Exploratory multivariate analysis of multi-omics data to decipher the mechanisms of anaerobic digestion inhibition

Internship location: **Irstea – Hydrosystems and Bioprocesses research unit (HBAN) - BIOMIC Team (Antony, 92, FRANCE)**

Duration: **approximately 6 months**

Supervision: **Dr O. Chapleur, microbial ecology Research scientist**
olivier.chapleur@irstea.fr, +33 (0) 1 40 96 65 06

Irstea is a multidisciplinary research institute with the aim of developing projects of high scientific level with applications. One of the main objectives of the BIOMIC team (Bioprocesses and microbial biotechnologies for waste valorisation) is to understand and optimise the exploitation of microbial ecosystems within bioprocesses for the treatment and valorisation of organic waste; and to set the basic principles for the ecological engineering of microbial bioprocesses to maximise services and foster innovation.

http://www.irstea.fr/en/research/research-fields/ted/biomic

Anaerobic digestion (AD) is a natural microbial process of degradation of the organic matter which ultimately produces biogas rich in methane. It is more and more widely used in industrial digesters to recover different types of organic waste and at the same time produce energy through the conversion of methane in electrical and thermal energy. However AD microbiota is very sensitive to different types of perturbations that can lead to process failure with important economic and environmental issues.

To better understand inhibition mechanisms and to propose solutions to increase the resistance of anaerobic digestion, different experiment simulating digesters disruption were performed in our lab. A specific focus was given to two key known perturbations of AD (temperature modification and presence of nitrogen ammonia). For that purpose, lab digesters were set-up in parallel and different levels of perturbation were applied. The degradation performance of the digesters was monitored during time, before and after perturbation, and during recovery. Liquid samples taken regularly in the digesters were used to analyse the associated microbial dynamics through sequencing of 16S RNA (RNA sequencing – active microorganisms) or 16S RNA gene (DNA sequencing – present microorganisms). Waste degradation pathways were also monitored through a non-targeted metabolomic approach (dynamics of the metabolites during time measured with high-resolution

mass spectrometer - LTQ-Orbitrap XL). The bioinformatics analysis of this data generated important datasets describing the microbial and metabolomic dynamics during the perturbation events.

The aim of the present internship is to investigate these datasets with statistical methods to extract the most relevant information. A particular focus will be given to the modifications of the microbial dynamics and metabolic pathways associated with the inhibition of AD or with its resistance to the perturbation. The objective is to identify the causes of performance variations (bottlenecks) and to propose strategies of successful management. Microbial and metabolic indicators of optimal performance as well as warning indicators of process failure will also be sought.

For that purpose, different multivariate statistical analysis methods will be used. Multivariate analysis is a group of statistical methods specifically designed to analyse data that arises from more than one variable. They include unsupervised multivariate analysis such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), Partial Least Squares Regression (PLS), etc. with the aim of finding hidden structure within the data; and supervised statistical analysis such as Partial Least Squares Discriminant Analysis (PLS-DA) used to sharpen the separation between groups of observations when structuration of the data is already inferred. All these statistical methods have the appealing properties of reducing the dimension of the data by combining or extracting variables and have proven to be a powerful tool to take into account the common and complementary information contained in the datasets. All the analysis will be performed with R software.

The candidate should be interested in data analysis applied to biological questions. The internship will give the opportunity to work at the interface between data mining and applied research questions. A good knowledge of biostatistics and/or programming and/or R software will be considered favourably.

Key words: biostatistics; multi-omics; 16S sequencing; metabolomics; anaerobic digestion

References:

Chapleur, O., Madigou, C., Civade, R., Rodolphe, Y., Mazéas, L., Bouchez, T. 2016a. Increasing concentrations of phenol progressively affect anaerobic digestion of cellulose and associated microbial communities. *Biodegradation*, **27**(1), 15-27.

Chapleur, O., Mazeas, L., Godon, J.J., Bouchez, T. 2016b. Asymmetrical response of anaerobic digestion microbiota to temperature changes. *Applied Microbiology and Biotechnology*, **100**(3), 1445-1457.

Madigou, C., Poirier, S., Bureau, C., Chapleur, O. 2016. Acclimation strategy to increase phenol tolerance of an anaerobic microbiota. *Bioresour Technol*, **216**, 77-86.

Poirier, S., Bize, A., Bureau, C., Bouchez, T., Chapleur, O. 2016a. Community shifts within anaerobic digestion microbiota facing phenol inhibition: Towards early warning microbial indicators? *Water Research*, **100**, 296-305.

Poirier, S., Desmond-Le Quéméner, E., Madigou, C., Bouchez, T., Chapleur, O. 2016b. Anaerobic digestion of biowaste under extreme ammonia concentration: Identification of key microbial phylotypes. *Bioresource Technology*, **207**, 92-101.

Poirier, S., Madigou, C., Bouchez, T., Chapleur, O. 2017. Improving anaerobic digestion with support media: Mitigation of ammonia inhibition and effect on microbial communities. *Bioresource Technology*, **235**, 229-239.