From describing to managing environmental biotechnologies, through a combination of molecular ecology and biostatistics

Olivier Chapleur

HABILITATION À DIRIGER DES RECHERCHES

PRÉSENTÉE À

L'UNIVERSITÉ PARIS SUD

Par Olivier CHAPLEUR

De la description à la gestion des biotechnologies environnementales en combinant des approches d'écologie moléculaire et de biostatistique

---

Soutenue le 27 novembre 2019

Devant la commission d'examen formée de :

Dr Patricia BONIN	Directrice de Recherche, Institut Méditerranéen d'Océanologie	Examinateur
Dr Jérôme HAMELIN	Directeur de Recherche, INRA Narbonne	Rapporteur
Dr Guillermina HERNANDEZ-RAQUET	Directrice de Recherche, INRA Toulouse	Rapporteur
Dr Ludwig JARDILLIER	Professeur, Université Paris Sud	Examinateur
Dr Marion LECLERC	Chargée de Recherche, INRA Jouy-en-Josas	Rapporteur
Dr Lionel RANJARD	Directeur de Recherche, INRA Dijon	Examinateur

Ce manuscrit résume les travaux de recherche que j'ai effectués depuis ma thèse, en incluant cette dernière. Il est composé de deux grandes parties.

La première partie est intitulée notice individuelle. Cette notice résume sous forme de listes le déroulement de ma carrière de chercheur, les différentes valorisations associées à mon travail de recherche et les différentes activités en lien. Les sections successives reprennent le découpage suggéré dans le 'formulaire HDR' de l'Université Paris Sud.

La seconde partie est une synthèse scientifique des activités que j'ai menées jusqu'à présent. Elle est divisée en différents chapitres permettant de marquer et d'illustrer l'évolution de mes activités. Elle présente également mon projet de recherche pour les années à venir. Cette synthèse est rédigée en anglais.

Table des matières

NOTICE INDIVIDUELLE

Déroulement de la carrière

1. Postes occupés

2008-2012 :	**Doctorat** en Sciences de l'Environnement, Cemagref, Antony
2012-2017 :	**Chercheur** en écologie microbienne des bioprocédés anaérobies de traitement de déchets, Irstea, Antony Unité de Recherche HBAN (Hydrosystèmes et Bioprocédés Antony) Equipe BIOMIC (Bioprocédés et bIOtechnologies MICrobiennes pour la valorisation des déchets)
Février-Mai 2017 :	**Séjour scientifique** auprès du Dr Déjean (https://perso.math.univ-toulouse.fr/dejean/), Université de Toulouse, Toulouse Institut de Mathématiques (https://www.math.univ-toulouse.fr/) Équipe de recherche Statistique et Probabilités
Août-Décembre 2017 et Mars-Mai 2019	**Visitor scientist** auprès du Dr Lê Cao (https://lecao-lab.science.unimelb.edu.au/), Université de Melbourne, Melbourne, Australie Melbourne Integrated Genomics Research team (https://research.unimelb.edu.au/integrative-genomics/MIG-home)
Depuis 2018 :	**Chercheur** en écologie microbienne des bioprocédés anaérobies de traitement de déchets, Irstea, Antony Unité de Recherche PROSE (PRocédés biOtechnologiques au Service de l'Environnement https://www.irstea.fr/fr/recherche/unites-de-recherche/prose) Equipe PROSE (Procédés microbiens au service de l'environnement) **Responsable du pôle analytique** de l'équipe PROSE (7 agents : 1 TR, 1 AI, 5 IE en charge des analyses, https://www.irstea.fr/fr/pole-analytique-antony)

2. Master

Nom du diplôme :	Master 2 Ecologie, Biodiversité, Evolution
Etablissement d'inscription :	Institut National Agronomique Paris Grignon
Date d'obtention :	2007
Titre du mémoire :	Identification de groupes microbiens fonctionnels impliqués dans la dégradation anaérobie de sous-produits de la cellulose dans des méthaniseurs de déchets

3. Doctorat

Date de début de doctorat :	2008
Date de soutenance :	18 juin 2012
Etablissement(s) d'inscription :	AgroParisTech
Titre de la thèse :	Ingénierie écologique des communautés microbiennes de méthanisation des déchets ligno-cellulosiques
Directeur de thèse :	Jean-Jacques Godon
Encadrement :	Théodore Bouchez
Laboratoire où s'est déroulé le doctorat :	Equipe SOWASTE, UR HBAN, Cemagref Antony, 4 PARC de Tourvoie, 92160, Antony

Production scientifique associée aux travaux de recherche

1. Articles de revue

	Année	Référence	Revue et impact factor	Nombre de citations
1	2013	**Chapleur O, Wu T-D, Guerquin-Kern J-L, Mazéas L, Bouchez T.** 2013. SIMSISH technique does not alter the apparent isotopic composition of bacterial cells. PloS one **8:**e77522.	PLoS ONE IF=3.540	4
2	2014	**Chapleur O, Bize A, Serain T, Mazéas L, Bouchez T.** 2014. Co-inoculating ruminal content neither provides active hydrolytic microbes nor improves methanization of 13C-cellulose in batch digesters. FEMS Microbiol Ecol **87:**616-629.	FEMS Microbiology Ecology IF=3.960	17
3	2014	**Lü F, Bize A, Guillot A, Monnet V, Madigou C, Chapleur O, Mazéas L, He P, Bouchez T.** 2014. Metaproteomics of cellulose methanisation under thermophilic conditions reveals a surprisingly high proteolytic activity. ISME J **8:**88-102.	The ISME Journal IF=9.520	85
4	2016	**Chapleur O, Mazeas L, Godon JJ, Bouchez T.** 2016. Asymmetrical response of anaerobic digestion microbiota to temperature changes. Appl Microbiol Biotechnol **100:**1445-1457.	Applied Microbiology and Biotechnology IF=3.420	12
5	2016	**Poirier S, Desmond-Le Quéméner E, Madigou C, Bouchez T, Chapleur O.** 2016. Anaerobic digestion of biowaste under extreme ammonia concentration: Identification of key microbial phylotypes. Bioresour Technol **207:**92-101.	Bioresource Technology IF=5.651	60
6	2016	**Chapleur O, Madigou C, Civade R, Rodolphe Y, Mazéas L, Bouchez T.** 2016. Increasing concentrations of phenol progressively affect anaerobic digestion of cellulose and associated microbial communities. Biodegradation **27:**15-27.	Biodegra-dation IF=2.410	17

7	2016	**Madigou C, Poirier S, Bureau C, Chapleur O.** 2016. Acclimation strategy to increase phenol tolerance of an anaerobic microbiota. Bioresour Technol **216:**77-86.	Bioresource Technology IF=5.651	19
8	2016	**Poirier S, Bize A, Bureau C, Bouchez T, Chapleur O.** 2016. Community shifts within anaerobic digestion microbiota facing phenol inhibition: Towards early warning microbial indicators? Water Res **100:**296-305.	Water Research IF=6.942	33
9	2016	**Hao L, Bize A, Conteau D, Chapleur O, Courtois S, Kroff P, Desmond-Le Quéméner E, Bouchez T, Mazéas L.** 2016. New insights into the key microbial phylotypes of anaerobic sludge digesters under different operational conditions. Water Res **102:**158-169.	Water Research IF=6.942	28
10	2017	**Poirier S, Madigou C, Bouchez T, Chapleur O.** 2017. Improving anaerobic digestion with support media: Mitigation of ammonia inhibition and effect on microbial communities. Bioresour Technol **235:**229-239.	Bioresource Technology IF=5.651	22
11	2017	**Connan R, Dabert P, Le Roux S, Chapleur O, Bridoux G, Vanotti MB, Beline F, Magri A.** 2017. Characterization of a combined batch-continuous procedure for the culture of anammox biomass. Ecological Engineering **106:**231-241.	Ecological Engineering IF=3.023	5
12	2018	**Poirier S, Déjean S, Chapleur O.** 2018. Support media can steer methanogenesis in the presence of phenol through biotic and abiotic effects. Water Res **140:**24-33.	Water Research IF=6.942	4
13	2019	**Cardona L, Guenne A, Chapleur O, Mazéas L.** Co-digestion of wastewater sludge: choosing the optimal blend. Waste Management **81**, 772-781.	Waste Management IF=5.431	0
14	2019	**Madigou C, Lê Cao K-A, Bureau C, Mazéas L, Déjean S, Chapleur O.** Ecological consequences of abrupt temperature changes in anaerobic digesters. Chemical Engineering Journal **361**: 266-277.	Chemical Engineering Journal IF=6.735	2
	/	**Cardona L, Lê Cao K-A, Puig Castellví F, Bureau C, Madigou C, Mazéas L, Chapleur O.** Integrative analyses to investigate the link between microbial activity and molecules degradation during anaerobic digestion. Submitted to Chemical Engineering Journal		
	/	**Bodein A*, Chapleur O*, Droit A, Lê Cao K-A.** A generic multivariate framework for the integration of microbiome longitudinal studies with other data types. Submitted to Frontiers in genetics, deposited on bioRxiv (https://www.biorxiv.org/content/10.1101/585802v2)		

	/	**Puig-Castellví F, Cardona L Jouan-Rimbaud Bouveresse D, Cordella C, Mazéas L, Rutledge D, Chapleur O** Community Interactions in Anaerobic Digesters by Common Components Analysis and Phylogenetics. Submitted to Science of the Total Environment		
	/	**Poirier S, Déjean S, Midoux C, Lê Cao KA, Chapleur O.** Predictive models of AD inhibition can be built by integrating independent microbial studies. Submitted to the ISME Journal		
	/	**Puig-Castellví F, Cardona L Jouan-Rimbaud Bouveresse D, Cordella C, Mazéas L, Rutledge D, Chapleur O.** Time-resolved Metabolomic fingerprint of the anaerobic co-digestion of sludge. Submitted to Water research		
	/	**Cardona L, Mazéas L, Chapleur O.** Zeolite mitigates the ammonia inhibition of anaerobic digestion by favouring the development of syntrophy for propionate degradation Submitted to Bioressource Technology		

2. Data papers

	Année	Référence
1	2018	**Poirier S, Chapleur O**. 2018. Influence of support media supplementation to reduce the inhibition of anaerobic digestion by phenol and ammonia: Effect on degradation performances and microbial dynamics. Data in Brief. **19**:1733-1754.
2	2018	**Poirier S, Chapleur O.** 2018. Inhibition of anaerobic digestion by phenol and ammonia: Effect on degradation performances and microbial dynamics. Data in Brief. **19**:2235:2239.

3. Chapitre d'ouvrages scientifiques

	Année	Référence
1	2015	« La méthanisation », éditions Lavoisier, 2015, René Moletta coordonnateur. Co-auteur du chapitre « L'élimination et la méthanisation des déchets non dangereux en installation de stockage » avec Théodore Bouchez.

Pour les communications, l'auteur présentateur est indiqué avec un astérisque*.

4. Communications scientifiques avec actes

	Année	Référence
1	2009	**Chapleur O*, Li T, An S, Wu T-D, Vavilin V, Guerquin-Kern J-L, Mazéas L, Bouchez T.** 2009. Linking of identity and function in complex communities of uncultured microbes using a coupled SIP-NanoSIMS approach, 5th IWA Activated Sludge Population Dynamics Specialist Conference - Microbial Population Dynamics in Biological Wastewater Treatment (ASPD 5), 24-27 May 2009, Aalborg, Denmark. Vol. Abstract Book Aalborg:146-148.
2	2009	**Mazéas L*, Li T, Wu T-D, Chapleur O, An S, Grossin-Debattista J, Vavilin V, Guerquin-Kern J-L, Bouchez T.** 2009. Simultaneous analysis of microbial identity and function using NanoSIMS: Application to anaerobic degradation of methanol. Goldschmidt Conference, 21-26 June 2009, Davos, Suisse. Geochimica Et Cosmochimica Acta. 73:A853-A853.
3	2011	**Chapleur O*, Mazéas L, Bouchez T.** 2011. Consequences of temperature changes on cellulose anaerobic degradation and functional microbial communities dynamics: toward an ecological engineering of anaerobic digesters. Microbes in Wastewater and Waste treatment conference - Bioremediation and Energy Production (MWT 1), 23-25 January 2011, Goa, India.
4	2011	**Grossin-Debattista J, Bouchez T. Fan L. Wu TD, Chapleur O, Bize A, Cardona L, Mazéas L*.** 2011. Syntrophic acetate oxidation in methanogenic reactors: an unexpected metabolic behaviour analysed by a polyphasic approach. First International Symposium on Microbial resource management in biotechnology: Concepts & Applications, 30 June-1 July 2011, Ghent, Belgium.
5	2011	**Mazéas L*, Li T, Wu T-D, Chapleur O, An S, Grossin-Debattista J, Guerquin-Kern J-L, Bouchez T.** 2011. Importance of syntrophic acetate oxidation during thermophilic municipal solid wastes anaerobic digestion. Goldschmidt Conference, 14-19 August 2011, Prague, Czech Republic.
6	2012	**Chapleur O*, Bize A, Serain T, Mazéas L, Bouchez T.** 2012. Co-inoculation of ruminal content and anaerobic sludge in 13C cellulose fed anaerobic batch digesters. Global Assessment for Organic Resources and Waste Management Conference (ORBIT 8), 12-15 June 2012, Rennes, France.
7	2013	**Chapleur O*, Civade R, Hoyos C, Mazeas L, Bouchez, T.** 2013. Growing concentrations of phenol increasingly modify microbial communities' dynamics and performances' stability of anaerobic digesters. World Congress on Anaerobic Digestion (AD 13), 25-28 June 2013, Santiago de Compostella, Spain
8	2016	**Poirier S*, Chapleur O.** 2016. Microbial community shifts within anaerobic digesters facing ammonia inhibition: Towards early warning bioindicators? Microbial Ecology in Water Engineering & Biofilms joint specialist conference (MEWE 2016), 4-7 September 2016, Copenhagen, Denmark.
9	2016	**Madigou C, Poirier S, Chapleur O*.** 2016. Characterization of the acclimatization of an anaerobic microbiota subjected to inhibitory phenol concentration. Microbial Ecology in Water Engineering & Biofilms joint specialist conference (MEWE 2016), 4-7 September 2016, Copenhagen, Denmark.

10	2017	**Conteau D*, Franchi Morales O, Chapleur O, Gaval G, Traba Lago G, Mazéas L, Araya P, Kroff P, Barillon B.** 2017. Low temperature sewage sludge anaerobic digestion: full-scale proof of interest and study of microbial adaptation. World Congress on Anaerobic Digestion (AD 15), 17-20 October 2017, Beijing, China
11	2019	**Poirier S, Déjean S, Midoux C, Lê Cao KA, Chapleur O*.** Predictive models of AD inhibition can be built by integrating independent microbial studies. World Congress on Anaerobic Digestion (AD 16), 23-27 June 2019, Delft, Netherlandss

5. Communications scientifiques internationales

	Année	Référence
1	2010	**Chapleur O*, Wu T-D, Guerquin-Kern J-L, Mazéas L, Bouchez T.** SIMSISH methodology: a very accurate and quantitative technique to penetrate inside the functional networks of microorganisms in complex ecosystems. International Symposium on Microbial Ecology (ISME 13), 22-27 August 2010, Seattle, United States of America.
2	2011	**Chapleur O*, Mazéas L, Bouchez T.** 2011. Towards an ecological engineering of anaerobic digesters: shaping structure of complex communities degrading cellulose through substrate adaptation. International Conference on Biogas Microbiology (ICBM 1), 13-15 September 2011, Leipzig, Germany.
3	2012	**Chapleur O*, Mazéas L, Bouchez T.** 2012. Asymmetrical behavior of cellulose anaerobic digestion towards temperature changes: a trigger for managing anaerobic communities in digesters? International Symposium on Microbial Ecology (ISME 14), 19-24 August 2012, Copenhagen, Denmark.
4	2017	**Dabert P*, Buffet J, Le Roux S, Chapleur O, Bize A, Trémier A.** 2017. Stability of chemical and microbial composition of digestates along time in agricultural and urban full-scale anaerobic digesters. International Ramiran conference, 4-6 September 2017, Wexford, Ireland.
5	2017	**Poirier S, Madigou C, Déjean S. Chapleur O*.** 2017. Towards the development of microbial indicators of anaerobic digestion inhibition. International conference on biogas microbiology (ICBM 3), 1-3 May 2017, Wageningen, Netherlands
6	2017	**Madigou C, Mazéas L., Bureau C, Déjean S, Chapleur O*.** 2017. Identifying the bottlenecks that limit temperature modification in anaerobic digesters. International conference on biogas microbiology (ICBM 3), 1-3 May 2017, Wageningen, Netherlands
7	2017	**Chapleur O*, Poirier S, Madigou C, Lê Cao K-A, Déjean S**. 2018. Development of microbial indicators of anaerobic digestion inhibition with omics data integration. International workshop, Environmental Omics, Integration and Modelling, 18-20 October 2018, Barcelona, Spain
8	2017	**Chapleur O*, Déjean S, Lê Cao K-A.** 2018. Identifying biomarkers by integrating multiple-omics datasets to improve anaerobic digestion (flash presentation). International workshop, Environmental Omics, Integration and

		Modelling, 18-20 October 2018, Barcelona, Spain
9	2018	**Chapleur O*, Poirier S, Madigou C, Lê Cao K-A, Déjean S.** 2018. Development of microbial indicators of anaerobic digestion inhibition with omics data integration. Biostatistics workshop, Mathematical Institute, Toulouse, France

6. Communications scientifiques nationales

	Année	Référence
1	2009	**Chapleur O*, Bouchez T.** 2009. Ingénierie écologique des communautés microbiennes dans les bioprocédés de traitement anaérobie des déchets ligno-cellulosiques, 2ème Journée Thématique Microbiologie Environnementale FIRE, 14 janvier 2009, Paris, France.
2	2009	**Chapleur O*, Mazéas L, Bouchez T.** 2009. Ingénierie écologique des communautés microbiennes dans les bioprocédés anaérobies de traitement des déchets. Colloque Groupe des Acteurs de l'Ingénierie Ecologique (GAIE), 9 décembre 2009, Paris, France.
3	2012	**Chapleur O*, Bize A, Serain T, Mazéas L, Bouchez T.** 2012. Co-inoculation d'un contenu ruminal et d'un digestat d'ordures ménagères dans un réacteur batch anaérobie dégradant de la cellulose : une caractérisation fonctionnelle, Conférence du réseau national biofilm, 24 janvier 2012, Narbonne, France
4	2013	**Chapleur O*, Wu T-D, Guerquin-Kern J-L, Mazéas L, Bouchez T.** 2013. La technique SIMSISH, une méthode précise et quantitative pour décrypter les réseaux multifonctionnels des communautés microbiennes complexes. Journées thématiques du Réseau National Biofilm, 20 novembre 2013, Pau, France.
5	2013	**Chapleur O*, Bize A, Bouchez T, Bureau C, Madigou C, Mazéas L.** 2013. Développement d'outils analytiques innovants pour les bioprocédés et biotechnologies microbiennes de valorisation des déchets. Colloque BIOMINNOV, 9 décembre 2013, Romainville, France
6	2015	**Poirier S*, Bouchez T, Chapleur O**. Les composés inhibiteurs de la digestion anaérobie : état de l'art et stratégies d'acclimatation. Journées Recherche et Innovation biogaz méthanisation, 3 – 5 Février 2015, Rennes, France
7	2016	**Madigou C, Poirier S, Chapleur O*.** 2016. Amélioration de la résistance au phénol d'un microbiote anaérobie par une stratégie d'acclimatation. Journées thématiques de l'association française d'écologie microbienne (AFEM), 31 mai -1 juin 2016, Marseille, France
8	2016	**Poirier S*, Chapleur O**. 2016. Influence de l'azote ammoniacal sur la composition d'un écosystème anaérobie issu d'un digesteur industriel. Journées thématiques de l'association française d'écologie microbienne (AFEM), 31 mai -1 juin 2016, Marseille, France
9	2018	**Cardona L*, Mazéas L, Chapleur O**. 2018. Improving waste valorisation: the data omics integration intake. Journées de l'Ecole Doctorale ABIES

7. Communications internationales par affiche

	Année	Référence
1	2010	**Chapleur O*, Mazéas L, Bouchez T.** 2010. Ecological engineering of microbial communities through substrate adaptation: consequences on anaerobic digestion of cellulose in municipal solid waste. World Congress on Anaerobic Digestion (AD 12), 31 October-4 November 2010, Guadalajara, Mexico.
2	2011	**Chapleur O*, Mazéas L, Bouchez T.** Ecological engineering of microbial communities through substrate adaptation: consequences on anaerobic digestion of cellulose. 2011. Microbes in Wastewater and Waste treatment conference - Bioremediation and Energy Production (MWT 1), 23-25 January 2011, Goa, India.
3	2011	**Chapleur O*, Mazéas L, Bouchez T.** 2011. Asymmetrical behavior of cellulose anaerobic digestion towards temperature changes: a trigger for ecological engineering of anaerobic digesters? Conference on Biogas Microbiology (ICBM 1), 13-15 September 2011, Leipzig, Germany.
4	2012	**Chapleur O*, Bize A, Serain T, Mazéas L, Bouchez T.** 2012. Dynamic of 13C cellulose biodegradation in batch digesters co-inoculated with rumen content and anaerobic sludge. International Symposium on Microbial Ecology (ISME 14), 19-24 August 2012, Copenhagen, Denmark.
5	2016	**Poirier S*, Bouchez T, Chapleur O.** 2016. Enhancing anaerobic digestion of biowaste under extreme ammonia concentration with support media : Identification of key microbial phylotypes. International Symposium on Microbial Ecology (ISME 16), 21-26 August 2016, Montreal, Canada.
6	2016	**Madigou C, Poirier S, Chapleur O*.** 2016. Acclimation of an anaerobic microbiota subjected to inhibitory phenol concentration. International Symposium on Microbial Ecology (ISME 16), 21-26 August 2016, Montreal, Canada.
7	2017	**Cardona L*, Poirier S, Madigou C, Bouchez T, Mazéas L, Chapleur O.** 2017. Understanding the effect of ammonia on anaerobic microbiota during biowaste anaerobic digestion. International conference on biogas microbiology (ICBM 3), 1-3 May 2017, Wageningen, Netherlands
8	2017	**Connan R*, Magri A, Chapleur O, Bridoux G, Béline F.** 2017. Impact of the inoculum source and nitrite concentration on anaerobic ammonium oxidation bacteria enrichment. International Ramiran conference, 4-6 September 2017, Wexford, Ireland.
9	2017	**Chapleur O*, Déjean S, Lê Cao K-A.** 2017. Identifying biomarkers by integrating multiple-omics datasets to improve anaerobic digestion. International workshop, Environmental Omics, Integration and Modelling, 18-20 October 2018, Barcelona, Spain
10	2018	**Chapleur O*, Poirier S, Lê Cao K-A.** 2018. Effect of ammonia on the dynamics of anaerobic digestion microbiome: omics data integration in a time course context. International Symposium on Microbial Ecology (ISME 17), 12-17 August 2016, Leipzig, Germany.

11	2018	**Cardona L*, Lê Cao K-A, Bureau C, Madigou C, Rouillac L, Lê Cao, Mazeas L, Chapleur O.** 2018. Multi-omics data integration to decipher the impact of feeding composition on the microbiota of anaerobic digestion. International Symposium on Microbial Ecology (ISME 17), 12-17 August 2016, Leipzig, Germany.

8. Actions de vulgarisation

	Année	Référence
1	2010	Invité de l'émission « Pose ta question » diffusée sur « Terre d'infos tv » en 2010, thème « Déchets et Recyclage ». Emission de vulgarisation réunissant un groupe d'enfants, une journaliste et un expert sur des sujets liés à l'environnement.
2	2011	Rapport annuel 2011 Irstea « 30 questions pour comprendre la recherche environnementale », réponse à la question 17 : « Comment travaille un chercheur »
3	2012	Participation à l'ouvrage « Quelle énergie durable pour demain ? », édité par Irstea, partie « optimiser les processus microbiens de dégradation des déchets »
4	2013	Présentation « Recherche pour le développement d'outils pour améliorer la gestion des bioprocédés anaérobies. » lors des « rencontres éco-technologiques 2013 » du réseau PEXE, éco-entreprises de France et des Instituts Carnot, 2 juillet 2013, Nanterre, France.
5	2014	Présentation « Processus de dépollution grâce aux écosystèmes microbiens » lors des « rencontres de l'INRA », 28 février 2014, Salon International de l'Agriculture, Paris, France
6	2016-2017	Encadrement d'un groupe d'étudiants de classe préparatoire pour un projet collectif sur la méthanisation

Collaborations et contrats de recherche

1. Contrats de recherches obtenus

- **Coordination du projet METHARESIST** (EC2CO INSU CNRS 42k€ 2015-2016): "Approches méta-omiques pour établir des stratégies de résistance des digesteurs anaérobies aux micropolluants, cas du phénol".
- **Co-coordination du projet DIGESTOMIC** (ANR, 450k€, 2016-2020): "Elaboration de nouvelles stratégies opératoires pour lever les verrous de la digestion anaérobie et élargir ses domaines d'application à l'aide d'approches méta-omiques". Responsable d'un des trois work-package du projet.
- **Coordination d'un projet d'échange scientifique bilatéral avec l'Australie** (financement PHC FASIC, IRSTEA, ED ABIES, AFRAN : 17k€, 2017-2019).
- **Coordination du projet STABILICS** (ANR, 220k€, 2019-2023): "Nouvelles perspectives dans les déterminants de la stabilité des bioprocédés anaérobies en couplant des approches multi-omiques et statistiques".

2. Collaborations nationales ou internationales soutenues

- **Collaboration avec des chercheurs de l'UR OPAALE d'Irstea Rennes :** un article de revue, une présentation orale et une présentation affichée cosignés (thèse R. Connan 2013-2016) participation au comité thèse de S. Piveteau (encadrement ML Daumer 2015-2018)
- **Collaboration avec le laboratoire LGPM de Centrale-Supélec :** co-encadrement d'un stage avec le Dr. B. Taidi (A. AshDSouza) en 2016
- **Collaboration avec Suez Environnement :** partenaires du projet ANR DIGESTOMIC (2016-2020)
- **Collaboration avec Pontificia Universidad Católica de Valparaíso, Chili :** accueil et encadrement d'un doctorant (O. Franchi) pendant 3 mois en 2017.
- **Collaboration avec l'équipe IAQA, AgroParisTech, UMR GENIAL :** partenaire des projets METHARESIST (2015-2016) et DIGESTOMIC (2016-2020).
- **Collaboration avec l'Institut de Mathématiques de Toulouse :** séjour scientifique de plusieurs mois en 2017 auprès du Dr S. Déjean, divers articles cosignés, partenaires du projet ANR STABILICS.
- **Collaboration avec l'équipe du Dr K-A Lê Cao, Melbourne Integrative Genomics, à l'Université de Melbourne, Australie :** séjour scientifique de plusieurs mois en 2017 puis 2019, divers articles cosignés, partenaires du projet ANR STABILICS. Co-organisation d'un workshop sur l'analyse des données omiques en juin 2018 à l'université Paris Saclay (25 participants). Co-organisation d'un workshop

« Advancing environmental biotechnologies through advanced microbiology and cutting-edge computational statistics » à l'Université de Melbourne en avril 2019 (35 participants).

- **Collaboration avec le laboratoire de chimie moléculaire (LCM) de l'école polytechnique:** analyse métabolomique d'échantillons, la préparation d'un article co-signé est en cours, partenaires du projet ANR STABILICS.

Expériences d'encadrement

1. Tableau récapitulatif des expériences d'encadrement

	Total
Encadrement de doctorants	2 (une thèse soutenue en 2016 et une thèse en cours, soutenance janvier 2020)
Encadrement de stages de niveau M2	5
Encadrement de stages de niveau L3 ou M1	2
Encadrement autre	Stage ingénieur fin d'étude : 7 Doctorant en séjour court (3 mois) : 1 Postdoctorant: 1

2. Encadrement de stagiaires

Année	Nom	Titre du stage	Niveau	Pourcentage d'encadrement
2008	An Shu	Identification de groupes fonctionnels au sein de communautés microbiennes complexes dans les bioprocédés de traitement anaérobie des déchets ligno-cellulosiques	M2	100%
2009	Charlotte Richard	Influence de la température sur les groupes microbiens fonctionnels responsables de la dégradation de la cellulose	M2	100%
2009	Adeline Desprez	Evaluation de l'Effet du Protocole de la Technique NanoSIMS-ISH	M2	100%
2009	Vincent Hébrail	Etude de l'effet d'un changement brutal de température (passage de 55°C à 35°C puis de 35°C à 55°C) sur la réponse des groupes microbiens fonctionnels de dégradation de la cellulose	Stage Ingénieur	100%
2010	Thibaut Serain	Influence of the introduction of an exogenous biomass on anaerobic degradation of cellulose in municipal solid waste digesters	Stage Ingénieur	100%

2010	Florian Almeras	Influence d'une biomasse exogène (sur le démarrage des fonctions microbiennes et les potentialités de dégradation de la cellulose dans les bioprocédés de méthanisation – Cas de la bactérie Fibrobacter succinogenes	Stage Ingénieur	100%
2012	Raphaël Civade	Impact du phénol sur la diversité et la stabilité des communautés microbiennes de dégradation de la cellulose dans les bioprocédés anaérobies	M2	100%
2013	Yohan Rodolphe	Impact of growing concentrations of phenol on microbial communities degrading cellulose in anaerobic conditions	4ème année ingénieur	100%
2016	Alexia AshDSouza	Developing a method of measuring Cell quantity and vitality in a culture of Arthrospira platensis	4ème année ingénieur	80%
2017	Camille Levrard	Effet de l'ajout de co-substrats sur les performances d'un digesteur anaérobie et les dynamiques microbiennes associées	L3	75%
2018	Delphine Cirederf	Suivi des dynamiques microbiennes de la digestion anaérobie à l'aide d'approche ADN et ARN	L3	75%
2018	Sophie Renault	Exploratory statistical analysis of multi-omics data to decipher the mechanisms of anaerobic digestion inhibition	4ème année ingénieur	100%
2018	Pierre-Antoine BAR	Ammonia inhibition of anaerobic digestion: A Research on the mitigation mechanisms of ammonia inhibition by zeolite	4ème année ingénieur	80%
2019	Alexandra Claudin	Effet de la codigestion sur le processus de digestion anaérobie en termes de performance des digesteurs et de dynamique des populations microbiennes	M2	80%

3. Autres expériences d'encadrement

2017 : accueil d'Oscar Franchi (doctorant chilien) pendant trois mois au laboratoire pour un travail collaboratif (un article en cours de rédaction « Low temperature sewage sludge anaerobic digestion: full-scale proof of interest and study of microbial adaptation. »). Pourcentage d'encadrement lors de son séjour : 75%.

2014-2017 : accompagnement de Céline Madigou au cours d'un projet de montée en compétence pour le passage du grade d'Assistant Ingénieur à Ingénieur d'Etude. Promotion réalisée en 2017. Valorisation du travail sous forme de deux articles (2016 et 2019) pour lesquels je suis dernier auteur.

2018-2020: postdoctorat de Francesc Puig-Castellví dans le cadre du projet DIGESTOMIC (Accueil conjoint AgroParisTech-Irstea). Deux articles dont je suis dernier auteur sont en cours de révision.

4. Encadrement de doctorants

4.1. Simon Poirier

- Dates : 2013-2016
- Titre de la thèse : Exploration écologique des phénomènes d'inhibition associés à la digestion anaérobie pour le développement d'indicateurs microbiens et de stratégies de résistance
- Directeur de thèse HDR : Théodore Bouchez
- Pourcentage d'encadrement :
    - Conception du dispositif expérimental : 90%
    - Suivi du travail expérimental : 90%
    - Analyse et interprétation des résultats : 80%
    - Valorisation des résultats : 80%
- Production scientifique : 4 articles de revue et 2 data paper, 1 article en cours de révision (je suis dernier auteur de l'ensemble de cette production)
- Devenir : contrat postdoctoral obtenu en sortie de thèse (INRA Micalis, Jouy-en-Josas) (2 ans). Pour suivre, Simon débutera un nouveau postdoctorat à l'EPLF, en Suisse (septembre 2019).

4.2. Laëtitia Cardona

- Dates : 2017-2020
- Titre de la thèse : Exploration écologique des phénomènes d'inhibition associés à la digestion anaérobie pour le développement d'indicateurs microbiens et de stratégies de résistance
- Directeur de thèse HDR : Laurent Mazéas
- Pourcentage d'encadrement :
    - Conception du dispositif expérimental : 80%
    - Suivi du travail expérimental : 90%
    - Analyse et interprétation des résultats : 80%
    - Valorisation des résultats : 80%
- Production scientifique : Au moins 4 articles sont prévus pendant la thèse (1 est accepté, 2 sont actuellement en révision et 1 en écriture)

Activités administratives et d'intérêt collectif

1. Contribution à la vie et au fonctionnement du collectif de recherche

Depuis le 1er janvier 2018 je suis responsable du pôle analytique de l'équipe PROSE. Le pôle analytique de l'équipe regroupe les agents en charge de l'analyse des échantillons produits au cours des expérimentations. Il est composé d'un technicien de la recherche, un assistant ingénieur et cinq ingénieurs d'étude. Il concentre les compétences associées aux différentes méthodologies analytiques développées dans l'équipe, notamment la chimie analytique, la microbiologie, la biologie moléculaire, la bioinformatique et la biostatistique.

J'ai pour missions d'assurer l'animation et la coordination des activités du pôle et de participer au management de l'équipe en lien avec le Directeur de l'unité (en particulier gestion des ressources humaines du pôle analytique, conduite des EAP, suivi des congés et des missions,…).

2. Contribution à l'enseignement ou aux formations

Depuis 2016 j'ai pris part à la construction d'un nouvel enseignement à l'Ecole Polytechnique (Graduate Degree ECOSEM : « Ecotechnologies for Sustainability and Environment Management Master »). Cette formation a accueilli ses premiers élèves en septembre 2018. Au sein de cet enseignement je suis responsable et interviens dans le module « Exploration and statistical analysis of complex datasets » (20h). J'interviens également dans le module « Microbial ecology for environmental sciences » (11h).

J'interviens ou suis intervenu dans d'autres enseignements en écologie microbienne (master 2) (notamment Master 2 Microbiologie Appliquée et Génie Biologique, AgroParisTech).

J'ai participé ou participe à différent comités de thèse (Diane Plouchart – INRA LBE, Tristan Cerisy - Genoscope, Simon Piveteau – Irstea Rennes, Nelly Badalato – Irstea Antony, Gregory Marandat – Irstea Antony, Thomas Jouen – Irstea Antony).

3. Relecture d'articles

Relecture d'articles pour différents journaux : Biosystems Engineering, Biotechnology for Biofuels, Bioressource Technology, Journal of Environmental Quality, Journal of Microbiological Methods, Microorganisms, Plos One, Science of the Total Environment, Waste Management, Water Research, Water Science and Technology.

SCIENTIFIC SYNTHESIS

From describing to managing environmental biotechnologies, through a combination of molecular ecology and biostatistics

Preamble

As a research scientist, I was sometimes asked what my field of expertise was. Even though my research is very multidisciplinary I would say that I am a microbial ecologist, with a strong focus on microbial bioprocesses and more specifically anaerobic digestion using molecular biology. In other words a molecular microbial ecologist in the service of biodegradation bioprocesses. Not sure that makes the things clearer…

I've been exploring this field since I started my PhD. Some people could find that it is dramatic, but on the contrary I think that I was really lucky.

Since I started my PhD, methods used to analyse the microbial communities have changed and evolved a lot. It brought a constant renewal in the research I do. New techniques in microbial ecology have been invented and deployed. New perspectives have been opened. New questions have been addressed. It required taming the new methodologies, adapting them to our samples, formalising the new research issues and learning how to make the most of the data obtained. All of this mostly in the service of one process, the anaerobic digestion. From year to year, my activities evolved continuously; I created new collaborations with scientist from new fields, constantly renewing my satisfaction.

One of the first paper I read as a master student exploring the world of microbial ecology, and that probably influenced my work a lot, was a recent (at that time) opinion letter by Curtis in Nature Reviews Microbiology entitled "Microbial ecologists: It's time to 'go large'" (Curtis, 2006). Curtis began his paper as follows: *« Microorganisms are small, they are extremely important and there is an awful lot of them — 10^{30} to be inexact. But we have only a vague idea of the extent of microbial diversity, with estimates differing by factors of 1,000 or more. We have even less of an idea about the proportional abundance of microorganisms, and not much of a clue how these patterns do and do not change. This is not simply a search for trivial facts. These concepts underlie all explorations in, and exploitations of, the microbial world"*. That paper was really important to me. It made me realise that the exploration of the microbial world was only in its infancy, despite the key importance of microorganisms in most of the natural and anthropic biological processes. It also made me understand that all these processes constituted great ecosystems at the microbial scale, with characteristics and functioning very similar to the 'macroscopic ecosystems' I had been studying during my master: composed of multiple species, with competitions and interactions, dynamic and changing, etc. I understood that focusing on only one or a few selected species was probably not the best option for describing and improving the processes.

The interest of studying the system as a whole was well known from microbial ecologist who tried to include it as much as possible in their studies. *"So why are we still in the dark?"* asked Curtis. *"There are many considerations, but the single factor that outweighs all others is sample size — we are looking at vast systems with pathetically small samples. A clone library of 1,000 might sound fine but in a community of 10^{18} it is modest: one clone for every 10^{12} individuals."* What to do then? Stop studying these systems? Focus only on the most dominant microorganisms? Combine all strengths to target only one selected microbial ecosystem or process? Change methods? Responses arrived quickly. The 13 years period between now and that paper has indeed seen the advent of high-throughput methodologies, in

particular new generation sequencing (NGS), and the drop of sequencing cost and the generalization of the use of omics technologies. Along with others, these new developments allowed to take a giant step in the exploration of microbial ecosystems, and to start to *"face up to the scale of the microbial world, retool and 'go large'"* as suggested by Curtis. They enabled a spectacular progress in a wide variety of fields.

How to go further? *"To escape from the Captain Cook phase of one-off heroic explorations, we must industrialize and automate sample processing and analysis"* said Curtis. That's also more or less what microbial ecologists have started to do in the last years, once they successfully managed the new techniques. More and more samples, time points, conditions, replicates are processed. Of course, microbial interactions are still not crystal clear. Microbial management is still at its infancy. Industrializing is still difficult for technical and financial reasons. On a certain point of view we are still in the dark, but the dark is not as dark as before and I believe that the tremendous amount of data we are now able to generate at a 'reasonable' cost will help us to bring more and more light to the fascinating microbial world. However, according to me, a crucial aspect, which is not detailed in Curtis's article, is the need to find methods to organise these data, analyse them and extract essential information. Curtis mentions mathematical modelling. How exactly? To answer which questions? That will probably be the challenge of the next 10 years.

Of course, microbial bioprocesses and specifically anaerobic digestion benefited a lot from these developments. The microbial ecosystem of anaerobic digestion is both very complex and dynamic. It is composed of an intricate network of microorganisms interacting together to progressively degrade complex organic molecules into small one carbon molecules: the methane and the carbon dioxide. It provides an answer to strong environmental issues as it enables valorisation of waste through the production of biogas, a renewable energy. However, despite it has been used and empirically improved for a century, the underlying microbial processes are still not mastered. The process still misses stability and management tools to be fully exploited. Anaerobic digestion has therefore provided and still provides a highly relevant field to apply previously mentioned methods to answer unexplored questions. The last developments of microbial ecology enabled a better description of the degradation processes, of the function at stake, of the role of the microorganisms… They enabled to understand that the performances of the process are strongly dependant of the dynamics of the microbial ecosystem. It is now time to organize these data, and to complete them in order to implement new microbial management strategies to improve the functioning of bioprocesses.

This document presents my contribution to enlighten on the functioning of microbial communities of anaerobic bioprocesses. It is divided in 5 parts. The first part introduces briefly the topic of my research. Three chapters summarize the main works carried out and results obtained. The last part presents the perspectives envisaged to extend and enhance the work already carried out.

General introduction

My research is at the crossroads of different disciplinary fields. Society issues (sustainable waste management, production of renewable energy, etc.) and various scientific disciplines (microbial ecology, molecular biology, omics, biostatistics, etc.) are mixed together. To put my work into perspective, this introduction will briefly present the issues of waste management and treatment in a global context, with a focus on anaerobic digestion. I will also introduce several microbial ecology methods used to further study and understand microbial ecosystems. I will conclude by mentioning different ways of dealing with the increasing amounts of data generated by these methods.

1. Microbial bioprocesses for waste treatment

1.1. Waste production and waste treatment in France

In 2015, waste production in France represented 324.5 million tonnes, of which 228 million tonnes for the construction sector, 62.5 million tonnes for economic activities outside construction, 30.6 million tonnes for households and 4 million tonnes for communities.

If production of household waste (waste generated by households and collected by the public service) had more than doubled in France between 1960 and 2000 (reaching more than one kilo per inhabitant per day), it started to decrease in 2002 thanks to incentive procedures. The current production however still forms a very important flow to manage. Four main elimination channels have been developed in France to deal with this flow (Figure 1): sorting for recycling, thermal treatments (incineration), biological treatments and landfilling.

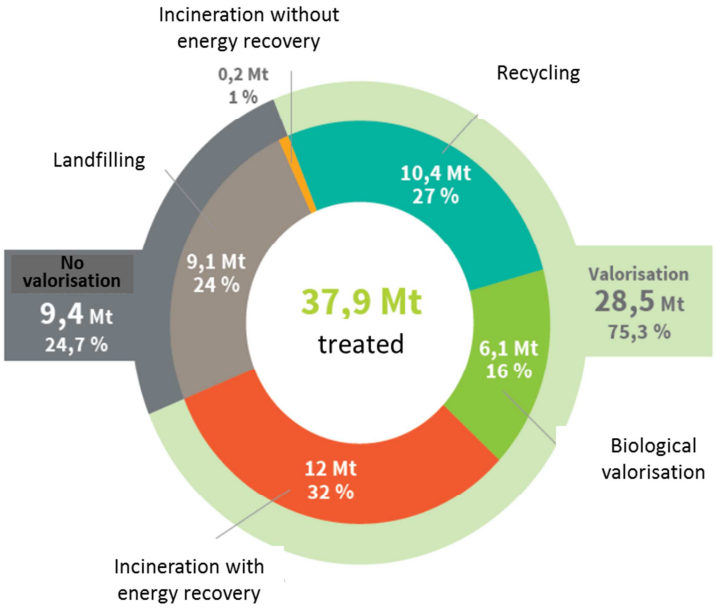


Figure 1: Treatment of household waste in France, in 2015 (from ADEME: déchets chiffres clés 2017).

In 2014, the sorting centres absorbed 27% of household waste. Incineration treated 33%. Biological treatments enabled to eliminate 16% of this mass. Finally, and thanks to the concomitant development of three other techniques, the use of landfill storage remained necessary to cope with only 24% of this flow of waste.

A greater attention to environmental issues has prompted decision-makers to reorient their waste management policy. More stringent discharge standards have been imposed. Waste is now considered as a resource that can potentially be valued physically, energetically and economically. There has been a growing interest in biological waste treatment. Indeed, an important fraction of household waste is composed of organic matter (Figure 2) and forms an ideal substrate for biological treatment processes. In total organic waste, excluding agricultural waste, amounts for 46.4 million tonnes, of which 20.2 million tonnes are collected by the public service every year in France.

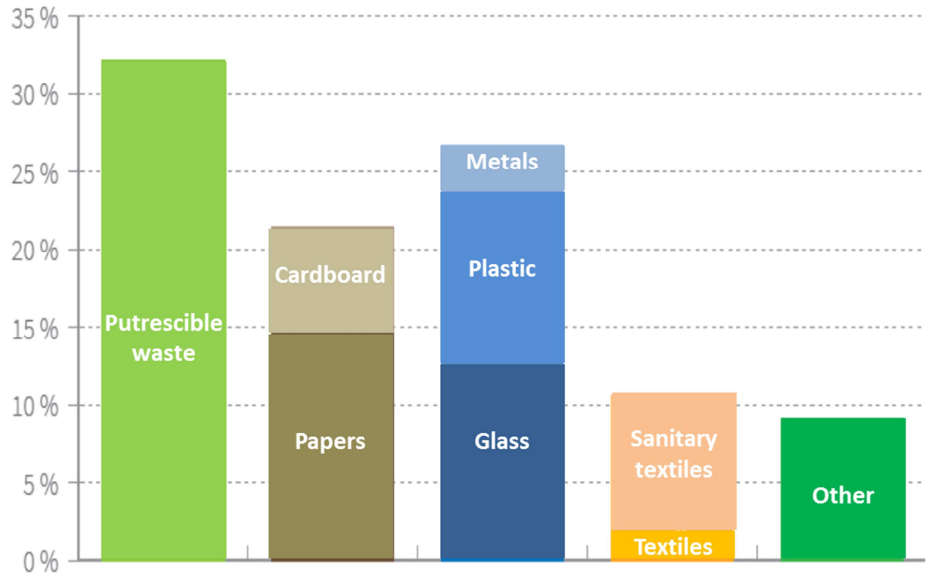


Figure 2: Household waste composition in France, in 2017 (from ADEME. MODECOM 2017).

1.2. Biological treatment of waste

There are two types of biological waste treatment: composting and anaerobic digestion.

Composting is an aerobic process of transformation of fermentable matter, under controlled conditions (aeration/mixing) in order to obtain a stabilized fertilizer rich in humic compounds, the compost. *I will not describe it further in this document as it is not the process I have been working on.*

Conversely, anaerobic digestion (or methanization) is an anaerobic process that can be implemented in anaerobic digesters, or biogas plant[1]. This bioprocess is based on the microbial degradation of the organic matter contained in waste. It produces on the one hand

[1] Anaerobic digestion also occurs spontaneously in natural ecosystems, where organic matter is present and oxygen is scarce, such as marshes, lakes, rice paddies, lacustrine and marine sediments, soil, mammalian gut, intestinal tract of ruminants and termites, etc.

biogas rich in methane that can be converted into valuable electrical and thermal energy; and on the other hand the digestate, a stabilized residue that can be in some cases used as a fertilizer (Figure 3). As it allows converting waste into energy resource, anaerobic digestion (AD) is highly relevant for environmental protection and for our quest to increase energy efficiency.

Almost any organic material can be processed with anaerobic digestion; however, the level of biodegradability is the key factor in the successful application of anaerobic digestion. For example, lignocellulosic material rich in cellulose and hemicellulose polymers as well as recalcitrant lignin are not broken down easily by anaerobic microorganisms and do not always provide a sufficient methane yield in anaerobic digesters. On the contrary, biowaste, such as leftover food, sewage sludge, animal waste, green waste, etc. are very adapted as they enable to achieve quickly high levels of degradation and biogas production. According to the type, source and volume of feeding, anaerobic digesters are designed and engineered using different configurations. Their size can vary from a few dozen of litres up to several dozen of cubic meters. They can be fed in batch or continuous mode, with waste containing a high or low portion of solids. They can be operated under different temperature conditions and with difference residence time of the waste in the digester. They are sensitive to different inhibitors and to the modification of the operational parameters, as well as the composition of waste. *The influence of some of these parameters is detailed in the paragraph 1.4.*

Figure 3: Anaerobic digestion is the microbial conversion of organic matter into valuable biogas and digestate that can be used as fertilizer.

In France in 2016, more than 600 biogas plants supplied nearly 4 TWh of heat and electricity, including methanization of household waste, methanization of sludge from wastewater treatment plants, industrial and agricultural methanization. Agricultural methanization

currently accounts for more than 50% of the plants and 30% of the energy produced. They are spread all over France (Figure 4). Biogas has been recognized as a renewable energy under the EC Directive (2001/77 / EC) since September 2001.

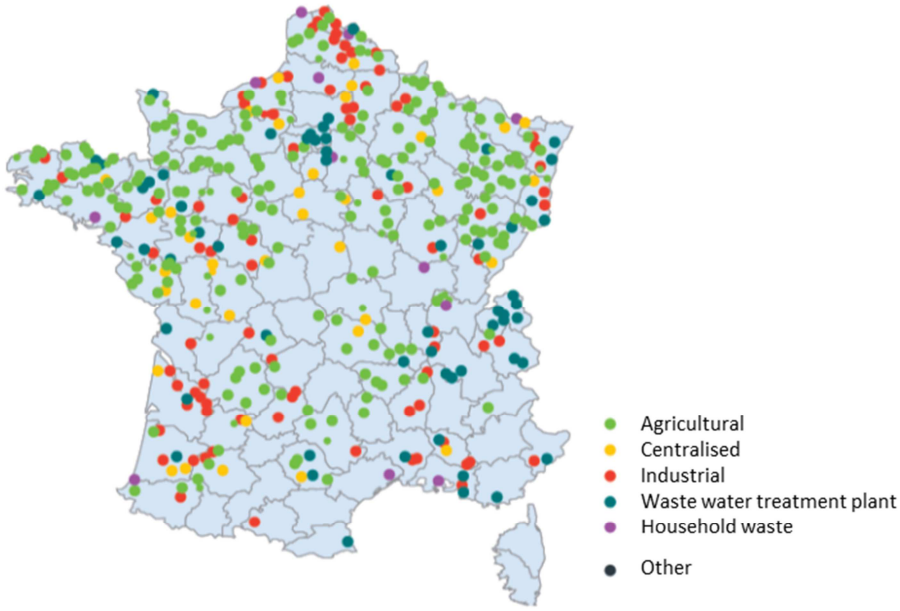


Figure 4: French park of methanization centres in 2016 (from ADEME: déchets chiffres clés 2017).

1.3. The anaerobic digestion

As illustrated on Figure 3, anaerobic digestion consists of the microbial degradation of the organic matter contained in the waste into biogas, composed mainly of carbon dioxide and methane. This simplistic representation does not illustrate the complexity of anaerobic digestion. Indeed, organic residues are composed of a wide variety of different materials meaning different types of molecules. Their degradation goes through many intermediates between the initial organic matter and the final biogas. It is catalyzed by a wide variety of microorganisms acting simultaneously, some of them feeding only on a very specific type of molecules, others being more ubiquitous. Because the biodegradation products of a reaction serve as substrates for the following reactions, important interactions between microbial populations exist. They form a complex microbial ecosystem within bioprocesses with an intricate network of interactions. To describe it more precisely, anaerobic digestion is often divided into four main stages: the hydrolysis, the acidogenesis, the acetogenesis and the methanogenesis (Figure 5).

1.3.1. The four main steps of anaerobic digestion

1.3.1.1. Hydrolysis

Hydrolysis consists of destructuring a substrate made of complex organic matter into simpler molecules that are easily assimilated by microorganisms. In this step, macromolecules such as nucleic acids as well as biopolymers (polysaccharides, lipids, proteins) are degraded (hydrolysed) into water-soluble fragments (monomers) such as monosaccharides, fatty acids, amino acids and nitrogen bases (Batstone et al., 2002; Cirne et al., 2007). These reactions are

catalyzed by exo-enzymes (Goel et al., 1998; Sanders et al., 2000; Zhang et al., 2007). The degradation rates depend on the nature of the substrates. This step is generally considered as limiting for the entire degradation process and especially when the organic matter is composed of complex molecules such as lignin or cellulose.

1.3.1.2. Acidogenesis

During acidogenesis, the monomers resulting from the hydrolysis step are mainly fermented into organic acids and alcohols. Volatile fatty acids (VFA) such as formate, acetate, propionate, butyrate or valerate, alcohols (propanol, butanol, ethanol ...) as well as other organic acids (caproic, succinic ...) are synthesized (Batstone et al., 2002). This step also leads to the formation of dihydrogen and carbon dioxide. It can be 30 to 40 times faster than hydrolysis (Moletta, 1993).

1.3.1.3. Acetogenesis

Acetogenesis allows the transformation of the various metabolic intermediates from the previous phases into acetate, dihydrogen and carbon dioxide. Two metabolic pathways have been identified:

- The heteroacetogenic pathway results in the production of acetate, dihydrogen and carbon dioxide. This pathway can also induce the production of VFA (propionate, butyrate ...).
- The homoacetogenic pathway leads only to the production of acetate. It can be synthesized either from organic molecules such as sugars, or degradation products such as formate, dihydrogen and carbon dioxide.

The reaction rates of acetogenesis are generally slow and subject to inhibition problems due to the presence of dihydrogen which modifies the thermodynamic equilibrium of the global kinetics. A very low partial pressure of dihydrogen is therefore necessary to make the production of acetic acid thermodynamically possible, the accumulation of dihydrogen leading to the interruption of acetogenesis (Siriwongrungson et al., 2007).

This phase is one of the key steps in anaerobic digestion. Indeed, any dysfunction at this level may result in an accumulation of hydrogen and VFA associated with a significant decrease in pH, resulting in the inhibition of the next step of methanogenesis (Stams, 1994).

1.3.1.4. Methanogenesis

Methanogenesis is the final step in the process of anaerobic digestion. It consists in the conversion of the products of acetogenesis (mainly acetate, dihydrogen, carbon dioxide and formate) in methane. Two metabolic pathways are involved during this last transformation:

- Acetoclastic methanogenesis produces methane and carbon dioxide from acetate.
- Hydrogenotrophic methanogenesis uses dihydrogen and carbon dioxide to produce methane

Other metabolites such as methanol or methylamine can also be precursors of methane (methylotrophic methanogenesis) (Lovley and Klug, 1983).

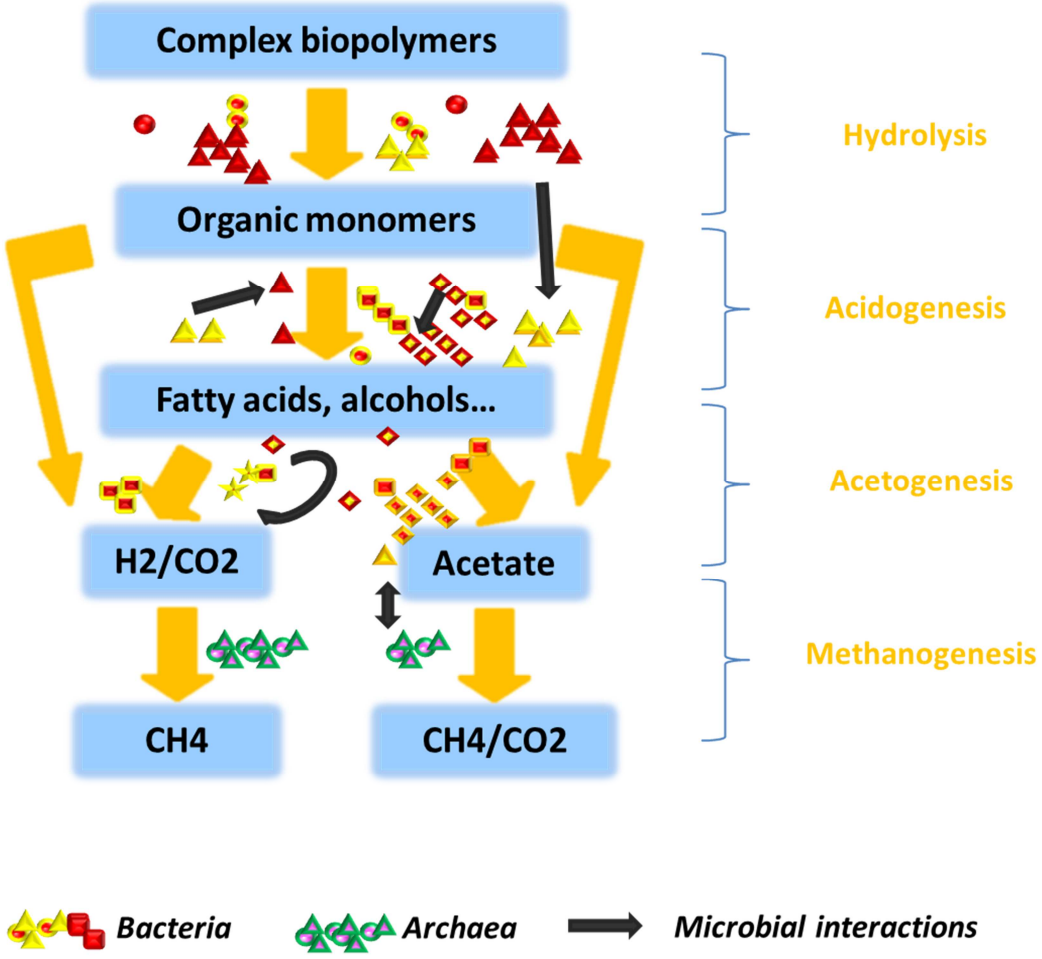


Figure 5: Anaerobic digestion is usually divided into four main steps. It is performed by a complex microbial community composed of interacting bacteria and archaea.

1.3.2. Multiples clades of microorganisms are involved

Key microorganisms driving the AD process form an extremely complex microbial community, mainly composed of bacteria and archaea, able to degrade the organic matter across multiple pathways, with lots of functional redundancies[2]. Those microorganisms collectively constitute the AD microbiome. They are generally specific of one of the previously described steps. The next paragraphs present the most important players of each step, based on the current knowledge.

1.3.2.1. Hydrolytic microorganisms

Hydrolysis is mainly performed by hydrolytic bacteria from various phylogenetic groups. In addition, some eukaryotic fungal groups found in the rumen are also able to hydrolyse complex organic matter (Nsereko et al., 2000). Hydrolytic bacteria are anaerobic, strict or optional. Their growth rate is relatively fast and their doubling time can reach a few hours although this step is often limiting in the overall methanization process. The extreme diversity

[2] Ecological phenomena that multiple species representing various taxonomic groups can share similar roles in the ecosystem.

of the microbial communities involved in this step makes an exhaustive inventory impossible. Nevertheless, among the many microbial genera involved in hydrolysis, it is possible to mention *Acetivibrio, Anaerovibrio, Bacillus, Bacteroides, Butyrivibrio, Cillobacterium, Clostridium, Lachnospira, Micrococcus, Mocrococcus, Peptococcus, Pseudomonas, Ruminococcus, Succinomonas, Staphylococcus* or *Syntrophomonas* (Moletta, 2015).

1.3.2.2. Acidogenic microorganisms

Acidogenic or fermentative microorganisms are strictly anaerobic. They have a rapid metabolism and growth rate, and can survive extreme pH or temperature conditions (Bayard and Gourdon, 2001; Moletta, 2015). The rapid metabolism of this trophic group compared to the later stages of anaerobic digestion (acetogenesis and methanogenesis) can be a source of inhibition. Indeed, an excess of acidogenic species compared to methanogens, caused for example by an overload of organic compounds, can lead to an accumulation of intermediate compounds such as acetate and dihydrogen, unfavourable to the growth of methanogens (Siegert and Banks, 2005; Moletta, 2015). Therefore, acidogenesis is a key step in anaerobic digestion whose reaction products must be degraded or extracted as quickly as possible. Acidogens are mainly represented by the genus *Clostridium*. However, the genera *Acetobacterium, Bacillus, Bacteroides, Pelobacter* and *Ulyobacter* as well as the family *Enterobacteriaceae* also include acidogenic bacteria (Ali Shah et al., 2014; Moletta, 2015).

1.3.2.3. Acetogenic microorganisms

Acetogenic bacteria are strictly anaerobic, very sensitive to pH and have very slow growth (Ali Shah et al., 2014). The two metabolic pathways of acetogenesis are carried out by two distinct bacterial populations.

- The heteroacetogenic pathway is performed by syntrophic microorganisms that necessarily produce dihydrogen (called OHPA for "Obligate Hydrogen Producing Acetogens"). The growth of OHPA requires a very low partial pressure of dihydrogen ($<10^{-4}$ atm). Indeed, dihydrogen has an inhibitory action on microbial growth and biochemical reactions (Moletta, 2015). It must therefore be constantly removed by a second microorganism in order to make acetate consumption possible (Schink, 1997; Huiliñir et al., 2008; Nie et al., 2008). OHPAs necessarily grow with microorganisms consuming dihydrogen. They have a rather slow growth (doubling time from 1 to 7.5 days) (Girguis et al., 2005) and are mainly found in the genera: *Syntrophobacter, Syntrophomonas, Syntrophus, Syntrophococcus or Syntrophosphora* (Ali Shah et al., 2014).
- Homoacetogenic acetate-producing bacteria are divided into two groups according to the origin of the acetate. The first group uses a carbon substrate (volatile fatty acids, ethanol, etc.) and are found in genera such as *Butyribacterium* and *Peptococcus.* The second group reduces carbon dioxide with dihydrogen. These microorganisms allow in particular the maintenance of non-inhibitory concentrations of dihydrogen. They are found within the genera *Acetoanaerobacterium, Acetobacterium, Acetofilamentum, Acetohalobium, Acetomicrobium, Acetothermus, Acetotitomaculum, Clostridium, Eubacterium, Sporomusa* or *Thermoanaerobacter* (Ali Shah et al., 2014; Moletta, 2015).

1.3.2.4. Methanogenic microorganisms

The production of methane is carried out by strictly anaerobic archaea. Methanogenic archaea belong to 7 phylogenetic orders: *Methanococcales, Methanopyrals, Methanobacteriales,*

Methanosarcinales, Methanomicrobiales, Methanocellales and *Methanoplasmatales* (Sakai et al., 2008; Paul et al., 2012; Borrel et al., 2013).

- Hydrogenotrophic methanogenic archaea live in syntrophic association with the hydrogen producing species (OPHA). They help to maintain a low pressure of dihydrogen necessary for the good progress of anaerobic digestion. Apart from the genus *Methanosaeta*, all the methanogenic archaea are able of carrying out this reaction. The most represented genera are *Methanosarcina*, *Methanobacterium*, *Methanobrevibacter*, *Methanospirillum*, *Methanogenium* and *Methanocorspusculum* (Ali Shah et al., 2014; Moletta, 2015).
- Acetoclastic methanogenic archaea are dominated by the order *Methanosarcinales* and in particular by two genera: *Methanosaeta* and *Methanosarcina*. *Methanosarcina* usually becomes dominant when acetate concentrations are above 5mM.
- The main methylotrophic archaea belong to the genus *Methanosarcina* (Ali Shah et al., 2014; Moletta, 2015), to the order *Methanobacteriales* (in particular *Methanosphaera stadtmanae* species) and to the order *Methanoplasmatales* (in particular *Methanomassiliicoccus luminyensis* and *Candidatus Methanomethylophilus alvus* species) (Borrel et al., 2013).

1.4. Parameters influencing anaerobic digestion equilibrium

The stability and efficiency of biogas production in anaerobic digesters highly relies on the stability of the complex microbial community described above (Pap et al., 2015). This equilibrium can be strongly influenced by the operational parameters of the AD. We describe some of them below.

1.4.1. Redox potential

The first steps of anaerobic digestion are performed by strict or facultative anaerobic bacteria. However, methanogenesis is performed by strictly anaerobic archaea. Oxygen has a bacteriostatic or even bactericidal effect on the methanogens. The redox potential represents the state of reduction of the system and can directly be related to the oxygen concentration of the medium. For anaerobic digestion to take place, it is generally assumed that the redox potential must be less than -300 mV.

1.4.2. The temperature

Despite the wide range of temperatures observed in natural AD, industrial anaerobic digesters are mostly operated at three levels of temperature: psychrophilic (4-15 °C), mesophilic (20-40 °C) and thermophilic (45-70 °C) (Van Lier et al., 1997), mesophilic and thermophilic temperatures being the most frequent. The operating strategies are chosen according to the treatment objectives but also the environment in which the digester is implanted. The advantages and disadvantages of thermophilic versus mesophilic conditions were described in many studies, e.g. (Kim et al., 2002; Labatut et al., 2014). Mesophilic processes were shown to be more stable and robust, requiring less energy for heating and generating fewer odours. They have less probability of inhibition by ammonium or long chain fatty acids (Fernández-Rodríguez et al., 2015) . However, thermophilic processes have the potential to produce biogas with a higher methane yield, to increase the degradation rate of organic solids (Lü et al., 2014) and to destroy more pathogenic organisms (Moset et al., 2015). For example thermophilic sludge digestion can improve energy balance and nutrient recovery potential in full-scale municipal wastewater treatment plants (De Vrieze et al., 2016a). Psychrophilic

digestion can be performed more easily, as no heating of the digesters is required, but has lower yields.

Temperature can affect the growth of microorganisms and enzymatic activity. An important part of the microorganisms can grow at only one temperature condition, and modifying the temperature of the process can strongly affect the microbiome and its functioning (Zhu et al., 2018). Numerous examples of process failure after temperature modification are available in the literature. For example, in response to temperature shocks in reactor, a decrease in specific anaerobic activity and reactor efficiency can occur, including lower organic carbon removal efficiency, accumulation of volatile fatty acids and decrease in biogas yield (Lau and Fang, 1997; Chae et al., 2008; Lindorfer et al., 2008; Kundu et al., 2014; Zhu et al., 2017). In all cases, authors observed strong modifications of digesters microbial communities and ecological dynamics after changes in the operating temperature (Noll et al., 2010; Pap et al., 2015; Chapleur et al., 2016a; Lin et al., 2016; Westerholm et al., 2017) and after restoration of temperature (Luo et al., 2015; Chapleur et al., 2016a).

1.4.3. pH and alkalinity

pH affects growth of microorganisms and enzymatic activity (Boone and Xun, 1987; Veeken et al., 2000). Methanogens are extremely sensitive to changes in this parameter. The pH modifies in particular the physicochemical equilibrium of acids and bases such as volatile fatty acids and ammonia nitrogen (Mata-Alvarez et al., 2000) or the liquid / gas equilibrium (for example carbonate/carbon dioxide). Optimum conditions for anaerobic digestion are around neutrality, between 6.5 and 8.5.

1.4.4. Inhibitors

Anaerobic microorganisms are vulnerable to a wide variety of inhibitory substances that can be either contained in organic waste or formed during their degradation within digesters (Chen et al., 2008a). Most reported inhibitors are ammonia, sulphides, heavy metals (Cu, Zn, Cr, Cd, Ni, Pb…), different ions (Na^+, K^+, Ca^{2+}, Mg^{2+}…), long chain fatty acids, N-substituted aromatics, halogenated aliphatics, phenols. Literature shows considerable variations in the inhibition/toxicity levels reported for most substances. The major reason for these variations is on the one hand the complexity of the anaerobic digestion process where mechanisms such as antagonism, synergism, acclimation, and complexing could significantly affect the phenomenon of inhibition (Chen et al., 2008a) ; and on the other hand the specificity of the studies, using different types of reactors, different inocula, not always taking into account the synergistic effect of multiple inhibitors on the process…

1.5. What remains to be done

In a context of environmental protection and research for increasing energy efficiency, methanization arouses a renewed interest because it allows converting waste to both an energy resource and an organic amendment. However, some obstacles still hamper its development and prevent it to be commercialized as widely as it could. AD is not fully controlled and still has an important potential for improvement (Carballa et al., 2015). A major limitation is the important susceptibility of the microbial communities to changes in operational conditions of the digesters (some of them listed above) that can lead to unstable methane formation. Despite the functional redundancy within digesters, the inhibition of certain key sensitive populations such as methanogenic archaea automatically results in decreased production yields and potentially in process failure leading to significant economic losses. Furthermore, the dynamics of the populations and industries cause sudden and unpredictable changes of

quantity and composition of waste to treat. Until recently, the methanization was considered not being well enough mastered to face these changes and present the desired flexibility.

Controlling AD microbial community stability is not a trivial task. Experience feedback from industrial operators shows that digesters remain little instrumented. Their successful operation relies on the know-how of the developer rather than on objective criteria. It strongly limits the standardisation, transfer and broad use of successful operational strategies. Such situation is mainly due to the limitation of microbial-based management of anaerobic reactors. Indeed, the microbiome, which is the key player of the AD process, has remained largely unknown for a long time and was a "Black box". However, efficiency and good functioning of AD totally depend on the complex balance of AD microbiome. Microbial-based management of anaerobic reactors is a major hurdle to better control and improve AD.

To build such management strategies, a better knowledge of the functioning of AD microbiome, of the influence of operational parameters on microbial equilibrium and of the processes leading from microbial balance disruption to process macroscopic failure is required. The rather recently developed microbial ecology methodologies now allow to better describe complex microbial communities. They provide the information necessary to perform functional microbiome diagnostics (Vanwonterghem et al., 2014a) which could help to build real microbial management strategies of anaerobic digestion and to set microbial indicators of optimal performance and warning indicators of process failure (Koch et al., 2014). Those indicators could be key enzymes, inhibitors or key microbial population for which on-line abundance monitoring techniques are currently being developed, for instance based on flow cytometry analyses or sequencing (Broger et al., 2011; Hammes et al., 2012; Koch et al., 2013). Better consideration of the functioning of the microbial ecosystem would enable to refine the management and control of microbial processes for waste degradation and biogas production and to improve their stability.

My research work aims at this objective.

2. Microbial ecology

2.1. The first approaches

The first attempts to study microbial complex communities (including microbiome from AD) relied on conventional cultivation techniques, implying the isolation of pure cultures. This process was time-consuming. All isolated microorganism required numerous physiological and biochemical tests for full characterization. Moreover, it is now estimated that less than 1% of microbes are cultivable (Amann et al., 1995). The study of complex microbial ecosystems was therefore biased. The advent of PCR (polymerase chain reaction) and the identification of several genetic biomarkers have has shattered the methodological barriers associated with culture-based methods, as microorganisms can now be studied in-situ, without isolation and cultivation.

The gene coding for the small 16S subunit of the ribosomal RNA (16S ribosomal RNA gene, Figure 6) is one of the most used biomarker to study prokaryotes (bacteria and archaea). Its sequence is used to identify microorganisms and to reconstruct phylogenies. Indeed, ribosomal RNA is a highly conserved molecule, which means that it is present in all prokaryotes. 16S rRNA gene also contains hypervariable regions that can provide species-specific signature sequences useful for identification (they form a molecular identity map to identify microorganisms). Highly conserved sequences between hypervariable regions can be

used to design universal PCR primers. It enables to reliably produce the same sections of the 16S sequence across different taxa for comparison purpose.

Carl Woese and George E. Fox pioneered the use of 16S rRNA in phylogenetics. Using this tool, they demonstrated in 1987 that organisms were divided into three major domains (Eubacteria, archaea, Eukarya), thus rejecting the classical dichotomy between prokaryotes and eukaryotes.

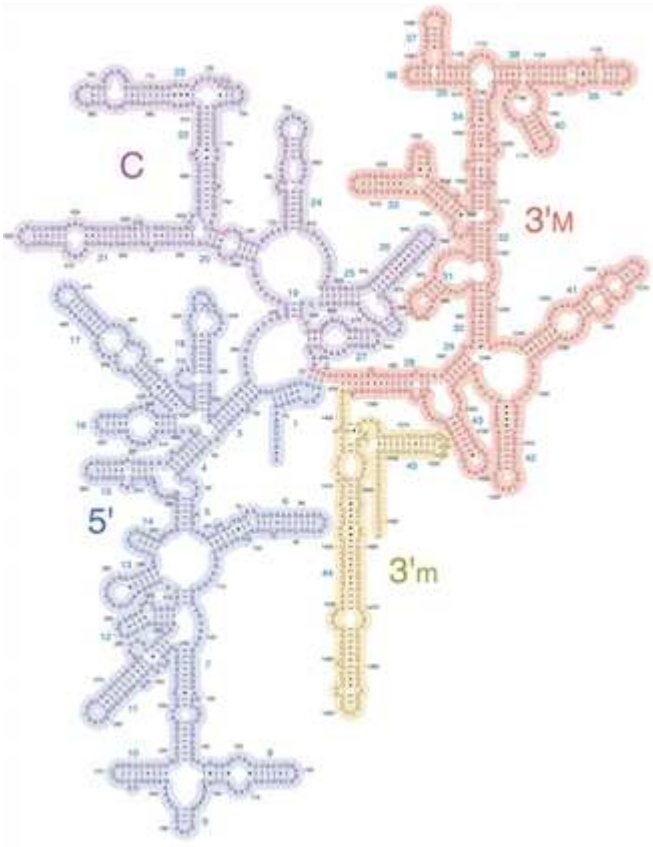


Figure 6: Secondary structure of the 16S RNA molecule (from rna.ucsc.edu/).

During the last thirty years, new techniques have been developed, bringing together different molecular tools to study the diversity and ecology of microbial populations without going through isolation petri dishes. All these techniques, based on the genetic characterisation of the microorganisms composing the medium, rely to a large extent on the properties of the ribosomal RNA molecule and make it possible to analyse complex and natural samples directly. *I detail some of them in the next paragraphs, mainly those I had the opportunity to use during my work.*

2.1.1. Fingerprinting

Complex microbial ecosystems consist of a large diversity of interdependent individuals interacting with each other. It is important to study the microorganisms of these systems in a global way, especially their relative dynamics over time. To this end, several techniques have been developed to monitor the general evolution of microbial communities. They are grouped under the name of fingerprinting techniques.

These methods are generally based on the differential electrophoresis migration of ribosomal DNA fragments previously amplified by PCR. The variability of the microorganism within the community leads to a variability of the PCR products (size or sequence of the amplified fragment) that enables to differentiate the members of the microbial community. Resolution of the PCR product with electrophoresis produces a community profile, or genetic fingerprint, a kind of barcode that characterizes the diversity of a sample at a given moment. These barcodes can be compared to evaluate the dynamics of microbial diversity within a microbial ecosystem, or the diversity of one sample relative to another. They can be associated with variations in operating conditions.

Numerous fingerprinting techniques were developed, such as for example:

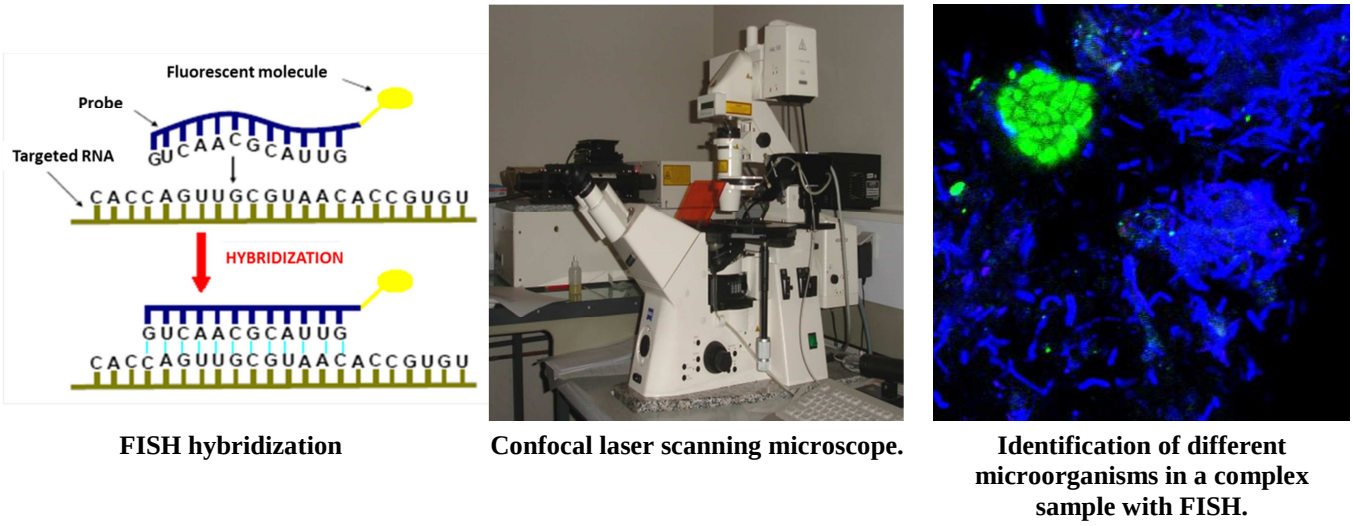
- denaturing gradient gel electrophoresis (DGGE) (Muyzer et al., 1993) which is based on the migration of amplicons in an acrylamide gel containing a denaturing gradient. It separates amplicons of the same size based on their different denaturing ability which is determined by their base pair sequence. The resulting bands can be cut and sequenced to identify the corresponding microorganisms.
- single strand conformation polymorphism (SSCP) (Orita et al., 1989) which is based on the denaturation of double stranded amplicons into single strands of DNA. Depending on their sequence, the single strands will adopt different 3D conformations and will have a differential migration in electrophoresis.
- restriction fragment length polymorphism (RFLP) (Laguerre et al., 1994; Moyer et al., 1994) which is based on the variability in the size of rRNA fragments generated by a restriction enzyme. After amplification of the 16S gene of a complex microbial sample, enzymatic digestion is performed by one or several restriction enzymes, chosen for their cut-off frequency. The fragments obtained can be analysed by gel or capillary electrophoresis, thus giving the fingerprint of the community.
- automated ribosomal intergenic spacer analysis (ARISA) (Garcia-Martinez et al., 1999; Ranjard et al., 2000) which is based on the observation of the size of the intergenic transcribed spacer (ITS) present between the genes encoding 16S rRNA and 23S rRNA. This region has a very high heterogeneity in terms of size and composition of nucleotide bases. After PCR amplification of this zone, the amplicons are separated by electrophoresis. The heterogeneity of the lengths of the ITS from one species of microorganism to another makes it possible to reveal small modifications in the genetic structure and thus to distinguish phylogenetically closely related species.

Even if they do not directly give information on the identity of the microorganisms involved, these techniques have been used for many years for their speed and repeatability, but also because they offered at a reasonable cost, an overview of the community of a sample.

2.1.2. Fluorescent in-situ hybridization

Fluorescent in-situ hybridization (Amann et al., 1995) enables to visualize in-situ specific populations. This technique is based on the hybridization of oligonucleotides, called probes, on a complementary nucleotidic sequence directly inside the cell. The probes can be more or less generalist, i.e. attach to a wider or more restricted number of microorganisms (species, orders, phyla ...). A fluorescent dye is attached to the probes. After hybridization, the samples are observed using an epifluorescence microscope or with a confocal laser scanning microscope. Microorganisms hybridized within the sample can thus be identified (Figure 7). Different probes targeting different signature sequences and labelled with different fluorescent molecules can be used in the same hybridization to study several types of

microorganisms at the same time and observe, for example, colocalisation of syntrophic microorganisms.



FISH hybridization

Confocal laser scanning microscope.

Identification of different microorganisms in a complex sample with FISH.

Figure 7: Fluorescent *in situ* hybridization.

2.1.3. 16S metabarcoding

The identification of members of microbial communities with sequencing techniques provides access to a higher level of information than fingerprinting techniques. For cost reasons, sequencing used to be reserved to very specific and selected samples. In recent years, it has benefited from significant progresses that have drastically reduced its cost while exponentially increasing the amount of sequences produced. High-throughput sequencing (HTS) methods, also called NGS (for next-generation sequencing) are more and more used as a standard, and progressively replace fingerprinting approaches.

Targeted metabarcoding is the most widely used high-throughput methodology to infer the structure and composition of microbial communities. For archaea and bacteria, it is based on the amplification of a fraction the 16S gene (several hundreds of base pairs), followed by sequencing. Amplified region can vary according to the type of microorganisms targeted. The microbial 16S gene contains nine hypervariable regions (V1-V9) with a length of approximately 30-100 base pairs each, as illustrated on Figure 8.

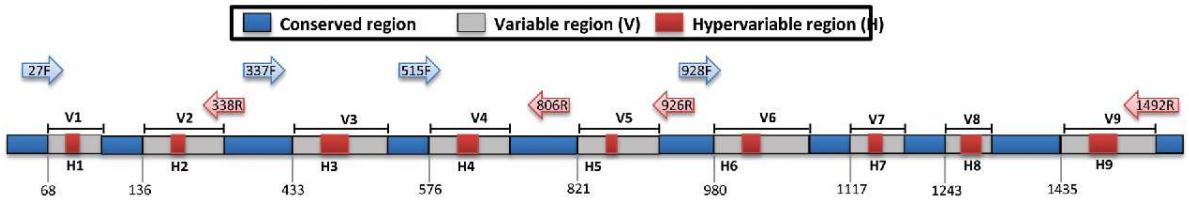


Figure 8: Conserved, variable, hypervariable regions within the 16S rRNA gene and the various primers used for amplification. Conserved regions are represented in blue, variable regions in gray, and hypervariable regions in red. (adapted from (Shahi et al., 2017)).

The degree of conservation of these different regions varies widely. The most conserved regions are common within high levels of taxonomy while the less conserved regions are specific of genus or species. Although no hypervariable region can accurately classify all bacteria from Domain to Species, some can reliably predict specific taxonomic levels. Many community studies select semi-conserved hypervariable regions like the V4 for this reason, as it can provide resolution at the phylum level as accurately as the full 16S gene (Yang et al., 2016).

16S metabarcoding nowadays constitutes a standard tool for microbial ecologists. It still has several limits. It struggles to differentiate between closely related species (Větrovský and Baldrian, 2013). For example, in the families *Enterobacteriaceae*, *Clostridiaceae*, and *Peptostreptococcaceae*, species can share up to 99% sequence similarity across the full 16S gene (Jovel et al., 2016). As a result, the V4 sequences can differ by only a few nucleotides, leaving reference databases unable to reliably classify these bacteria at lower taxonomic levels (Jovel et al., 2016). In these cases, the use of other markers such as DNA gyrase subunit B (*gyrB*) (Poirier et al., 2018b), RNA polymerase subunit B (*rpoB*) (Case et al., 2007), DNA recombinase protein (*recA*), protein synthesis elongation factor-G (fusA), or dinitrogenase protein subunit D (*nifD*) (Holmes et al., 2004) could improve the assignation and abundance estimation for crucial taxonomic groups.

Metabarcoding is often based on DNA amplification (16S RNA gene) rather than 16S RNA. It can lead to incorrect interpretation of the community parameters, as microbial abundance does not necessarily reflect activity (De Vrieze et al., 2016b). In different studies, I observed that the difference between microbial community response on 16S RNA (active community) and 16S RNA gene (microbial diversity present in the digester, even if not active) could be important, the response time of RNA being much faster than that of DNA regarding both increase and decrease of abundance, in particular under stress conditions. In some case, RNA-based approach can thus be preferred to obtain reliable information on actual community parameters (De Vrieze et al., 2016b).

Finally, even if archaea and bacteria are often referred as the most abundant microorganisms in AD, authors evidenced that fungi can also play an important role in AD (Sun et al., 2015) even if less investigated (Theuerl et al., 2019). Similar approaches using 18S RNA gene or ribosomal ITS regions as biomarker instead of 16S can target their abundance (Sun et al., 2015).

2.2. New developments linked to high-throughput omics

"Omics" approaches have recently been developed to increase the level of information obtained through conventional microbial ecology techniques. They consist in apprehending the complexity of the microbial ecosystem as a whole, using the least restrictive methodologies possible. They can target the different ‘layers’ of the biological systems: DNA, RNA, proteins, metabolites. The four tools which are commonly encompassed under the heading (meta)omics are metagenomics, metaproteomics, metatranscriptomics and metabolomics. The advent of these methodologies will provide microbiologists with the tools to represent portraits of a community’s genes, gene expression, and metabolite production in a single sample, providing great insights into the understanding of microbial communities of complex environments, in particular bioprocesses. For example, Sales et al. reviewed different application of metaomics to enhance the recovery of carbon and phosphorus as well as energy from wastewater (Sales and Lee, 2015). Methodological frameworks for linking bioreactor function to microbial communities and environmental conditions are progressively being proposed (De los Reyes et al., 2015; Sales and Lee, 2015).

2.2.1. Metagenomics

Metagenomics is the random sequencing of genomic DNA fragments extracted directly from a microbial community inhabiting a natural or engineered environment. It allows access to the encoded functional traits and identities of the members of the community. It goes beyond 16S rRNA gene based characterization by providing insight into the physiological potential of the community. It enables to recover new enzymes through function-based analyses.

In AD systems, current studies aim at reconstructing important metabolic pathways and genomes (Badalato, 2014; Vanwonterghem et al., 2014a; Maus et al., 2016). It does not only reveals great phylogenetic and functional diversity (Maus et al., 2016), but it can also help understand the observations made with metabarcoding on how reactor set-up and operational conditions influence the community composition and function (Wong et al., 2013).

2.2.2. Metatranscriptomics

Metatranscriptomics involves the sequencing of reverse transcribed mRNA extracted from a microbial community, to measure *in situ* gene expression. Compared to metagenomics it focusses only on community members inferred to be metabolically active. Only a few studies have been published so far in the field of AD. A PhD was completed in our lab on the "Development and implementation of metatranscriptomic tools" and set-up the different steps of the protocol (Dr Marandat, 2012-2015). In more recent work, Delforno et al. used a bioinformatics pipeline developed at Irstea to compare the microbial active community in anaerobic digesters treating anionic surfactant contaminated wastewater (Delforno et al., 2019). Interestingly, Maus et al. observed that high abundance of microorganisms as deduced from metagenome analysis did not necessarily indicate high transcriptional or metabolic activity, and vice versa, stressing the complementarity of both approaches (Maus et al., 2016).

2.2.3. Metaproteomics

Metaproteomics is the characterization of the proteins extracted from a microbial community under a given set of conditions at a specific time point. Metaproteomics can be used to detect catalytic enzymes, entire metabolic pathways and novel functional proteins. Metaproteomics can also be used to identify the distribution of metabolic activities among a community and how populations cooperate or compete (Lü et al., 2014; Bize et al., 2015).

Identification of the proteins still relies on the presence of their sequence in reference databases but the exponential increase in database size should reinforce the attractiveness of this approach in the future. Moreover, metaproteomics can be used in combination with metagenomics to complement or create database of enzymatic and other protein functions (Hanreich et al., 2013).

2.2.4. Metabolomics

Metabolomics approaches provide a qualitative and quantitative measure of all low molecular-weight molecules contained in a sample. These molecules can be involved in cellular metabolic reactions and required for the maintenance, growth and normal function of microbial community. In the case of AD, they can arise from the degradation of organic matter by the microorganisms. Metabolomics gives access to informative details about key metabolic pathways (Vanwonterghem et al., 2014a). Application of whole community metabolomics to anaerobic digesters is difficult due to the very large range of metabolites with limited a priori knowledge (Vanwonterghem et al., 2014a). Until now, mainly specific

metabolites or regulatory compounds were analysed (De Kok et al., 2013). Development of new high-resolution instrumentation enables untargeted approaches.

The data generated by these broadband technologies opens the way for new approaches no longer based on an initial hypothesis that needs to be verified (so-called "Hypothesis driven") but guided by the data generated ("Data Driven" approach). It is from the analysis of these data generated without a priori that hypotheses should be taken into account and studied. The probability of highlighting an operation that would not have been thought of as the hypothesis is thus higher and opens the way to the development of disruptive technologies.

3. Treating big data sets with biostatistics approaches

The decrease of the cost of sequencing and the advent of omics studies generate an increasing amount of massive data obtained from different analytical devices. Specific bioinformatics tools are required to convert this raw data into processed data ready to be analysed. Databases of sequences or proteins are also required to identify the microorganisms or functions at stake. These tools are constantly evolving and improving. I will not present them in this manuscript.

The data obtained after bioinformatics treatments are usually high dimensional and extracting information from them is challenging. Appropriate tools are therefore needed to handle these datasets suitably. Classical multivariate ordination methods have been widely used to treat fingerprinting, 16S or omics data. They enable to reduce the dimension of the data, and to summarize hundreds of initial variables into a few, easier to explore and visualize. They highlight the level of similarity or dissimilarity of individuals within a dataset. Among these methods we can cite:

- The principal component analysis (PCA) which is a dimensionality-reduction method. It identifies the largest sources of variation in a dataset. It uses orthogonal linear transformation of data to form principal components. The first principal component explains as much of the variability in the data as possible, and each following principal component explains as much of the remaining variability as possible. The results of the PCA enable to highlight the covariances which exist between the different individuals by a simple observation of the Euclidean distance.
- Principal coordinates analysis (PCoA, also known as metric multidimensional scaling MDS) is a generalization of principal component analysis. It attempts to represent the distances between samples in a low-dimensional, Euclidean space. It first calculates a distance matrix between the samples and then maximizes the linear correlation between the distances in the distance matrix, and the distances in a space of low dimension, just as PCA. Any dissimilarity coefficient or distance measure may be used to build the distance matrix used as input (ex: Bray-Curtis distance). When the distance metric used is Euclidean, PCoA is equivalent to Principal Components Analysis.
- The non-metric multidimensional scaling (NMDS) is a ranking method. It also uses a matrix of distances within samples to collapse information from multiple dimensions. Contrary to PCA and PCoA, NMDS is an iterative algorithm. NMDS routines often begin by random placement of data objects in ordination space. The algorithm then begins to refine this placement by an iterative process, attempting to find an ordination in which ordinated object distances closely match the order of object dissimilarities in the original distance matrix. The stress value reflects how well the ordination summarizes the observed distances among the samples.

Many other methods such as correspondence analysis (CFA), independent component analysis (ACI) can also be used to address multivariate datasets from bioprocesses (Glassey, 2013).

However, above described methods are limited in their findings. For example they do not permit to account for time or experimental effect, or to treat several types of data at the same time. New methods are being developed. Some of them and the possibility they offer will be presented in chapter 3.

Chapter 1: Targeting specific degradation processes with stable isotopes

One of the biggest challenges that microbial ecologists face is to identify which microorganisms are carrying out a specific set of metabolic processes in the natural environments; that is, who is doing what and under which conditions. Anaerobic digesters, as many other microbial processes, rely on synergistic interactions of a huge variety of microorganisms. In this complexity, understanding which functional guilds are involved in the degradation processes of specific molecules is not straightforward, and identifying the members of these different guilds is even more complex.

At the end of the 1990s, the best way to address the question of the link between identity and function remained culture. Microorganisms of a targeted environment were first isolated and cultivated in the laboratory, using growth media that contained a desired substrate. A subsequent characterization of the isolated microorganisms was performed at the physiological, biochemical and genetic levels. The metabolic properties of these microbial isolates could then be used to infer their potential roles *in situ* in the environment as well as the role of related microorganisms. An important limitation of this approach was that most prokaryotic species that are present in the natural environment are not cultivable in the laboratory using traditional cultivation methods. Indeed it was shown that only 0.1 to 10% of the phylogenetic groups that are widely distributed in the environment are cultivable (Amann et al., 1995) and this fraction could not be reflective of microbial diversity as some microbial families are less culturable than others (Johnson, 1998).

Advances in molecular techniques provided alternative strategies for microbial ecologists to characterize organisms within particular habitats (Ranjard et al., 2000). PCR-based methods enable to get information on the microorganisms independently of cultivation. In particular, targeting of the 16S ribosomal RNA subunit or the corresponding genes became an established and robust mean to describe the phylogenetic diversity of microbial communities and extended considerably the knowledge of complex microbial environment such as anaerobic digesters. However, the rRNA sequences that revealed a remarkably vast microbial diversity generally provided few direct clues regarding the interactions and metabolic capabilities of the microorganisms that these sequences represent. Thus, the fundamental question remained as pertinent as ever: which functions are attributable to which microorganisms in the natural environment?

Use of stable isotopes enabled a giant step to evaluate the fate of specific compounds and to probe functional microorganisms. Several methods using them were specifically developed to target specific mechanisms within complex environments. The first part of my research work, mainly during my PhD, was dedicated to the application of these methodologies to further describe and understand anaerobic digestion, in particular after modifications of operating parameters. This chapter presents the state of the art of these methods, as well as the main results I obtained.

1. Stable isotope labelling to trace functionality

1.1. Principle and state of the art

1.1.1. Stable isotopes

Of the known chemical elements, 54 elements have more than one stable nuclide, or stable isotope (not radioactive). The stable isotopes of one element differ by the number of neutrons they contain. Elements found in molecules of natural systems consist in a mixture of the different isotopes in very reproducible proportions, one of them being generally more abundant than the others. Main elements composing organic matter (carbon, hydrogen, nitrogen, oxygen, sulfur…) have more than one stable isotope. For example, carbon has two stable isotopes: $_{12}C$ (natural abundance of 98.9%) and $_{13}C$ (natural abundance of 1.1%). Nitrogen has 2 stable isotopes: $_{14}N$ (natural abundance of 99.6%) and $_{15}N$ (natural abundance of 0.4%). Oxygen has 3 stable isotopes: $_{16}O$ (natural abundance 99.7%), $_{17}O$ (natural abundance 0.2%) and $_{18}O$ (natural abundance 0.037%). The property that one of the isotopes is predominant is of great interest, as it opens the possibility to use the others as tracers.

Indeed, numerous molecules can be synthesized and artificially enriched in one rare isotope. Stable isotopes do not suffer from legal restrictions and health problems associated with radioisotopes. A labelled molecule can thus be placed in the presence of a complex community of microorganisms directly in their environment. Its assimilation or degradation will result in the gradual enrichment of the microorganisms involved in its metabolism in stable isotopes, and the production of degradation by-products themselves enriched.

Different techniques were developed to trace the fate of the labelled molecules in microbial ecosystems. Isotope-ratio mass spectrometry (IRMS) allows the precise measurement of the relative abundance of isotopes in molecules of a given sample. It is a powerful analytical technique to monitor the degradation of enriched molecules and identify the degradation products. Molecular biology methods, grouped under the name of ‘Stable isotope probing’ (SIP), aim at identifying active microorganisms by the selective recovery and analysis of isotope-enriched cellular components. They are presented in the next section.

1.1.2. Stable isotope probing

According to the cellular component targeted, different approaches are used.

1.1.2.1. Different molecules, different methods

The SIP-DNA technique has been formalised by Radajewski et al. in 2000. It is based on the fact that carbon or nitrogen isotope labelling increases the density of DNA (Meselson and Stahl, 1958). Recovering denser DNA can thus give access to the identity of the active microorganisms. For that purpose, after isolation, DNA is subjected to caesium chloride (CsCl) buoyant density-gradient centrifugation (Radajewski et al., 2002). The solution of caesium chloride subjected to a strong centrifugal force has the property of forming spontaneously an extremely resolving density gradient. Nucleic acids placed in this gradient will be distributed according to their density, which makes it possible to separate the heavy nucleic acids from the light. Labelled nucleic acids, from microorganisms involved in the degradation of the labelled substrate, can then be sequenced. It is therefore possible to identify microorganisms responsible for a particular process in a complex environment.

Manefield et al. reasoned that RNA could serve as a more responsive biomarker than DNA for use in SIP. Indeed RNA molecules are present in higher copy numbers, show higher turnover rates, and are produced independently of cellular replication. (Manefield et al., 2002). This increased sensitivity enables to reduce the amount of (costly) labelled substrate used or the length of the incubation required to obtain a link between metabolic function and taxonomic identity. Native RNA having a higher buoyant density than DNA, caesium trifluoroacetate (CsTFA) is preferred to create the density gradient.

Microbial populations responsible for sulphate-reduction coupled to acetate oxidation in estuarine sediments were identified by targeting the ^{13}C enrichment of the polar lipid derived fatty acids (PLFAs) with an isotope ratio mass spectrometer after growth on ^{13}C labelled acetate (Boschker et al., 1998).

SIP was also applied to proteins (Jehmlich et al., 2008). The authors developed a method to analyse the specific metabolic activity of a single bacterial species within a consortium making use of [^{13}C]-toluene for metabolic labelling of proteins. Labelled proteins were subsequently analysed by 2D gel electrophoresis and mass spectrometry to characterize their identity as well as their ^{13}C content as an indicator for function and activity of the host organism.

1.1.2.2. Limits

Despite these methods are really powerful to target specific processes, they present several limits.

The principal consideration for determining whether SIP will be suitable for investigating a specific process is whether the nucleic acids (or other biological molecule) of the target organisms will contain a sufficient proportion of labelled atoms to permit the collection of the heavy nucleic acid fraction. For example, microbial non-labelled DNA with a G/C content of 35–70% ranges in density of approximately 1.69–1.73 $g.cm^{-3}$ (Radajewski et al., 2003), whereas the calculated buoyant density for the same DNA with a ^{13}C content of 100% is approximately 1.75–1.79 $g.cm^{-3}$. The more enriched the DNA is, the more resolved the separation will be.

In this way, long incubation periods (>40 days) have been necessary before labelled DNA extracted from soil samples can be correctly observed in a CsCl gradient (Radajewski et al., 2000). However, any such incubation results in the formation of labelled products and intermediates of substrate metabolism, which can then be assimilated by non-target microorganisms (cross-feeding) and bias the results. In order to help circumvent these weaknesses, careful experimental design with different incubation times can help detect and minimize cross-feeding (Radajewski et al., 2000).

Another risk is that target organisms grow on both labelled and naturally occurring (non-labelled) substrates, resulting in moderate stable isotope enrichment.

1.2. Examples of research outputs associated to these methodologies

1.2.1. General objectives and methods

I mainly applied Stable Isotope Probing methodologies during my PhD. Research questions targeted were associated to the methanization of cellulose, used as a model substrate. Lignocellulose is indeed a major component of the biosphere (Lynd et al., 2002). An important amount of this bioresource is available and it is also the most voluminous type of

waste produced by our societies (Bayer et al., 2007). The objective of my work was to better describe the microbial communities responsible for cellulose degradation and to assess the possibilities to modify these communities with environmental levers. For example 1) I evaluated the possibility to add specialized microbial flora in anaerobic digesters (bioaugmentation) to improve cellulose degradation (Chapleur et al., 2014) 2) I evaluated the influence of temperature changes on the cellulolytic community (Chapleur et al., 2016a), 3) I evaluated the possibility to model an initial inoculum into different cellulolytic communities through preliminary incubations with simple substrates (Chapleur, 2012).

In all these experiments, ^{13}C-labelled cellulose was used to probe microorganisms responsible for cellulose degradation. Experiments were performed in laboratory batch digesters (50 mL of liquid phase). I mainly used DNA-SIP protocol. I successfully applied the RNA-SIP protocol on mixed ^{13}C-labelled pure cultures, but the amount of RNA recovered after the processing of environmental samples was never sufficient to perform the downstream applications. ^{13}C-labeled cellulose degradation and reaction pathways were monitored over time through Gas Chromatography–Combustion–Isotope Ratio Mass Spectrometry (GC-C-IRMS) of biogas (methane and carbon dioxide) and volatile fatty acids (acetic and propionic acids). Analytical framework is summarized in Figure 9.

In this manuscript I will focus on one example (bioaugmentation experiment (Chapleur et al., 2014) and illustrate the different contributions of SIP to target ^{13}C cellulose degradation.

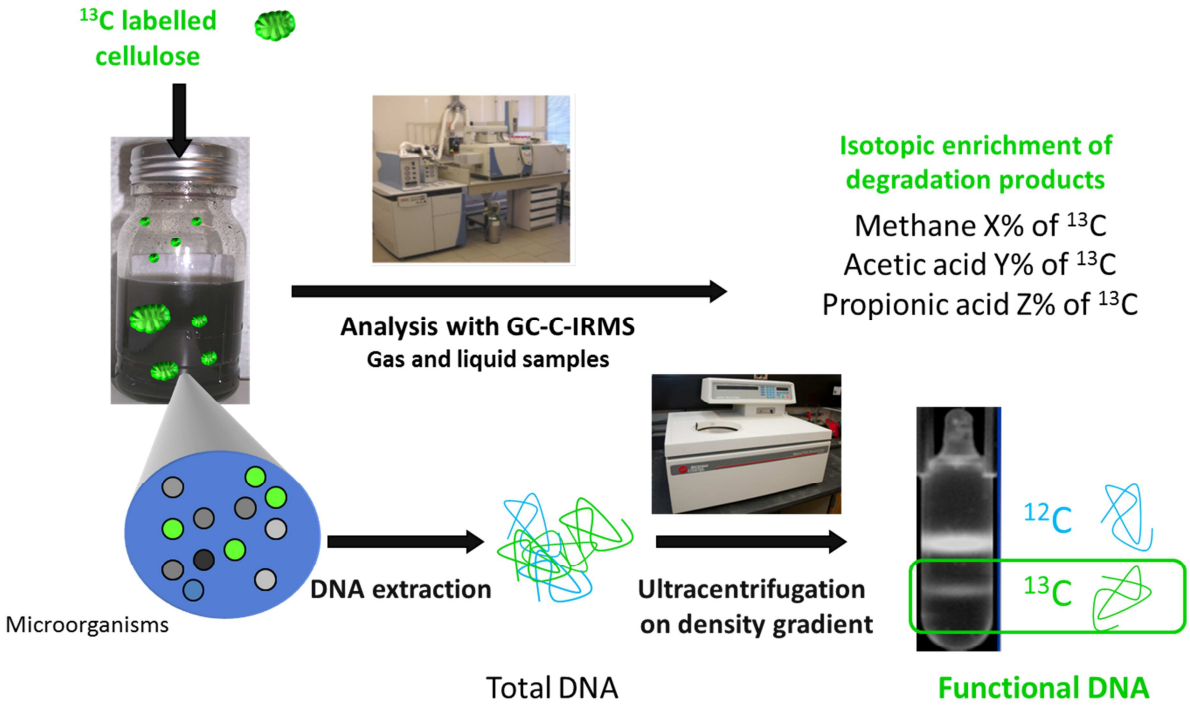


Figure 9: Analytical framework employed for experiment targeting the degradation of ^{13}C labelled cellulose. Cellulose was introduced in anaerobic digesters. Gas and liquid samples were analysed with GC-C-IRMS to monitor the fate of the cellulose (degradation products). Functional microorganisms progressively enriched in stable isotope (green). After DNA extraction, labelled DNA was isolated with ultracentrifugation on density gradient.

1.2.2. Bioaugmentation of anaerobic digesters with ruminal content

1.2.2.1. Objectives of the research

Cellulose hydrolysis often limits the kinetics and efficiency of anaerobic degradation in industrial digesters. In animal digestive systems (in particular ruminants or termites), specialized microorganisms enable cellulose biodegradation at significantly higher rates despite much lower residence times than in digesters (Weimer et al., 2009). Several studies involving ruminal microorganisms in anaerobic digesters have therefore been conducted to enhance methane production and the rate of cellulose degradation (Gijzen et al., 1987; Blasig et al., 1992; Barnes and Keller, 2003). They theorized that inoculating reactors with ruminal content increased cellulose solubilisation rates in the digestion processes of lignocellulosic waste. However, macroscopic observations suggest that ruminal microbiota may not easily settle in industrial anaerobic digesters and outcompete native solid waste microorganisms (O'Sullivan and Burrell, 2007). In this context I performed a study with the objective to assess the potential of ruminal microbial communities to settle and to express their cellulolytic properties in anaerobic digesters.

Replicated cellulose-degrading batch incubations were set up and run in parallel. They were co-inoculated with municipal digester sludge and ruminal content (Figure 10). Microbial synthesized ^{13}C cellulose was added in several digesters. Non-labelled cellulose was used in control experiments.

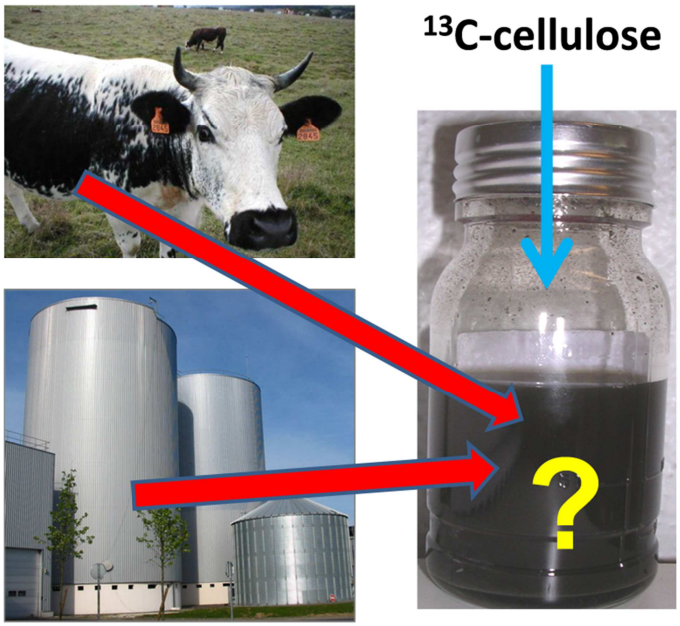


Figure 10: Experimental design of bioaugmentation experiment.

1.2.2.2. Monitoring of the degradation performances after bioaugmentation

Briefly, we observed that cellulose degradation in co-inoculated incubations was efficient but not significantly improved compared to controls inoculated with municipal digester sludge

only. The main difference was the start of methane production, which occurred later in bottles that had not been co-inoculated (not shown).

The isotopic composition of VFAs and gas was measured across time with GC-C-IRMS. It enabled to estimate the amount of metabolites degraded from cellulose and from other substrates (present in the inoculum). Results are illustrated on Figure 11 for VFAs. We observed that an important fraction of the VFAs was not ^{13}C-enriched, which suggested that they came from ruminal content degradation. However, an important accumulation of ^{13}C enriched VFAs was observed from day 7. It confirmed that cellulose was significantly degraded in these incubations.

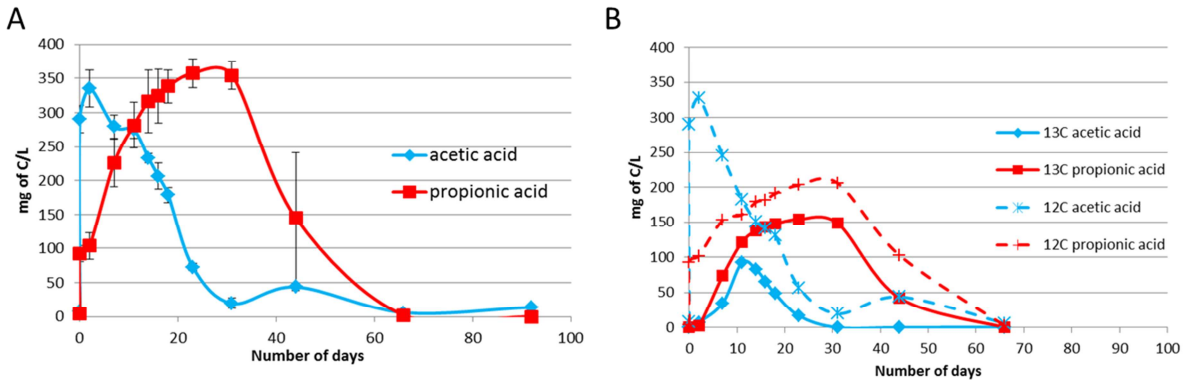


Figure 11: Volatile fatty acids accumulation: (A) total (B) according to the enrichment type.

Microbial populations were monitored to identify those responsible of cellulose degradation.

1.2.2.3. Monitoring of the functional microbial dynamics after bioaugmentation

Parallel incubations had been performed with cellulose that was either ^{13}C-labeled or unlabelled (control incubation). Total DNA from both sets of incubations was extracted from samples recovered at different incubation times and centrifuged in a CsCl gradient in specific tubes. After buoyant density gradient ultracentrifugation, content of each tube was divided in 24 fractions of different densities. DNA concentration/density profiles from ^{13}C-labeled experiments were compared over time (Figure 12). Total DNAs extracted from the control experiments with unlabelled cellulose were also processed in parallel as negative controls.

Control DNAs extracted from all incubations performed with unlabelled substrates exhibited a density ranging from 1.68 to 1.71 g/mL. In DNA density profiles obtained on Days 2, 7, 16 and 44, the peak of density profiles shifted to the highest values from Day 2 to Day 44, indicating an increase in the proportion of heavy DNA. At Day 44, this peak was significantly shifted in labelled experiments when compared with that of the controls. These results suggested a progressive assimilation of labelled carbon by the microbial biomass. No clear ‘heavy peak’ could be distinguished from a ‘light peak’, probably because the density of the heavy DNA was not different enough. Moreover, cellulose is at the beginning of multiple pathways and its degradation into biogas implies a wide number of microorganisms. DNA recovered in our incubation probably originated from microorganisms enriched at different levels of ^{13}C resulting in a continuum of density.

Following these observations, the various DNA fractions in each sample were pooled, based on their density, into three fractions designated as Light, Medium and Heavy, characterized by a mean density of 1.69, 1.71 and 1.73 g/mL respectively, as illustrated in Figure 12. Heavy and light fractions were supposed to contain DNA of microorganism respectively involved or not involved in the cellulose degradation process. Medium was supposed to contain DNA of microorganisms weakly enriched in ^{13}C, possibly involved in cellulose degradation but also involved in the last steps of the degradation or in cross-feeding.

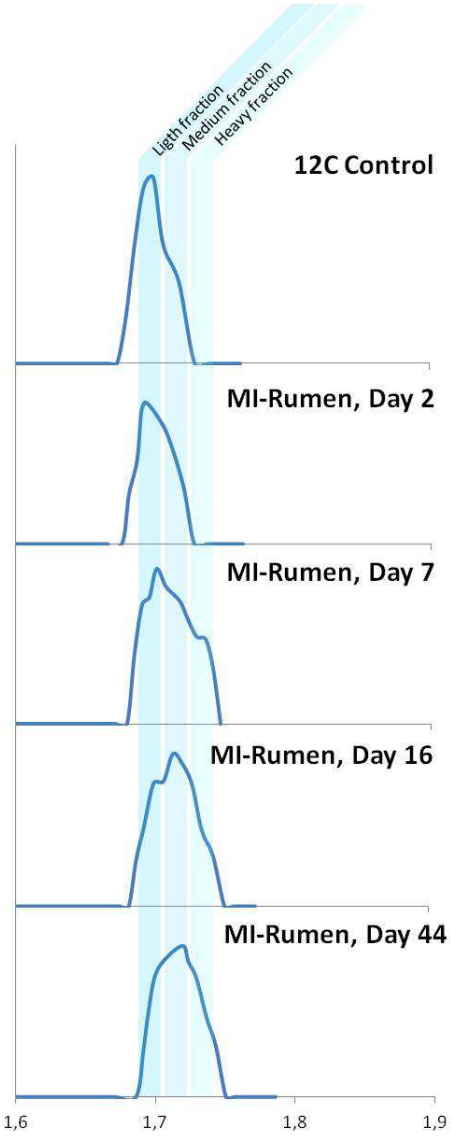


Figure 12: CsCl density gradient profiles of DNA extracted from ^{13}C-labeled cellulose incubations after 2, 7, 16 and 44 days. The first graph depicts the CsCl density gradient profile of DNA extracted from a negative control with unlabelled cellulose. Total DNA distribution into gradient fractions was quantified fluorometrically.

To complement this data, and in order to check that DNA had actually been separated according to its isotopic enrichment through ultracentrifugation, ARISA was performed on the different fractions of DNA recovered after SIP in one sample of an incubation with ^{13}C and one sample of the unlabelled control incubation. Figure 13 A and B shows the results obtained at Day 16 for archaea and bacteria respectively.

On archaeal profiles, bands A and B were recovered in the heavier DNA fractions of the labelled sample, whereas they were recovered in the lighter DNA fractions of the unlabelled control. It confirmed that a shift in archaeal DNA density due to ^{13}C-labeling was observed in this experiment and that labelled DNA could be recovered in specific fractions. The same observations were made on bacterial profiles (Figure 13 B). Bands C, D, E and F were recovered in the heavier DNA fractions of the labelled sample whereas they were recovered in the lighter DNA fractions of the control. It therefore appears that they corresponded to microorganisms involved in ^{13}C-cellulose degradation and that ultracentrifugation enabled to correctly resolve enriched and not enriched DNA.

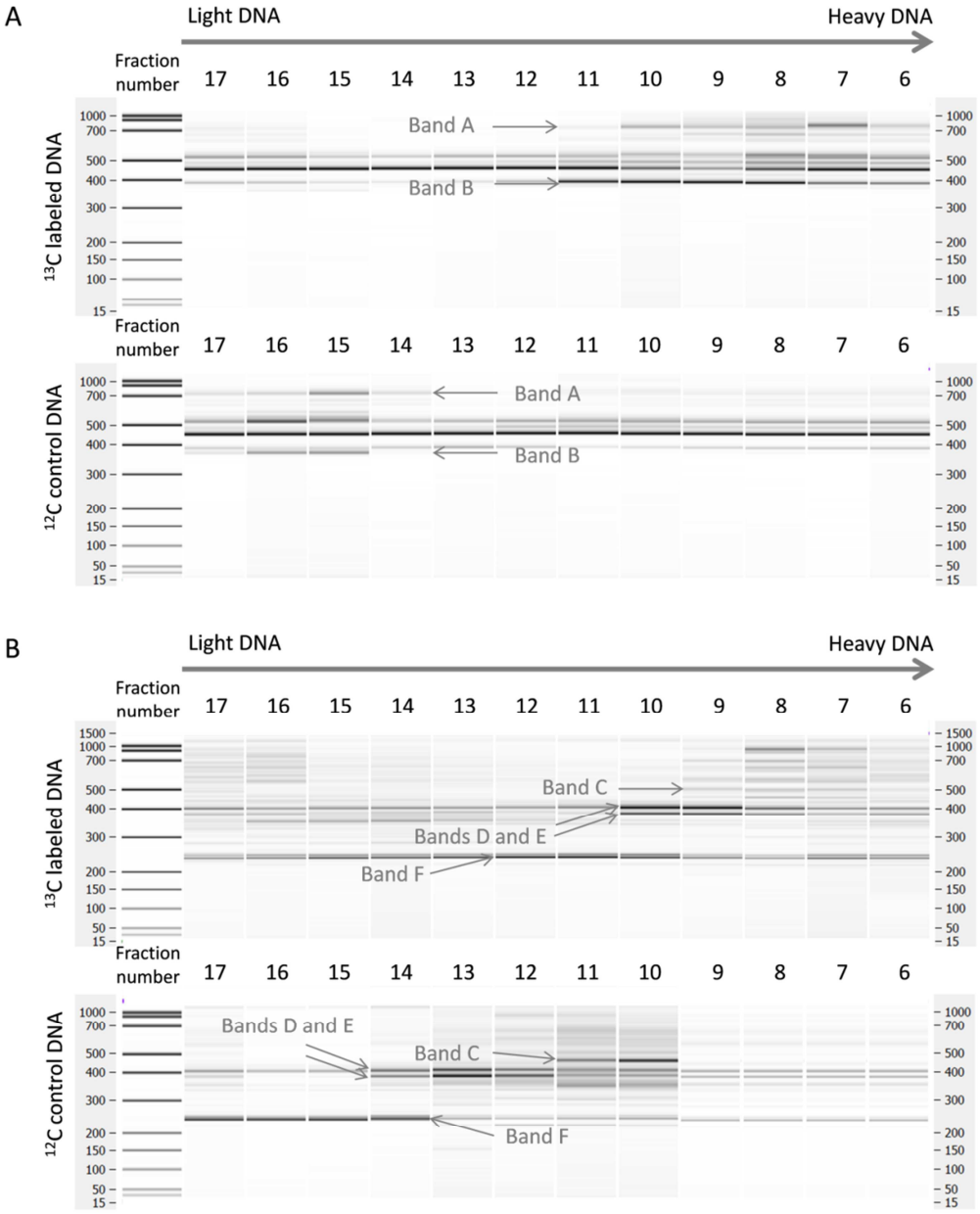


Figure 13: Archaeal (A) and bacterial (B) ARISA profiles of the different DNA fractions recovered after ultracentrifugation for ^{13}C-labeled cellulose incubation and negative control with unlabeled cellulose after 16 days.

To evaluate the involvement of ruminal microorganisms in cellulose degradation, DNA from the heavy fraction was sequenced to target 16S rRNA gene. Analysis was focused on the day 16. This date was selected because it corresponded to a stage at which approximately 1/3 of the final cumulated methane had been produced, enriched VFA were at their maximum level (Figure 11), and it was associated with a clear shift in the DNA density profiles (Figure 12) and fractioned DNA ARISA profiles (Figure 13 A and B). This sampling point thus seemed to represent the best available compromise to clearly identify the microbes involved in the processes while at the same time minimizing possible cross-feeding biases. Ruminal content and industrial mesophilic digester sludge used as inocula were also sequenced.

The analysis showed that none of the microorganism of the ruminal content settled in the digesters and was involved in cellulose degradation. Co-inoculation did not enable the expression of ruminal microbial community during cellulose degradation in an anaerobic digester. We hypothesized that the inocula competed for the carrying out of this functional process. The industrial microbial community was more flexible than the natural one and seemed to exhibit greater resistance to changes in incubation parameters. On the contrary, the rumen's highly specialized species did not adapt and settle in the microcosms. Ruminal microbiota was probably consumed by other microorganisms coming from the industrial sludge. Our conclusion was that other specific parameters were probably needed for ruminal community to settle in the reactor. This study showed that exploiting the rumen's cellulolytic properties in anaerobic digesters is not straightforward. Co-inoculation can only be successful if ruminal microorganisms manage to thrive in the anaerobic digester and out-compete native microorganisms, which requires specific nutritional and environmental parameters, and a meticulous reproduction of the selection pressure encountered in the rumen. SIP was a powerful method to really focus on the degradation of a specific substrate and obtain these conclusions. In the same way I used these approaches to evidence an asymmetrical response of anaerobic digestion microbiome to temperature changes (Chapleur et al., 2016a).

1.3. Perspectives

The use of biomarkers in combination with stable isotope analysis was a very new approach in microbial ecology when I started my PhD. In 2007 in their paper entitled "Who eats what, where and when? Isotope labelling experiments are coming of age", Neufeld et al. wrote that Isotope-labelling experiments would change the way microbial ecologists investigate the ecophysiology of microbial populations and cells in the environment. They anticipated a shift from discovery-driven science to hypothesis-testing experimentation (Neufeld et al., 2007). Since then, a great number of papers applying this technique to a variety of subjects have appeared (Ge et al., 2006; Chen et al., 2008b; Pinjing et al., 2009; Ito et al., 2012; Kleinsteuber et al., 2012). Stable isotope probing (SIP) techniques have become state-of-the-art in microbial ecology over the last ten years.

Dumont and Murrell suggested to go further than only identify active microorganisms and also clone ^{13}C-labelled DNA from a DNA-SIP experiment to generate a metagenomic library of the microorganisms and get insights in the functions of a particular environmental process (Dumont and Murrell, 2005). Development of sequencing and high-throughput omics enabled to do so without the tedious process of cloning multiple DNA fragments. SIP has evolved going from identification of active microorganisms with 16S sequencing to targeted metagenomics of active microbial populations with stable-isotope probing (Chen and Murrell, 2010; Coyotzi et al., 2016). For example, a recent study combined both approaches to target key long-chain fatty acid-degrading populations in anaerobic digesters (Ziels et al., 2018). Recent studies detecting labelled mRNA demonstrated that RNA-SIP was not limited to the analysis of rRNA, but could enable targeted transcriptomics (Lueders et al., 2016).

I'm currently not using SIP in the research projects I coordinate, but I'm associated to the VIRAME ANR project (coordination Dr Ariane Bize, Irstea) which evaluates the possibility to use SIP approaches to indirectly label viruses.

2. Associated methodologies for isotope incorporation visualisation and quantification

Methods described in the previous section enable to identify microorganisms involved in a degradation process. However, they do not link identity to a specific biochemical process and do not enable to measure levels of isotope enrichment. Different methods coupling isotope-labelling experiments and *in situ* hybridization have been developed to investigate the ecophysiology of microbial populations. They reveal the specific uptake of isotopically labelled substrates within individual cells and the phylogenetic affiliation of the very same cells.

Among them, MarFISH (Lee et al., 1999) combines fluorescence *in situ* hybridization and microautoradiography (mar) but is limited to radioactive isotopes. Raman FISH (Huang et al., 2007) allows the detection of isotopes incorporated into cells using Raman microscopy but has a lateral resolution of about 1 µm and requires that the minimum amount of label within a cell is 10% to detect a spectral shift. Development of secondary ion mass spectrometry (SIMS) made possible to measure precisely the isotopic composition of microbes after isotopic enrichment (Orphan et al., 2001). In particular, nanoSIMS instrument, with an important resolution (down to 50 nm) and a great sensitivity (less than 0.1% of difference in the labelling can be detected) enables to measure the incorporation of stable isotopes in individual cells.

When I arrived at Irstea, a very innovative method, based on the combined use of stable isotopes labelling, FISH and nanoSIMS had just been developed (Li et al., 2008). It has been called SIMSISH (Simultaneous analysis of microbial identity and function using NanoSIMS). I had the opportunity to carry out a technical validation of the accuracy of this method (Chapleur et al., 2013). I also applied this technique to measure the isotopic enrichment of cells involved in methanol degradation. The method and the results obtained are presented in the next sections.

2.1. NanoSIMS and SIMSISH technique

NanoSIMS is a nanoscopic scale resolution chemical imaging mass spectrometer based on secondary ion mass spectrometry. Briefly, an ion beam is emitted inside the instrument. It collides with the surface of the sample. The collision that occurs causes atoms to sputter from the sample surface. They become secondary ions, which are then detected after transfer through a mass spectrometer (see Figure 14). Resolution of the instrument allows precise isotopic and elemental measurements of the surface of various types of samples (geological, biological, etc.). Analysis can be realized in two dimensions to map the surface of the samples.

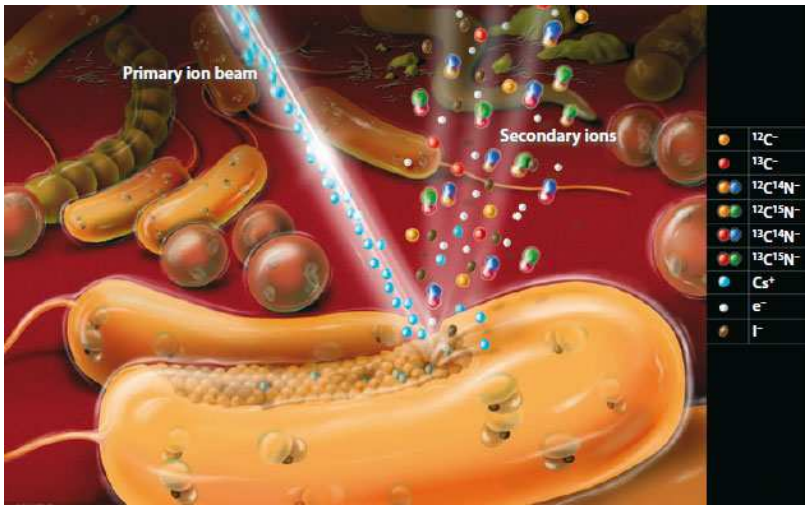


Figure 14: Illustration of the principle of nanoSIMS from (Wagner, 2009). Primary ion beam sputters the sample. Secondary ions are emitted and will be detected in the mass spectrometer.

The concept of SIMSISH (Li et al., 2008) relies on performing FISH experiments on samples containing microorganisms enriched in a stable isotope (typically samples from a SIP experiment). In the probes used for FISH, the fluorescent dyes are replaced by molecules containing elements rarely present in biomass, like halogens. Analysis in the nanoSIMS of samples hybridized this way allows on the one hand the detection of probes distribution in the samples (halogens signal) and on the other hand the mapping of the isotopic composition on the sample (for example ^{12}C and ^{13}C signal). Using a single instrument, it is thus possible to simultaneously detect the hybridization of the oligonucleotide probe revealing the phylogenetic identity of the targeted microbe and monitor *in situ*, at the single-cell level, its isotopic enrichment in various elements of biological interest. Concept of the method is illustrated in Figure 15.

Measures of isotopic composition of microorganisms can be used in combination with degradation models to get information on the role of the different microorganisms in the degradation chain (see Figure 16). Briefly, the idea is that, according to their function in the digestion process, the microorganisms feeding on labelled substrates or on the degradation products of the labelled substrates will increase their isotopic composition in the atom of interest. By comparing on the one hand the modelled theoretical isotopic composition of microorganisms performing the different functions across time, and on the other hand the measures of the actual isotopic composition of the microorganisms suspected to carry out these functions, it is possible to confirm or deny their role.

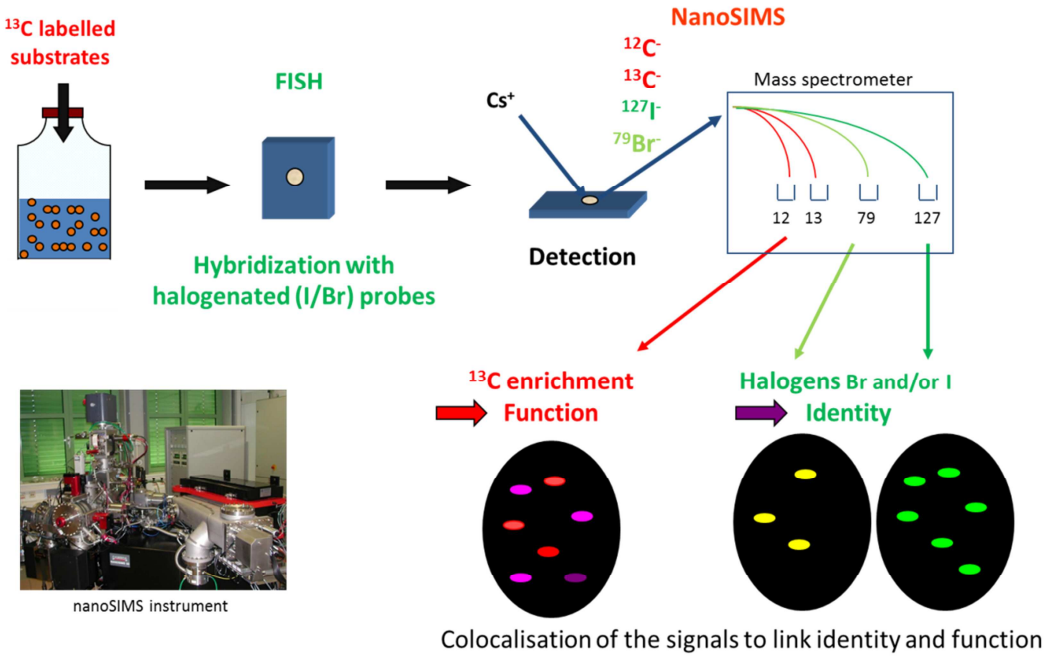


Figure 15: Schematic representation of the SIMSISH methodology. After stable isotope labelling, samples are spread on silicon wafers and hybridized with halogenated probes targeting specific microorganisms. Samples are analysed in a nanoSIMS instrument. A caesium ion beam extracts secondary ions from the surface of the samples. Analysis of the ions in a mass spectrometer enables to map the stable isotope enrichment and the halogens distribution. By comparing the different images, identified microorganisms can be associated to a specific isotope enrichment that can be used to determine the function of the microorganisms.

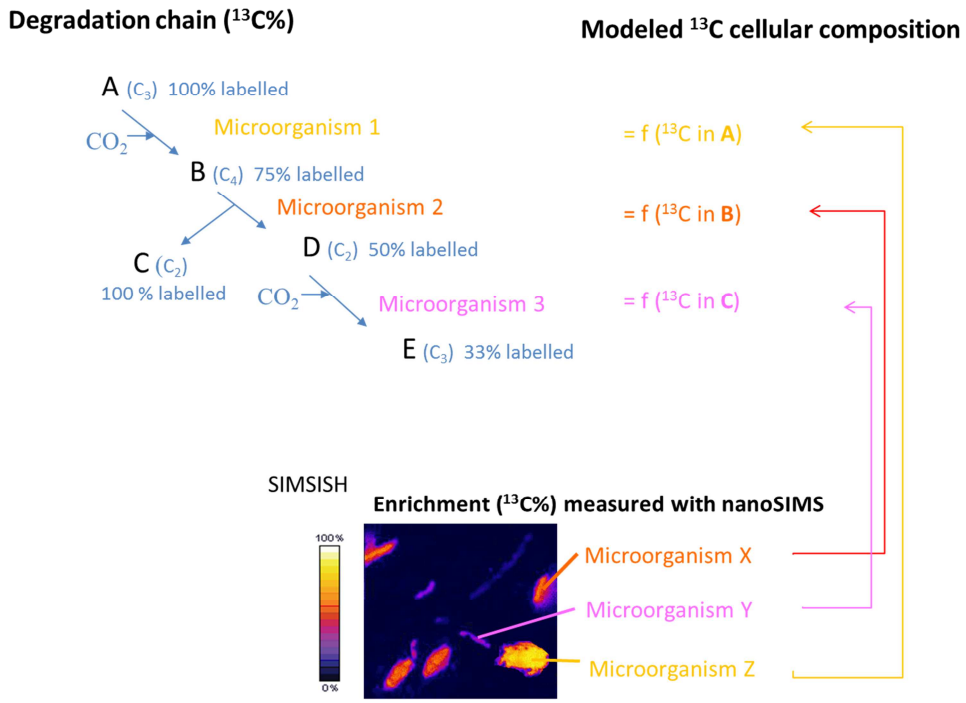


Figure 16: How to link function and identity with SIMSISH? Degradation of a labelled molecule (A) can be modelled to estimate the theoretical isotopic enrichment of its degradation products (B, C, D, E) and the isotopic composition of the microorganisms that would be involved in the different steps of the process (microorganisms 1, 2, 3). Actual isotopic composition of the microorganisms identified in the samples (microorganisms A, B, C) can be measured with SIMSISH. The measured isotopic composition can be compared to the predictions of the model to hypothesize the role of the microorganisms.

2.2. Evaluation of the influence of the experimental procedure of SIMSISH on enrichment measures with nanoSIMS

Treatment and preparation of samples prior to nanoSIMS analysis with SIMSISH are factors that can modify the initial composition of the microbial cells. Indeed, hybridization is performed on PFA-fixed and ethanol-permeabilized cells. Halogenated probes used have a natural isotopic carbon composition. Even if only a limited number of probes and exogenous atoms are introduced (and endogenous atoms removed) into cells during the fixation and hybridization procedures, the samples might be modified resulting in differences between the original isotopic content and the apparent isotopic composition of cells measured by nanoSIMS.

To evaluate the influence of these procedures, I realized an experiment where the isotopic enrichment of E. coli cells was measured before treatment, after fixation, and after hybridization. Single cell level analyses, realized with nanoSIMS at Institut Curie, were compared to reference analyses, at population level on cells pellets, and realized with elemental analyser-isotopic ratio mass spectrometer (EA-IRMS). Analyses were performed on a set of *E. coli* cells isotopically labelled at different levels of ^{13}C enrichment (1.10%, 2.08%, 3.06%, 6.00%, 10.89%, 20.68%, 40.26%, 59.84%, 79.42% and 99.00% of ^{13}C for EA-IRMS and 10.9, 20.7, 40.3 and 79.4 % of ^{13}C for nanoSIMS).

Results of the different measures are presented on Figure 17. EA-IRMS and nanoSIMS measurements showed independently that the fixation and hybridization procedures used in SIMSISH technique had little influence on the carbon isotopic composition of *E. coli* treated-cells. This indicated that only a very small amount of exogenous carbon was introduced in cells during SIMSISH procedure (a theoretical calculation suggested that the amount of carbon introduced with probes during SIMSISH procedure was less than 0.22% of the total amount of carbon in *E.coli* cells).

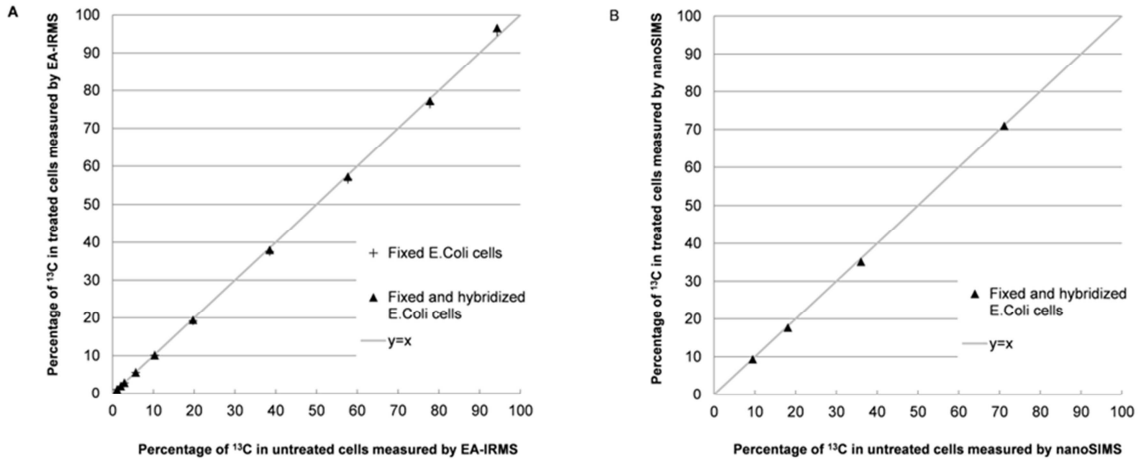


Figure 17: Isotopic composition of untreated, fixed and fixed hybridized cells measured with EA-IRMS and nanoSIMS. (a) ^{13}C isotopic composition of fixed, and fixed/hybridized E.coli cells measured with EA-IRMS compared to ^{13}C isotopic composition of untreated cells – (b) ^{13}C isotopic composition of fixed/hybridized E.coli cells measured with nanoSIMS compared to ^{13}C isotopic composition of untreated cells.

Additionally, untreated and fixed/hybridized cells were analysed in mixture with nanoSIMS for the enrichment values of 10.9, 20.7, 40.3 and 79.4 % of ^{13}C. A set of images showing

sulfur and iodine emission as well as corresponding ^{13}C abundance map is presented in Figure 18. Sulfur is an indication of the total biomass. Iodine signal enables an identification of hybridized cells from the non-hybridized (untreated) ones.

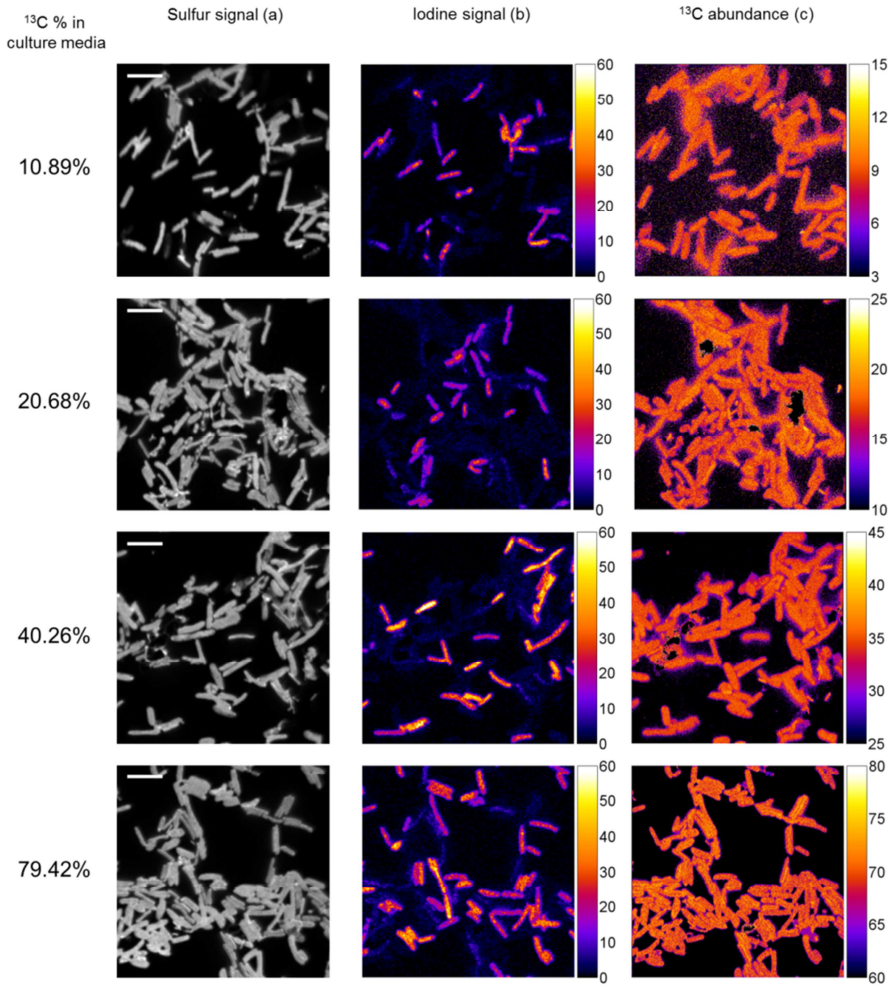


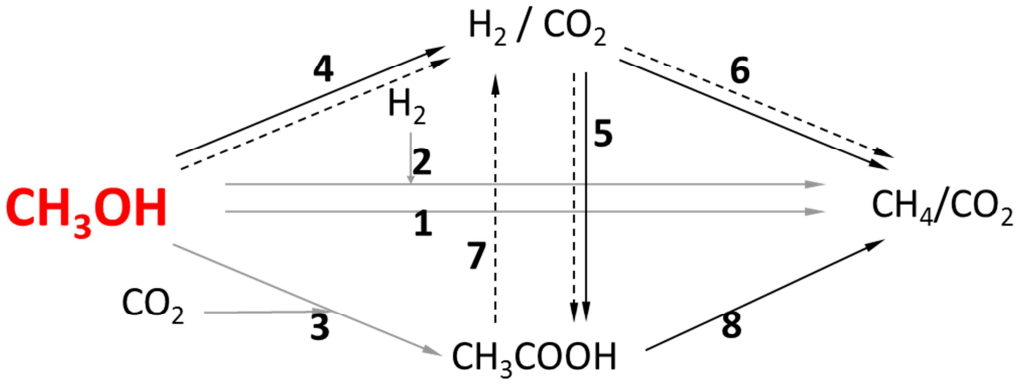
Figure 18: NanoSIMS images obtained for mixed untreated and fixed/hybridized cells grown in enriched culture media with nominal ^{13}C abundance of 10.9, 20.7, 40.3 and 79.4%. Panel (a) shows the secondary ion of $^{32}S^-$ image as an image of total biomass (scale bar : 5 µm). Panel (b) shows the secondary ion of $^{127}I^-$ image as an indication of hybridized cells. (c) ^{13}C Isotopic abundance map of corresponding area.

We observed that nanoSIMS provided an efficient discrimination between lateral neighbour cells. It revealed that incorporation of labelled carbon was homogeneous in all the cells of the same type, and within the cells. Iodine signal enabled to clearly identify the labelled cells. We concluded that isotopic composition of single cells could be determined precisely with SIMSISH.

We also observed that multiple labelling was possible to perform several isotopic measurements and identification at the same time or to study specific syntrophic interactions (Figure 22). Combined with stable isotope probing, SIMSISH constitute an elegant tool to decipher networks of biogeochemical processes catalysed by uncultured microorganisms within complex environments.

2.3. Combination of SIMSISH with modelling to assess the function of microorganisms

To illustrate the potential of SIMSISH, an experiment was set up to identify the microorganisms responsible of the different pathways activated during mesophilic anaerobic degradation of methanol. Methanol is a small molecule, but under anaerobic conditions, it has a complex web of possible degradation routes, summarized in Figure 19.



	Reactions	Microorganisms involved
1	$4CH_3OH \rightarrow 3CH_4 + HCO_3^- + H^+ + H_2O$	Methylotrophic methanogen
2	$CH_3OH + H_2 \rightarrow CH_4 + H_2O$	Methylotrophic methanogen
3	$4CH_3OH + 2HCO_3^- \rightarrow 3CH_3COO^- + H^+ + 4H_2O$	Acetogenic bacteria
4	$CH_3OH + 2H_2O \rightarrow 3H_2 + HCO_3^- + H^+$	Acetogenic bacteria
5	$4H_2 + 2HCO_3^- + H^+ \rightarrow CH_3COO^- + 4H_2O$	Acetogenic bacteria
6	$4H_2 + HCO_3^- + H^+ \rightarrow CH_4 + 3H_2O$	Hydrogenotrophic methanogen
7	$CH_3COO^- + 4H_2O \rightarrow 2HCO_3^- + 4H_2 + H^+$	Syntrophic bacteria
8	$CH_3COO^- + H_2O \rightarrow CH_4 + HCO_3^-$	Acetoclastic methanogen

Figure 19: Summary of the degradation pathways possible for the anaerobic digestion of methanol, and main microorganisms involved.

Anaerobic batch digesters were set-up and fed with ^{13}C-labeled methanol (PhD of Tianlun Li). Concentration and isotopic enrichment of degradation intermediates and products were analysed. Based on the values measured, it was demonstrated that only a part of the degradation routes shown in Figure 19 were active. They are illustrated on Figure 20. Based on this diagram, a mathematical model describing the functioning of the ecosystem was developed (by Pr Vassily Vavilin during a scientific stay in the lab). It was fitted on macroscopic data (gas, VFA, methanol, dissolved inorganic carbon). Mass balanced equations were computed both for ^{12}C and ^{13}C. Equations describing the temporal isotopic enrichment pattern of microbial functional groups were designed based on hypothesis made on their *in situ* ecophysiology.

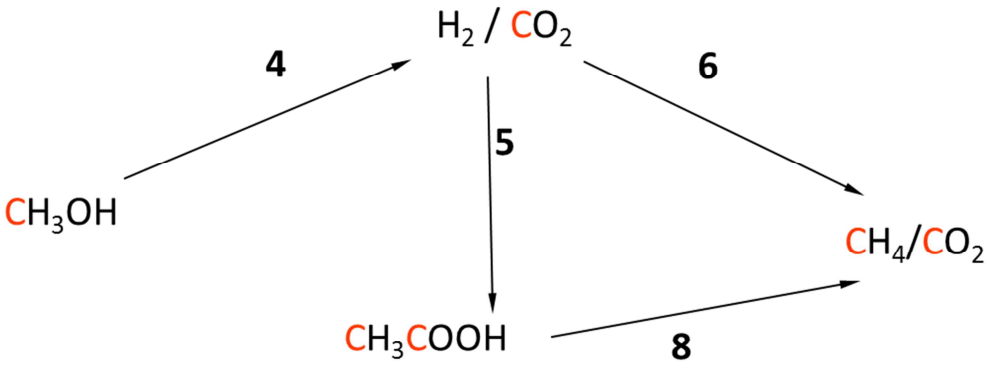


Figure 20: Degradation pathways activated during the mesophilic anaerobic digestion of methanol.

A SIP strategy allowed to identify the functional microorganisms. Specific FISH probes labelled with halogens were designed to target these microorganisms. For example, archaea of *Methanosarcina* genus can use the hydrogenotrophic or acetoclastic pathways to produce methane (reactions 6 and 8 on Figure 20). Their isotopic enrichment across time according to the type of metabolism was modelled by V. Vavilin and is presented on Figure 21.

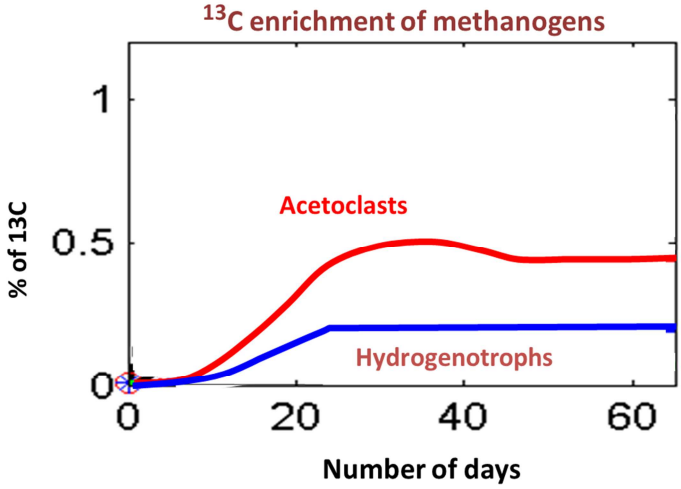


Figure 21: Modelled ^{13}C enrichment of acetoclastic and hydrogenotrophic archaea across time in the conditions of the anaerobic digester degrading ^{13}C-methanol.

I prepared samples with the SIMSISH technique to target *Methanosarcina* cells with a brominated probe and bacteria in general with an iodinated probe. Samples were analysed with nanoSIMS at Institut Curie. Images of one sample taken after 25 days are presented on Figure 22. Sulfur image showed a big round form surrounded by smaller forms with a typical shape of cell. We hypothesized that the round form was a cluster of archaea that was surrounded by bacteria. Indeed, clear bromine and iodine signals, corresponding to probes hybridization, were observed and enabled the identification of both *Methanosarcina* archaea and bacterial cells at the same time. NanoSIMS also enabled to measure the isotopic enrichment of the archaeal cluster and of the nearby bacterial cells. The mean ^{13}C enrichment

of archaeal cluster was 30.0% (standard error = 1.3%), and the mean enrichment of bacteria was 24.4% (standard error = 2.5%), showing that both types of microorganisms were involved in methanol degradation at different levels.

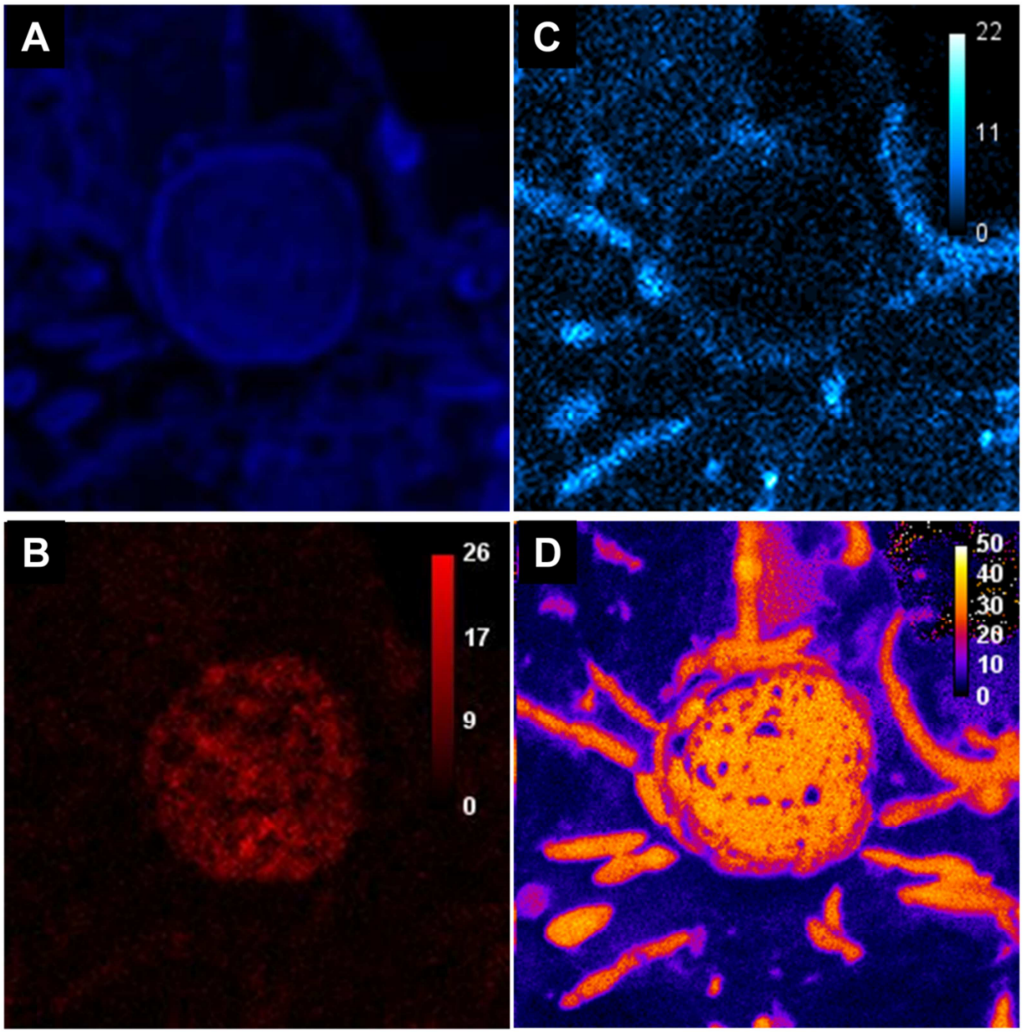


Figure 22: NanoSIMS images of a sample taken at day 25 from an anaerobic digester fed with ^{13}C-labeled methanol labelled with a generalist bacterial iodinated probe (EUBI) and a *Methanosarcina* genus specific brominated probe (MS1414). Panel (a) shows the secondary ion of $^{32}S^-$ image as an image of total biomass. Panel (b) shows the secondary ion of $^{81}Br^-$ image as an indication of *Methanosarcina* cells identity. Panel (c) shows the secondary ion of $^{127}I^-$ image as an indication of total bacteria. Panel (d) shows the ^{13}C Isotopic abundance map.

Another measure performed later, at day 42, but this time with only an iodinated probe for archaea, resulted in the images presented on Figure 23. Another cluster of archaea was evidenced with the iodine signal. This time the isotopic enrichment of the cells of the cluster was 39.7%. Modelling predictions were compared to isotopic enrichment measured at the single cell level (Figure 21). It showed that, in the conditions of this experiment, archaea of *Methanosarcina* genus behaved as acetoclasts.

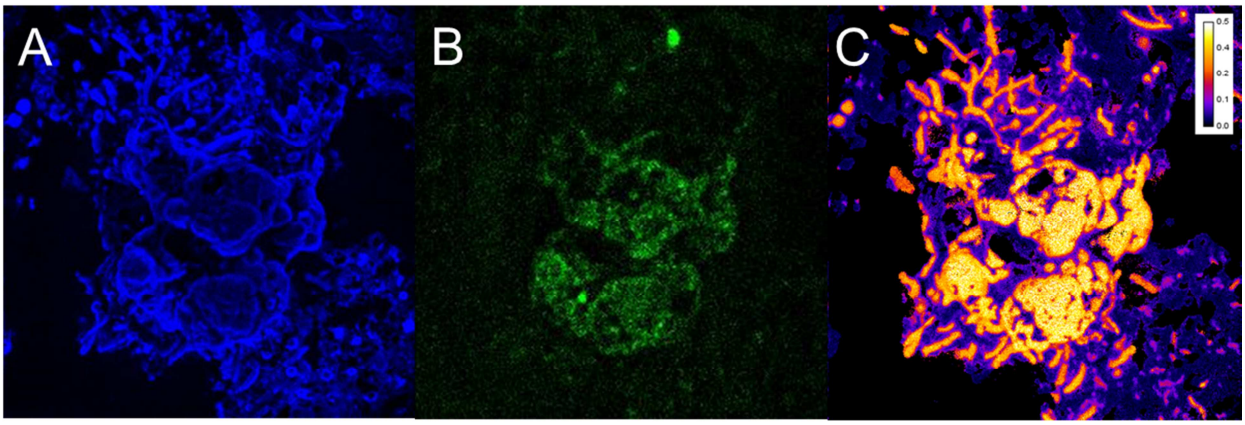


Figure 23: NanoSIMS images of a sample taken at day 42 from an anaerobic digester fed with ^{13}C-labeled methanol labelled with an archaeal iodinated probe. Panel (a) shows the secondary ion of $^{32}S^-$ image as an image of total biomass. Panel (b) shows the secondary ion of $^{127}I^-$ of archaeal cells. Panel (c) shows the ^{13}C Isotopic abundance map.

Similarly, we questioned the identity of microorganisms responsible for methanol oxidation (reaction 4, Figure 20). Based on the sequences of microorganisms identified with SIP, several FISH probes were designed. One of them, GAM489, targeted bacteria from the class *Gammaproteobacteria*. NanoSIMS analysis showed that these cells exhibited an enrichment similar to the one modelled for autotrophic methanol oxidizers (not shown here), suggesting that they were involved in this reaction. SIMSISH procedure therefore provided meaningful indication to understand the *in situ* functional behaviour of uncultured microbes.

3. Conclusion

Methods making use of stable isotope showed up to be very performant to target specific and complex questions, provided that they were used meticulously. According to me, their main limitation is that they are very tedious and costly, especially if the labelled molecule used must be synthesized specifically. They require that the question targeted is very specific and that the experimental set-up is properly designed.

For example, nanoSIMS instrument is powerful to measure the isotopic enrichment of single cell. However, this technique requires a lot of time of observation. Spotting specific microorganisms, even if labelled, in an area of one square millimetre is very complex. To realise accurate measurement it is important that cells are either deposited in single or homogenous layers. Analysis requires a long implantation of the ion beam to observe a surface important enough. Sometimes successive analyses at different resolution or with different parameters are required to obtain image showing both halogens signal and an accurate measure of the isotopic enrichment. Instruments are very costly, no automatization is possible. Coupling this approach with modelling is complex. A single carbon organic molecule such as methanol is already at the origin of multiple reactions, carried out by multiple microorganisms. With this in mind, nanoSIMS enabled to evidence various microbial interactions that had never been proven before, or to understand the fate of substrate in the microbial community and to visualise food network in microbial spatial organisation (Jiang et al., 2016; Musat et al., 2016; Nuñez et al., 2018; Tao et al., 2019).

Finally, I think that stable isotope probing, used in combination with high-throughput untargeted methods, will probably enable to gain more complex information. They have already been combined with metagenomics, as said before (Ziels et al., 2018). Use of untargeted metabolomics to analyse sample from SIP experiments would open new perspectives to decipher the pathways of degradation of a specific compound in a complex environment.

Chapter 2: Screening of numerous experimental conditions with molecular tools to study the inhibition of anaerobic digestion

After spending several years analysing in detail particular situations using SIP methods, I decided to explore other approaches. I took advantage of the possibility to analyse multiple samples in parallel with fingerprinting or metabarcoding (provided by the reduction of analytical costs) to set-up experiments testing simultaneously multiple conditions. My new research questions did not aim at deciphering one specific mechanism anymore, but aimed at comparing the dynamics of the microbial community as a whole under various conditions. The objective was to get insight into the drivers or the bottlenecks of the process under various conditions (1) to better understand how these conditions impact AD process and (2) to propose management strategies to optimize the functioning of AD. I chose to explore the topic of the inhibition of anaerobic digestion with this approach. Indeed, anaerobic digestion process still remains vulnerable to a wide variety of inhibitory substances that induce poor operational stability, preventing AD wide commercialization (see General Introduction, 1.4.4 page 43).

In a first series of experiments, I evaluated the effect of different levels of inhibitors on AD performance and microbial dynamics. In a second series of experiments, I assessed the potential of different operational strategies to mitigate the inhibition of AD. Testing multiple experimental conditions required using appropriate methods to compare the effects, organise the generated data and extract the relevant information. The results presented in this chapter were produced by me or during the PhD of Simon Poirier and Laëtitia Cardona that I supervised, or during a joint work with Céline Madigou (assistant engineer) who wished to develop new skills to become a study engineer (position obtained in 2017).

1. Effect of different inhibitors of the anaerobic digestion

Exploring the literature related to AD inhibition, I observed that most of the time, the studies assessed the effect of an inhibitor by testing only one or a few levels of inhibitor (for example, low and high concentration). I also observed that, from one study to the other, adding the same concentration of an inhibitor often resulted in different effects, both on the degradation performances and on the microbial community, because of the specificity of the experimental

set-up (type of inoculum, feeding, reactor etc.). It limited the possibilities to assess globally the effect of one inhibitor on AD microbiome.

To overpass this limitation, I decided to assess within the same experiment the effect of a wide range of concentration of an inhibitor. Levels of inhibition were selected to obtain an amplitude of inhibition going from no inhibition to total inhibition of the process (i.e. no methane production at extremely high level of inhibitor). The objective was to describe the impact of the different levels of inhibitor on microbial communities' dynamics and to identify the key phylotypes involved into the resistance to the inhibitor. Using different levels of the same inhibitor enabled to evidence different 'thresholds' of inhibition.

Three published papers are related to this research topic.

- Increasing concentrations of phenol progressively affect anaerobic digestion of cellulose and associated microbial communities (Chapleur et al., 2016b)
- Anaerobic digestion of biowaste under extreme ammonia concentration: identification of key microbial phylotypes (Poirier et al., 2016b)
- Community shifts within anaerobic digestion microbiota facing phenol inhibition: Towards early warning microbial indicators? (Poirier et al., 2016a)

The influence of two inhibitors (ammonia or phenol) on the anaerobic degradation of two types of substrate (cellulose or biowaste) was tested in batch reactors. The following paragraphs present the reasons for selecting these inhibitors, the effect of the inhibitors on the degradation performances and on the microbial communities based on a selection of the results I obtained.

1.1. State of the art related to the studied inhibitors

Ammonia and phenol were selected from a wide variety of potential inhibitors. They are often present in waste or can be produced during its degradation. They are therefore frequently encountered in industrial digesters. As described in the next paragraphs, even if they have already been studied, there is still no consensus on the levels of concentrations which are inhibitory and no consensus on their effects on the microbial communities. These inhibitors were complementary. They are of different nature (inorganic and organic molecules). Phenol can be degraded during the process while ammonia cannot. Both could be easily manipulated and quantified in our lab.

1.1.1. Ammonia

Ammonia is regularly mentioned as the primary cause of digester failure (Yenigün and Demirel, 2013), even though it is an essential nutrient for the growth of microorganisms at low level. Many traditional AD substrates such as livestock manure, slaughterhouse by-products and food industrial residues contain a high nitrogen concentration. It is due to the presence of organic nitrogen such as urea and proteins which readily release ammonia upon digesters during their anaerobic degradation (Rajagopal et al., 2013).

In liquid media, ammonia can be found into two distinctive forms: the unionized free ammonia nitrogen (FAN or NH_3) and the ammonium ion (NH_4^+). Total ammonia nitrogen, abbreviated as TAN, is a combination of unionized free ammonia and ammonium ion. The chemical equilibrium between FAN and TAN concentrations mainly depends on pH and temperature (Anthonisen et al., 1976). FAN is considered more toxic to anaerobes than TAN. The most widely accepted mechanism explaining FAN inhibition is based on change of intracellular pH due to its passive diffusion through cell walls, increasing the maintenance energy requirement and inhibiting specific enzyme reactions (Wittmann et al., 1995). TAN

has been reported to induce 50% reduction in CH_4 production at a wide range concentrations varying from 1.7 g/L to 14.0 g/L (Chen et al., 2008a). According to Sung and Liu (2003), total inhibition occurs in the range of 8.0-13.0 g/L of TAN depending on pH and inoculum. Similarly, toxic effect of FAN was observed from 150 mg/L to 1200 mg/L (Rajagopal et al., 2013; Yenigün and Demirel, 2013). The wide range of inhibiting FAN or TAN concentrations can be ascribed to the differences in nature of substrates, environmental conditions (pH, temperature, hydraulic retention time) and acclimation of inoculum (Van Velsen, 1979).

The inhibitory effect of ammonia is known to mainly influence the methanogenesis (Calli et al., 2005b). However there is no consensus on its exact role. For example, under ammonia stress, acetoclastic methanogens are usually considered to be more vulnerable than the hydrogenotrophic ones (Koster and Lettinga, 1984; Angelidaki and Ahring, 1993). Thanks to stable isotope-based analytical techniques, a metabolic shift in methanogenesis from the acetoclastic mechanism to the syntrophic pathway (in which hydrogenotrophic methanogens are involved) was pointed out when ammonia concentration is increased (Schnürer and Nordberg, 2008). Other studies also showed that increasing concentrations of ammonia induced a growth of Syntrophic Acetate Oxidation (SAO) Bacteria and hydrogenotrophic methanogens (Karakashev et al., 2006; Westerholm et al., 2012). However, in these latter studies, acetoclastic methanogens were still detected. Conflicting results tended to demonstrate that acetoclastic methanogenesis was predominant under ammonia stress (Calli et al., 2005a; Fotidis et al., 2014). Furthermore, Hao et al., (2015) indicated that besides the SAO pathway, acetoclastic methanogenesis catalyzed by *Methanosarcina spp* was still active at high ammonia levels.

1.1.2. Phenol

Phenols can be produced from the biodegradation of naturally occurring aromatic polymers such as humic acids and tannins or from degradation of xenobiotic compounds such as pesticides (Fang et al., 2006). Different pre-treatments applied to increase biogas production efficiency from lignocellulosic materials are also known to result in the production of phenolic compounds (Monlau et al., 2014). They are regularly detected at concentrations reaching up to several grams per litre in different types of effluents from petrochemical or paper industry (Veeresh et al., 2005; Rosenkranz et al., 2013) or in contaminated sewage sludge or municipal solid waste (Hoyos-Hernandez et al., 2014). Phenol concentrations up to 4288 mg/kg were measured in digestion sludge of anaerobic digesters degrading different types of organic waste such as pig manure, biowaste and plant materials (Levén et al., 2012). Although phenol can be biodegraded to harmless compounds under methanogenic conditions, high concentrations are toxic to different groups of microorganisms involved in methane production (Olguin-Lora et al., 2003; Rosenkranz et al., 2013). Their detrimental effects on anaerobic micro-organisms have been observed in anaerobic systems treating municipal solid waste, olive mill wastes and wine distillery wastewater (Busca et al., 2008). Phenol is indeed a membrane-damaging microbiocide that affect membrane proteins and alter the cell wall permeability, inducing progressive leakage of intracellular constituents (McDonnell and Russell, 1999). In particular, it was observed that unexpected phenol shock loadings within not acclimated anaerobic digesters can induce major disruptions in AD bioprocess, leading to the decrease of biogas production rate and longer treatment durations (Veeresh et al., 2005).

1.2. Effect of phenol and ammonia on the degradation performances

1.2.1. Screening of multiple inhibitory levels: experimental design

To test the influence of different levels of inhibition, a similar experimental design was used in (Chapleur et al., 2016b; Poirier et al., 2016a; Poirier et al., 2016b), as illustrated in Figure 24. In these three experiments, batch anaerobic digesters of 1L were set-up, inoculated with a sludge coming from an anaerobic digester, fed with either biowaste or cellulose (chosen as two representative substrates for AD) in biochemical methane potential buffer (BMP). Ammonia or phenol was added in order to reach 10 different concentrations. The different concentrations were chosen based on the literature to go from non-inhibited reactors to totally inhibited reactors. All incubations were performed in triplicates. Time zero samples were collected before starting incubation of the reactors and one sample was taken every week until biogas production was completed. Degradation performances were monitored (biogas production and composition, VFA, total organic carbon).

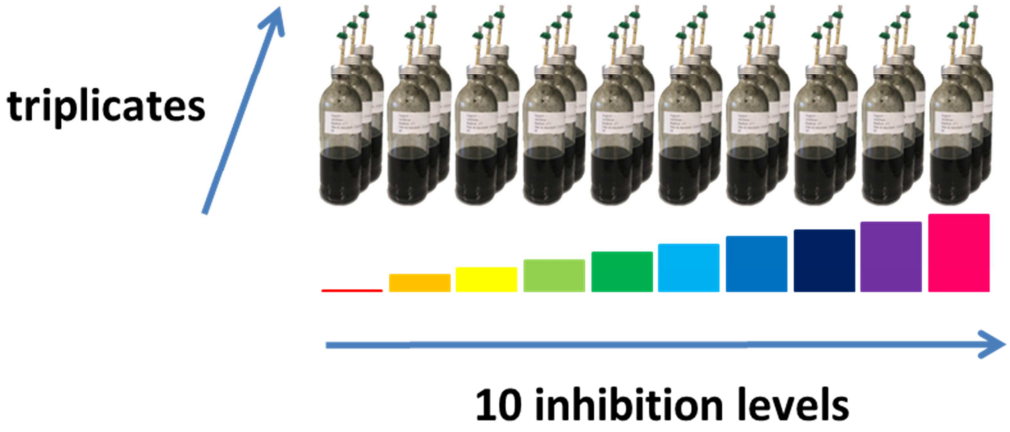


Figure 24: Typical experimental design used in my experiments to assess the influence of one inhibitor on AD.

1.2.2. Effect of the inhibitors on the degradation performances

The average cumulated methane production of the different triplicate reactors of the different experiments are presented in Figure 25A, Figure 26A and Figure 27A. In all the experiments, as inhibitor concentration increased, rates and levels of total methane production decreased progressively. For both experiments with phenol (Figure 25A and Figure 26A), independently of the type of substrate (cellulose or biowaste) phenol concentrations under 1.00 g/L resulted in a moderate inhibition of methane production. Methanogenesis occurred up to 2.00 g/L of initial phenol concentration, despite importantly inhibited. At higher phenol concentration, no CH_4 was produced. For the experiment with ammonia, a moderate effect was observed under 10.0 g/L of TAN, corresponding in this experiment to 242 mg/L of FAN. Digesters were only importantly inhibited at concentrations higher than 25.0 g/L of TAN (387 mg/L of FAN).

The experimental design used gave us access to a progressive gradient of effects. No sharp inhibitory threshold was observed.

In general, in all the experiments, methane production was more affected than carbon dioxide production (not illustrated in this document), presumably because archaeal activity was more sensitive to inhibition than bacterial activity.

We noted that ammonia levels remained stable all along the experiment in (Poirier et al., 2016b). On the contrary, the biogas data revealed that total biogas production in all the bioreactors containing less than 1.25 g/L of phenol exceeded that in the control, suggesting that phenol degradation to biogas occurred as already observed by (Fedorak and Hrudey, 1984; Wang et al., 1991). It was confirmed by measuring the concentration of phenol in the reactors across time (see figures 3 and 2 in the full texts of (Chapleur et al., 2016b; Poirier et al., 2016a)). Phenol was converted to biogas once the degradation of the main substrate was completed, suggesting that its degradation was less favourable or required the growth of specific microorganisms.

In all these experiments, performance of degradation was also assessed by monitoring the VFA accumulation dynamics. These data are not illustrated in this manuscript, but, in general, we observed that as the level of inhibitors increases, VFA consumption was progressively delayed, resulting in a bigger accumulation.

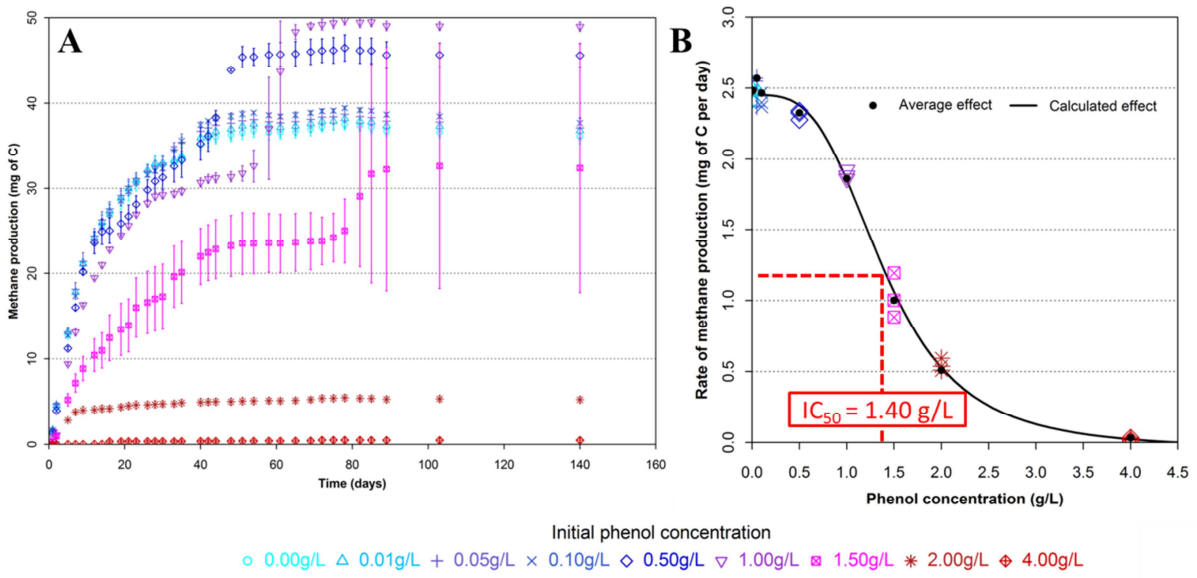


Figure 25: (A) Cumulated CH_4 production (mg of C) over time (number of days) and (B) specific methanogenic activity for the different initial concentrations of phenol in reactors degrading cellulose (Chapleur et al., 2016b). Mean values of triplicate bioreactors are presented and error bars represent standard deviation within triplicates. All experimental specific methanogenic activities (SMA) values are shown for each triplicate. Calculated fitting of experimental values to the Hill model is shown as well as IC50.

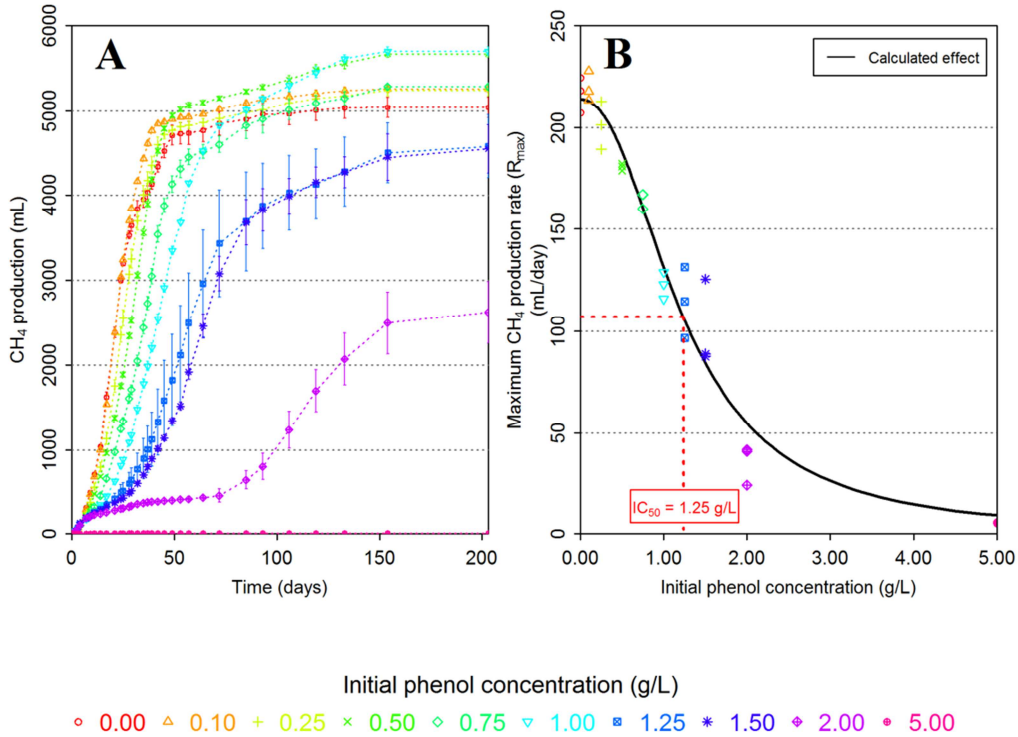


Figure 26: (A) Cumulated CH_4 production (mL) over time (number of days) and (B) specific methanogenic activity for the different initial concentrations of phenol in reactors degrading biowaste (Poirier et al., 2016a). Mean values of triplicate bioreactors are presented and error bars represent standard deviation within triplicates. All experimental specific methanogenic activities (SMA) values are shown for each triplicate. Calculated fitting of experimental values to the Hill **model** is shown as well as IC50.

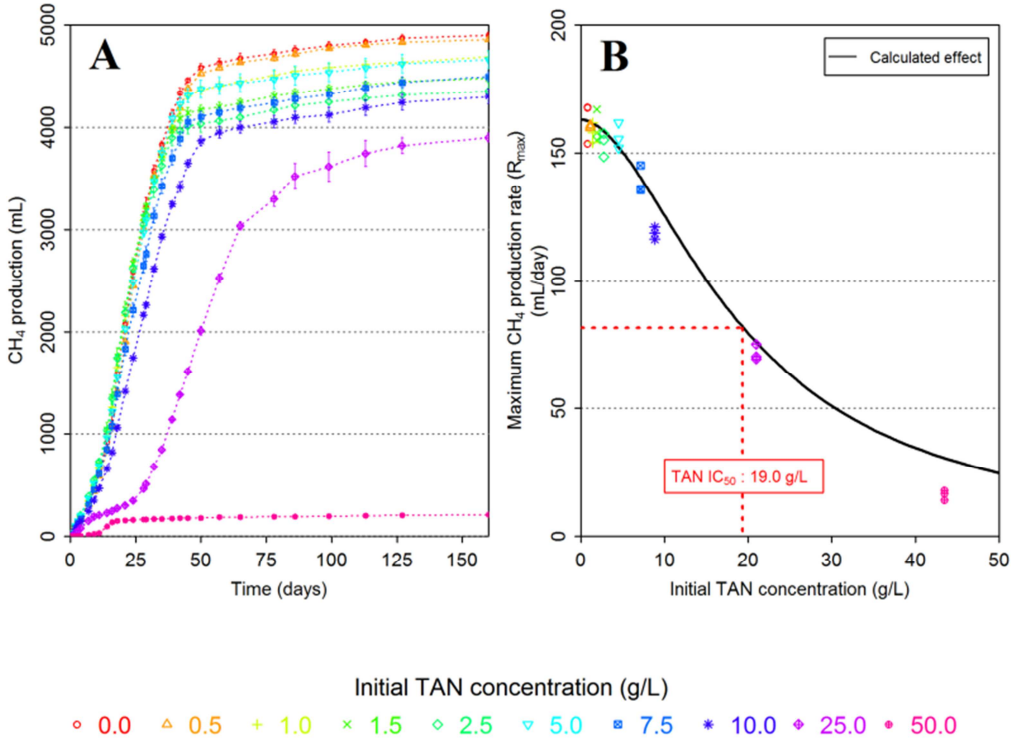


Figure 27: (A) Cumulated CH_4 production (mL) over time (number of days) and (B) specific methanogenic activity for the different initial concentrations of ammonia in reactors degrading biowaste (Poirier et al., 2016b). Mean values of triplicate bioreactors are presented and error bars represent standard deviation within triplicates. All experimental specific methanogenic activities (SMA) values are shown for each triplicate. Calculated fitting of experimental values to the Hill model is shown as well as IC50.

1.2.3. Modelling of the gas production and comparison of the half maximal inhibitory concentrations

To facilitate the comparison of the different conditions tested, cumulative production curves were analytically compared by fitting experimental data with a modified Gompertz three-parameter model (Equation 1) where M(t) is the cumulative methane production (mL) at time t (days); P is the ultimate methane yield (mL); R_m is the maximum methane production rate (mL/days); λ is the lag phase (days); e is the exponential constant.

$$\boldsymbol{M(t) = P \times exp\left\{-exp\left[\frac{R_m \times e}{P} \times (\lambda - t) + 1\right]\right\}}$$

Equation 1: Gompertz three-parameter model

It provided three quantitative indicators to compare accurately the different parameters of gas production, as illustrated Figure 28. This method is commonly used to analyse gas production data (Ware and Power, 2017). Logistic regression can also be used for the same purpose (Ware and Power, 2017).

For example, it confirmed that in both experiments with phenol, the inhibitory effect of the inhibitor on ultimate biogas production was very limited up to initial concentrations of 1.00 g/L. However, at these concentrations, maximum production rate values (R_{max}) gradually decreased along with increasing phenol levels. In parallel lag times (λ) increased. The detrimental effect of phenol, at these concentrations, was therefore a slowdown in the degradation.

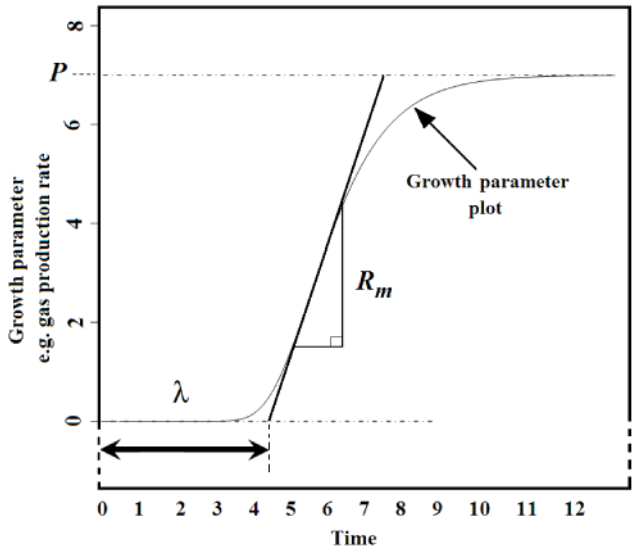


Figure 28: Gompertz equation enables to summarize gas production curves with 3 parameters.

In order to compare the inhibitory effect of ammonia or phenol in our studies with other studies, AD inhibition was more precisely quantified by calculating a half maximal inhibitory concentration value (IC_{50}). IC_{50} is a quantitative measure that indicates how much of a particular inhibitory substance is needed to inhibit a given biological process by 50%. Specific methanogenic activity (slope of the methane production curve during maximum

production, Rm parameter of Gompertz model) was selected as an indicator of AD efficiency. IC_{50} was determined from the correlation plot between Rm and the level of inhibitor. The experimental values were fitted to the Hill model by non-linear regression using Marquard's algorithm and a non-linear regression program (Fechner et al., 2010). The concentration of inhibitor required for a 50% reduction of the methane production rate was estimated (Figure 25B, Figure 26B and Figure 27B).

In experiments using phenol as an inhibitor, IC_{50} of respectively 1.40 g/L phenol (Chapleur et al., 2016b) and 1.25 g/L of phenol (Poirier et al., 2016a) were determined (see Figure 25B and Figure 26B). They were in keeping with values reported by other authors for different substrates: (Wang et al., 1991) observed an IC_{50} of 1.25 g/L during experiments on acetate degradation, (Sierra-Alvarez and Lettinga, 1991) determined that acetoclastic methanogenesis in granular sludge was half-inhibited at 1.10 g/L of phenol etc. Various inhibiting phenol concentrations can be attributed to the differences in nature of substrates, environmental conditions (pH, temperature) and acclimation of inoculum. Interestingly, we noticed important standard deviations for the methane production of bioreactor inhibited with 1.50 g/L of phenol in (Chapleur et al., 2016b) while all the other triplicates behaved very similarly. As this value was close to the IC_{50}, microbiome may have evolved stochastically toward strong inhibition or moderate inhibition.

IC_{50} for experiment with ammonia was 19.0 g/L of TAN. Considering inhibiting TAN concentrations reported in literature, it was surprisingly high (Chen et al., 2008a). For example, similar batch toxicity assays showed a 50% decrease in SMA at TAN range of 8.0 - 13.0 g/L with acclimated biomass (Sung and Liu, 2003). However FAN corresponding to this IC_{50} in our experiment was calculated and was circa only 333 mg/L, due to the high buffering capacity of BMP solution. It was comparable to the values of the literature (Rajagopal et al., 2013). This observation confirmed previous results which demonstrated that FAN is more toxic to methanogenesis than TAN (Kayhanian, 1994).

1.3. Effect of phenol and ammonia on the microbial dynamics

1.3.1. Thresholds of inhibition determined with fingerprinting

ARISA was initially selected to study the dynamics of microbial communities because it made it possible to process a large number of samples efficiently in order to obtain a well-replicated dataset. This large dataset was appropriate to correctly identify ecological patterns and correlate environmental variables with microbial community structure (Prosser, 2010).

For (Chapleur et al., 2016b) (cellulose digestion inhibited by phenol), 324 samples were processed for both archaea and bacteria. They correspond to 12 different sampling dates for all triplicates of bioreactors (between Days 0 and 69). As it was impossible to compare all of them visually, they were compared statistically with a PCA (Figure 29).

For both archaea and bacteria, it was observed that up to 0.10 g/L, the effect of phenol on microbial diversity in general was not significant, in accordance with performance data (diversity profiles were similar to diversity profiles of non-inhibited digesters). The dispersion of profiles at each level of phenol was mainly correlated to the date of sampling (not shown). At higher concentrations (0.50 to 1.00 g/L), profile distribution shifted for both archaea and bacteria. Phenol induced structural modifications in the community's composition. Several microorganisms were inhibited by phenol and were replaced by more tolerant ones. However, at 0.50 g/L, no decrease of the degradation performances had been observed, indicating that, at this concentration, performance stability was linked to a structural modification of microbial communities, revealing a process of adaptation to the inhibitor. *The possibility to*

acclimate the microbiome of AD to the presence of phenol was subsequently tested, see 2.3 and (Madigou et al., 2016). At still higher concentrations, the same communities remained dominant but performances decreased, meaning that microorganisms were progressively inhibited, especially archaea (from 1.00 g/L upwards). Digesters reached their limits of adaptability to the inhibitor and progressively lost efficiency. Beyond 1.50 g/L, rather similar profiles were observed for bacteria for all the concentrations. They were close to the profile of the inoculum, suggesting that even if bacteria remained active, none of them gained a significant advantage over the others compared to the initial inoculum (as the community did not change drastically, based on ARISA data). For archaea, we observed that profiles of reactors with 2.00 and 4.00 g/L of phenol were very different from profiles of lower concentrations. It was linked to the strong archaeal activity decrease observed with methane production.

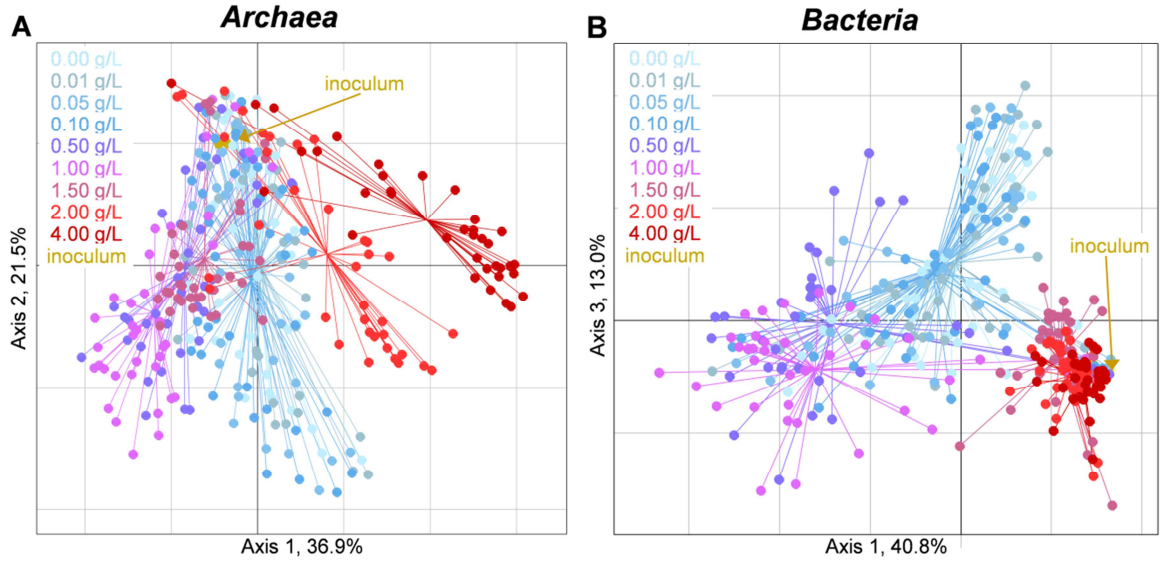


Figure 29: Principal Component Analysis (PCA) of archaeal (A) and bacterial (B) microbial diversity profiles generated by ARISA in (Chapleur et al., 2016b). The colour scale represents the initial phenol concentration in the bioreactors. For archaea, the first and second axes of the PCA provided the clearest separation of ARISA profiles, with resp. 36.9% and 21.5% of the total variance, while for bacteria, the first and third axes provided the clearest separation of ARISA profiles, with resp. 40.8% and 13.0% of the total variance.

Same approach was applied to the other studies (not illustrated in this document). Briefly, ARISA analysis in (Poirier et al., 2016a) (biowaste digestion inhibited by phenol) indicated that phenol induced microbial reorganizations from 0.50 g/L for archaea (gradual influence of phenol up to 2.00 g/L), and from 0.75 g/L for bacteria whereas AD final performances were not impaired up to 1.00 g/L of phenol. Above 1.25 g/L, all bacterial ARISA profiles were rather similar. These results were in compliance with (Chapleur et al., 2016b) despite the differences between the experiences (inoculum, substrate).

In (Poirier et al., 2016b) (biowaste digestion inhibited by ammonia), ARISA evidenced that from 0.0 g/L to 10.0 g/L of TAN, both archaeal and bacterial populations were progressively affected by ammonia. It seems that they were able to rapidly and gradually modify their composition to resist to increasing levels of TAN. 10.0 g/L appeared as a threshold beyond which bacterial communities were strongly modified (25.0 and 50.0 g/L of TAN) and archaea could not adapt.

Even if they cannot yield the specific phylogenetic identities of key phenol-tolerant or ammonia-sensitive microorganisms, ARISA fingerprinting profiles provided qualitative information on the response of archaeal and bacterial communities to different levels of inhibitors. Used at different time points, they also provided a powerful tool to evidence the dynamics of microbial communities in my work.

1.3.2. Identification of the key phylotypes at stake

The purchase of a benchtop sequencer in our lab gave me the opportunity to use 16S metabarcoding on samples from (Poirier et al., 2016a; Poirier et al., 2016b) (biowaste inhibition by phenol or ammonia). We were able to go further than the analysis of ARISA fingerprinting profiles that had initially been planned and in particular we identified the key phylotypes at stake during the inhibition. For both experiment, 44 samples were sequenced with 16S metabarcoding of DNA for each condition tested (different time points within one replicate).

In this document I illustrate the approach with the results of (Poirier et al., 2016a) (biowaste inhibition by phenol). Relative abundances of the archaeal and bacterial OTUs are presented in Figure 30 and Figure 31. This data was processed with PCA to identify the key phylotypes and their threshold of inhibition (not illustrated in this document).

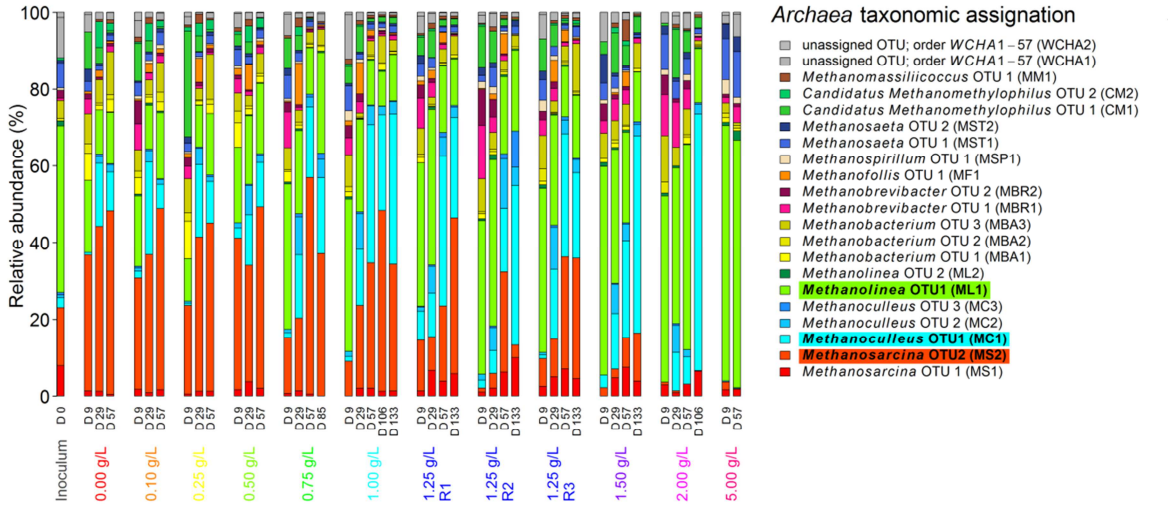


Figure 30: Relative abundances of the archaeal genera generated by 16S rRNA gene sequencing in (Poirier et al., 2016a). Samples were clustered by initial phenol concentration (g/L) and by collection date (number of days). Different shades of the same colour were used to represent archaeal OTUs belonging to the same genus.

Up to 0.50 g/L, phenol did not influenced archaeal composition much. Then a gradual shift between one OTU from *Methanosarcina* genus (MS2 in red) and one OTU from *Methanoculleus* genus (MC1 in blue) was observed along with increasing phenol concentrations from 0.50 g/L up to 2.00 g/L. We hypothesized that *Methanoculleus* played a key role in the maintenance of active methanogenesis through the establishment of specific syntrophic interactions probably with bacterial syntrophic acetate oxidizers. *Methanoculleus* was indeed shown to be the predominant group of methanogens in digesters with syntrophic acetate oxidation (SAO) (Schnürer et al., 1999). However, when inhibitory pressure became

too important, *Methanoculleus* was in turn inhibited. Archaeal community was then dominated by *Methanolinea* (ML1 in green) probably not much active but highly abundant in the inoculum.

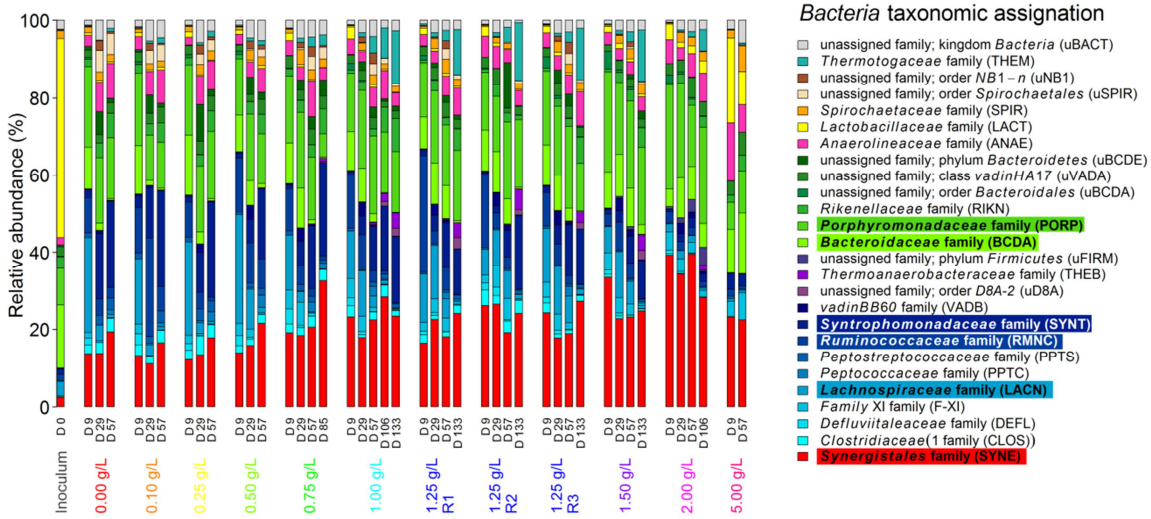


Figure 31: Relative abundances of the bacterial families generated by 16S rRNA gene sequencing in (Poirier et al., 2016a). Samples were clustered by initial phenol concentration (g/L) and by collection date (number of days). Different shades of the same colour were used to represent bacterial OTUs belonging to the same family.

Similarly to archaea, from 0.00 g/L up to 0.50 g/L, bacterial composition dynamics was not influenced by phenol. From 0.75 g/L up to 1.50 g/L, increasing phenol concentrations caused important modifications within dominant communities. Relative abundance of family *Synergistaceae* (SYNE in red) increased along with increasing initial phenol concentrations until reaching 24%-33%. Bacteria affiliated to this family have been described as anaerobic amino-acid degraders and syntrophic acetate oxidizers able to cooperate with *Methanosaeta* (Ito et al., 2011). We hypothesized that they had a potential role in SAO with *Methanoculleus* (not described yet), which was the predominant methanogen at high inhibitory concentrations. In the meantime abundance of *Syntrophomonadaceae* (SYNT in dark blue) dropped. *Syntrophomonadaceae* have also already been associated with SAO (Schnürer et al., 1999; Carballa et al., 2015).We hypothesized that the increasing levels of phenol lead to a shift of syntrophic interaction and that SAO could be maintained due to a possible functional redundancy between *Syntrophomonadaceae* and *Synergistaceae*. At 2.00 g/L, we also noticed that the relative abundance of families belonging to order *Clostridiales* dropped (in particular families *Lachnospiraceae* (LACN in cyan) and *Ruminococcaceae* (RMNC in blue)), leading to the predominance of populations assigned to order *Bacteroidales* (in particular families *Bacteroidaceae* (BCDA in light green) and *Porphyromonadaceae* (PORP in green). We suggested that they could be part of the bacterial key players at high phenol concentrations.

Same approach was applied to ammonia inhibition study (Poirier et al., 2016b). We observed a gradual methanogenic shift between two OTUs from genus *Methanosarcina* when TAN concentration increased up to 25.0 g/L. Proportion of potential syntrophic microorganisms such as *Methanoculleus* and *Treponema* progressively rose with increasing TAN up to 10.0 and 25.0 g/L respectively, while *Syntrophomonas* and *Ruminococcus* groups declined. In 25.0 g/L assays, *Caldicoprobacter* were dominant.

Among other things, both studies presented in this paragraph emphasized the key importance of establishing or maintaining syntrophic interactions to resist to the inhibitors and maintain methane production.

1.4. To be continued

To carry on with this topic, new experiments have been set-up by Laëtitia Cardona during her PhD. In particular, as the inhibition involves many rearrangements of the microbial community that are limited by the growth of microorganisms, we questioned the effect of adding ammonia at different speeds in the reactors. These experiments were performed in 6 semi-continuous reactors, more complex to operate and monitor than batch reactors but closer from industrial systems. Microbial dynamics were targeted with both 16S RNA gene metabarcoding, and 16S RNA metabarcoding, that better reflects the activity of the microorganisms. Patterns of waste degradation across time and under the different types of inhibition were monitored with metabolomics. Data of this experiment are still being analysed (see also chapter 3, section 3.2) but the new experimental design and the new analytical methods that have been used will enable to get further insights in the fate of AD microbiome under inhibitory pressure.

2. Mitigation of the inhibition

In parallel to the work presented in the first section of this chapter, I also evaluated the possibility to mitigate the inhibition of AD by phenol and ammonia. Based on the observations made in the experiments presented above, in particular that (1) establishing or maintaining syntrophic interactions was essential to resist to the inhibitors and that (2) the microbiome could adapt to resist to the inhibitor up to a certain concentration, I designed 3 experiments.

Two of them evaluated the possibility to use different support media to facilitate the resistance of the microbiome to ammonia and phenol.

- Improving anaerobic digestion with support media: Mitigation of ammonia inhibition and effect on microbial communities (Poirier et al., 2017)
- Support media can steer methanogenesis in the presence of phenol through biotic and abiotic effects (Poirier et al., 2018a)

One of them evaluated the possibility to acclimate the microbiome to the presence of phenol.

- Acclimation strategy to increase phenol tolerance of an anaerobic microbiota (Madigou et al., 2016)

The next paragraphs present a brief review of the strategies used to mitigate inhibition by ammonia and phenol, with a focus on the use of support media. The main findings of my work related to this topic are then presented.

2.1. State of the art related to the strategies used to overcome inhibition

2.1.1. Multiple strategies have been considered

Different strategies were studied to mitigate the ammonia or phenol inhibition at different steps of the process.

Some of them consist in reducing the level of the inhibitors in the reactors, for example by dilution (Yun et al., 2016). In the case of livestock effluents, many studies use an air stripping process to remove excess ammonia nitrogen (Lei et al., 2007; Abouelenien et al., 2010; Serna-Maza et al., 2015). Desalination cells (Zhang and Angelidaki, 2015) or electrochemical systems are used to remove ammonia from recycled effluent or digestate (Desloover et al., 2015).

Some of them consist in modifying the operating conditions to reduce the inhibition pressure or facilitate the action of the microorganisms. For example, trace elements necessary for the growth of the microorganism can be added (Uludag-Demirer et al., 2008; Westerholm et al., 2015). In order to counteract ammonia inhibitory effect, pH and temperature adjustment have also been implemented (Strik et al., 2006; Wang et al., 2014). Anaerobic co-digestion allows to balance the C/N ratio by mixing at least two substrates, one with a low protein content (Zhang et al., 2013), resulting in a lower level of ammonia in the reactors.

Some of them consist in modifying the microbial community so that it becomes more resistant. It includes microbial acclimation (Abouelenien et al., 2009; Gao et al., 2015) and bioaugmentation to enhance the resistance of the microbial community by the addition of adapted preground cultures (Hajji et al., 2000; Westerholm et al., 2012; Fotidis et al., 2014).

More recently, many studies underlined that the addition of different materials or polymers to AD process led to significant increase in organic matter degradation, as well as in biogas production, in particular under inhibitory conditions. They are detailed in the next paragraph.

2.1.2. Use of active support media

Recent studies highlighted that adding support media in anaerobic digesters could result in an increased biogas production despite the presence of inhibitors by specifically shaping the microbial ecosystem (Milan et al., 2010; Mumme et al., 2014). One of the hypotheses is that microbial colonization of these supports could enhance proximity of microorganisms in aggregates and favour cross-feeding, metabolic end products of one species becoming substrates for nearby species until the final end products. In particular, syntrophic associations of anaerobic digestion could benefit a lot from this type of structures as they require that the pool size of the shuttling intermediate is kept low to allow efficient cooperation. However, even if the positive effect of these supports was evidenced, clear explanations on the underlying mechanisms are still missing. In particular, to go further in the optimization of this operating strategy in industrial processes, a better understanding of the influence of supports on key phylotypes steering the microbial community towards enhanced biodegradation performance is needed.

Among these materials, zeolite is a natural rock known for its natural ion-exchange properties, absorptive capacity and to be a support for the biomass (Montalvo et al., 2012). It has been proven to be an effective way to mitigate ammonia inhibition due to its high adsorption capacity and selectivity for ammonium (Ho and Ho, 2012; Montalvo et al., 2014). It was also proven to favour biomass growth due to microorganisms immobilization on the surface of zeolite particles (Weiss et al., 2013). Authors also reported that zeolites, presenting a high local conductivity, could facilitate the interactions between microorganism (Montalvo et al., 2012).

Electrically conductive materials such as biochar or activated carbon are known to promote interspecies electron exchange and to enhance specific electro-active microorganisms improving AD performance (Lee et al., 2016; Lu et al., 2016). They were proven to promote Direct Interspecies Electron Transfer (DIET) between *Geobacter* and *Methanosarcinales* (Liu

et al., 2012). DIET is particularly interesting as it may be a more effective mechanism for interspecies electron exchange than indirect transfer via the production of reduced molecules such as H_2 and formate (Subramanyam, 2013). Biochars demonstrated significant improvements within AD process facing ammonia inhibition (Mumme et al., 2014; Lu et al., 2016). Finally activated carbon is known to adsorb organic molecules such as phenol.

The supplementation of the reactors with biopolymers such as chitosan can also enhance granulation (Hudayah et al., 2016) and specific methanogenic activity (Lertsittichai et al., 2007).

2.2. Effect of active support media to mitigate the inhibition of AD

2.2.1. Experimental design

To evaluate the effect of support media on anaerobic digestion under inhibitory conditions, three experiments were set-up in parallel, in batch digesters as described in 1.2.1. In the first series, no inhibitor was added. In the second series, 1.3 g/L of phenol was added. In the third series, 19 g/L of TAN was added. These values corresponded to the IC50 values determined in the experiment described in 1.2.3. The objective was to bring a high, but not too important inhibitory pressure.

For each series, 6 conditions were tested in triplicate. Five support media were separately added within batch digesters (2 zeolites, 2 activated carbons and 1 chitosan). One control without support media was also set-up.

Experimental design is presented in Figure 32.

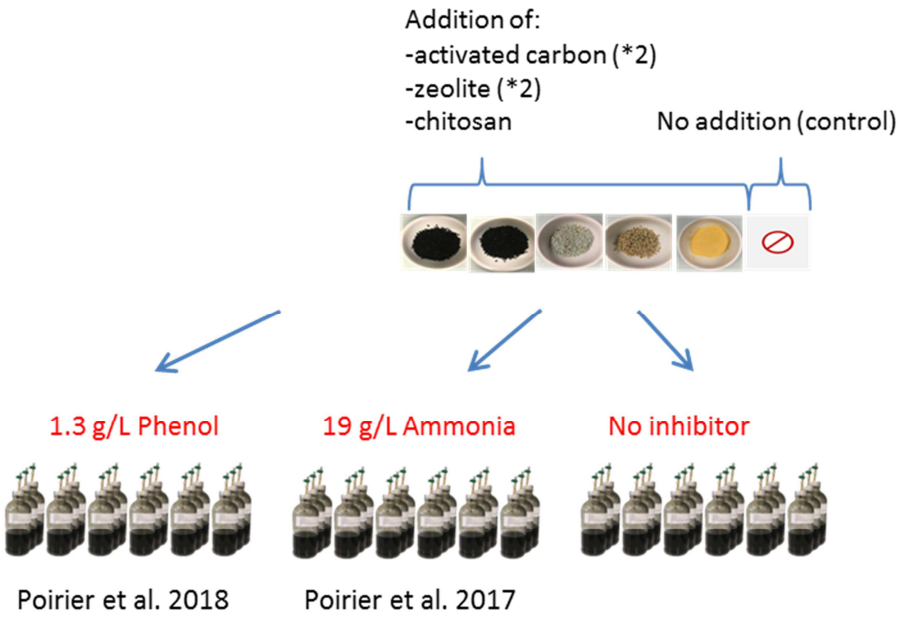


Figure 32: Experimental design used to test the influence of support media to mitigate the inhibition.

2.2.2. *Effect of the support media on biogas production*

First observation was that, in the experiment without inhibitor, CH_4 and CO_2 production kinetics remained relatively similar and were not affected by the addition of support media (not illustrated).

On the contrary, when the digesters were inhibited, an important effect of the support media on gas production was observed (Figure 33 and Figure 34), especially for methane.

For digesters inhibited by phenol (Figure 33), in all cases, addition of support media considerably reduced the lag phase observed for methane production in digesters inhibited without support media. It also increased the maximum production rate of biogas in general. Digesters with activated carbon exhibited a biogas production close to the production of non-inhibited digesters. It was explained by the rapid adsorption of phenol on the activated carbon (abiotic effect). The other support media (zeolite and chitosan) also improved the performances despite they had no influence on the phenol concentration in the reactor. We hypothesized that they had a direct effect on the microbial communities (biotic effect). Ultimate gas production was increased compared to the control without inhibitor as phenol was degraded and metabolized into biogas.

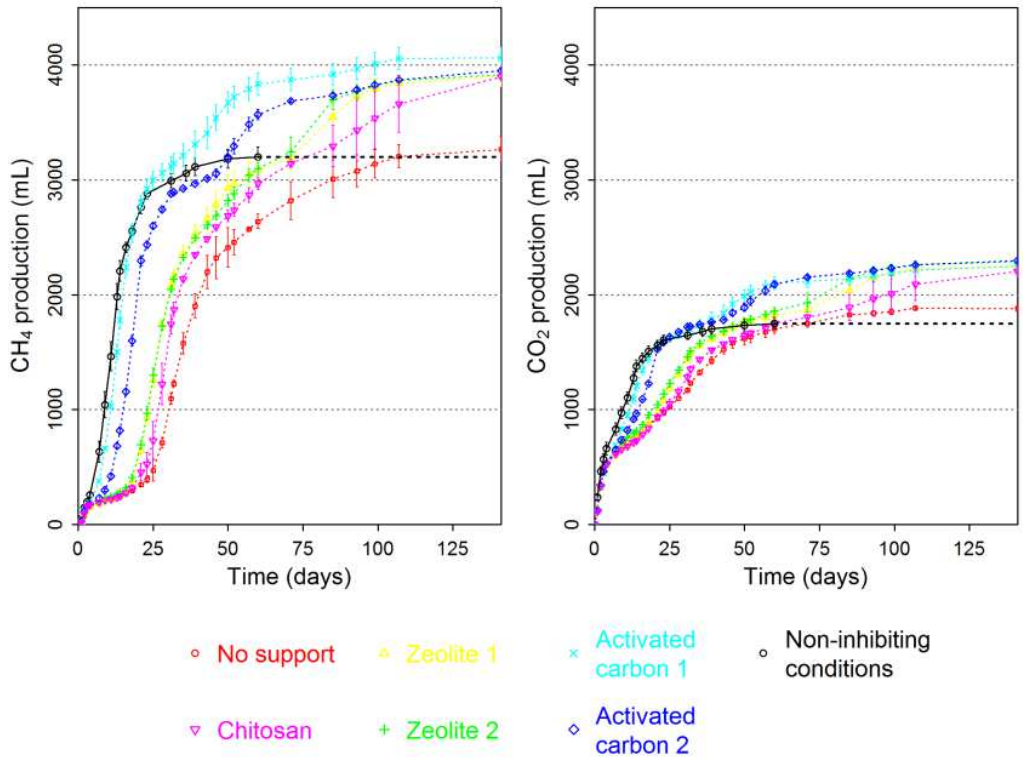


Figure 33: Cumulated CH_4 and CO_2 production (mL) over time (number of days) for the different support media initially added in the presence of 1.3 g/L of phenol. (Poirier et al., 2018a). Mean values of each triplicate of bioreactors are presented for CH_4 and CO_2 production and error bars represent standard deviation within triplicates. The black line represents the mean values of all bioreactors for CH_4 and CO_2 production under non-inhibiting conditions.

For digesters inhibited by ammonia (Figure 34), both zeolites mitigated ammonia inhibition and enhanced anaerobic digestion compared to inhibited digesters without support. Surprisingly, the distinct additions of activated carbons revealed contrasted results. In presence of activated carbon 1, lag time was reduced by 10 days while activated carbon 2 led to an increase of 21 days of the lag time compared to inhibited digesters without support. Similarly, the addition of chitosan also led to an extra inhibitory effect while no effect of chitosan was observed during the control experiment without ammonia. The additional inhibitory effect brought by chitosan and activated carbon 2 was not clear but was probably due to a synergistic effect between a molecule contained in these media and ammonia. Indeed no adverse effect had been observed in bioreactors supplemented with these support media in reactors not inhibited.

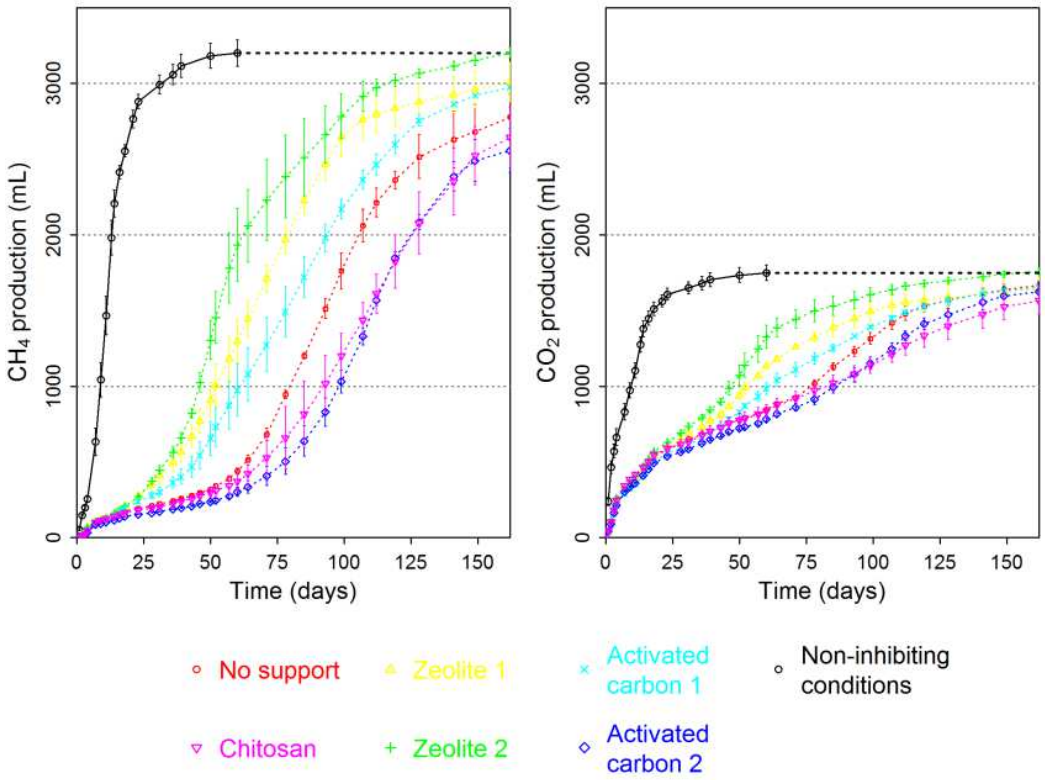


Figure 34: Cumulated CH_4 and CO_2 production (mL) over time (number of days) for the different support media initially added in the presence of 19 g/L of TAN. (Poirier et al., 2017). Mean values of each triplicate of bioreactors are presented for CH_4 and CO_2 production and error bars represent standard deviation within triplicates. The black line represents the mean values of all bioreactors for CH_4 and CO_2 production under non-inhibiting conditions.

Once the positive effect of the support media to mitigate inhibition of AD was confirmed, we focused on understanding their effect on AD microbiome.

2.2.3. Effect of active support media on the microbial community

For the different experiments, microbial dynamics were investigated with 16S RNA gene metabarcoding. Several samples were analysed for each condition. Consistently with

performance results, we observed that archaeal and bacterial compositions were not significantly influenced by supports media addition in not inhibited reactors. On the contrary, in presence of inhibitors, support media had an influence. Two different approaches were used to analyse the data and highlight the key players among the important number of OTUs identified in each experiment.

For the experiment with phenol (Poirier et al., 2018a), effect of the support media was mainly observed on the bacterial population. A partial least squares-discriminant analysis (PLSDA) was performed to identify the OTUs characteristics of the different support media. PLSDA is a discriminant analysis used to sharpen the separation between groups. It enables to understand which variables (OTUs in our case) best describe the differences between the groups. This method is particularly useful to treat high-dimensional data (Lee et al., 2018). Sparse version of the PLSDA (Sparse-PLSDA) (Lê Cao et al., 2011) was used to perform variable selection into the large data set in order to identify the most discriminant OTUs.

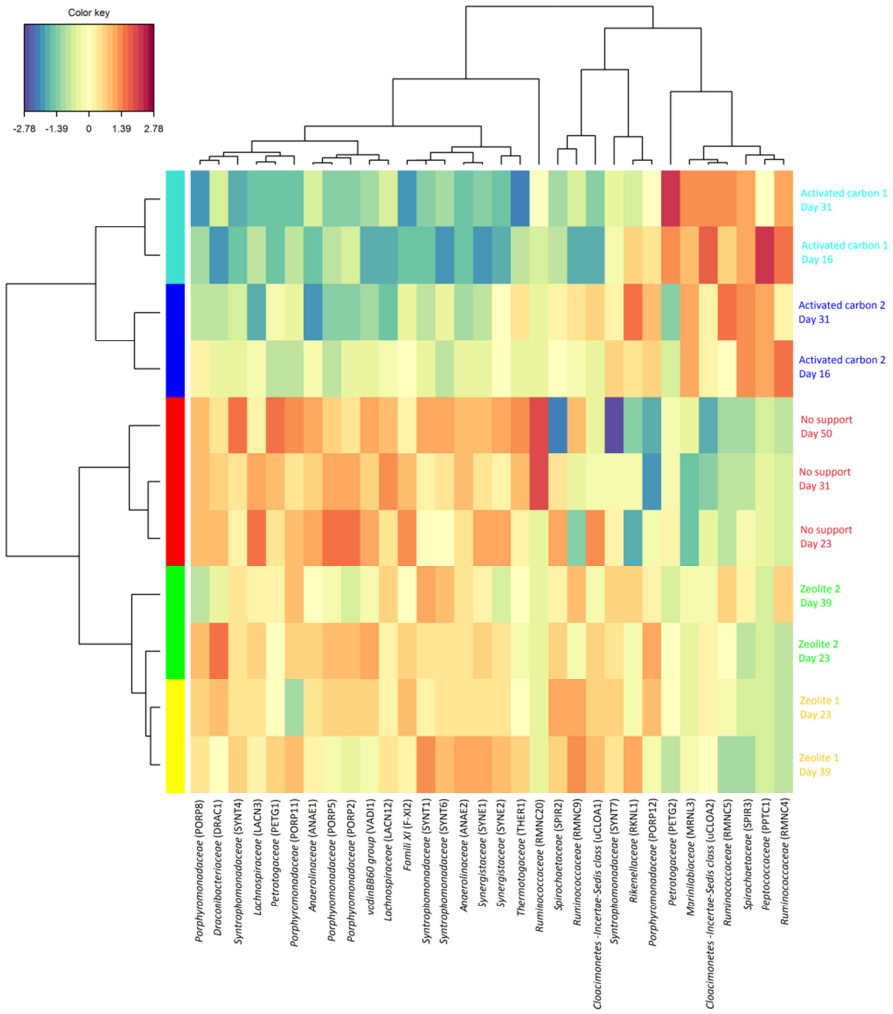


Figure 35: Heatmap of the most discriminant bacterial OTUs (in column) determined after sparse Partial Least Squares Discriminant Analysis (sparse-PLSDA) of all OTUs generated by 16 S rRNA gene sequencing for the different support media initially added in the presence of 1.3 g/L of phenol. (Poirier et al., 2018a). Type of support and date of sampling is indicated on the right. Name of the OTUs is indicated at the bottom. The colour scale on the left represents the grouping of samples used to compute the sparse-PLS-DA (type of support media). The colour key of the heatmap shows the abundance of the OTUs (from blue=low abundance to brown red=high abundance).

Heatmap of the most discriminant bacterial OTUs identified with Sparse-PLSDA is presented in Figure 35 (for technical reasons, samples from digesters with chitosan were not included in the analysis). In the lines, the samples were clustered based on the similarity of their abundance in the selected OTUs. The different samples corresponding to the same support media were clustered together, and characterized by specific OTUs. For example, digesters with the two types of activated carbon were rather similar. They were associated to the presence of OTUs assigned to genera *Ruminofilibacter* (MRNL3) and *Ruminococcus* (RMNC4 and RMNC5) but also to genera *Cryptanaerobacter* (PPTC1) and *Spirochaeta* (SPIR3). Genera *Ruminofilibacter* and *Ruminococcus* contains rumen species related to the degradation of xylan (Nissilä et al., 2012) while *Cryptanaerobacter* consists of species involved in phenol degradation into benzoate via 4-hydroxybenzoate (Juteau et al., 2005). This figure also highlighted that activated carbon 1 could specifically be related to the presence of an unassigned OTU belonging to family *Petrotogaceae* (PETG) that is known to be involved in phenol degradation (Na et al., 2016). Similar types of conclusions were drawn for the digesters incubated with zeolites and without support. sPLSDA enabled to focus on the most discriminant OTUs of each condition. Specific microbial populations could be related to all types of support evidencing their biotic influence on the microbiome.

For the experience with ammonia (Poirier et al., 2017) a different approach was used. Ratios of abundances were calculated to assess the over/subrepresentation of the different archaeal OTUs and bacterial families in presence of the different support media compared to a 'control' situation. 'TAN ratio' compared the inhibited digesters to the non-inhibited digesters. 'Support ratio' compared the inhibited digesters with support media to the inhibited digesters without support media. As the sPLSDA, this method enabled to highlight the taxons differentially expressed between the conditions.

More precisely, for each sample, the relative abundance of the main archaeal OTUs or bacterial families was divided by the corresponding relative abundance in samples from the control. These ratios are presented in Figure 36 and Figure 37 for bacterial families, taking respectively non-inhibited digesters and inhibited digesters without support media as 'controls'. Colour scale corresponds to the type of support media and a size corresponds to the number of days of incubation.

Figure 36 showed that TAN mainly inhibited bacterial families affiliated to orders *Bacteroidales, Spirochaetales, Anaerolinales* and *Synergistales*. By contrast, most of the families assigned to order *Clostridiales* and *Thermotogales* seemed to be more resistant to this inhibitor. *Syntrophomonadaceae* and *Clostridiaceae 1* were the two only principal families belonging to order *Clostridiales* which were also impaired by TAN. Therefore it was hypothesized that TAN induced strong reorganizations within bacterial community by favouring the presence of order *Clostridiales* and *Thermotogales* at the expense of other orders. An influence of the type of support media was observed and was assessed more precisely by calculating other ratios (Figure 37).

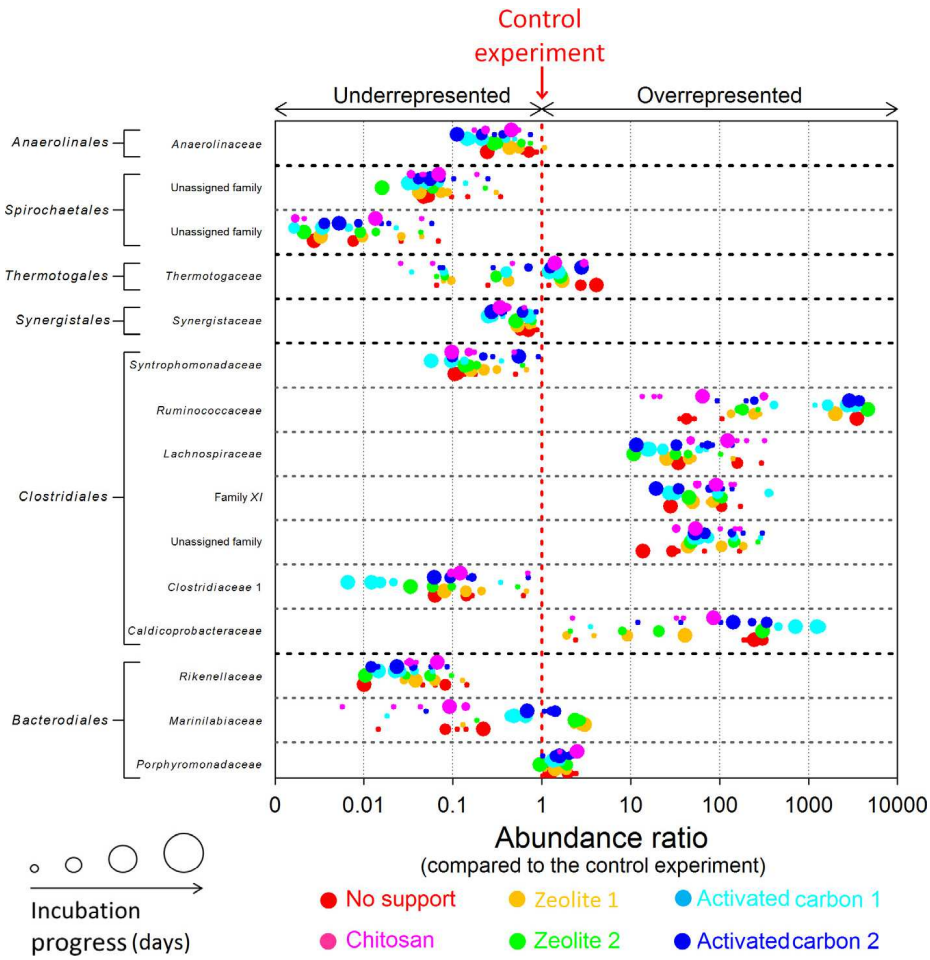


Figure 36: ‘‘TAN ratios” calculated for main bacterial OTUs. ‘‘TAN Ratio” was obtained by dividing, for each sampling date, the relative abundance of the main bacterial OTUs in the inhibited experiment by their corresponding relative abundance in the non-inhibited experiment. The size of the symbols is correlated to the collection dates: the more advanced the bigger.

Figure 37 showed the specific effect of the support media on the bacterial community, under inhibitory conditions, in comparison to an inhibited situation without support media. The effect of support media addition was minimal on bacterial order *Anaerolinales*, *Spirochaetales*, *Thermotogales* and *Synergistales*. However, inside *Clostridiales* order, both zeolites and activated carbons favoured the development of *Ruminococcaceae* family, at different levels. Activated carbon 1 favoured the development of *Caldicoprobacteraceae* family while this family was underrepresented in digesters containing zeolites. Zeolites seemed to lead to the emergence of families belonging to order *Bacteroidales* such as *Marinilabiaceae*. Activated carbons also favoured this family but less importantly.

Similar approaches were applied to archaea. They showed that under ammonia inhibition zeolite preserved *Methanosarcina* genus and enhanced *Methanobacterium* genus while activated carbon enhanced *Methanoculleus* genus emergence.

The adverse effect of activated carbon 2 and chitosan on biogas production could not be explained clearly based on microbial community patterns.

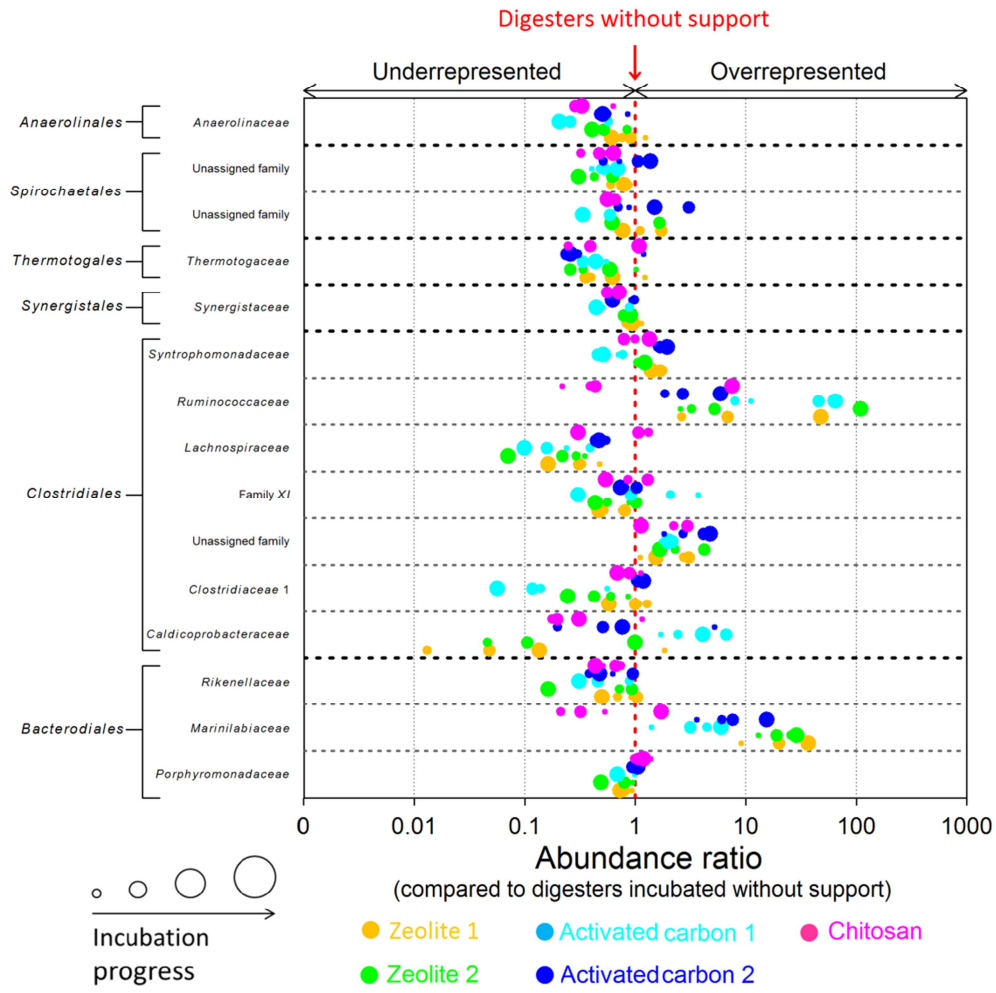


Figure 37: ‘‘Support ratios” calculated for main bacterial OTUs. ‘‘Support Ratio” was obtained by dividing, for each sampling date, the relative abundance of the main bacterial OTUs in the inhibited experiment with support media by their corresponding relative abundance in the inhibited experiment without support media. The size of the symbols is correlated to the collection dates: the more advanced the bigger.

These results highlighted potentialities to use supports to enhance the process stability of anaerobic digesters subjected to phenol or ammonia inhibition. Microorganisms positively correlated with the presence of the support media were identified. However, the exact role of the support media was not clear. The structure and the microbial colonization of the supporting materials after AD operation was assessed using fluorescent *in situ* hybridization. However, observations did not evidence clear colonization patterns. We also planned to isolate support media from the samples taken in the digesters to sequence specifically the DNA of the possibly ‘attached’ microorganisms but were not able to obtain enough material.

I recently set up an experiment (PhD of Laëtitia Cardona) to test different assumptions about the mechanisms of action of the zeolite during ammonia inhibition. Analysis of the data is still in progress. In addition to the characterisation of the microbial dynamics with 16S, methanogenic pathway was monitored by measuring the isotopic fractionation of the methane (Grossin-Debattista, 2011). Patterns of waste degradation under the different conditions will be monitored with metabolomics. It will provide additional information to decipher the role of support media in mitigating the inhibition. With the data obtained so far, the schematic pathways presented in the preliminary Figure 38 were drawn for digesters without ammonia,

or with ammonia and with or without zeolite. Figure 38 highlights the main differences between the three conditions, in particular degradation products which accumulate more, methanogenic pathways which is dominant, and specific microorganisms.

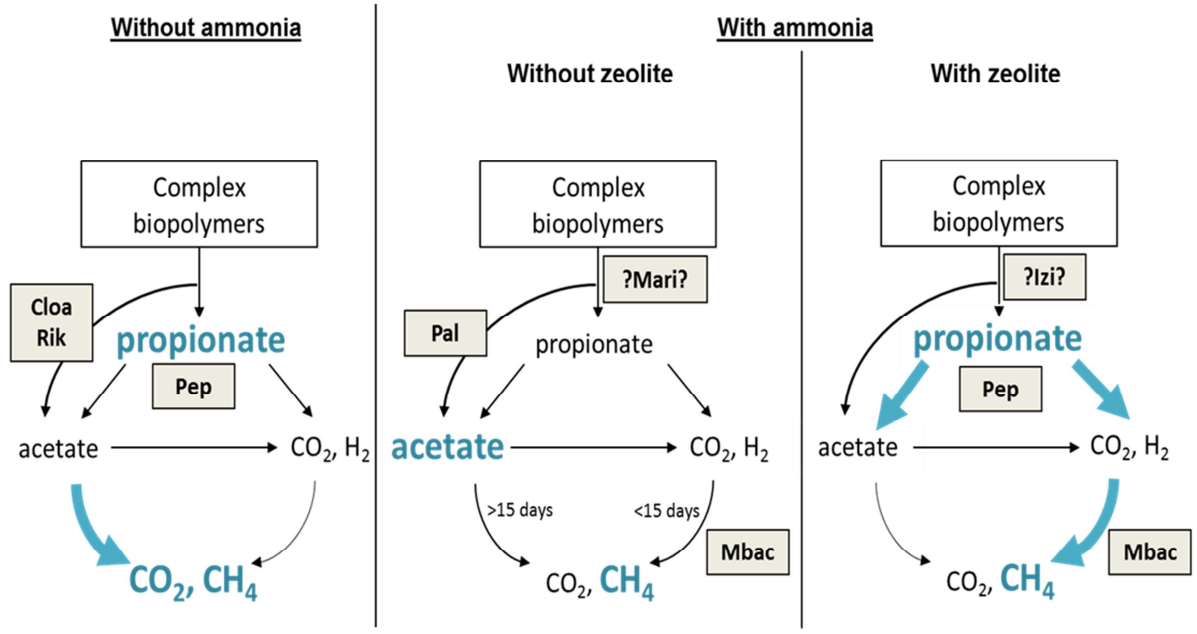


Figure 38: Summary of the hypothetical pathways for the degradation of biowaste, without ammonia and in presence of ammonia, with or without zeolite. The molecules which accumulated the more and dominant methanogenesis pathway are highlighted in blue. Specific microorganisms identified by the sparse PLS-DA are represented by grey rectangles that were placed at the position where we hypothesize they are active. The role of the microorganisms between question marks is not known. Genus Mbac: Methanobacterium; Orders Spi: Spirochaetaceae; Cloa: Cloacimonadaceae; Rik: Rikenellaceae; Pal: Paludibacteraceae; Pep: Peptococcaceae; Mari: Marinilabiliaceae; Family Izi: Izimaplasmatales.

2.3. Is the acclimation to the inhibitor possible?

I tested another strategy to mitigate the effect of the inhibitors that I summarize briefly in this section.

To confront future disturbances, McMahon suggests that the best approach could consist in provoking a preliminary adaptation of the microbial communities to stressful conditions (McMahon et al., 2004). I noticed that my batch tests with multiple concentrations of inhibitor suggested that a process of adaptation of the microbiome was possible (see 1.3.1). Based on these observations, I set up an experiment to question the possibility to acclimate AD microbiome to phenol (Madigou et al., 2016).

Our work was based on the presupposition that a regular step-wise adaptation of the microbial community to stressful conditions could strengthen the microbial ecosystem against the disturbance. Based on this presupposition, the impact of increasing phenol levels on the performances of a semi-continuous anaerobic bioreactor degrading cellulose was analysed.

This experimental device was not used in our team and it was an opportunity to initiate a work in these more dynamics systems. It allowed getting conditions closer to industrial situations than the batch digesters used in the other experiments.

Degradation performances were monitored throughout the experiment and molecular tools (16S sequencing and ARISA fingerprinting technique) were used to track changes in the microbial community. We observed that the acclimation strategy progressively minimized the effect of phenol on degradation performances. The application of 3 successive disturbance episodes enabled a considerable improvement of the microbiota resistance to phenol and total inhibition thresholds were significantly augmented, from 895 to 1942 mg/L of phenol. Microbiota adaptation was characterized by the selection of the most resistant archaea OTU from *Methanobacterium* genus and an important elasticity of bacteria, especially within *Clostridiales* and *Bacteroidales* orders, that probably enabled the adaptation to more and more stressful conditions.

The acclimation strategy minimized the effect of phenol on AD performances. It showed that adaptation of AD microbiota to increase the resistance was possible. Further studies should aim at investigating the mechanisms of this acclimation with other techniques, such as metabolomics and metagenomics or metatranscriptomics. It would enable to see for example if the microbial community adapted thanks to a functional redundancy, or through an evolution of the pathways of degradation.

3. Conclusion

Inhibition of the anaerobic digestion is a broad topic. My research activities are still strongly linked to this subject. DIGESTOMIC ANR project and PhD of Laëtitia Cardona focus in particular on the effect of ammonia, variation of substrates and modifications of temperature on the performances of anaerobic digestion. STABILICS project that I recently submitted to the ANR (see last chapter of this manuscript) aims at getting new insights in the determinants of AD stability).

My new experiments are now mainly performed in semi-continuous digesters that were acquired after I performed the experiments detailed in this chapter. They better mimic industrial digesters. Regular input of waste enables to reach steady states, while digestate outlet progressively removes non active microorganisms. New methodologies are now available to explore other ‘layers’ of information (in particular the suite of omics methods) to enrich the information already obtained. They will deepen the understanding of the inhibition process. Indeed, due to the high functional redundancy within bacterial populations, 16S characterization only does not enable to assess their respective role during anaerobic digestion.

So far I explored metabolomics in several experiments, and I just obtained new metagenomics data. Analysing and integrating these different layers of data is complex and requires appropriate statistical methods. I spent 8 months in 2017 in two institutions (Université de Toulouse and University of Melbourne) to acquire the necessary analytical and computational skills to apply these methods. The first results obtained are presented in the next chapter.

Chapter 3: Advancing environmental biotechnologies through combined use of advanced molecular ecology and cutting-edge computational statistics

To face the increasing quantities of data I was generating through experiments of the kind presented in chapter 2, I started to include different statistical analytical approaches in my research. The objectives were both to extract the most accurate and relevant information from the important datasets already produced, but also to anticipate the use of omics methodologies in my future studies.

In 2017, I had the opportunity to make scientific visits of several months in two laboratories specialist of the analysis of high-throughput biological data (Institut de Mathématiques de Toulouse at the University of Toulouse and Melbourne Integrated Genomics at the University of Melbourne). During these stays I've been through my data with different methods, to answer different questions. I benefited a lot from the help of the scientists I was visiting, but also from the others statisticians or users of computational methods I met during my stay. Some of the results obtained are presented in this chapter, which is divided in three sections.

- The first section presents a method of integration of different types of data measured on the same samples, illustrated with an analysis where we draw links between microbial activity and molecules degradation rate during a co-digestion experiment (manuscript submitted to Chemical Engineering Journal (Cardona et al., submitted)).
- The second section presents an integration of 16S data generated during different and independent studies, to identify ubiquitous bioindicators signatures of AD inhibition (independent of the studies) and build predictive models of AD inhibition (manuscript submitted to the ISME Journal (Poirier et al., submitted)).
- The third section presents different examples of analysis of longitudinal data using time course modelling: the effect of an abrupt temperature increase on the dynamics of the microbiome of AD (ongoing project), the effect of different intensities of inhibition on the microbial dynamics (ongoing project) and a new analytical framework to integrate different types of longitudinal data (manuscript under revision in Frontiers in Genetics (Bodein et al., submitted)).

1. Integration of different types of data

1.1. Objectives

As no single analytical technique is sufficient to decipher complex microbial relationships, the combination of information from several data sources has become a decisive issue. However, current statistical methods mainly enable the analysis of single microbiome dataset, and generally do not allow for the analysis of several types of data simultaneously. Consequently, data derived from each omics technique are typically studied in isolation, and disregard the correlation structure that may be present between the multiple data types. Appropriate tools are needed to handle these datasets more suitably.

The coupling of data with inherently very different structures, such as chemical analyses and metaomic responses is particularly challenging. The challenge is how to organize, analyse, gain insight from, and use the data for predictive, design, and operational purposes, such as improving the function of specific engineered bioprocesses. This is not an easy challenge, given the complexity of microbial communities of dynamic environments such as anaerobic reactors. However different integrative methods have been developed to treat such multi-table data. For example 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 different matrices (Boccard and Rutledge, 2013). It has also been used for the joint treatment of chemical and sensory data (Blackman et al., 2010). DIABLO (Singh et al., 2016) based on sparse Generalised Canonical Correlation Analysis (Tenenhaus et al., 2014) is available for the integration of multiple data sets in a supervised analysis. The methods aim to extract complementary information from several data sets to gain a better understanding of the interplay between the levels of information measured. For example, it enables to identify correlated (or coexpressed) biological features measured from the different omics while also investigating and modelling the relationship between meta-omics and performance stability of the digester. An equivalent method based on Projection to Latent Structures (PLS) enables to integrate different data in an unsupervised framework (Lê Cao et al., 2008). More precisely, PLS regression is a multivariate methodology which relates (integrates) two data matrices to evidence correlations. To go further, Sparse PLS performs simultaneously variables selection to identify the variables explaining the most the correlation between the datasets (Lê Cao et al., 2008).

These methods have so far not been used to treat data from AD. The next paragraph presents an example of the use of sparse PLS (or sPLS) to integrate two types of data. An example of the integration of three types of data is presented in section 3.3.

1.2. Linking microbial activity and molecules degradation during codigestion

We used a sPLS to integrate 16S and metabolomics data in an unsupervised framework. The objective was to investigate the link between microbial activity and molecules degradation during AD. *This work was submitted to Chemical Engineering (Cardona et al., submitted).*

Data were taken from a co-digestion experiment (Cardona et al., 2019). In this experiment the influence of different mixtures of sewage sludge with fish waste or green waste (grass) on the microbiome dynamics was investigated in batch digesters. Digesters were fed with nine different mixtures of waste as illustrated Figure 39. For all the digesters, samples were taken at two different dates during the active methane production.

RNA from the samples was extracted. Corresponding cDNA was used for 16S metabarcoding, to target the active microorganisms (microbial activity dataset). The same samples were also analysed with an untargeted LCMS approach to target the degradation of molecules from the feeding. The rate of molecules degradation in the bioreactors was estimated by dividing the molecules intensity at day 0 by their intensity at days of interest (molecule degradation rate dataset). It resulted in two datasets describing the same samples.

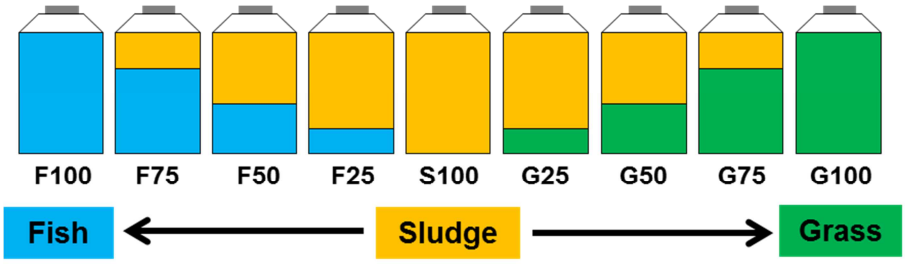


Figure 39: Experimental design used in the co-digestion study. S100 stands for wastewater sludge alone, F25, F50, F75, F100 stands for respectively 25, 50, 75 or 100% of fish (F) in co-digestion with sludge, G25, G50, G75, G100 stands for respectively 25, 50, 75 or 100% of Grass (G) in co-digestion with sludge.

A sparse PLS analysis was performed to integrate the data relative to the microbial activity with the data of the molecules degradation. The ordination plots from sparse PLS (Figure 40 A and B) showed similar dispersion patterns for the two datasets. It suggested the existence of an underlying correlation structure between the two datasets. The correlation circle plot (Figure 40 C) allowed visualizing at the same time the variables responsible for the ordination of the samples in the two datasets, namely the variables correlated between both datasets. It corresponded to the groups of active microorganisms correlated to the molecules degradation rate. To further identify the microorganisms potentially responsible of the molecules degradation, a hierarchical clustering based on the loadings of the microbial and metabolic from the PLS was performed. Five groups of correlated microorganisms and molecules were identified. The dynamics of their abundance was plotted (Figure 41).

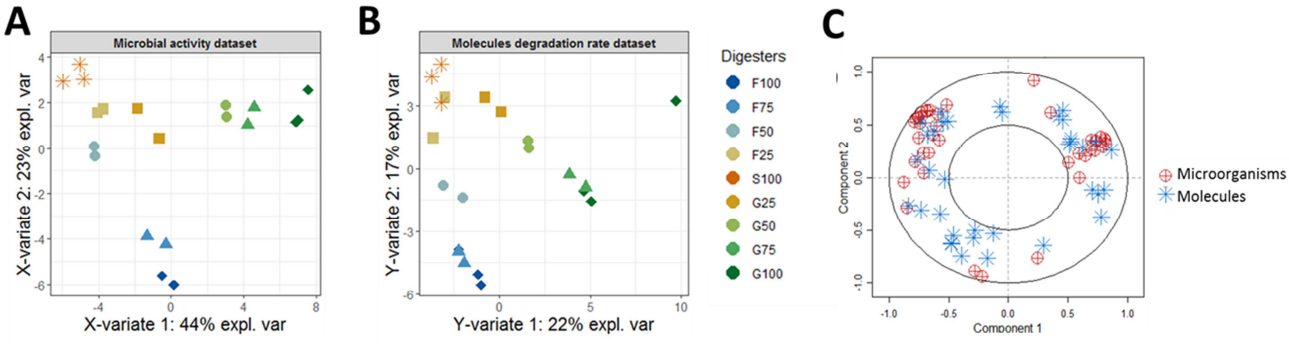


Figure 40: sPLS data integration of the microbial activity and molecules degradation rate. Sample ordination plots according to the microbial activity (A) and molecules degradation dataset (B) show a similar influence of the feeding composition on the microorganisms and on the degradation of the molecules. The correlation circle plot (C) shows the microorganisms and the molecules with similar dynamics across the different samples.

More specifically, Figure 41 depicts the mean values and standard deviation of the microbial activity and molecule degradation rate according to the feeding types for the 5 groups. Group 1 included microorganisms and molecules with a high microbial activity and high molecule

degradation rate during the digestion of sludge and grass. Groups 2-4 included microorganisms and molecules that were specific of either sludge, grass, or fish, respectively. Finally, group 5 included the microorganisms and molecules that were highly active and highly degraded, respectively, in fish and sludge bioreactors.

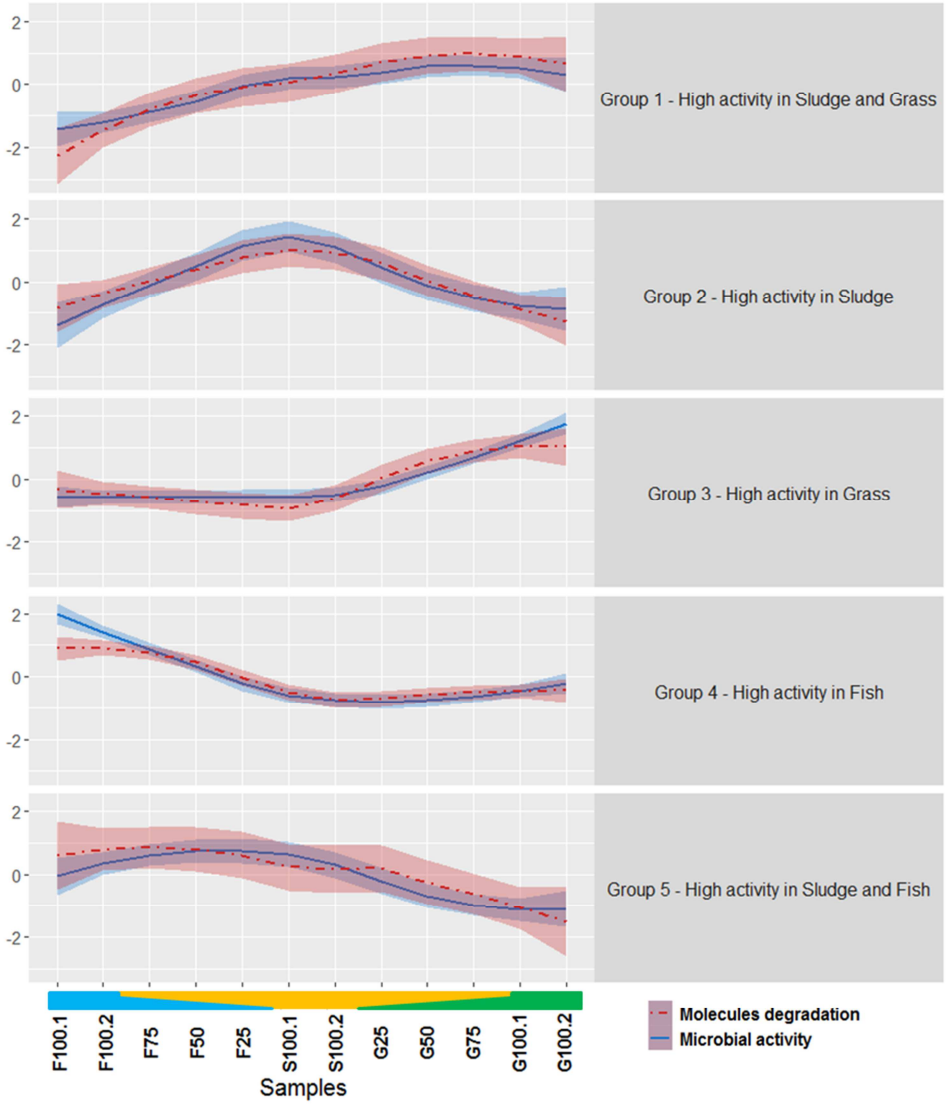


Figure 41: Dynamics of the active microorganisms correlated to the molecules degradation among the samples at day 21. Duplicates were carried out on the bioreactors containing only fish, grass or sludge. Lines represent the mean values of the microbial activity (solid blue line) or molecules degradation rates (dashed red line) within each cluster, the shadowing traces represent their standard deviation.

Correlations between microbial activity and molecules degradation within the different digesters suggested that within each group the microorganisms were potentially responsible of the molecules degradation. Based on this hypothesis, we sought for biological interpretation.

For example, group 1 included two genera of archaea, *Methanosarcina* and *Methanospirullum*, and the molecules diethylthiophosphate and N-(3-methylbutyl)acetamide. Diethylthiophosphate is a pesticide degradation product and a urine metabolite (Ueyama et al., 2014). Conversely, N-(3-methylbutyl)acetamide is a metabolite found in alcoholic beverages obtained by fermentation such as beer and wine. One hypothesis explaining the correlation of these archaea and molecules could be the indirect role of the archaea in the molecules degradation through a syntrophic interaction with bacteria.

In group 4, *Methanosarcina* and 2 OTUs from the order *Clostridiales* were correlated to molecules that can be classified as amino acids degradation products. Cadaverine and 5-aminopentanoic acid are obtained from L-lysine degradation, histamine from L-histidine degradation, and phenylpyruvic acid from L-phenylalanine degradation. The presumed role of these microorganisms in the degradation of cadaverine and L-histidine was supported by previous studies. Roeder and Schink described a new strain close to *Clostridium aminobutyricum*, able to degrade cadaverine, in co-culture with the archaea *Methanospirullum* (Roeder and Schink, 2009). On the other hand, some *Clostridium* were also identified to be involved in the histamine degradation (Pugin et al., 2017).

The statistical method developed in this study allowed to posit hypotheses on the degradation of molecules by different microorganisms. These hypotheses were consistent in regards with the literature. *Another example of multiple types of data integration is presented in section 3.3.*

2. Integration of data from independent studies

2.1. Context

In the biomedical field, high-throughput technologies, based on DNA or RNA-sequencing or omics approaches, are now being used to identify biomarkers or gene signatures that distinguish disease subgroups, predict cell phenotypes or classify responses to treatments (Lee et al., 2019; Sun et al., 2019). Same types of approaches could be applied to microbial ecology of bioprocess to identify biomarkers signatures (based on a 16S or omic data) that could for example characterize a type of inhibition, be used to forecast digesters performances, to evidence that a specific pathways is active etc.

To be robust, these biomarkers should be reproducible from a study to another. However, it is observed that few of the findings are reproduced when assessed in subsequent studies in general (Rohart et al., 2017b). The poor reproducibility of identified signatures is most likely a consequence of high-dimensional data, in which the number of OTUs, genes or transcripts being analysed is very high (often several thousands) relative to the comparatively small sample size being used (generally less than 20). In that case, identified signatures are likely to be very specific of the samples and of the study. More ubiquitous biomarkers could be identified by increasing the number of samples. One easy way to increase sample size is to combine raw data from independent experiments in an integrative analysis. This can improve both the statistical power of the analysis and the reproducibility of the signatures that are identified.

However, integrating studies with the aim of classifying biological samples based on an outcome of interest has a number of challenges. Omics or metabarcoding studies often differ from each other in a number of ways, such as their experimental protocols or the technological platform used. These differences can lead to so-called ‘batch-effects’, or systematic variation across studies, which is an important source of confounding factors (Gagnon-Bartsch and Speed, 2012). They must be accounted for when combining independent studies to enable genuine biological variation to be identified. Different approaches have been developed for that, including Multivariate INTegrative method, MINT in mixOmics R package (Rohart et al., 2017b). It integrates independent data sets while simultaneously accounting for unwanted study variation, classifying samples and identifying key discriminant variables. It also enables to predict the class of new samples from external studies. More precisely, MINT seeks for a common projection space for all studies that is defined on a small subset of discriminative variables and that display an analogous

discrimination of the samples across studies. The identified variables share common information across all studies and therefore represent a reproducible signature that helps characterising biological systems.

I applied this tool to identify biomarkers of AD inhibition and build predictive models of AD inhibition by ammonia and phenol.

2.2. Integration of independent studies to identify microbial bioindicators

2.2.1. Description of the data

I selected four studies assessing the influence of ammonia or phenol on anaerobic digestion to identify biomarkers, build predictive models and assess their performance.

- Study 1 (Poirier and Chapleur, 2018a) corresponds to the experiments testing the influence of different ammonia and phenol levels in digesters described in chapter 2 (papers: (Poirier et al., 2016a; Poirier et al., 2016b)).
- Study 2 (Poirier and Chapleur, 2018b) corresponds to the experiments testing the influence of different support media to mitigate the inhibition by ammonia and phenol described in chapter 2 (papers:(Poirier et al., 2017; Poirier et al., 2018a)).
- Study 3 (Lu et al., 2016) was an external study testing the influence of support media (biochar) to reduce the inhibition of ammonia in digesters fed with glucose.
- Study 4 (Peng et al., 2018) was an external study testing the influence of ammonia on the AD of food waste.

In all the studies, samples were taken across time and under different inhibitory conditions. All studies included non-inhibited controls or period. Based on the performance data (biogas production, VFA accumulation), and conclusions of the authors, we determined if the samples had been collected in bioreactors inhibited by ammonia, phenol, or in non-inhibited bioreactors.

In all these studies, DNA was extracted from the samples and 16S rRNA gene was sequenced providing datasets of raw sequences associated to different inhibitory conditions. As described in Table 1, the four studies differed by different parameters. Different types of substrates and inoculum were used, 16S targeted region, sequencing platform and technology also differed. We analysed the raw sequences from studies 1 and 2, study 3 and study 4 independently with the same bioinformatics approach, using FROGS pipeline (Find, Rapidly, OTUs with Galaxy Solution) (Escudié et al., 2018).

Table 1: Summary of the main characteristics of the different studies used in the analysis, to illustrate the possible confounding factors. Same colour was used when the characteristics were similar.

Study	Inhibitors	Lab	Substrate	Inoculum	Use of support media	Sequencing technology, 16S targeted region,
1	Ammonia or Phenol	A	Foodwaste A	Inoculum A	No	IonTorrent – PGM, V4-V5
2	Ammonia or Phenol	A	Foodwaste A	Inoculum B	Yes	IonTorrent – PGM, V4-V5
3	Ammonia	B	Glucose	Inoculum C	Yes	Illumina MiSeq, V4-V5
4	Ammonia	C	Foodwaste B	Inoculum D	No	Illumina MiSeq, V3-V4

2.2.2. Accounting for study effect

Our first objective was to identify biomarkers of ammonia and phenol inhibition and was focused on studies 1 and 2, as studies 3 and 4 did not contain samples inhibited by phenol.

A PCA was performed (Figure 42A) for a first exploration of the major sources of variation in the data. Sample distribution highlighted a strong study effect. Samples on the left part of the factorial plane were related to study 1 conducted with the inoculum A while samples collected during study 2 conducted with inoculum B were on the right side of the factorial plane. However, a clear influence of the type of inhibition on microbial community could still be observed within the samples of each study.

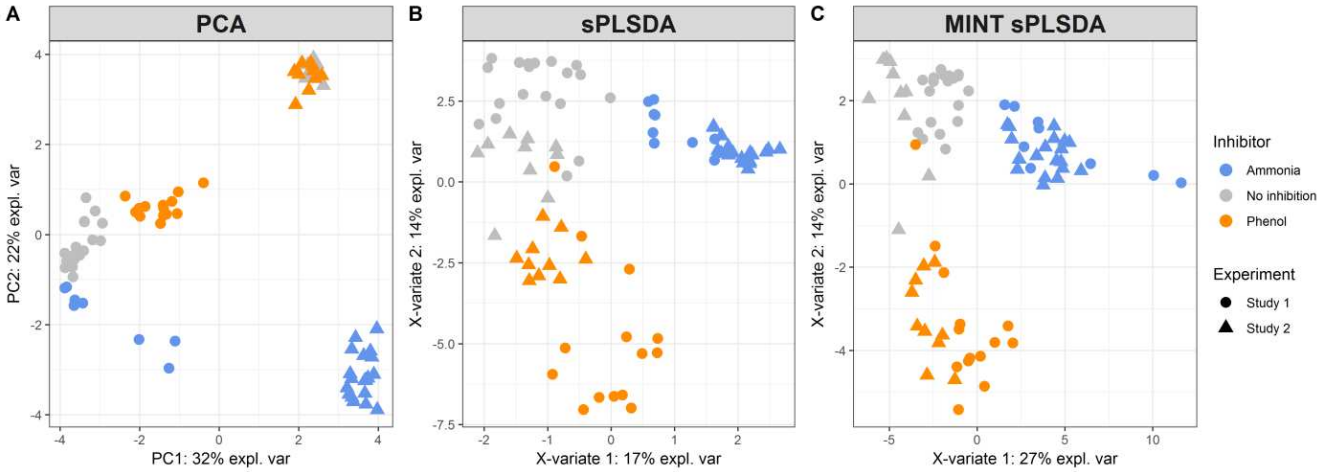


Figure 42: (A) Principal Component Analyses (PCA), (B) Sparse Partial Least Squares Discriminant Analysis, (C) Multivariate Integrative Sparse Partial Least Squares Discriminant Analysis of OTUs distribution in samples from studies 1 and 2, inhibited by phenol, ammonia or not inhibited.
On the factorial maps each sample is represented with a coloured marker. The colour scale represents the type of inhibitor. The type of marker represents the study. OTU data was generated by 16S rRNA gene sequencing.

A supervised PLS-DA model was then fitted on the data. Sparse version of the method was applied to select features and identify discriminative OTUs that best described the difference between the groups of samples (Lê Cao et al., 2011). Samples distribution based on the first two components is presented on Figure 42B. sPLS-DA model enabled to mitigate the study effect compared to the unsupervised PCA. However, within each condition, the study effect was still present: each sample collected in Study 1 was clearly separated from the ones collected in Study 2.

In order to counteract this bias, we applied MINT (Rohart et al., 2017b). Tuning of MINT indicated that an optimal number of 45 OTUs should be selected to achieve the lowest classification error rate. Samples representation from MINT is presented in Figure 42C. It evidenced that the study effect was accounted for, with the strongest separation observed according to inhibiting condition rather than studies.

2.2.3. Identification of microbial bioindicators of inhibition

Microbial signatures identified with MINT were output in a clustered image maps (81 samples and 45 OTUs) in Figure 43. This representation confirmed that, based on their microbial community composition, samples could be grouped by inhibition type (non-

inhibited samples, samples inhibited by ammonia and samples inhibited by phenol), and that this classification was independent of the study. The 45 OTUs selected by MINT were clustered into five different groups (A to E).

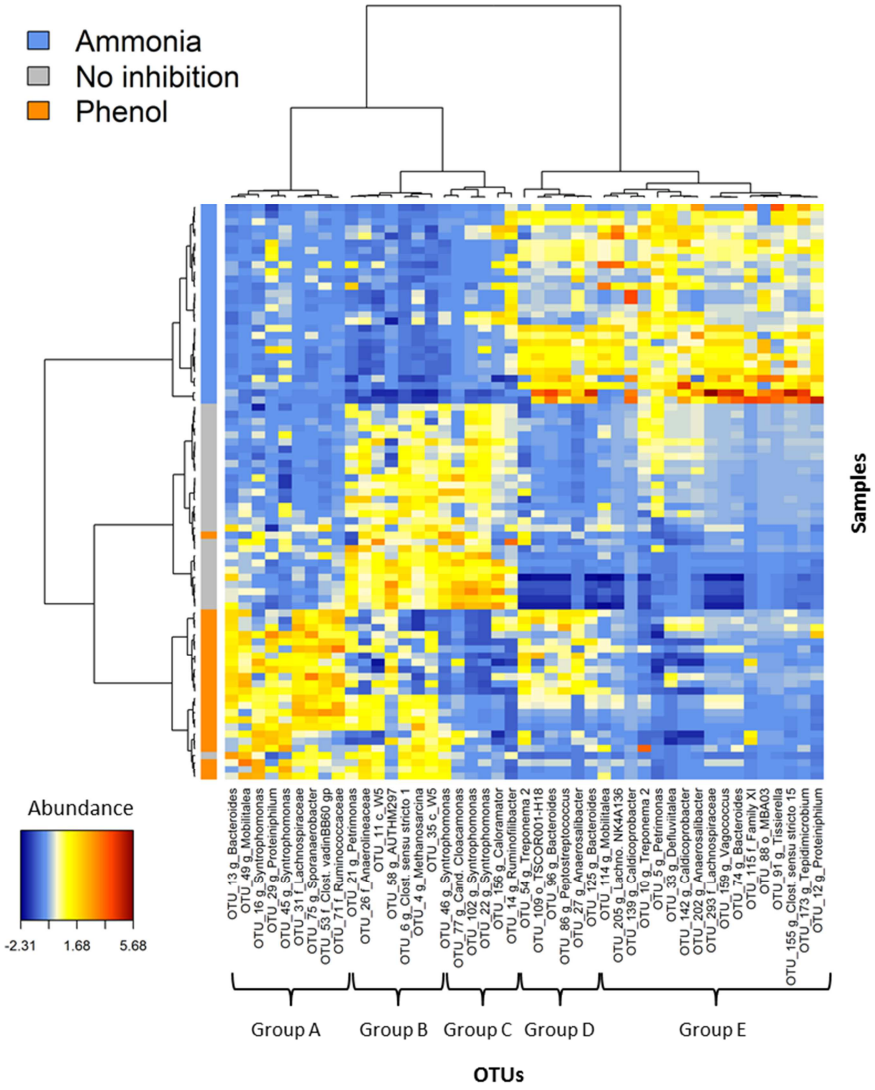


Figure 43: Heatmap of the most discriminant OTUs. Heatmap was built after selection of the most discriminant OTUs with Multivariate Integrative Sparse Partial Least Squares Discriminant Analysis of all OTUs generated by 16S rRNA gene sequencing for the different samples of studies 1 and 2, inhibited by phenol, ammonia or not inhibited. Name of the OTUs is indicated at the bottom. The colour scale on the left represents the type of inhibitor. The colour key of the heatmap shows the abundance of the OTUs after CLR transformation (from blue = low abundance to brown red = high abundance).

Groups A to E were composed of microorganisms respectively:

a) Highly abundant in digesters inhibited by phenol.
b) Lowly abundant in digesters inhibited by ammonia
c) Lowly abundant in inhibited digesters in general.
d) Highly abundant in inhibited digesters in general.
e) Highly abundant in digesters inhibited by ammonia.

These microorganisms constituted a first series of general biomarkers of ammonia and phenol inhibition, identified despite the strong initial study effect observed in Figure 42A. Their

specific identity was checked and made sense with fragmented information of the literature. They can be considered as valuable early warning signals announcing the dysfunction of the process in anaerobic digesters. These first results are encouraging because they suggest potential applications for a wide diversity of AD inhibitors and or other questions related to biotechnologies in general.

2.2.4. Building of a predictive model for ammonia inhibition

Our second objective was to evaluate the possibility to build models predicting the inhibitory status of digesters based on 16S metabarcoding data. We selected from the literature two studies analysing digesters inhibited by ammonia (studies 3 and 4). Our criteria for selecting the studies were: recent studies, raw 16S sequencing data easily accessible, digesters inhibited by phenol or ammonia, inhibition status of the digesters clearly defined, possibility to assign with no doubt the raw sequences to the samples described in the papers. *We were not able to find studies meeting these criteria for phenol.*

Our aim was to build a predictive model with studies 1 and 2 and test it with studies 3 and 4. Based on the original papers, samples from study 3 were categorized into four groups depending on the sampling time and on the inhibitory pressure: “No inhibition”, “Ammonia moderate concentration”, “Ammonia inhibition, early days”, “Ammonia inhibition, final days”. Similarly, samples from study 4 were categorized into four groups: “No inhibition”, “Ammonia inhibition start”, “Ammonia inhibition”, “Ammonia inhibition decrease”.

As samples of the different studies had been analysed with distinct sequencing techniques, different bioinformatics treatments had to be applied to the raw sequences. It was therefore not possible to compare directly OTU tables. Data were first aggregated at the genus level before merging the different datasets. Since studies 3 and 4 were focused on inhibition by ammonia only, we trained a new MINT model with studies 1 and 2 where we removed the phenol condition. Seventeen genera were selected by MINT to discriminate samples inhibited by ammonia from samples not inhibited in studies 1 and 2. The biomarkers highlighted by this second model were consistent with those evidenced by the model integrating phenol inhibition (not illustrated here). Data from studies 3 and 4 were predicted by the model (namely, based on their abundance in the 17 selected genus biomarkers, their inhibitory status was predicted). In order to visualize external samples distribution in the model, Figure 44 presents the test samples from studies 3 and 4 projected on the two first components of the trained model, as well as prediction areas that were calculated (Rohart et al., 2017a).

As expected, inhibited samples were separated against the non-inhibited samples on the first component. Two samples (11 and 39) were misclassified as ‘inhibited’. Samples of reactors that just started inhibition (“early days” in Lü et al. and ‘start of the inhibition’ in Peng et al.) were mostly classified at an intermediary position and classified as either inhibited or non-inhibited. We hypothesize that microbial community had started to change but was not yet totally characteristic of inhibited reactors. Interestingly, sample 45 (Peng et al., day 223, just after inhibitory pressure was lowered) was classified as inhibited while sample 46 (Peng et al. day 232, several days after inhibitory pressure was lowered) was classified as non-inhibited. This result illustrated the progressive resilience of the microbial community after the inhibition. Sample 14 (moderate ammonia, final days, no addition of activated charcoal) was classified as inhibited while samples 13, 15, 16 (moderate ammonia, but early days or addition of activated charcoal) were predicted as non-inhibited, in agreement with the conclusions of the authors.

Taking into account all the samples from digesters clearly non-inhibited (15) or clearly inhibited by ammonia (13) we estimated that the model predicted the inhibitory status of external samples with an accuracy of 93% as only two samples were incorrectly classified.

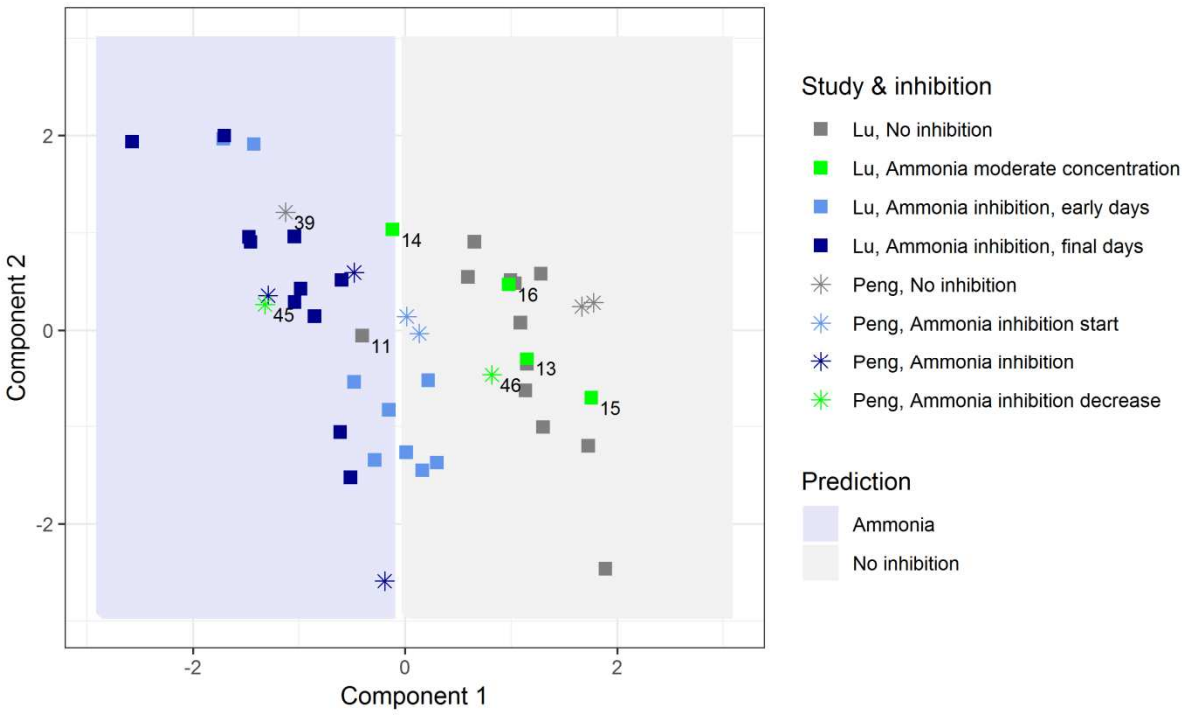


Figure 44: Projection of samples from studies 3 and 4 in the factorial plan determined after Multivariate Integrative Sparse Partial Least Squares Discriminant Analysis of samples from studies 1 and 2.
Each sample from studies 3 and 4 is represented by a marker. Type of marker indicates the study. Colour of the marker indicates the inhibition status in the reactor where the sample was taken. A prediction area, based on studies 1 and 2 was calculated and is plotted on the graph. The different figures indicate remarkable samples.

Our results evidenced that multivariate integrative methods could be efficiently implemented to predict the inhibitory status of samples collected from two independent studies focusing on ammonia inhibition, and despite differences in sequencing primers and targeted regions. It enables data sharing across research communities and re-use of existing data deposited in public databases while identifying a reproducible biomarker signature. It can be useful notably to pave the way for digester microbial management, as in the meantime miniaturized sequencers are being developed and could be in the future implemented in biogas plants at a reasonable cost. Combination of both approaches would enable to establish a functional diagnosis of the digesters provided that other models are developed to target other inhibitors of parameters characterising the functioning of the digesters. Other types of biomarkers than 16S could also be used with the same approach.

This work is not published yet but is currently under review in the ISME Journal (Poirier et al., submitted).

3. Time-course modelling to take into account the temporal variability

Microbial communities are highly dynamic biological systems. Snapshot analyses only focus on one time point and are not always sufficient to characterize microbial communities and reflect the ongoing processes. The decreasing cost of DNA sequencing and samples analysis in general has enabled to address that problem. Longitudinal and time-course studies to record

the temporal variation of microbial communities and associated processes are now possible (Knight et al., 2012; Faust et al., 2015). These studies can inform us about the stability and dynamics of microbial communities in response to perturbations or different conditions. They can capture the dynamics of microbial interactions (Bucci et al., 2016; Ridenhour et al., 2017) or associate changes of microbial features, such as taxonomies or genes, to a 'phenotypic' group (Metwally et al., 2018) such as degradation performances in the case of AD.

The overall aim of such studies is to investigate relationships between longitudinal measures in a holistic manner to further decipher the link between molecular mechanisms and microbial community structure. However, comparing at the same time multiple conditions and multiple time points characterized with one or several types of high-dimensional data is challenging. In particular, analytical frameworks enabling an integrated analysis between microbial communities and other types of biological, chemical or performance data are still at their infancy.

My collaboration with Dr Lê Cao at the University of Melbourne enabled me to pave the way for the analysis and integration of longitudinal data measured within anaerobic digesters. We used a combination of different statistical methods to analyse complex longitudinal data sets. Our approach includes modelling of profiles of the different datasets across time with spline smoothing (Déjean et al., 2007) and use of classification or multivariate ordination methods to identify correlated sets of variables across the data types, and across time. It enables us to visualize patterns of time course evolution, compare them across different conditions, infer relationships between different microorganisms or different types of data. The fitted splines enable us to predict or interpolate time points that might be missing within the time interval (e.g. inconsistent time points between different types of data or covariates).

I have applied these time-course modelling approaches on different types of data, to answer different types of questions:

- To evaluate the effect of an abrupt temperature increase on the abundance and activity of AD microbiome. Experimental set-up is presented in (Madigou et al., 2019) but analysis is based on additional data not presented in this paper.
- To compare the influence of different inhibition speeds on the activity of AD microbiome and identify biomarkers of inhibition by ammonia (PhD of Laëtitia Cardona, writing in progress).
- To draw links between microbial dynamics measured with 16S sequencing, patterns of waste degradation measured with a GCMS metabolomic approach and performances of the process (Bodein et al., submitted).

3.1. Effect of an abrupt temperature increase on the dynamics of AD microbiome

In this study, the influence of a rapid increase of temperature (35°C to 55°C) on the dynamics of AD microbiome was assessed in a semi-continuous digester. After stable functioning in mesophilic conditions for several hydraulic retention times, the digester was submitted to sudden temperature increase. Our objective was to better document the consequences of temperature modifications on the microbiome of AD to have clues to limit their detrimental effect on the process.

Samples were taken regularly in the digesters. Both RNA and DNA of the samples was extracted, and used for 16S metabarcoding to target activity and abundance of the microorganisms. Treatment of the data obtain is summarized on Figure 45.

Focus was first given to RNA data. 545 OTUs were detected. This dataset was preprocessed. Low abundant OTUs were filtered, data were normalised and transformed with centered log ratio (CLR) (Tsilimigras and Fodor, 2016). A second filter, based on the differential expression across time was applied, to keep only the OTUs influenced by the temperature modification (Straube et al., 2015). After these treatments only 30 OTUs were kept. However, in all the samples they accounted for approximately 90% of the total number of sequences. The time-course profile of each OTU was modelled with Linear mixed model splines (R lmms package, lmmSpline() function) (Déjean et al., 2007).

In order to group together the microorganisms sharing a similar temporal evolution of their activity just after temperature increase, a clustering was carried out on the first derivative of the predicted fitted curves. We focussed the analysis on a period of 20 days (going from -2 to 18, day 0 being the day where temperature was increased), represented in white on Figure 45 (abundance in samples taken before and after (blue zones of the graphs) was not used for the clustering). Hierarchical clustering, with Euclidean distance and ward aggregation was applied for the clustering. We decided to separate the profiles in 4 clusters. Time course profiles of the OTUs of the different clusters are presented on Figure 45.

- Cluster A grouped 5 OTUs. Their activity increased importantly just after temperature shock. They were the key players that helped to sustain biogas production just after the perturbation.
- Cluster B grouped the 6 most sensitive OTUs. Their activity decreased importantly just after temperature shock. High temperature had fatal consequences on their activity.
- Cluster C grouped 14 OTUs. Their activity decreased smoothly after temperature increased. They were probably not the most efficient in thermophilic conditions.
- Cluster D grouped 5 OTUs. Their activity increased only after several days under thermophilic conditions. It could be due to their slow growth rate or because they benefited from the progressive disappearance of other microorganisms.

OTUs of the different clusters were identified. For example, OTUs of cluster A seemed to be mainly ‘mesophilic microorganisms’ that adapted (at least temporarily) to the new conditions (for example *Methanobacterium beijingense*, OTU from *Mesotoga* genus, various OTUs from *Bacteroidales* order). These thermotolerant OTUs probably enabled to maintain bacterial activity just after temperature change. On the contrary some of the OTUs of cluster D were ‘thermophilic microorganisms’ that developed progressively to take advantage of the ecological ‘niches’ freed (for example *Methanoculleus thermophilus.*…).

The employed method was very powerful to collapse the dimension of the data and group together the OTUs sharing similar patterns across time and in response to the temperature increase.

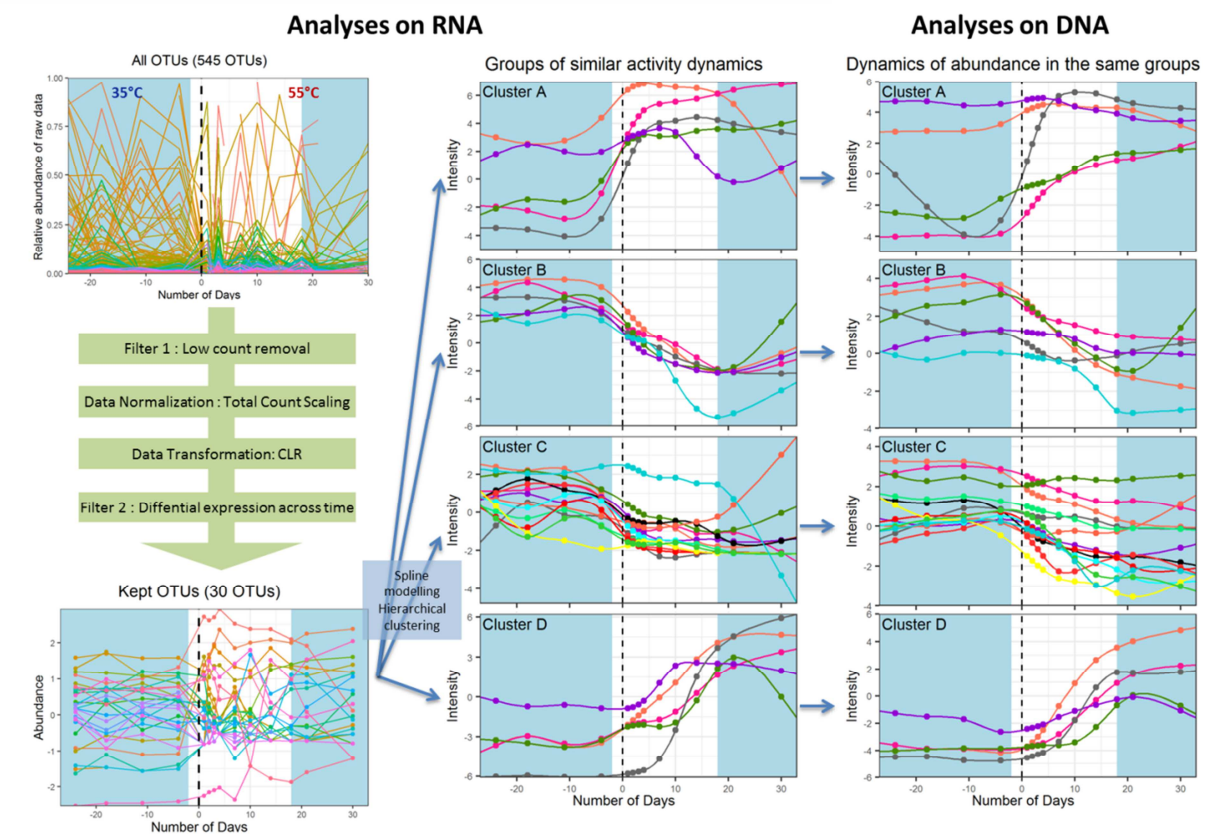


Figure 45: Time-course data analytical framework used to analyse AD microbiome data after an abrupt temperature increase. Dashed line represents the increase of temperature at day 0. Raw RNA data (OTU table) was pre-processed and filtered. Kept OTUs were modelled with spline smoothing and clustered into 4 groups with similar time course evolution between days -2 and 18 (represented in white on the figure). The time course evolutions of the OTUs seen through DNA analysis within each cluster were also plotted.

For each cluster, time course profiles of the same OTUs modelled from 16S RNA gene (DNA) metabarcoding data were also plotted (Figure 45). It enabled to see if abundance and activity of the microorganisms were strongly related or not. For example, for cluster A, we observed that the increase of the activity of 2 of the OTUs was not linked to an increase of their relative abundance in the digester (purple and orange curves), while the increase of activity of the three others OTUs was associated with a growth of the microorganisms (increase of the relative abundance).

For respectively clusters B and C and cluster D, we observed that the relative activity of the microorganisms respectively decreased or increased more rapidly than their relative abundance, confirming that 16S RNA was more accurate to reflect the AD process evolution after a shock than 16S RNA gene (De Vrieze et al., 2018). We used an approach based on fast Fourier transform (Straube et al., 2017) to estimate the delays between the evolution of the activity and the evolution of microbial abundances, based on the modelled curves. The concept is illustrated in Figure 46 (first column) for a few OTUs. Briefly, by combining the fast Fourier transform angular difference between reference (RNA in our case) and query trajectories (DNA in our case) with lagged Pearson correlation, the method characterises the magnitude and direction of delay (number of days early or late). Delays were calculated for all OTUs and are summarized in the second column of the figure with box-plots. They were of a few days.

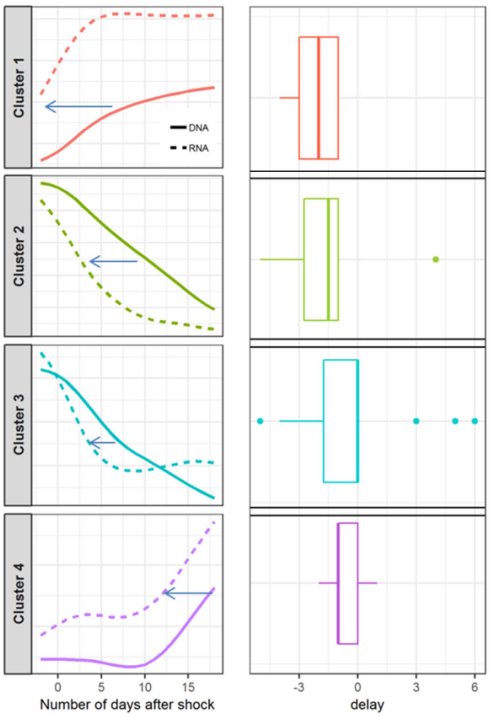


Figure 46: Delays between RNA and DNA time course profiles can be observed and calculated with an approach based on Fourier transform. They are summarized with boxplots for the different clusters.

More generally, the approach described in this section can be useful to highlight the connections and interdependency between the applied conditions, and different 'levels' of biological information such as 'DNA', 'RNA', 'proteins', 'metabolites' and performance data (macroscopic output). Calculating delays between the different types of information can be used to realign the trajectories and identify those which show a high degree of correlation despite a shift in time.

3.2. Influence of the addition of ammonia at different speeds on the dynamics of the microbiome

To evaluate the possibility to acclimate AD microbiome to the presence of ammonia, and identify potential bioindicators of inhibition, ammonia level was increased at different speeds in 6 semi-continuous digesters (Figure 47).

More precisely, after a stabilisation phase of more than 3 HRT without ammonia, NH_4Cl was added in order to reach a final NH_3 concentration of 183 mg/L, estimated to be partially inhibiting thanks to previous batch experiments (data not shown). This concentration was reached at different dates in the different reactors. P0 reactor was not perturbed (control). In P1 the target ammonia concentration was reached almost directly (2 days), in P2 it was reached after 0.5 HRT (14 days), in P3 after 1 HRT (25 days), in P4 after 2 HRT (50 days) and in P5 after 3 HRT (75 days). For all the bioreactors, once reached, the inhibitory stress was maintained during 3 HRT.

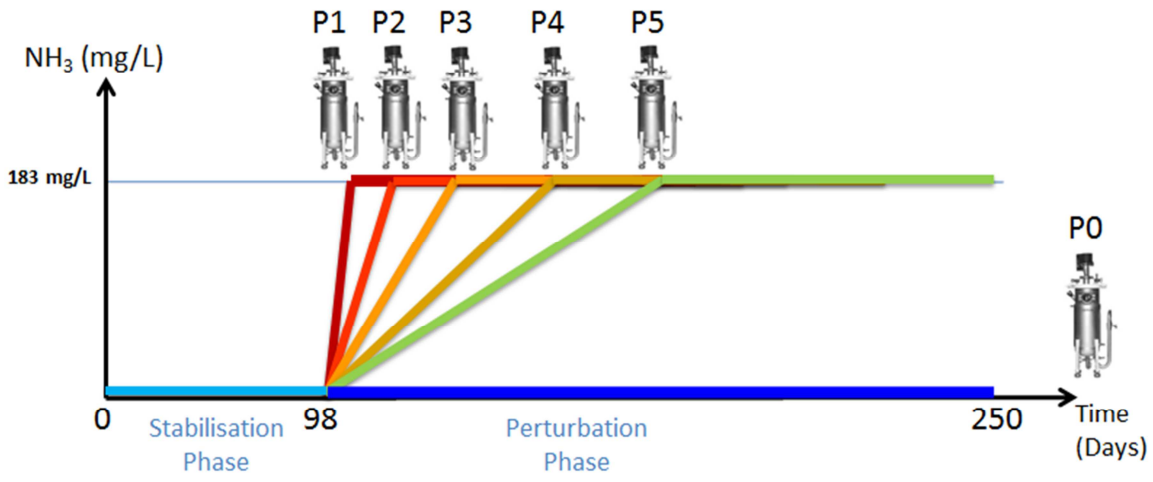


Figure 47: Experimental design used to evaluate the influence of inhibition speed in semi-continuous bioreactors.

Process parameters were measured across time to compare the effect of the different inhibition speeds on the performances of the process (Figure 48). It showed that in all case the level of VFA (acetate and propionate) increased after the perturbation. A decrease of biogas production was also observed, but its intensity decreased from P1 to P5. Proportion of methane in the biogas diminished only in P1 and P2. It suggested that an acclimation of the microbiome was possible provided that the inhibition speed was not too important.

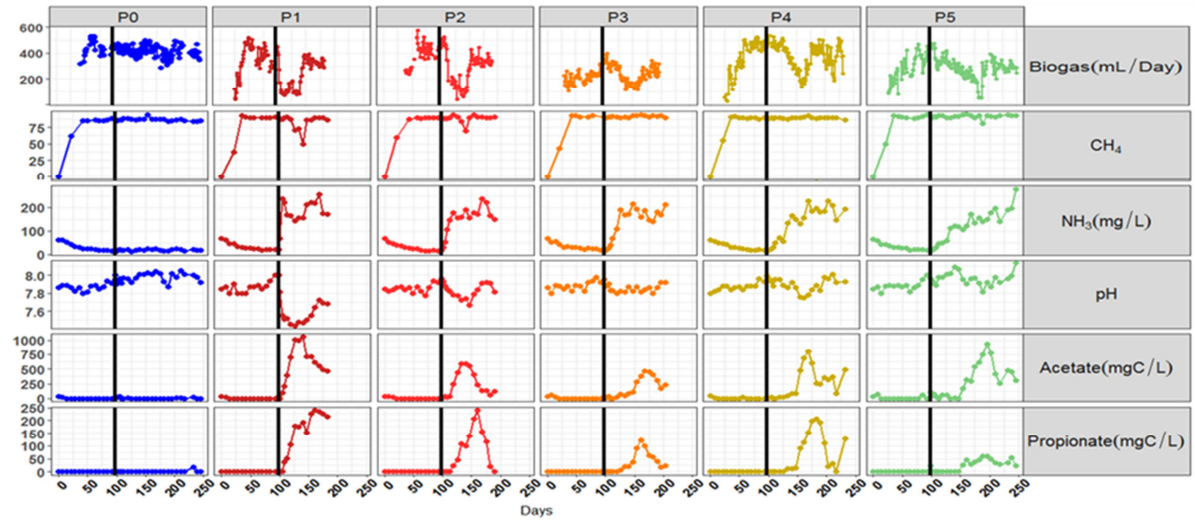


Figure 48: Main performances indicators measured in the different reactors across time. Black line shows when the inhibition was started.

Samples were taken regularly for RNA extraction and subsequent 16S metabarcoding of the corresponding cDNA was performed. Time course profiles of the different OTUs in the different reactors were modelled as described previously. Focus was given to the period between days 70 and 154 corresponding to the end of the stabilisation phase and the inhibition phase.

Data of reactor P1 was analysed first. Data was filtered to remove microorganisms whose activity did not evolved over time. In order to group together the variables sharing a similar temporal evolution a clustering was carried out on the first derivatives of the predicted fitted curves (Déjean et al., 2007). The kml method (k-means for longitudinal data, (Genolini and Falissard, 2011)) was applied for the clustering. 8 clusters of profiles were obtained (see line P1 in Figure 49) corresponding to OTUs whose activity decreased or increased more or less rapidly after the addition of ammonia.

The evolutions of the time-course profiles in P1 were compared to the evolution of the time course profiles in the other reactors. For that purpose, the time course profiles of these clusters of OTUs in the other reactors was also plotted. In P0 (not inhibited) profiles remained rather constant, showing that their evolution in P1 was likely linked to the addition of ammonia. From P2 to P5 contrasted results were observed. For example, several OTUs from cluster 2 that were strongly inhibited in P1 were not or less inhibited when ammonia addition was more progressive. Most of the OTUs from cluster 5 that appeared in P1 just after the inhibition also appeared in P2, but less and less appeared in P3, P4 and P5. It confirmed that an acclimation of the digesters to the ammonia seemed to be possible, limiting the consequences on microbial community. On the contrary activity of other microorganisms evolved in all the cases meaning that several OTUs were very sensitive to ammonia, independently of the inhibition speed. These results are currently analysed more in depth to identify which microorganism acclimatized to the inhibitor and which did not (PhD Laëtitia Cardona).

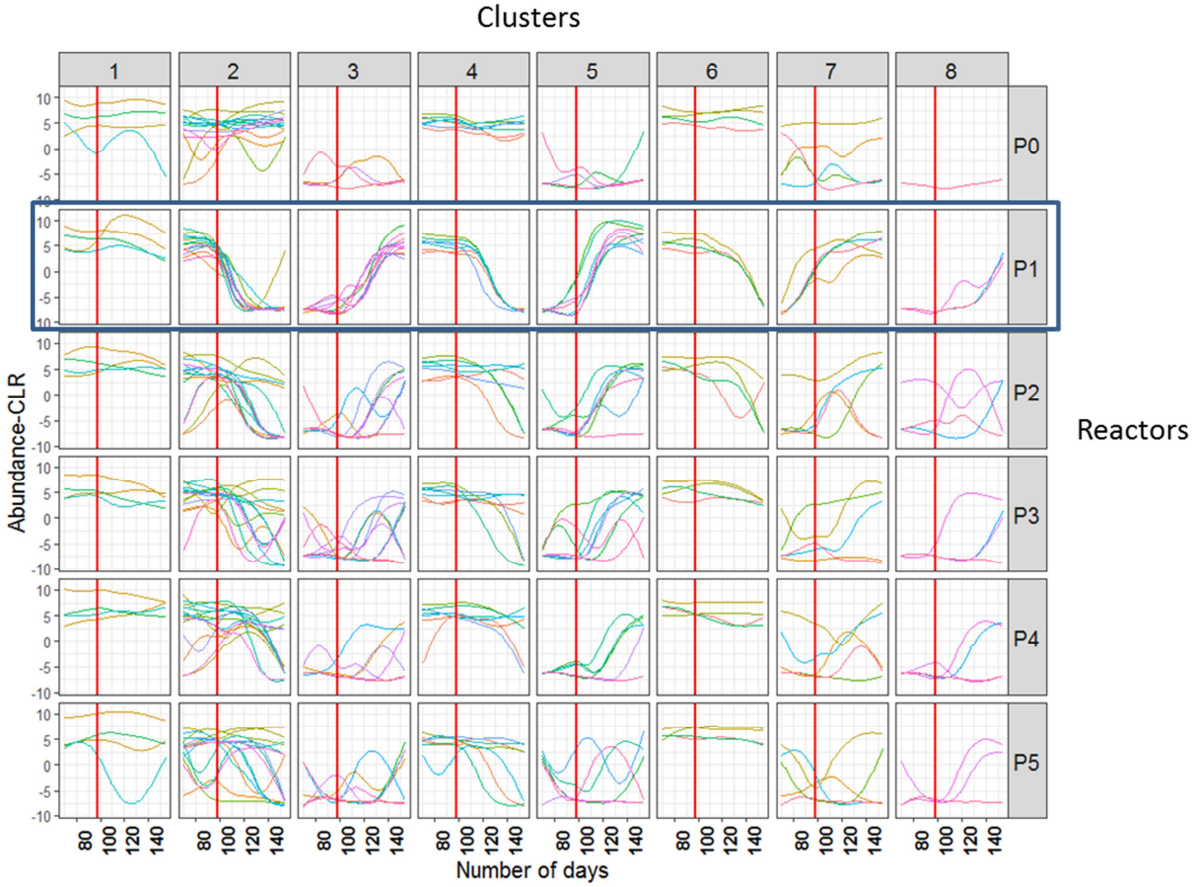


Figure 49: Time course profiles of the different OTUs in the different bioreactors (lines). OTUs were grouped in clusters (column) based on their similar evolution in P1 digester. Red lines correspond to the day when inhibition was started.

For example, bioindicators of the presence of ammonia were sought by identifying the microorganisms whose activity evolved in all the reactors where ammonia was added, independently of the inhibition speed. For that purpose, for each OTU, time delays were estimated between the time profiles in P1 and in the others inhibited reactors. The microorganisms with a similar time profile evolution between the different inhibited reactors, showing a delay in accordance with the ammonia addition speed, were considered as potential bioindicators of ammonia inhibition. For example, Figure 50 illustrates the time course profile of two OTUs in the 6 digesters. Their activity remained stable in P0 but gradually decreased across time, in accordance with the level of ammonia in the digesters P2 to P5 and independently of the inhibition speed.

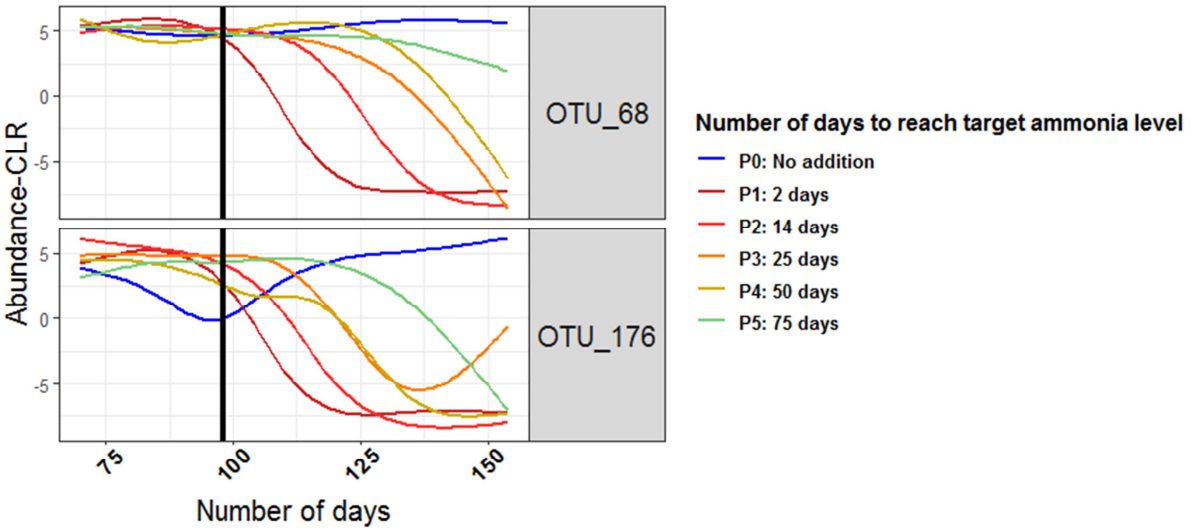


Figure 50: Time course profiles of two OTUs in the different bioreactors. OTU 68 is from the family *Syntrophaceae* and OTU 176 from family *Leptospiraceae.*

Finally, with the same samples we are also currently assessing the effect of the ammonia perturbation on the pathways of degradation. Untargeted metabolomics analyses were performed with FTICR (Fourier-transform ion cyclotron resonance). Several thousands of metabolites were detected. A similar treatment was applied and is illustrated in Figure 51. In this case only four groups of profiles were evidenced, that reacted more or less quickly to ammonia inhibition in P1. As observed with microorganisms, contrasted results were observed in the other reactors. This data was obtained very recently (collaboration with the Ecole Polytechnique) but will undoubtedly provide further insights in the inhibition of AD by ammonia.

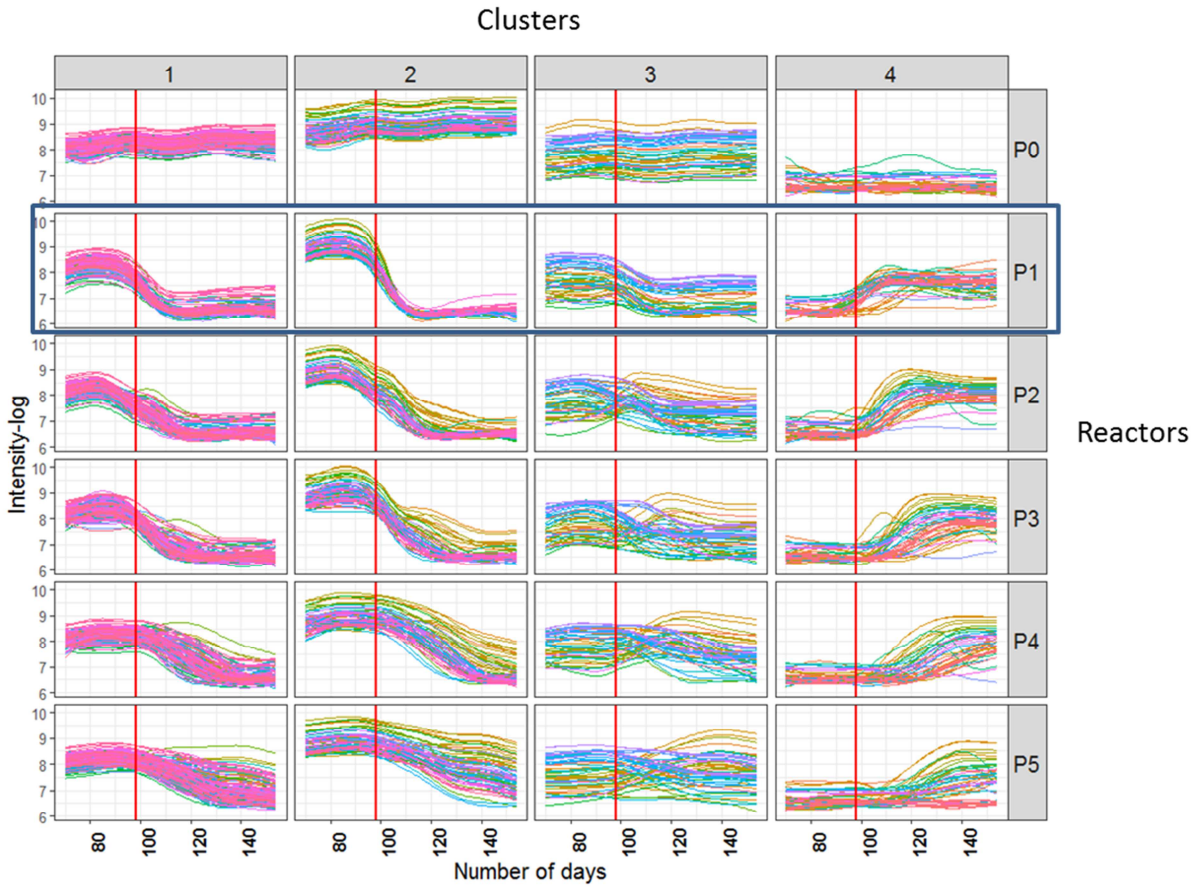


Figure 51: Time course profiles of the different metabolites in the different bioreactors (lines). Metabolites were grouped in clusters (column) based on their similar evolution in P1 digester. Red line corresponds to the day when inhibition was started.

3.3. Integration of different types of data in a time-course context

The approaches presented in the two previous paragraphs did not seek for correlation between data of different types. However, as illustrated in particular in section 1 of this chapter, the same samples are sometimes analysed with different techniques. During a collaborative work with Kim-Anh Lê Cao, we set-up a specific analytical framework for the integration of microbiome longitudinal studies with other data types (Bodein et al., submitted) (preprint was deposited on biorxiv (10.1101/585802v2.full)).

The general workflow of the method is presented Figure 52. In brief, the different types of data measured on the same samples are first pre-processed (filtering, normalisation etc.). For each type of data, the 'variables' (for example OTUs, metabolites, performance measures…) are modelled with spline smoothing. Modelled data are filtered (noisy variables are often modelled as straight lines that do not represent a biological reality and are therefore discarded). Resulting time course profiles are integrated with different methods according to the number of types of data to integrate. Sparse PCA is used to treat single datasets. Sparse Projection to Latent Structures (sPLS) method (Lê Cao et al., 2008) was described in section 1. Multiblock sPLS is a generalisation of sPLS to more than 2 blocks of data (Singh et al., 2016). In all the cases, the objective is to identify correlated (or co-expressed) variables measured on one or several heterogeneous data sets.

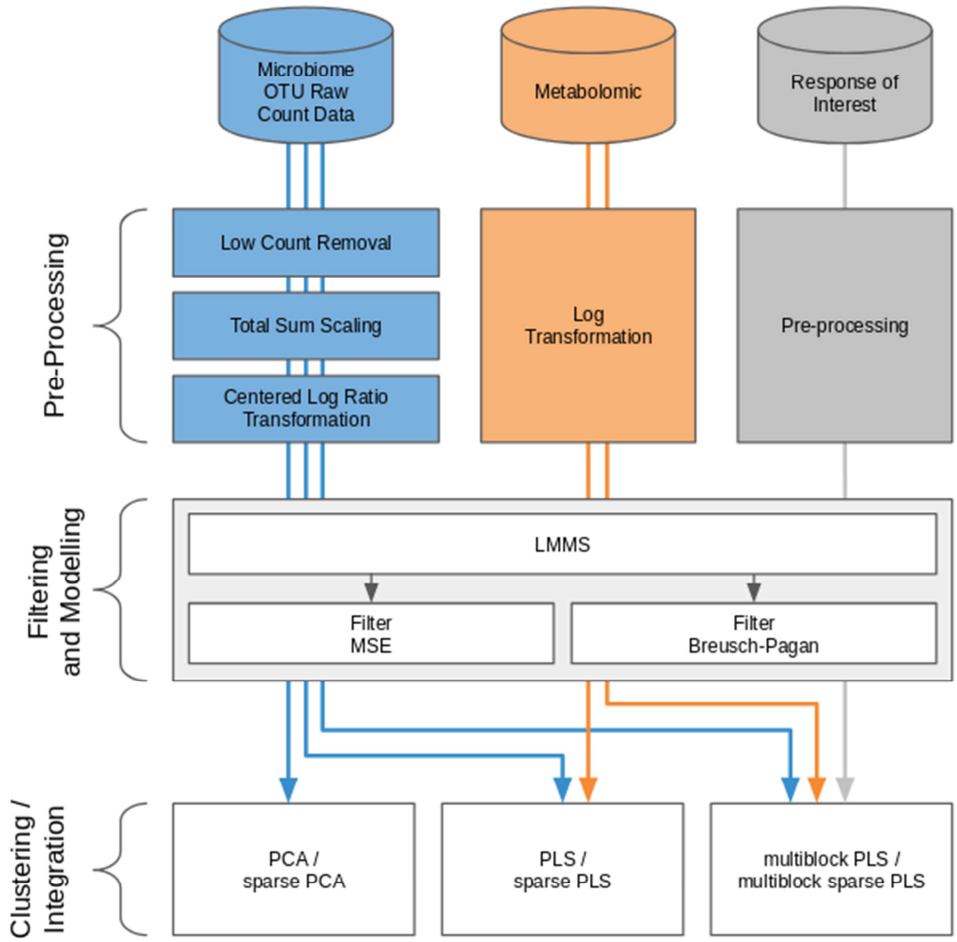


Figure 52: Workflow diagram for longitudinal integration of microbiome studies. We consider studies for the analysis of the microbiome through OTU (16S amplicon) or gene (whole genome shotgun) counts. This information can be complemented by additional information at the microbiome level, such as metabolic pathways measured with metabolomics, or information measured at a macroscopic level resulting from the aggregated actions of the microbiome.

More precisely, after computing the sPCA, the sPLS or the multiblock sPLS, a set of correlated variables is selected on each component (axis) of the analysis. This set can be divided into two groups of variables: variables that have a positive loading on the component and variables that have a negative loading. They are negatively correlated. Each component thus enables to evidence two groups of correlated profiles, as illustrated on Figure 53.

This method was assessed on different types of data, including a digester study which is presented below.

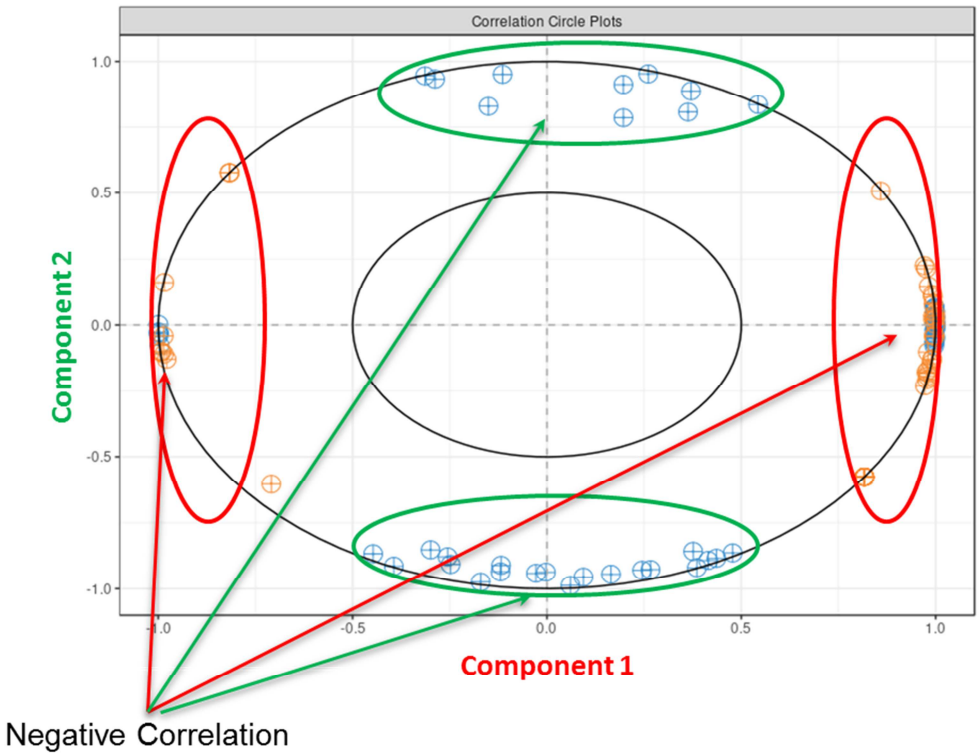


Figure 53: Identification of the correlated variables with sPCA, sPLS or multiblock sPLS on the correlation circle plot. After computing a sPCA, sPLS or multiblock sPLS a correlation circle plot is obtained. It shows the loadings of the different variables (here variables are presented with orange and blue markers) on the different components of the analysis (here components 1 and 2). Variables with a negative loading on component 1 can be grouped together. They are highly correlated. Variables with a positive loading on component 1 are also correlated with each other, and negatively correlated to the variables with a negative loading on component 1. They form a second group of correlated variables. Process can be iterated on component 2 and subsequent components if necessary to evidence groups of correlated variables.

To evaluate this method, we used the data from an experiment where three anaerobic batch digesters fed with biowaste had been monitored across time. Degradation performance was monitored through 4 parameters: methane and carbon dioxide production (16 time points) and accumulation of acetic and propionic acid in the bioreactors (5 time points). Microbial dynamics were profiled with 16S RNA gene metabarcoding and included 4 time points and 90 OTUs. A metabolomic assay was conducted on the same biological samples on 4 time points with gas chromatography coupled to mass spectrometry (GC-MS) after solid phase extraction to monitor substrates degradation. The different types of data were pre-processed, modelled with smoothing splines and filtered as described in Figure 52.

The three datasets were first analysed one by one and two by two (not illustrated here). Finally, the three types of data (16S, metabolites and performance data) were integrated together with multiblock sPLS as described above. Correlation between the different variables was correctly described by two components. On each component, variables with respectively negative and positive loadings were grouped together to form 4 cluster. The time course evolution of the variables of the 4 clusters is illustrated on Figure 54.

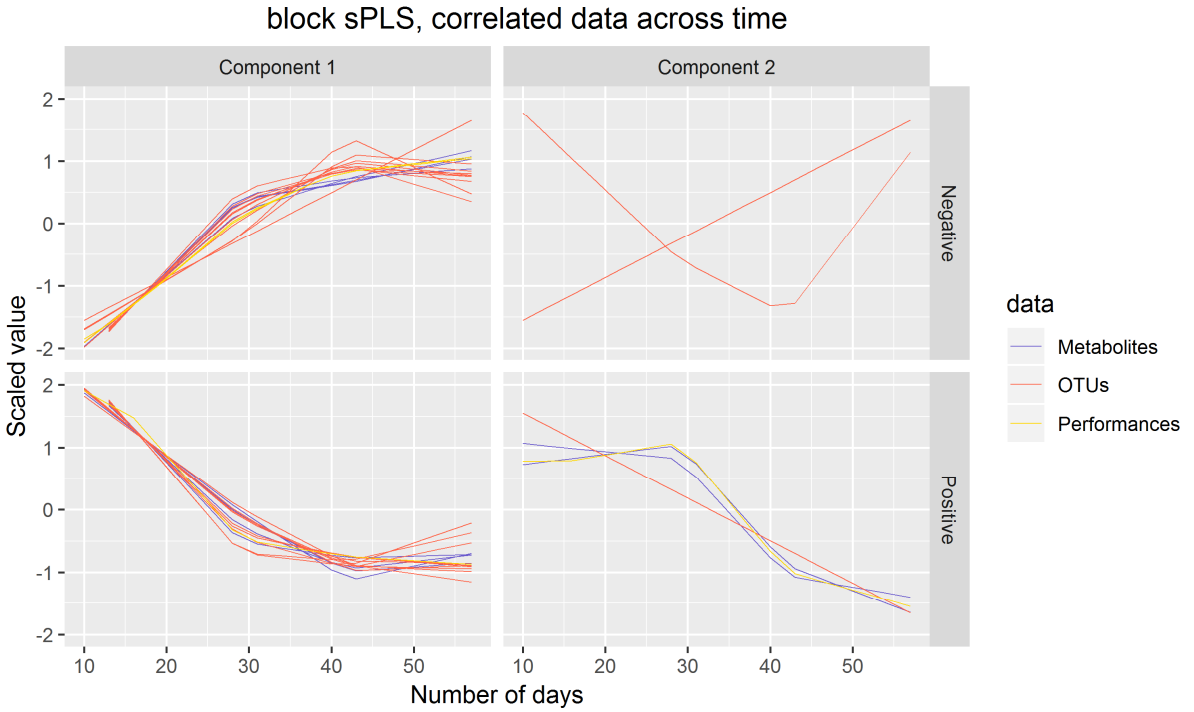


Figure 54: Integration of OTUs, metabolites and performance measures with block sPLS. Each line represents the scaled abundance of selected OTUs, metabolites and performance measures across time. OTUs, metabolites and performance measures were clustered according to their contribution on each component for block sPLS. The clusters were further separated into profiles with a positive or negative correlation.

The first cluster (denoted `component 1 negative') included 10 OTUs, 4 metabolites and 2 performance variables. The variables showed increasing level until a plateau was reached at approximately 40 days. The OTUs were microorganisms often recovered during anaerobic digestion of biowaste, such as methanogenic archaea of *Methanosarcina* genus or bacteria of *Clostridiales*, *Acholeplasmatales*, and *Anaerolineales* orders. Their abundance increased while biowaste was degraded, until there was no more biowaste available in the bioreactor. Their abundance was correlated to the intensity of various metabolites produced during the AD process, such as benzoic acid that is formed during the degradation of phenolic compounds (Hoyos-Hernandez et al., 2014), or phytanic acid, known to be produced during the fermentation of plant materials in the ruminant gut (Watkins et al., 2010), as well as indole-2-carboxylic acid. Two performance variables (methane and carbon dioxide production) were assigned to cluster 1. This result is biologically relevant, as biogas is the final output of the AD reaction and is known to be associated with microbial activity and growth. Moreover, it is produced by archaea, such as *Methanosarcina*, also selected in this cluster.

Cluster 2 (component 1 positive) included 10 OTUs, 4 metabolites and 1 performance variable. These profiles were negatively correlated to Cluster 1, and their abundance decreased with time. OTUs mainly belonged to the *Bacteroidales* order. They were present in the initial inoculum but did not survive in this experiment, as the operating conditions or the substrate were not optimal for their growth, as observed in other studies (Madigou et al., 2019). Metabolites identified in Cluster 2 were present in the biowaste and were degraded during the experiment. They included fatty acids (decanoic and tetradecanoic acids) that can be found in oil, or 3-(3-Hydroxyphenyl)propionic acid, arising from digestion of aromatic amino-acids or breakdown product of lignin or other plant-derived phenylpropanoids (Torres

et al., 2003). As their profile was negatively correlated to those from cluster 1, it is likely that these metabolites were consumed by OTUs assigned to cluster 1. Acetate was also associated to this cluster.

Cluster 3 (component 2 negative) included 11 OTUs and 3 metabolites with slow abundance increase. OTUs of this group were very varied with 8 orders represented. They may have slower growth rates than OTUs of cluster 1 or were involved in the last steps of the degradation. Metabolites included N-Acetylanthranilic acid and Dehydroabietic acid that were likely produced by microorganisms and accumulated during the anaerobic digestion process.

Finally, Cluster 4 (component 2 positive) included 1 OTU and 5 metabolites. Profiles decreased slowly with time. One OTU of order *Clostridiales* appears to have been out-competed by other OTUs during the first days of the degradation. Among the metabolites of this cluster, Hydrocinnamic and 3,4-Dihydroxyhydrocinnamic acids are commonly found in plant biomass and its residues (Boerjan et al., 2003). Their molecular structure may have contributed to their slower degradation compared to other molecules. Propionate was assigned to this cluster, which made sense as its degradation only starts when all acetate is degraded (Chapleur et al., 2014).

Multiblock PLS enabled to identify microorganisms and metabolites with correlated time course profiles. The computational framework we have proposed is one of the first to integrate longitudinal microbiome data with other omics data or other variables generated on the same biological samples or material. The identification of highly-correlated key omics features can help generate novel hypotheses to better understand the dynamics of biological and biosystem interactions. Thus, our data-driven approach will open new avenues for the exploration and analyses of multi-omics studies.

4. Conclusion

Advances in technology and reduced sequencing costs have resulted in the emergence of new and more complex experimental designs that combine multiple omic datasets and several sampling times from the same biological material. Thus, the challenge is to integrate longitudinal, multi-omic data to capture the complex interactions between these omic layers and obtain a holistic view of biological systems. These approaches enable to propose data-driven analytical frameworks.

The data I have been working on so far come from experiments that were not designed on purpose for this type of analysis. Working with statisticians opened new perspectives and will help me to improve the experimental design to answer future research questions. I now also plan to include other types of omics data in these analyses. These objectives are presented with more details in the next and last part of this manuscript.

Future research project

As said in the preamble and described in this manuscript, during the first decade of my research career, I had the opportunity to use multiple techniques to target different questions related to the topic of AD. The concomitant development of microbial ecology significantly boosted the possibilities. I'm now also generating omics data. So far it was mostly metabolomics, at Irstea and at the Ecole polytechnique through collaboration I started in 2018. In the framework of the Digestomic project, I generated in 2019 a first set of metagenomics data. Metatranscriptomics data will also be produced soon for this project. These data put back into perspective the analysis that can be done of microbial ecosystems of AD digesters.

To face the growing volume of data I was generating, I initiated in 2017 collaboration with statisticians. This collaboration was decisive. Of course, it provided me with the different tools I needed to treat my data. I learnt how and when to use the methods. But it also opened new perspectives. The close link with the statisticians, greatly facilitated by different stays of several months within their research team, allowed me to both learn how to explain my research questions so that they can be associated to statistical methods, and discover the potential of the field through the results obtained by my collaborators on other types of data. Based on these exchanges, we started to assemble specific analytical frameworks for the treatment of my data (for example for the analysis of longitudinal data).

As a future research project, I want to carry on the work engaged to improve the environmental bioprocesses. I want to combine together the new methodological strengths mentioned above and also mix them with the methods I used earlier in my research. The following pages describe these objectives. Some of them are illustrated with examples of the work planned in the STABILICS project submitted to the ANR call in 2019 ("Coupling statistics and multi-omics to gain new insights in the determinants of anaerobic microbial bioprocess stability", in collaboration with the different partners mentioned above) and that has just been accepted.

1. Intended research focus for the next years

1.1. Toward the development of anaerobic digestion

Despite it has been studied for a long time, AD is still not fully mastered and has a potential for improvement. In particular, it still lacks stability and flexibility. In the next paragraph I present different topics that could be targeted in future projects to favour AD development.

1.1.1. Optimizing the stability of the anaerobic digestion

Key microorganisms driving the AD process form an extremely complex microbial community or microbiome, able to degrade the organic matter across multiple pathways, with lots of functional redundancies. One of the major limitations of AD is the important susceptibility of this microbiome to changes in operational conditions of the digesters, including inhibitors (see introduction, section 1.4). It can lead to unstable methane formation and have important economic and environmental consequences.

Controlling AD microbial community stability, though, is not a trivial task. Its exact role in terms of "elasticity" of the digesters, i.e., its robustness or ability to rapidly adapt to changing conditions, is unclear (De Vrieze et al., 2013). Numerous studies were attempted to understand the influence of inhibitors and identify key operational parameters playing a role in the stability of AD microbiome (Carballa et al., 2011; Poirier et al., 2016b). However, they were often limited to a 'before' and 'after' comparison. No consensus on the underlying effect of perturbation was obtained. Moreover, experience feedback shows that instability is encountered even when the operating conditions of the digester seem highly regular (Fernandez et al., 1999; Zumstein et al., 2000; Briones and Raskin, 2003; Rivière et al., 2009).

Therefore, to facilitate sustainability and AD development, I think that it is essential to deepen our understanding of the determinants of bioprocess stability over a long time period. They consist in the conditions and the succession of microbial events that allow maintaining a balance after a disruption or, on the contrary, that generate a domino effect leading to total failure. *That's one of the objectives of the STABILICS project (described with more details in 1.2).* Once the causal chain leading to the disruption of AD digesters will be better understood, it will be easier to propose adequate solutions to improve AD stability.

1.1.2. Broaden the fields of application of AD

To address the increasing need for energy, independent from the use of fossil fuels, AD appears as a promising process. Increasing the number of biogas plant requires that the organic matter resource is widen and that as many substrates as possible can be used for methane generation. In the last years, both municipal and industrial production of organic waste has shown increased metrics in production and quality of the organic waste (higher collect efficiency, improved source separation). Numerous new substrates are showing interesting potential as for example fats and algae. Before they can be used routinely, several aspects need to be better understood, such as of course their impact on the process performances, but also the changes they induce on microbial dynamics in digesters (Mata-Alvarez et al., 2014). I think it will be another interesting area of research in the following years.

In the same way, processes need to be as adaptable as possible to the volume and organic content of waste. Indeed, the production of several types of waste is highly seasonal. Among others, temperature could be an operational lever to adjust the efficiency of bioreactors according to the fluxes of waste. It was shown that decrease in temperature results in a decrease of the metabolic rate of the microorganisms, and in the substrate utilization (Bowen et al., 2014), which can be useful in digesters treating low or irregular volumes of waste. However, works are still needed to better understand and control this lever. For example, temperature decrease usually causes an accumulation of intermediate products, such as VFA, which may in turn inhibit microbial members, eventually reducing the process efficiency (Regueiro et al., 2014). Temperature modification can also result in important changes in the community structure (McKeown et al., 2009; Chapleur et al., 2016a; Madigou et al., 2019).

Links between the operational parameters and the degradation pathways employed by the microbial community could also be sought. It could for example enable to define possible microbial management strategies to orientate the functioning of the ecosystem toward the production of molecules with high added value such as biofuels or synthons usable in green chemistry.

1.2. Example of the STABILICS project to address these topics

As an illustration of my approach to address these research focuses, this section describes more precisely the experimental set-up of the STABILICS project. The aim of the project is to

gain new insights in the determinants of anaerobic microbial bioprocess stability. For that purpose, the microbial dynamics in semi-continuous anaerobic digesters under constant environmental parameters or subject to different model perturbations created by the addition of NaCl will be monitored in the long run (more than one year).

1.2.1. Influence of different patterns of perturbation

Perturbation can greatly influence the evenness, dynamics and diversity of AD microbial community and associated functions. The microbial community can maintain a stable composition (resistance), temporarily change in composition (resilience) or shift to a new composition (redundancy) in response to a disturbance, depending on the type of perturbation (Theuerl et al., 2019). To explore diverse scenarios, different model disturbances will be created in STABILICS (abrupt, gradual, repeated shock). They are depicted in schematic format in Figure 55. It will enable to induce various ecological responses of the microbiome, cascades of effects at the taxonomic or functional levels and evidence mechanisms such as adaptation, resistance, resilience or functional redundancy (Theuerl et al., 2019). Dynamics of the microbiome under disturbance will be compared to the baseline level under stable conditions.

Abrupt addition of salt (1) will immediately create an important stress, giving no opportunity to the microorganisms to adapt. It will evidence clearly the microorganisms and the functions or metabolic pathways most sensitive to the stress or most eager to take over freed ecological niches. Recovery after inhibition may require multiple microbial, functional and metabolic changes (Madigou et al., 2019).

On the contrary, progressive addition of salt (2) could enable a smoother adaptation of the microbial community. Main functions and degradation pathways could be maintained through functional redundancy despite microbial changes. For example, acclimation of methanogens to high concentrations of sodium over prolonged periods of time was already described (Mottet et al., 2014).

Pulse increasing addition of salt (3) will also allow an adaptation of AD microbiome. In that case, inhibiting pressure release between each pulse event will ensure the total recovery of the microbiome and functions between each stress. It will avoid drift toward a simplified and less robust community (De Vrieze et al., 2013).

A control condition (0) will be of crucial interest to measure the dynamics of the performance and of the microbiome when all the parameters are kept constant.

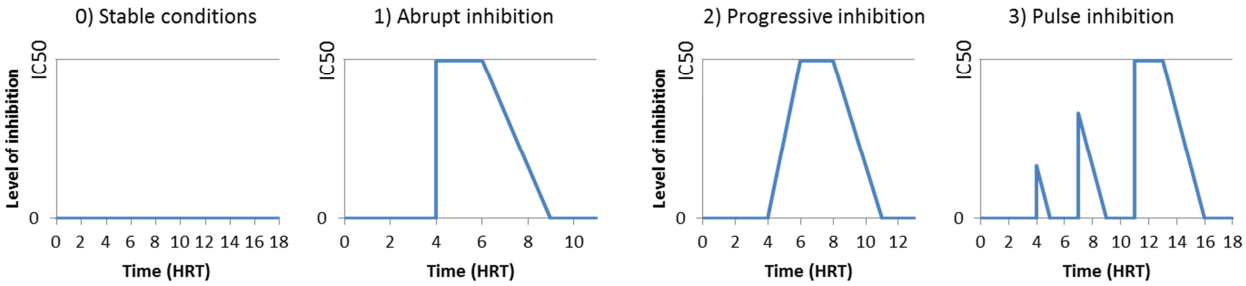


Figure 55: Perturbations that will be applied in the digesters.

Specific focus will be given to the transitional states and to the regeneration phase after perturbation as we assume they are highly dynamic.

For each condition tested, triplicate reactors will be set-up. It is essential to assess the reproducibility of the observations, too seldom done in the literature (Prosser, 2010).

1.2.1. Salt as a model inhibitor

Stress will be created by the addition of salt (NaCl), regularly mentioned as an AD inhibitor. NaCl was selected because it is found in various types of waste treated with AD (Lefebvre et al., 2007). It is found in various residues from the food processing industry (Chen et al., 2008a), in aquaculture sludge, molasses or vinasses wastewaters, kitchen waste (De Vrieze et al., 2017), microalgae (Mottet et al., 2014). While moderate concentrations stimulate microbial growth, excessive amounts slow it down, and cause severe inhibition or toxicity (Mottet et al., 2014; De Vrieze et al., 2017).

Salt is also very easy to manipulate and control, as it is soluble, non-degradable, non-toxic (contrary for example to micropollutants), not impacted by other changes in the reactor (contrary for example to ammonia whose effect depends on the pH). Its concentration can be easily assessed with conductivity and measured with ionic chromatography.

2. Methodologies

The methodologies I plan to use in my future projects rely on the combination of different innovative methods to study complex microbial communities. In line with the first results presented in Chapter 3, I want to apply high-throughput methodologies more widely in combination with statistics. I would also like to combine stable isotope probing with omics methodologies in future projects.

2.1. Application of omics in combination with statistics

Emerging omics high-throughput approaches can now lead to unprecedented data to portray AD microbiome at different levels (Vanwonterghem et al., 2014a). Metagenomics, metatranscriptomics, metaproteomics and metabolomics provide the information necessary to represent portraits of a community's genes, gene expression, and metabolite production. They reveal the potential and the activated functions of the microbiome and the outcome of their expression. They can allow to unravel the intricate networks of functional processes of AD (Lü et al., 2014; Vanwonterghem et al., 2014b; Carballa et al., 2015; De Vrieze et al., 2016b), provided that appropriate analytical methods are applied to decipher these big datasets. In particular, combining omics information with data on reactor performance during different operational conditions (especially transitional states) could help elucidate the mechanisms of process instability (Li et al., 2018). Computational analytical methods have a promising potential to capitalise on this rich data. However they are still at their infancy and were not broadly used for this type of problems.

STABILICS project totally fits in this category. It was specifically designed to produce repeated longitudinal omics data to be analysed with computational statistics to generate more information on the reasons leading to a disruption of digesters.

2.1.1. Illustration of the approach with STABILICS project

2.1.1.1. Research hypotheses

To build the project and imagine the different experiments of STABILICS, I formulated the hypotheses that:

- AD is managed by an intricate microbial community, highly diverse and functionally redundant, resulting in an important dynamism across time in general and elasticity under perturbation.
- Perturbation results from a succession of events at different omics levels leading to a disruption of the digesters if microbiome equilibrium cannot be maintained despite its important elasticity.
- The different levels of omics provide different types of information that describe more or less accurately the state of the digesters regarding operational stability or dysfunction. For example, despite stable performances of the digester, one or several levels of omics can be highly variable.
- Different types of stress (abrupt, progressive, repeated…) can result in different types of effects on the microbiome, the different levels of omics, and ultimately on the performances of the digesters.
- Deciphering the determinants of AD stability requires longitudinal studies across multiple time points, at several levels of omics and under stable and stress conditions.

2.1.1.2. Extensive use of omics methodologies

To test these hypotheses, and get insights in the determinant of AD stability, I proposed to conduct sets of high-throughput multi-omics longitudinal analyses, with an unprecedented sampling depth, in anaerobic digesters described in section 1.2. Obtained data will be analysed with innovative computational methods. More precisely, the objective will be to evaluate at different omics levels the dynamics of AD microbiome in long term and replicated time course experiments, under stable conditions or subject to different types of stress. Two levels of analysis will be applied.

A first level of analysis will consist in a high-frequency monitoring (weekly) of different descriptors of microbiome activity, easily accessible at the lab. Non-targeted metabolomics will characterise the degradation pathways. Associated microbial dynamics will be monitored with metabarcoding. Both RNA and DNA methods will be used, to target both active and present microorganisms (De Vrieze et al., 2016b), i.e. actual and potential actors of the process. Links between the conditions applied, microbial dynamics, stability of the degradation pathways and digester's performance across time will be sought to get new insights in the determinants of AD stability. This frequency of analysis is high and will be a strength of the project. It will enable to follow in details the time course evolution of the different 'variables'. Our previous results showed that this sampling time step was necessary in order not to lose precious information as microbiome evolution can be very quick (Poirier et al., 2016b). Reproducibility of the observation will be assessed.

A second level of analysis will consist in an in-depth monitoring of microbiome functioning with both shotgun metagenomics and metatranscriptomics that give complementary information (Maus et al., 2016). Selected samples will be analysed, as these methods are more expensive and time consuming. Focus will be given to one disturbed condition and to the stable condition. A deeper access to digesters microbial functioning will be gained. Information on the expressed and potential functions will be integrated to the other levels of information with statistical methods. First hypotheses on microbial dynamics will be validated and new hypotheses on the association with functional dynamics will be posit, still with the view to getting insights in the determinant of AD stability across time. These approaches will provide precious information on the expression of the different functions during time and in response to the stress. Analytical strategy (selection of samples) will be based on the results of high-frequency monitoring. Four time points will be analysed in triplicates: before stress, during stress establishment, during stress maintenance and after recovery. Control analysis in stable conditions will also be performed (2 time points analysed in triplicates) to evaluate the evolution of the functions despite stability.

2.1.1.3. Integrative statistical analyses and novel computational developments

A methodological challenge will be to integrate the different types of data and the temporal information. Data will be analysed independently with multivariate methods to assess the influence of the stress on the stability of the different parameters measured. Supervised statistical approaches will be used to identify discriminant features across time, between the conditions or the replicates. It will provide us with a subset of bioindicators of process stability/forthcoming instability, helping us to generate novel hypothesis. Longitudinal analytical frameworks will compare the time course evolution of the variables within each data set and cluster them into groups of similar behaviour. This novel approach will help us identify features with similar patterns of activity, and link microorganisms with degradation pathways. It will enable to gain a better understanding of the dynamics and the interplay between the different levels of information measured.

2.1.1.4. Questions targeted

The following question may typically be addressed:

- How dynamic is the microbiome under stable operation?
- What are the consequences of microbial community changes on the degradation pathways?
- Is the stability of the digesters more dependent on stability of present or active microorganisms?
- What is the succession of events that, under stress, leads to microbiome equilibrium unbalance and digester disruption or on the contrary microbiome equilibrium preservation and maintenance of stability?
- Which level of omics data is the more appropriate to identify bioindicators of optimal performance or stress?

It will give insights on the respective role of potential and really expressed functions on microbiome equilibrium. The combination of all these data will enable us to identify precisely the limiting stages of the processes at the origin of the varying performances (for example instability of one specific pathway, one type of microorganisms, one function or one microbial functional group etc.). The project will deliver generic knowledge to understand the determinants of perturbations. It will enable to propose microbial management strategies to improve the stability of anaerobic digesters.

2.2. Use of omics and isotope labelling in combination

Experiments done during my PhD revealed that stable isotope probing was very powerful to targeted specific functional information. However, they were limited by the volume of data generated. The news omics methodologies give access to more information and can be advantageously combined to SIP.

Only a few examples of research targeting the fate of labelled substrates in AD systems with metagenomics were published (Ziels et al., 2018). However, as a limitation of metagenomic studies is that sequencing of bulk DNA focuses primarily on genomic sequences from relatively abundant microorganisms, coupling both approaches can give access to a more resolved information (Coyotzi et al., 2016).

With the development of metabolomics and fluxomics (metabolomic approach that seek to determine the rates of metabolic reactions within a biological entity), a new perspective could be to monitor the degradation of a complex labelled substrate in a totally untargeted way (Llufrio et al., 2019). So far, I could not find any published results on the combination of

isotope labelling and metabolomics in the AD context. In combination with metagenomics it would enable to resolve complex pathways and get further information to improve the understanding of anaerobic digesters and propose microbial management methods. Similarly, in a perspective of biorefinery, the functioning of bioelectrochemical systems relying on microbial electrosynthesis could also be targeted more thoughtfully with a combined SIP – omic approach, by feeding the microbial community of the cathode with labelled carbon dioxide (Batlle-Vilanova et al., 2019; Jiang et al., 2019).

3. Conclusion

Curtis concluded the paper I mentioned in the preamble (Curtis, 2006) as follows: "A 'Microbial Survey' could run, or subcontract, sequencing factories for the community at large, and explore the microbial world systematically itself using a judicious mix of deep-sequencing and mathematical modelling". That's more of less where we are arriving now. All the ingredients are available "to start thinking big" and address pending issues related to various microbial environments. The challenge is now to combine these ingredients adequately.

In parallel to the increased description of microbial ecosystems, which is under way, further development and miniaturization of analytical technologies coupled to deep-learning methodologies will undoubtedly favour the development of rapid microbial diagnosis solutions. They will enable to establish real microbial management strategies of environmental bioprocesses and more generally provide tools to evaluate the good functioning of microbial ecosystems. A new dimension of information will be accessed and will open new possibilities.

I hope my future research will participate to the development of such microbial management strategies for AD, and more generally provide frameworks for the analysis of microbial communities. It is now difficult to forecast what will be my research in ten years. Techniques evolve quickly and will open further possibilities. Other fields than AD could be targeted with the methods I'm using. Lines of research are highly dependent on funding but also on collaborations. My future research will depend on the combination of these different parameters.

References

Abouelenien, F., Nakashimada, Y., and Nishio, N. (2009) Dry mesophilic fermentation of chicken manure for production of methane by repeated batch culture. *J Biosci Bioeng* **107**: 293-295.

Abouelenien, F., Fujiwara, W., Namba, Y., Kosseva, M., Nishio, N., and Nakashimada, Y. (2010) Improved methane fermentation of chicken manure via ammonia removal by biogas recycle. *Bioresour Technol* **101**: 6368-6373.

Ali Shah, F., Mahmood, Q., Maroof Shah, M., Pervez, A., and Ahmad Asad, S. (2014) Microbial ecology of anaerobic digesters: the key players of anaerobiosis. *The Scientific World Journal* **2014**: 183752.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* **59**: 143-169.

Angelidaki, I., and Ahring, B.K. (1993) Thermophilic anaerobic digestion of livestock waste: the effect of ammonia. *Appl Microbiol Biotechnol* **38**: 560-564.

Anthonisen, A.C., Loehr, R.C., Prakasam, T.B.S., and Srinath, E.G. (1976) Inhibition of Nitrification by Ammonia and Nitrous Acid. *J Water Pollut Control Fed* **48**.

Badalato, N. (2014) Structure de déchets lignocellulosiques : effets sur la colonisation, les communautés microbiennes et les performances de méthanisation, caractérisés par des approches fonctionnelles et haut-débit. *Thèse de doctorat* **PhD supervisor : A. Bize**.

Barnes, S.P., and Keller, J. (2003) Cellulosic waste degradation by rumen-enhanced anaerobic digestion. *Water Sci Technol* **48**: 155-162.

Batlle-Vilanova, P., Rovira-Alsina, L., Puig, S., Balaguer, M.D., Icaran, P., Monsalvo, V.M. et al. (2019) Biogas upgrading, CO2 valorisation and economic revaluation of bioelectrochemical systems through anodic chlorine production in the framework of wastewater treatment plants. *Sci Total Environ* **690**: 352-360.

Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S.V., Pavlostathis, S.G., Rozzi, A. et al. (2002) The IWA Anaerobic Digestion Model No 1 (ADM1). *Water Sci Technol* **45**: 65-73.

Bayard, R., and Gourdon, R. (2001) Traitement biologique des déchets. *Techniques de l'ingénieur*.

Bayer, E.A., Lamed, R., and Himmel, M.E. (2007) The potential of cellulases and cellulosomes for cellulosic waste management. *Curr Opin Biotechnol* **18**: 237-245.

Bize, A., Cardona, L., Desmond-Le Quemener, E., Battimelli, A., Badalato, N., Bureau, C. et al. (2015) Shotgun metaproteomic profiling of biomimetic anaerobic digestion processes treating sewage sludge. *Proteomics* **15**: 3532-3543.

Blackman, J., Rutledge, D.N., Tesic, D., Saliba, A., and Scollary, G.R. (2010) Examination of the potential for using chemical analysis as a surrogate for sensory analysis. *Anal Chim Acta* **660**: 2-7.

Blasig, J.D., Holtzapple, M.T., Dale, B.E., Engler, C.R., and Byers, F.M. (1992) Volatile fatty acid fermentation of AFEX-treated bagasse and newspaper by rumen microorganisms. *Resources, Conservation and Recycling* **7**: 95-114.

Boccard, J., and Rutledge, D.N. (2013) A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion. *Anal Chim Acta* **769**: 30-39.

Bodein, A., Chapleur, O., Droit, A., and Lê Cao, K. (submitted) A generic multivariate framework for the integration of microbiome longitudinal studies with other data types. *Frontiers in Genetics*.

Boerjan, W., Ralph, J., and Baucher, M. (2003) Lignin biosynthesis. *Annu Rev Plant Biol* **54**: 519-546.

Boone, D.R., and Xun, L. (1987) Effects of pH, Temperature, and Nutrients on Propionate Degradation by a Methanogenic Enrichment Culture. *Appl Environ Microbiol* **53**: 1589-1592.

Borrel, G., O'Toole, P.W., Harris, H.M., Peyret, P., Brugere, J.F., and Gribaldo, S. (2013) Phylogenomic data support a seventh order of Methylotrophic methanogens and provide insights into the evolution of Methanogenesis. *Genome Biology and Evolution* **5**: 1769-1780.

Boschker, H.T.S., Nold, S.C., Wellsbury, P., Bos, D., de Graaf, W., Pel, R. et al. (1998) Direct linking of microbial populations to specific biogeochemical processes by 13C-labelling of biomarkers. *Nature* **392**: 801.

Bowen, E.J., Dolfing, J., Davenport, R.J., Read, F.L., and Curtis, T.P. (2014) Low-temperature limitation of bioreactor sludge in anaerobic treatment of domestic wastewater. *Water Sci Technol* **69**: 1004-1013.

Briones, A., and Raskin, L. (2003) Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Curr Opin Biotechnol* **14**: 270-276.

Broger, T., Odermatt, R.P., Huber, P., and Sonnleitner, B. (2011) Real-time on-line flow cytometry for bioprocess monitoring. *J Biotechnol* **154**: 240-247.

Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E. et al. (2016) MDSINE: Microbial Dynamical Systems INference Engine for microbiome time-series analyses. *Genome Biology* **17**: 121.

Busca, G., Berardinelli, S., Resini, C., and Arrighi, L. (2008) Technologies for the removal of phenol from fluid streams: a short review of recent developments. *J Hazard Mater* **160**: 265-288.

Calli, B., Mertoglu, B., Inanc, B., and Yenigun, O. (2005a) Methanogenic diversity in anaerobic bioreactors under extremely high ammonia levels. *Enzyme and Microbial Technology* **37**: 448-455.

Calli, B., Mertoglu, B., Inanc, B., and Yenigun, O. (2005b) Community changes during star-up in methanogenic bioreactors exposed to increasing levels of ammonia. *Environ Technol* **26**: 85-91.

Carballa, M., Regueiro, L., and Lema, J.M. (2015) Microbial management of anaerobic digestion: Exploiting the microbiome-functionality nexus. *Curr Opin Biotechnol* **33**: 103-111.

Carballa, M., Smits, M., Etchebehere, C., Boon, N., and Verstraete, W. (2011) Correlations between molecular and operational parameters in continuous lab-scale anaerobic reactors. *Appl Microbiol Biotechnol* **89**.

Cardona, L., Cao, K.A., and Chapleur, O. (submitted) Integrative analyses to investigate the link between microbial activity and molecules degradation during anaerobic digestion *Chem Eng J*.

Cardona, L., Levrard, C., Guenne, A., Chapleur, O., and Mazéas, L. (2019) Co-digestion of wastewater sludge: Choosing the optimal blend. *Waste Manag* **87**: 772-781.

Case, R.J., Boucher, Y., Dahllof, I., Holmstrom, C., Doolittle, W.F., and Kjelleberg, S. (2007) Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl Environ Microbiol* **73**: 278-288.

Chae, K.J., Jang, A., Yim, S.K., and Kim, I.S. (2008) The effects of digestion temperature and temperature shock on the biogas yields from the mesophilic anaerobic digestion of swine manure. *Bioresour Technol* **99**: 1-6.

Chapleur, O. (2012) Ingénierie écologique des communautés microbiennes de méthanisation des déchets ligno-cellulosiques. *Thèse de doctorat*.

Chapleur, O., Mazeas, L., Godon, J.J., and Bouchez, T. (2016a) Asymmetrical response of anaerobic digestion microbiota to temperature changes. *Appl Microbiol Biotechnol* **100**: 1445-1457.

Chapleur, O., Wu, T.-D., Guerquin-Kern, J.-L., Mazéas, L., and Bouchez, T. (2013) SIMSISH technique does not alter the apparent isotopic composition of bacterial cells. *PloS one* **8**: e77522.

Chapleur, O., Bize, A., Serain, T., Mazéas, L., and Bouchez, T. (2014) Co-inoculating ruminal content neither provides active hydrolytic microbes nor improves methanization of 13C-cellulose in batch digesters. *FEMS Microbiol Ecol* **87**: 616-629.

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

Chen, Y., and Murrell, J.C. (2010) When metagenomics meets stable-isotope probing: Progress and perspectives. *Trends Microbiol* **18**: 157-163.

Chen, Y., Cheng, J.J., and Creamer, K.S. (2008a) Inhibition of anaerobic digestion process: A review. *Bioresour Technol* **99**: 4044-4064.

Chen, Y., Dumont, M.G., Neufeld, J.D., Bodrossy, L., Stralis-Pavese, N., McNamara, N.P. et al. (2008b) Revealing the uncultivated majority: Combining DNA stable-isotope probing, multiple displacement amplification and metagenomic analyses of uncultivated Methylocystis in acidic peatlands. *Environmental Microbiology* **10**: 2609-2622.

Cirne, D.G., Paloumet, X., Björnsson, L., Alves, M.M., and Mattiasson, B. (2007) Anaerobic digestion of lipid-rich waste—Effects of lipid concentration. *Renewable Energy* **32**: 965-975.

Coyotzi, S., Pratscher, J., Murrell, J.C., and Neufeld, J.D. (2016) Targeted metagenomics of active microbial populations with stable-isotope probing. *Curr Opin Biotechnol* **41**: 1-8.

Curtis, T. (2006) Microbial ecologists: It's time to 'go large'. *Nat Rev Microbiol* **4**: 488.

De Kok, S., Meijer, J., Van Loosdrecht, M.C.M., and Kleerebezem, R. (2013) Impact of dissolved hydrogen partial pressure on mixed culture fermentations. *Appl Microbiol Biotechnol* **97**: 2617-2625.

De los Reyes, F.L., Weaver, J.E., and Wang, L. (2015) A methodological framework for linking bioreactor function to microbial communities and environmental conditions. *Curr Opin Biotechnol* **33**: 112-118.

De Vrieze, J., Verstraete, W., and Boon, N. (2013) Repeated pulse feeding induces functional stability in anaerobic digestion. *Microb Biotechnol* **6**: 414-424.

De Vrieze, J., Pinto, A.J., Sloan, W.T., and Ijaz, U.Z. (2018) The active microbial community more accurately reflects the anaerobic digestion process: 16S rRNA (gene) sequencing as a predictive tool. *Microbiome* **6**: 63.

De Vrieze, J., Smet, D., Klok, J., Colsen, J., Angenent, L.T., and Vlaeminck, S.E. (2016a) Thermophilic sludge digestion improves energy balance and nutrient recovery potential in full-scale municipal wastewater treatment plants. *Bioresour Technol* **218**: 1237-1245.

De Vrieze, J., Christiaens, M.E.R., Walraedt, D., Devooght, A., Ijaz, U.Z., and Boon, N. (2017) Microbial community redundancy in anaerobic digestion drives process recovery after salinity exposure. *Water Res* **111**: 109-117.

De Vrieze, J., Regueiro, L., Props, R., Vilchez-Vargas, R., Jáuregui, R., Pieper, D.H. et al. (2016b) Presence does not imply activity: DNA and RNA patterns differ in response to salt perturbation in anaerobic digestion. *Biotechnology for Biofuels* **9**: 244.

Déjean, S., Martin, P.G.P., Baccini, A., and Besse, P. (2007) Clustering time-series gene expression data using smoothing spline derivatives. *Eurasip Journal on Bioinformatics and Systems Biology* **2007**.

Delforno, T.P., Macedo, T.Z., Midoux, C., Lacerda, G.V., Jr., Rué, O., Mariadassou, M. et al. (2019) Comparative metatranscriptomic analysis of anaerobic digesters treating anionic surfactant contaminated wastewater. *Sci Total Environ* **649**: 482-494.

Desloover, J., De Vrieze, J., Van de Vijver, M., Mortelmans, J., Rozendal, R., and Rabaey, K. (2015) Electrochemical nutrient recovery enables ammonia toxicity control and biogas desulfurization in anaerobic digestion. *Environ Sci Technol* **49**: 948-955.

Dumont, M.G., and Murrell, J.C. (2005) Stable isotope probing - linking microbial identity to function. *Nat Rev Microbiol* **3**: 499-504.

Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K. et al. (2018) FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics* **34**: 1287-1294.

Fang, H.H.P., Liang, D.W., Zhang, T., and Liu, Y. (2006) Anaerobic treatment of phenol in wastewater under thermophilic condition. *Water Res* **40**: 427-434.

Faust, K., Lahti, L., Gonze, D., de Vos, W.M., and Raes, J. (2015) Metagenomics meets time series analysis: unraveling microbial community dynamics. *Curr Opin Microbiol* **25**: 56-66.

Fechner, L.C., Gourlay-Francé, C., Uher, E., and Tusseau-Vuillemin, M.H. (2010) Adapting an enzymatic toxicity test to allow comparative evaluation of natural freshwater biofilms' tolerance to metals. *Ecotoxicology* **19**: 1302-1311.

Fedorak, P.M., and Hrudey, S.E. (1984) The effects of phenol and some alkyl phenolics on batch anaerobic methanogenesis. *Water Res* **18**: 361-367.

Fernández-Rodríguez, J., Pérez, M., and Romero, L.I. (2015) Temperature-phased anaerobic digestion of Industrial Organic Fraction of Municipal Solid Waste: A batch study. *Chem Eng J* **270**: 597-604.

Fernandez, A., Huang, S., Seston, S., Xing, J., Hickey, R., Criddle, C., and Tiedje, J. (1999) How stable is stable? Function versus community composition. *Appl Environ Microbiol* **65**: 3697-3704.

Fotidis, I.A., Wang, H., Fiedel, N.R., Luo, G., Karakashev, D.B., and Angelidaki, I. (2014) Bioaugmentation as a Solution To Increase Methane Production from an Ammonia-Rich Substrate. *Environ Sci Technol* **48**: 7669-7676.

Gagnon-Bartsch, J.A., and Speed, T.P. (2012) Using control genes to correct for unwanted variation in microarray data. *Biostatistics* **13**: 539-552.

Gao, S., Zhao, M., Chen, Y., Yu, M., and Ruan, W. (2015) Tolerance response to in situ ammonia stress in a pilot-scale anaerobic digestion reactor for alleviating ammonia inhibition. *Bioresour Technol* **198**: 372-379.

Garcia-Martinez, J., Acinas, S.G., Antan, A.I., and Rodriguez-Valera, F. (1999) Use of the 16S-23S ribosomal genes spacer region in studies of prokaryotic diversity. *J Microbiol Methods* **36**: 55-64.

Ge, Y., He, J.Z., Zheng, Y.M., Zhang, L.M., and Zhu, Y.G. (2006) Stable isotope probing and its applications in microbial ecology. *Acta Ecologica Sinica* **26**: 1574-1582.

Gijzen, H.J., Lubberding, H.J., Verhagen, F.J., Zwart, K.B., and Vogels, G.D. (1987) Application of rumen microorganisms for an enhanced anaerobic degradation of solid organic waste materials. *Biological Wastes* **22**: 81-95.

Girguis, P.R., Cozen, A.E., and DeLong, E.F. (2005) Growth and Population Dynamics of Anaerobic Methane-Oxidizing Archaea and Sulfate-Reducing Bacteria in a Continuous-Flow Bioreactor. *Appl Environ Microbiol* **71**: 3725-3733.

Goel, R., Mino, T., Satoh, H., and Matsuo, T. (1998) Comparison of hydrolytic enzyme systems in pure culture and activated sludge under different electron acceptor conditions. *Water Sci Technol* **37**: 335-343.

Grossin-Debattista, J. (2011) Fractionnements isotopiques (13C/12C) engendrés par la méthanogenèse : apports pour la compréhension des processus de biodégradation lors de la digestion anaérobie. *Thèse de doctorat* **PhD supervisor: L. Mazéas**.

Hajji, K.T., Lepine, F., Bisaillon, J.G., Beaudet, R., Hawari, J., and Guiot, S.R. (2000) Effects of bioaugmentation strategies in UASB reactors with a methanogenic consortium for removal of phenolic compounds. *Biotechnol Bioeng* **67**: 417-423.

Hammes, F., Broger, T., Weilenmann, H.U., Vital, M., Helbing, J., Bosshart, U. et al. (2012) Development and laboratory-scale testing of a fully automated online flow cytometer for drinking water analysis. *Cytometry Part A* **81 A**: 508-516.

Hanreich, A., Schimpf, U., Zakrzewski, M., Schlüter, A., Benndorf, D., Heyer, R. et al. (2013) Metagenome and metaproteome analyses of microbial communities in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation. *Syst Appl Microbiol* **36**: 330-338.

Ho, L., and Ho, G. (2012) Mitigating ammonia inhibition of thermophilic anaerobic treatment of digested piggery wastewater: use of pH reduction, zeolite, biomass and humic acid. *Water Res* **46**: 4339-4350.

Holmes, D.E., Nevin, K.P., and Lovley, D.R. (2004) Comparison of 16S rRNA, nifD, recA, gyrB, rpoB and fusA genes within the family Geobacteraceae fam. nov. *Int J Syst Evol Microbiol* **54**: 1591-1599.

Hoyos-Hernandez, C., Hoffmann, M., Guenne, A., and Mazeas, L. (2014) Elucidation of the thermophilic phenol biodegradation pathway via benzoate during the anaerobic digestion of municipal solid waste. *Chemosphere* **97**: 115-119.

Huang, W.E., Stoecker, K., Griffiths, R., Newbold, L., Daims, H., Whiteley, A.S., and Wagner, M. (2007) Raman-FISH: Combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. *Environmental Microbiology* **9**: 1878-1889.

Hudayah, N., Suraraksa, B., and Chaiprasert, P. (2016) Synergistic effects of the chitosan addition and polysaccharides-EPS on the formation of anaerobic granules. *Environ Technol* **37**: 2713-2722.

Huiliñir, C., Roa, E., Vargas, D., Roeckel, M., and Aspé, E. (2008) Kinetics of syntrophic acetogenesis in a saline medium. *J Chem Technol Biotechnol* **83**: 1433-1440.

Ito, T., Yoshiguchi, K., Ariesyady, H.D., and Okabe, S. (2011) Identification of a novel acetate-utilizing bacterium belonging to Synergistes group 4 in anaerobic digester sludge. *ISME J* **5**: 1844-1856.

Ito, T., Yoshiguchi, K., Ariesyady, H.D., and Okabe, S. (2012) Identification and quantification of key microbial trophic groups of methanogenic glucose degradation in an anaerobic digester sludge. *Bioresour Technol* **123**: 599-607.

Jehmlich, N., Schmidt, F., Von Bergen, M., Richnow, H.H., and Vogt, C. (2008) Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. *ISME Journal* **2**: 1122-1133.

Jiang, H., Kilburn, M.R., Decelle, J., and Musat, N. (2016) NanoSIMS chemical imaging combined with correlative microscopy for biological sample analysis. *Curr Opin Biotechnol* **41**: 130-135.

Jiang, Y., May, H.D., Lu, L., Liang, P., Huang, X., and Ren, Z.J. (2019) Carbon dioxide and organic waste valorization by microbial electrosynthesis and electro-fermentation. *Water Res*: 42-55.

Johnson, D.B. (1998) Biodiversity and ecology of acidophilic microorganisms. *FEMS Microbiol Ecol* **27**: 307-317.

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T. et al. (2016) Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Frontiers in Microbiology* **7**.

Juteau, P., Cote, V., Duckett, M.F., Beaudet, R., Lepine, F., Villemur, R., and Bisaillon, J.G. (2005) Cryptanaerobacter phenolicus gen. nov., sp. nov., an anaerobe that transforms phenol into benzoate via 4-hydroxybenzoate. *Int J Syst Evol Microbiol* **55**: 245-250.

Karakashev, D., Batstone, D.J., Trably, E., and Angelidaki, I. (2006) Acetate oxidation is the dominant methanogenic pathway from acetate in the absence of Methanosaetaceae. *Appl Environ Microbiol* **72**: 5138-5141.

Kayhanian, M. (1994) Performance of a high-solids anaerobic digestion process under various ammonia concentrations. *J Chem Technol Biotechnol* **59**: 349-352.

Kim, M., Ahn, Y.H., and Speece, R.E. (2002) Comparative process stability and efficiency of anaerobic digestion; mesophilic vs. thermophilic. *Water Res* **36**: 4369-4385.

Kleinsteuber, S., Schleinitz, K.M., and Vogt, C. (2012) Key players and team play: Anaerobic microbial communities in hydrocarbon-contaminated aquifers. *Appl Microbiol Biotechnol* **94**: 851-873.

Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J.A. et al. (2012) Unlocking the potential of metagenomics through replicated experimental design. *Nat Biotechnol* **30**: 513.

Koch, C., Müller, S., Harms, H., and Harnisch, F. (2014) Microbiomes in bioenergy production: From analysis to management. *Curr Opin Biotechnol* **27**: 65-72.

Koch, C., Fetzer, I., Schmidt, T., Harms, H., and Müller, S. (2013) Monitoring functions in managed microbial systems by cytometric bar coding. *Environmental Science and Technology* **47**: 1753-1760.

Koster, I.W., and Lettinga, G. (1984) The influence of ammonium-nitrogen on the specific activity of pelletized methanogenic sludge. *Agricultural Wastes* **9**: 205-216.

Kundu, K., Bergmann, I., Klocke, M., Sharma, S., and Sreekrishnan, T.R. (2014) Impact of abrupt temperature increase on the performance of an anaerobic hybrid bioreactor and its intrinsic microbial community. *Bioresour Technol* **168**: 72-79.

Labatut, R.A., Angenent, L.T., and Scott, N.R. (2014) Conventional mesophilic vs. thermophilic anaerobic digestion: Atrade-off between performance and stability? *Water Res* **53**: 249-258.

Laguerre, G., Allard, M.R., Revoy, F., and Amarger, N. (1994) Rapid identification of rhizobia by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl Environ Microbiol* **60**: 56-63.

Lau, I.W.C., and Fang, H.H.P. (1997) Effect of temperature shock to thermophilic granules. *Water Res* **31**: 2626-2632.

Lê Cao, K.A., Boitard, S., and Besse, P. (2011) Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* **12**.

Lê Cao, K.A., Rossouw, D., Robert-Granié, C., and Besse, P. (2008) A sparse PLS for variable selection when integrating omics data. *Statistical Applications in Genetics and Molecular Biology* **7**.

Lee, G., Nho, K., Kang, B., Sohn, K.A., Kim, D., Weiner, M.W. et al. (2019) Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Scientific Reports* **9**.

Lee, J.Y., Lee, S.H., and Park, H.D. (2016) Enrichment of specific electro-active microorganisms and enhancement of methane production by adding granular activated carbon in anaerobic reactors. *Bioresour Technol* **205**: 205-212.

Lee, L.C., Liong, C.-Y., and Jemain, A.A. (2018) Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* **143**: 3526-3539.

Lee, N., Nielsen, P.H., Andreasen, K.H., Juretschko, S., Nielsen, J.L., Schleifer, K.H., and Wagner, M. (1999) Combination of fluorescent in situ hybridization and microautoradiography - A new tool for structure-function analyses in microbial ecology. *Appl Environ Microbiol* **65**: 1289-1297.

Lefebvre, O., Quentin, S., Torrijos, M., Godon, J.J., Delgenes, J.P., and Moletta, R. (2007) Impact of increasing NaCl concentrations on the performance and community composition of two anaerobic reactors. *Appl Microbiol Biotechnol* **75**.

Lei, X., Sugiura, N., Feng, C., and Maekawa, T. (2007) Pretreatment of anaerobic digestion effluent with ammonia stripping and biogas purification. *J Hazard Mater* **145**: 391-397.

Lertsittichai, S., Lertsutthiwong, P., and Phalakornkule, C. (2007) Improvement of upflow anaerobic sludge bed performance using chitosan. *Water Environ Res* **79**: 801-807.

Levén, L., Nyberg, K., and Schnurer, A. (2012) Conversion of phenols during anaerobic digestion of organic solid waste--a review of important microorganisms and impact of temperature. *J Environ Manage* **95 Suppl**: S99-103.

Li, L., Peng, X., Wang, X., and Wu, D. (2018) Anaerobic digestion of food waste: A review focusing on process stability. *Bioresour Technol* **248**: 20-28.

Li, T., Wu, T.D., Mazéas, L., Toffin, L., Guerquin-Kern, J.L., Leblon, G., and Bouchez, T. (2008) Simultaneous analysis of microbial identity and function using NanoSIMS. *Environmental Microbiology* **10**: 580-588.

Lin, Q., De Vrieze, J., Li, J., and Li, X. (2016) Temperature affects microbial abundance, activity and interactions in anaerobic digestion. *Bioresour Technol* **209**: 228-236.

Lindorfer, H., Waltenberger, R., Kollner, K., Braun, R., and Kirchmayr, R. (2008) New data on temperature optimum and temperature changes in energy crop digesters. *Bioresour Technol* **99**: 7011-7019.

Liu, F., Rotaru, A.-E., Shrestha, P.M., Malvankar, N.S., Nevin, K.P., and Lovley, D.R. (2012) Promoting direct interspecies electron transfer with activated carbon. *Energy & Environmental Science* **5**: 8982-8989.

Llufrio, E.M., Cho, K., and Patti, G.J. (2019) Systems-level analysis of isotopic labeling in untargeted metabolomic data by X13CMS. *Nature Protocols* **14**: 1970-1990.

Lovley, D.R., and Klug, M.J. (1983) Methanogenesis from Methanol and Methylamines and Acetogenesis from Hydrogen and Carbon Dioxide in the Sediments of a Eutrophic Lake. *Appl Environ Microbiol* **45**: 1310-1315.

Lu, F., Luo, C., Shao, L., and He, P. (2016) Biochar alleviates combined stress of ammonium and acids by firstly enriching Methanosaeta and then Methanosarcina. *Water Res* **90**: 34-43.

Lü, F., Bize, A., Guillot, A., Monnet, V., Madigou, C., Chapleur, O. et al. (2014) Metaproteomics of cellulose methanisation under thermophilic conditions reveals a surprisingly high proteolytic activity. *ISME Journal* **8**: 88-102.

Lueders, T., Dumont, M.G., Bradford, L., and Manefield, M. (2016) RNA-stable isotope probing: From carbon flow within key microbiota to targeted transcriptomes. *Curr Opin Biotechnol* **41**: 83-89.

Luo, G., De Francisci, D., Kougias, P.G., Laura, T., Zhu, X., and Angelidaki, I. (2015) New steady-state microbial community compositions and process performances in biogas reactors induced by temperature disturbances. *Biotechnology for Biofuels* **8**.

Lynd, L.R., Weimer, P.J., Van Zyl, W.H., and Pretorius, I.S. (2002) Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol Mol Biol Rev* **66**: 506-577.

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

Madigou, C., Lê Cao, K.-A., Bureau, C., Mazéas, L., Déjean, S., and Chapleur, O. (2019) Ecological consequences of abrupt temperature changes in anaerobic digesters. *Chem Eng J* **361**: 266-277.

Manefield, M., Whiteley, A.S., Griffiths, R.I., and Bailey, M.J. (2002) RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Applied Environmental Microbiology* **68**: 5367-5373.

Mata-Alvarez, J., Macé, S., and Llabrés, P. (2000) Anaerobic digestion of organic solid wastes. An overview of research achievements and perspectives. *Bioresour Technol* **74**: 3-16.

Mata-Alvarez, J., Dosta, J., Romero-Güiza, M.S., Fonoll, X., Peces, M., and Astals, S. (2014) A critical review on anaerobic co-digestion achievements between 2010 and 2013. *Renewable and Sustainable Energy Reviews* **36**: 412-427.

Maus, I., Koeck, D.E., Cibis, K.G., Hahnke, S., Kim, Y.S., Langer, T. et al. (2016) Unraveling the microbiome of a thermophilic biogas plant by metagenome and metatranscriptome analysis complemented by characterization of bacterial and archaeal isolates. *Biotechnology for Biofuels* **9**.

McDonnell, G., and Russell, A.D. (1999) Antiseptics and Disinfectants: Activity, Action, and Resistance. *Clin Microbiol Rev* **12**: 147-179.

McKeown, R.M., Scully, C., Enright, A.M., Chinalia, F.A., Lee, C., Mahony, T. et al. (2009) Psychrophilic methanogenic community development during long-term cultivation of anaerobic granular biofilms. *ISME Journal* **3**: 1231-1242.

McMahon, K.D., Zheng, D., Stams, A.J.M., Mackie, R.I., and Raskin, L. (2004) Microbial population dynamics during start-up and overload conditions of anaerobic digesters treating municipal solid waste and sewage sludge. *Biotechnol Bioeng* **87**: 823-834.

Metwally, A.A., Yang, J., Ascoli, C., Dai, Y., Finn, P.W., and Perkins, D.L. (2018) MetaLonDA: a flexible R package for identifying time intervals of differentially abundant features in metagenomic longitudinal studies. *Microbiome* **6**: 32.

Milan, Z., Montalvo, S., Ilangovan, K., Monroy, O., Chamy, R., Weiland, P. et al. (2010) The impact of ammonia nitrogen concentration and zeolite addition on the specific methanogenic activity of granular and flocculent anaerobic sludges. *Journal of environmental science and health Part A, Toxic/hazardous substances & environmental engineering* **45**: 883-889.

Moletta, R. (1993) Anaerobic processing: From the smallest to the largest. *La digestion anaérobie : du plus petit au plus grand*: 16-25.

Moletta, R. (2015) La méthanisation. 3ème édition. *Tec & Doc Editions Lavoisier Paris.*

Monlau, F., Sambusiti, C., Barakat, A., Quemeneur, M., Trably, E., Steyer, J.P., and Carrere, H. (2014) Do furanic and phenolic compounds of lignocellulosic and algae biomass hydrolyzate inhibit anaerobic mixed cultures? A comprehensive review. *Biotechnol Adv* **32**: 934-951.

Montalvo, S., Martin, J.S., Huiliñir, C., Guerrero, L., and Borja, R. (2014) Assessment of a UASB reactor with high ammonia concentrations: Effect of zeolite addition on process performance. *Process Biochem* **49**: 2220-2227.

Montalvo, S., Guerrero, L., Borja, R., Sánchez, E., Milán, Z., Cortés, I., and Angeles de la la Rubia, M. (2012) Application of natural zeolites in anaerobic digestion processes: A review. *Applied Clay Science* **58**: 125-133.

Moset, V., Poulsen, M., Wahid, R., Højberg, O., and Møller, H.B. (2015) Mesophilic versus thermophilic anaerobic digestion of cattle manure: Methane productivity and microbial ecology. *Microbial Biotechnology* **8**: 787-800.

Mottet, A., Habouzit, F., and Steyer, J.P. (2014) Anaerobic digestion of marine microalgae in different salinity levels. *Bioresour Technol* **158**: 300-306.

Moyer, C.L., Dobbs, F.C., and Karl, D.M. (1994) Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Appl Environ Microbiol* **60**: 871-879.

Mumme, J., Srocke, F., Heeg, K., and Werner, M. (2014) Use of biochars in anaerobic digestion. *Bioresour Technol* **164**: 189-197.

Musat, N., Musat, F., Weber, P.K., and Pett-Ridge, J. (2016) Tracking microbial interactions with NanoSIMS. *Curr Opin Biotechnol* **41**: 114-121.

Muyzer, G., De Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695-700.

Neufeld, J.D., Wagner, M., and Murrell, J.C. (2007) Who eats what, where and when? Isotope-labelling experiments are coming of age. *ISME Journal* **1**: 103-110.

Nie, Y., Liu, H., Du, G., and Chen, J. (2008) Acetate yield increased by gas circulation and fed-batch fermentation in a novel syntrophic acetogenesis and homoacetogenesis coupling system. *Bioresour Technol* **99**: 2989-2995.

Nissilä, M.E., Li, Y.-C., Wu, S.-Y., Lin, C.-Y., and Puhakka, J.A. (2012) Hydrogenic and methanogenic fermentation of birch and conifer pulps. *Applied Energy* **100**: 58-65.

Noll, M., Klose, M., and Conrad, R. (2010) Effect of temperature change on the composition of the bacterial and archaeal community potentially involved in the turnover of acetate and propionate in methanogenic rice field soil. *FEMS Microbiol Ecol* **73**: 215-225.

Nsereko, V.L., Morgavi, D.P., Rode, L.L.M., Beauchemin, K.A., and McAllister, T.A. (2000) Effects of fungal enzyme preparations on hydrolysis and subsequent degradation of alfalfa hay fiber by mixed rumen microorganisms in vitro. *Anim Feed Sci Technol* **88**: 153-170.

Nuñez, J., Renslow, R., Cliff, J.B., and Anderton, C.R. (2018) NanoSIMS for biological applications: Current practices and analyses. *Biointerphases* **13**.

O'Sullivan, C.A., and Burrell, P.C. (2007) The effect of media changes on the rate of cellulose solubilisation by rumen and digester derived microbial communities. *Waste Manag* **27**: 1808-1814.

Olguin-Lora, P., Puig-Grajales, L., and Razo-Flores, E. (2003) Inhibition of the acetoclastic methanogenic activity by phenol and alkyl phenols. *Environ Technol* **24**: 999-1006.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**: 874-879.

Orphan, V.J., House, C.H., Hinrichs, K.U., McKeegan, K.D., and DeLong, E.F. (2001) Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**: 484-487.

Pap, B., Györkei, Á., Boboescu, I.Z., Nagy, I.K., Bíró, T., Kondorosi, É., and Maróti, G. (2015) Temperature-dependent transformation of biogas-producing microbial communities points to the increased importance of hydrogenotrophic methanogenesis under thermophilic operation. *Bioresour Technol* **177**: 375-380.

Paul, K., Nonoh, J.O., Mikulski, L., and Brune, A. (2012) "Methanoplasmatales," Thermoplasmatales-related archaea in termite guts and other environments, are the seventh order of methanogens. *Appl Environ Microbiol* **78**: 8245-8253.

Peng, X., Zhang, S., Li, L., Zhao, X., Ma, Y., and Shi, D. (2018) Long-term high-solids anaerobic digestion of food waste: Effects of ammonia on process performance and microbial community. *Bioresour Technol* **262**: 148-158.

Pinjing, H., Fan, L., Liming, S., and Hua, Z. (2009) Characterization of methanogenesis using stable isotopic probing. *Progress in Chemistry* **21**: 540-549.

Poirier, S., and Chapleur, O. (2018a) Inhibition of anaerobic digestion by phenol and ammonia: Effect on degradation performances and microbial dynamics. *Data in Brief*: 2235-2239.

Poirier, S., and Chapleur, O. (2018b) Influence of support media supplementation to reduce the inhibition of anaerobic digestion by phenol and ammonia: Effect on degradation performances and microbial dynamics. *Data in Brief* **19**: 1733-1754.

Poirier, S., Déjean, S., and Chapleur, O. (2018a) Support media can steer methanogenesis in the presence of phenol through biotic and abiotic effects. *Water Res* **140**: 24-33.

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

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

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

Poirier, S., Déjean, S., Midoux, C., Lê Cao, K., and Chapleur, O. (submitted) Predictive models of AD inhibition can be built by integrating microbial patterns from independent studies. *The Isme Journal*.

Poirier, S., Rué, O., Peguilhan, R., Coeuret, G., Zagorec, M., Champomier-Vergès, M.-C. et al. (2018b) Deciphering intra-species bacterial diversity of meat and seafood spoilage microbiota using gyrB amplicon sequencing: A comparative analysis with 16S rDNA V3-V4 amplicon sequencing. *PLOS ONE* **13**: e0204629.

Prosser, J.I. (2010) Replicate or lie. *Environmental Microbiology* **12**: 1806-1810.

Pugin, B., Barcik, W., Westermann, P., Heider, A., Wawrzyniak, M., Hellings, P. et al. (2017) A wide diversity of bacteria from the human gut produces and degrades biogenic amines. *Microb Ecol Health Dis* **28**: 1353881.

Radajewski, S., McDonald, I.R., and Murrell, J.C. (2003) Stable-isotope probing of nucleic acids: A window to the function of uncultured microorganisms. *Curr Opin Biotechnol* **14**: 296-302.

Radajewski, S., Ineson, P., Parekh, N.R., and Murrell, J.C. (2000) Stable-isotope probing as a tool in microbial ecology. *Nature* **403**: 646-649.

Rajagopal, R., Massé, D.I., and Singh, G. (2013) A critical review on inhibition of anaerobic digestion process by excess ammonia. *Bioresour Technol* **143**: 632-641.

Ranjard, L., Poly, F., and Nazaret, S. (2000) Monitoring complex bacterial communities using culture-independent molecular techniques: Application to soil environment. *Res Microbiol* **151**: 167-177.

Regueiro, L., Carballa, M., and Lema, J.M. (2014) Outlining microbial community dynamics during temperature drop and subsequent recovery period in anaerobic co-digestion systems. *J Biotechnol* **192**: 179-186.

Ridenhour, B.J., Brooker, S.L., Williams, J.E., Van Leuven, J.T., Miller, A.W., Dearing, M.D., and Remien, C.H. (2017) Modeling time-series data from microbial communities. *ISME J* **11**: 2526-2537.

Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J. et al. (2009) Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME Journal* **3**: 700-714.

Roeder, J., and Schink, B. (2009) Syntrophic degradation of cadaverine by a defined methanogenic coculture. *Appl Environ Microbiol* **75**: 4821-4828.

Rohart, F., Gautier, B., Singh, A., and Lê Cao, K.-A. (2017a) mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLOS Computational Biology* **13**: e1005752.

Rohart, F., Eslami, A., Matigian, N., Bougeard, S., and Lê Cao, K.A. (2017b) MINT: A multivariate integrative method to identify reproducible molecular signatures across independent experiments and platforms. *BMC Bioinformatics* **18**.

Rosenkranz, F., Cabrol, L., Carballa, M., Donoso-Bravo, A., Cruz, L., Ruiz-Filippi, G. et al. (2013) Relationship between phenol degradation efficiency and microbial community structure in an anaerobic SBR. *Water Res* **47**: 6739-6749.

Sakai, S., Imachi, H., Hanada, S., Ohashi, A., Harada, H., and Kamagata, Y. (2008) Methanocella paludicola gen. nov., sp. nov., a methane-producing archaeon, the first isolate of the lineage 'Rice Cluster I', and proposal of the new archaeal order Methanocellales ord. nov. *Int J Syst Evol Microbiol* **58**: 929-936.

Sales, C.M., and Lee, P.K.H. (2015) Resource recovery from wastewater: Application of meta-omics to phosphorus and carbon management. *Curr Opin Biotechnol* **33**: 260-267.

Sanders, W.T., Geerink, M., Zeeman, G., and Lettinga, G. (2000) Anaerobic hydrolysis kinetics of particulate substrates. *Water Sci Technol* **41**: 17-24.

Schink, B. (1997) Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol Mol Biol Rev* **61**: 262-280.

Schnürer, A., and Nordberg, A. (2008) Ammonia, a selective agent for methane production by syntrophic acetate oxidation at mesophilic temperature. *Water Sci Technol* **57**: 735-740.

Schnürer, A., Zellner, G., and Svensson, B.H. (1999) Mesophilic syntrophic acetate oxidation during methane formation in biogas reactors. *FEMS Microbiol Ecol* **29**: 249-261.

Serna-Maza, A., Heaven, S., and Banks, C.J. (2015) Biogas stripping of ammonia from fresh digestate from a food waste digester. *Bioresour Technol* **190**: 66-75.

Shahi, S.K., Freedman, S.N., and Mangalam, A.K. (2017) Gut microbiome in multiple sclerosis: The players involved and the roles they play. *Gut Microbes* **8**: 607-615.

Siegert, I., and Banks, C. (2005) The effect of volatile fatty acid additions on the anaerobic digestion of cellulose and glucose in batch reactors. *Process Biochem* **40**: 3412-3418.

Sierra-Alvarez, R., and Lettinga, G. (1991) The effect of aromatic structure on the inhibition of acetoclastic methanogenesis in granular sludge. *Appl Microbiol Biotechnol* **34**: 544-550.

Singh, A., Gautier, B., Shannon, C.P., Vacher, M., Rohart, F., Tebutt, S.J., and Le Cao, K.-A. (2016) DIABLO - an integrative, multi-omics, multivariate method for multi-group classification. *bioRxiv*.

Siriwongrungson, V., Zeng, R.J., and Angelidaki, I. (2007) Homoacetogenesis as the alternative pathway for H2 sink during thermophilic anaerobic degradation of butyrate under suppressed methanogenesis. *Water Res* **41**: 4204-4210.

Stams, A.J.M. (1994) Metabolic interactions between anaerobic bacteria in methanogenic environments. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* **66**: 271-294.

Straube, J., Huang, B.E., and Cao, K.A.L. (2017) DynOmics to identify delays and co-expression patterns across time course experiments. *Scientific Reports* **7**.

Straube, J., Gorse, A.D., Huang, B.E., and Le Cao, K.A. (2015) A Linear Mixed Model Spline Framework for Analysing Time Course 'Omics' Data. *PLoS One* **10**: e0134540.

Strik, D.P.B.T.B., Domnanovich, A.M., and Holubar, P. (2006) A pH-based control of ammonia in biogas during anaerobic digestion of artificial pig manure and maize silage. *Process Biochem* **41**: 1235-1238.

Subramanyam, R. (2013) Physicochemical and morphological characteristics of granular sludge in upflow anaerobic sludge blanket reactors. *Environ Eng Sci* **30**: 201-212.

Sun, Q., Zhao, W., Wang, L., Guo, F., Song, D., Zhang, Q. et al. (2019) Integration of metabolomic and transcriptomic profiles to identify biomarkers in serum of lung cancer. *J Cell Biochem* **120**: 11981-11989.

Sun, W., Yu, G., Louie, T., Liu, T., Zhu, C., Xue, G., and Gao, P. (2015) From mesophilic to thermophilic digestion: the transitions of anaerobic bacterial, archaeal, and fungal community structures in sludge and manure samples. *Appl Microbiol Biotechnol* **99**: 10271-10282.

Sung, S., and Liu, T. (2003) Ammonia inhibition on thermophilic anaerobic digestion. *Chemosphere* **53**: 43-52.

Tao, Y., Huang, X., Gao, D., Wang, X., Chen, C., Liang, H., and van Loosdrecht, M.C.M. (2019) NanoSIMS reveals unusual enrichment of acetate and propionate by an anammox consortium dominated by Jettenia asiatica. *Water Res* **159**: 223-232.

Tenenhaus, A., Philippe, C., Guillemot, V., Le Cao, K.A., Grill, J., and Frouin, V. (2014) Variable selection for generalized canonical correlation analysis. *Biostatistics* **15**: 569-583.

Theuerl, S., Klang, J., and Prochnow, A. (2019) Process disturbances in agricultural biogas production—causes, mechanisms and effects on the biogas microbiome: A review. *Energies* **12**.

Torres, B., Porras, G., Garcia, J.L., and Diaz, E. (2003) Regulation of the mhp cluster responsible for 3-(3-hydroxyphenyl)propionic acid degradation in Escherichia coli. *J Biol Chem* **278**: 27575-27585.

Tsilimigras, M.C.B., and Fodor, A.A. (2016) Compositional data analysis of the microbiome: fundamentals, tools, and challenges. *Ann Epidemiol* **26**: 330-335.

Ueyama, J., Saito, I., Takaishi, A., Nomura, H., Inoue, M., Osaka, A. et al. (2014) A revised method for determination of dialkylphosphate levels in human urine by solid-phase extraction and liquid chromatography with tandem mass spectrometry: application to human urine samples from Japanese children. *Environ Health Prev Med* **19**: 405-413.

Uludag-Demirer, S., Demirer, G.N., Frear, C., and Chen, S. (2008) Anaerobic digestion of dairy manure with enhanced ammonia removal. *J Environ Manag* **86**: 193-200.

Van Lier, J.B., Rebac, S., and Lettinga, G. (1997) High-rate anaerobic wastewater treatment under psychrophilic and thermophilic conditions. *Water Sci Technol* **35**: 199-206.

Van Velsen, A.F.M. (1979) Adaptation of methanogenic sludge to high ammonia-nitrogen concentrations. *Water Res* **13**: 995-999.

Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J., and Tyson, G.W. (2014a) Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. *Curr Opin Biotechnol* **27**: 55-64.

Vanwonterghem, I., Jensen, P.D., Dennis, P.G., Hugenholtz, P., Rabaey, K., and Tyson, G.W. (2014b) Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters. *ISME J* **8**: 2015-2028.

Veeken, A., Kalyuzhnyi, S., Scharff, H., and Hamelers, B. (2000) Effect of pH and VFA on hydrolysis of organic solid waste. *Journal of Environmental Engineering* **126**: 1076-1081.

Veeresh, G.S., Kumar, P., and Mehrotra, I. (2005) Treatment of phenol and cresols in upflow anaerobic sludge blanket (UASB) process: A review. *Water Res* **39**: 154-170.

Větrovský, T., and Baldrian, P. (2013) The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLOS ONE* **8**: e57923.

Wagner, M. (2009) Single-Cell Ecophysiology of Microbes as Revealed by Raman Microspectroscopy or Secondary Ion Mass Spectrometry Imaging. *Annu Rev Microbiol* **63**: 411-429.

Wang, X., Lu, X., Li, F., and Yang, G. (2014) Effects of Temperature and Carbon-Nitrogen (C/N) Ratio on the Performance of Anaerobic Co-Digestion of Dairy Manure, Chicken Manure and Rice Straw: Focusing on Ammonia Inhibition. *PLoS ONE* **9**: e97265.

Wang, Y.T., Gabbard, H.D., and Pai, P.C. (1991) Inhibition of acetate methanogenesis by phenols. *Journal of Environmental Engineering* **117**: 487-500.

Ware, A., and Power, N. (2017) Modelling methane production kinetics of complex poultry slaughterhouse wastes using sigmoidal growth functions. *Renewable Energy* **104**: 50-59.

Watkins, P.A., Moser, A.B., Toomer, C.B., Steinberg, S.J., Moser, H.W., Karaman, M.W. et al. (2010) Identification of differences in human and great ape phytanic acid metabolism that could influence gene expression profiles and physiological functions. *BMC Physiol* **10**: 19.

Weimer, P.J., Russell, J.B., and Muck, R.E. (2009) Lessons from the cow: What the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass. *Bioresour Technol* **100**: 5323-5331.

Weiss, S., Lebuhn, M., Andrade, D., Zankel, A., Cardinale, M., Birner-Gruenberger, R. et al. (2013) Activated zeolite--suitable carriers for microorganisms in anaerobic digestion processes? *Appl Microbiol Biotechnol* **97**: 3225-3238.

Westerholm, M., Leven, L., and Schnurer, A. (2012) Bioaugmentation of syntrophic acetate-oxidizing culture in biogas reactors exposed to increasing levels of ammonia. *Appl Environ Microbiol* **78**: 7619-7625.

Westerholm, M., Müller, B., Isaksson, S., and Schnürer, A. (2015) Trace element and temperature effects on microbial communities and links to biogas digester performance at high ammonia levels. *Biotechnology for Biofuels*.

Westerholm, M., Isaksson, S., Sun, L., and Schnürer, A. (2017) Microbial Community Ability to Adapt to Altered Temperature Conditions Influences Operating Stability in Anaerobic Digestion. In *8th International Conference on Applied Energy, ICAE 2016*. Li, H., Yan, J., Sun, F., Desideri, U., and Chou, S.K. (eds): Elsevier Ltd, pp. 895-900.

Wittmann, C., Zeng, A.P., and Deckwer, W.D. (1995) Growth inhibition by ammonia and use of a pH-controlled feeding strategy for the effective cultivation of Mycobacterium chlorophenolicum. *Appl Microbiol Biotechnol* **44**: 519-525.

Wong, M.T., Zhang, D., Li, J., Hui, R.K.H., Tun, H.M., Brar, M.S. et al. (2013) Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH 10 pretreatment. *Biotechnology for Biofuels* **6**.

Yang, B., Wang, Y., and Qian, P.-Y. (2016) Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* **17**: 135.

Yenigün, O., and Demirel, B. (2013) Ammonia inhibition in anaerobic digestion: A review. *Process Biochem* **48**: 901-911.

Yun, Y.M., Kim, D.H., Cho, S.K., Shin, H.S., Jung, K.W., and Kim, H.W. (2016) Mitigation of ammonia inhibition by internal dilution in high-rate anaerobic digestion of food waste leachate and evidences of microbial community response. *Biotechnol Bioeng* **113**: 1892-1901.

Zhang, B., He, P.-j., Lü, F., Shao, L.-m., and Wang, P. (2007) Extracellular enzyme activities during regulated hydrolysis of high-solid organic wastes. *Water Res* **41**: 4468-4478.

Zhang, C., Xiao, G., Peng, L., Su, H., and Tan, T. (2013) The anaerobic co-digestion of food waste and cattle manure. *Bioresour Technol* **129**: 170-176.

Zhang, Y., and Angelidaki, I. (2015) Submersible microbial desalination cell for simultaneous ammonia recovery and electricity production from anaerobic reactors containing high levels of ammonia. *Bioresour Technol* **177**: 233-239.

Zhu, X., Kougias, P.G., Treu, L., Campanaro, S., and Angelidaki, I. (2017) Microbial community changes in methanogenic granules during the transition from mesophilic to thermophilic conditions. *Appl Microbiol Biotechnol* **101**: 1313-1322.

Zhu, X., Treu, L., Kougias, P.G., Campanaro, S., and Angelidaki, I. (2018) Converting mesophilic upflow sludge blanket (UASB) reactors to thermophilic by applying axenic methanogenic culture bioaugmentation. *Chem Eng J* **332**: 508-516.

Ziels, R.M., Sousa, D.Z., Stensel, H.D., and Beck, D.A.C. (2018) DNA-SIP based genome-centric metagenomics identifies key long-chain fatty acid-degrading populations in anaerobic digesters with different feeding frequencies. *ISME Journal* **12**: 112-123.

Zumstein, E., Moletta, R., and Godon, J.J. (2000) Examination of two years of community dynamics in an anaerobic bioreactor using fluorescence polymerase chain reaction (PCR) single-strand conformation polymorphism analysis. *Environmental Microbiology* **2**: 69-78.

Résumé

La digestion anaérobie (DA) est un processus microbiologique de dégradation de la matière organique qui produit du biogaz. Ce biogaz est riche en méthane qui peut être converti en énergie électrique et thermique. La DA est couramment utilisée pour traiter différents types de déchets organiques à l'échelle industrielle dans des digesteurs anaérobies. Dans un contexte de protection de l'environnement et d'accroissement de l'efficacité énergétique, elle suscite un regain d'intérêt car elle permet de transformer les déchets en ressource énergétique. Cependant, ce bioprocédé n'est pas complètement maîtrisé et présente encore un potentiel d'amélioration important. Cette situation est due principalement à l'absence de méthodes de gestion prenant explicitement en compte le microbiome de la DA. Les microorganismes clés responsables du processus de la DA restent en effet méconnus. Ils forment une communauté microbienne extrêmement complexe, composée principalement de bactéries et d'archées, capable de dégrader la matière organique par de multiples voies. L'une des principales limites de la DA est la grande vulnérabilité de cette communauté microbienne aux modifications des conditions de fonctionnement des digesteurs. Cela peut avoir pour conséquence une production de méthane instable. Ce manuscrit présente ma contribution pour améliorer les connaissances sur le fonctionnement des communautés microbiennes de cette biotechnologie environnementale. Il est divisé en 5 parties. La première partie présente brièvement le sujet et les enjeux de ma recherche, qui fait appel à différentes techniques d'écologie moléculaire et, depuis peu, aux biostatistiques. Trois chapitres résument les principaux travaux réalisés et les résultats obtenus. Le premier chapitre porte sur l'utilisation d'isotopes stables pour tracer la dégradation de substrats spécifiques et identifier les microorganismes fonctionnels dans les digesteurs anaérobies. Le deuxième chapitre présente mes contributions pour une meilleure compréhension de l'effet d'inhibiteurs sur la dynamique et les performances du microbiome de la DA, ainsi que diverses stratégies de gestion permettant d'atténuer l'inhibition. Le troisième chapitre décrit des travaux récents associant des statistiques à des méthodologies à haut débit afin de tirer le meilleur parti des informations générées et de développer des approches *data-driven*. La dernière partie présente les perspectives envisagées pour étendre et améliorer le travail déjà effectué.

Summary

Anaerobic digestion (AD) is a microbiological process of degradation of the organic matter which produces biogas rich in methane that can be converted into valuable electrical and thermal energy. It is commonly used to manage different types of organic waste at industrial scale using anaerobic digesters. In a context of environmental protection and research for increasing energy efficiency, AD arouses a renewed interest because it allows converting waste into an energy resource. However, this bioprocess is not fully mastered and still has an important potential for improvement. Such situation is mainly due to the limitation of microbial-based management of anaerobic reactors as the microbiome, which is the key player of the AD process, still remains largely unknown. Indeed, key microorganisms driving the AD process form an extremely complex microbial community, mainly composed of bacteria and archaea, able to degrade the organic matter across multiple pathways. One of the major limitations of AD is the important susceptibility of the microbial community to changes in operational conditions of the digesters. It can lead to unstable methane formation. This manuscript presents my contribution to enlighten on the functioning of microbial communities of this environmental biotechnology. It is divided in 5 parts. The first part introduces briefly the topic and the issues of my research, which calls on the use of different molecular ecology techniques and, more recently, biostatistics. Three chapters summarize the main works carried out and the results obtained. The first chapter focuses on the use of stable isotopes to describe the degradation of specific substrates and identify functional microorganisms in anaerobic digesters. The second chapter presents my contributions to the better understanding of the effect of different inhibitors on AD microbiome dynamics and performances, as well as management strategies tested to mitigate the inhibition. The third chapter describes recent works combining computational statistics with high-throughput methodologies to make the most of the generated information, and develop data-driven approaches. The last part presents the perspectives envisaged to extend and enhance the work already carried out.