

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** **Large-scale sequence abundances indexing.**
Since May 2023 **Supervisor:** *Eric Pelletier*
CEA/Genoscope
LAGE team
Development of bioinformatic methods and tools for large-scale indexing of biological sequences with abundances, including powerful interfaces for end-users.
- Ph.D thesis** **Indexing and analyses of large sequencing collections using k -mer matrices.**
2019 - 2022 **Supervisors:** *Pierre Peterlongo, Rayan Chikhi*
Inria Rennes
Genscale team
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** **Search engine for genomic sequencing data.**
January - July 2019 **Supervisor:** *Pierre Peterlongo*
Inria Rennes
Genscale team
Study and development of algorithms and methods for indexing sequencing data.
Key words: Algorithmic, probabilistic data structures, indexing, C++ implementation, metagenomic
- MSc1 Internship** **Curation of the metabolic network of *Tisochrysis lutea* using heterogeneous methods.**
April - July 2018 **Supervisor:** *Jeanne Got*
Inria Rennes
Dyliss team
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




- Under review **Téo Lemane**, Nolan Lezsoche, Julien Lecubin, Eric Pelletier, Magali Lescot, Rayan Chikhi, and Pierre Peterlongo. kmindex and ora: indexing and real-time user-friendly queries in terabytes-sized complex genomic datasets. *Nature Computational Science*, Under review.
- In press Beatriz Willink, Kalle Tunström, Sogie Nilén, Rayan Chikhi, **Teo Lemane**, M. Takahashi, Yuma Takahashi, Erik Svensson, and Christopher Wheat. The genomics and evolution of inter-sexual mimicry and female-limited polymorphisms in damselflies. *Nature Ecology and Evolution*, In press.
- In press Camila Duitama Gonzàles, Riccardo Vicedomini, **Téo Lemane**, Nicolas Rascovan, Hugues Richard, and Rayan Chikhi. Microbial source tracking for contamination assessment of ancient oral samples using k-mer-based methods. *Microbiome*, In press.
- 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*, 2022.
- 2022 **Téo Lemane**, Rayan Chikhi, and Pierre Peterlongo. kmdiff, large-scale and user-friendly differential k-mer analyses. *Bioinformatics*, Oct 2022.

- 2022 Yoann Dufresne, **Téo 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*, Jul 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 654. Microbiology Society, 2021.

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:  tlemanekmtricks
- 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:  tlemanekmdiff
- 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:  tlemanekmindex

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>