

Offre de stage M2

Exploitation de données multidimensionnelles longitudinales pour le monitoring avancé de bioprocédé

Contexte et enjeux

La digestion anaérobie, ou méthanisation, est un bioprocédé permettant de convertir la matière organique en biogaz, riche en méthane. Ce bioprocédé est utilisé de plus en plus largement à l'échelle industrielle pour la valorisation des déchets organiques en énergie renouvelable. Il repose sur l'activité d'une communauté microbienne complexe, dynamique, et sensible aux variations des conditions opératoires. Cette sensibilité peut avoir pour conséquence une instabilité du procédé, et des répercussions importantes pour les opérateurs. Afin d'optimiser la digestion anaérobie, une compréhension approfondie du fonctionnement des communautés microbiennes, en particulier face à des stress, est essentielle. Les méthodologies haut-débit omiques s'avèrent des outils particulièrement pertinents pour cela. La métagénomique, la métatranscriptomique, la métabolomique et la métataxonomique permettent en effet de caractériser une communauté microbienne à différents niveaux de son fonctionnement. Ces approches sont particulièrement intéressantes lorsqu'elles sont utilisées dans un contexte longitudinal, c'est-à-dire en analysant une suite temporelle d'échantillons plutôt que des échantillons uniques. Ainsi, il est possible de mieux saisir la dynamique de l'écosystème microbien suite à un changement de paramètre opératoire ou à un stress, et de proposer des solutions pour limiter les répercussions sur le procédé.

Cependant la plupart des méthodes statistiques utilisées pour analyser les données omiques ne tiennent pas compte explicitement de la temporalité, et traitent les différents points de temps comme des échantillons indépendants. Pour exploiter pleinement le potentiel des données longitudinales et prendre en compte le lien temporel entre les échantillons, la création de pipelines analytiques spécifiques est requise. Ces pipelines permettraient d'identifier les marqueurs biologiques associés à l'évolution des conditions opératoires et à l'évolution des performances des procédés. On pourrait chercher des liens entre les dynamiques temporelles des microorganismes et faire des hypothèses biologiques. On pourrait également envisager de déconvoluer l'effet des différents facteurs sur les dynamiques microbiennes et par exemple distinguer la dynamique naturelle de l'écosystème de l'effet spécifique d'un paramètre donné.

Objectif

Dans ce contexte, l'objectif du stage sera de mettre en place un pipeline analytique spécifique pour la valorisation des données omiques longitudinales issues de bioprocédés. Différents défis statistiques seront à résoudre (Kodikara et al. 2022). Des jeux de données omiques (métataxonomique, métagénomique, métabolomique) issues d'échantillons prélevés dans des bioréacteurs de laboratoire sont disponibles. Elles proviennent d'expériences réalisées dans l'unité PROSE pour évaluer les conséquences de différents stress salins sur les performances de la digestion anaérobie. On pourra s'appuyer sur des travaux préliminaires précédents (Bodein et al. 2019, Chapleur et al. 2021).

Bodein, A., Chapleur, O., Droit, A. and Lê Cao, K.-A. (2019) A Generic Multivariate Framework for the Integration of Microbiome Longitudinal Studies With Other Data Types. *Frontiers in Genetics* 10(963).

Chapleur, O., Poirier, S., Guenne, A. and Lê Cao, K.-A. (2021) Time-course analysis of metabolomic and microbial responses in anaerobic digesters exposed to ammonia. *Chemosphere* 283, 131309.

Kodikara, S., Ellul, S. and Lê Cao, K.-A. (2022) Statistical challenges in longitudinal microbiome data analysis. *Briefings in Bioinformatics* 23(4).

Compétences recherchées

- Connaissances solides en statistiques et en particulier en analyse de données,
- Aptitude à développer des scripts de traitement de données sous R,
- Capacités rédactionnelles et lecture d'articles scientifiques en anglais,
- Aptitudes au travail en équipe et à la communication.

Equipes et encadrement

Le stage se déroulera au sein de l'unité de recherche StatSC (Oniris VetAgroBio/INRAE) à Nantes.

Il sera co-encadré par l'unité de recherche PROSE (Procédés biotechnologiques au Service de l'Environnement) de l'INRAE à Jouy-en-Josas.

Il bénéficie d'un financement du département Transform de l'INRAE.

Le stage pourra démarrer en mars 2024, pour une durée de 5 à 6 mois.

Il sera rémunéré à hauteur d'environ 600€/mois.

Candidatures à envoyer à olivier.chapleur@inrae.fr et veronique.cariou@oniris-nantes.fr