



HAL
open science

Monitoring of cellulose-rich biowaste co-digestion with 3D fluorescence spectroscopy and mass spectrometry-based metabolomics

Olivier Chapleur, Angéline Guenne, Douglas N Rutledge, Francesc Puig-Castellví

► **To cite this version:**

Olivier Chapleur, Angéline Guenne, Douglas N Rutledge, Francesc Puig-Castellví. Monitoring of cellulose-rich biowaste co-digestion with 3D fluorescence spectroscopy and mass spectrometry-based metabolomics. *Chemosphere*, 2023, 349, pp.140824. <10.1016/j.chemosphere.2023.140824>. <hal-04330239>

HAL Id: hal-04330239

<https://hal.inrae.fr/hal-04330239v1>

Submitted on 6 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 **Monitoring of cellulose-rich biowaste co-digestion with 3D fluorescence**
2 **spectroscopy and mass spectrometry-based metabolomics**

3 *Olivier Chapleur,¹ Angéline Guenne,¹ Douglas N. Rutledge,^{2,3} Francesc Puig-Castellví^{1,4,5,*}*

4 ¹Université Paris-Saclay, INRAE, PRocédés biOtechnologiques au Service de l'Environnement, 92761 Antony,
5 France

6 ²Faculté de Pharmacie, Université Paris-Saclay, 91400 Orsay, France

7 ³Muséum National d'Histoire Naturelle, 75005 Paris, France

8 ⁴Université Paris-Saclay, INRAE AgroParisTech, UMR SayFood, 75005 Paris, France

9 ⁵Present address: EGENODIA Inserm U1283, CNRS UMR8199, Institut Pasteur de Lille, Lille University
10 Hospital, European Genomic Institute for Diabetes (EGID), Université de Lille, 59045, Lille, France

11 **Corresponding Author. Tel: +33 0320974242, email address: francesc.puig-castellvi@inserm.fr*

12

13 **Abbreviations**

14 3D EEM: Three dimensional excitation emission matrix

15 AD: Anaerobic digestion

16 AcoD: Anaerobic co-digestion

17 FW: Food waste

18 HPLC: High performance liquid chromatography

19 MS: Mass spectrometry

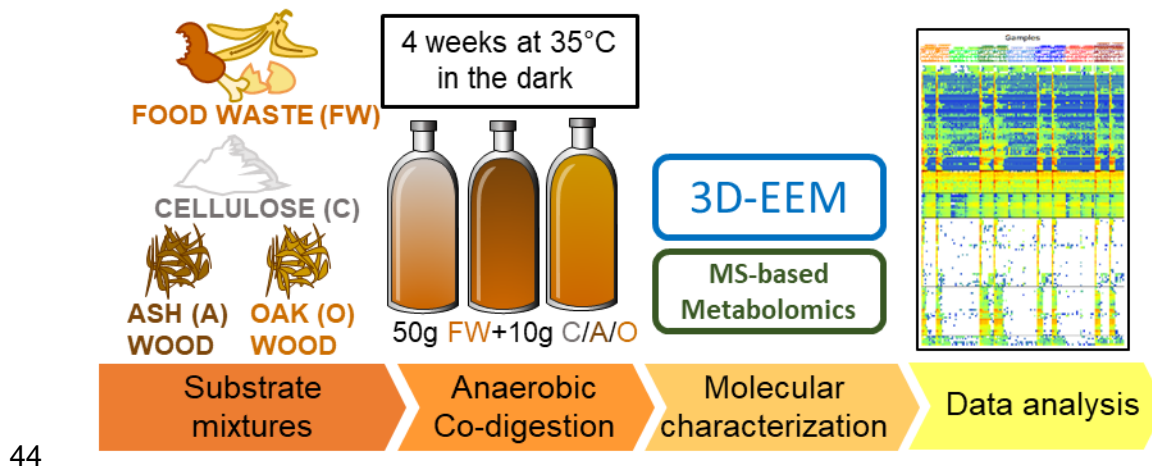
20 **Abstract**

21 Anaerobic digestion (AD) is a promising waste management strategy that reduces landfilling while
22 generating biogas. Anaerobic co-digestion involves mixing two or more substrates to enhance the
23 nutrient balance required for microorganism growth and thus improve the degradation. Monitoring
24 AD is crucial for comprehending the biological process, optimizing process stability, and achieving
25 efficient biogas production. In this work, we have used three dimensional excitation emission
26 fluorescence spectroscopy and mass spectrometry metabolomics, two complementary techniques, to
27 monitor the anaerobic co-digestion (AcoD) of cellulose, ash wood or oak wood with food waste. The
28 two approaches were compared together and to the biogas production records. Results of this
29 experiment demonstrated the complementarity of both analytical techniques with the measurement of
30 the biogas production since 3D fluorescence spectroscopy and MS metabolomics revealed the earlier
31 molecular changes occurring in the bioreactors, mainly associated with the hydrolysis step, whereas
32 the biogas production data reflected the biological activity in the last step of the digestion. Moreover,
33 in all cases, the three data sets effectively delineated the differences among the substrates. While the
34 two wood substrates were poorly degradable as they were richer in aromatic compounds, cellulose
35 was highly degradable and was characterized by the production of several glycolipids. Then, the three
36 tested AcoDs resulted in a similar 3D EEM fluorescence and metabolomics profiles, close to the one
37 observed for the AD of food waste alone, indicating that the incorporation of the food waste drove the
38 molecular degradation events in the AcoDs. Substrate-specific differences were appreciated from the
39 biogas production data. The overall results of this research are expected to provide insight into the
40 design of guidelines for monitoring AcoD.

41 **Keywords**

42 anaerobic digestion, food waste, wood, 3D EEM, metabolomics

43 Graphical abstract



44 45 1. Introduction

46 The production of municipal solid waste grows around the world (Braguglia et al., 2018) . In 2017,
47 biowaste in the 28 European Union member states accounted for more than 34 % of the municipal
48 solid waste generated, or 86 million tonnes (EEA, 2020). Biowaste includes garden and park waste,
49 food products and kitchen waste from households, restaurants and supermarkets, and comparable
50 waste from food processing plants (EEA, 2020). Recycling biowaste is essential to achieve the EU's
51 goal of recycling 65% of municipal waste by 2035 (EEA, 2020). Composting and anaerobic digestion
52 (AD) are currently the two most widely applied techniques to reduce the amount of biowaste and
53 minimize its landfilling (EEA, 2020). While composting is more extended than AD, the use of the
54 latter is increasing because it is a source of renewable energy as it produces biogas and because AD
55 tends to deliver higher environmental benefits (Slorach et al., 2019).

56 Biowaste contains, depending on its origin, a mixture of readily biodegradable organic constituents
57 (e.g., simple sugars, starch, proteins and lipids), more recalcitrant biopolymers (e.g., lignocellulose)
58 and undesirable content (e.g., plastic, packaging) (Moretti et al., 2020). During anaerobic digestion,
59 complex organic molecules are broken down into simpler compounds that are ultimately converted to
60 methane and CO₂ by the microbial community living in the anaerobic digesters (Chatterjee and

61 Mazumder, 2019). Biogas production in anaerobic digesters depends on the capacity of the microbial
62 community to degrade the waste. To increase the degradability of the more recalcitrant waste, it can be
63 digested in combination with more biodegradable substrates. This strategy is known as anaerobic co-
64 digestion (AcoD) and is considered to be very environmentally efficient as it allows for the
65 simultaneous processing of multiple types of waste (Borowski and Kubacki, 2015).

66 A current challenge in AcoD is its optimization for a maximal methane yield while also maintaining
67 process stability and better handling of wastes. In this aim, it is important to monitor anaerobic
68 digesters during AcoD for understanding their functional behavior, that has a direct impact for
69 efficient biogas production (Cruz et al., 2021). So far, AD is mainly assessed by the inspection of
70 several parameters such as pH, alkalinity, volatile fatty acid concentrations, biogas composition, and
71 the composition of the microbial community (Awhangbo et al., 2020). While these parameters can
72 provide information about the digester stability, they are not descriptive of the specific molecular
73 events occurring during the waste degradation and, thus, they cannot fully explain the underlying
74 cause that resulted on the measured reactor's disturbance (i.e., pH acidification, low methane yield
75 production). To do so, several analytical techniques have been proposed to characterize the molecular
76 composition of biological samples. For example, spectroscopic techniques can be used to pinpoint the
77 functional groups in the compounds present in the studied samples (e.g., infrared, ultraviolet and
78 fluorescence spectroscopies) (Madsen et al., 2011). Mass spectrometry can also be employed to
79 identify the small compounds in the samples (Du et al., 2023).

80 Three dimensional excitation emission fluorescence spectroscopy and mass spectrometry are two
81 techniques whose combined use for the study of AD is promising. 3D fluorescence spectroscopy
82 generates 3D excitation-emission matrices (3D EEM) which can provide information about the
83 structural complexity of the organic matter, characterizing its degradation capabilities (Maynaud et al.,
84 2017). Then, mass spectrometry provides a comprehensive view of the specific compositional changes
85 occurring during biowaste degradation. 3D EEM fluorescence spectroscopy can be used to
86 characterize the signature of macromolecules such as fulvic acid and proteins (Muller et al., 2014),
87 while mass spectrometry reveals the changes in the smaller molecules (Puig-Castellví et al., 2020).

88 This application of mass spectrometry to characterize the small molecules in biological samples is
89 known as mass spectrometry metabolomics (Dettmer et al., 2007).

90 The potential of these techniques for the waste management field has resulted in a few recent
91 publications on their use in anaerobic digestion. In Li et al. (2023) and Zhao et al. (2023), the
92 anaerobic digestion of wastewater sludge was studied with 3D EEM spectroscopy and Fourier
93 transform ion cyclotron resonance mass spectrometry. Having said this, wastewater sludge and
94 biowaste are two very different substrates. In addition to biomass (e.g. fibers, proteins, fats) found in
95 both substrates (Bhatia et al., 2018; Rorat et al., 2019), wastewater sludge can contain considerable
96 amounts of emerging contaminants such as phthalates, organic nanoparticles and surfactants, among
97 other chemicals (Rodriguez-Narvaez et al., 2017). Due to the presence of these compounds in
98 wastewater sludge and to a lower protein content than in biowaste, the latter is generally more
99 efficiently degraded than the former (Iacovidou et al., 2012). In this context, the use of these two
100 analytical techniques could result in a better understanding of each specific AcoD process (e.g., to
101 examine the inhibitory effects of heavy metals and other compounds, and to characterize the
102 degradation profile of contaminants in the sludge). This knowledge could then lead proposals for new
103 process strategies to improve overall process efficiency (e.g., changing the substrate/co-substrate ratio
104 (Puig-Castellví et al., 2020) or other operational parameters such as the reactor's temperature (Puig-
105 Castellví et al., 2022), or using bio-augmentation or additional substrate pretreatments to facilitate the
106 degradation of a specific family of compounds). To our knowledge, these two techniques have never
107 been used simultaneously to study the AD of biowaste.

108 In order to obtain insight about the molecular processes in the context of the AcoD of biowaste and
109 their implications in biogas production, we have explored the digestion of three cellulosic substrates
110 (cellulose, and sawdust from ash (*Fraxinus spp.*) and oak (*Quercus spp.*) trees employing food waste
111 (FW) as their co-substrate with 3D fluorescence spectroscopy and mass spectrometry metabolomics.
112 Therefore, the main objectives of this study were to (1) evaluate the performance of the investigated
113 AcoD conditions, (2) analyze the temporal changes in the molecular composition of the substrates and
114 assess the complementarity of the two techniques, and (3) relate the observed changes to the variations

115 in biogas production. The results from this study would provide significant guides for studying
116 biowaste degradation from a molecular perspective, with potential applications for the development of
117 a monitoring tool aiming to improve process efficiency in waste management.

118 **2. Material and methods**

119 **2.1. Feedstock preparation and experimental set-up**

120 FW was provided by an industrial food waste collector (Valdis Energie, Issé). Cellulose was alpha
121 cellulose (Sigma-Aldrich). Ash and oak wood was shredded sawdust collected at a local sawmill
122 processing these two types of wood. The inoculum used in the study was obtained from a mesophilic
123 60 L laboratory anaerobic bioreactor fed with similar FW. For the experimental set-up, a total of 14
124 anaerobic batch bioreactors, each consisting of 1 L glass bottles, were used. In this study, 7 different
125 substrate compositions are tested in duplicate: the four mentioned substrates, and three binary
126 mixtures containing FW with either cellulose, ash wood or oak wood. Bioreactors dedicated to mono-
127 digestions contained 52.2 g of FW, 10 g of cellulose, 10 g of ash wood or 10 g of oak wood,
128 corresponding respectively to 12, 4.6, 5.8 or 6.3 grams of Chemical Oxygen Demand (gCOD)
129 (LCK514 DCO test kit, Hach Lange). Binary mixtures used in the remaining bioreactors were
130 prepared by blending together a quantity of each substrate similar to the one used for mono-digestion.
131 Then, all 14 bioreactors were inoculated with 60 g (1.2 gCOD) of anaerobic sludge. All the
132 bioreactors were complemented with a biochemical potential buffer (International Standard ISO
133 11734 (1995) (International Organization for Standardization, Geneva, 1995)) up to a final volume of
134 700 mL. Bioreactors were sealed with a screw cap and a rubber septum and headspaces were flushed
135 with N₂ (purity > 99.99%, Linde gas SA), and incubated for 4 weeks at 35°C in the dark without
136 agitation. Daily biogas productions were measured with a Gas Endeavour® (BCP Instrument).

137

138 **2.2. Sampling**

139 Sampling was carried out at days 0, 3, 7, 10, 14, 17, 21, 24, 28 and 31. Sampling was performed by
140 collecting 10 mL of liquid phase from the digester through the septum using a syringe. Then samples
141 were centrifuged at 10,000 g for 10 min at 4°C to collect the supernatants, which were then snap
142 frozen in liquid nitrogen and kept at -20°C prior the 3D EEM fluorescence and metabolomic analyses.
143 The 14 bioreactors and the 10 sampling dates resulted in a total of 140 samples being collected.

144 **2.3. 3D fluorescence analysis**

145 3D fluorescence spectra were measured directly on the thawed samples without prior chemical
146 treatment, using a Xenius spectrofluorometer (SAFAS, Monaco) equipped with a xenon lamp source,
147 excitation and emission monochromators and a front face sample-cell holder. Measurements were
148 carried out in triplicate using a quartz cuvette. The instrumental settings were: excitation range 240–
149 600 nm, emission range 210–620 nm, wavelength increments 3 nm, spectral slits 3 nm, automatic
150 (EX) and open (EM) filter positions, and PMT voltage 420 V.

151 **2.4. MS metabolomic analysis**

152 One mL of each sample was diluted with 4 mL of double-distilled H₂O. The resulting samples were
153 then filtered with 0.45 µm Nylon filters and stored at -80°C until the HPLC-MS analyses were
154 conducted. HPLC-MS samples were prepared by mixing 500 µL of each sample with an equal amount
155 of acetonitrile (ACN). Quality Control samples (QCs), formed by pooling all individual samples, were
156 also prepared. The instrumental setup consisted of an Accela 1250 pump system connected to an
157 LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Waltham, MA, USA) operated in positive
158 electrospray ionization mode (ESI+). Detection was performed in full scan over an m/z range from 50
159 to 800 at a resolution of 60,000. A NUCLEODUR HILIC column (100 mm x 2.1 mm ID, particle size
160 1.8 µm) provided with a NUCLEODUR HILIC guard column (4 mm x 2.1 mm ID, particle size 1.8
161 µm) (Macherey-Nagel, Düren, Germany) was used. Mobile phases used were milli-Q ACN (phase A)
162 and water containing 50 mM ammonium acetate at pH 5.0 (phase B). For each sample, 10 µL were

163 injected. The flow rate was set at 0.4 mL min⁻¹. The HPLC gradient started at 5 % B and held for 3
164 min, then increased to 40 % B during 11 min. From minutes 14 to 24, B was linearly decreased to 5
165 %. Finally, initial conditions were re-equilibrated in 5 minutes, resulting in a total run time of 29 min.
166 To remove possible batch effects, samples were injected in random order. Moreover, Quality Controls
167 (QCs) and blanks were injected every 8 samples. In total, 181 samples were injected (140 samples, 19
168 QCs and 22 blanks).
169 Raw HPLC-MS data were converted into mzXML-format files using MSConvert (ProteoWizard 3.0).
170 Then, chromatographic features were extracted using the XCMS Online platform.

171

172 **2.5. Data pretreatment and analysis**

173 **2.5.1. Fluorescence dataset**

174 Each 3D fluorescence spectrum can be regarded as a matrix, where the columns and rows describe the
175 excitation (121 wavelengths) and corresponding emission (125 wavelengths) spectra, respectively.
176 For every fluorescence spectrum, Rayleigh and Raman scatterings were removed and replaced by
177 interpolation of the surrounding data using a three-dimensional Delaunay triangulation method, as
178 described in Zepp et al., 2004. Each set of triplicate technical spectra were averaged and unfolded into
179 a vector, and the vectors for the 140 samples (14 bioreactors times 10 time-points) were concatenated
180 column-wise keeping in common the same emission and excitation wavelengths. Spectral points
181 containing intensity values lower than 40 for all the samples were removed. After this variable
182 removal, the dataset had 140 rows and 7331 columns. Next, this dataset was SNV-transformed and
183 analyzed by Independent Components Analysis (ICA) (Rutledge and Jouan-Rimbaud Bouveresse,
184 2013).
185 ICA is a blind source chemometric method used for separating mixed signals into their underlying
186 source signals and the proportions (or scores). In this context, the JADE algorithm, one of the several
187 methods available for calculating ICA models, was employed. (Rutledge and Jouan-Rimbaud
188 Bouveresse, 2013). The determination of the optimal number of Independent Components (ICs) to use

189 is a key step for building an ICA model. This number was determined using both the Durbin–Watson
190 criterion (Jouan-Rimbaud Bouveresse et al., 2012) and the random ICA method (Kassouf et al., 2018).
191 In order to explore the ICA results at the variable level, the ICA signals were refolded (filling the
192 empty spaces due to the variable removal with 0s) so as to recover the matrix structure of the original
193 3D fluorescence spectra.

194 **2.5.2. LC-MS metabolomics dataset**

195 The metabolomics dataset, with 140 samples in rows and 1947 LC-MS features in columns, was
196 examined using two complementary strategies.

197 A first strategy was applied to identify the main metabolite degradation trends occurring in the
198 bioreactors. In this strategy, the matrix was auto-scaled and analyzed by Principal Components
199 Analysis (PCA) (Peris-Díaz and Krężel, 2021). Then, for each principal component (PC),
200 chromatographic features with loading values beyond 2 standard deviations were considered to be
201 representative of these degradation trends (Puig-Castellví et al., 2020).

202 A second strategy was used to determine the features that enable the discrimination of the two wood
203 types. In this case, a subset of the metabolomics dataset was created by selecting those samples
204 containing either oak or ash wood (in presence or not of FW) that had been collected on the first
205 sampling day. This dataset was auto-scaled and subsequently analyzed with Orthogonal Partial Least
206 Squares-Discriminant Analysis (O-PLS-DA). This analysis was performed using the *opls* function
207 from the *ropls* package in R (Thévenot et al., 2015). LC-MS features with Variable Importance in
208 Projection (VIP) values above 2 were considered to be drivers of the sample discrimination and their
209 temporal profile over the 31 days was inspected in a heatmap.

210 For both strategies, features were tentatively annotated using the LC-MS search tool from HMDB
211 (Wishart et al., 2022).

212 **3. Results**

213 **3.1. Biogas production**

214 The daily biogas production profile in **Figure 1** reflects the efficiency and performance of the
215 anaerobic digestion process. Early peaks in this profile correspond to highly degradable compounds,
216 while later peaks are associated with less accessible ones. Profiles without significant peaks indicate a
217 highly recalcitrant composition.

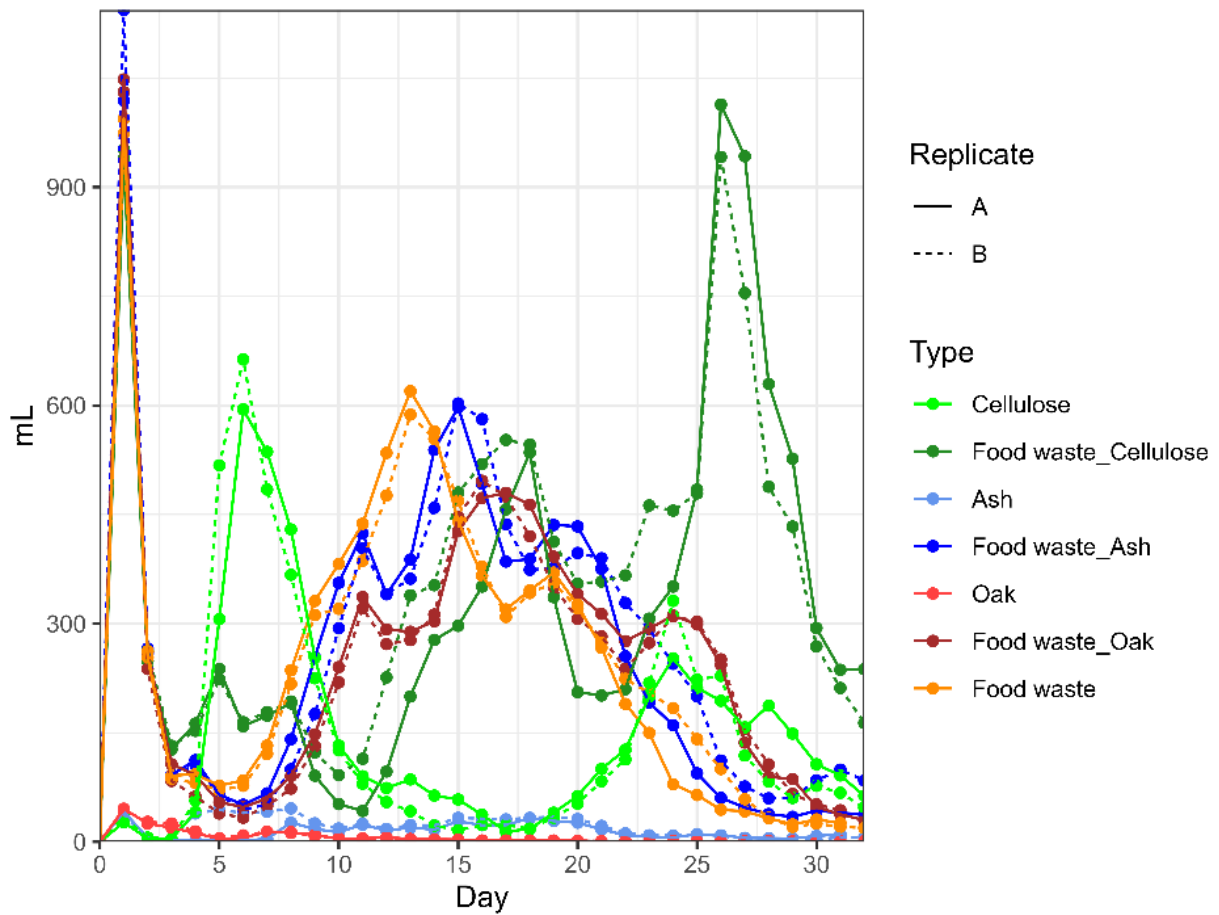
218 The biogas production varied over time and could be categorized into four temporal periods: early
219 degradation (days 1 to 4), the immediate period following early degradation (approximately days 4 to
220 10), mid-degradation (approximately weeks 2 to 3), and late degradation (after 3 weeks).

221 During the early degradation stage, biogas production exhibited a prominent peak in FW-containing
222 digestions, approximately 15-25 times larger than in equivalent mono-substrate digestions. In the
223 second degradation period, cellulose had the highest biogas production peak (600 mL/day at day 7),
224 followed by FW-cellulose (approximately 230 mL/day at day 5). Oak and ash mono-digestions also
225 had peaks in this period (14 mL/day at day 7 for oak AD and 36 mL/day at day 10 for ash AD).

226 The mid-degradation phase was primarily observed in reactors containing FW, showing a wide,
227 continuous peak throughout this period. FW AD reached a peak of 626 mL/day on day 13, while the
228 corresponding peaks for FW co-digestions appeared slightly later (day 15 for FW-ash AD and day 17
229 for FW-oak and FW-cellulose ADs). Ash AD also had a peak, but it was approximately 10 times
230 smaller than that of FW-containing digestions.

231 In the final degradation period, production peaks were observed in cellulose-containing digesters,
232 similar to the second degradation period. However, the maximum biogas production was higher in co-
233 digestion systems (the contrary of the observed in the second degradation period), reaching around
234 1000 mL/day, while the corresponding mono-digestions achieved levels of approximately 300
235 mL/day.

236 Considering the total amount of biogas produced in the AcoD systems compared to the sum of the
237 mono-digestion counterparts, FW-oak resulted in slightly higher biogas production (8407 mL for
238 AcoD versus 8163 mL for the sum of the FW and oak mono-digestions). In contrast, FW-ash and FW-
239 cellulose produced less biogas compared to when each substrate was degraded separately, averaging
240 7900 mL versus 8525 mL and 11034 mL versus 12459 mL, respectively, for the FW-ash and FW-
241 cellulose comparisons.



242

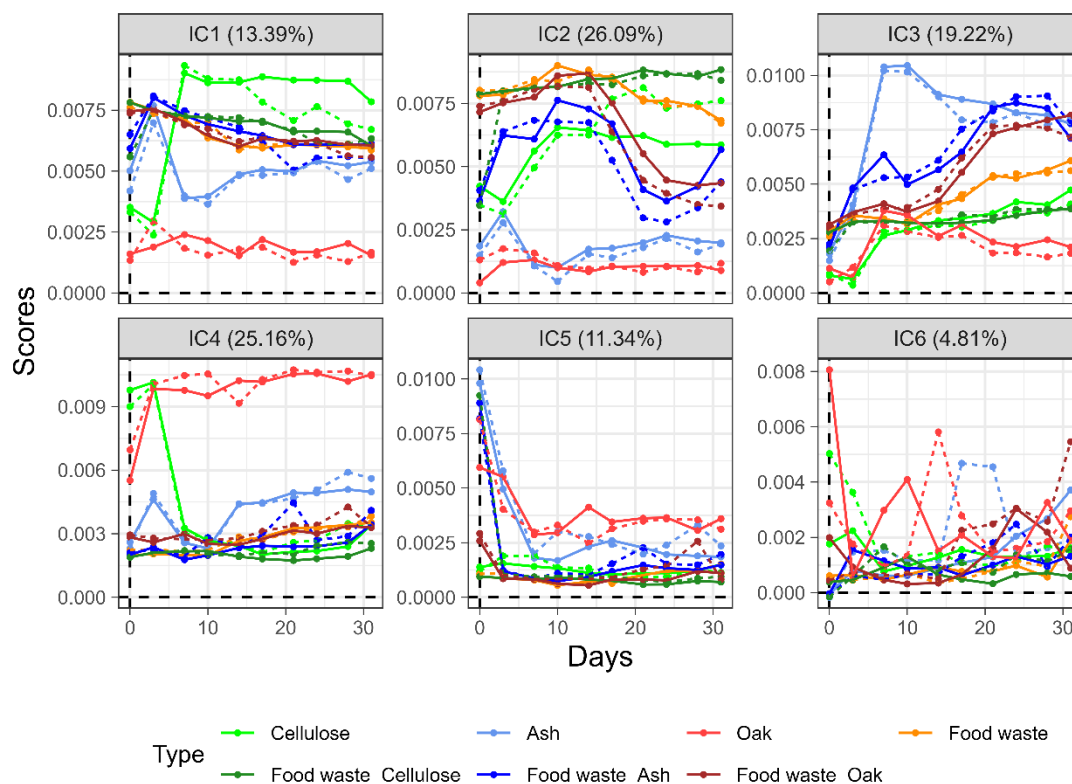
243

244 **Figure 1.** Daily biogas production.

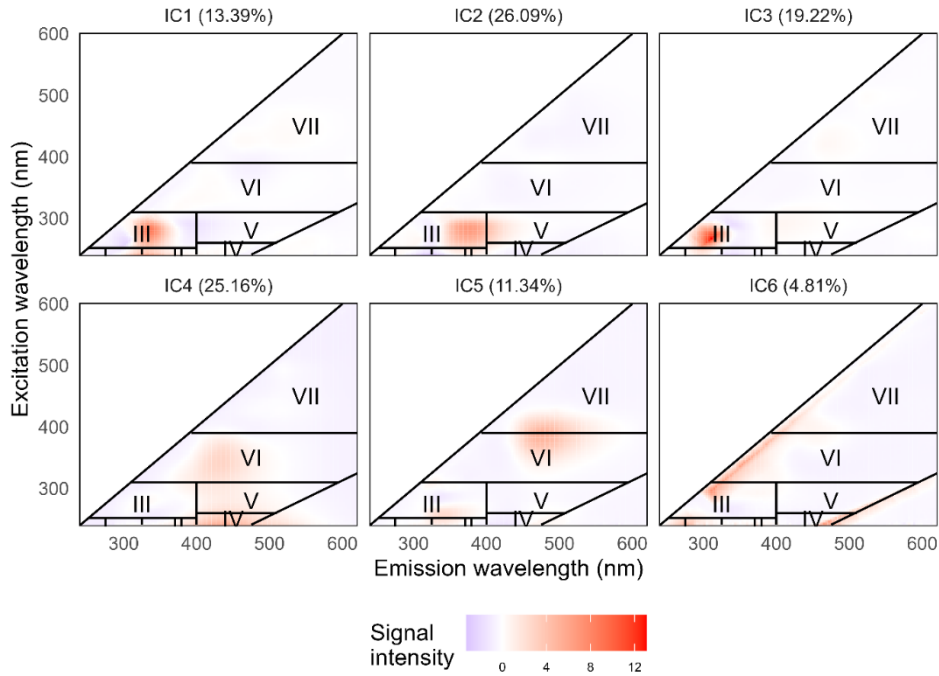
245 3.2. 3D fluorescence spectroscopy

246 The fluorescence data provides information about the degradation of fluorescent compounds in the
247 bioreactors, such as proteins and other more recalcitrant substances like humic acids (Muller et al.,
248 2014). These data were analyzed by ICA to extract the degradation profiles of these compounds

249 (Figure 2 and 3), using 6 ICs for building the ICA model. Each independent component (IC) is
 250 composed of a vector of IC proportions, or scores (Figure 2), describing the temporal evolution of a
 251 compound, and a vector of IC signals (Figure 3), which contains the 3D EEM signature of that
 252 compound.
 253



254
 255 **Figure 2.** Temporal evolution of the ICA scores of the 3D fluorescence data. Replicates are depicted
 256 with different line types.



257

258 **Figure 3.** ICA signals of the 3D fluorescence data. Regions I-III: protein-like content; regions IV-V:
 259 fulvic-acid-like content; region VI: glycated protein-like molecules, lignocellulose-like and humic
 260 acid-like compounds; region VII: humic-like and lipofuscin-like materials.

261

262 Each of the components can be linked to one or a few molecular groups of compounds depending on
 263 the spectral fingerprint associated with them (Muller et al., 2014). Components IC1 (13.39% of the
 264 explained variance), IC2 (26.09%) and IC3 (19.22%) are characteristic of signals from region III,
 265 attributed to the protein-like content in the samples. The maxima of these signals are at wavelengths
 266 of $\lambda_{EX}=282\text{nm}$ and $\lambda_{EM}=334\text{nm}$ for IC1, $\lambda_{EX}=279\text{nm}$ and $\lambda_{EM}=373\text{nm}$ for IC2, and $\lambda_{EX}=267\text{nm}$ and
 267 $\lambda_{EM}=311\text{nm}$ for IC3, respectively (top panels in Figure 3). IC4 (25.16% of the explained variance) is
 268 descriptive of two signals, one from region IV ($\lambda_{EX}=240\text{nm}$ and $\lambda_{EM}=439\text{nm}$) and another from region
 269 VI ($\lambda_{EX}=345\text{nm}$ and $\lambda_{EM}=436\text{nm}$) (bottom left panel in Figure 3). The signal from region IV is
 270 attributed to fulvic-acid-like material while that from region VI can represent glycated protein-like
 271 molecules, lignocellulose-like and humic acid-like compounds. IC5 (11.34%) is defined by a peak
 272 comprised between regions VI and VII ($\lambda_{EX}=387\text{nm}$ and $\lambda_{EM}=475\text{nm}$) (bottom middle panel in Figure
 273 3). Region VII represents humic-like and lipofuscin-like materials (e.g., oxidized proteins and lipids).
 274 Lastly, IC6 (4.81%) captured the residual signals of the Rayleigh and Raman scatterings (bottom right

275 panel in Figure 3). The residual nature of IC6, rather than chemical, can also be appreciated from the
276 erratic behavior of the scores over time (bottom right panel in Figure 2), as opposed to the more
277 structured score profiles in the other 5 ICs.

278 The temporal evolution of the ADs can be studied by examining the scores (Figure 2):

279 Regarding IC1, cellulose AD followed a sudden increase after the 4th day and it remained the AD with
280 the largest scores for IC1. Then, the IC1 scores related to all FW-containing AD followed the same
281 trend, consisting of a constant slow decline over time. For ash AD, the temporal profile showed a first
282 maxima at the 4th day, followed by a strong decrease, and a continued increase after that. Oak is the
283 substrate with the lowest IC1 scores, which remained stable during all the experiment.

284 Ash, oak and cellulose ADs followed a similar evolution in IC2, albeit at a lower magnitude,
285 compared to their corresponding IC1 scores. Then, the largest IC2 scores were found for FW, FW-oak
286 and FW-cellulose ADs. In all these ADs, they increased monotonically. This rise was maintained for
287 all the experiment for FW-cellulose while they started to decline at the 14th day for the other two ADs,
288 more abruptly in FW-oak than in FW. The FW-ash AD followed a waving profile, starting at low
289 scores and with change of progression trends in day 14 and once again in day 21.

290 IC3 scores exhibited a steady pattern for the first 10 days of the experiments and started to increase
291 after that for most FW-containing ADs (except for FW-ash AD). From the slightest to the most
292 marked increase, the sequence of escalation was as follows: FW-cellulose, FW and FW-oak. A second
293 trend was followed by cellulose and oak ADs. These two ADs started at very low IC3 scores,
294 increasing after day 4 to reach the levels for FW AD. After day 7, oak AD decreased while cellulose
295 AD mimicked the profile of FW-cellulose AD. Ash AD manifested a strong increase during the first
296 week and was followed by a slower decline that reached the levels observed for FW-oak ADs in this
297 component. FW-ash temporal evolution shows a profile that includes trends already observed in the
298 corresponding ash and FW mono-digestions, with a first increase during the first days of the
299 experiment (as in ash AD), and another increase after day 10 (as in FW AD).

300 At the beginning of the experiment, IC4 has important contributions from cellulose and oak ADs
301 (although more in the former), with very low contributions for the rest. After the day 4, IC4 scores
302 rapidly decreased for cellulose AD, reaching the levels observed for the FW AcoDs, and rapidly
303 increased for oak AD. The trend observed for cellulose AD was opposed to that seen in IC1 for the
304 same substrate. After that, IC4 remained steady in all conditions except ash AD until the end of the
305 experiment. In ash AD, the temporal profile resembled that of the same substrate observed in IC1 and
306 IC2, with a small peak at day 4 and a continued increase after day 10. At the end of the experiment,
307 IC4 scores for ash AD were about half of those obtained for oak AD.

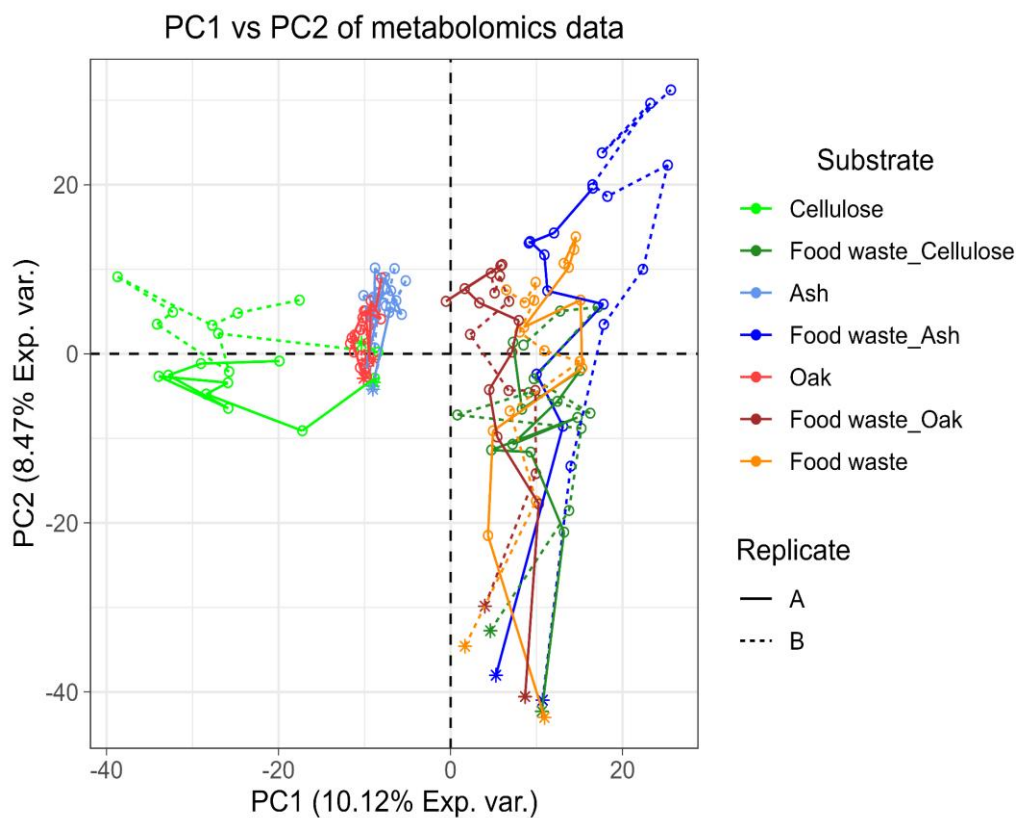
308 IC5 temporal evolution is mostly specific to ash- and oak-containing ADs, although more important in
309 the former. In all these digesters, IC5 scores decreased rapidly at the beginning and remained stable
310 for the rest of the experiment.

311 **3.3. LC-MS metabolomics**

312 Metabolomics, similarly to 3D fluorescence spectroscopy, can provide information about the
313 degradation processes in waste materials. Rather than identifying the general groups of molecules
314 being modified during the degradation, metabolomics highlights the specific molecules involved in
315 the degradation process, to the extent that is possible to pinpoint the main degradation processes as
316 well as determining how they are for each of the tested substrates.

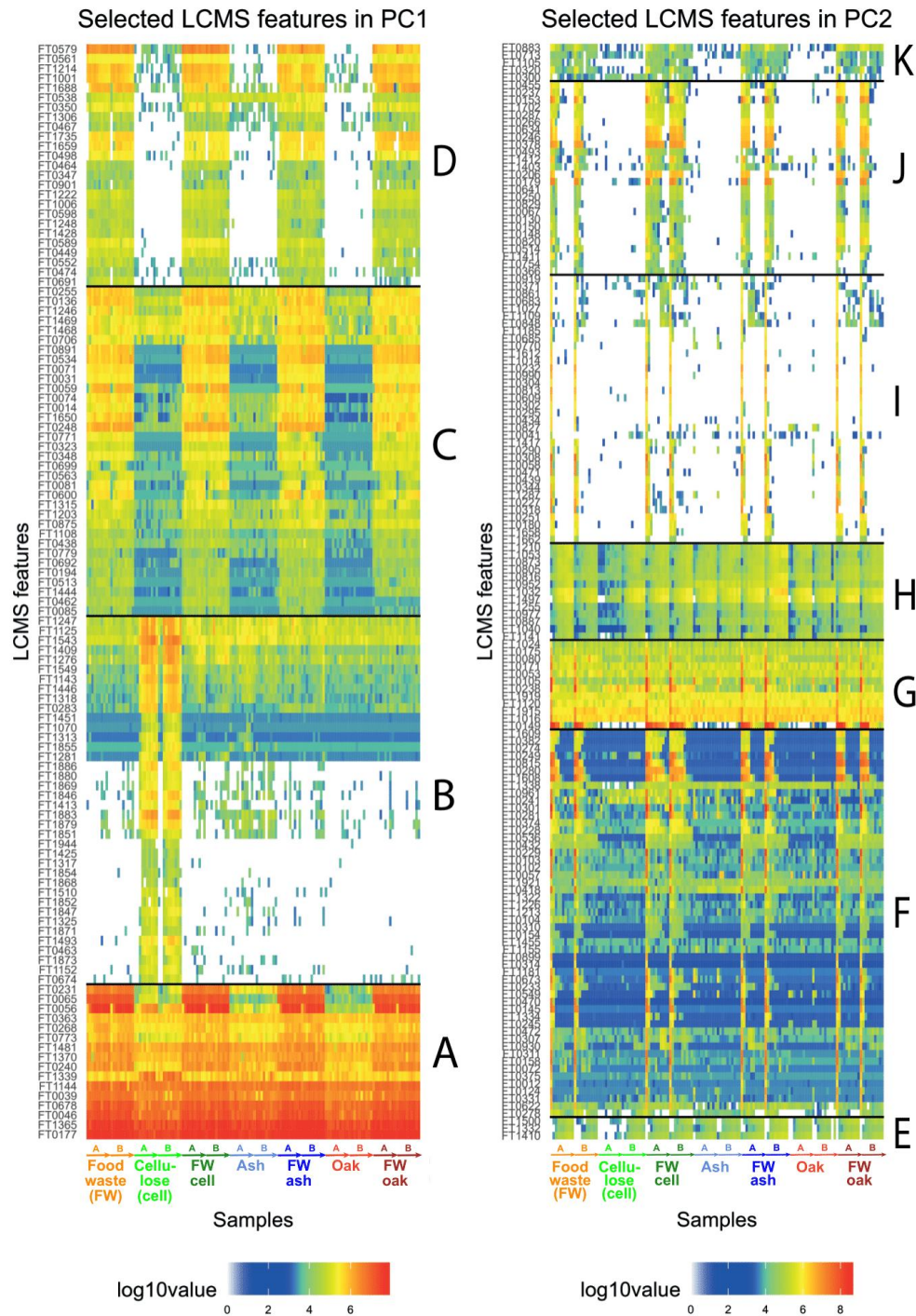
317 As in the case of the 3D fluorescence data, the metabolomics data was first explored using PCA. The
318 first two components of the PCA on the metabolomics data (**Figure 4**) explained 18.6% of the total
319 dataset variance, meaning that the metabolomics data is the result of more biological processes than
320 the 3D fluorescence data. In the PCA scores plot, PC1 separates samples containing FW (positive PC1
321 scores) from those that do not (negative PC1 scores). PC1 scores from cellulose, ash and oak samples
322 collected at the beginning of the digestion are clustered together. The PC1 scores of the oak and ash
323 samples did not substantially change over time, while the PC1 scores of the cellulose digestion
324 samples evolved towards more negative values. On the other hand, PC2 mainly describes the temporal

325 evolution of samples containing FW, which go from negative to positive scores over time. Among
326 them, the addition of oak or cellulose to FW did not produce an evident change in the metabolomics
327 profile of the digestion. However, in the case of FW-ash AcoD, this caused a more dilated evolution
328 in the PC2 dimension. At later time-points, the PC2 values became significantly more positive
329 compared to the other digestion setups containing FW. PC2 does not reveal any consistent metabolic
330 changes in the cellulose, ash and oak woods bioreactors.



331
332 **Figure 4.** PCA scores of the metabolomics data. Samples from the first day are denoted with stars
333 (instead of circles). Chronologically consecutive samples are connected with lines.

334
335 In order to further investigate the metabolic changes observed, the most representative LCMS features
336 from PC1 and PC2 were plotted using heatmaps (**Figure 5**). 113 LCMS features were selected from
337 PC1 and 145 features from PC2, based on their VIP.



338

339 **Figure 5.** Heatmaps of the selected features in PC1 and PC2. Samples from each reactor are grouped

340 and sorted by time. A and B stand for the replicates.

341

342 LCMS PC1 features can be classified into four different groups (A-D, **Figure 5A**). Metabolites in

343 group B are more abundant in cellulose-only containing samples, while metabolites in groups A, C

344 and D are more abundant in the FW-containing reactors. From a structural level, metabolic diversity

345 varies across the three groups. Among other compounds, group B contains glycosylated compounds
 346 (e.g., ceramides), group A has several adducts of phosphoric acid and short alkenes, group C has
 347 oxidized fatty acids, and group D contains quinolines, polycyclic structures (naphthyls, phenantrenes)
 348 and some glycerolipids). More detail about the annotated metabolites can be consulted in **Table A.1**.
 349 Similarly, PC2 features can be classified into 7 groups (E-K, **Figure 5B**), depending on whether they
 350 are being accumulated or consumed over time, their presence at the first sampling date, how long their
 351 accumulation or consumption lasted, and on whether they were specific to biowaste-containing
 352 bioreactors. A summary of the main differences across groups is given in **Table1**. More detail about
 353 the annotated metabolites in PC2 can be consulted in **Table A.1**.

354 **Table1.** Groups of metabolites according to their evolution in PC2.

Group	Temporal evolution	Detection at day 0	Progression stopped after	FW-specific	Metabolite examples
E	Accumulation	No	3 days	No	oxidized C20-fatty acids
F	Consumption	Yes	3-14 days	No	taurine, lysine
G	Consumption	Yes	3 days	No	oxazine, adenine
H	Accumulation	Yes	7 days	No	oxidized C14-fatty acids
I	Consumption	Yes	7 days	Yes	creatine, dipeptides
J	Consumption	Yes	7-21 days	Yes	quinolines, cholines
K	Consumption	Yes	14 days	Yes	coumaryl acetate, hydrocinnamic acid

355 According to **Figure 5B** and **Table 1**, it can be seen that, in all cases, some of the metabolites that
 356 accumulate over time correspond to oxidized fatty acids. Among these, longer-chain fatty acids
 357 accumulated more rapidly than the shorter-chain ones. At the same time, small metabolites were
 358 rapidly consumed. Metabolites from group G were less consumed than those from group F,
 359 suggesting the latter to be more easily assimilated by the microorganisms in the digesters. Regarding
 360 the FW-specific metabolites (groups I-K), these were consumed at different paces. Dipeptides were

361 first exhausted, followed by molecules with choline groups or quinolines, and finally
362 phenylpropanoids.

363 **3.3.1. Metabolomics of oak and ash**

364 Since the previous analysis did not reveal differences at the metabolomics level among ash and oak
365 bioreactors, not even at higher PCs (results not shown), these were further sought for using the O-
366 PLS-DA chemometric method (see Methods). The O-PLS-DA model composed of one orthogonal
367 and one predictive component explained the 99.8% of the Y-variance (i.e., the wood type) with 27.3%
368 of the X-variance (i.e., the metabolomic data). This strategy highlighted 55 metabolites associated
369 with ash-containing substrate and 18 with oak-containing substrate. On the one hand, oak samples
370 were found to be more enriched in furans, quinolines, and phenanthrenes. On the other hand, ash
371 samples contained higher levels of sesquiterpenoids, alkaloids, coumaric acids, and plant hormones
372 (e.g., abscisic acid), among others. In those samples, biologically active compounds such as
373 isoflavones and coumaric acid derivatives are rapidly degraded while others such as benzaldehyde or
374 sesquiterpenes are more recalcitrant (**FigS1** and **Table A.1**).

375 **4. Discussion**

376 The three types of studied data gave information about three different groups of molecules: 3D
377 fluorescence spectroscopy revealed the signature of macromolecules, HPLC metabolomics revealed
378 changes in the smaller molecules, and finally the biogas production data related to the final products
379 of the degradation of the compounds detected in those two analytical techniques. Although both 3D
380 fluorescence spectroscopy and HPLC metabolomics were used in this work because of their
381 complementarity, there is an overlap between the metabolites measured with these two techniques,
382 corresponding to the small metabolites with aromatic moieties (e.g., some amino acids and short
383 peptides, polyphenols (Ku et al., 2019)).

384 In terms of the experimental setup, all anaerobic digesters were found to be very stable since a good
385 reproducibility of the results was observed for the three types of measurements, even one month after
386 the start of the experiment.

387 Among the three data sets, the biogas production data provides the most informative results regarding
388 the overall performance of the anaerobic digesters. This may be due to the fact that the biogas
389 production data reflects the cascade of all individual reactions occurring in the digesters while both
390 the HPLC-MS metabolomics data and the 3D fluorescence data captured the alterations of more
391 limited groups of molecules (small soluble compounds for HPLC-MS metabolomics and proteins and
392 humic/fulvic-like molecules for fluorescence spectroscopy).

393 Oak is more recalcitrant than ash since biogas production was lower for oak than for ash. This was
394 also observed in the co-digestion experiment as the second peak appeared later for FW-oak than for
395 FW-ash. This is in agreement with the fact that, in terms of molecular content, oak has more lignin,
396 flavonoid and phenolic content and less polysaccharides than ash (Madrera et al., 2010; Sillero et al.,
397 2019). Despite this aspect, the biogas production of oak was slightly increased (103%) by using the
398 co-digestion approach.

399 The cellulose substrate was more degradable than oak and ash woods, as seen by the two peaks of gas
400 production appearing before the 10th day. Mixing cellulose with FW seems to have altered the order of
401 the compounds being degraded. We postulate that the degradation of cellulose alone is driven by a
402 microbial community specific for this substrate (i.e., cellulolytic) while in presence of FW, the
403 microbial community resembled more that involved in FW degradation. As a result of that,
404 compounds associated with the 2nd peak in cellulose digestion were not fully degraded in FW-
405 cellulose as the microbial community was turning to degrade the more recalcitrant compounds from
406 the FW fraction. After these compounds were consumed, compounds from cellulose continued to be
407 degraded, producing the 4th and final peak in FW-cellulose digestion. This phenomenon of continued
408 adaptation of the community to degrade the FW-cellulose mixture may have caused the levels of
409 biogas production to be around 89 % of the combined amount obtained by FW- and cellulose-mono-

410 digestions. Similarly, the daily biogas production of FW-ash AcoD as well as its overall lower biogas
411 production if compared to the combined mono-digestion counterparts (93 %) suggests that both
412 substrates are degraded by two different populations of microorganisms that are in competition with
413 each other, that the microorganisms more adapted to degrade FW dominate in the FW-ash AcoD, and
414 that at the end of the experiment the degradation of FW-ash was not completed.

415 The previous interpretations of the gas production can be confirmed by some observations made on
416 the 3D fluorescence and the HPLCMS metabolomics data.

417 Regarding the 3D fluorescence data: 1) the least degradable substrates were oak, followed by ash, as
418 hinted by the lower scores shown in IC1 and IC2 and by the higher scores in IC4 and IC5; 2) ash has
419 more protein-like content than oak (as seen by the larger scores for IC1, IC2 and IC3), as well as more
420 humic-like and lipofuscin-like content (depicted by IC5). The almost complete reduction of IC5 in ash
421 AD might have been the cause that led to the production of the big peak in IC3. 3) the opposite trend
422 observed in IC4 for cellulose AD and oak AD suggests that alpha cellulose (as in cellulose AD) can
423 be rapidly degraded but not lignocellulose (as in oak), which led to different AD outcomes; 4) AcoDs
424 gave very low scores in IC4 and IC5, suggesting that the co-digestion approach effectively aided in
425 depleting the lignocellulosic material and other compounds of low degradability from the digester; 5)
426 about the protein-like signature of the AcoDs, IC1 may represent the overall consummation of the
427 protein content in the digesters, while IC2 and IC3 may together reflect the partial degradation of the
428 protein-like content, occurring more prominently after the 14th day, as observed decreases in IC2 seem
429 to translate to equivalent increases in IC3 (Figure 2). As bigger increases occurred in IC3 for FW-ash
430 and FW-oak, we can argue that this protein-like signature comes from compounds that were not
431 accessible for the microbial community at the beginning of the digestion. The fact that these
432 signatures are not associated with the initial degradation stage suggests that the corresponding
433 compounds were not present in the liquid fraction at the beginning of the degradation, and that they
434 were gradually released during the digestion.

435 Regarding the metabolomics data: 1) ash and oak are very poorly degraded as seen from the small
436 time evolution of their scores; 2) cellulose degradation differs greatly from that of FW-cellulose since
437 their PCA trajectories evolve in different sectors of the PCA plot; 3) the degradation profile of FW-
438 cellulose, along with the other AcoD degradation profiles, resembles more the degradation profile of
439 FW, indicating that adding this more degradable co-substrate promotes the growth of the
440 microorganisms more adapted to degrade FW; 4) the addition of ash to FW AD produced a delay in
441 the molecular events leading to an incomplete degradation of the substrate, as seen by the continued
442 evolution of the PC2 scores for FW-ash AcoD even at the latest time-points.

443 In addition to information on the substrates' temporal degradabilities, the metabolomics data
444 illustrated the chemical differences of the substrates as well as some of the underlying molecular
445 degradation mechanisms occurring during AD: glycosylated compounds and other compounds from
446 group B in **Figure 5** were specific to cellulose mono-digestions, detected from day 7. Compounds
447 from group D were mainly specific to FW-containing bioreactors. Compounds from group A and C,
448 relative to more ubiquitous compounds (e.g., phosphoric acid, oxidized fatty acids) were detected in
449 all the tested conditions. Then, as seen in the analysis of section 3.3.1 regarding oak and ash woods,
450 the former was richer in compounds of high aromaticity, in line with the observed recalcitrance
451 (Marschner et al., 2008).

452 On a temporal scale, the observed metabolomics changes, either monotonic consumptions or
453 accumulations, occurred from the start of the digestion and their duration depended on the metabolite
454 levels in the substrate and on their degradability. For instance, compounds from groups F, G, I, J and
455 K were consumed over time, reaching undetectable levels for groups I and J but not for the other 3
456 groups. Besides consumption, another metabolic alteration that is suggested by the metabolomics data
457 is (lipid) oxidation.

458 Altogether, we can conclude that the investigation of the liquid fraction of the bioreactor allowed us to
459 capture the metabolic signature of the hydrolysis step in AD, where extracellular enzymes (such as
460 cellulases and proteases) produced by bacteria act on the substrates, converting water-insoluble

461 polymeric structures into smaller soluble metabolites, which microorganisms can subsequently take in
462 and use as nutrients. Metabolites described by PC1 vary less over time than those related to PC2,
463 denoting a more recalcitrant profile for the former metabolites. Regarding PC2, differences among the
464 cellulose-rich substrates were small, as shown in the PCA score plot of **Figure 4**, due to their very
465 similar chemical nature.

466 In this work we focused on the soluble fraction, but alternative metabolomics approaches could have
467 been employed for different purposes as well. For example, the metabolomics analysis of the solid
468 fraction, which contains both the substrate being degraded and the microbial load, would have
469 provided insights into how the metabolites, after being internalized, were directed towards biogas
470 production through the AD stages of acidogenesis, acetogenesis, and methanogenesis (Puig-Castellví
471 et al., 2022).

472 While we cannot directly establish a link between the metabolites detected in this study and biogas
473 production because of this lack of information about the internalized metabolites, data suggests that
474 metabolites used as substrate during early biogas production were from a more diverse mixture of
475 compounds (specifically, metabolites from groups F, G, I, J, K in **Table 1**), and that later biogas
476 production resulted from the degradation of less biodegradable compounds (metabolites from groups J
477 and K in **Table 1**, as those from groups F, G, and I had already been exhausted). Still, for some
478 substrates, the link between the metabolites and the biogas production could be drawn: The peak
479 occurring between days 4 to 10 in cellulose AD might be associated to compounds described by IC4
480 from the analysis of the 3D fluorescence data (Figure 2) and to compounds from groups G, and I-K in
481 the metabolomics data (Table 1). In both cases, compounds were consumed right before the peaks'
482 formation. Then, for the two wood ADs, their widest peak of biogas production could be mainly
483 associated to IC3 and IC5 from the analysis of the 3D fluorescence data (Figure 2) as these ICs are the
484 most important for these substrates.

485 In this paper, we have demonstrated the applicability of two molecular characterization techniques to
486 study the progress of anaerobic digestion from a biochemical perspective. These two techniques not

487 only have revealed the dynamics of the metabolite levels inside the reactors (relative to the
488 macromolecules, with 3D fluorescence spectroscopy; and to the small metabolites, with
489 metabolomics), but they also allowed us to identify the more recalcitrant metabolites. Similarly,
490 metabolomics and 3D fluorescence can be explanatory of the metabolic changes induced by
491 modifications in the operational conditions (e.g., reactor temperatures (Liu et al., 2015; Puig-Castellví
492 et al., 2022), presence of inhibitory molecules such as ammonia (Chapleur et al., 2021; Su et al.,
493 2019)). This knowledge can guide the selection of a more suitable microbial inoculum to these
494 operational conditions, or more suitable conditions in general, resulting on a more efficient substrate
495 degradation.

496 Compared to studies in the existing literature, we have proposed a data analysis strategy that allowed
497 us to have a deeper understanding of the molecular changes. On the one hand, we used ICA instead
498 of manually integrating the peaks from the 3D data (Chen et al., 2022) to identify alterations in the
499 levels of compounds with overlapping fingerprints that might have otherwise gone unnoticed.
500 Specifically, we separated the peaks related to the protein-like content (region III) into three
501 components (IC1, IC2, and IC3). On the other hand, despite the chemical complexity of the
502 substrates, we showed that the metabolomics could still be studied at the metabolite level to get
503 insight into the biochemical changes, rather than using other less informative approaches such as the
504 compounds' H/C and O/C ratios and plotting them in van Krevelen diagrams (Li et al., 2023).

505 3D fluorescence spectroscopy, due to its relatively low instrumentation cost and its high data output
506 rate (in the order of seconds), has the potential to be used in-line to routinely monitor AD (especially for
507 monitoring degradations of substrates exhibiting strong fluorescent signatures, e.g., rich in proteins
508 and/or with an abundance of humic acids) and to complement the information obtained with the
509 traditional measurements (e.g., temperature, pH, COD). Although HPLC-MS metabolomics can
510 characterize a larger number of compounds, the technique is more expensive and the time required to
511 acquire and analyze the data is greater. Because of this, metabolomics should be applied more to
512 investigate in-depth anomalous situations in digesters, that cannot be fully understood using the
513 routine measurements. This type of application is currently covered with other omics approaches,

514 such as 16S DNA metabarcoding (Cardona et al., 2022). Having said this, HPLC-MS metabolomics
515 gives a more direct insight into the degradation process than the omics approaches based on genetic
516 information since, for the latter, the metabolic alterations are not measured but inferred from the
517 modifications in the microbial community in terms of the DNA or RNA levels. Nevertheless, at
518 present, these gene-based omics approaches are considered the gold standard in the field as their
519 experimental protocols are already established. Conversely, for HPLC-MS metabolomics, their
520 application in AD is very recent and further research is crucial to characterize the chain of degradation
521 events linked to each substrate and molecular species, to better understand these processes from a
522 biological and ecological perspective, and to use this knowledge for operational purposes, such as for
523 finding monitorable metabolite biomarkers of process stability and (in)effective degradation.

524

525 **5. Conclusions**

526 3D fluorescence spectroscopy and HPLC-MS metabolomics are two complementary techniques that
527 can be used to decipher substrate degradability. We found that cellulose was more degradable than ash
528 and oak, oak being the most recalcitrant, possibly due to the presence of compounds with a high
529 aromaticity (e.g., fulvic acids).

530 Regarding AcoD, the two approaches allowed us to obtain insight into how the substrate's chemical
531 composition determines the outcome of the degradation. Adding FW to cellulose, ash wood, or oak
532 wood resulted in several modifications in their molecular fingerprints, compared to their
533 corresponding mono-digestion without the FW substrate. Specifically, the degradation profiles of the
534 molecules in the AcoDs resembled those observed in FW AD (due to the higher abundance of
535 glycerolipids, among other compounds) more than that in the digestion of the three other cellulose-
536 rich substrates (containing more glycosylated compounds, among others).

537 Both 3D fluorescence spectroscopy and HPLC-MS metabolomics captured the early molecular
538 changes in the bioreactors, which cannot be inferred from the biogas production data alone. These

539 changes could be grouped into 7 dynamics: 2 related to metabolite accumulation and 5 to metabolite
540 exhaustion, and metabolites within each group presented similarities in their chemical structure. For
541 this reason, to further investigate AcoD, more studies combining several high-throughput techniques
542 should be undertaken in the future.

543 **6. Appendices**

544 **Table A.1.** Annotation of significant metabolomics features.

545 **Fig S1.** Heatmap of the selected features in the O-PLS-DA of ash vs oak samples.

546 **Acknowledgements**

547 We thank Christophe Cordella and Nathalie Locquet from AgroParisTech for providing us with
548 access to the spectrofluorometer used in this study.

549 **Declaration of competing interest**

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

552

553 **References**

- 554 Awhangbo, L., Bendoula, R., Roger, J.M., Béline, F., 2020. Detection of early imbalances in semi-
555 continuous anaerobic co-digestion process based on instantaneous biogas production rate.
556 *Water Research* 171, 115444. <https://doi.org/10.1016/j.watres.2019.115444>
557 Bhatia, S.K., Joo, H.-S., Yang, Y.-H., 2018. Biowaste-to-bioenergy using biological methods – A
558 mini-review. *Energy Conversion and Management* 177, 640–660.
559 <https://doi.org/10.1016/j.enconman.2018.09.090>
560 Borowski, S., Kubacki, P., 2015. Co-digestion of pig slaughterhouse waste with sewage sludge. *Waste*
561 *Management* 40, 119–126. <https://doi.org/10.1016/j.wasman.2015.03.021>

- 562 Braguglia, C.M., Gallipoli, A., Gianico, A., Pagliaccia, P., 2018. Anaerobic bioconversion of food
563 waste into energy: A critical review. *Bioresource Technology*, Bioconversion of Food Wastes
564 248, 37–56. <https://doi.org/10.1016/j.biortech.2017.06.145>
- 565 Cardona, L., Mazéas, L., Chapleur, O., 2022. Deterministic processes drive the microbial assembly
566 during the recovery of an anaerobic digester after a severe ammonia shock. *Bioresource*
567 *Technology* 347, 126432. <https://doi.org/10.1016/j.biortech.2021.126432>
- 568 Chapleur, O., Poirier, S., Guenne, A., Lê Cao, K.-A., 2021. Time-course analysis of metabolomic and
569 microbial responses in anaerobic digesters exposed to ammonia. *Chemosphere* 283, 131309.
570 <https://doi.org/10.1016/j.chemosphere.2021.131309>
- 571 Chatterjee, B., Mazumder, D., 2019. Role of stage-separation in the ubiquitous development of
572 Anaerobic Digestion of Organic Fraction of Municipal Solid Waste: A critical review.
573 *Renewable and Sustainable Energy Reviews* 104, 439–469.
574 <https://doi.org/10.1016/j.rser.2019.01.026>
- 575 Chen, Y., Ping, Q., Li, D., Dai, X., Li, Y., 2022. Comprehensive insights into the impact of
576 pretreatment on anaerobic digestion of waste active sludge from perspectives of organic
577 matter composition, thermodynamics, and multi-omics. *Water Research* 226, 119240.
578 <https://doi.org/10.1016/j.watres.2022.119240>
- 579 Cruz, I.A., Andrade, L.R.S., Bharagava, R.N., Nadda, A.K., Bilal, M., Figueiredo, R.T., Ferreira,
580 L.F.R., 2021. An overview of process monitoring for anaerobic digestion. *Biosystems*
581 *Engineering* 207, 106–119. <https://doi.org/10.1016/j.biosystemseng.2021.04.008>
- 582 Dettmer, K., Aronov, P.A., Hammock, B.D., 2007. Mass spectrometry-based metabolomics. *Mass*
583 *Spectrometry Reviews* 26, 51–78. <https://doi.org/10.1002/mas.20108>
- 584 Du, Z., Zhao, P., Fu, Q., Wang, Q., Hu, A., Zhang, W., Wang, D., 2023. Biostimulants in dissolved
585 organic matters recovered from anaerobic digestion sludge with alkali-hydrothermal
586 treatment: Nontarget identification by ultrahigh-resolution mass spectrometry. *Environment*
587 *International* 173, 107813. <https://doi.org/10.1016/j.envint.2023.107813>
- 588 EEA, 2020. Bio-waste in Europe — turning challenges into opportunities. EEA Report No 4/2020,
589 European Environment Agency.
- 590 Iacovidou, E., Ohandja, D.-G., Voulvoulis, N., 2012. Food waste co-digestion with sewage sludge –
591 Realising its potential in the UK. *Journal of Environmental Management* 112, 267–274.
592 <https://doi.org/10.1016/j.jenvman.2012.07.029>
- 593 Jouan-Rimbaud Bouveresse, D., Moya-González, A., Ammari, F., Rutledge, D.N., 2012. Two novel
594 methods for the determination of the number of components in independent components
595 analysis models. *Chemometrics and Intelligent Laboratory Systems* 112, 24–32.
596 <https://doi.org/10.1016/j.chemolab.2011.12.005>
- 597 Kassouf, A., Jouan-Rimbaud Bouveresse, D., Rutledge, D.N., 2018. Determination of the optimal
598 number of components in independent components analysis. *Talanta* 179, 538–545.
599 <https://doi.org/10.1016/j.talanta.2017.11.051>
- 600 Ku, Y.G., Kim, H.C., Bae, J.H., Kang, B.S., Nemirovski, A., Barasch, D., Gorinstein, S., 2019.
601 Antioxidant capacities and polyphenols in autumn-growing cultivar of Chinese cabbage
602 (*Brassica rapa* L. ssp. *pekinensis* cv. Bulam Plus). *Eur Food Res Technol* 245, 1871–1879.
603 <https://doi.org/10.1007/s00217-019-03294-0>
- 604 Li, X., Huang, X., Zhao, C., Wang, X., Dong, B., Goonetilleke, A., Kim, K.-H., 2023. Characterizing
605 molecular transformation of dissolved organic matter during high-solid anaerobic digestion of
606 dewatered sludge using ESI FT-ICR MS. *Chemosphere* 320, 138101.
607 <https://doi.org/10.1016/j.chemosphere.2023.138101>
- 608 Liu, Y., Li, X., Kang, X., 2015. Effect of volume ratio on anaerobic co-digestion of thermal
609 hydrolysis of food waste with activated sludge. *International Biodeterioration &*
610 *Biodegradation*, CESE-2014 – Challenges in Environmental Science and Engineering Series
611 Conference 102, 154–158. <https://doi.org/10.1016/j.ibiod.2015.02.015>
- 612 Madrera, R.R., Valles, B.S., García, Y.D., del Valle Argüelles, P., Lobo, A.P., 2010. Alternative
613 woods for aging distillates-an insight into their phenolic profiles and antioxidant activities.
614 *Food Sci Biotechnol* 19, 1129–1134. <https://doi.org/10.1007/s10068-010-0161-4>

615 Madsen, M., Holm-Nielsen, J.B., Esbensen, K.H., 2011. Monitoring of anaerobic digestion processes:
616 A review perspective. *Renewable and Sustainable Energy Reviews* 15, 3141–3155.
617 <https://doi.org/10.1016/j.rser.2011.04.026>

618 Marschner, B., Brodowski, S., Dreves, A., Gleixner, G., Gude, A., Grootes, P.M., Hamer, U., Heim,
619 A., Jandl, G., Ji, R., Kaiser, K., Kalbitz, K., Kramer, C., Leinweber, P., Rethemeyer, J.,
620 Schäffer, A., Schmidt, M.W.I., Schwark, L., Wiesenberg, G.L.B., 2008. How relevant is
621 recalcitrance for the stabilization of organic matter in soils? *Journal of Plant Nutrition and*
622 *Soil Science* 171, 91–110. <https://doi.org/10.1002/jpln.200700049>

623 Maynaud, G., Druilhe, C., Daumoin, M., Jimenez, J., Patureau, D., Torrijos, M., Pourcher, A.-M.,
624 Wéry, N., 2017. Characterisation of the biodegradability of post-treated digestates via the
625 chemical accessibility and complexity of organic matter. *Bioresource Technology* 231, 65–74.
626 <https://doi.org/10.1016/j.biortech.2017.01.057>

627 Moretti, P., Morais de Araujo, J., Borges de Castilhos, A., Buffière, P., Gourdon, R., Bayard, R.,
628 2020. Characterization of municipal biowaste categories for their capacity to be converted
629 into a feedstock aqueous slurry to produce methane by anaerobic digestion. *Science of The*
630 *Total Environment* 716, 137084. <https://doi.org/10.1016/j.scitotenv.2020.137084>

631 Muller, M., Jimenez, J., Antonini, M., Dudal, Y., Latrille, E., Vedrenne, F., Steyer, J.-P., Patureau, D.,
632 2014. Combining chemical sequential extractions with 3D fluorescence spectroscopy to
633 characterize sludge organic matter. *Waste Management* 34, 2572–2580.
634 <https://doi.org/10.1016/j.wasman.2014.07.028>

635 Peris-Díaz, M.D., Krężel, A., 2021. A guide to good practice in chemometric methods for vibrational
636 spectroscopy, electrochemistry, and hyphenated mass spectrometry. *TrAC Trends in*
637 *Analytical Chemistry* 135, 116157. <https://doi.org/10.1016/j.trac.2020.116157>

638 Puig-Castellví, F., Cardona, L., Jouan-Rimbaud Bouveresse, D., Cordella, C.B.Y., Mazéas, L.,
639 Rutledge, D.N., Chapleur, O., 2020. Assessment of substrate biodegradability improvement in
640 anaerobic Co-digestion using a chemometrics-based metabolomic approach. *Chemosphere*
641 254, 126812. <https://doi.org/10.1016/j.chemosphere.2020.126812>

642 Puig-Castellví, F., Midoux, C., Guenne, A., Conteau, D., Franchi, O., Bureau, C., Madigou, C., Jouan-
643 Rimbaud Bouveresse, D., Kroff, P., Mazéas, L., Rutledge, D.N., Gaval, G., Chapleur, O.,
644 2022. Metataxonomics, metagenomics and metabolomics analysis of the influence of
645 temperature modification in full-scale anaerobic digesters. *Bioresource Technology* 346,
646 126612. <https://doi.org/10.1016/j.biortech.2021.126612>

647 Rodriguez-Narvaez, O.M., Peralta-Hernandez, J.M., Goonetilleke, A., Bandala, E.R., 2017. Treatment
648 technologies for emerging contaminants in water: A review. *Chemical Engineering Journal*
649 323, 361–380. <https://doi.org/10.1016/j.cej.2017.04.106>

650 Rorat, A., Courtois, P., Vandenbulcke, F., Lemiere, S., 2019. Sanitary and environmental aspects of
651 sewage sludge management. *Industrial and Municipal Sludge* 155–180.
652 <https://doi.org/10.1016/B978-0-12-815907-1.00008-8>

653 Rutledge, D.N., Jouan-Rimbaud Bouveresse, D., 2013. Independent Components Analysis with the
654 JADE algorithm. *TrAC Trends in Analytical Chemistry* 50, 22–32.
655 <https://doi.org/10.1016/j.trac.2013.03.013>

656 Sillero, L., Prado, R., Andrés, M.A., Labidi, J., 2019. Characterisation of bark of six species from
657 mixed Atlantic forest. *Industrial Crops and Products* 137, 276–284.
658 <https://doi.org/10.1016/j.indcrop.2019.05.033>

659 Slorach, P.C., Jeswani, H.K., Cuéllar-Franca, R., Azapagic, A., 2019. Environmental sustainability of
660 anaerobic digestion of household food waste. *Journal of Environmental Management* 236,
661 798–814. <https://doi.org/10.1016/j.jenvman.2019.02.001>

662 Su, C., Zhao, L., Liao, L., Qin, J., Lu, Y., Deng, Q., Chen, M., Huang, Z., 2019. Application of
663 biochar in a CIC reactor to relieve ammonia nitrogen stress and promote microbial
664 community during food waste treatment. *Journal of Cleaner Production* 209, 353–362.
665 <https://doi.org/10.1016/j.jclepro.2018.10.269>

666 Thévenot, E.A., Roux, A., Xu, Y., Ezan, E., Junot, C., 2015. Analysis of the Human Adult Urinary
667 Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a
668 Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* 14,
669 3322–3335. <https://doi.org/10.1021/acs.jproteome.5b00354>

670 Wishart, D.S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S.,
671 Lee, B.L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-
672 Giesbrecht, M., Lui, V.W., Varshavi, Dorna, Varshavi, Dorsa, Allen, D., Arndt, D.,
673 Khetarpal, N., Sivakumaran, A., Harford, K., Sanford, S., Yee, K., Cao, X., Budinski, Z.,
674 Liigand, J., Zhang, L., Zheng, J., Mandal, R., Karu, N., Dambrova, M., Schiöth, H.B.,
675 Greiner, R., Gautam, V., 2022. HMDB 5.0: the Human Metabolome Database for 2022.
676 *Nucleic Acids Research* 50, D622–D631. <https://doi.org/10.1093/nar/gkab1062>
677 Zepp, R.G., Sheldon, W.M., Moran, M.A., 2004. Dissolved organic fluorophores in southeastern US
678 coastal waters: correction method for eliminating Rayleigh and Raman scattering peaks in
679 excitation–emission matrices. *Marine Chemistry, CDOM in the Ocean: Characterization,*
680 *Distribution and Transformation* 89, 15–36. <https://doi.org/10.1016/j.marchem.2004.02.006>
681 Zhao, P., Du, Z., Fu, Q., Ai, J., Hu, A., Wang, D., Zhang, W., 2023. Molecular composition and
682 chemodiversity of dissolved organic matter in wastewater sludge via Fourier transform ion
683 cyclotron resonance mass spectrometry: Effects of extraction methods and electrospray
684 ionization modes. *Water Research* 232, 119687. <https://doi.org/10.1016/j.watres.2023.119687>

685

686

687

688