Time-course analysis of metabolomic and microbial responses in anaerobic digesters exposed to ammonia

Olivier Chapleur, Simon Poirier, Angéline Guenne, Kim-Anh Le Cao

Time-course analysis of metabolomic and microbial responses in anaerobic digesters exposed to ammonia

Author names and Affiliations

Olivier Chapleur [a*], Simon Poirier [a,$], Angéline Guenne [a], Kim-Anh Lê Cao [b]

[a] Université Paris-Saclay, INRAE, PRocédés biOtechnologiques au Service de l'Environnement, 92761, Antony, France

[b] Melbourne Integrative Genomics and the School of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria, Australia

$ Present address: IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, Rueil-Malmaison, 92852, France

olivier.chapleur@inrae.fr

simon.poirier@ifpen.fr

angeline.guenne@inrae.fr

kimanh.lecao@unimelb.edu.au

***Corresponding author**

olivier.chapleur@inrae.fr

Université Paris-Saclay, INRAE, PRocédés biOtechnologiques au Service de l'Environnement, 92761, Antony, France

Highlights

- A novel time-course statistical framework was used for omics longitudinal data
- Similar temporal dynamics between microorganisms and metabolites were identified
- Successive growth of guilds of microorganisms across time was revealed
- Microbial dynamics under different ammonia levels were compared
- The influence of ammonia on degradation rate differs between identified metabolites

Abstract

Omics longitudinal studies are effective experimental designs to inform on the stability and dynamics of microbial communities in response to perturbations, but time-course analytical frameworks are required to fully exploit the temporal information acquired in this context. In this study we investigate the influence of ammonia on the stability of anaerobic digestion (AD) microbiome with a new statistical framework. Ammonia can severely reduce AD performance. Understanding how it affects microbial communities development and the degradation progress is a key operational issue to propose more stable processes. Thirty batch digesters were set-up with different levels of ammonia. Microbial community structure and metabolomic profiles were monitored with 16S-metabarcoding and GCMS (gas-chromatography-mass-spectrometry). Digesters were first grouped according to similar degradation performances. Within each group, time profiles of OTUs and metabolites were modelled, then clustered into similar time trajectories, evidencing for example a syntrophic interaction between *Syntrophomonas* and *Methanoculleus* that was maintained up to 387 mg FAN/L. Metabolites resulting from organic matter fermentation, such as dehydroabietic or phytanic acid, decreased with increasing ammonia levels. Our analytical framework enabled to fully account for time variability and integrate this parameter in data analysis.

Keywords

Metabolomics, inhibition, statistics, longitudinal data, 16S

1 Introduction

Microorganisms can be found across all environments on Earth and are parts of multiple microbial ecosystems. In these ecosystems they interact in various ways that can result in an important dynamics of the microbial populations (Braga et al., 2016). The ecosystem global functioning reflects the collective activities of the microorganisms and its stability can be related to the dynamics of the interaction network between microorganisms. If these dynamics were better understood, a better description and control of microbial ecosystems could be achieved.

Sequencing and omics methods in general are becoming classical tools to take snapshots of microbial communities and have been used extensively to describe microbial ecosystems. However, while such snapshots yield a large amount of information regarding the presence or absence of specific microorganisms, functions or metabolites, they are not sufficient to explain why these microorganisms, functions or metabolites are there, how they evolve across time or how they interact within the ecosystem (Ridenhour et al., 2017). As analytical cost is decreasing, more samples can be processed, especially time series, to record the temporal variation of microbial communities. Longitudinal studies can go further than snapshots and inform about the stability and dynamics of the microbiome in response to perturbations or different conditions (Bodein et al., 2019). Thus, they enable to capture more precisely the consequences of disturbances and could accelerate our progress in understandingmicrobial sensitivity. Still, they are not employed so commonly.

The statistical analysis of high-throughput longitudinal studies do not always fully exploit temporal information. Times-series are often analysed as independent samples and experimental design is not always fully taken into account. So far, only a few computational methods have been proposed to examine longitudinal studies with different omics measured under different conditions (Bodein et al., 2019; Ridenhour et al., 2017; Straube et al., 2015). In this study we propose a statistical framework to evidence groups of microorganisms or metabolites exhibiting similar temporal dynamics. We applied this methodology to a case study related to anaerobic digestion (AD) process. Our objective was to

explore the consequence of ammonia on AD based on microbial and metabolomic temporal dynamics.

Anaerobic digestion (AD) is one of the major bioprocesses for converting organic waste into energy. It is commonly used at industrial scale in anaerobic digesters. During this process, large organic molecules are successively broken down into smaller molecules and ultimately into biogas, mainly composed of methane and carbon dioxide. Biogas provides a versatile carrier of renewable energy, as methane can be used for replacement of fossil fuels in both heat and power generation. Allowing waste conversion into energy resource, AD is highly relevant for environmental protection and for our quest to increase energy efficiency. However, it is estimated that only 50% of the potential energy contained in organic waste is currently recovered during AD (Liu et al., 2015). This low energy recovery is related to the poor biodegradability of some waste fractions but also to the presence of several inhibitors in digesters (Amha et al., 2018; Azman et al., 2015; Chen et al., 2008). Many compounds can affect AD microbiome and cause bioreactor instability resulting in low methane yield. In particular, ammonia is considered to be the major toxicant of commercial AD reactors (Rajagopal et al., 2013; Tian et al., 2018). Ammonia inhibition generally occurs in anaerobic digesters treating wastewater or protein-rich waste, such as slaughterhouse wastewater, food waste or manure. In these digesters, ammonia is released throughout the anaerobic degradation of organic nitrogen contained in proteins or urea and is not further degraded.

Ammonia in solution, also called total ammonia nitrogen (TAN), results in dissolved NH_3, the free ammonia nitrogen (FAN), and its ionized form, the ammonium ion (NH_4^+). Equilibrium between NH_3 and NH_4^+ depends on pH and temperature (Anthonisen et al., 1976). FAN is considered as the most toxic, due to its permeability into cell membrane (Wittmann et al., 1995). However, the sensitivity of anaerobic microbiome to ammonia varies by orders of magnitude depending on operating parameters and the composition of the microbial communities (Rajagopal et al., 2013). NH_3 concentrations ranging from 27 to 1450 mg NH_3/L (Capson-Tojo et al., 2020) have been reported to

inhibit AD microbiome. All phases of AD, from hydrolysis to methanogenesis, are influenced by the presence of high ammonia levels. The advancement of molecular tools such as the next-generation sequencing (NGS) technology allows descriptions of microbial dynamics in digester inhibited by ammonia. For example, De Vrieze *et al*. observed that in full-scale digesters, *Firmicutes* and especially *Clostridiaceae* positively correlate with increased ammonia conditions, while *Bacteroidales* seem to be more abundant under lower ammonium concentrations (De Vrieze et al., 2015). Similarly, under ammonia stress, acetoclastic methanogens are usually considered to be vulnerable while the hydrogenotrophs are more resistant (Poirier et al., 2016b). Shift from acetoclastic to hydrogenotrophic methanogenesis is reported frequently (Capson-Tojo et al., 2020). In that case, syntrophic acetate oxidation (SAO) is the predominant acetate-consuming pathway. For example, important growth of syntrophic acetate oxidation bacteria (SAOB), syntrophic partners of hydrogenotrophic methanogens, was already observed (Westerholm et al., 2015). However, contradictory results showed that acetoclastic methanogens could partially resist to the increase of ammonia level (Hao et al., 2015). Although this topic has already been deeply explored, no consensus has been found yet. Differences probably arise from the various environmental conditions applied in the studies, and how AD microbiome is balanced by these conditions. Therefore, a detailed understanding of AD microbiome remains of great importance to facilitate further optimization of this bioprocess (Rajagopal et al., 2013).

To complement microbial dynamics data, metabolomics appears to be an effective approach to study bioprocesses (Vanwonterghem et al., 2014). Untargeted metabolomics consists in the study of all low molecular-weight molecules. These molecules can be involved in cellular metabolic reactions or arise from the organic matter degradation by the microorganisms. Monitoring degradation fate with molecular fingerprints using analytical chemistry methods can provide informative details about key metabolic pathways, in particular when the microbiome is exposed to stress (Beale et al., 2016). These methods include gas or liquid chromatographic techniques coupled to mass spectrometry (GC-MS or LC-MS), and nuclear magnetic resonance (NMR). However, application of metabolomics to

anaerobic digesters remains challenging due to the extreme variability of metabolites with limited *a priori* knowledge (Vanwonterghem et al., 2014). So far, only a few examples of metabolomics application in the AD context have been reported, and they were not related to ammonia inhibition. Beale *et al*. used GC-MS to characterise the impact of operational shocks in lab-scale digesters (Beale et al., 2016). Puig-Castellví *et al.* used LC-MS to assess substrate biodegradability improvement during a co-digestion experiment (Puig-Castellví et al., 2020b). Cardona *et al.* evidenced links between microbial activity and the degradation of metabolites identified with LC-MS (Cardona et al., 2020). Murovec *et al.* used ^{1}H NMR spectroscopic profiling to provide a more comprehensive view of microbial metabolites associated with poor reactor performance in a full-scale mesophilic agricultural biogas plant (Murovec et al., 2018).

In this study we used time series of 16S rRNA gene sequencing data from AD (Poirier & Chapleur, 2018) to investigate the influence of different levels of ammonia, as well as GCMS data newly acquired on the same samples. An original longitudinal analytical framework inspired from (Straube et al., 2015) was used to integrate temporal aspects while comparing the different conditions in a data driven approach.

2 Material and methods

2.1 Lab-scale digesters set-up

Thirty bioreactors (1000 mL) were set-up and inoculated with 20 g of centrifuged methanogenic sludge as inoculum and fed with 50 g of mashed biowaste (corresponding to an initial organic loading of 12 g COD as substrate vs 1.2 g COD as inoculum). Mashed biowaste was provided by an industrial food waste deconditioning unit (Chemaudin, France) (pH=4.1, dry mass = 12.5%, volatile solid = 11.0%, $C_{[wt\% \text{ dry solids}]}$ = 49.5%, $N_{[wt\% \text{ dry solids}]}$ = 3.6%; K^+ = 1.8 mg/g; NH_4^+ = 0.2 mg/g; Na^+ = 2.7 mg/g, Mg^{2+} = 0.2 mg/g; Ca^{2+} = 2.4 mg/g, acetic acid = 3.0 mg/g, propionic acid = 0 mg/g, butyric acid = 0 mg/g, lactic acid = 25 mg/g). Inoculum was sampled from a 50 L laboratory anaerobic bioreactor (pH=7.7,

dry mass = 2.5%, volatile solid = 1.2%, $C_{[wt\% \text{ dry solids}]}$ = 41.7%, $N_{[wt\% \text{ dry solids}]}$ = 2.5%; K^+ = 0.8 mg/g; NH_4^+ = 1.7 mg/g; Na^+ = 6.4 mg/g, acetic acid = 1.1 mg/g, propionic acid = 0.6 mg/g, butyric acid = 0.2 mg/g, lactic acid = 0.1 mg/g). A volume of 430 mL of biochemical methane potential buffer (International Standard ISO 11734 (1995)) was added to reach a total working volume of 500 mL). NH_4Cl (99.998%, Sigma Aldrich) was added in order to reach 10 different TAN concentrations (0.0, 0.5, 1.0, 1.5, 2.5, 5.0, 7.5, 10.0, 25.0 and 50.0 g/L). pH was measured in order to determine FAN concentration (47, 55, 72, 99, 128, 145, 214, 242, 387 and 494 mg/L). All incubations were performed in triplicates. All reactors were incubated without agitation, in the dark, at 35°C. Liquid samples (8 mL) were periodically taken and centrifuged at 10,000 g for 10 min. The pellets and supernatant thus obtained were stored separately at -20°C for analysis of biomass and chemical indicators respectively. In total 9 samples were taken. Digestion tests were run for 160 days, as no biogas production had been observed for one month in the different bioreactors. Detailed experimental procedures are described in (Poirier & Chapleur, 2018; Poirier et al., 2016b).

2.2 Degradation monitoring

Biogas accumulation in the headspace was measured using a differential manometer (Digitron 2082P). Headspace gas analysis was performed using a micro GC (CP4900, Varian) as described in (Chapleur et al., 2016). These data were used to calculate gas production and composition, at standard temperature and pressure, taking into account the extracted volume of liquid samples. Biogas was assimilated to an ideal gas.

Volatile Fatty Acids (VFA) concentrations were measured by ionic chromatography coupled to conductimetric detection, using a Dionex 120 equipped with IonPAc ICE-AS1 column (9 mm x 250 mm). The mobile phases were heptafluorobutyric acid (0.4 mmol/L) and tetrabutylammonium hydroxide (5 mmol/L) as described in (Cardona et al., 2021). Acetate, propionate, butyrate, lactate, formate, valerate and caproate were quantified, but mainly acetate and propionate were detected in the incubations.

TAN concentration was measured in the supernatants using Nessler's colorimetric assay following the French standard (NF T 90-105) and a spectrophotometer (DR-3900, Hach). pH was also measured using a pH meter (IQ160). From those values, FAN was also calculated from the equilibrium of Eq. 1 (Anthonisen et al., 1976) as described in (Cardona et al., 2019).

$FAN = \frac{10^{pH}}{exp\left(\frac{6334}{T}\right)+10^{pH}} \times TAN$ (Eq. 1) where T is the temperature in Kelvin.

Biowaste degradation was monitored with non-targeted gas chromatography mass spectrometry (GCMS) of the liquid phase to produce molecular fingerprints. In brief, after thawing, supernatants were centrifuged for 5 min to remove the precipitate that may have formed during freezing. 1 mL of the obtained liquid was diluted with 1 mL of ultrapure water and acidified with 10 µL of hydrochloric acid (37%). Mixture was loaded onto a 60-mg OASIS® HLB cartridge previously conditioned using 2 mL of methanol and 2 mL of ultrapure water. The cartridge was then washed with 2 mL of ultrapure water. Analytes were eluted in 2 mL of methanol, which was subsequently evaporated under a stream of nitrogen at 40°C. 80 µL of BSTFA was added to the dried extract and mixture was heated in an oven at 60°C for 1h.

GC-MS analysis was performed on a Trace/DSQ II (Thermo Fisher Scientific, Bremen, Germany) equipped with a CombiPAL autosampler and Xcalibur acquisition software. Separation was done by using a ZB-5MS capillary column (60 m X 0.25 mm X 0.25 µm). Oven temperature was maintained at 50°C for 3 minutes, then increased to 250 °C at the rate of 8 °C min^{-1} and held for 10 min. Helium was used as carrier gas with a flow rate of 1.5 ml min^{-1}. 1 µl of sample mixture was injected in splitless mode. All temperatures (i.e. injector, transfer line and source) were set at 280 °C. Electron impact mode (EI) at 70eV was used. Data were acquired using full scan mode from 50 to 650 amu with a solvent delay of 6.5 min. Each sample was analysed three times.

For data processing, scans were averaged using MetAlign. Data were processed using the xcms R package (version 1.52.0) (Smith et al., 2006). The MatchedFilter method was applied to select peaks

on chromatogram with a mzdiff of 1 and a signal-to-noise ratio greater than 3. Similar peaks identified in different samples were grouped using the group method using a bandwidth of 5. Retention time between samples was corrected using the peakgroup method. A second grouping was carried out with the same parameters. Finally, the missing peaks between samples were filled using the fillPeaks method.

To identify the peaks of interest, the mass spectrum of each peak was compared to spectrums from the library of the National Institute of Standards and Technology (NIST, USA). All samples were analysed in triplicate.

2.3 Microbial dynamics monitoring

Total DNA from samples' pellet was extracted using Power Soil DNA Isolation Kit (Mobio Laboratories Inc. Carlsbad) according to the manufacturer's instructions. Extracted DNA was quantified by Qubit (dsDNA HS Assay Kit, Invitrogen, Eugene). Extracted DNA was used for the amplification of the bacterial and archaeal hypervariable region V4-V5 of the 16S rRNA genes with the primers 515F (5'-GTGYCAGCMGCCGCGGTA-3') and 928R (5'-CCCCGYCAATTCMTTTRAGT-3') as described in (Poirier et al., 2016a). Library preparation is described in (Poirier et al., 2016a). Sequencing was performed on Ion Torrent Personal Genome Machine using Ion 316 chip and the Ion PGM Sequencing 400 Kit according to the manufacturer's instructions. For the 48 sequenced samples, total high-quality reads varied between 10,000 and 30,000. The sequencing data have been deposited with links to BioProject accession number PRJNA450311, in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA450311).

FROGS pipeline was used to analyse 16S rRNA tags reads. FROGS (Find Rapidly OTU with Galaxy solution) is a galaxy workflow designed to produce an OTU count matrix from high depth sequencing amplicon data (Escudié et al., 2018). Briefly, after merging the reads comprised between 100 and 500 bp, the software denoised the dataset. It was then clustered with Swarm algorithm. Chimera were

removed with vsearch. The dataset was further filtered by removing singletons. Taxonomic affiliation was performed using Silva128 16S as reference database.

2.4 Data analysis

2.4.1 Clustering of the bioreactors into groups according to the performance data

Based on cumulative methane production, bioreactors were grouped according to the similarity of their performance. This was achieved by first modelling temporal evolution of methane cumulative production for each bioreactor using smoothing splines (lmeSplines R package, smooth.spline() function) (Déjean et al., 2007). For each bioreactor, values of modelled methane production profile were then predicted on a daily time step. K-means clustering (R stats library, kmeans() function) was applied on this data and enabled to define 5 groups of bioreactors showing similar methane production profiles. Selection of the number of groups was based on a visual appreciation of the results (clustering in 4 groups resulted in grouping of bioreactors with different methane profiles, while clustering in 6 groups did not enable to identify another category of reactors). These 5 groups were then compared at the omics level.

2.4.2 Filtering and transformation of the omics data

OTUs with more than 1% of relative abundance in at least one sample were retained for the analysis. To take into account the dispersion in the total number of sequence reads identified in each sample, microbial OTUs abundances were scaled using centered log ratio (CLR) transformation (Lê Cao et al., 2016). The GCMS dataset was filtered based on blanks. Peaks with abundance more than ten times superior to the maximum value in the blanks in at least one sample were kept. For each retention time, only the dominant peak was kept. Data were log transformed.

A fold change filter was then applied on both 16S and metabolomic datasets to remove noisy time profiles and select only informative variables within a given performance group. The fold change of a variable was calculated as the difference of the maximum and minimum mean for each of the time points. A threshold of 3 and 0.5 was applied to filter 16S and metabolomic datasets, respectively.

Within a given performance group, for each time point and each variable, values were averaged across the different reactors. A second filter was then applied to assess the significance of time effect for each profile (16S or metabolite) within a given performance group using a linear mixed model spline framework from the lmmsDE method (lmms R package, (Straube et al., 2015). Variables that were differentially expressed in time ($p<0.050$) were retained.

2.4.3 Omics profiles modelling and clustering

The profile of each variable within each performance group was modelled with Linear mixed model splines (R lmms package, lmmSpline() function). Briefly, each variable profile (16S or metabolite) was modelled with the best fitting LMMS model, as described in (Straube et al., 2015) and (Bodein et al., 2019).

Derivatives of the modelled profiles after discretization were calculated (lmmSpline() function). They contain valuable information about the range of change of expression over time and were particularly relevant in our study. Hierarchical clustering (Euclidean distance, Ward aggregation) was applied on the derivative of the modelled profiles to identify groups of correlated profiles over time and to get insight into the variables that shared similar patterns of time-course trajectories.

3 Results and discussion

3.1 Degradation performance under different levels of ammonia inhibition

Performance indicators (*i.e.* biogas production and volatile fatty acids (VFA) accumulation) and ammonia concentration (TAN and FAN) were monitored for all bioreactors. TAN, FAN and pH values stabilized quickly and remained stable throughout the experiment (supplementary material) indicating that neither important amount of additional ammonia was produced during biowaste degradation nor that ammonia was released in the headspace. However, the presence of ammonia reduced the performances of degradation and in particular the biogas production. As seen by various

267 authors, we did not observe a sharp threshold between no inhibition and total inhibition of the
268 degradation (Capson-Tojo et al., 2020). On the contrary, the different initial levels of ammonia
269 resulted in several patterns of reduced performances. To identify these different patterns, the time-
270 course trajectories of one performance indicator (methane cumulative production) was first
271 modelled in each bioreactor. Bioreactors were then clustered into a few groups with overall similar
272 methane cumulative production profiles.

273 A popular modelling approach for time-course data is smoothing splines (Déjean et al., 2007). It
274 handles varying numbers of time points per bioreactors or different dates of sampling and enables
275 the interpolation of missing values and to denoise the data if necessary. Methane production
276 trajectories modelled with this method were clustered into groups with K-means. Five groups of
277 bioreactors with similar profiles of cumulative methane production were identified as illustrated in
278 Fig. 1. Group 1 corresponds to the non-inhibited bioreactors, with an initial FAN concentration
279 between 47 and 72 mg/L. As FAN concentration increased (group 2 to 4), ultimate methane
280 production declined, as well as production rate. Groups 2, 3 and 4 corresponded to FAN
281 concentrations of respectively 99 to 145 mg/L, 214 to 242 mg/L and 387 mg/L. In group 5 (494 mg/L
282 of FAN) nearly no methane production was observed, and methanogenesis was severely inhibited.

283 Relevance of this clustering was then assessed using three other performance indicators measured in
284 this study (carbon dioxide cumulative production, acetate and propionate accumulation). To do so,
285 the time-course trajectories of these three parameters were first modelled in each bioreactor, then
286 plotted according to the clustering obtained from methane (Fig. 1). Grouping of bioreactors made
287 with cumulative methane production data was consistent for the other indicators illustrating the
288 relevance of this approach. We observed that cumulative carbon dioxide production was also
289 affected by ammonia. From group 1 to group 5 its production declined, as well as production rate.
290 However, nearly 800 mL of carbon dioxide was produced in group 5, suggesting that even if
291 methanogenesis was probably strongly inhibited, bacteria were still active as already observed (Puig-

292 Castellví et al., 2020a). More details on methane and carbon dioxide yield and maximum production
293 rate are presented in the supplementary material. Profiles of acetate and propionate accumulation
294 (see detailed values in supplementary material) were rather similar in groups 1 and 2, showing an
295 important and rapid activity of VFA producing bacteria. In groups 3 to 5, acetate accumulation was
296 slightly delayed but the main difference was evidenced with the consumption rate. In particular,
297 propionate was consumed very late in group 4 while acetate and propionate were never consumed
298 in bioreactors of group 5. These observations suggest that VFA producing bacteria were not totally
299 inhibited even in bioreactors containing 494 mg/L of FAN. However, acetoclastic archaea or bacteria
300 responsible for VFA oxidation were potentially impaired by high ammonia levels. Inhibition of
301 acetoclastic archaea by ammonia was reported and is known to induce a switch towards
302 hydrogenotrophic methanogenesis. It involves syntrophic pathways with acetate or propionate
303 oxidizing bacteria (Hao et al., 2011; Schnürer & Nordberg, 2008). Despite this metabolic possibility,
304 no rapid VFA consumption was observed at high ammonia levels, suggesting that one or both of
305 these populations were inhibited in our experiment, in particular in group 5. From group 2 to group
306 5, propionate consumption was also delayed, probably for similar reasons, namely inhibition of
307 syntrophic propionate oxidizers by ammonia, or because of high level of acetate. In all groups,
308 comparatively to acetate, propionate consumption phase occurred later as already observed
309 (Chapleur et al., 2014). It is classically hypothesised that propionate degradation is less
310 thermodynamically favourable than acetate degradation and occurs preferentially after its
311 consumption (Capson-Tojo et al., 2017). Regardless VFA accumulation, pH remained above 7 in all
312 conditions (supplementary material) suggesting that main inhibitory factor to VFA degradation was
313 ammonia or VFA themselves.

314 Performance measurements confirmed that the reactors could be clustered in five groups showing
315 specific patterns of inhibition despite a more important number of initial ammonia concentrations.
316 Microbial and metabolomics dynamics in these five groups were then explored.

3.2 Effect of ammonia on microbial dynamics

DNA of 48 samples from the different performance groups and different sampling dates (days 0, 9, 29, 42 and 57) were sequenced for 16S identification. Archaea and bacteria were analysed together. In total 1813 OTUs were identified. Samples were divided into different groups corresponding to the bioreactor performance groups. Group 1 was considered as a non-inhibited reference for the study and was examined first. Filters based on abundance in group one (83 OTUs remaining) then fold change and differential expression analysis to identify profiles varying over time within all the reactors of group 1 were performed (30 OTUs left). Smoothing splines to model time course trajectories were then applied on these 30 selected OTUs in the group 1 bioreactors. Derivative of the predicted curves was calculated to capture the rate of change of relative abundance over time. Hierarchical clustering on the derivatives trajectories identified 6 clusters of correlated profiles over time (Fig. 2). This approach enabled to identify similar trajectories in terms of rates and speed of change as well as shapes and magnitude. We hypothesised that microorganisms involved in similar biological processes behaved similarly across time, as we investigated further.

Time-course trajectories of OTUs abundance for the different clusters of group 1 are presented in Fig. 3 for the different groups. Taxonomic affiliation is presented in table 1. In group 1, abundance of OTUs from cluster 1, 2, 5 and 6 increased across time while abundance of OTUs from clusters 3 and 4 decreased.

Abundance of OTUs from cluster 1 increased sharply immediately after the start of the experiment. They belong to *Ruminococcaceae* family, *Ruminococcus* genus and *Porphyromonadaceae* family. These taxa have been acknowledged to play an important role in degrading complex carbohydrates (La Reau & Suen, 2018) or carbohydrates and proteins (Krieg et al., 2010) and catalysing the production of VFAs, ethanol and carbon dioxide. They were thus likely involved in the first steps of the degradation.

Similarly, abundance of OTUs from cluster 2 increased rapidly. These very diverse OTUs belonged to *Porphyromonadaceae*, *Bacteroidaceae*, *Lachnospiraceae, Acholeplasmataceae*, *Clostridiaceae* 1 and *Methanomicrobiaceae* families. Similarly to *Porphyromonadaceae*, *Bacteroidaceae* and genus *Bacteroides* can degrade carbohydrates and some proteins (Krieg et al., 2010). *Lachnospiraceae* and particularly genus *Mobilitalea* has been suggested to play important roles in fermentation of mono-, di- and polysaccharides, including microcrystalline cellulose (Podosokorskaya et al., 2014). *Acholeplasmataceae* and genus *Acholeplasma* are glucose and simple sugars fermenters and produce acids (Krieg et al., 2010). Species *Clostridium butyricum* from *Clostridiaceae 1* family can also consume sugars and proteins (Vos et al., 2009) while producing different VFA including butyrate. Consequently, all these families were probably involved in the early steps of biowaste degradation and their relative abundance increased until there was no more biowaste available in the bioreactor. *Methanoculleus* genus, from *Methanomicrobiaceae* is an hydrogenotrophic methanogen and probably consumed carbon dioxide and hydrogen resulting from the fermentation which explains that its growth was correlated to that of the other microorganisms of group 2.

On the contrary, the abundance of the 5 OTUs from cluster 3 decreased immediately. Among them, 3 OTUs of *Lactobacillus* genus likely originated from the biowaste (Probst et al., 2013) and could not grow in the anaerobic digesters. One OTU was affiliated to *Pseudomonas caeni* sp. that was isolated from the sludge of an anaerobic ammonium-oxidizing bioreactor (Xiao et al., 2009). One OTU belonged to *Paenalcaligenes* genus isolated in different types of guts (Lee et al., 2013). The last two OTUs were probably abundant in the inoculum. However, none of them survived in the conditions of our bioreactors.

Abundance of OTUs of cluster 4 also decreased, but less rapidly than cluster 3. These OTUs belonged to various taxa commonly found in anaerobic digesters, such as *Bacteroidales*, *Clostridiales*, *Methanobacteriales* and *Thermoplasmatales* orders (Madigou et al., 2019). We hypothesized that even if the environmental conditions of the digester was not unfavourable to their growth (contrary

366 to OTUs of cluster 6), they were progressively outcompeted by microorganisms from cluster 1, 2, 5
367 and 6 and could not settle in our digesters.

368 Cluster 5 included 4 OTUs with abundance increasing between day 9 and 29, but not before day 9.
369 They belonged to *Syntrophomonadaceae* (2 OTUs) and *Peptococcaceae* families and to
370 *Armatimonadetes* phylum. Members of the *Syntrophomonas* genus, identified in this system, are
371 fatty acids (C4–C18) degraders (Narihiro et al., 2016; Zou et al., 2003). Specifically, *Syntrophomonas*
372 *wolfei* mainly uses straight-chain fatty acids containing 4 to 8 carbon atoms as energy sources (Vos et
373 al., 2009). Members of the family *Peptococcaceae* are known as propionate oxidizers (Cardona et al.,
374 2021). In particular *Pelotomaculum* genus has been described as a syntrophic microorganisms that
375 can have different archaea as partners, including *Methanoculleus* identified in cluster 2 (Chen et al.,
376 2020). Its growth was in accordance with the degradation pattern of the propionate.
377 *Armatimonadetes* phylum is one of the most recently recognised bacterial phylum. It has already
378 been identified in anaerobic digesters (Campanaro et al., 2020; Puig-Castellví et al., 2020c) but its
379 role remains unclear. These microorganisms are VFA degraders thus explaining that their abundance
380 only increased once biowaste degradation already results in the production of large amounts of fatty
381 acids as intermediates as observed in figure 1.

382 Abundance of OTUs of cluster 6 increased progressively with time. They belonged to *Bacteroidales*
383 order (2 OTUs) and *Syntrophomonadaceae* and *Peptococcaceae* families. In particular OTU from
384 genus *Alkaliflexus* and *Marinilabiaceae* family could contribute to hydrolysis by secreting a
385 cellulolytic enzyme for the degradation of cellulose (Gao et al., 2014). OTU from *Peptococcaceae*
386 family is from *Cryptoanaerobacter* genus, which has been characterized in a methanogenic
387 consortium derived from a waste mixture. Phenol or 4-hydroxybenzoate (4-OHB) could be required
388 for the growth of this genus (Juteau et al., 2005). They may have been involved in the degradation of
389 less accessible waste or degradation products.

Our clustering approach identified several groups of microorganisms with similar time-course trajectories. These different groups were relevant, as they included microorganisms involved in the successive steps of the degradation. Based on this non-inhibited reference, trends relative to the time-course response of the same clusters of OTUs in group 2 to 5 were examined (Fig. 3). As ammonia concentration increased, we observed delays in the trajectories of the time-course profiles.

Although different inhibitions thresholds could be evidenced for each OTUs, similar trends were observed for clusters 1 and 2, which both consisted of populations emerging from the inoculum. OTUs rumi_2 from *Ruminococcus* genus and bact_4 from genus *Bacteroides* were inhibited as soon as FAN initial concentration reached 99 mg/l (group 2). Meanwhile, OTUs lach_1 (*Mobilitalea* genus), porp_6 (*Porphyromonadaceae* family) and clos1_1 (*Clostridium butyricum* species) relative abundance still increased in group 2 suggesting that they could survive despite FAN concentrations up to 145 mg/L. The porp_8 (*Porphyromonadaceae* family) and acho_1 (*Acheoplasma* genus) OTUs appeared more resistant to ammonia levels up to 242 mg/L (group 3). Interestingly, *Methanoculleus* was the only OTU from both clusters to resist and thus steer methanogenesis up to 387mg/L of FAN (group 4).

Clusters 3 and 4, which consisted of dominant populations respectively present in the biowaste and in the inoculum also demonstrated comparable behaviours between groups. OTUs from the biowaste (cluster 3) which were not adapted to the environment and thus rapidly disappeared in groups 1, 2, 3 seemed to benefit from ammonia inhibition in group 4 (387 mg FAN/L). Since active populations did not emerge at this FAN level, they could remain dominant in the ecosystem. Similarly, the decrease of OTUs present in the inoculum (cluster 4) was more progressive along increasing FAN levels. Some opportunist populations such as FX1_1 (*Anaerosalibacter* genus) probably resistant to ammonia even managed to grow in group 4.

Within cluster 5, different inhibition thresholds could be noticed. Syntrophic populations such as synt_2 and synt_4 could resist up to group 4 and thus steered the interaction with *Methanoculleus*.

Syntrophomonas was already shown to be resistant to the increase of ammonia in CSTR experiment (Bonk et al., 2018). On the other hand, arma_1 (*Armatimonadetes* phylum) and pept_5 (*Pelotomaculum* genus) were respectively inhibited since group 3 (214 mg FAN/L) and group 4 (387 mg FAN/L). Syntrophic OTU from cluster 6 (synt_6, *Syntrophomonas* genus) showed the same trend, suggesting that three populations could interact with *Methanoculleus* at 387 mg FAN/L. Nevertheless, hydrolytic OTUs from cluster 6 such as pept6 (*Cryptanaerobacter* genus) and GZK_1 (*Bacteroidales* order) were inhibited in group 4.

As we considered the bioreactors in group 1 as a reference, the filtering, selection and clustering of OTUs was based only their trajectories and abundance in group 1. Thus, some OTUs with significant time effects in groups 2, 3, 4 or 5 may have been omitted in the selection. We evaluated this potential bias by considering each group as a reference group and compared the OTUs selected. Details are provided in the supplementary material. To summarise, another 14 OTUs would have been selected in groups 2, 3, 4 or 5. We observed two categories of OTUs: 7 OTUs decreased over. They were present in the inoculum but were not adapted to our experimental set-up and were replaced more or less rapidly by other microorganisms. As their fold change was small, they were not selected when considering group 1 as a reference. The abundance of the remaining 7 OTUs increased more sharply over time in groups 2 to 5 compared to group 1. These microorganisms were probably resistant to ammonia and benefited from the inhibition of other microorganisms sharing the same ecological niche. Among them, 4 can be highlighted. The abundance of an OTU from *Methanosarcina* genus increased with ammonia level (except in group 5). It is in agreement with previous studies results which revealed that *Methanosarcina* was able to resist to high TAN concentrations (Hao et al., 2015). Similarly, the abundance of an OTU from *Synergistaceae* family, *Aminobacterium* genus increased in digester with low or medium ammonia level. It is an amino acid degrader (He et al., 2017) and for this reason is probably moderately sensitive to ammonia. OTUs from this family have also been described as acetate oxidizing bacteria and their growth in presence of ammonia has already been observed (Puig-Castellví et al., 2020a). The abundance of an OTU from *Proteiniphilum*

441 genus increased in groups 4 and 5. Growth of this OTU is promoted by the thermal hydrolysis of
442 proteins, which could be a source of ammonia (Chen et al., 2019). It can also adapt to the presence of
443 ammonia (Puig-Castellví et al., 2020a). Finally, abundance of an OTU from *Ruminiclostridium* genus
444 increased in group 3. Increase of the abundance of this genus in presence of ammonia has already
445 been observed (Fernandez-Gonzalez et al., 2019).

446 To summarize, the selection of OTUs with respect to the reference group 1 selected the most
447 dynamic OTUs. While this approach facilitated the comparison with the different groups, an OTU
448 selection bias was observed and similar procedure should be applied to other groups considered as
449 reference.

450 3.3 Effect of ammonia on biowaste degradation

451 The evolution of organic matter in the digesters was characterised with GCMS. In total 44 samples
452 from the different performance groups and sampling dates (days 9, 29, 42 and 57) were analysed in
453 triplicates. Only metabolites with an abundance exceeding 10^5 were taken into account (92 ions).
454 Samples were divided into different groups corresponding to the bioreactor performance group. In
455 order to remove non-informative molecules, a fold change filter was applied within group 1 sub-
456 dataset, as well as a differential expression filter. A total 20 ions were selected for time-course
457 analysis. The time-course trajectories of the selected metabolites was modelled as described for
458 OTUs. Derivative trajectories were calculated and clustered with hierarchical clustering (Euclidean
459 distance and Ward method). Five clusters of ions were identified (Fig. 2). Time-course trajectories of
460 metabolites abundance for the different clusters from group 1 are presented in Fig. 4. Identification
461 of the different metabolites is presented in Table 2. In group 1, the abundance of ions from clusters 1
462 and 2 increased over time while the abundance of ions from clusters 3 to 5 decreased (Fig. 4).

463 In cluster 1, the abundance of the identified metabolites slowly increased. These metabolites were
464 likely produced by microorganisms and accumulated during the anaerobic digestion process,
465 suggesting they could not be metabolised by other microorganisms. Only anthranilic acid, N-

466 acetylanthranilic acid and dehydroabietic acid were identified. In organisms capable of tryptophan
467 synthesis, anthranilic acid is a precursor to the amino acid tryptophan. It is also a metabolite of
468 pigments and dyes anaerobic breakdown (Razo-Flores et al., 1999). N-acetylanthranilic acid can have
469 similar origins and is metabolised in anthranilic acid. Dehydroabietic acid is found in different plants
470 and particularly conifers and used in different industrial applications (Jia et al., 2019). These
471 metabolites were released during the breakdown or hydrolysis of organic matter by the
472 microorganisms. In general, as the level of inhibition increased, their accumulation rate decreased,
473 suggesting that ammonia slowed down hydrolysis of organic matter, even if it did not stop totally.
474 Microorganisms from OTUs cluster 6 could be involved in their production since they are hydrolytic
475 microorganisms potentially involved in the degradation of recalcitrant organic matter and behaved
476 similarly toward ammonia.
477 Metabolites identified in cluster 2 exhibited a higher increase in abundance over time compared to
478 cluster 1. Three metabolites were identified. Benzoic acid is formed during the degradation of
479 phenolic compounds (Hoyos-Hernandez et al., 2014) but also from lignin (Zhu et al., 2017). Phytanic
480 acid is known to be produced during the fermentation of plant materials in the ruminant gut
481 (Watkins et al., 2010), as well as indole-2-carboxylic acid. It can be hypothesized that they originated
482 from the slow degradation of complex organic polymers from the biowaste as they progressively
483 accumulated in the digesters. Their abundance increased more rapidly than that of metabolites from
484 cluster 1 suggesting that they were more abundant or could be extracted more easily from organic
485 matter. As observed for metabolites from cluster 1, presence of ammonia reduced their
486 accumulation rate. They could also be related to microorganisms from OTUs clusters 6, or 1.
487 The abundance of metabolites from cluster 3 decreased rapidly after the start of the incubation. Two
488 molecules were identified. Decanoic acid is found in different type of oils and also in the milk of
489 various mammals. 3-(3-Hydroxyphenyl)propionic acid is one of the major metabolites of ingested
490 caffeic acid and of the phenolic degradation products of proanthocyanidins (the most abundant
491 polyphenol present in chocolate) by the microflora in the colon. (Konishi & Kobayashi, 2004; Rios et

al., 2003). It can also arise from digestion of aromatic amino-acids or breakdown product of lignin or other plant-derived phenylpropanoids (Torres et al., 2003). These molecules we already abundant at the beginning of the experiment and were metabolised relatively easily by the microbial community. Ammonia had a moderate effect on their degradation, till group 3, but their degradation was stopped in group 4 and 5. Fermentative microorganisms from cluster 2 could be responsible for the degradation of these compounds.

The abundance of the metabolites identified in cluster 4 also progressively decreased over time as they were consumed. Among them we identified 3,4-dihydroxyhydrocinnamic acid also commonly found in plant biomass and its residues (Boerjan et al., 2003) and a fatty acid (tetradecanoic acid) that can be found in a great variety of oils and fats from vegetal and animal origins. They were degraded less efficiently than metabolites from cluster 3, and influence of ammonia seemed more important as degradation of one metabolite (not identified) already stopped in group 2.

The abundance of metabolites from cluster 5 remained stable for a long time, but their degradation started after 26 days. In this cluster, only hydrocinnamic acid was identified. As 3,4-dihydroxyhydrocinnamic acid, this molecule is commonly found in plant biomass and its residues (Boerjan et al., 2003). Its molecular structure may have contributed to their slower degradation compared to other molecules. As ammonia level increased, the degradation of one of the molecules (not identified) was rapidly slowed down, while hydrocinnamic acid was degraded similarly in group 1 to 3. However, in groups 4 and 5, it was not degraded. Metabolites from clusters 4 and 5 could have been degraded by microorganisms from cluster 1, 2 or 6, but direct link cannot be established.

Similar the analysis conducted for the OTUs, here we also considered the bioreactor group 1 as a reference. By repeating the analysis for other groups as reference, we identified additional metabolites, as fully described in the supplementary material.

To summarise, GCMS enabled to visualize the time-course response of various molecules throughout AD and under different ammonia stresses. Mainly complex molecules from biomass and plant degradation were identified. However, this snapshot is not exhaustive. Several parameters should be

518 considered when applying metabolomics to AD studies. Depending on the technique used (GC-MS,
519 LC-MS, NMR…) and experimental preparation, specific categories of molecules can be targeted and
520 observed.
521 Additionally, the sampling frequency selected in our study mainly enabled to identify molecules
522 slowly degraded or produced and persisting in the batch bioreactors. Biowaste also contains multiple
523 easily degradable metabolites that were probably present only in the first days of the experiment.
524 Similarly, some metabolites were probably degraded very rapidly after their production and could
525 thus not be observed or filtered by our analytical pipeline as they appeared transiently.
526 Finally, even though metabolomics is more and more widely used, it has not been applied extensively
527 to AD yet. Thus, the identification of AD molecules remains limited to their description in databases
528 and to the few studies available, leading to possible bias in data interpretation.

529 ### 3.4 Perspectives for the analysis of time course series of omics
530 samples in bioprocesses

531 Microbiome studies have been limited to small sample sizes due to the high cost and complexity of
532 experimental. Longitudinal replicated studies are now possible and new analytical designs can be
533 implemented. In particular, they enable studies of temporal dynamics of AD microbiome coupled to
534 experimental interventions. These experimental designs are essential to move beyond descriptive
535 associations and attempt to decipher causal mechanisms. It will ultimately enable to rationalize
536 approaches to manipulate AD microbiome and achieve durable benefits. However, data analysis from
537 such studies has been hampered by a lack of appropriate computational tools.

538 In this work, we developed an analytical framework to integrate temporal dynamics of the microbial
539 ecosystem while exploiting time information. The comparison of time-course profiles enables the
540 identification of co-evolving biological features that are potentially related. It also allows to identify
541 succession of events that occur in the digesters, and to draw hypotheses on the chain of events
542 occurring in AD. As such, in depth-understanding of microbial and degradation dynamics can be
543 achieved.

Compared to the few frameworks proposed for omics longitudinal data, our analytical approach presents multiple advantages. Firstly we focus on relevant and reproducible patterns through the filtering of relevant biological variables. Smoothing splines denoises experimental values and manages replicate variability as the spline aggregates signal obtained from different replicates. Moreover, temporally adjacent samples can compensate for errors. Splines enable the interpolation of missing values (for example missing date or missing variables) or varying number of time points per bioreactor, and different number of replicates per time point. Finally, the clustering reduces dataset complexity by grouping variables into a small number of time-course trajectory types (clusters) that are likely to be biologically related. It can thus address the high dimensionality of omics longitudinal data and associated constraints.

Additionally, our pipeline is a very generalist approach, widely applicable, if adapted to the data type. For example, a relevant clustering approach must take into account the data specificity and question targeted. For performance data, we used k-means on modelled data, to group the bioreactors based on both the shape and magnitude of methane production. For omics data we applied hierarchical clustering to the derivative of the modelled profiles to account for the rates and speeds of changes in the omics profiles. The absolute level of expression of the molecules or microorganisms was of limited interest as it can be biased due to the experimental procedure (PCR amplification during 16S metabarcoding – different ionization capability during mass spectrometry metabolomics analysis). Instead, curves shapes (accessed through the derivative) could provide meaningful information on coordinate growth, production or consumption.

Despite these various advantages, a limitation of our pipeline is that it does not enable simultaneous integration of various data type. For example, we analysed degradation performances, 16S and metabolomics data independently and did not seek for correlation links between performances microorganisms and molecules. Correlating several types of data is a challenge as these data are not generated by the same techniques and do not reflect the same biological phenomena. However,

several studies attempted to evidence such correlations with different data integration approaches (Bodein et al., 2019; Cardona et al., 2020). Further development should be undertaken to address these questions.

For the field of AD and the general field of bioprocess, longitudinal profiling technologies have a great potential. They can uncover complex relations between variations across microbiome variables (Jendoubi & Ebbels, 2020), and enlightening complex microbiota interactions in AD. The clustering of longitudinal profiles helps identifying groups of biological entities that may be functionally related generate novel hypotheses about the interaction mechanisms that take place within the processes. Similar framework can be applied to multiple other omics, allowing an in-depth characterization of ecological niches over time. It can provide enhanced understanding of the underlying biology of the system that will help the design of optimized processes.

Another benefit of such approaches is that they naturally yield a list of perturbed microbial dynamics or metabolic pathways, either by comparison of conditions, as illustrated in this study, or by looking at long term dynamics in continuous processes submitted to disturbance (Herold et al., 2020). It can thus foster the discovery of early warning bioindicators, namely microorganisms or metabolites that have significant association patterns with a particular situation, such as the presence of an inhibitor, overloading, dysfunction, change of substrate. For a diagnosis perspective, these bioindicators can anticipate failure (Poirier et al., 2020). This knowledge can be used to increase the robustness of the process and limit dysfunction.

4 Conclusion

Our study provides a novel time-course perspective of the effect of ammonia on AD at the performance, microbial and metabolic levels. Our analysis framework accounted for variability between reactors of the same group of performance. Successive growth of guilds of microorganisms across time was revealed as well as influence of ammonia on the production and degradation of some metabolites. With increasing availability of high-throughput sequencing time-course data, this

analytical framework is a powerful tool for characterizing complex datasets and revealing novel insights into microbiome related issues.

5 Funding sources

Kim Anh Lê Cao and Olivier Chapleur scientific travels were supported in part by the France-Australia Science Innovation Collaboration (FASIC) Program Early Career Fellowships from Campus France and the Australian Academy of Science. KALC was supported in part by the National Health and Medical Research Council (NHMRC) Career Development fellowship (GNT1159458). This work was conducted as part of the STABILICS project supported by the National Research Agency (ANR-19-CE43-0003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

6 Data Availability

The sequencing data have been deposited with links to BioProject accession number PRJNA450311, in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA450311). GCMS raw spectra are available on request.

7 References

Amha, Y.M., Anwar, M.Z., Brower, A., Jacobsen, C.S., Stadler, L.B., Webster, T.M., Smith, A.L. 2018. Inhibition of anaerobic digestion processes: Applications of molecular tools. *Bioresource Technology*, **247**, 999-1014.

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

Azman, S., Khadem, A.F., van Lier, J.B., Zeeman, G., Plugge, C.M. 2015. Presence and Role of Anaerobic Hydrolytic Microbes in Conversion of Lignocellulosic Biomass for Biogas Production. *Critical Reviews in Environmental Science and Technology*, **45**(23), 2523-2564.

Beale, D.J., Karpe, A.V., McLeod, J.D., Gondalia, S.V., Muster, T.H., Othman, M.Z., Palombo, E.A., Joshi, D. 2016. An 'omics' approach towards the characterisation of laboratory scale anaerobic digesters treating municipal sewage sludge. *Water Res*, **88**, 346-357.

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

Boerjan, W., Ralph, J., Baucher, M. 2003. Lignin biosynthesis. *Annu Rev Plant Biol*, **54**, 519-46.

Braga, R.M., Dourado, M.N., Araújo, W.L. 2016. Microbial interactions: ecology in a molecular perspective. *Brazilian Journal of Microbiology*, **47**, 86-98.

Campanaro, S., Treu, L., Rodriguez-R, L.M., Kovalovszki, A., Ziels, R.M., Maus, I., Zhu, X., Kougias, P.G., Basile, A., Luo, G., Schlüter, A., Konstantinidis, K.T., Angelidaki, I. 2020. New insights from the biogas microbiome by comprehensive genome-resolved metagenomics of nearly 1600 species originating from multiple anaerobic digesters. *Biotechnology for Biofuels*, **13**(1), 25.

Capson-Tojo, G., Moscoviz, R., Astals, S., Robles, Á., Steyer, J.P. 2020. Unraveling the literature chaos around free ammonia inhibition in anaerobic digestion. *Renewable and Sustainable Energy Reviews*, **117**, 109487.

Capson-Tojo, G., Ruiz, D., Rouez, M., Crest, M., Steyer, J.-P., Bernet, N., Delgenès, J.-P., Escudié, R. 2017. Accumulation of propionic acid during consecutive batch anaerobic digestion of commercial food waste. *Bioresource Technology*, **245**, 724-733.

Cardona, L., Cao, K.A.L., Puig-Castellví, F., Bureau, C., Madigou, C., Mazéas, L., Chapleur, O. 2020. Integrative Analyses to Investigate the Link between Microbial Activity and Metabolite Degradation during Anaerobic Digestion. *Journal of Proteome Research*.

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

Cardona, L., Mazéas, L., Chapleur, O. 2021. Zeolite favours propionate syntrophic degradation during anaerobic digestion of food waste under low ammonia stress. *Chemosphere*, **262**, 127932.

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 Microbiology Ecology*, **87**(3), 616-629.

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

Chen, S., Dong, B., Dai, X., Wang, H., Li, N., Yang, D. 2019. Effects of thermal hydrolysis on the metabolism of amino acids in sewage sludge in anaerobic digestion. *Waste Management*, **88**, 309-318.

Chen, Y., Cheng, J.J., Creamer, K.S. 2008. Inhibition of anaerobic digestion process: A review. *Bioresource Technology*, **99**(10), 4044-4064.

Chen, Y.T., Zeng, Y., Wang, H.Z., Zheng, D., Kamagata, Y., Narihiro, T., Nobu, M.K., Tang, Y.Q. 2020. Different Interspecies Electron Transfer Patterns during Mesophilic and Thermophilic Syntrophic Propionate Degradation in Chemostats. *Microbial Ecology*, **80**(1), 120-132.

De Vrieze, J., Saunders, A.M., He, Y., Fang, J., Nielsen, P.H., Verstraete, W., Boon, N. 2015. Ammonia and temperature determine potential clustering in the anaerobic digestion microbiome. *Water Research*, **75**, 312-323.

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

Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S., Hernandez-Raquet, G., Combes, S., Pascal, G. 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics*, **34**(8), 1287-1294.

Fernandez-Gonzalez, N., Pedizzi, C., Lema, J.M., Carballa, M. 2019. Air-side ammonia stripping coupled to anaerobic digestion indirectly impacts anaerobic microbiome. *Microbial Biotechnology*, **12**(6), 1403-1416.

Gao, Z.M., Xu, X., Ruan, L.W. 2014. Enrichment and characterization of an anaerobic cellulolytic microbial consortium SQD-1.1 from mangrove soil. *Applied Microbiology and Biotechnology*, **98**(1), 465-474.

Hao, L., Lü, F., Mazéas, L., Desmond-Le Quéméner, E., Madigou, C., Guenne, A., Shao, L., Bouchez, T., He, P. 2015. Stable isotope probing of acetate fed anaerobic batch incubations shows a partial resistance of acetoclastic methanogenesis catalyzed by Methanosarcina to sudden increase of ammonia level. *Water Research*, **69**, 90-99.

Hao, L.P., Lü, F., He, P.J., Li, L., Shao, L.M. 2011. Predominant contribution of syntrophic acetate oxidation to thermophilic methane formation at high acetate concentrations. *Environmental Science and Technology*, **45**(2), 508-513.

He, Q., Li, L., Zhao, X., Qu, L., Wu, D., Peng, X. 2017. Investigation of foaming causes in three mesophilic food waste digesters: Reactor performance and microbial analysis. *Scientific Reports*, **7**(1).

Herold, M., Martínez Arbas, S., Narayanasamy, S., Sheik, A.R., Kleine-Borgmann, L.A.K., Lebrun, L.A., Kunath, B.J., Roume, H., Bessarab, I., Williams, R.B.H., Gillece, J.D., Schupp, J.M., Keim, P.S., Jäger, C., Hoopmann, M.R., Moritz, R.L., Ye, Y., Li, S., Tang, H., Heintz-Buschart, A., May, P., Muller, E.E.L., Laczny, C.C., Wilmes, P. 2020. Integration of time-series meta-omics data reveals how microbial ecosystems respond to disturbance. *Nature Communications*, **11**(1), 5281.

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

Jendoubi, T., Ebbels, T.M.D. 2020. Integrative analysis of time course metabolic data and biomarker discovery. *BMC Bioinformatics*, **21**(1).

Jia, P., Ma, Y., Feng, G., Hu, L., Zhou, Y. 2019. High-value utilization of forest resources: Dehydroabietic acid as a chemical platform for producing non-toxic and environment-friendly polymer materials. *Journal of Cleaner Production*, **227**, 662-674.

Juteau, P., Côté, V., Duckett, M.F., Beaudet, R., Lépine, F., Villemur, R., 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**(Pt 1), 245-250.

Konishi, Y., Kobayashi, S. 2004. Microbial metabolites of ingested caffeic acid are absorbed by the monocarboxylic acid transporter (MCT) in intestinal Caco-2 cell monolayers. *J Agric Food Chem*, **52**(21), 6418-24.

Krieg, N.R., Ludwig, W., Whitman, W.B., Hedlund, B.P., Paster, B.J., Staley, J.T., Ward, N., Brown, D. 2010. Bergey's Manual of Systematic Bacteriology, 2nd ed., **4**.

La Reau, A.J., Suen, G. 2018. The Ruminococci: key symbionts of the gut ecosystem. *J Microbiol*, **56**(3), 199-208.

Lê Cao, K.A., Costello, M.E., Lakis, V.A., Bartolo, F., Chua, X.Y., Brazeilles, R., Rondeau, P. 2016. MixMC: A multivariate statistical framework to gain insight into microbial communities. *PLoS ONE*, **11**(8).

Lee, Y.Y., Lee, J.K., Park, K.H., Kim, S.Y., Roh, S.W., Lee, S.B., Choi, Y., Lee, S.J. 2013. Paenalcaligenes hermetiae sp. nov., isolated from the larval gut of Hermetia illucens (Diptera: Stratiomyidae), and emended description of the genus Paenalcaligenes. *Int J Syst Evol Microbiol*, **63**(Pt 11), 4224-9.

Liu, X., Bayard, R., Benbelkacem, H., Buffière, P., Gourdon, R. 2015. Evaluation of the correlations between biodegradability of lignocellulosic feedstocks in anaerobic digestion process and their biochemical characteristics. *Biomass and Bioenergy*, **81**, 534-543.

Madigou, C., Lê Cao, K.-A., Bureau, C., Mazéas, L., Déjean, S., Chapleur, O. 2019. Ecological consequences of abrupt temperature changes in anaerobic digesters. *Chemical Engineering Journal*, **361**, 266-277.

Murovec, B., Makuc, D., Kolbl Repinc, S., Prevoršek, Z., Zavec, D., Šket, R., Pečnik, K., Plavec, J., Stres, B. 2018. (1)H NMR metabolomics of microbial metabolites in the four MW agricultural biogas plant reactors: A case study of inhibition mirroring the acute rumen acidosis symptoms. *J Environ Manage*, **222**, 428-435.

Narihiro, T., Nobu, M.K., Tamaki, H., Kamagata, Y., Liu, W.-T. 2016. Draft Genome Sequence of Syntrophomonas wolfei subsp. methylbutyratica Strain 4J5T (JCM 14075), a Mesophilic Butyrate- and 2-Methylbutyrate-Degrading Syntroph. *Genome announcements*, **4**(2), e00047-16.

727 Podosokorskaya, O.A., Bonch-Osmolovskaya, E.A., Beskorovaynyy, A.V., Toshchakov, S.V., Kolganova,
728 T.V., Kublanov, I.V. 2014. Mobilitalea sibirica gen. nov., sp. nov., a halotolerant
729 polysaccharide-degrading bacterium. *International Journal of Systematic and Evolutionary*
730 *Microbiology*, **64**(PART 8), 2657-2661.
731 Poirier, S., Bize, A., Bureau, C., Bouchez, T., Chapleur, O. 2016a. Community shifts within anaerobic
732 digestion microbiota facing phenol inhibition: Towards early warning microbial indicators?
733 *Water Research*, **100**, 296-305.
734 Poirier, S., Chapleur, O. 2018. Inhibition of anaerobic digestion by phenol and ammonia: Effect on
735 degradation performances and microbial dynamics. *Data in Brief*, 2235-2239.
736 Poirier, S., Déjean, S., Midoux, C., Lê Cao, K.-A., Chapleur, O. 2020. Integrating independent microbial
737 studies to build predictive models of anaerobic digestion inhibition by ammonia and phenol.
738 *Bioresource Technology*, **316**, 123952.
739 Poirier, S., Desmond-Le Quéméner, E., Madigou, C., Bouchez, T., Chapleur, O. 2016b. Anaerobic
740 digestion of biowaste under extreme ammonia concentration: Identification of key microbial
741 phylotypes. *Bioresource Technology*, **207**, 92-101.
742 Probst, M., Fritschi, A., Wagner, A., Insam, H. 2013. Biowaste: a Lactobacillus habitat and lactic acid
743 fermentation substrate. *Bioresour Technol*, **143**, 647-52.
744 Puig-Castellví, F., Cardona, L., Bureau, C., Bouveresse, D.J.-R., Cordella, C.B.Y., Mazéas, L., Rutledge,
745 D.N., Chapleur, O. 2020a. Effect of ammonia exposure and acclimation on the performance
746 and the microbiome of anaerobic digestion. *Bioresource Technology Reports*, **11**, 100488.
747 Puig-Castellví, F., Cardona, L., Jouan-Rimbaud Bouveresse, D., Cordella, C.B.Y., Mazéas, L., Rutledge,
748 D.N., Chapleur, O. 2020b. Assessment of substrate biodegradability improvement in
749 anaerobic Co-digestion using a chemometrics-based metabolomic approach. *Chemosphere*,
750 **254**, 126812.
751 Puig-Castellví, F., Cardona, L., Jouan-Rimbaud Bouveresse, D., Cordella, C.B.Y., Mazéas, L., Rutledge,
752 D.N., Chapleur, O. 2020c. Assessment of the microbial interplay during anaerobic co-
753 digestion of wastewater sludge using common components analysis. *PLOS ONE*, **15**(5),
754 e0232324.
755 Rajagopal, R., Massé, D.I., Singh, G. 2013. A critical review on inhibition of anaerobic digestion
756 process by excess ammonia. *Bioresource Technology*, **143**, 632-641.
757 Razo-Flores, E., Smulders, P., Prenafeta-Boldú, F., Lettinga, G., Field, J.A. 1999. Treatment of
758 anthranilic acid in an anaerobic expanded granular sludge bed reactor at low concentrations.
759 *Water Science and Technology*, **40**(8), 187-194.
760 Ridenhour, B.J., Brooker, S.L., Williams, J.E., Van Leuven, J.T., Miller, A.W., Dearing, M.D., Remien,
761 C.H. 2017. Modeling time-series data from microbial communities. *ISME J*, **11**(11), 2526-
762 2537.
763 Rios, L.Y., Gonthier, M.P., Rémésy, C., Mila, I., Lapierre, C., Lazarus, S.A., Williamson, G., Scalbert, A.
764 2003. Chocolate intake increases urinary excretion of polyphenol-derived phenolic acids in
765 healthy human subjects. *Am J Clin Nutr*, **77**(4), 912-8.
766 Schnürer, A., Nordberg, A. 2008. Ammonia, a selective agent for methane production by syntrophic
767 acetate oxidation at mesophilic temperature. *Water Sci Technol*, **57**(5), 735-40.
768 Straube, J., Gorse, A.D., Huang, B.E., Le Cao, K.A. 2015. A Linear Mixed Model Spline Framework for
769 Analysing Time Course 'Omics' Data. *PLoS One*, **10**(8), e0134540.
770 Tian, H., Fotidis, I.A., Mancini, E., Treu, L., Mahdy, A., Ballesteros, M., González-Fernández, C.,
771 Angelidaki, I. 2018. Acclimation to extremely high ammonia levels in continuous
772 biomethanation process and the associated microbial community dynamics. *Bioresource*
773 *Technology*, **247**, 616-623.
774 Torres, B., Porras, G., Garcia, J.L., Diaz, E. 2003. Regulation of the mhp cluster responsible for 3-(3-
775 hydroxyphenyl)propionic acid degradation in Escherichia coli. *J Biol Chem*, **278**(30), 27575-85.
776 Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J., Tyson, G.W. 2014. Linking microbial
777 community structure, interactions and function in anaerobic digesters using new molecular
778 techniques. *Current Opinion in Biotechnology*, **27**, 55-64.

Vos, P., G., G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.-H., Whitman, W.B. 2009. Bergey's Manual of Systematic Bacteriology, 2nd ed., **3**.

Watkins, P.A., Moser, A.B., Toomer, C.B., Steinberg, S.J., Moser, H.W., Karaman, M.W., Ramaswamy, K., Siegmund, K.D., Lee, D.R., Ely, J.J., Ryder, O.A., Hacia, J.G. 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.

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

Wittmann, C., Zeng, A.P., Deckwer, W.D. 1995. Growth inhibition by ammonia and use of a pH-controlled feeding strategy for the effective cultivation of Mycobacterium chlorophenolicum. *Applied Microbiology and Biotechnology*, **44**(3), 519-525.

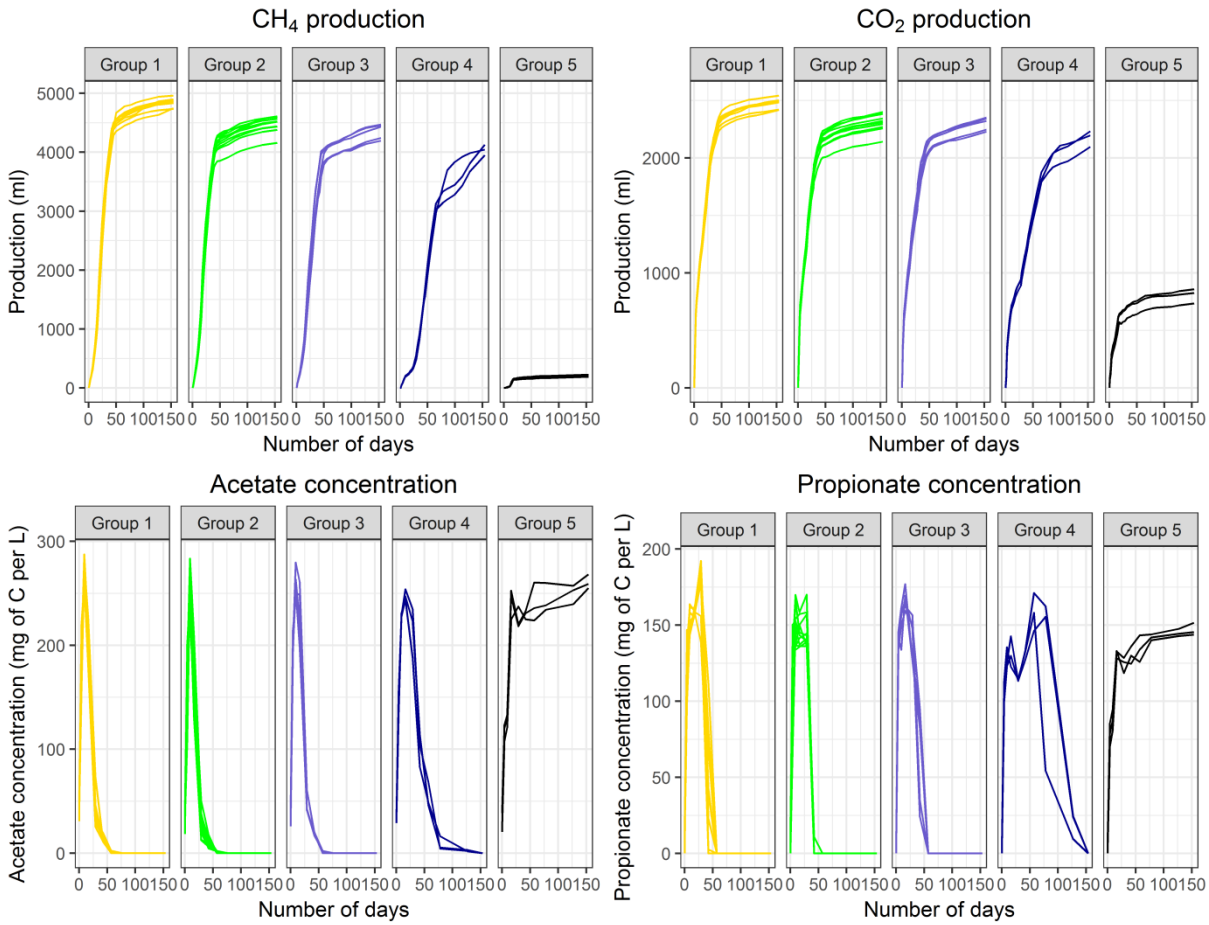
Xiao, Y.-P., Hui, W., Wang, Q., Roh, S.W., Shi, X.-Q., Shi, J.-H., Quan, Z.-X. 2009. Pseudomonas caeni sp. nov., a denitrifying bacterium isolated from the sludge of an anaerobic ammonium-oxidizing bioreactor. *International Journal of Systematic and Evolutionary Microbiology*, **59**(10), 2594-2598.

Zhu, D., Zhang, P., Xie, C., Zhang, W., Sun, J., Qian, W.-J., Yang, B. 2017. Biodegradation of alkaline lignin by Bacillus ligniniphilus L1. *Biotechnology for Biofuels*, **10**(1), 44.

Zou, B.-Z., Takeda, K., Tonouchi, A., Akada, S., Fujita, T. 2003. Characteristics of an Anaerobic, Syntrophic, Butyrate-degrading Bacterium in Paddy Field Soil. *Bioscience, Biotechnology, and Biochemistry*, **67**(10), 2059-2067.

802 # Figures:

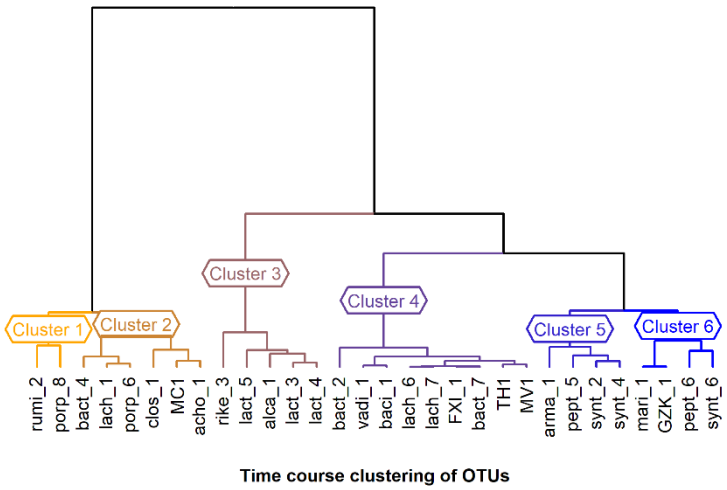
803



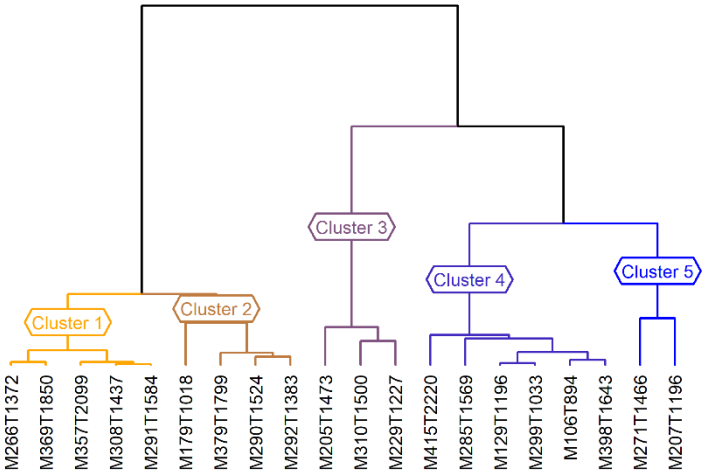
804

805 **Figure 1:** cumulated methane, carbon dioxide production (mL of gas) and acetate and propionate
806 concentration (mg of C per L) over time (number of days) for the different groups of bioreactors
807 obtained after clustering based on methane production curves.

808



Time course clustering of OTUs



Time course clustering of GCMS ions

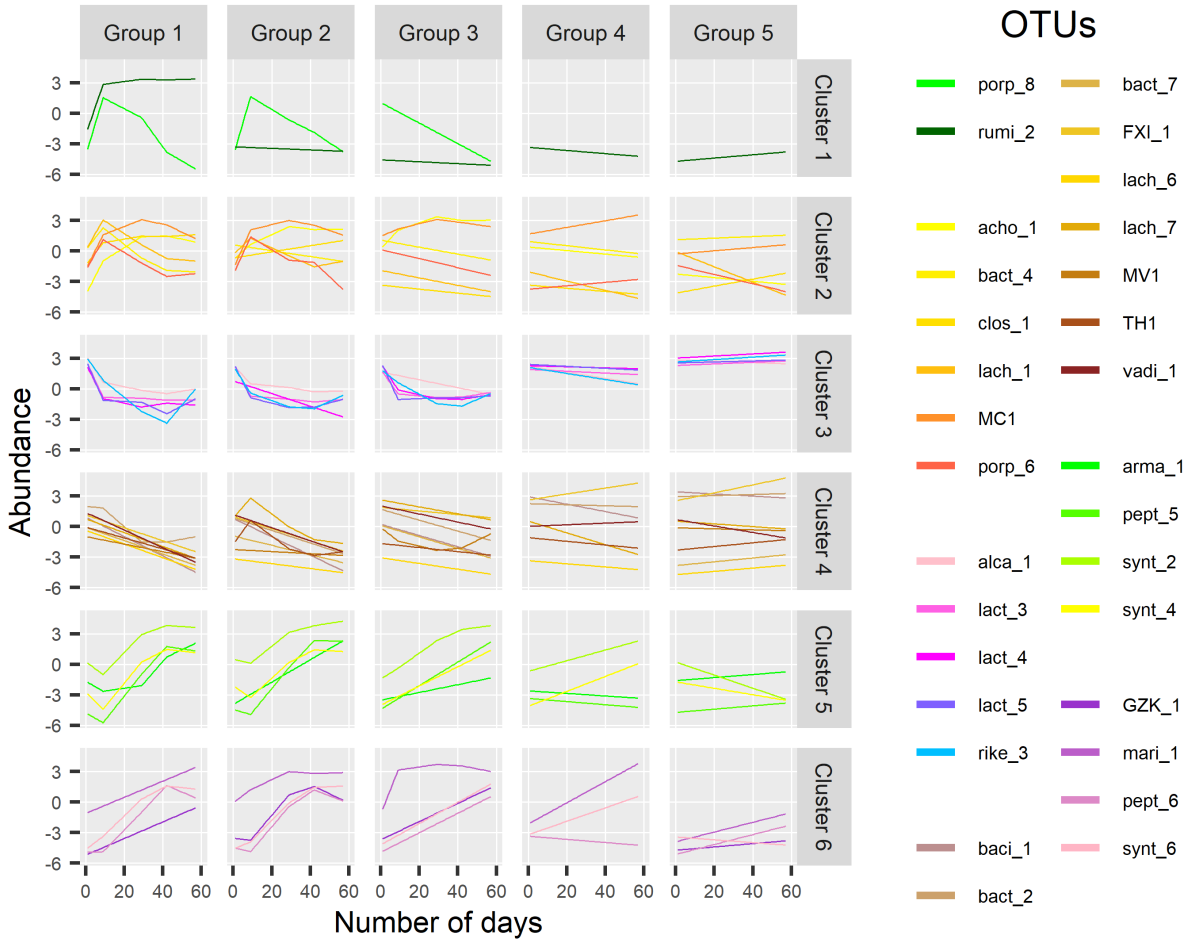
809

810 **Figure 2:** hierarchical clustering of the OTUs identified with sequencing and ions detected with

811 GCMS, based on their time-course trajectories, obtained after filtering, modelling and derivative of

812 the profiles. Colours indicate the grouping of the variables into clusters.

813



814

815 **Figure 3:** Time-course trajectories of the abundance of the different OTUs selected after filtering in
816 the 5 groups of bioreactors (in column). The different OTUs are grouped based on the time-course
817 trajectories of their abundance in group 1 (in row). Original spline fitted values.

818

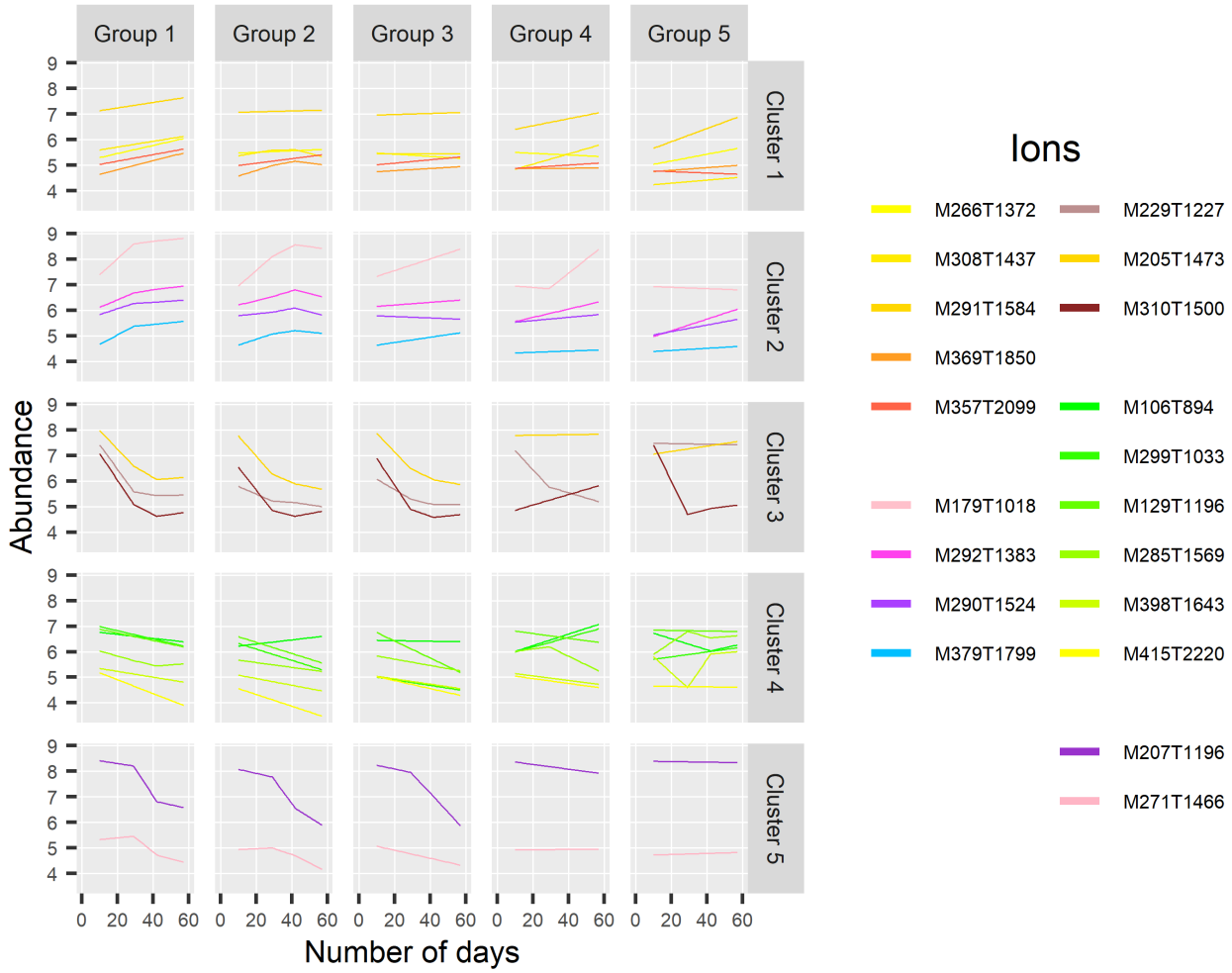


Figure 4: Time-course trajectories of the abundance of the different metabolites selected after filtering in the 5 groups of bioreactors (in column). The different metabolites are grouped based on the time-course trajectories of their abundance in group 1 (in row). Original spline fitted values

827 **Table 1:** Taxonomic affiliation of the OTUs selected after fold change filter and differentially expressed during time in group 1. OTUs are grouped by clusters

828 with similar time-course trajectories in group 1.

Cluster	OTU name	Domain	Phylum	Class	Order	Family	Genus	Species
1	porp_8	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	unknown	unknown
	rumi_2	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus 1	unknown
2	acho_1	Bacteria	Tenericutes	Mollicutes	Acholeplasmatales	Acholeplasmataceae	Acholeplasma	unknown
	bact_4	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	Bacteroides intestinalis
	clos_1	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae 1	Clostridium sensu stricto 1	Clostridium butyricum
	lach_1	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Mobilitalea	unknown
	MC1	Archaea	Euryarchaeota	Methanomicrobia	Methanomicrobiales	Methanomicrobiaceae	Methanoculleus	unknown
	porp_6	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	unknown	unknown
3	alca_1	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Paenalcaligenes	unknown
	lact_3	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	Lactobacillus parabuchneri
	lact_4	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	Lactobacillus guizhouensis
	lact_5	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	Lactobacillus parabrevis
	rike_3	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	Pseudomonas caeni
4	baci_1	Bacteria	Firmicutes	Bacilli	TSCOR001-H18	unknown	unknown	unknown
	bact_2	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	Bacteroides uniformis
	bact_7	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	Bacteroides ovatus
	FXI_1	Bacteria	Firmicutes	Clostridia	Clostridiales	Family XI	Anaerosalibacter	unknown
	lach_6	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Tyzzerella	unknown
	lach_7	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Mobilitalea	unknown species
	MV1	Archaea	Euryarchaeota	Methanobacteria	Methanobacteriales	Methanobacteriaceae	Methanobrevibacter	Methanobrevibacter smithii
	TH1	Archaea	Euryarchaeota	Thermoplasmata	Thermoplasmatales	Thermoplasmatales Incertae Sedis	Candidatus Methanogranum	unknown

	vadi_1	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiales vadinBB60 group	unknown	unknown
5	arma_1	Bacteria	Armatimonadetes	unknown	unknown	unknown	unknown	unknown
	pept_5	Bacteria	Firmicutes	Clostridia	Clostridiales	Peptococcaceae	Pelotomaculum	unknown
	synt_2	Bacteria	Firmicutes	Clostridia	Clostridiales	Syntrophomonadaceae	Syntrophomonas	unknown
	synt_4	Bacteria	Firmicutes	Clostridia	Clostridiales	Syntrophomonadaceae	Syntrophomonas	Syntrophomonas wolfei
6	GZK_1	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	GZKB124	unknown	unknown
	mari_1	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Marinilabiaceae	Alkaliflexus	unknown
	pept_6	Bacteria	Firmicutes	Clostridia	Clostridiales	Peptococcaceae	Cryptanaerobacter	unknown
	synt_6	Bacteria	Firmicutes	Clostridia	Clostridiales	Syntrophomonadaceae	Syntrophomonas	unknown

829

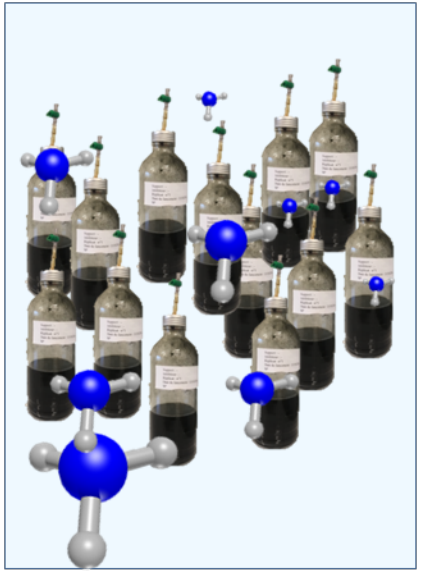
830

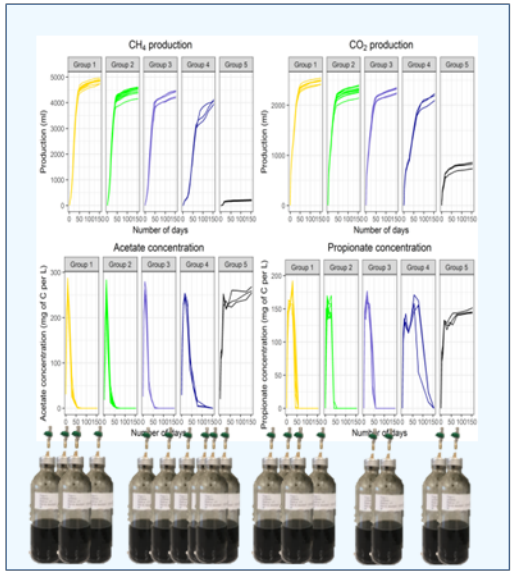
831 **Table 2:** Identification of the ions detected with GCMS and selected after fold change filter and differentially expressed during time in group 1. Ions are

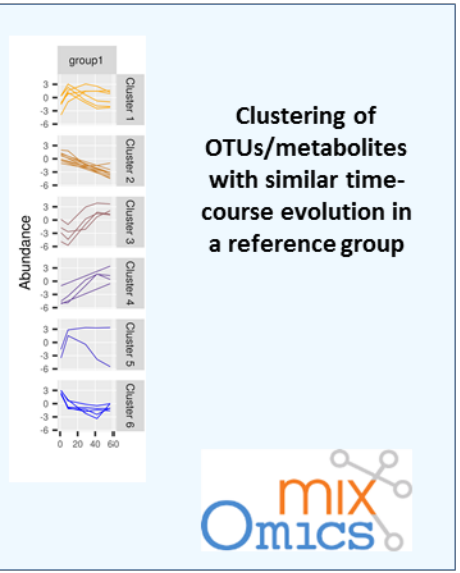
832 grouped by clusters with similar time-course trajectories in group 1.

Cluster number	Ion name (mass and retention time)	Identification
1	M266T1372	Anthranilic acid (2-aminobenzoic acid)
	M308T1437	*N*-Acetylanthranilic acid (2-Acetamidobenzoic acid)
	M291T1584	Not identified
	M369T1850	Not identified
	M357T2099	Dehydroabietic acid (Abieta-8,11,13-trien-18-oic acid)
2	M179T1018	Benzoic acid
	M292T1383	Not identified
	M290T1524	Indole-2-carboxylic acid
	M379T1799	Phytanic acid (3,7,11,15-tetramethyl hexadecanoic acid)
3	M229T1227	Decanoic acid
	M205T1473	Not identified
	M310T1500	3-(3-Hydroxyphenyl)propionic acid
4	M106T894	Not identified
	M299T1033	Phosphoric acid
	M129T1196	Not identified
	M285T1569	Myristic acid (1-tetradecanoic acid)
	M398T1643	3,4-Dihydroxyhydrocinnamic acid (3-(3,4-Dihydroxyphenyl)propionic acid)
	M415T2220	Not identified
5	M207T1196	Hydrocinnamic acid (Phenylpropanoic acid)
	M271T1466	Not identified

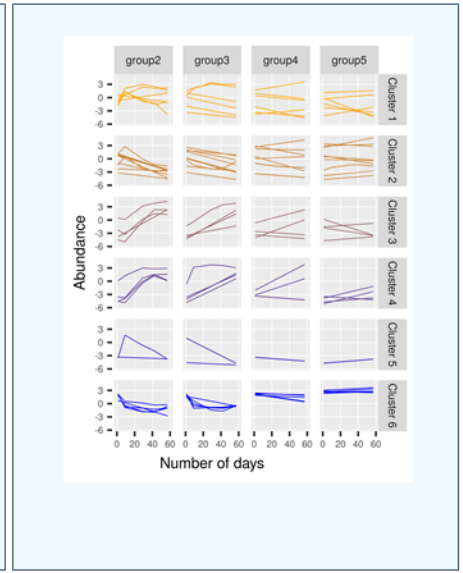
833







Clustering of OTUs/metabolites with similar time-course evolution in a reference group



30 batch digesters with various levels of ammonia

Clustering of the digesters in groups with similar degradation performances

Time-course evolution of omics in a reference group

Time-course evolution of omics in the other groups