

# Postdoctoral position (2 years)

## Integration of multiple omics datasets to decipher the mechanisms of anaerobic digestion inhibition

---

*Post-doc duration: 24 months*

*Starting date: position available immediately*

### **Laboratory:**

*-AgroParisTech, Laboratoire de Chimie Analytique  
16, rue Claude Bernard, 75005 Paris*

The AgroParisTech team develops and uses new multivariate, multi-table and multi-way methods of analysis, to extract relevant information from the broadband signals generated by instrumental techniques such as the LC-MS, NMR, GC-MS and 3D Fluorescence.

*-Irstea, HBAN research Unit*

*1 rue PG de Gennes, CS10030, 92761 Antony*

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>

### **Contacts:**

AgroParisTech : Dr. Christophe Cordella : [christophe.cordella@agroparistech.fr](mailto:christophe.cordella@agroparistech.fr)

AgroParisTech : Dr. Delphine Bouveresse: [delphine.bouveresse@agroparistech.fr](mailto:delphine.bouveresse@agroparistech.fr)

Irstea : Dr. Mazeas Laurent: [laurent.mazeas@irstea.fr](mailto:laurent.mazeas@irstea.fr)

Irstea : Dr. Olivier Chapleur: [olivier.chapleur@irstea.fr](mailto:olivier.chapleur@irstea.fr)

---

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 into 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 consequences.

To better understand inhibition mechanisms and to propose solutions to increase the resilience of anaerobic digestion, different experiments simulating digester disruptions were performed in our lab and on an industrial scale. A specific focus was given to three well-known key perturbations of AD (temperature modification, presence of ammonia and modification of feeding composition). For that purpose, lab digesters were set-up in parallel and different

levels of perturbation were applied, and industrial digesters were monitored during modification of operating parameters. The degradation performance of the digesters was monitored over 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). On selected samples, high-throughput metagenomics and metatranscriptomics methodologies will be applied (shotgun sequencing) to identify the effects of the inhibition on the degradation functions. The bioinformatic analysis of this data will generate very large datasets characterising the microbial and metabolomic dynamics during the perturbation events at different levels. Extracting information from these multiple data sets is a challenging task that requires appropriate statistical integrative methods.

**The aim of the postdoctoral position is to investigate the datasets with such methods.**

A wide range of chemometric tools are available and can be adapted to extract relevant information from individual data sets; and also to detect common information among several data sets. Many of these methods are implemented in chemometric software tools which render them accessible for use by analytical chemists, such as the freely available Matlab toolbox "SAISIR" (Toolbox available at: [http://www.chimiometrie.fr/saisir\\_webpage.html](http://www.chimiometrie.fr/saisir_webpage.html)). Different tools will be used in this project. Independent Components Analysis (ICA) aims to extract 'source signals' from complex instrumental data. The calculated proportions of these 'source signals' in each observed signal will facilitate the analysis of the individual data sets of complex signals. The results of these individual analyses may in some cases be used as a relevant input for a subsequent multi-block data fusion. It is clear that a single analytical technique cannot meet all of the issues raised by the study of a complex system. Therefore, the results obtained by the different instrumental, physico-chemical and biological methods will be integrated together so as to bring out their complementarities and communalities. The coupling of data with inherently very different structure, such as chemical analyses and metagenomic responses is particularly challenging. To treat such data tables, Common Components and Specific Weights Analysis (CCSWA or 'ComDim') has proven to be a powerful tool to take into account the common and complementary information contained in these matrices. ComDim identifies directions of samples dispersion which are common to several data tables. 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 powerful tools to take into account the common and complementary information contained in the datasets. The overall objective of the analyses will be to identify the causes of performance variations during AD inhibition (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.

*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 biological research questions.*

The candidate will have also the opportunity to work in collaboration with two teams focussed on different but complementary research questions.

The position is a part of a project funded by the ANR (DIGESTOMIC project <http://www.agence-nationale-recherche.fr/Projet-ANR-16-CE05-0014>).

A successful candidate will

- Analyze data from different omics experiments
- Identify biomarkers of inhibition and optimal performances
- Develop and validate dedicated pipelines to integrate different types of omics data
- Work with different inter-disciplinary teams

Qualifications

- A Ph.D. in biostatistics, chemometrics or environmental microbiology
- Demonstrated competencies in statistics
- Good skills in scripting and programming
- Experience in treatment and statistical analysis of biological data
- Knowledge in biology and sequencing data analysis
- Ability to work collaboratively with members from different teams and different disciplinary fields

Please send CV and motivation letter

Gross salary : 2800 Euros