

Avis de Soutenance

Madame REFKA BEN HAMOUDA

Biologie-Santé - Spécialité Bioinformatique et Génomique

Soutiendra publiquement ses travaux de thèse intitulés

Développement d'un outil bioinformatique innovant pour l'analyse des spectres MS MALDI-TOF d'arthropodes

dirigés par Monsieur Lionel ALMERAS

Soutenance prévue le **vendredi 29 novembre 2024** à 9h00

Lieu : 19-21 Bd Jean Moulin, 13005 Marseille

Salle : amphithéâtre IHU

Composition du jury proposé

M. Lionel ALMERAS	Unité de Parasitologie et Entomologie (U.P.E.) Institut de Recherche Biomédicale des Armées (I.R.B.A.) UMR RITMES – Risques Infectieux Tropicaux et Microorganismes Emergents, Aix Marseille Université, Institut Hospitalo-Universitaire – Méditerranée Infection (IHU-MI)	Directeur de thèse
Mme Nadia CREIGNOU	Logique, interaction, raisonnement et inférence, complexité et applications (LYRICA)	Examinatrice
M. Oussama SOUIAI	Institut Pasteur de Tunis	Rapporteur
M. Christophe BRULEY	Laboratoire d'Etude de la Dynamique des Protéomes (EDyP) / UA13 Inserm/CEA/UGA – Biosciences et Bioingénierie pour la sante (BGE) – CEA Grenoble	Rapporteur
M. Anthony LEVASSEUR	Equipe 7 UMR D-258, Microbes, Evolution, Phylogénie et Infection (MEPHI) Aix-Marseille Université (AMU), Institut de Recherche pour le Développement (IRD)	Président
M. Ameer CHERIF	University of Manoub - Laboratoire de Biotechnologie et de Valorisation des Bio-Géo Ressources	Examineur
M. Bertrand ESTELLON	Laboratoire d'Informatique et Systèmes, AMU	Invité

Mots-clés : Bio-informatique, Package R, Spectrométrie de masse, Empreintes protéiques, Arthropodes, Base de données

Résumé :

Au cours de la dernière décennie, le profilage MALDI-TOF MS est apparu comme une approche innovante et pertinente pour l'identification des arthropodes et la détection de certains traits de vie. Malgré le succès de cette nouvelle méthode d'analyse des arthropodes, son application dans le domaine de l'entomologie médicale reste confidentielle. Parmi les facteurs qui entravent l'utilisation généralisée de cette approche, on peut citer les restrictions des logiciels commerciaux dédiés à

l'analyse des spectres MS, qui ne permettent pas d'étendre de nouvelles fonctionnalités, et l'absence de base de données de spectres de référence des arthropodes. Pour surmonter ces limitations, le présent projet visait à créer un outil bioinformatique plus rapide, fiable, évolutif et innovant pour identifier plusieurs traits de vie des moustiques en utilisant la stratégie de MALDI-TOF profiling. Les objectifs spécifiques étaient de i) développer un logiciel en accès libre dédié à l'analyse de spectres MS complexes, ii) créer une base de données de spectres organisés pour une identification rapide des espèces d'arthropodes, de leur préférences trophiques, de leur statut d'infection et de leur sensibilité aux insecticides, et iii) valider l'outil à l'aide d'échantillons d'arthropodes d'origines multiples. L'outil bioinformatique a été organisé en étapes successives, permettant un développement et des tests indépendants avant inclusion dans la version finale. Cet outil a été construit à l'aide de l'environnement de programmation R, appelant des packages spécifiques pour le traitement des spectres, avec une interface implémentée via R shiny. L'outil est disponible sur Microsoft Windows, accessible à l'aide du lien suivant sur GitHub (<https://github.com/Almeras-Lionel/MSPProfileR>). Pour pallier l'absence de bases de données, une base de données de référence pour les spectres d'arthropodes a été développée en utilisant le système SQL (Structured Query Language) pour sa compatibilité avec R, facilitée par la bibliothèque RMySQL. Cette configuration offre de vastes possibilités de structuration et d'organisation de la base de données et propose de nombreuses fonctions d'interrogation de la base de données. La base de données a également été développée avec R Shiny, offrant une interface web conviviale pour ajouter/supprimer des données ou pour des requêtes filtrées. La base de données de référence contient actuellement un total de 1927 échantillons correspondant à 7451 spectres de 7 familles d'arthropodes et 102 espèces distinctes. Cependant, la liste est encore extensible, représentant une archive de données d'arthropodes validées dans le but d'améliorer l'identification des espèces à l'aide d'algorithmes de mise en correspondance, l'une des futures directions de ce projet. Enfin, MSPProfileR a été validé à l'aide de deux jeux de données, comprenant exclusivement des spectres MS d'arthropodes. Ce logiciel open source avec une interface Shiny peut être utilisé par n'importe qui sans nécessiter de compétences en programmation. Le langage de programmation R permet des modifications, permettant l'évolution continue de l'outil et l'extension de ses fonctionnalités. Collectivement, le développement du package MSPProfileR sera décisif pour une adoption plus large de l'outil de profilage MS au sein de la communauté scientifique, en particulier parmi les groupes de recherche en entomologie.

LE DOVEN

Georges LEONETTI