

# Antoine Limasset

Chargé de recherche CNRS

CRISAL, Université de Lille  
Lille, France  
✉ antoine.limasset@univ-lille.fr  
🌐 malfoy.github.io  
👤 Malfoy  
ORCID: 0000-0002-0669-4141  
Google Scholar  
Bluesky: @npmalfoy.bsky.social

## Education

- 2025 **HDR (Habilitation to direct research)**, Université de Lille  
Defended on 04/09/2025
- 2017 **Ph.D. in Computer Science**, Université de Rennes 1, Supervised by Pierre Peterlongo and Dominique Lavenier, Defended on 12/07/2017  
Novel approaches for the exploitation of high throughput sequencing data
- 2014 **M.Sc. in Computer Science**, École Normale Supérieure de Rennes
- 2012 **B.Sc. in Computer Science**, École Normale Supérieure de Cachan

## Experience

- 2018–present **Researcher**, CNRS, CRISAL, Lille  
Chargé de recherche in the Bonsai team
- 2017 **Postdoctoral researcher**, Université Libre de Bruxelles, Belgium  
Working on de novo assembly of heterozygous genomes
- 2014 **Trainee**, INRIA Rennes, France  
Five-month internship in the GENSCALE research group, working on read mapping on assembly graphs
- 2013 **Trainee**, Pennsylvania State University, US  
Ten-week internship in the Computer Science department, working on compaction of de Bruijn graphs
- 2012 **Trainee**, LIRMM, Montpellier, France  
Six-week internship in the MAB research group, on improving the Gk-arrays

## Publications

- 2026 **SEA, ZOR Filters: Fast and Smaller Than Fuse Filters**  
Sole author
- 2026 **RECOMB/Genome Research, Compressed inverted indexes for scalable sequence similarity**  
Last author
- 2026 **RECOMB / Journal of Computational Biology, Minimizer Density revisited: Models and Multiminimizers**  
Last author
- 2026 **RECOMB-Seq/Journal of Computational Biology, Super Bloom: Fast and precise filter for streaming k-mer queries**  
Last author
- 2026 **Peer Community Journal, Accelerating k-mer-based sequence filtering**  
Last author
- 2026 **PeerJ, Inverted colored de Bruijn Graph for practical k-mer sets storage**  
Last author
- 2025 **Algorithms for Molecular Biology, Fractional hitting sets for efficient multiset sketching**  
Last author
- 2025 **SPIRE, REINDEER2: practical abundance index at scale**  
Last author
- 2025 **Bioinformatics Advances, OReO: Optimizing Read Order for practical compression**  
Last author

- 2025 **Bioinformatics Advances**, *K2R: Tinted de Bruijn Graphs implementation for efficient read extraction from sequencing datasets*  
Last author
- 2025 **RECOMB**, *Hyper-k-mers: efficient streaming k-mers representation*  
Last author
- 2024 **PeerJ**, *Automated evaluation of multiple sequence alignment methods to handle third-generation sequencing errors*  
Last author
- 2024 **ISMB/Bioinformatics**, *Conway–Bromage–Lyndon (CBL): an exact, dynamic representation of k-mer sets*
- 2023 **ISMB/Bioinformatics**, *Locality-Preserving Minimal Perfect Hashing of k-mers*  
Last author
- 2023 **ISMB/Recomb-Seq**, *Scalable sequence database search using Partitioned Aggregated Bloom Comb-Trees*  
Last author
- 2023 **WABI**, *Fractional Hitting Sets for Efficient and Lightweight Genomic Data Sketching*  
Last author
- 2022 **WABI**, *Toward optimal fingerprint indexing for large scale genomics*  
Last author
- 2022 **Nature Methods**, *Critical Assessment of Metagenome Interpretation: the second round of challenges*
- 2021 **Science Advances**, *Chromosome-level genome assembly reveals homologous chromosomes and recombination in asexual rotifer *Adineta vaga**
- 2021 **Genome Biology**, *STRONG: metagenomics strain resolution on assembly graphs*
- 2021 **RECOMB-Seq/Bioinformatics**, *BLight: Efficient exact associative structure for k-mers*  
Last author
- 2021 **RECOMB-Seq/Scientific Reports**, *Scalable long read self-correction and assembly polishing with multiple sequence alignment*
- 2020 **RECOMB-Seq/Bioinformatics**, *Toward perfect reads: self-correction of short reads via mapping on de Bruijn graphs*  
First author
- 2020 **NAR Genomics and Bioinformatics**, *ELECTOR: evaluator for long-read correction methods*  
Last author
- 2020 **Discrete Applied Mathematics**, *A resource-frugal probabilistic dictionary and applications in bioinformatics*
- 2017 **Symposium on Experimental Algorithms**, *Fast and scalable minimal perfect hashing for massive key sets*  
First author
- 2016 **ISMB/Bioinformatics**, *Compacting de Bruijn graphs from sequencing data quickly and in low memory*
- 2016 **BMC Bioinformatics**, *Read mapping on de Bruijn graphs*  
First author
- 2014 **RECOMB**, *On the representation of De Bruijn graph*

## Supervision

- 2026–present **Ph.D. student**, Étienne Conchon-Kerjan, Director
- 2026–present **Research engineer**, Lucas Robidou, Supervisor  
Working on scalable genomic sequence analysis
- 2024–present **Ph.D. student**, Yohan Hernandez-Courbevoie, Director  
Working on indexing global transcriptomic databases

- 2023–present **Ph.D. student**, *Timothé Rouzé*, Supervisor  
Working on large-scale compression of sequencing collections
- 2025–2026 **Postdoctoral researcher**, *Florian Ingels*, Director  
Working on minimizer schemes
- 2023–2024 **Engineer**, *Caleb Smith*, Supervisor  
Working on large-scale compression of sequencing collections
- 2022–2025 **Ph.D. student**, *Léa Vandamme*, Director  
Working on indexing third-generation sequencing datasets
- 2019–2023 **Ph.D. student**, *Coralie Rohmer*, Supervisor  
Working on multiple sequence alignment algorithms adapted to third-generation sequencing

## Conference committees

- 2021 **SPIRE**, *Organizing committee*, International
- 2020-2026 **RECOMB**, *Program committee*, International
- 2020-2024 **ACM-BCB**, *Program committee*, International
- 2020-2026 **ECCB/ISMB**, *Program committee*, International
- 2020-2025 **SeqBim**, *Program committee*, National (France)

## Grants

- CDP – PIE**, *Protein-Interaction-Evolution*, 1.5M€ (total project funding)  
Université de Lille Initiative d'Excellence; four-year project (renewable)
- 2026 **ANR PRC – GRANDSMERS**, *Graph-based Research on Accurate Nucleotide Data via Scalable, Multi-scale, and Efficient RepresentationS*, ≈585k€, Principal investigator
- 2026 **ANR PRC – PRO-K-MER**, *PRObabilistic K-MERs for environmental sequence analysis*  
Member
- 2024 **MIC INSERM**, *Analyse efficace et évolutive du cancer par exploration transcriptomique avancée à grande échelle*, 554k€, Principal investigator
- 2024 **ANR**, *Shannon x Cray*, 500k€, Member
- 2021 **ANR JCJC**, *Adequate graph structures for third-generation sequencing data exploration*, 227k€, Principal investigator
- 2019 **Hauts-de-France Region PhD Grant**, 150k€, Principal investigator

## Reviewing activities

Nature Communications

Nature Methods

Genome Research

Bioinformatics

ACM-BCB

RECOMB

ISMB

WABI

ALMOB

Frontiers in Biology

Frontiers in Bioinformatics

SPIRE

Journal of Molecular Biology

Journal of Biological Macromolecules

GigaScience

Scientific Reports

---

## Thesis committees

- 2023-2025 **Nastasija Mijovic**, *PhD committee*, France
- 2022 **Riku Walve**, *Examiner*, Finland
- 2021 **Svitlana Lukicheva**, *PhD jury*, Belgium
- 2019-2020 **Nadege Guiglielmoni**, *PhD committee*, Belgium
- 2020-2021 **Théo Lemane**, *PhD committee*, France

---

## Teaching

- 2026 **EMBO Practical Course on Pangenomics**, *Instructor*, Naples, Italy, International course
- 2026 **Scalable Genomics and Pangenomics**, *Training team and organiser*, Wellcome Genome Campus, Hinxton, UK, International course
- 2020-2027 **Master's students**, *Genome assembly course*, France,  $\approx 20$  students
- 2019-2026 **Workshop Evomics**, *Genome assembly course*, Czech Republic,  $\approx 60$  students
- 2023-2026 **CNRS Formation**, *Genome assembly course*, France,  $\approx 20$  students
- 2015-2017 **Bachelor's students**, *Functional programming*, France,  $\approx 20$  students

---

## Software

- 2026 **Onika**, *Compressed inverted index for scalable sequence similarity search*
- 2026 **KLOE**, *Targeted compression and retrieval of large k-mer set collections*
- 2026 **SuperBloom**, *Fast and precise Bloom filter for streaming k-mer queries*
- 2026 **ZOR**, *Memory-efficient deterministic approximate membership filter*
- 2025 **REINDEER2**, *K-mer abundance index at scale*
- 2025 **K2R**, *K-mer-to-read index adapting de Bruijn graphs to long reads*
- 2025 **OReO**, *Read ordering to boost universal compressors*
- 2025 **KFC**, *Hyper-k-mers: efficient streaming k-mers representation*
- 2024 **MSALimit**, *Automated evaluation of multiple sequence alignment methods to handle third-generation sequencing errors*
- 2024 **CBL**, *Conway–Bromage–Lyndon: an exact, dynamic representation of k-mer sets*
- 2023 **LPHASH**, *Locality-Preserving Minimal Perfect Hashing of k-mers*
- 2023 **PAC**, *Scalable sequence database search using Partitioned Aggregated Bloom Comb-Trees*
- 2023 **SuperSampler**, *Fractional Hitting Sets for Efficient and Lightweight Genomic Data Sketching*
- 2022 **NIQKI**, *Toward optimal fingerprint indexing for large scale genomics*
- 2021 **STRONG**, *Metagenomics strain resolution on assembly graphs*
- 2021 **BLight**, *Efficient exact associative structure for k-mers*
- 2021 **CONSENT**, *Scalable long read self-correction and assembly polishing with multiple sequence alignment*
- 2020 **BCOOL**, *Toward perfect reads: self-correction of short reads via mapping on de Bruijn graphs*
- 2020 **ELECTOR**, *Evaluator for long-read correction methods*
- 2017 **BBHASH**, *Fast and scalable minimal perfect hashing for massive key sets*
- 2016 **Bcalm2**, *Compacting de Bruijn graphs from sequencing data quickly and in low memory*

---

## Recent Communications

- 2026 **Scalable Genomics and Pangenomics**, *Wellcome Genome Campus, Hinxton (UK)*, International, Training team and organiser

- 2026 **JC2B – Junior Conference on Computational Biology**, *Gif-sur-Yvette (France)*, National, Invited keynote speaker
- 2026 **RECOMB**, *Thessaloniki (Greece)*, International, Proceedings
- 2026 **RECOMB-Seq**, *Thessaloniki (Greece)*, International, Proceedings
- 2026 **EMBO Practical Course on Pangenomics**, *Naples (Italy)*, International, Instructor
- 2024 **EMBL-EBI K-mer/sequence indexing workshop**, *Cambridge (UK)*, International, No proceedings
- 2024 **Kmer days**, *Dijon*, National, No proceedings
- 2023 **ISMB**, *Lyon*, International, Proceedings
- 2022 **RECOMB**, *San Diego*, International, Proceedings
- 2022 **DSB**, *Düsseldorf*, International, No proceedings
- 2022 **TUDASTIC**, *Lille*, National, No proceedings
- 2022 **Genopim kickoff**, *Rennes*, National, No proceedings
- 2021 **Kmer days**, *Marville*, National, No proceedings
- 2019 **Biata**, *Saint Petersburg*, International, Proceedings
- 2018 **RECOMB**, *Paris*, International, Proceedings