Inhibition of anaerobic digestion by various ammonia sources resulted in subtle differences in metabolite dynamics

Xiaoqing Wang, Stephany Campuzano, Angéline Guenne, Laurent Mazéas, Olivier Chapleur

Chemosphere 351 (2024) 141157

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Inhibition of anaerobic digestion by various ammonia sources resulted in subtle differences in metabolite dynamics

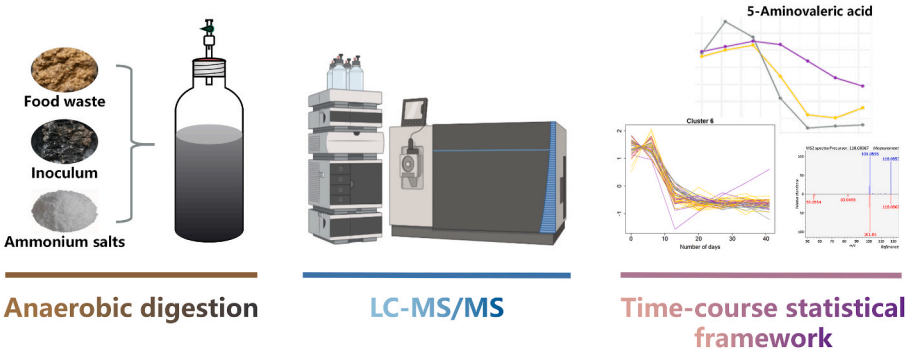
Xiaoqing Wang, Stephany Campuzano, Angéline Guenne, Laurent Mazéas, Olivier Chapleur *

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

HIGHLIGHTS

- Ammonium salt anions affected microbial metabolite dynamics in anaerobic digestion.
- 19 clusters of metabolites sharing similar time-course trajectories were identified.
- Differential analyses selected 48 metabolites typically affected by anions and time.
- 28 metabolites were annotated by MS-DIAL software with confident dot product scores.
- Phosphates slowed down the consumption of 5-aminovaleric acid.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling editor: A Adalberto Noyola

Keywords:
Anaerobic digestion
Ammonia inhibition
Anion
Non-targeted metabolomics
Time-course analysis
Statistics

ABSTRACT

The impact of ammonia on anaerobic digestion performance and microbial dynamics has been extensively studied, but the concurrent effect of anions brought by ammonium salt should not be neglected. This paper studied this effect using metabolomics and a time-course statistical framework. Metabolomics provides novel perspectives to study microbial processes and facilitates a more profound understanding at the metabolic level. The advanced statistical framework enables deciphering the complexity of large metabolomics data sets. More specifically, a series of lab-scale batch reactors were set up with different ammonia sources added. Samples of nine time points over the degradation were analyzed with liquid chromatography-mass spectrometry. A filtering procedure was applied to select the promising metabolomic peaks from 1262 peaks, followed by modeling their intensities across time. The metabolomic peaks with similar time profiles were clustered, evidencing the correlation of different biological processes. Differential analysis was performed to seek the differences in metabolite dynamics caused by different anions. Finally, tandem mass spectrometry and metabolite annotation provided further information on the molecular structure and possible metabolic pathways. For example, the consumption of 5-aminovaleric acid, a short-chain fatty acid obtained from L-lysine degradation, was slowed down by phosphates. Overall, by investigating the effect of anions on anaerobic digestion, our study demonstrated the effectiveness of metabolomics in providing detailed information in a set of samples from different experimental conditions. With the statistical framework, the approach enables capturing subtle differences in metabolite dynamics between samples while accounting for the differences caused by time variations.

* Corresponding author.
E-mail address: olivier.chapleur@inrae.fr (O. Chapleur).

https://doi.org/10.1016/j.chemosphere.2024.141157
Received 11 September 2023; Received in revised form 22 December 2023; Accepted 7 January 2024
Available online 11 January 2024

1. Introduction

Anaerobic digestion (AD) is a promising strategy for treating biodegradable organic wastes in circular waste management (Cazaudehore et al., 2022). Organic waste such as food waste, manure, and sludge from wastewater treatment plants are introduced in the anaerobic digesters and progressively degraded by the microbes to produce methane and digestate. Methane is used to produce renewable energy, while the digestate can be implemented in soil fertilization (Samoraj et al., 2022). However, the microorganisms that perform the degradation are very sensitive to various inhibitors, which can decrease the production yields and lead to instability or even failure in industrial systems (Czatzkowska et al., 2020).

In the process of AD, although nitrogen is an essential nutrient for the growth of microbes at a low level, a high concentration of ammonia is usually reported to cause inhibition (Yenigün and Demirel, 2013). Free ammonia (FAN, NH_3) and ammonium ions (NH_4^+) compose the total ammonia nitrogen (TAN), and both are considered coexisting forms of ammonia inhibition (Jiang et al., 2019). However, FAN is regarded as the first cause of AD failure due to its free penetration through the cell membranes of microbes, especially methanogens. The potential mechanisms of ammonia toxicity could be the imbalance of protons, changes in intracellular pH, and inhibition of specific enzymatic activity (Sun et al., 2022).

Previous studies have utilized nitrogen-rich substrates such as livestock manure to investigate the impact of ammonia on AD, but struggled with controlling the FAN level due to high concentrations of ammonia releasing gradually during the degradation (Yellezuome et al., 2022). Thus, using ammonium salts remains the simplest way to mimic the ammonia inhibition in lab-scale studies. Although ammonium salts may not accurately simulate the nitrogen source structure of real wastes, it allows to achieve the desired ammonia inhibition level quickly (Tian et al., 2018). However, when adding ammonium salts, the introduction of anions brought by the salts (chlorides, phosphates, carbonates, etc., hereafter referred to as anions) must be considered, as they may directly or indirectly influence the microbes and their metabolism, potentially leading to biased laboratory results. The effect of anions on AD has been documented in several previous studies. For example, Viana et al. (2012) reported that the presence of chlorides can cause co-inhibition in AD due to the high osmotic pressure they created, which lead to cell plasmolysis of microorganisms. Or else, it has been reported that the phosphates released from waste-activated sludge can inhibit the methanogenic activity as they may form phosphate precipitate and clog the internal channels of the anaerobic granular sludge (Xu et al., 2021). High phosphate concentration can also inhibit the acetoclastic *Methanosarcina* species and reduce methane production (Lackner et al., 2020). Therefore, the effect of anions is non-negligible for lab studies of AD. However, in the case of ammonia inhibition, their effect could be relatively minor compared to the inhibitory effects caused by ammonia, making it challenging to detect and analyze. Consequently, it is imperative to find an effective approach for studying the effect of anions in AD.

Metabolites are the final products of gene expression. Changes in microbial metabolite profiles provide chemical fingerprints of microbial functional states and variations (Chen et al., 2020). The field of metabolomics has emerged as a powerful tool to quantitatively describe biological systems by detecting and analyzing metabolites produced by microbes under specific environmental conditions at a specific time (Li et al., 2020; Ding et al., 2022). In particular, non-targeted metabolomics can capture all low molecular weight molecules arising from cellular metabolism and organic matter degradation by the microbes (Chapleur et al., 2021). Palama et al. (2016) introduced a novel NMR-based metabolomics strategy for rapid discrimination of bacterial species based on non-targeted metabolic profiling of the metabolites in bacterial culture media. Kasuga et al. (2020) employed non-targeted screening analysis using high-resolution Orbitrap MS to investigate the changes in the molecular-level composition of dissolved organic matter in drinking water. Pei et al. (2021) utilized 16S rRNA sequencing and GC-MS to explore the correlation between core bacteria and metabolites in aerobic composting. Metabolomics has also been applied to characterize the gut microbes and their metabolites associated with human diseases (Ye et al., 2022). Even though the application of metabolomics in AD is still limited due to the narrow prior knowledge regarding the metabolites involved in AD (Puig-Castellví et al., 2020), it introduces a fresh perspective for investigating the AD process at a metabolomic level, which enables the observation that can not be acquired through conventional taxonomic or genomic studies, such as the degradation or accumulation of microbial metabolites. Metabolomics stands as a promising tool to explore the effect of anions by focusing on the metabolites affected by different anions.

Furthermore, the longitudinal study enables to capture more information from time series data (Chapleur et al., 2021), which is particularly beneficial for analyses involving long-term sampling such as in the case of AD, facilitating the elucidation of the metabolite evolving patterns over time. The combination of metabolomics and longitudinal analysis allows to explore the dynamic changes in microbiome metabolite profiles in response to environmental stimuli over time (Beale et al., 2016; Klassen et al., 2017).

In our study, a series of lab-scale batch reactors were set up added with different ammonia sources (ammonium chloride, ammonium phosphate, and ammonium carbonate). Samples of 9 time points over the degradation were collected and analyzed with liquid chromatography-mass spectrometry (LC-MS). The purposes of our study were (1) to process the metabolomic data obtained by LC-MS and extract essential metabolomic information; (2) to apply a longitudinal statistical framework on the data and focus on the time-course evolution of the metabolites; (3) to capture the subtle differences in metabolites temporal dynamics caused by different ammonia sources; (4) to identify key metabolites using MS/MS and provide concrete metabolic information.

2. Material and methods

2.1. Experimental setup

The experimental setup is presented in Fig. 1A. Twelve bottles of 1 L were designed as lab-scale batch anaerobic digesters with a working volume of 700 mL. They were filled with 6.3 g of methanogenic sludge (0.139 g COD per g of sludge) as inoculum and 52.2 g of mashed food waste (1.173 g COD per g of food waste) as feeding to reach a substrate/inoculum ratio of 8.4 to facilitate the robust development of the microbiota. The sludge came from a 60 L laboratory anaerobic bioreactor fed with the same biowaste, and the food waste was provided by an industrial food waste collector (Valdis Energie, Issé). A biochemical potential buffer (ISO 11734 (1995)) was added in all digesters to reach the final working volume.

Three replicate digesters served as control and were not inhibited. In nine others, for every three digesters, 14.9 g of $(NH_4)_2CO_3$ (Alfa Aeser), 16.6 g of NH_4Cl (Acros Organics), and 20.5 g of $(NH_4)_2HPO_4$ (Acros Organics) were added, respectively, to reach a TAN (total ammonia nitrogen) concentration of 8 g/L. This concentration was determined based on previous tests assessing ammonia inhibition. All digesters were sealed with a rubber septum and incubated at 35 °C without agitation.

2.2. Sampling and sample preparation

The experiment lasted for nine weeks. Liquid samples were taken once a week from each digester at day 0, 6, 13, 20, 27, 34, 41, 48, and 55. The samples were then centrifuged at 10,000 g for 10 min to separate the supernatant and pellet and stored at −80 °C.

For analysis, the samples were removed from −80 °C storage and thawed in ice at room temperature. For each sample, 1 mL of supernatant was filtered by a 0.45 μm Nylon filter, followed by freeze-drying for

one night. The freeze-dried powder was then dissolved in 580 μL MilliQ water and centrifuged at 15500 rcf for 3 min at 4 °C. After centrifugation, 530 μL of supernatant was transferred to HPLC vials. 30 μL of supernatant from each vial was collected to prepare a pool of samples for quality control (QC). Finally, 500 μL of acetonitrile (Optima LC/MS grade, Fisher Chemical) was added in each vial. Additionally, a blank was prepared with equal volumes of MilliQ water and acetonitrile.

2.3. Metabolomic analysis

The analytical method is LC-MS. The instrument comprised an Accela 1250 pump system connected to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, MA, US). The analytical column is HPLC Column EC 100/2 NUCLEODUR HILIC (Macherey Nagel, length 100 mm, inner diameter 2 mm, particle size 1.8 μm). The two mobile phases were 50 mM ammonium acetate (pH 5.0, solvent A) and acetonitrile (solvent B). The chromatographic gradient program used a linear gradient of solvents A and B at a flow rate of 200 μL/min, changing from a ratio of 20:80 to 10:90, with each ratio maintained for at least 10 min. This was followed by a flow rate of 400 μL/min for an additional 10 min at a ratio of 10:90. Finally, the ratio was changed to 5:95, and samples were injected and analyzed under this condition.

10 μL was injected into the analytical system for each sample at a 400 μL/min flow rate. Samples were injected in random order to remove possible batch effects. A blank and a QC sample were injected every nine experimental samples. The samples were analyzed by Orbitrap in positive electrospray ionization mode (ESI+). The detection was performed in full scan over an m/z range from 50 to 800 at a resolution of 60,000 during an acquisition time of 25 min.

2.4. LC-MS data preprocessing

The metabolomics data have been deposited to the EMBL-EBI MetaboLights database with the identifier MTBLS7859 (https://www.ebi.ac.uk/metabolights/MTBLS7859) (Haug et al., 2020). Raw LC-MS data were converted into mzXML format files by MSConvert (ProteoWizard 3.0.21193) using the Peak Picking filter and Vendor algorithm. Then, the files were treated with XCMS Online to detect and extract the

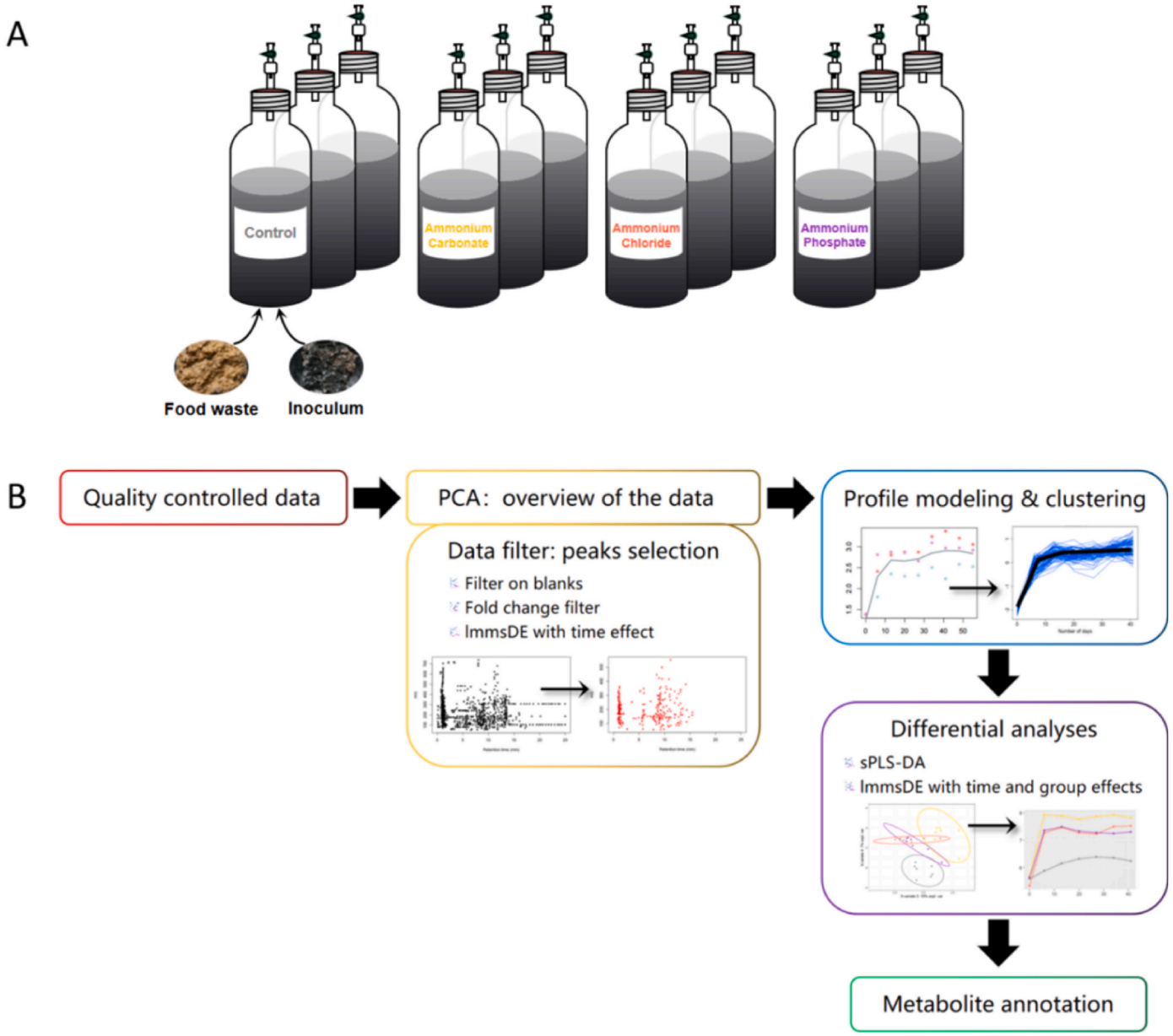


Fig. 1. (A) Experimental setup. (B) Overview of the biostatistical analysis framework.

chromatographic peaks. The parameters setting used in XCMS Online is shown in Supplementary Material (Smith et al., 2006; Tautenhahn et al., 2012). A table containing the peak intensities in each sample was then created and imported into the statTarget tool to evaluate the data quality and remove unwanted variations at the feature level (Luan et al., 2018). Drifts in signal intensities were corrected with QC-RLSC (LOESS) method and KNN (K-nearest neighbors) method was used to imputation missing values (Puig-Castellví et al., 2020).

2.5. Biostatistical analysis

The biostatistical analysis framework is briefly illustrated in Fig. 1B. The data matrix containing the peak intensities in each sample was first log-transformed and investigated by PCA computed with mixOmics R package to get an overview of the obtained data (Rohart et al., 2017). Then peaks that exhibited significant variations across time were selected using three stepwise filtering methods to investigate the time-course evolution of potential metabolites (details are provided in Supplementary Material).

Time-course profiles of kept peak intensities were modeled using spline-smoothing with linear mixed models (lmmSpline function in lmms R package), as described in Straube et al. (2015) and Bodein et al. (2019). The lmms.filter.lines function in timeOmics R package was used to filter linear models that contained less biological information (Bodein et al., 2019). Then, hierarchical clustering with ward.D agglomeration method was executed on the derivatives of the modeled profiles to identify clusters of profiles exhibiting similar patterns in their time-course trajectories, which could correspond to correlated biological processes across time.

Considering the large number of peaks and their varying presence across conditions, studying their relationships was challenging. Therefore, only the "common" peaks retained in all four conditions were targeted to compare their profiles using sparse Partial Least Squares Discriminant Analysis (sPLS-DA) (Lê Cao et al., 2011) in conjunction with lmmsDE function. Detailed description can be found in Supplementary Material.

2.6. Metabolite annotation

The molecular structure of the metabolite was first explored using a script in R based on Kendrick mass defect calculation (see the Supplementary Material for details). Tentative molecular formulae were proposed by searching the exact m/z values in METLIN and HMDB libraries (Wishart et al., 2018), using a mass tolerance of 30 ppm and a limited number of elements. The molecular formula associated with the lowest mass tolerance was proposed (Puig-Castellví et al., 2020).

To obtain the exact m/z of the potential metabolites, a pooled sample was prepared for each experimental condition to perform tandem mass spectrometry (MS/MS) analysis using LTQ-Orbitrap. The LC method was the same as described in Section 2.3. For the MS/MS method, data-dependent analysis (DDA) was performed over the m/z range from 50 to 800 at a resolution of 30,000 using higher energy collisional dissociation (HCD) activation type. The default charge state was 1, and the isolation width was 2.0 m/z during an activation time of 30 ms. The DDA was performed twice on each pool at normalized collision energy (NCE) of 50.0 and 70.0, respectively, to increase the number of fragmented peaks as much as possible.

The MS/MS raw data were converted to binary abf format using ABF Converter (Reifycs Inc., Tokyo, Japan) and imported to MS-DIAL (version 4.90, http://prime.psc.riken.jp/compms/msdial/main.html) for non-targeted data processing. The metabolite annotation was realized by matching the deconvolved experimental spectra against the reference spectra from the GNPS spectra library (All Public Spectra at GNPS) for precursor and fragment ions (Wang et al., 2016). The annotation was manually assessed and confirmed by evaluating the accurate mass score, (reverse) dot product score, and total identification score calculated by the software (Wasito et al., 2022). The parameters of each step used in MS-DIAL and the criteria for assessment are described in Supplementary Material.

3. Results and discussion

To assess the state of ammonia inhibition and the performance of the digesters, gas production and the accumulation of VFAs were closely monitored during the experiment. To avoid excessive length, detailed monitoring results are shown in the Supplementary Material. The results and discussion presented below will focus on the metabolomic data, which enabled to reveal the subtle differences induced by different anions at a new metabolomic level.

3.1. Unsupervised analysis of the metabolomics data

A total of 1262 features were acquired after the data preprocessing in Section 2.4. The data containing peak intensities and sample names were first analyzed with PCA to investigate the variance of the metabolomics data among samples.

The curves of PCA scores across time on each component are illustrated in Fig. 2. For each condition and sampling date, the mean score of the triplicates was applied with the error bars. The conventional PCA score plots are available in Supplementary Material. For PC1, all curves dropped across time, showing a significant time effect on the samples irrespective of the ammonia sources. However, a delay in the decline of the curve was observed at the initial stage, specifically for the chloride condition, resulting in its separation on PC1. Additionally, PC2 separated chloride and phosphate conditions from the other conditions. Likewise, PC3 and PC4 distinguished control and carbonate conditions, respectively, from the other conditions. PC2 to PC4 indicate the variations between conditions independent of time, proving that the anions influenced the degradation process during AD. The metabolite types or concentrations of certain metabolites may vary over time due to the alteration of dominated metabolic pathways affected by the anions, which is particularly relevant in our study. However, the information provided by PCA was limited due to the large number of peaks. Their distribution in the correlation circle plots (Supplementary Material) shows complicated correlations, and no clear patterns could be discerned. It was thus requisite to select the peaks of interest, i.e., peaks with intensities varying significantly across time or in different conditions, which could be the degradation or production of certain metabolites.

3.2. Clusters of selected peaks

The peaks of interest were selected using the data filter steps in Section 2.5. After filtering, 414, 350, 358, and 367 peaks were retained in the control, carbonate, chloride, and phosphate conditions, respectively. The intensities of these peaks were then modeled across time using smoothing splines (Déjean et al., 2007), a popular modeling approach for time-course data, which handles various sampling dates and interpolates missing values (Chapleur et al., 2021). The modeling process is illustrated in Supplementary Material. The last two time points were excluded from our final models because of the low variation in peak intensity at the last three time points, which could affect the accuracy of the modeling. The modeled profiles were filtered (Section 2.5) to remove straight-line profiles resulting from poor repeatability of the triplicates or limitations of the modeling method (Bodein et al., 2019). After profile filtering, 359, 314, 310, and 329 peaks were finally retained in the control, carbonate, chloride, and phosphate conditions, respectively.

Derivatives of the predicted profiles at each time point were computed to obtain the rate of change of the relative peak intensity over time. Hierarchical clustering was then performed based on the derivatives, which grouped similar trajectories in the rate of change

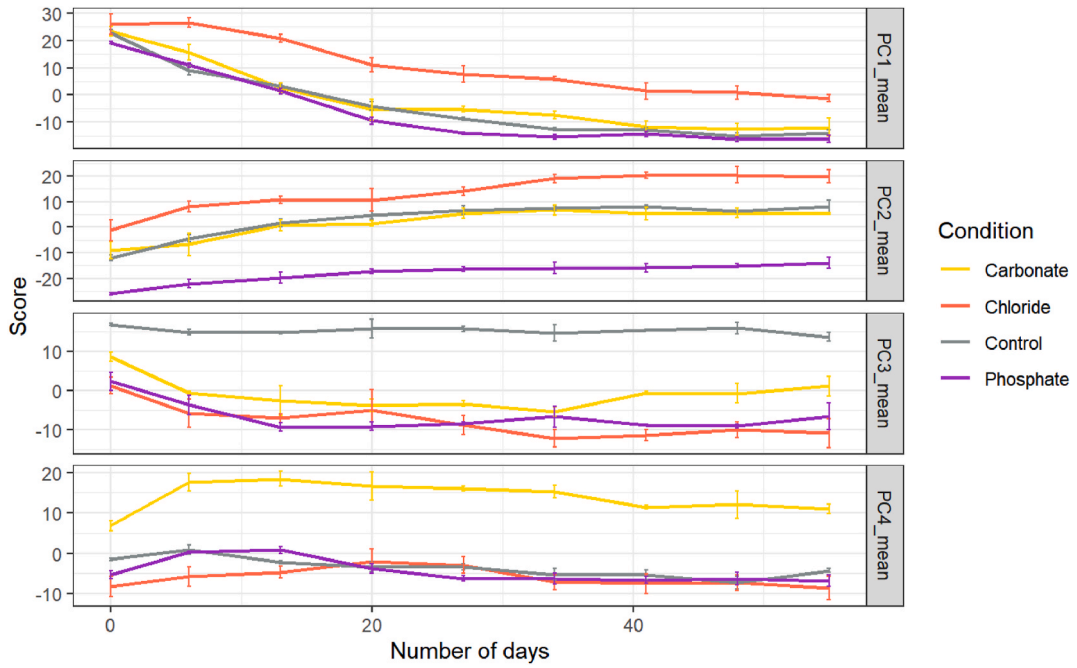


Fig. 2. Scores in PCA across time on each component. For each condition and sampling date, the mean score of the triplicates was calculated with the error bars.

(Chapleur et al., 2021), as different biological processes could be correlated across time. To investigate whether the peaks were differentially expressed in different conditions, a global clustering was carried out by merging the four conditions as illustrated in Fig. 3. Nineteen clusters of peaks with correlated time-course trajectories were identified. Each line corresponds to one peak under a specific condition (1312 peaks in total, with some peaks occurring multiple times if they were kept in more than one condition). The shapes of the trajectories were broadly grouped into four categories across time: decrease, increase, decrease followed by an increase, and increase followed by a decrease. Each category was further divided into several clusters with different kinetics. The different patterns of trajectories consisted in different degradation or production bioprocesses influenced by the inhibitory effect of ammonia or different anions. For example, the metabolites in cluster 1 were consumed much more rapidly at the beginning stage than in cluster 3. Additionally, each cluster contained trajectories from almost all conditions, although the number of trajectories for each condition varied slightly. In cluster 5, for example, the trajectories were mainly from phosphate condition, while more trajectories from carbonate condition were found in cluster 14, implying the effect of anions. So far, our statistical framework identified groups of metabolites that exhibited similar temporal dynamics, providing a novel time-course perspective on the effect of anions on AD. Importantly, the framework highlighted the time variability between conditions, which is often overlooked in classical analysis that treats time-series data as independent samples (Chapleur et al., 2021).

Out of the total 1262 *m/z*, a tentative molecular formula was assigned to 765 of them using KMD calculation and a database search to confirm the formulae. To extract information from the vast number of assigned formulae, they were categorized into individual regions based on their O/C and H/C ratios in the van Krevelen diagram depicted in Fig. 4A (Rivas-Ubach et al., 2018; Maccelli et al., 2019). Overall, the major classes of identified metabolites belonged to lipids, condensed hydrocarbons, protein, and lignin. Carbohydrates were not broadly denoted, probably due to the composition of the food waste-derived substrate.

Among the 765 compounds, 386 were retained after applying the filter procedure in the statistical analysis. For these retained compounds, Fig. 4B re-demonstrates the van Krevelen diagram showcasing the clustering results. Each facet of the figure represents a van Krevelen diagram that plots the compounds belonging to a particular cluster. The compounds from different categories are marked with their corresponding colors. The size of dots increases with the frequency of occurrences to minimize the impact of overlapping dots in a specific region. Note that, unlike the general van Krevelen diagram, our plot may feature a single compound several times if it was present in more than one condition, similar to the principle used in Fig. 3. In general, there were no significant differences in the distribution of compound categories across each cluster, but the number of compounds belonging to each category varied. For example, compared to clusters 1 and 2, clusters 3 and 4 contained a higher proportion of condensed hydrocarbons, which could account for the slower decline in the trajectories observed in these clusters.

3.3. *Profiles of metabolites influenced by anions*

3.3.1. *Differential analyses on common metabolites*

Considering the large number of peaks and their diverse trajectories, the focus was shifted to the 96 common peaks (as shown in the Venn diagram in Supplementary Material) present in all conditions to better understand their behavior in different conditions.

The supervised sPLS-DA was first performed on the relative abundance data after spline-smoothing of the 96 metabolites. The sparse variant of the method enabled the selection of the peaks that discriminated each condition and best described the difference between them across time compared to the unsupervised PCA (Poirier et al., 2020). The number of components and variables to select was specified based on the classification performance of sPLS-DA by Mfold cross-validation. Four components and 21 peaks were thus selected. The score plots for comp 1 to comp 4 and time-course profiles of the selected peaks on each component were plotted in Supplementary Material. Comp 1 to 4 discriminated the phosphate, chloride, carbonate, and control condition, respectively. To complement the results from sPLS-DA, the lmmsDE function was applied again but assigned with p-values (Group_Time) $\leq$0.05, which selected peaks with intensities varying over time and affected by different conditions. Time-course profiles of 35 peaks selected by this method were plotted in the same way in Supplementary Material. Combining the two methods, 48 distinct peaks were selected (8

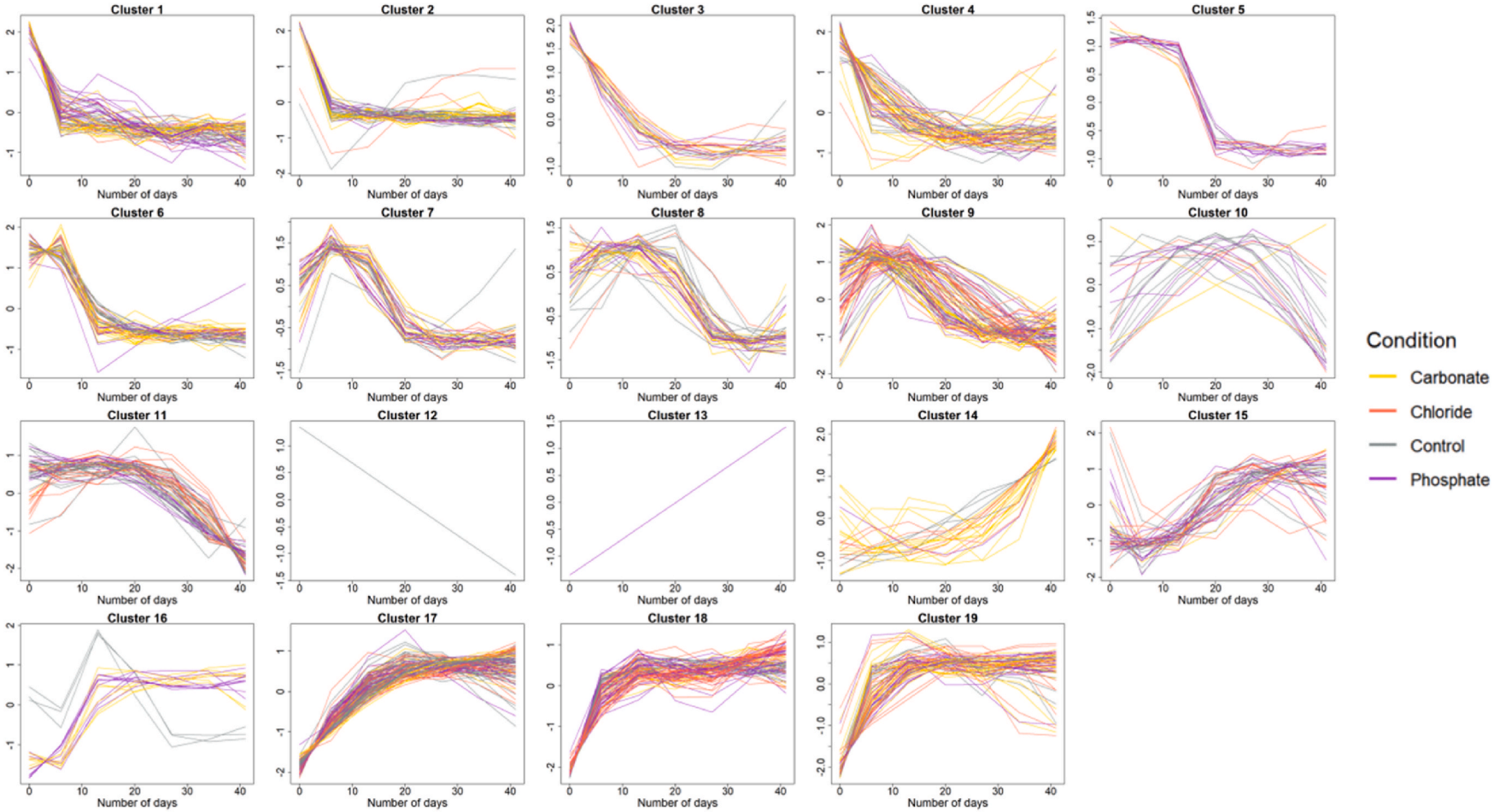


Fig. 3. 19 clusters of time-course trajectories of the peak intensities performed among four conditions. Each line corresponds to one peak under a specific condition. One peak could occur several times if kept in more than one condition.

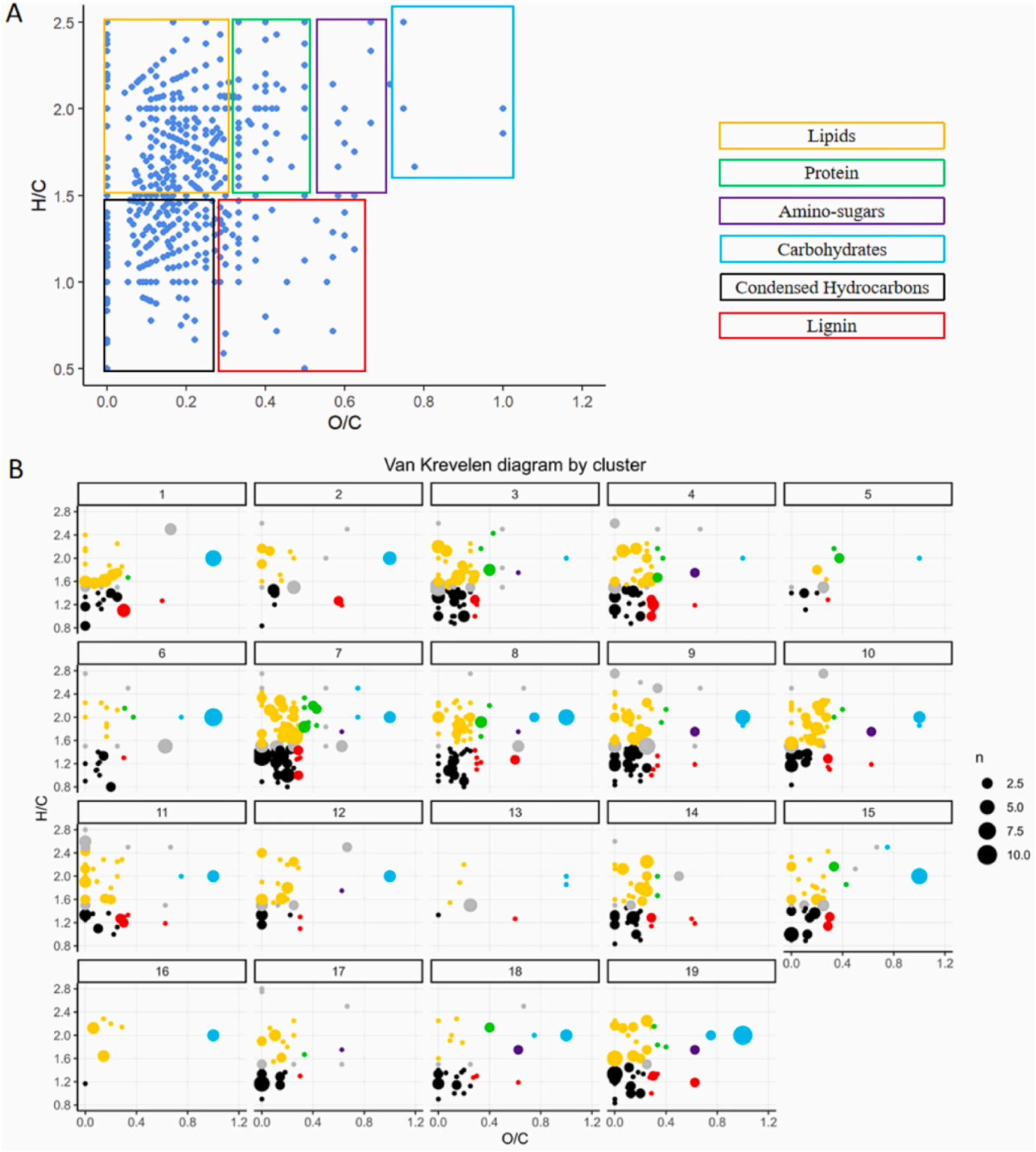


Fig. 4. (A) Van Krevelen diagram of all *m/z* values with a tentative formula. The populated metabolite classes are lipids, condensed hydrocarbons, protein, and lignin. (B) Van Krevelen diagram plotted by cluster. Each facet represents a van Krevelen diagram involving the compounds in this cluster. The compounds from different categories are marked with their corresponding colors (gray color represents the undetermined). The size of dots increases with the frequency of occurrences. A single compound could occur several times if included in more than one condition. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

peaks in common) that typically demonstrated the effect of anions on the variation of metabolites over time. The distinction of these peaks was exhibited intuitively through the heat map plotted with their spline-smoothing abundance in Fig. 5. The 0, cl, c, and p in the row names represent control, chloride, carbonate, and phosphate conditions, respectively, followed by numbers indicating the days. The color changes from day 0 to day 41 within a condition indicates the accumulation (from blue to red) or consumption (from red to blue) of the

metabolites. Different patterns of color change between conditions imply the effect of anions. For example, the abundance of m340 did not change much except in control condition, which was consumed a lot in the late stage and showed medium blue in the heat map. m96 was produced during the process, but less in control condition, which was still yellow while others were orange. m599 was accumulated in phosphate condition but degraded in the other conditions. On the contrary, m1091 was consumed only in phosphate condition while accumulated in the others. In addition, m601 was produced during the process and was least produced in phosphate condition. Similarly, the abundance of m203 increased over time and carbonate condition had the fastest production rate.

It can be concluded that the differential analyses unearthed the critical information from the large data set by focusing on the peaks with time or group (condition) effect. The intensities of the selected peaks varied significantly over time or were differently expressed under different conditions, indicating the effect of anions on the metabolites involved in this AD process.

3.3.2. *Annotation of potential metabolites by MS-DIAL*

For many users, metabolite identification remains the most time-consuming work in metabolomics analysis. Identification of LC-MS features is particularly difficult since mass spectrometry only provides limited structural information, making identifying unknown features challenging (Weber et al., 2017). According to the five criteria of confidence levels for identifying small molecules proposed by Schymanski et al. (2014), the identification attained using software could only reach Level 2, which gives a probable structure with a matched library spectrum or with diagnostic evidence when no literature information is available but no other structure fits the experimental information. The same is true for our annotation by MS-DIAL, as no authentic chemical standard was analyzed to confirm the structure.

The acquisition of MS/MS data was carried out as described in Section 2.6. The data were then processed with MS-DIAL using the settings listed in Supplementary Material. By comparing the deconvolved measurement spectra with the reference spectra from All Public Spectra at GNPS, and manual assessment and evaluation of the scores of the proposed candidates, 51 annotations were finally determined, including one compound with two different adducts (L-Tryptophan). Details of the annotation (m/z tolerance, adduct, formula, scores, peak height, etc.) can be found in Supplementary Material. According to the thresholds of the scores studied in the literature (Section 2.6), 28 annotations are deemed confident with a dot product score >800, among which three annotations are considered highly confident with a total identification score >700 (full score is 1000). The deconvolved MS/MS spectra of these confident annotations are displayed in Supplementary Material. The molecular formulae of the annotated compounds were in accordance with the tentative Kendrick formulae except the m/z = 152.05602, which was annotated as $C_5H_5N_5O$ (guanine) by MS-DIAL but $C_4H_9NO_5$ in the Kendrick table. Although our Kendrick method still needs improvement to retain as many m/z as possible, the consistency rate of the annotation by MS-DIAL with the Kendrick formulae proved

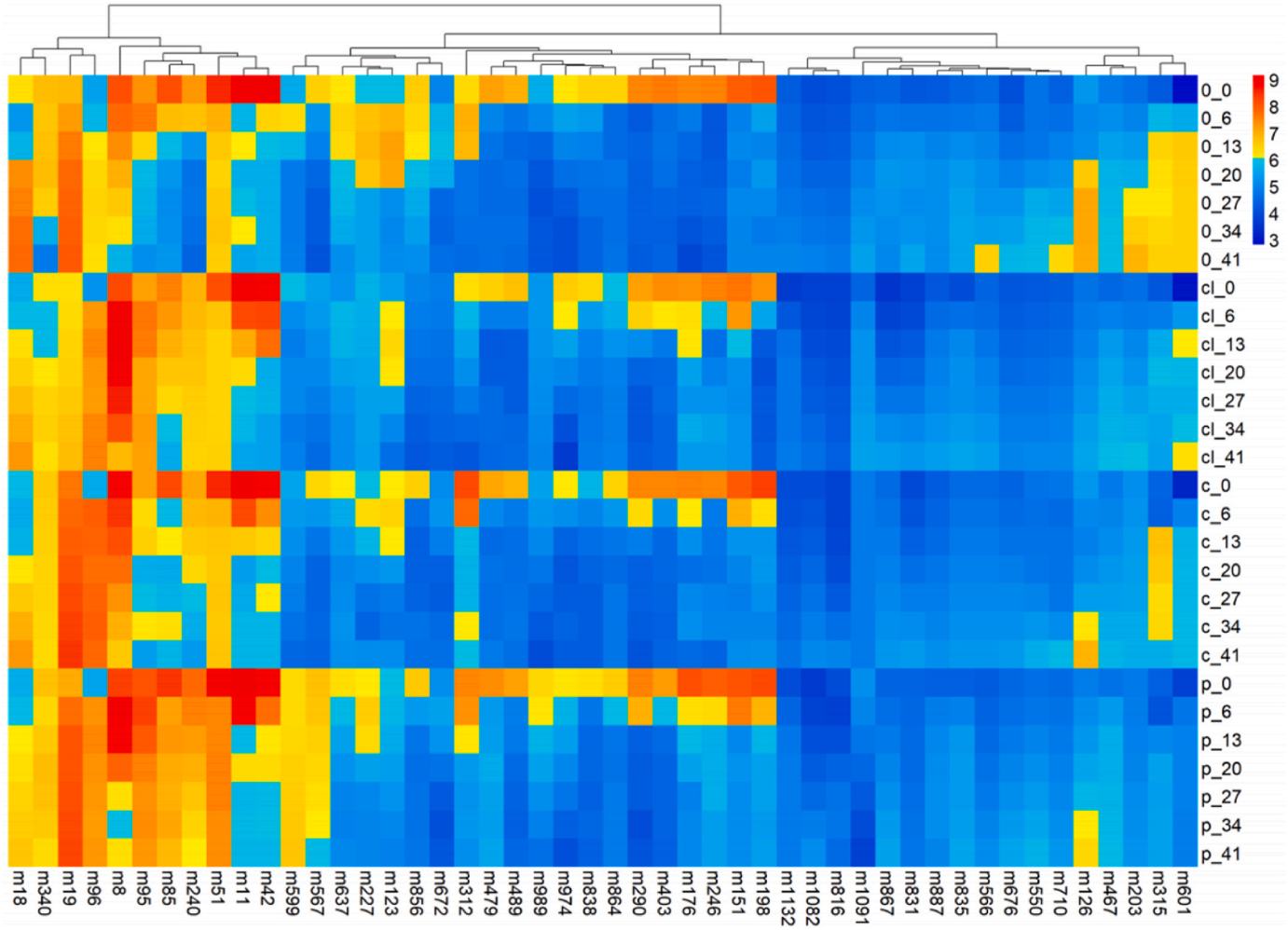


Fig. 5. Heat map of the peaks selected by sPLS-DA and lmmsDE, plotted with their spline-smoothing abundance. The row names are written in the form of condition_number of days. The 0, cl, c, and p represent control, chloride, carbonate, and phosphate conditions, respectively.

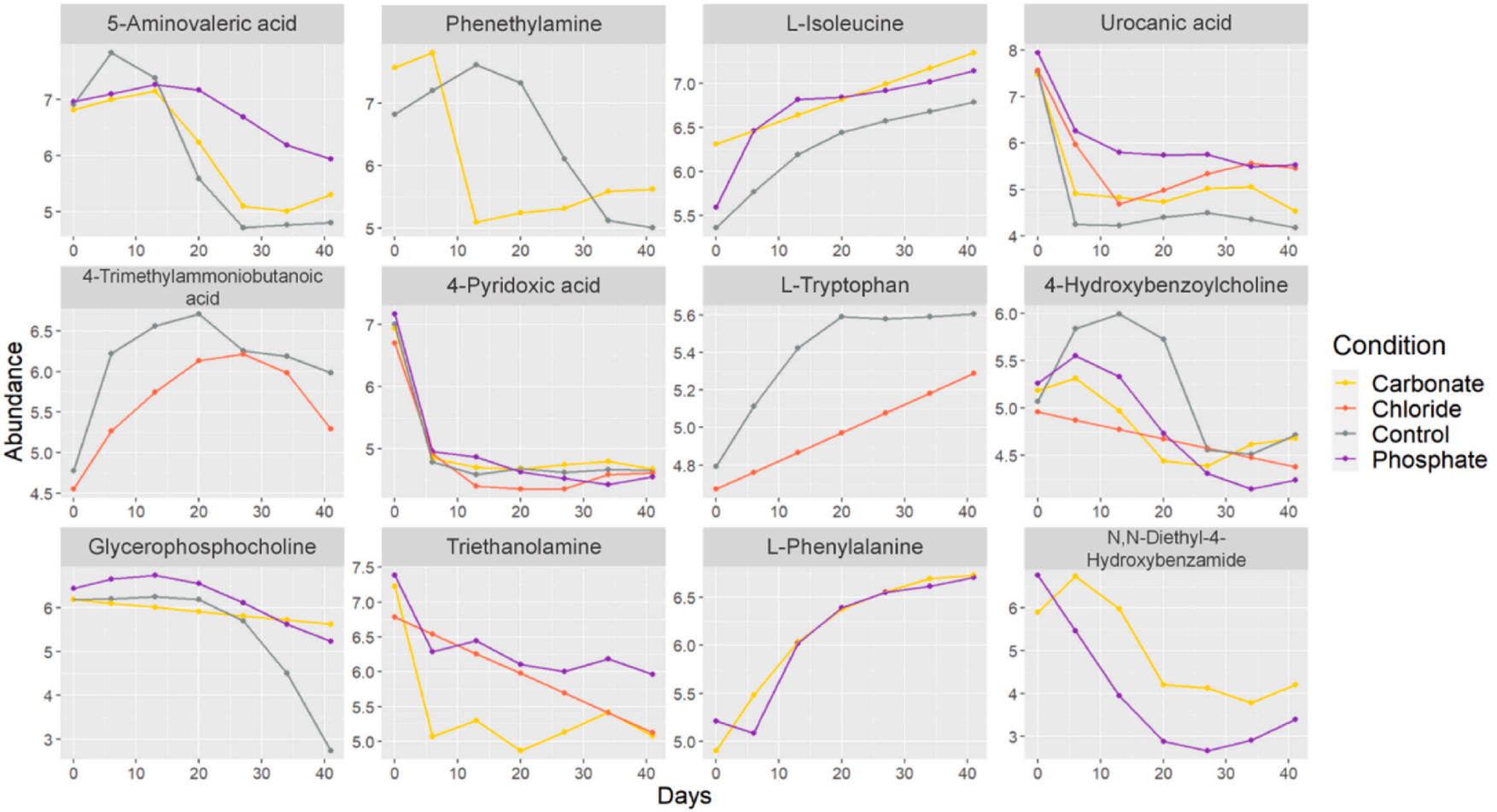


Fig. 6. Time-course profiles of the metabolites retained in at least two conditions after the filter procedure and annotated in the MS/MS analysis.

the feasibility and correctness of our method.

Relating the annotation results to the statistical analysis results, we found that among the 51 annotated metabolites, 20 were retained by our previous filter procedure, and 12 were further kept in at least two conditions. Time-course profiles of these 12 metabolites are shown in Fig. 6. Particularly, 4-hydroxybenzoylcholine was selected by sPLS-DA on comp 4, the component discriminant for the control condition. In this case, 4-hydroxybenzoylcholine was produced a lot at the beginning and remained at a much higher concentration in control condition during the first four weeks, as ammonia did not inhibit the process. Besides, urocanic acid and 4-pyridoxic acid were selected by lmmsDE analysis. The evolution of 4-pyridoxic acid mainly reflected the time effect, with a sharp decrease in all conditions, while the trajectories of urocanic acid showed time and group effects simultaneously since the degradation was most significant in control condition and less so in phosphate condition. Looking for a more obvious group effect, the consumption of 5-aminovaleric acid was slowed down by phosphates. Triethanolamine was consumed more rapidly with the presence of carbonates. In addition to indicating the effect of anions, the annotated metabolites with various trajectories under different conditions could serve as biomarkers to indicate the instability of AD. For example, glycerophosphocholine may signify instability as its concentration was considerably higher in phosphate and carbonate conditions than in the control after day 30 and did not decrease due to the inhibition by ammonium salts. Similarly, if phenethylamine was present at a very low concentration in the middle stage of the process, it may imply that inhibition is occurring.

3.3.3. *Metabolite set enrichment and pathway analysis*

The MetaboAnalyst 5.0 platform (https://www.metaboanalyst.ca/) was utilized for metabolite set enrichment analysis (MSEA) to identify metabolic pathways and gain mechanistic insights into differential regulation of biological processes (Blaber et al., 2017). In our analysis, the pathway-associated metabolite sets based on normal metabolic pathways were used to reveal the enriched metabolic pathways in the AD process.

The analysis was conducted for all annotated metabolites or a limited subset of the 12 selected metabolites in Fig. 6. A hypergeometric test was used to assess the enriched significance (p-value). Out of 51 queried compounds, 36 were successfully matched with the databases, and 40 metabolite sets were identified. Only three metabolite sets (histidine metabolism, phenylalanine and tyrosine metabolism, and tryptophan metabolism) had a p-value <0.05. Moreover, the identified metabolite sets from the 12 metabolites were in the top eleven for all annotated metabolites (Supplementary Material), except for retinol metabolism, indicating the distinctiveness of the selected metabolites.

The hit metabolites matched for each metabolite set were listed in Supplementary Material. 5-aminovaleric acid, a short-chain fatty acid, failed to match the database but has been reported to be obtained from L-lysine degradation (Cardona et al., 2020). The amino acid, L-isoleucine was also found to metabolize into acetyl-CoA directly during the AD process without converting to other amino acids (Wang et al., 2022). Consistent with the enrichment analysis results, urocanic acid was produced by the deamination of histidine catalyzed by specific enzymes (Bogachev et al., 2012). 4-Trimethylammoniobutanoic acid, also known as gamabutyrobetaine, is an end-product of the gamabutyrobetaine-crotonobetaine-carnitine cycle, found in some anaerobic bacteria (Meadows and Wargo, 2015). Associated with the time-course evolution in Fig. 6, ammonium chloride may inhibit this metabolic cycle. It's worth noting that glycerophosphocholine variations reflect non-steady-state changes in membrane turnover (i.e., enhanced membrane synthesis or breakdown) or changes in cell density (Rae, 2014). Therefore, for phosphate and carbonate conditions, the higher concentration of glycerophosphocholine in the late stage implied that these two anions might affect the synthesis of the bacterial membrane.

To conclude, the pathway analysis faces the challenge of insufficient information since the metabolite annotation is still a bottleneck. However, despite the lack of databases (spectra and pathways) specific to AD and the limited literature (Longnecker et al., 2015; Cardona et al., 2020), the advancement in metabolomics and high resolution mass spectrometry will enable our research to be carried out further in the future.

4. Conclusion

Our study employed metabolomics and a time-course statistical framework to investigate the effect of anions brought by ammonium salts on AD at the metabolic level. The critical metabolites influenced by anions over time were revealed, and further information on the molecular structure was provided by MS/MS analysis and metabolite annotation. Phosphates were found to slow down the consumption of 5-aminovaleric acid, while carbonates accelerated the consumption of triethanolamine. The identified metabolites have the potential to serve as biomarkers for indicating instability in the AD process. Overall, by investigating the effect of anions on AD, our study proved that metabolomics is a powerful tool to provide detailed information in a set of samples from different experimental conditions. In combination with the statistical framework, the approach enables capturing subtle differences in metabolite dynamics between samples while taking account of the time variability of samples.

Fundings

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, publication decision, or manuscript preparation.

CRediT authorship contribution statement

Xiaoqing Wang: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft. **Stephany Campuzano:** Conceptualization, Data curation, Investigation, Methodology. **Angéline Guenne:** Data curation, Methodology, Validation. **Laurent Mazéas:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Olivier Chapleur:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The metabolomics data have been deposited to the EMBL-EBI MetaboLights database with the identifier MTBLS7859 (https://www.ebi.ac.uk/metabolights/MTBLS7859)

Acknowledgments

We are grateful to Francesc Puig-Castellví for his exceptional guidance and support in resolving complex issues encountered during this study. Xiaoqing Wang thanks the China Scholarship Council (CSC) for her Ph.D. fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2024.141157.

References

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

Blaber, E.A., Pecaut, M.J., Jonscher, K.R., 2017. Spaceflight activates autophagy programs and the proteasome in mouse liver. Int. J. Mol. Sci. 18, 2062.

Bodein, A., Chapleur, O., Droit, A., Lê Cao, K., 2019. A generic multivariate framework for the integration of microbiome longitudinal studies with other data types. Front. Genet. 10, 963.

Bogachev, A.V., Bertsova, Y.V., Bloch, D.A., Verkhovsky, M.I., 2012. Urocanate reductase: identification of a novel anaerobic respiratory pathway in S hewanella oneidensis MR-1. Mol. Microbiol. 86, 1452–1463.

Cardona, L., Lê Cao, K.A., 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. J. Proteome Res. 19, 3981–3992.

Cazaudehore, G., Guyoneaud, R., Evon, P., Martin-Closas, L., Pelacho, A.M., Raynaud, C., Monlau, F., 2022. Can anaerobic digestion be a suitable end-of-life scenario for biodegradable plastics? a critical review of the current situation, hurdles, and challenges. Biotechnol. Adv. 56, 107916.

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

Chen, H., Hao, S., Chen, Z., Sompong, O., Fan, J., Clark, J., Luo, G., Zhang, S., 2020. Mesophilic and thermophilic anaerobic digestion of aqueous phase generated from hydrothermal liquefaction of cornstalk: molecular and metabolic insights. Water Res. 168, 115199.

Czatzkowska, M., Harnisz, M., Korzeniewska, E., Koniuszewska, I., 2020. Inhibitors of the methane fermentation process with particular emphasis on the microbiological aspect: a review. Energy Sci. Eng. 8, 1880–1897.

Déjean, S., Martin, P.G., Baccini, A., Besse, P., 2007. Clustering time-series gene expression data using smoothing spline derivatives. Eurasip J. Bioinf. Syst. Biol. 1–10, 2007.

Ding, R., Li, M., Zou, Y., Wang, Y., Yan, C., Zhang, H., Wu, R., Wu, J., 2022. Effect of normal and strict anaerobic fermentation on physicochemical quality and metabolomics of yogurt. Food Biosci. 46, 101368.

Haug, K., Cochrane, K., Nainala, V.C., Williams, M., Chang, J., Jayaseelan, K.V., O Donovan, C., 2020. MetaboLights: a resource evolving in response to the needs of its scientific community. Nucleic Acids Res. 48, D440–D444.

Jiang, Y., McAdam, E., Zhang, Y., Heaven, S., Banks, C., Longhurst, P., 2019. Ammonia inhibition and toxicity in anaerobic digestion: a critical review. J. Water Process Eng. 32, 100899.

Kasuga, I., Suzuki, M., Kurisu, F., Furumai, H., 2020. Molecular-level characterization of biodegradable organic matter causing microbial regrowth in drinking water by non-target screening using Orbitrap mass spectrometry. Water Res. 184, 116130.

Klassen, A., Faccio, A.T., Canuto, G.A.B., Cruz, P.L.R.D., Ribeiro, H.C., Tavares, M.F.M., Sussulini, A., 2017. Metabolomics: definitions and significance in systems biology. In: Metabolomics: from Fundamentals to Clinical Applications, pp. 3–17.

Lackner, N., Wagner, A.O., Markt, R., Illmer, P., 2020. pH and phosphate induced shifts in carbon flow and microbial community during thermophilic anaerobic digestion. Microorganisms 8, 286.

Lê Cao, K., Boitard, S., Besse, P., 2011. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. BMC Bioinf. 12, 1–17.

Li, M., Li, W., Wu, J., Zheng, Y., Shao, J., Li, Q., Kang, S., Zhang, Z., Yue, X., Yang, M., 2020. Quantitative lipidomics reveals alterations in donkey milk lipids according to lactation. Food Chem. 310, 125866.

Longnecker, K., Futrelle, J., Coburn, E., Soule, M.C.K., Kujawinski, E.B., 2015. Environmental metabolomics: databases and tools for data analysis. Mar. Chem. 177, 366–373.

Luan, H., Ji, F., Chen, Y., Cai, Z., 2018. statTarget: a streamlined tool for signal drift correction and interpretations of quantitative mass spectrometry-based omics data. Anal. Chim. Acta 1036, 66–72.

Maccelli, A., Vitanza, L., Imbriano, A., Fraschetti, C., Filippi, A., Goldoni, P., Maurizi, L., Ammendolia, M.G., Crestoni, M.E., Fornarini, S., 2019. Satureja Montana L. Essential oils: chemical profiles/phytochemical screening, antimicrobial activity and o/w nanoemulsion formulations. Pharmaceutics 12, 7.

Meadows, J.A., Wargo, M.J., 2015. Carnitine in bacterial physiology and metabolism. Microbiology+ 161, 1161.

Palama, T.L., Canard, I., Rautureau, G.J., Mirande, C., Chatellier, S., Elena-Herrmann, B., 2016. Identification of bacterial species by untargeted NMR spectroscopy of the exo-metabolome. Analyst 141, 4558–4561.

Pei, F., Sun, Y., Kang, J., Ye, Z., Yin, Z., Ge, J., 2021. Links between microbial compositions and metabolites during aerobic composting under amoxicillin stress was evaluated by 16S rRNA sequencing and gas chromatography-mass spectrometry: benefit for the plant growth. Biores. Technol 340, 125687.

Poirier, S., Déjean, S., Midoux, C., Lê Cao, K., Chapleur, O., 2020. Integrating independent microbial studies to build predictive models of anaerobic digestion inhibition by ammonia and phenol. Biores. Technol 316, 123952.

Puig-Castellví, F., Cardona, L., Bouveresse, D.J., Cordella, C.B.Y., Mazéas, L., Rutledge, D.N., Chapleur, O., 2020. Assessment of substrate biodegradability improvement in anaerobic co-digestion using a chemometrics-based metabolomic approach. Chemosphere 254, 126812.

Rae, C.D., 2014. A guide to the metabolic pathways and function of metabolites observed in human brain 1 H magnetic resonance spectra. Neurochem. Res. 39, 1–36.

Rivas-Ubach, A., Liu, Y., Bianchi, T.S., Tolic, N., Jansson, C., Pasa-Tolic, L., 2018. Moving beyond the van Krevelen diagram: a new stoichiometric approach for compound classification in organisms. Anal. Chem. 90, 6152–6160.

Rohart, F., Gautier, B., Singh, A., Lê Cao, K., 2017. mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13, e1005752.

Samoraj, M., Mironiuk, M., Izydorczyk, G., Witek-Krowiak, A., Szopa, D., Moustakas, K., Chojnacka, K., 2022. The challenges and perspectives for anaerobic digestion of animal waste and fertilizer application of the digestate. Chemosphere 295, 133799.

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ. Sci. Technol. 48, 2097–2098.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., Siuzdak, G., 2006. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal. Chem. 78, 779–787.

Straube, J., Gorse, A., Team, P.C.O.E., Huang, B.E., Lê Cao, K., 2015. A linear mixed model spline framework for analysing time course 'omics' data. PLoS One 10, e134540.

Sun, H., Yang, Z., Zhou, L., Papadakis, V.G., Goula, M.A., Liu, G., Zhang, Y., Wang, W., 2022. Calcium ion can alleviate ammonia inhibition on anaerobic digestion via balanced-strengthening dehydrogenases and reinforcing protein-binding structure: model evaluation and microbial characterization. Bioresource Technol. 354, 127165.

Tautenhahn, R., Patti, G.J., Rinehart, D., Siuzdak, G., 2012. XCMS online: a web-based platform to process untargeted metabolomic data. Anal. Chem. 84, 5035–5039.

Tian, H., Fotidis, I.A., Kissas, K., Angelidaki, I., 2018. Effect of different ammonia sources on aceticlastic and hydrogenotrophic methanogens. Bioresour. Technol. 250, 390–397.

Viana, M.B., Freitas, A.V., Leitão, R.C., Pinto, G.A.S., Santaella, S.T., 2012. Anaerobic digestion of crude glycerol: a review. Environ. Technol. Rev. 1, 81–92.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., 2016. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. Nat. Biotechnol. 34, 828–837.

Wang, S., Ping, Q., Li, Y., 2022. Comprehensively understanding metabolic pathways of protein during the anaerobic digestion of waste activated sludge. Chemosphere 297, 134117.

Wasito, H., Causon, T., Hann, S., 2022. Alternating in-source fragmentation with single-stage high-resolution mass spectrometry with high annotation confidence in non-targeted metabolomics. Talanta 236, 122828.

Weber, R.J., Lawson, T.N., Salek, R.M., Ebbels, T.M., Glen, R.C., Goodacre, R., Griffin, J. L., Haug, K., Koulman, A., Moreno, P., 2017. Computational tools and workflows in metabolomics: an international survey highlights the opportunity for harmonisation through Galaxy. Metabolomics 13, 1–5.

Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., 2018. Hmdb 4.0: the human metabolome database for 2018. Nucleic Acids Res. 46, D608–D617.

Xu, H., Liu, Y., Yang, B., Jia, L., Li, X., Li, F., Song, X., Cao, X., Sand, W., 2021. Inhibitory effect of released phosphate on the ability of nano zero valent iron to boost anaerobic digestion of waste-activated sludge and the remediation method. Chem. Eng. J. 405, 126506.

Ye, D., Li, X., Shen, J., Xia, X., 2022. Microbial metabolomics: from novel technologies to diversified applications. TrAC, Trends Anal. Chem., 116540

Yellezuome, D., Zhu, X., Wang, Z., Liu, R., 2022. Mitigation of ammonia inhibition in anaerobic digestion of nitrogen-rich substrates for biogas production by ammonia stripping: a review. Renew. Sustain. Energy Rev. 157, 112043.

Yenigün, O., Demirel, B., 2013. Ammonia inhibition in anaerobic digestion: a review. Process Biochem. 48, 901–911.