Téo Lemane

5B rue de l'arcade
91540 Mennecy
France
☎ *0623314420*
✉ *teo.lemane@genoscope.cns.fr*
tlemane.github.io
tlemane

Education

2022 **Ph.D in computer science**, *University of Rennes 1, France*.

2019 **Master's Degree in bioinformatic, with high honours**, *University of Rennes 1, France*.
Subjects: Algorithmic, sequence analysis, programming, semantic web, combinatorial optimization, machine learning, biological networks, structural biology

2017 **Bachelor's Degree: biology of organisms**, *University of Rennes 1, France*.

2013 **High School Diploma, scientific option**, *Lycée André Malraux, Allonnes, France*.

Research experiences

Postdoc
Since May 2023
CEA/Genoscope
LAGE team

Large-scale sequence abundances indexing.
Supervisor: *Eric Pelletier*

Development of bioinformatic methods and tools for large-scale indexing of biological sequences with abundances, including powerful interfaces for end-users.

Ph.D thesis
2019 - 2022
Inria Rennes
Genscale team

Indexing and analyses of large sequencing collections using k-mer matrices.
Supervisors: *Pierre Peterlongo, Rayan Chikhi*

Development of bioinformatic methods for constructing k-mer matrices from large sequencing collections. This project focused on making the computational tools necessary to scale up and enable indexing and downstream analyses of these collections.

MSc2 Internship
January - July 2019
Inria Rennes
Genscale team

Search engine for genomic sequencing data.
Supervisor: *Pierre Peterlongo*

Study and development of algorithms and methods for indexing sequencing data.
Key words: Algorithmic, probabilistic data structures, indexing, C++ implementation, metagenomic

MSc1 Internship
April - July 2018
Inria Rennes
Dyliss team

Curation of the metabolic network of *Tisochrysis lutea* using heterogeneous methods.
Supervisor: *Jeanne Got*

Improvement of the metabolic network quality of Tisochrysis lutea, especially the metabolic pathways related to vitamin B12.
Key words: GEMs (Genome-scale metabolic models), flux balance analysis, topological analysis

Publications

2024 Beatriz Willink, Kalle Tunström, Sofie Nilén, Rayan Chikhi, Téo Lemane, Michihiko Takahashi, Yuma Takahashi, Erik I. Svensson, and Christopher West Wheat. The genomics and evolution of inter-sexual mimicry and female-limited polymorphisms in damselflies. *Nature Ecology & Evolution*, volume 8, pages 83–97, January 2024. Publisher: Nature Publishing Group.

2024 Téo Lemane, Nolan Lezzoche, Julien Lecubin, Eric Pelletier, Magali Lescot, Rayan Chikhi, and Pierre Peterlongo. Indexing and real-time user-friendly queries in terabyte-sized complex genomic datasets with kmindex and ORA. *Nature Computational Science*, volume 4, pages 104–109, February 2024. Publisher: Nature Publishing Group.

2023 Camila Duitama González, Riccardo Vicedomini, Téo Lemane, Nicolas Rascovan, Hugues Richard, and Rayan Chikhi. decOM: similarity-based microbial source tracking of ancient oral samples using k-mer-based methods. *Microbiome*, volume 11, page 243, November 2023.

2022 Téo Lemane, Paul Medvedev, Rayan Chikhi, and Pierre Peterlongo. kmtricks: efficient and flexible construction of Bloom filters for large sequencing data collections. *Bioinformatics Advances*, volume 2, page vbac029, January 2022.

2022 Téo Lemane, Rayan Chikhi, and Pierre Peterlongo. kmdiff, large-scale and user-friendly differential k-mer analyses. *Bioinformatics*, volume 38, pages 5443–5445, December 2022.

2022 Yoann Dufresne, Teo Lemane, Pierre Marijon, Pierre Peterlongo, Amatur Rahman, Marek Kokot, Paul Medvedev, Sebastian Deorowicz, and Rayan Chikhi. The K-mer File Format: a standardized and compact disk representation of sets of k-mers. *Bioinformatics*, volume 38, pages 4423–4425, September 2022.

2021 Grégoire Siekaniec, Emeline Roux, Téo Lemane, Eric Guédon, and Jacques Nicolas. Identification of isolated or mixed strains from long reads: a challenge met on Streptococcus thermophilus using a MinION sequencer. *Microbial Genomics*, volume 7, page 000654, 2021. Publisher: Microbiology Society,.

Talks

2022 **Téo Lemane**, Paul Medvedev, Rayan Chikhi, and Pierre Peterlongo. kmtricks: creating bloom filters for indexing large sequencing data collections. JOBIM, 2022.

2021 **Téo Lemane**, Rayan Chikhi, and Pierre Peterlongo. Advances in k-mer matrix construction for analysis of large sequencing collections. SeqBIM, 2021.

2020 **Téo Lemane**, Paul Medvedev, Rayan Chikhi, and Pierre Peterlongo. kmtricks: Modular k-mer count matrix and Bloom filter construction for large read collections. SeqBIM, 2020.

2020 **Téo Lemane**, Rayan Chikhi, and Pierre Peterlongo. HowDeSBT and Simka wedding: what has been done and what we plan to do? DSB, 2020.

Softwares

kmtricks **Modular k-mer count matrix and Bloom filter construction for large read collections**.
kmtricks is a generic tool suite along with a library for the construction of k-mer matrices and Bloom filters from large sequencing collections. It provides also a C++ plugin system to extend its features.
Availability: tlemane/kmtricks

kmdiff **Large-scale and user-friendly differential k-mer analyses**.
kmdiff is a tool allowing differential k-mer analysis on large cohorts (hundreds of individuals). It uses the concept of k-mer matrix to find differentially represented k-mers between two cohorts, opening the door to large-scale downstream analysis like structural variant detection.
Availability: tlemane/kmdiff

kmindex **Real-time k-mers query in TB-sized banks**.
kmindex is a tool for real-time querying of sequencing samples indexed using kmtricks. It allows to compute the percentage of shared k-mers between a query and each indexed sample.
Availability: tlemane/kmindex

Teaching

2019 - 2022 **String algorithms and bioinformatics**, *MSc1 Software Engineering, Istic Univ. Rennes 1, France*.

Computer skills

Systems	GNU/Linux (CentOS, Arch, NixOS, Debian)
Programming	Advanced C/C++, Python, Java
CI/CD	Github Actions, GitLab CI, Travis CI, Jenkins
Packaging	Nix, Conda, Docker
Version control	Git
Office	LaTeX, LibreOffice

Languages

French	Native
English	Advanced

Popular science

2021 **Organization of "Sciences en cour[t]s"**.
`http://sciences-en-courts.fr`

2020 **Production of an animated short film for "Sciences en cour[t]s".**
`http://sciences-en-courts.fr`

Referees

Dr. Pierre Peterlongo
Research Associate, Head of GenScale team
Inria, Rennes, France
✉ pierre.peterlongo@inria.fr
🖱 `http://people.rennes.inria.fr/Pierre.Peterlongo/`

Dr. Rayan Chikhi
Research Associate, Head of Sequence Bioinformatics team
Institut Pasteur, Paris, France
✉ rayan.chikhi@pasteur.fr
🖱 `http://rayan.chikhi.name`