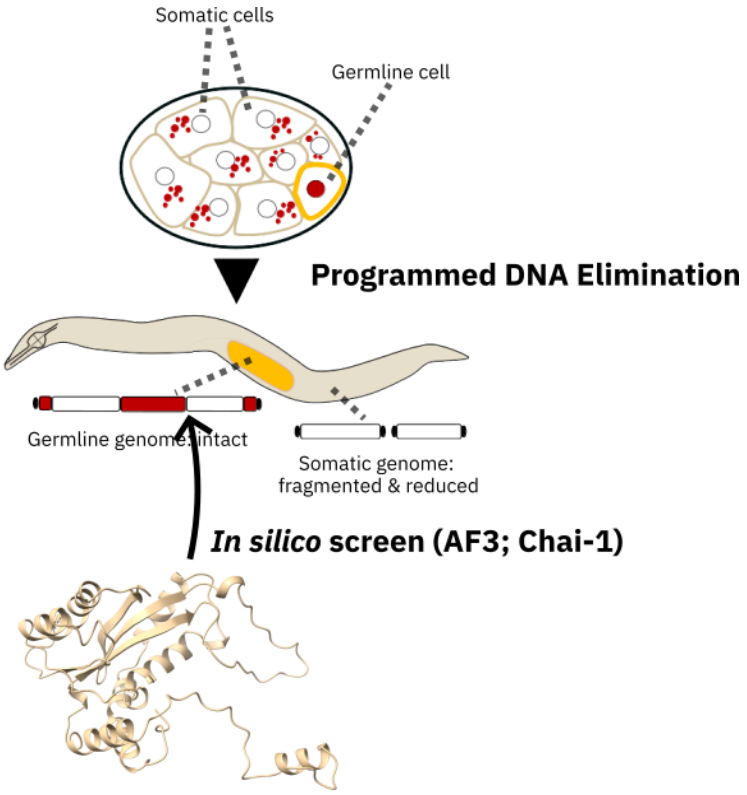
Internship Offer / Offre de Stage
(Level / Niveau: Master Bioinformatics)

Host Laboratory / Laboratoire d'accueil

Host Team: **Teams BPGE (LBBE) and WORM (LBMC)**
Unit Names: LBBE: Laboratoire de Biométrie et Biologie Evolutive, Campus La Doua.
LBMC: Laboratoire de Biologie et Modélisation de la Cellule, Campus ENS de Lyon.
Address: Your desk/workplace will be at LBBE (Campus La Doua).
Team Leaders: LBBE: Laurent Duret laurent.duret@univ-lyon1.fr; LBMC: Marie Delattre, marie.delattre@ens-lyon.fr
Supervisor: Brice Letcher brice.letcher@ens-lyon.fr
Contact: brice.letcher@ens-lyon.fr; Tel: 04 72 72 86 27

Project Description / Description du Projet

Title: ***In silico* screen for proteins involved in Programmed DNA Elimination in nematodes**



Description: Programmed DNA Elimination is a process in which specific parts of the genome are systematically destroyed during early embryonic development, in all the cells that will give rise to the soma. The genome of germline cells by contrast remains intact. While this process is in fact widespread among animals, its biological function and mechanisms remain largely unknown.

We have recently discovered Programmed DNA Elimination (PDE) in free-living nematodes and established them as a new model system to probe the function and mechanisms of PDE (1, 2). Using comparative genomics and RNA-Seq data, we recently produced a list of 21 candidate proteins that could be involved in chromosome breakage.

In this project, you will analyse the structural features of these proteins using state-of-the-art large language-based models. Your goal is to assess whether any of our

21 candidates could be truly involved in PDE (we can then test these in the lab!). To start off, we have generated the following data, using AlphaFold3 (3):

- Individual structures for the 21 candidates
- Homodimers of each of the 21 candidates, as we hypothesise the responsible protein(s) act as dimers
- Heterodimers of each candidate with an interactor that we hypothesise is in complex with the responsible protein(s)

Objective 1 (relatively easy)
In this objective, you will parse the AlphaFold3 outputs and analyse the different confidence scores (PAE, pTM, ipTM, ranking score) to assess which structures are well predicted and test the two hypotheses above. You can use ChimeraX for structure visualisation and further analyses (e.g., structural homology searches).

Objective 2 (moderate)
We additionally know that at least one of the responsible protein(s) binds to DNA, and we know the sequence of the targeted DNA. In this objective, you will predict DNA-protein complexes, starting with the best candidates from Objective 1.

For this you will use Chai-1 (4), an open-source model that performs well and that we have deployed at ENS de Lyon (we do not have AlphaFold3 deployed). You will have access to dedicated machines with GPUs to perform predictions. We will additionally use positive (known DNA-protein interactions) and negative (known DNA sequences not targeted for PDE) controls to calibrate and evaluate confidence scores.

Keywords: Structural bioinformatics; Programmed DNA Elimination; AlphaFold3; Chai-1
Techniques: Scripting in bash, R, Python; Workflow development using Nextflow

Publications

1. Rey C, Launay C, Wenger E, Delattre M. Programmed DNA elimination in *Mesorhabditis* nematodes. Current Biology. 2023 Sep;33(17):3711-3721.e5. https://doi.org/10.1016/j.cub.2023.07.058
2. Launay C, Wenger E, Letcher B, Delattre M. Somatic Programmed DNA Elimination is widespread in free-living Rhabditidae nematodes. bioRxiv; 2025 https://www.biorxiv.org/content/10.1101/2025.08.21.671558v1
3. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024 Jun;630(8016):493–500. https://doi.org/10.1038/s41586-024-07487-w
4. Chai Discovery, Boitreaud J, Dent J, McPartlon M, Meier J, Reis V, et al. Chai-1: Decoding the molecular interactions of life. bioRxiv; 2024 http://biorxiv.org/lookup/doi/10.1101/2024.10.10.615955