

Ecole doctorale Environnements-Santé
Dossier de projet de thèse « Contrat doctoral Etablissements »
ANNEE 2025

Ce projet de thèse concerne-t-il le sous-jury environnement oui non
Ce projet de thèse concerne-t-il le sous-jury santé oui non

Le/ les demandants ont-ils déposés un sujet à la région en 2025 ? oui non
Rappel : Un même directeur ou codirecteur HDR de thèse ne pourra pas prétendre la même année à deux contrats doctoraux « Etablissement » ou « Région » (article 3 du RI de l'ED E-S).

TITRE DU PROJET : Génomique comparative et phylogénomique des clones à haut risque épidémique du pathogène bactérien *Pseudomonas aeruginosa* pour la caractérisation génétique de leur succès.

(Une fois votre projet accepté, si vous faites une mise en ligne par vos propres moyens, merci de ne pas changer ce titre ou le simplifier. Lors de la mise en ligne faite par l'ED, le titre ci-dessus est utilisé. Si vous en changez, cela entraîne un doublon de projet)

Ce projet est aussi proposé au titre de la bourse générale FRM oui¹ non

1) Renseignements administratifs sur la direction de thèse² (1 page maximum) :

Directeur de thèse HDR :

Nom : Valot

Prénom : Benoit

Section CNU : 64

Grade : *Ingénieur de Recherche*

HDR : Date de soutenance : 19/11/2020 Discipline : *Biochimie et Biologie Moléculaire*

l'HDR devra être soutenue, ou sa soutenance autorisée, au moment du dépôt du présent projet.

Coordonnées (adresse, courriel, téléphone) :

- UFR Santé, Haut du Chazal, 19 rue Ambroise Paré, 25030 Besançon.

- benoit.valot@univ-fcomte.fr

- 03.63.08.22.42

Unité d'appartenance (intitulé, label, n°, directeur) : UMR CNRS 6249 Chrono-Environnement, Émilie Gauthier

2) Descriptif du projet de thèse (devra inclure les rubriques suivantes) :

- nom et label de l'unité de recherche (ainsi que l'équipe interne s'il y a lieu)
- localisation
- nom du directeur de thèse et du co-directeur s'il y a lieu
- adresse courriel du contact scientifique

¹ Le candidat doit être connu au moment du dépôt du projet, et son CV doit être joint à la présente demande. Le CV du candidat sera l'un des critères de sélection.

² ATTENTION : depuis le texte de loi de mai 2016, le total d'encadrants ne peut pas dépasser 2, sauf si l'un des encadrants appartient au monde économique, qui peut venir en supplément, ou en cas de co-tutelle; Le décompte des co-encadrements se fera au prorata du nombre d'encadrants : 1 pour 1 encadrant, ½ pour deux encadrants.

- titre du projet
- description du projet (2 pages maximum)
- Financement du projet – partie Recherche (montants acquis, type de contrat)
- connaissances et compétences requises

TITRE DU PROJET : Génomique comparative et phylogénomique des clones à haut risque épidémique du pathogène bactérien *Pseudomonas aeruginosa* pour la caractérisation génétique de leur succès.

Laboratoire : UMR CNRS 6249 Chrono-Environnement ; Thème Agents Pathogènes

Adresse : UMR CNRS 6249 Chrono-Environnement, UFR Santé, Haut de Chazal, 19 rue Ambroise Paré, 25030 Besançon

Encadrement : Directeur de thèse : Benoit Valot (benoit.valot@univ-fcomte.fr)

Description du Projet :

Pseudomonas aeruginosa est une bactérie à Gram négatif opportuniste impliquée particulièrement dans des infections pulmonaires et dans des bactériémies de patients immunodéprimés. La mortalité de ces infections atteint 30% chez les patients fragiles. Naturellement présente dans l'environnement, notamment hydrique, cette espèce bactérienne se distingue par un génome de grande taille (6-7 Mb) lui permettant de coloniser des milieux biotiques et abiotiques variés.

Il est admis que cette espèce a une structure de population non clonale épidémique. Cela signifie que la population est composée d'un nombre limité de clones très répandus qui sont sélectionnés à partir d'un grand nombre de clones rares et non reliés qui recombinent fortement (Pirnay *et al.* PLoS One 2009). Ces clones très répandus sont aussi appelés clones 'à haut risque épidémique' (HRE). Ils sont souvent à l'origine d'épidémies hospitalières et porteur de résistances multiples aux antibiotiques. Ces clones HRE sont identifiés par génotypage en utilisant la méthode *Multi-Locus Sequence Typing* (MLST).

Parmi ces clones à HRE, le *Sequence Type* 235 (ST235) est le plus répandu avec une distribution mondiale et une acquisition de nombreux gènes codant pour des mécanismes de résistances aux β -lactamines. Particulièrement virulent, il est associé à une plus grande mortalité des patients infectés. Lors d'une thèse précédente au laboratoire (Panisa Treepong, 2015-2018), nous avons analysé l'histoire évolutive de ce clone dans lequel le gène *dprA* était enrichi. Ce gène pouvait faciliter l'acquisition de matériel génétique par transfert horizontal, incluant ceux codant des mécanismes de résistance (Treepong *et al.* Clin Microbiol Infect 2018). De plus, une autre étude a également montré une faible présence de systèmes immunitaire bactérien CRISPR-cas dans les génomes de ce clone HRE ST235, pouvant faciliter l'acquisition des gènes étrangers (van Belkum A *et al.* mBio 2015). De manière identique, le clone HRE ST111 présente une distribution mondiale et des acquisitions variées de résistance par transfert horizontal. A l'inverse, le clone HRE ST175 est principalement retrouvé en Europe et la résistance est plus fréquemment associée à des mutations chromosomiques. D'autres clones HRE secondaires sont également décrits comme le ST308, ST244 ou le ST395 (Oliver *et al.* Drug Resist Updat 2015).

Une récente publication (Del Barrion-Tofiño *et al.* Int J Antimicrob Agents 2020) a décrit la distribution et les caractéristiques cliniques de 10 clones à HRE (ST235, ST111, ST233, ST244, ST357, ST308, ST175, ST277, ST654 et ST298). Dans cette étude, les auteurs soulignaient le

manque de connaissance des facteurs génétiques associés au succès de ces clones ainsi qu'à leur stratégie adaptative et évolutive différente. Dans ce contexte, l'objectif de cette thèse est d'identifier ces facteurs génétiques associés aux clones à HRE de *P. aeruginosa*.

Dans un premier temps, le(la) doctorant(e) regroupera l'ensemble des données génomiques pour chacun de ces clones à HRE ainsi qu'un panel de souches non épidémiques à partir des bases de données publiques (*e.g.*, NCBI, Patric). Les bases de données contenant beaucoup de redondances (multiples séquences d'un même clone lors d'épidémies) et des séquences de qualité hétérogène, des déréplications et nettoyage seront nécessaires.

Les génomes déposés sont principalement ceux de souches isolées de prélèvements cliniques. Ces données seront complétées avec des génomes de souches isolées d'environnements hospitaliers (lors de la thèse de Charlotte Couchoud 2018-2022 ; Couchoud *et al.* J Hosp Infect 2023) et d'environnements à différents degrés d'anthropisation (rivières, eaux de captage, eaux de boisson, égouts). Les souches dont les génomes ne sont pas encore disponibles seront séquencées durant la thèse. Toutes ces souches environnementales sont conservées au Centre de Ressource Biologique – Filière Microbiologique de Besançon (CRB-FMB, biobanque BB-0033-00090, certifiée norme NF S96-900, plateforme labélisée BFC). Au final, le(la) doctorant(e) travaillera sur environ 10 000 génomes de *P. aeruginosa*.

Dans un second temps, le(la) doctorant(e) réalisera des approches de phylogénomique et de génomique comparative à l'échelle de chaque clone. Celles-ci détermineront la structure de population de chaque clone, et identifieront les mécanismes de résistance (transfert horizontal et/ou mutation conférant la résistance) et de virulence. De plus, la recherche d'enrichissement de gènes, d'opérons ou de groupes fonctionnels sera réalisée afin de mettre en évidence le génome accessoire potentiellement impliqué dans le succès de ces clones à HRE.

En parallèle des approches de génomiques, des analyses transcriptomiques dans des conditions de stress abiotique seront réalisées pour mettre en évidence des différences de réponses entre des clones à HRE contrastés dans leur structure populationnelle et leur mécanisme évolutif.

Enfin, les fonctions biologiques conférées par les facteurs génétiques identifiés lors des analyses précédentes seront évaluées phénotypiquement sur ces clones et sur des mutants pour valider les hypothèses décrites.

L'identification de fonctions biologiques communes aux clones à HRE permettra de mieux comprendre leur nature épidémique et d'identifier de nouvelles cibles thérapeutiques et épidémiologiques.

Financement :

Travail de thèse essentiellement en bio-informatique, nécessitant peu de financement.

- Capacité de calcul : Serveur dédié 2B2S situé à et financé par l'UFR SMP et Mésocentre de calcul (UFR ST, accès pris en charge par l'UMR Chrono-environnement)
- Séquençage WGS de souches conservées au CRB-FMB demandées lors de la première année (~10 k€, CRB-FMB, CHU de Besançon)
- Analyses RNASeq (~ 10k€, CRB-FMB, CHU de Besançon) demandé en deuxième année.

Connaissances et compétences acquises :

- Analyse bioinformatique de génomes procaryotes
- Phylogénomique et génomique comparative
- Analyse différentielle de données RNASeq
- Tests statistiques associés

Résumé en français (limité chacun à 1800 caractères)

Pseudomonas aeruginosa, une bactérie à Gram négatif, est un pathogène opportuniste pour l'homme. Naturellement résistante à certains antibiotiques, elle peut également acquérir de nouveaux mécanismes de résistance par mutation ou acquisition de gènes extrinsèques par transfert horizontal. On parle alors de clones multi-résistant (MDR) or extrêmement résistante (XDR) aux antibiotiques. La structure de population de *P. aeruginosa* est non clonale épidémique. En d'autres termes, quelques clones prédominants sont sélectionnés d'un fond abondant de souches rares. Ces clones sont nommés à haut risque épidémique (HRE).

L'objectif de cette thèse est de mieux caractériser la structure évolutive de ces clones à HRE et leur contenu génétique en comparaison à un panel de souches non épidémiques. De plus, une approche expérimentale par RNASeq déterminera les différences de réponses de ces clones à HRE à un stress abiotique.

Pour ce faire, le(a) doctorant(e) utilisera les génomes déjà présents dans les bases de données (issues principalement de prélèvements cliniques), complétés par des génomes de souches environnementales conservées au CHU de Besançon.

Enfin, les fonctions biologiques enrichies dans les clones à HRE seront évaluées phénotypiquement sur ces clones et sur des mutants K-O pour valider les hypothèses décrites. L'identification de fonctions biologiques communes aux clones à HRE permettra de mieux comprendre leur nature épidémique et d'identifier de nouvelles cibles thérapeutiques et épidémiologiques.

Le(a) candidat(e) devra être à l'aise avec les outils informatique d'analyse de séquences et avoir des connaissances dans la génomique microbienne.

Résumé en anglais

Pseudomonas aeruginosa, a Gram-negative bacterium, is an opportunistic pathogen for human. It is naturally resistant to certain antibiotics, but can also acquire new resistance mechanisms through mutation or acquisition of extrinsic genes by horizontal transfer. These are known as multidrug resistant (MDR) or extremely resistant (XDR) clones. The population structure of *P. aeruginosa* is non-clonal epidemic. In other words, a few predominant clones are selected from an abundant background of rare strains. These clones are called 'epidemic high-risk clones' (EHR).

The aim of this PhD work is to better characterize the evolutionary structure of these EHR clones and their genetic content compared to a panel of non-epidemic strains. In addition, an experimental RNA-Seq approach will be used to determine differences in the responses of these EHR clones to abiotic stress.

To this end, the PhD student will use the genomes already available in the databases (mainly originating from clinical samples), complemented by the genomes of environmental strains kept at the University Hospital of Besançon.

Finally, the biological functions enriched in the ERH clones will be evaluated phenotypically on these clones and on K-O mutants to validate the hypotheses described. The identification of biological functions common to the ERH clones will allow a better understanding of their epidemic nature and the identification of new therapeutic and epidemiologic targets.

The candidate should be familiar with computational tools for sequence analysis and have a background in microbial genomics.

Préciser le domaine de compétence dans la liste ci-dessous (2 choix possibles maximum – ne pas modifier les intitulés : ils sont imposés par certains sites web) :

Informatique, électronique

Santé, médecine humaine, vétérinaire

Mots clés :

bio-informatique, génomique microbienne, clones à haut risque épidémique