
Stage M2 - Développement bioinformatique en C++

Description

SortMeRNA est un logiciel de bioinformatique qui permet de réaliser l'identification et le filtrage des lectures d'ARN ribosomiques dans les données de transcriptome, génome, métagenome, métatranscriptome. Ce logiciel a été très largement adopté par la communauté, totalisant plus de 1700 citations Google Scholar.

SortMeRNA est développé depuis 2012 par l'équipe Bonsai (CRISAL, Lille) avec l'aide de collaborations (plateforme Bilille, Knight Lab, ...). Une refonte récente du code, écrit en C++, a permis la stabilisation des aspects algorithmiques au cœur de l'outil. Mais cette nouvelle version demande encore des développements complémentaires autour de l'installation, la gestion des entrées et sorties, l'utilisation dans des environnements HPC, et l'expérience utilisateur en général.

Le but principal de ce stage de M2 sera de réaliser des développements ciblés, prioritairement en C++, permettant des améliorations cumulatives de l'expérience utilisateur en général. Un objectif secondaire sera l'amélioration de la base documentaire du logiciel afin de faciliter la prise en main par les utilisateurs novices.

Activités principales

- Prise en main du logiciel SortMeRNA → devenir un "power-user"
- Développements ciblés en C++ ou Python des couches haut niveau du logiciel SortMeRNA (<https://github.com/biocore/sortmerna>)
- Amélioration de la gestion des fichiers temporaires, des index et des répertoires de sortie
- Renforcement de la base documentaire

Compétences requises

- Programmation en C++ ou Python
- Utilisation courante de Git, Bash et Unix
- Connaissances générales sur le séquençage à haut débit et ses applications aux domaines de la génomique et transcriptomique
- Langue anglaise: B2 (cadre européen commun de référence pour les langues) anglais technique du domaine
- Goût pour la programmation

Encadrement et conditions d'accueil

Le ou la stagiaire sera accueilli.e au sein de l'équipe de la plateforme bilille (UAR 2014 PLBS, <https://wikis.univ-lille.fr/bilille/>). Il ou elle sera encadré.e par Pierre Pericard (Ingénieur de Recherche, Bilille, PLBS) et Hélène Touzet (Directrice de Recherche, Bonsai, CRISAL).

Le stage sera réalisé en présentiel sur le campus de Cité Scientifique à Villeneuve d'Ascq, avec la possibilité d'un jour de télétravail par semaine.

La date de démarrage du stage est à fixer en fonction des contraintes de calendrier du Master du ou de la stagiaire.

Contact

Pierre Pericard (pierre.pericard@univ-lille.fr) et Hélène Touzet (helene.touzet@univ-lille.fr)