

## SUJET DE THÈSE AU LEESU ET AU SIAAP

### Relier la pression chimique à ses effets sur la biodiversité dans les milieux aquatiques en couplant l'analyse non ciblée avec l'ADN environnemental à l'aide d'outils numériques : exemple de l'antibiorésistance

#### Allocation doctorale financée par le programme Paris région PhD 2024

#### Contexte

La qualité des milieux aquatiques et de la ressource en eau peut être appréhendée du point de vue de sa contamination chimique, en particulier par des contaminants émergents, ou de la biodiversité qu'elle accueille. La présence de contaminants (ex. pesticides, produits pharmaceutiques ou industriels), en grande partie rejetés avec les effluents traités de station d'épuration, est néfaste pour la santé des écosystèmes avec des effets nocifs sur les organismes aquatiques et une perte de biodiversité (Santos *et al.* 2010; Hamilton *et al.* 2016; Reid *et al.* 2019; Alderton *et al.* 2021), et pour la santé humaine avec une mauvaise qualité des ressources pour la production d'eau potable et l'augmentation des risques d'antibiorésistance. Afin de caractériser cette contamination chimique, des méthodes d'analyse non ciblée par spectrométrie de masse haute résolution (HRMS) ont été développées (Aurich *et al.* 2023; Hollender *et al.* 2023). Au Leesu, deux thèses ont été menées sur cette thématique. Une première a permis de développer la méthode analytique pour optimiser l'extraction et la détection d'une quantité importante de molécules organiques, et d'appliquer cette méthode à l'analyse des polluants dans différentes eaux urbaines, en particulier dans les filières de traitement des eaux usées (Huynh *et al.* 2021). La deuxième thèse a visé à développer des outils informatiques avancés de traitement des données (Sade *et al.* 2022). En parallèle, l'étude de l'ADN environnemental (ADNe), dans lequel les signaux des macro- et micro-organismes de l'environnement sont récupérés à partir d'échantillons d'eau est de plus en plus utilisée pour la surveillance de la biodiversité des milieux (Duarte *et al.* 2021; Altermatt *et al.* 2023) et de la diffusion des bactéries et gènes d'antibiorésistance (ARB/ARG). De telles méthodes par analyse de métagénomique ont été utilisées au Leesu dans une thèse récente, et ont permis par exemple de caractériser avec précision les différentes communautés bactériennes présentes dans la Seine et les dynamiques de ces communautés soumises à différentes pressions anthropiques (ex. rejets urbains, traitements de désinfection des eaux usées) et aux variations saisonnières/climatiques (Bagagnan 2024).

Les progrès techniques qui permettent aujourd'hui d'enregistrer des empreintes chimiques par des approches non ciblées ou de mesurer l'ADN environnemental pour étudier la biodiversité, fournissent d'énormes quantités de données. Le défi consiste à mettre en place les outils d'analyse biologique et chimique multivariée pour faciliter l'utilisation de ces données (Slobodnik *et al.* 2019). L'utilisation d'outils numériques comme l'apprentissage machine se développe pour traiter et interpréter les spectres HRMS de contaminants dans les eaux (Sade *et al.* 2022; Arturi and Hollender 2023; Hollender *et al.* 2023), de même que dans le domaine de l'ADN environnemental pour extraire des indicateurs de biodiversité ou expliquer les variabilités spatio-temporelles (Cordier *et al.* 2018; Hu *et al.* 2023).

Cependant, pour l'instant **ces deux approches HRMS et ADNe n'ont pas été couplées**. Or les mélanges complexes de produits chimiques doivent être pris en compte en même temps que leurs effets complexes et leurs impacts sur les écosystèmes. Le couplage entre l'empreinte chimique des contaminants et l'ADNe qui mesure la biodiversité est un enjeu majeur pour la surveillance de la qualité des milieux récepteurs.

## Objectifs et résultats attendus

L'**objectif général** de ce projet de thèse est d'évaluer la possibilité **d'interpréter conjointement les empreintes chimiques HRMS de contaminants** dans les eaux (des eaux usées brutes aux rejets traités dans l'environnement) avec les **informations sur la biodiversité données par l'ADNe** en développant et appliquant des **méthodes numériques**.

Plus spécifiquement, il s'agira de :

- Identifier pour chaque type de données (HRMS et ADNe) des traceurs ou indicateurs d'intérêt à partir de la littérature scientifique et automatiser leur traitement (comme les ARB/ARG, les biocides, les antibiotiques...)
- Appliquer des méthodes de traitement des données pour coupler les données HRMS et d'ADNe en utilisant des outils statistiques avancés ou numériques de type machine learning.
- Proposer une stratégie d'échantillonnage pertinente pour acquérir les deux types de données (HRMS et ADNe) dans les filières de traitement des eaux usées du SIAAP et effectuer un suivi régulier des différents points d'échantillonnage.
- Suivre et interpréter le devenir temporel de molécules/marqueurs d'intérêt identifiés précédemment (suivi à différentes fréquences : saisonnières, hebdomadaires, journalières et le long des différents points de prélèvement)

## Cadre et partenaires du projet

Cette thèse s'insère dans les actions de recherche du Leesu proposées dans le cadre de la phase 6 du programme OPUR, en collaboration étroite avec le Service public de l'assainissement francilien (SIAAP) dont un des objectifs est d'améliorer la compréhension du devenir des contaminants dans les filières de traitement des eaux et dans les rejets de stations d'épuration. Les analyses de contaminants émergents s'appuient sur les instruments analytiques de la plateforme Prammics (OSU Efluve), en particulier les instruments de chromatographie en phase liquide (Waters Vion – UPLC-IMS-QTOF et Shimadzu HPLC Fraction collector). L'UPLC-IMS-QTOF est un instrument de HRMS équipé d'une séparation par mobilité ionique (IMS). L'ADNe sera séquencé par un prestataire par des méthodes de métabarcoding ou métagénomique. Les séquences obtenues seront traitées par des outils bio-informatiques (*kaiju*, *eggNOG-mapper*, *BWA-MEM*, *rgi* et autres programmes bio-informatiques) permettant d'identifier taxonomiquement chaque groupe de macro ou micro-organismes séparément (bactéries, algues, champignons, poissons, etc.) et également les différentes fonctions/activités des microorganismes présents dans l'échantillon avec la possibilité de mettre en évidence les corrélations potentielles entre la présence de contaminants émergents et les groupes/espèces résistants/sensibles.

## Profil souhaité

M2 en bio-informatique. Le candidat devra posséder une expérience en analyses de données omiques, des compétences en statistiques, ainsi que des connaissances de base du logiciel R. Une bonne compréhension de la biologie générale, de l'écotoxicologie microbienne, et des problématiques liées aux données métagénomiques sera considérée comme un atout.

## Laboratoire d'accueil et encadrement

La thèse se déroulera au Laboratoire Eau Environnement et Systèmes Urbains (LEESU - Université Paris-Est Créteil et École des Ponts ParisTech), et au sein de la direction innovation du SIAAP. La thèse se déroulera à 20% au SIAAP pour les formations opérationnelles, les campagnes d'échantillonnage, les comités de pilotage du projet et les transferts de compétences opérationnelles.

- Adèle Bressy, chargée de recherche HDR de l'École des Ponts, chimie de l'environnement, hydrologie ([adele.bressy@enpc.fr](mailto:adele.bressy@enpc.fr))
- My Dung Jusselme, maîtresse de conférence UPEC, biologie moléculaire, microbiologie environnementale ([thi-my-dung.jusselme@u-pec.fr](mailto:thi-my-dung.jusselme@u-pec.fr))
- Julien Le Roux, maître de conférences UPEC, chimie analytique, outils statistiques et numériques, chimie de l'environnement ([julien.le-roux@u-pec.fr](mailto:julien.le-roux@u-pec.fr))
- Sabrina Guérin, responsable innovation au SIAAP, et Anthony Marconi, Responsable Service Rivière-Usine Métrologie Innovante

## Références bibliographiques

- Alderton I, Palmer BR, Heinemann JA, Pattis I, Weaver L, Gutiérrez-Ginés MJ, Horswell J, Tremblay LA (2021). The role of emerging organic contaminants in the development of antimicrobial resistance. *Emerging Contaminants* **7**, 160–171. doi:10.1016/j.emcon.2021.07.001
- Altermatt F, Carraro L, Antonetti M, Albouy C, Zhang Y, Lyet A, Zhang X, Pellissier L (2023). Quantifying biodiversity using eDNA from water bodies: General principles and recommendations for sampling designs. *Environmental DNA* **5**, 671–682. doi:10.1002/edn3.430
- Arturi K, Hollender J (2023). Machine Learning-Based Hazard-Driven Prioritization of Features in Nontarget Screening of Environmental High-Resolution Mass Spectrometry Data. *Environmental Science & Technology* **57**, 18067–18079. doi:10.1021/acs.est.3c00304
- Aurich D, Diderich P, Helmus R, Schymanski EL (2023). Non-target screening of surface water samples to identify exposome-related pollutants: a case study from Luxembourg. *Environmental Sciences Europe* **35**. doi:10.1186/s12302-023-00805-5
- Bagagnan S (2024). Désinfection des rejets de stations d'épuration par l'acide performique et suivi des communautés microbiennes du milieu récepteur en agglomération parisienne. Paris 12. Available at: <https://www.theses.fr/s233944> [accessed 12 March 2024]
- Cordier T, Forster D, Dufresne Y, Martins CIM, Stoeck T, Pawlowski J (2018). Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Molecular Ecology Resources* **18**, 1381–1391. doi:10.1111/1755-0998.12926
- Duarte S, Leite BR, Feio MJ, Costa FO, Filipe AF (2021). Integration of DNA-based approaches in aquatic ecological assessment using benthic macroinvertebrates. *Water (Switzerland)* **13**. doi:10.3390/w13030331
- Hamilton PB, Cowx IG, Oleksiak MF, Griffiths AM, Grahn M, Stevens JR, Carvalho GR, Nicol E, Tyler CR (2016). Population-level consequences for wild fish exposed to sublethal concentrations of chemicals – a critical review. *Fish and Fisheries* **17**, 545–566. doi:10.1111/faf.12125
- Hollender J, Schymanski EL, Ahrens L, Alygizakis N, Béen F, Bijlsma L, Brunner AM, Celma A, Fildier A, Fu Q, Gago-Ferrero P, Gil-Solsona R, Haglund P, Hansen M, Kaserzon S, Kruve A, Lamoree M, Margoum C, Meijer J, Merel S, Rauert C, Rostkowski P, Samanipour S, Schulze B, Schulze T, Singh RR, Slobodnik J, Steininger-Mairinger T, Thomaidis NS, Togola A, Vorkamp K, Vulliet E, Zhu L, Krauss M (2023). NORMAN guidance on suspect and non-target screening in environmental monitoring. *Environmental Sciences Europe* **35**. doi:10.1186/s12302-023-00779-4
- Hu H, Wei X-Y, Liu L, Wang Y-B, Jia H-J, Bu L-K, Pei D-S (2023). Supervised machine learning improves general applicability of eDNA metabarcoding for reservoir health monitoring. *Water Research* **246**. doi:10.1016/j.watres.2023.120686
- Huynh N, Caupos E, Soares Peirera C, Le Roux J, Bressy A, Moilleron R (2021). Evaluation of Sample Preparation Methods for Non-Target Screening of Organic Micropollutants in Urban Waters Using High-Resolution Mass Spectrometry. *Molecules* **26**, 7064. doi:10.3390/molecules26237064
- Reid AJ, Carlson AK, Creed IF, Eliason EJ, Gell PA, Johnson PTJ, Kidd KA, MacCormack TJ, Olden JD, Ormerod SJ, Smol JP, Taylor WW, Tockner K, Vermaire JC, Dudgeon D, Cooke SJ (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews* **94**, 849–873. doi:10.1111/brv.12480
- Sade J, Le Roux J, Mottelet S, Guérin-Rechdaoui S, Moilleron R (2022). Caractérisation de la contamination des eaux de surface par le couplage d'analyses non-ciblées en spectrométrie de masse avec des analyses d'écotoxicologie - Rapport d'avancement du projet MeSeine Innovation.
- Santos LHMLM, Araújo AN, Fachini A, Pena A, Delerue-Matos C, Montenegro MCBSM (2010). Ecotoxicological aspects related to the presence of pharmaceuticals in the aquatic environment. *Journal of Hazardous Materials* **175**, 45–95. doi:10.1016/j.jhazmat.2009.10.100
- Slobodnik J, Hollender J, Schulze T, Schymanski EL, Brack W (2019). Establish data infrastructure to compile and exchange environmental screening data on a European scale. *Environmental Sciences Europe* **31**, 65. doi:10.1186/s12302-019-0237-6