# Sandra Romain

♥ Rosemary Gardens 53, SW14 7HD London, UK 🛛 🖂 sandra.romain.pro@gmail.com 🛛 📞 +33647506168

• https://github.com/SandraLouise

# Education and degrees \_\_\_\_\_

PhD degree	<ul> <li>University of Rennes - Matisse Doctoral School, Bioinformatics</li> <li>Title: Identification, genotyping and representation of structural variants in pangenomes (online manuscript 2)</li> <li>Diploma awarded on Nov 8 2024</li> <li>Jury composition: <ul> <li>Claire Lemaitre, Research Director, Inria/IRISA, PhD Director</li> <li>Fabrice Legeai, Research Engineer, INRAe, PhD Co-supervisor</li> <li>Birte Kehr, Professor, University of Regensburg, Reviewer</li> <li>François Sabot, Research Director, IRD, Reviewer</li> <li>Sèverine Berard, Senior Lecturer, University of Montpellier, Examinator</li> </ul> </li> </ul>	Sept 2021 – Nov 2024
	<ul> <li>Pierre Peteriongo, Research Director, Inria/IRISA, Jury Presi- dent</li> </ul>	
Master's degree	<ul><li>University of Rennes, Bioinformatics</li><li>Major: Computer Science and Integrative Biology</li></ul>	2019 - 2021
<b>Master's</b> 1st year	<ul> <li>University of Rennes &amp; AgroCampus Rennes, Biology and Agrosciences</li> <li>Major: Plant Improvement, Production and Development</li> </ul>	2018 – 2019
Bachelor's degree	<ul><li>University of Rennes, Life Sciences</li><li>Major: Plant Sciences</li></ul>	2015 – 2018

# Research experience \_\_\_\_\_

<ul> <li>Inria (National Institute for Research in Digital Science and Technology), PhD student</li> <li>Identification, genotyping and representation of structural variants in pangenomes</li> <li>Supervised by: Claire Lemaitre (Inria/IRISA) and Fabrice Legeai (IGEPP, INRAe)</li> </ul>	Rennes, France Sept 2021 – Nov 2024
<ul> <li>Inria (National Institute for Research in Digital Science and Technology), M2 Intern</li> <li>Genotyping structural variants with long read sequencing data and variation graph</li> <li>Supervised by: Claire Lemaitre (Inria/IRISA)</li> </ul>	Rennes, France Feb 2021 – July 2021
<ul> <li>CNRS (National Centre for Scientific Research), M1 Intern</li> <li>Automating the rendering of circular representations of genomic data with the Circos software suite - Application to highly duplicated genomes (Spartina maritima)</li> </ul>	Rennes, France May 2020 – June 2020
<ul> <li>Supervised by: Armel Salmon (ECOBIO, CNRS) and Malika Ainouche (ECOBIO, CNRS)</li> </ul>	

**INRAE** (French National Research Institute for Agriculture, Food and Environment), M1 Intern

- Analysing the genetic diversity of *Brassica napus* and identification of QTLs involved in the control of glucosinolates tissue levels
- Supervised by: Anne Laperche (IGEPP, INRAe) and Antoine Gravot (IGEPP, INRAe)

# **Teaching experience**

#### Text algorithm and bioinformatics

- Level: First year of Computer Science Master
- Lectures (6h), tutorial sessions (8h), practical sessions (10h), student project supervision

#### Algorithms, sequences and structures

- Level: Second year of Bioinformatics Master
- Practical sessions (8h)

Publications \_

**Sandra Romain**, Thibaut Capblancq, Laurence Després, Mathieu Joron, Fabrice Legeai and Claire Lemaitre. (*in prep.*). Characterization of large inversions to investigate hybrid speciation in the four species-complex of alpine *Coenonympha* butterfly.

**Sandra Romain**, Siegfried Dubois, Fabrice Legeai and Claire Lemaitre. (**2025**). Investigating the topological motifs of inversions in pangenome graphs. *bioRxiv*. doi.org/10.1101/2025.03.14.643331

Fabrice Legeai, **Sandra Romain**, Thibaut Capblancq, Paul Doniol-Valcroze, Mathieu Joron, Claire Lemaitre and Laurence Després. (**2024**). Chromosome-level assembly and annotation of the pearly heath *Coenonympha arcania* butterfly genome. *Genome Biology and Evolution*, *16*(3), evae055. doi.org/10.1093/gbe/evae055

**Sandra Romain** and Claire Lemaitre. (**2023**). SVJedi-graph: improving the genotyping of close and overlapping structural variants with long reads using a variation graph. *Bioinformatics*, *39*(Supplement\_1), i270-i278. doi.org/10.1093/bioinformatics/btad237

# Oral communications \_

Sandra Romain and Claire Lemaitre. SVJedi-graph: improving the genotyping of close and overlapping structural variants with long reads using a variation graph. **ISMB/ECCB (Intelligent Systems for Molecular Biology / European Conference on Computational Biology)**, July 2023, Lyon. (*talk*)

Sandra Romain, Fabrice Legeai and Claire Lemaitre. Understanding the limits of pangenome graphs for the analysis of large inversions in a complex of butterfly species. **EAGS (International Environmental and Agronomical Genomics Symposium)**, Feb. 2024, Toulouse. *(talk)* 

Sandra Romain and Claire Lemaitre. SVJedi-graph: using a variation graph to improve structural variant genotyping with long reads. **DSB (Data Structures in Bioinformatics)**, June 2022, Düsseldorf. *(talk)* 

Sandra Romain and Claire Lemaitre. SVJedi-graph: genotyping close and overlapping structural variants with a variation graph and long reads. **Journées Ouvertes en Biologie, Informatique et Mathématiques**, July 2022, Rennes. (*poster*)

University of Rennes Sept 2022 – Dec 2022

University of Rennes Sept 2022 – Dec 2022

Rennes, France May 2019 – July 2019

# Softwares

### SVJedi-graph

- Structural variant genotyper for long read sequencing data, using with a variation graph representation.
- Implemented in Python
- Distributed under GNU Affero licence, Open Source

#### **INVPG-annot**

- Tool to annotate inversion variants from a pangenome-based VCF.
- Implemented in Python
- Distributed under GNU Affero licence, Open Source

# Skills .

### **Bioinformatics**

- Sequence algorithm
- Genomic sequence analysis
- Structural variant analysis

## **Programming languages**

- Python
- R
- Bash
- Awk

### Tools

- Git
- LaTeX

github.com/SandraLouise/INVPG\_annot

github.com/SandraLouise/SVJedigraph