

Candidature à l'appel à projets 2026 – InLife & IPV

Summary of the PhD project

Title: Predicting the mutation signature of double strand break repair

Damaged DNA, particularly double strand breaks (DSBs), poses a significant risk to genome integrity and can lead to mutations that allow cancer to develop. DSB repair is crucial for genome stability, but repair pathways often introduce mutations like insertions or deletions, each with a distinct signature. This PhD project aims to develop a data-driven, interpretable machine learning model to predict the mutational signatures resulting from DSB repair in *Saccharomyces cerevisiae*.

Using an inducible CRISPR/Cas9 DSB system and high-throughput sequencing, the team has already generated mutational signature data at 11 genomic loci, revealing that repair outcomes are influenced by local sequence context. The project will first statistically analyze these data to quantify relationships between sequence features (e.g., GC content, repeats, microhomologies, etc.) and mutational patterns (e.g., insertion/deletion size and sequence content, single nucleotide variants, their frequencies, etc.). Next, a machine learning model will be trained to predict mutational signatures from local sequence context, aiming to uncover both suspected and novel mechanistic determinants. The third phase will experimentally test model predictions by engineering new DSB sites and manipulating sequence features. The ultimate goal is to extend the model to mutant strains, providing mechanistic insights into how specific genes shape mutational signatures.

This interdisciplinary project bridges computational and experimental biology, with potential applications to understanding mutational signatures in cancer genomes. The successful candidate should have strong programming skills, with a background in biology to facilitate collaboration with the experimental team.

Contacts: elodie.laine@sorbonne-universite.fr, zhou.xu@sorbonne-universite.fr