematic Biology.

- Guillaume Beslon reviewed for the ALife Journal, for Molecular Systems Biology and for EMBO.

11.1.3 Invited talks

- Guillaume Beslon, "Computational and mathematical evidence of fine regulation of non-coding sequences in genomes", Symposium on the evolution of cellular and genomic complexity, Utrecht University (NL), December 2023

- Guillaume Beslon, "Spontaneous regulation of non-coding sequences through border effect duplications neutrality bias", Schrool NVTB school, Schrool (NL) June 2024

- Guillaume Beslon, "Modeling the interplay between structural variants and single nucleotide polymorphisms in molecular evolution ", Groupe Evolution de l'ENS-Lyon, Lyon (France), mai 2023

- Eric Tannier, "l'engagement chez les scientifiques", participation to a round table at the conference Evolyon, Lyon, november 2023

- Eric Tannier, "rebound effects in bioinformatics", big day of the bioinformatics doctoral school, Rennes, november 2023

- Eric Tannier, "ateliers sciences sociétés", big day of the bioinformatics doctoral school, Rennes, november 2023

- Eric Tannier, "ateliers sciences sociétés", IGFL Lyon, december 2023

- Eric Tannier, "ateliers sciences sociétés", journée de l'école doctorale des mines de Saint-Etienne, January 2023.

- Eric Tannier, "ateliers sciences sociétés", Masterclass Ecole Doctorale BMIC Lyon, june 2023

- Eric Tannier, "se réapproprier la production de connaissance", Ecopolien Ile de France, June 2023

- Eric Tannier, "se réapproprier la production de connaissance", Brainhack neurosciences Marseille, October 2023

content

11.1.4   Scientific expertise

- Guillaume Beslon served as an external expert for the interdisciplinary doctoral program i-Bio at Sorbonne University.

- Eric Tannier reviewed for the Belgian FNRS Fond National pour la Recherche Scientifique

- David Parsons provided expertise on source code management to Laboratoire d'Océanographie et du Climat

11.1.5   Research administration

- Guillaume Beslon is a member of the "Commission des Moyens Incitatifs", Centre Inria de Lyon

- Guillaume Beslon, Juliette Luiselli et Lisa Chabrier are members of the "Comité de Centre", Centre Inria de Lyon

- Carole Knibbe is the director of the Biosciences department at INSA Lyon

- Eric Tannier is an elected member of the administration council, Inria

- David Parsons has been an administrator of the local AGOS (social words for Inria) association

- Eric Tannier is a member of the "FSS", formation spécialisée de site, Centre Inria de Lyon

- Eric Tannier is a member of the scientific council of the "Science shop", Boutique des sciences, université de Lyon 2

- David Parsons is a member of the steering committee of the Aramis network

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, Algorithmics and Python programming, 48 h eqTD, L3,Bioinformatics and Modelling program of INSA-Lyon

- Licence: C.Knibbe, Introduction to automatic data processing, 16 h eqTD, L3, Biosciences program of INSA-Lyon

- Master: C.Knibbe,Careers in bioinformatics and modelling, 20 h eqTD, M1, Bioinformatics and Modelling program of INSA-Lyon

- License: David Parsons, Linux - Local and Remote, 6h, L3, Biosciences Department, INSA Lyon

- Master : David Parsons, Software Development, 36h, M1, Biosciences Department, 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, Research Ethics, 6h, M2,Bioinformatics UCBL

- Doctorat: Eric Tannier, Research Ethics, 8h, Inria.

- 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

11.2.2 Supervision

- PhD, Paul Banse, "Evolution beyond substitutions: Computational modeling of the impact of chromosomal rearrangements on evolutionary dynamics", supervised by Guillaume Beslon, defended December 2023

- PhD, Marco Foley, "Dynamique des génomes bactériens : une étude expérimentale in silico avec la plate-forme aevol", supervised by Guillaume Beslon and Jonathan Rouzaud-Cornabas, defended December 2023

- PhD in progress: Juliette Luiselli, "dynamic of genome evolution in an eukaryotic model", supervised by Guillaume Beslon and Nicolas Lartillot (LBBE). Started September 2022.

- PhD, Julie Etienne, "Modélisation et simulation de la captation des acides gras alimentaires à longue chaîne et de leur re-synthèse en triglycérides au sein des entérocytes", supervised by Carole Knibbe and Marie-Caroline Michalski (CarMeN), defended June 2023.

- 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.

- PhD in progress (PEPR Santé Numérique): Thibaut Peyric, "Single-cell multi-omics data integration for gene regulatory network inference", supervised by Anton Crombach and Thomas Guyet (AIstrosight), started November 2023.

- Postdoc in progress (Institut du Cancer, project CLAIRE): Hamza Chegraoui, "Molecular mechanisms of cancer cell adaptation to targeted therapies: novel insights into the biology of drug-tolerance.", supervised by Anton Crombach and Sandra Ortiz-Cuaran (CRCL), started November 2023.

- Post-doc: Jean-Sébastien Beaulne, "Suggestions of plants based on ecological indicators for citizens and scientists", supervised by Eric Tannier (until feb. 2024)

- Engineer: Hugo Daudey, "Evoluthon: a benchmark of evolutionary tools based on artificial life", supervised by Eric Tannier (until oct. 2024) in the LBBE lab.

- M1, Lionel Dalmau, "Inférence de liens de régulation de gènes dans les données d'expression single-cell RNA-seq à l'aide de réseaux de neurones artificiels." supervised by Lisa Chabrier and Anton Crombach, finished in August 2023.

- M1, Félicie Chaudron "Development of a graphical user interface for the aevol platform" supervised by Paul Banse, Guillaume Beslon and David Parsons, finished in August 2023.

- PhD in progress (AEx ExODE PhD grant): Arsene Marzorati, "Scaling the solving of Ordinary Differential Equation for Computational Biology (and Deep Learning)", supervised by Samuel Bernard (ICJ UMR 5208 – Inria Dracula) and Jonathan Rouzaud-Cornabas, started October 2022.

- PhD in progress (Inria PhD grant): Romain Gallé, " Impact et évolution des modèles de programmation et des systèmes d'exécution pour les machines à mémoire hétérogène : application à la biologie computationnelle.", supervised by Thierry Gauthier (LIP UMR 5668 – Inria Avalon) and Jonathan Rouzaud-Cornabas, started October 2023.

- PhD defense (ANR Evoluthon PhD grant): Marco Foley, "Dynamique des génomes bactériens : une étude expérimentale in-silico avec la plate-forme aevol", supervised by Guillaume Beslon and Jonathan Rouzaud-Cornabas, defended December 2023.

- Mouhama Al Sayed Ali (AEx ExODE senior engineer), "Scaling the solving of Ordinary Differential Equation for Computational Biology (and Deep Learning)", supervised by Samuel Bernard (ICJ UMR 5208 – Inria Dracula) and Jonathan Rouzaud-Cornabas, started July 2023.

11.2.3 Juries

- Jonathan Rouzaud-Cornabas member of the PhD Defense jury of Lisa Blum Moyse : Computational neuroscience models at different levels of abstraction for synaptic plasticity, astrocyte modulation of synchronization and systems memory consolidation

- Guillaume Beslon, reviewer for the PhD defence of Sam van der Dunk: Constructive evolution and the emergence of complex cells. Utrecht (NL), December 2023

11.3   Popularization

11.3.1   Articles and contents

- Eric Tannier wrote an article for AOC, an online general public journal, "se réapproprier la production de connaissance" [13]

11.3.2   Education

the team regularly welcomes middle and high school students for research internships. In 2023, we welcomed two high school students (Elisa Vederine et Else Alexandre) for an optional discovery internship. They conducted research on genome evolution under stressful conditions (small populations and high mutation rate). The results were presented at the Dutch Society for Computational Biology summer school.

11.3.3   Interventions

- Eric Tannier was interviewed for the general public journal Mediacite and this article appeared: Savants-militants : ces scientifiques lyonnais qui s'engagent pour le climat

- Eric Tannier was interviewed by the online podcast journal "L'âge de faire": Chercher... Pourquoi? Comment?

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]   P. Banse, J. Luiselli, D. P. Parsons, T. Grohens, M. Foley, L. Trujillo, J. Rouzaud-cornabas, C. Knibbe and G. Beslon. 'Forward-in-time simulation of chromosomal rearrangements: The invisible backbone that sustains long-term adaptation'. In: *Molecular Ecology* (11th Dec. 2023). DOI: 10.1111/mec.17234. URL: https://hal.science/hal-04350147.

[11]   A. Comte, T. Tricou, E. Tannier, J. Joseph, A. Siberchicot, S. Penel, R. Allio, F. Delsuc, S. Dray and D. de Vienne. 'PhylteR: Efficient Identification of Outlier Sequences in Phylogenomic Datasets'. In: *Molecular Biology and Evolution* 40.11 (2023), msad234. DOI: 10.1093/molbev/msad234. URL: https://hal.science/hal-03995366.

[12]   H. Menet, A. Nguyen Trung, V. Daubin and E. Tannier. 'Host-symbiont-gene phylogenetic reconciliation'. In: *Peer Community Journal* (22nd May 2023), pp. 1–22. DOI: 10.24072/pcjournal.273. URL: https://hal.science/hal-03781023.

[13]   A. Monnin, E. Tannier and M. Thomas. 'Se réapproprier la production de connaissance'. In: *AOC [Analyse Opinion Critique]* (2023), https://aoc.media/opinion/2023/05/17/se-reapproprier-la-production-de\bibrangedashconnaissance/. URL: https://hal.science/hal-04102099.

[14]   A. Mouttou, E. Bremaud, J. Noero, R. Dibsy, C. Arone, J. Mak, D. Muriaux, H. Berry and C. Favard. 'Quantifying membrane binding and diffusion with fluorescence correlation spectroscopy diffusion laws'. In: *Biophysical Journal* 122.11 (June 2023), pp. 2216–2229. DOI: 10.1016/j.bpj.2023.01.006. URL: https://inria.hal.science/hal-04165544.

Invited conferences

[15]   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://inria.hal.science/hal-03935333.

International peer-reviewed conferences

[16]   R. Kalhor, G. Beslon, M. Lafond and C. Scornavacca. 'Classifying the Post-duplication Fate of Paralogous Genes'. In: *Lecture Notes in Computer Science*. RECOMB-CG 2023 - 20th conference on Comparative Genomics. Vol. 13883. Lecture Notes in Computer Science. Istanbul, Turkey: Springer; Springer Nature Switzerland, July 2023, pp. 1–18. DOI: 10.1007/978-3-031-36911-7_1. URL: https://hal.science/hal-04239853.

National peer-reviewed Conferences

[17] 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://inria.hal.science/hal-03902300.

Conferences without proceedings

[18] S. Castellan, J. Käfer and E. Tannier. 'Back to the trees: Identifying plants with Human Intelligence'. In: LIMITS 2023 - Ninth Workshop on Computing within Limits. Virtual, France: LIMITS, 14th June 2023, pp. 1–11. DOI: 10.21428/bf6fb269.265c52ce. URL: https://hal.science/hal-04121511.

Reports & preprints

[19] R. Beau, C. Aubertin, A. Aish, A. Atlan, S. Barot, O. Barrière, N. Belaïdi, F. Criscuolo, C. Damesin, V. Daubin, J. Dellaux, F. Ducarme, J. Fernández-Manjarrés, R. Freitas Machado, J. Garcia de Silveira, T. Genissel, S. Guillerme, J. Lecomte, H. Melin, M. Pelé, F. F. Pinton, S. Pouteau, B. Reber, F. Sarrazin, C. Sueur and E. Tannier. *Approche globale des éthiques dans les relations sociétés/nature*. CNRS - INEE, 27th Nov. 2023. URL: https://hal.science/hal-04311761.

[20] M. Morin and L. Chabrier. *Reported speech detection in newspapers*. Semaines Études Entreprises en Data Sciences (SEEDS), 3rd Oct. 2023. URL: https://hal.science/hal-04227066.

12.3   Cited publications

[21] 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.