

Projet de recherche : Développement d'approches informatiques basées sur le Deep Learning pour l'évaluation conjointe des CNV (Copy Number Variant) et des SNV (Single Nucleotide Polymorphism) identifiés par séquençage de génome dans une cohorte de patients présentant une déficience intellectuelle sous l'hypothèse de digénisme.

Coordinateur du projet : Pr Valérie MALAN, Dr Antonio RAUSELL

Les troubles du neurodéveloppement concernent environ 2-3% de la population générale avec de nombreuses manifestations cliniques. Malgré l'importante prévalence collective de ces troubles, un diagnostic n'est atteint que pour environ la moitié des patients. A ce jour, plus de 1200 gènes ont été décrits dans ces atteintes cliniques.

Les technologies de séquençage de génome ont en effet permis de faciliter la détection des variations génétiques propres à chaque individu et de nombreux outils bio-informatiques ont été développés pour aider à leur interprétation. La plupart des stratégies diagnostiques actuelles sont basées sur:

- l'identification de variants (SNV, indel) codants et non-codants altérant l'expression d'un seul gène selon les modes d'hérédité mendélien
- la détection de variants de structure (SV, structural variant) équilibrés (point de cassure dans un gène ou sa séquence régulatrice) et déséquilibrés impliquant plusieurs gènes contigus

L'objectif de notre travail est d'évaluer la pathogénicité conjointe de variants (SNV, indel, SV) localisés dans des paires de gènes non contigus selon un mode digénique. Pour cela, nous utiliserons l'Intelligence Artificielle pour établir un score de pathogénicité de toutes les paires de gènes possibles à partir de bases de données génétiques publiques. Une fois que notre modèle aura été entraîné, nous l'appliquerons aux patients de la cohorte Defidiag pour lesquels l'analyse primaire n'a pas permis d'établir un diagnostic génétique. Nous espérons ainsi identifier des paires de variants affectant deux gènes différents à l'origine du phénotype du patient. Aucun de ces variants ne seraient prédits pathogènes par les outils actuels, mais leur occurrence simultanée entraînerait la déficience intellectuelle. Ce travail permettra d'augmenter le taux diagnostique du SHD de génome dans DI et sera une étape importante pour la compréhension des maladies multigéniques.