

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **18 décembre 2024**

Nom de famille et prénom de l'auteur. e : **Madame Sara MARTINS RIBEIRO**

Titre de la thèse : Analyse de données à grande échelle pour découvrir le talon d'Achille microbien

Résumé



La montée mondiale de la résistance aux antibiotiques, parallèlement à l'émergence de nouveaux agents pathogènes bactériens, met en évidence la nécessité pressante d'outils diagnostiques sophistiqués et de nouvelles stratégies thérapeutiques. L'augmentation massive des données génomiques, en particulier celles concernant les génomes bactériens, révèle que les méthodologies traditionnelles basées sur les séquences sont souvent insuffisantes pour aborder la complexité de la pathogénicité bactérienne. De plus, les études sur la pathogénicité bactérienne se concentrent généralement sur un nombre restreint d'espèces ou reposent sur des hypothèses excessivement simplistes pour annoter la pathogénicité. Ces facteurs compliquent l'identification de nouveaux déterminants de la pathogénicité à large échelle. Ma thèse se concentre sur l'analyse de données à grande échelle visant à surmonter ces défis et à identifier de nouveaux déterminants de la pathogénicité bactérienne susceptibles d'améliorer à la fois les méthodes diagnostiques et les traitements antimicrobiens. Pour soutenir cet objectif, nous avons d'abord développé une base de données sur la pathogénicité des souches bactériennes (BacSPaD), afin de fournir des annotations de pathogénicité au niveau des souches à grande échelle. Cela a impliqué l'intégration et la curation minutieuse de métadonnées génomiques provenant de sources fiables, combinant un appariement automatique de mots-clés avec une revue approfondie de la littérature. Un total de 5,992 génomes bactériens de haute qualité couvrant 532 espèces ont ainsi été annotés, créant une ressource importante pour comprendre les variations de virulence spécifiques aux souches. En utilisant cette ressource curée, nous avons ensuite effectué une analyse d'orthologie basée sur la phylogénie afin de comparer les protéomes des souches bactériennes pathogènes pour l'homme (HP) et non pathogènes pour l'homme (NHP) à travers 514 espèces. Grâce à un flux de travail bio-informatique adapté, nous avons identifié 4,383 groupes orthologues hiérarchiques (HOGs) significativement associés à la pathogénicité. Ces HOGs étaient liés à la tolérance au stress, à la flexibilité métabolique et à la résistance aux antibiotiques. De manière significative, notre analyse a non seulement confirmé des facteurs de virulence déjà reconnus, mais a également découvert de nouvelles protéines susceptibles de représenter des cibles prometteuses pour des interventions thérapeutiques ou des applications diagnostiques. En intégrant des données génomiques vastes à des techniques d'analyse innovantes, cette recherche pourrait ouvrir la voie à de nouvelles stratégies dans la lutte contre les infections bactériennes et la résistance aux antibiotiques.

Mots-clés : Bactériologie,Bioinformatique,Protéomique,