

Proposition de stage

Parcours Master 2 « Microbiologie, Environnement, Santé »

1. Laboratoire / Entreprise d'accueil : UMR 7205, Equipe BIPEM – ISYEB - MNHN

Intitulé : Equipe Biologie intégrative des populations et évolution moléculaire

Adresse : 16 rue Buffon, 75005 Paris

Responsable du Laboratoire / Entreprise : Thierry Wirth

Responsable de l'encadrement : Thierry Wirth (DE1) - HDR

Téléphone : 06 48 15 53 20

E-mail : wirth@mnhn.fr

Co-encadrant éventuel : Jean-Philippe Rasigade (Hospices Civils de Lyon, Institut des Agents Infectieux, Centre International de Recherche en Infectiologie).

2. Description du stage (2 pages maximum) :

Titre : Etude d'association pangénomique du succès épidémique de souches de *Mycobacterium tuberculosis* multirésistantes impliquées dans des épidémies à large échelle.

Mots clés : Etude d'association génomique, tuberculose, *Mycobacterium tuberculosis*, évolution, sélection

Contexte et objectifs généraux :

Mycobacterium tuberculosis est l'agent pathogène humain le plus meurtrier, devançant le paludisme et le VIH. On estime qu'il infecte un tiers de la population mondiale, causant 9 millions de nouveaux cas de tuberculose et près de 2 millions de décès par an. La dynamique de transmission de la tuberculose implique des interactions complexes de facteurs environnementaux, socio-économiques et génétiques. Cependant, en ce qui concerne le volet génétique, la vaste majorité des publications se concentrent sur l'étude de mutations ponctuelles qui confèrent une résistance à un antibiotique, négligeant dans une certaine mesure d'autres mutations tout aussi importantes comme par exemple celles liées au temps de génération, au taux de mutation ou à l'échappement du système immunitaire.

Au sein de notre équipe nous avons développé un outil mathématique et statistique qui nous permet d'inférer de manière quantitative la réussite épidémique souche-spécifique d'une bactérie. L'indice retenu (THD) repose sur les distances génétiques et il peut varier significativement entre les phases épidémiques et endémiques. Grâce à cet outil, nous sommes en mesure de déduire les facteurs de succès épidémique en nous basant sur l'analyse conjointe de données génétiques, épidémiologiques et cliniques.

Une des premières applications du THD vise donc à mettre en place une analyse de type GWAS, en prenant comme trait quantitatif les valeurs du THD inférées à partir de très gros jeux de données génomiques disponibles au sein de l'équipe. Cette approche est novatrice et devrait permettre de détecter des mutations convergentes entre différents outbreaks ou du moins touchant des familles de gènes voisines. La méthodologie disposera également d'un contrôle positif puisque nous avons en nos mains les données ancillaires de profil de résistance aux antibiotiques de toutes les souches considérées.

Projet de stage : L'objectif du projet proposé est de caractériser la base génétique du succès épidémique des souches, en s'éloignant ainsi de l'approche classique et monolithique centrée sur les seuls profils de résistance aux antibiotiques. La finalité ultime étant d'investiguer avec une plus grande acuité le paysage adaptatif de souches à succès avéré.

Les objectifs de ce stage M2 sont :

Pendant le stage, le stagiaire devra:

1) Appliquer l'indice THD (une approximation de l'aptitude épidémique de la souche) à trois grands ensembles de données indépendants couvrant plus de 1 000 génomes de *Mycobacterium tuberculosis*, liés aux principales épidémies récentes de tuberculose MDR en Asie centrale et en Europe occidentale ;

2) Développer une étude quantitative d'association pangénomique afin de détecter les SNPs associés à des scores THD élevés. Les SNPs bien connus impliqués dans la résistance aux antibiotiques seront utilisés comme contrôles positifs afin d'optimiser la méthodologie.

Nous anticipons la détection de corrélations positives avec des gènes impliqués dans la réponse immunitaire de l'hôte, la virulence, la réparation de l'ADN et les systèmes ESX. Enfin, nous espérons découvrir des traits adaptatifs spécifiques à certaines lignées, mais aussi des traits génétiques partagés, convergents, lors d'épidémies indépendantes.

Le stage est essentiellement bioinformatique, une appétence pour ce genre d'analyses est donc cardinal tout comme une solide base sous R.

Bibliographie :

Collins C & Didelot X (2018) A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. *PloS Computational Biology* 14: e1005958.

Farhat M et al. (2019) GWAS for quantitative resistance phenotypes in *Mycobacterium tuberculosis* reveals resistance genes and regulatory regions. *Nature Communications* 10: 2128.

Jaillar M et al. (2018) A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. *PloS Genetics* 14: e1007758.

Sélection d'autres publications de l'équipe sur le sujet:

Rasigade JP, Barbier M, Dumitrescu O, Pichat C, Carret G, Ronnaux-Baron AS, Blasquez G, Godin-Benham C, Boisset S, Carricajo A, Jacomo V, Fredenucci I, Pérouse de Montclos M, Flandrois JP, Ader F, Supply P, Lina G & Wirth T (2017) Strain-specific estimation of epidemic success provides insights into the transmission dynamics of tuberculosis. *Scientific Reports* 7: 45326.

Wirth T (2015). Massive lineage replacements and cryptic outbreaks of *Salmonella* Typhi in Eastern and South Africa. *Nature Genetics* 47: 565-567.

Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, Blum M, Rüscher-Gerdes S, Mokrousov I, Aleksic E, Allix-Béguec C, Antierens A, Augustynowicz-Kopeć E, Ballif M, Barletta F, Beck HP, Barry III CE, Bonnet M, Borroni E, Campos-Herrero I, Cirillo D, Cox H, Crowe S, Crudu V, Diel R, Drobniewski F, Fauville-Dufaux M, Gagneux S, Ghebremichael S, Hanekom M, Hoffner S, Jiao WW, Kalon S, Kohl TA, Kontsevaya I, Lillebæk T, Maeda S, Nikolayevskyy V, Rasmussen M, Rastogi N, Samper S, Sanchez-Padilla E, Savic B, Chola Shamputa I, Shen A, Sng LH, Stakenas P, Toit T, Varaine F, Vukovic D, Wahl C, Warren R, Supply P, Niemann S & Wirth T (2015) Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature Genetics* 47: 242-249.

Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, Wirth T, Jaenicke S, Schuback S, Rüscher-Gerdes S, Supply P, Kalinowski J & Niemann S (2013) Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: A longitudinal molecular epidemiological study. *PloS Medicine* 10: e1001387.

Ce stage peut-il se poursuivre par une thèse ? : Oui (dans le cadre du concours de l'ED de l'EPHE).

