Master Internship offer

Long read-based genome assembly and comparative genomics of species undergoing Programmed DNA Elimination

Laboratoire de Biologie et Modélisation de la Cellule, Lyon
http://www.ens-lyon.fr/LBMC/equipes/NematodeCell

Some species systematically eliminate parts of their genome in somatic cells, in a process called programmed-DNA elimination (PDE). The germline cells by contrast maintain an intact genome (Figure 1). Although PDE has emerged multiple times throughout evolution in animals, we still don't know how the genome is scanned and excised at specific locations, or what PDE's ultimate function is (*1*). We fortuitously discovered PDE in the free-living nematode worms *Mesorhabitis* (*2, 3*). In contrast to the other animal species so far described, *Mesorhabditis* are genetically tractable,offering a unique opportunity to finally understand the mechanism, role and evolution of PDE.

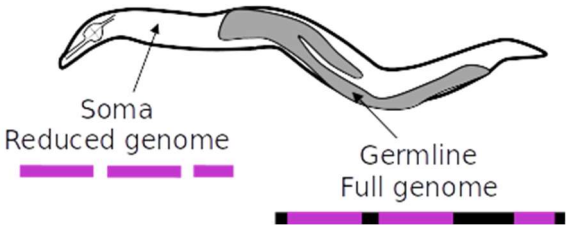

Figure 1. Programmed DNA Elimination (PDE) removes parts of the genome (black bars) in the somatic cells of certain species, while the germline genome remains intact (pink + black bars). We discovered PDE in *Mesorhabditis* nematodes (illustrated), and are characterising it using a mixture of bioinformatics and molecular biology.

In order to identify which sequences are eliminated in the soma and what specifies the breakpoints of chromosomes, we first need to separately assemble the somatic and the germline genomes (from mixed data, as germline- or soma-only sequencing is not possible). This is a challenge that we recently solved for one species using a combination of long read sequencing (Nanopore and PacBio HiFi) and Hi-C data. In this way we identified a specific motif occurring at breakpoints, and are now working on identifying the molecular machinery responsible for elimination.

During this internship, genomes of other nematode species undergoing PDE will be assembled following the same or newly-developed strategies (we have already collected the sequences). We will then ask if the different species eliminate the same type of sequences (genes, transposable elements) and whether the same motif is found at breakpoints. Through testing these various hypotheses, the project will also probe the evolutionary dynamics of PDE.

Prior experience: The successful candidate should be familiar with the basics of genome assembly and molecular evolution and be comfortable with bioinformatic scripting (some or all of python, R, shell). Prior experience in genomic bioinformatics is also a clear plus (any of genome assembly & annotation, comparative genomics, read mapping & variant calling). The ideal candidate should also be highly curious about biology – in particular genome biology and evolution.

Learning Objectives: Handling 3rd generation (PacBio, Nanopore) and Hi-C sequencing data; genome assembly & annotation; bioinformatic scripting; working in a biology-focussed environment; testing

evolutionary hypotheses. Depending on the candidate we can also explore lower-level programming (C, C++, Rust), test-driven development and statistics.

Supervision and contact: Brice Letcher (brice.letcher@ens-lyon.fr bioinformatics) and Marie Delattre (marie.delattre@ens-lyon.fr lab head). Send a CV and a cover letter to both of us.

References

1. J. Wang, R. E. Davis, Programmed DNA elimination in multicellular organisms. *Curr. Opin. Genet. Dev.* **27**, 26–34 (2014).
2. C. Rey, C. Launay, E. Wenger, M. Delattre, *bioRxiv*, in press, doi:10.1101/2022.03.19.484980.
3. M. Grosmaire, C. Launay, M. Siegwald, T. Brugière, L. Estrada-Virrueta, D. Berger, C. Burny, L. Modolo, M. Blaxter, P. Meister, M.-A. Félix, P.-H. Gouyon, M. Delattre, Males as somatic investment in a parthenogenetic nematode. *Science*. **363**, 1210–1213 (2019).