

## OFFRE DE STAGE DE RECHERCHE EN LABORATOIRE : M2 BIOINFORMATIQUE

**Sujet:** Analyse bioinformatique de données de séquençage génome entier dans des familles multiplexes avec scoliose idiopathique

**Structures d'accueil:** **1. Equipe GenMedStroke**

NeuroDiderot, Inserm U1141/Université de Paris  
Hôpital Robert Debré, 48 Boulevard Sérurier, 75019 Paris

**2. Equipe GenDev**

Centre de Recherche en Neurosciences de Lyon (CRNL)  
Inserm U1028/CNRS UMR5292/UCBL1  
Institut IDEE, 59 Boulevard Pinel, 69500 Bron

**Tuteurs:** LEUTENEGGER Anne-Louise, CR Inserm (Paris)  
[anne-louise.leutenegger@inserm.fr](mailto:anne-louise.leutenegger@inserm.fr) ; 01.40.03.19.70  
DELOUS Marion, CR Inserm (Lyon)  
[marion.delous@inserm.fr](mailto:marion.delous@inserm.fr) ; 04 72 13 89 69  
EDERY Patrick, PU-PH, HCL, Université Lyon1  
[patrick.edery@chu-lyon.fr](mailto:patrick.edery@chu-lyon.fr) ; 04 72 12 96 98

**Description:**

Notre équipe s'intéresse aux causes génétiques de la scoliose idiopathique, maladie fréquente au sein de la population adulte (2-3%) et dont pourtant l'étiologie reste méconnue. L'étude de grandes familles nous a permis d'identifier par analyse de liaison différents loci (Ederly et al, EJHG 2011)<sup>1</sup>. Pour une de ces familles (F2), deux régions sur les chromosomes 3 et 5 co-ségrègent avec la maladie. Combinée au séquençage d'exome de 3 individus atteints de la famille F2, l'analyse a mené à l'identification d'un variant dans le gène *POC5* (sur chr.5), responsable de la scoliose idiopathique (Patten et al, JCI 2015)<sup>2</sup>.

Comme l'analyse de liaison génétique était également en faveur de la présence d'un variant délétère sur le chromosome 3 et que *POC5* ne permettait pas d'expliquer l'ensemble des phénotypes observés au sein de la famille F2 (atteints/non atteints), nous cherchons maintenant à identifier l'autre variant situé sur le chromosome 3. Notre hypothèse est que, dans cette famille, des variants génétiques dans les 2 régions chromosomiques (3 et 5) sont nécessaires pour développer la maladie, et que le variant du chromosome 3 est en dehors des exons. De fait, nous avons réalisé le séquençage du génome entier des individus clés de la famille, et l'analyse des données a porté en premier lieu sur les variants de nucléotide simple (single nucleotide variants, SNVs) et insertions/délétions courtes (indels).

**La.e stagiaire sera responsable de la poursuite de l'analyse bioinformatique des génomes entiers** de la famille pour identifier les variants pathogènes candidats dans la région de liaison du chr.3 (~18Mb). Il s'agira de finaliser l'étude des SNVs et indels, puis dans un deuxième temps, de rechercher des variations de structure ou d'épissage non canonique.

En fonction du temps et de ses intérêts, la.e stagiaire pourra également travailler sur des projets plus en amont (analyse de liaison, ou sélection des individus en vue de séquençage), lui permettant ainsi de gagner en expérience sur les différentes étapes de l'analyse de données génétiques en présence de données familiales.

**Environnement :**

A Paris, une équipe ayant l'habitude de l'analyse de données de séquençage (1 chercheur, 1 post-doc bioinformaticien). A Lyon, une équipe connaissant bien les pathologies et la validation fonctionnelle des variants identifiés (1 médecin-chercheur et 1 chercheur).

**Compétences:**

Connaitre un langage de programmation (par exemple Perl, Python, C/C++) et être à l'aise sous Linux. Avoir été exposé.e aux données de séquençage haut-débit (exome ou génome) et à leurs outils d'analyse serait un plus (par exemple GATK best practices, score CADD, annotations VEP, visualisation IGV). Curiosité pour l'analyse de données familiales.

---

<sup>1</sup> Ederly P, et al. New disease gene location and high genetic heterogeneity in idiopathic scoliosis. Eur J Hum Genet. 2011, 19(8):865-9.

<sup>2</sup>Patten SA et al. Functional variants of *POC5* identified in patients with idiopathic scoliosis. J Clin Invest. 2015, 125(3):1124-8.