



Analyse génomique et phénotypique de patients atteints de leucémie myéloïde aiguë

Stagiaire en Intelligence Artificielle (Deep Learning) Affectation CRCT Oncopole

Durée : 5 à 6 mois

Rémunération : environ 600€ net mensuel

Le Centre de Recherche en Cancérologie de Toulouse

Le CRCT est une unité de recherche conjointe entre l'Inserm et l'Université Toulouse III Paul Sabatier. Au cœur de l'Oncopole de Toulouse, le CRCT, avec tous ses partenaires (institutionnels, universitaires, cliniques, industriels, caritatifs...), stimule l'innovation en termes de recherche et d'enseignement dans la lutte contre le cancer. Le CRCT conduit une approche intégrée entre la recherche, les soins et l'enseignement, dans une logique transversale et multidisciplinaire.

L'Institut de Recherche en Informatique de Toulouse

L'Institut de Recherche en Informatique de Toulouse (IRIT), une des plus imposantes Unités Mixtes de Recherche au niveau national, est l'un des piliers de la recherche en Occitanie avec ses 700 membres, permanents et non-permanents. De par son caractère multi-tutelle (CNRS, Universités toulousaines), son impact scientifique et ses interactions avec les autres domaines, le laboratoire constitue une des forces structurantes du paysage de l'informatique et de ses applications dans le monde du numérique, tant au niveau régional que national.

Contexte

La leucémie myéloïde aiguë (LMA) est un cancer des tissus hématopoïétiques (ou cancer du sang) fréquente chez les adultes, une maladie associée à un pronostic sombre. La pathogenèse moléculaire de la LAM est caractérisée par l'acquisition d'un petit nombre de mutations conduisant à une architecture clonale complexe. Les résultats génétiques limités combinés au caryotype sont utilisés pour définir le pronostic de la maladie et construire le projet clinique du patient.

Le génotype moléculaire seul a démontré son utilité pour la prise en charge individuelle des patients atteints de LAM mais, actuellement, seuls quelques gènes liés à l'apparition du cancer ont un traitement adapté. La fréquence des mutations pour chaque gène de cancer est très variable, et chaque LAM a typiquement plusieurs mutations drivers conduisant à un nombre très élevé de combinaisons de mutations. Nous voulons augmenter la prédiction de l'évolution des jeunes LAM traitées avec le nouveau standard de traitement en France et développer un support solide pour choisir parmi différentes options thérapeutiques (thérapies ciblées, allogreffe ou thérapie de maintenance) combinant le génotype de la maladie à son phénotype.

Travail demandé

Ce stage vise à mieux définir l'évolution de la maladie à partir d'analyses moléculaires au moment du diagnostic, en combinant le séquençage de l'ADN et le séquençage de l'ARNm.

Le séquençage de l'ADN permettra d'identifier les mutations ponctuelles, les petites insertions et les délétions combinées à l'analyse de l'ITD FLT3. Le séquençage de l'ARN identifiera également les mutations ponctuelles, les petites insertions et les délétions mais aussi les gènes de fusion. Il permettra également de générer le phénotype moléculaire défini par le transcriptome des cellules blastiques médullaires au moment du diagnostic.

A partir du transcriptome, nous pouvons étudier les propriétés spectrales et structurelles de la matrice obtenue pour détecter la structure des blocs cachés de manière non supervisée. Ces motifs extraits permettront de caractériser des classes de patients aux profils génomiques similaires. Ces méthodes seront basées sur la factorisation de matrices non négatives, des techniques d'apprentissage non supervisé. Ensuite un modèle prédictif non linéaire basé sur XgBoost et/ou réseaux de neurones sera conçu afin de prédire la survie des patients.

Formation

École d'ingénieur, de préférence avec spécialisation en informatique ou mathématiques (en année de césure ou stage long). Master 2 informatique ou mathématiques appliquées.

Compétences attendues

Des connaissances en science des données, apprentissage et réseaux de neurones seront appréciées.

Modalité de candidature

CV à envoyer à: LAM_GenoPheno@@stages-medecine-numerique.fr

Date limite de candidature : 31/01/2023

Encadrants: Eric Delabesse (CRCT), Sandrine Mouysset (IRIT), Philippe Leleux (IRIT/CRCT)