



**HAL**  
open science

# Benefit of the Use of GCxGC/MS Profiles for 1D GC/MS Data Treatment Illustrated by the Analysis of Pyrolysis Products from East Asian Handmade Papers

Bin Han, Silvia Lob, Michel Sablier

► **To cite this version:**

Bin Han, Silvia Lob, Michel Sablier. Benefit of the Use of GCxGC/MS Profiles for 1D GC/MS Data Treatment Illustrated by the Analysis of Pyrolysis Products from East Asian Handmade Papers. Journal of The American Society for Mass Spectrometry, 2018, 29 (8), pp.1582 - 1593. 10.1007/s13361-018-1953-7 . hal-01916660

**HAL Id: hal-01916660**

**<https://hal.science/hal-01916660>**

Submitted on 18 Oct 2023

**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.

Benefit of the use of GCxGC/MS profiles for 1D GC/MS data treatment  
illustrated by the analysis of pyrolysis products from East Asian  
handmade papers

Bin Han, Silvia Lob, Michel Sablier\*

Sorbonne Universités, Centre de Recherche sur la Conservation (USR 3224), Muséum national  
d'Histoire naturelle, Ministère de la Culture et de la Communication, CNRS. 75005 Paris, France

\*Corresponding authors

Email: [michel.sablier@mnhn.fr](mailto:michel.sablier@mnhn.fr)

Address reprint requests to Dr. Michel Sablier

Tel: +33 1 40 79 53 23

Fax: +33 1 40 59 53 12

E-mail: [michel.sablier@mnhn.fr](mailto:michel.sablier@mnhn.fr)

Address : Centre de recherche sur la conservation (CRC, USR 3224), Sorbonne Universités,  
Muséum national d'Histoire naturelle, Ministère de la Culture et de la Communication, CNRS ;  
CP21, 36 rue Geoffroy-Saint-Hilaire, 75005 Paris, France

## **Abstract**

In this study, we report the use of pyrolysis-GCxGC/MS profiles for an optimized treatment of data issued from pyrolysis-GC/MS combined with the automatic deconvolution software Automated Mass Spectral Deconvolution and Identification System (AMDIS). The method was illustrated by the characterization of marker compounds of East Asian handmade papers through the examination of pyrolysis-GCxGC/MS data to get information which was used for manually identifying low concentrated and co-eluting compounds in 1D GC/MS data. The results showed that the merits of a higher separation power for co-eluting compounds and a better sensitivity for low concentration compounds offered by a GCxGC system can be used effectively for AMDIS 1D GC/MS data treatment: i) the compound distribution in pyrolysis-GCxGC/MS profiles can be used as 'peak finder' for manual check of low concentration and co-eluting compound identification in 1D GC/MS data; ii) pyrolysis-GCxGC/MS profiles can provide better quality mass spectra with observed higher match factors in the AMDIS automatic match process. The combination of 2D profile with AMDIS was shown to contribute efficiently to a better characterization of compound profiles in the chromatograms obtained by 1D analysis in focusing on the mass spectral identification.

**Keywords:** pyrolysis; comprehensive two dimensional gas chromatography; AMDIS; data processing; East Asian handmade papers

# 1 **Introduction**

2 Two problems are currently pointed out for 1D GC/MS workflow analysis: one concerns the co-  
3 elution of compounds and the other concerns the detection of analytes of low concentration  
4 within the inherent background. One effective way to identify low concentrated and co-eluting  
5 compounds consists in mass spectra deconvolution. The deconvolution process allows the  
6 separation of overlapping mass spectra into individual pure-mass spectra, a practical way to  
7 unmix co-eluting compounds. Several applications have been developed for mass interpretation  
8 with the help of mathematical deconvolution employing different computing environment:  
9 Microsoft Visual Basic [1], R [2, 3] or Matlab [4]. Among them, the Automated Mass Spectral  
10 Deconvolution and Identification System (AMDIS), developed by the National Institute of  
11 Standards and Technology (NIST), is certainly the oldest stand-alone software package freely  
12 available for the routine processing of GC/MS data [5]. It was originally intended to provide  
13 unambiguous identification through deconvolution, algorithms-based quality matching,  
14 background subtraction as well as retention index comparison [6].

15 Deconvolution, the basis of AMDIS approach, is a critical step in the sense that deconvolution  
16 will reconstruct a pure mass spectrum for each eluted compound detected by the mass  
17 spectrometer. This deconvolution process is designed for the identification of target compounds  
18 in complex matrices. However, if extracted spectra do not match the target, the corresponding  
19 compound will be retained only if it shows dissimilar characteristics to the extracted nearby  
20 spectra. As a consequence, a risk consists in the fact that there are many extracted spectra that do  
21 not represent true chemical compounds but that are due to the effect of overlapping or the  
22 presence of noise in the collected chromatograms. Post-processing AMDIS filtering procedure

23 have been subsequently proposed to circumvent these drawbacks and to limit the problem of co-  
24 eluting peaks [7].

25 AMDIS works with an identity search approach: the unknowns are searched with a reference  
26 library, generally a home-made library in connection with the routine workflow. This procedure,  
27 at the opposite of a similarity search, requires the utmost effectiveness of the deconvolution of  
28 mass spectra issued from the chromatograms. A serious limitation resides in the difficulty to  
29 distinguish clearly between compounds belonging to the same family or isomer structures. A key  
30 feature of AMDIS is that it can automatically identify components from a complex spectrum by  
31 comparing the mass spectra with the so-called reference libraries to resolve the problem of  
32 complex mixtures [8]. However, AMDIS may produce false peaks and may require optimized  
33 settings for an efficient use [9, 10] and was reported to frequently misidentify peaks (in  
34 generating false positives) or do not recognize compounds resulting in false negative  
35 assignments [11].

36 Nevertheless, AMDIS can provide fast and accurate results for identification and quantification  
37 of GC/MS results [12, 13], and compares well for its ability in GC/MS routine analysis with  
38 other commercially or freely available deconvolution software platforms [9]. In particular, since  
39 its introduction, AMDIS has demonstrated its ability to detect target compounds in matrices with  
40 high backgrounds [9]. The protocols for its application in GC/MS workflow has been fully  
41 explained elsewhere [14]. One must concede that with the development of metabolomics in  
42 GC/MS, among software platforms for untargeted metabolomics based on GC-MS rapid  
43 annotation using in-house databases, AMDIS competes well with the other home-made software  
44 allowing for much tighter control and fine tuning than commercial software often characterized  
45 by their “black box” effects. Automatic searches for identification of unknowns from current

46 libraries are still a vivid domain of development since most algorithms require constant  
47 optimization for improving their effectiveness [8].

48 Another way to solve the problem of co-eluting and low concentration compounds is by adding  
49 another dimension for separation. GC/MS is now a well-established method, while the seeking  
50 for higher separation and resolution has been partly fulfilled by multidimensional gas  
51 chromatography. On the criterion of 1D GC separation (e.g., different volatilities), co-elution of  
52 compounds occurred when the vapor pressures of analytes do not differ sufficiently from each  
53 other. In practice, the peak capacity of column (e.g., 30 meters ) in a 1D GC is in hundreds range  
54 while values of peak capacity over 1000 is hard to attain even at the sacrifice of instrumental  
55 complexity and acquisition time. The introduction of 2D GC in 1991 largely accomplish the  
56 initial goal of separating these co-eluting compounds posed by 1D GC separation [15]. The setup  
57 of 2D GC is well know now which employs two columns of different polarity via a modulator.  
58 The entire eluent from the first column is subjected to reinjection in a second column through the  
59 modulator in a designed continuous and real-time mode. A comprehensive two-dimensional gas  
60 chromatography/mass spectrometry (GC×GC/MS) system offers much higher resolution, peak  
61 separation capacity, selectivity and lower detection limit for the analysis of volatile organic  
62 molecules [16, 17]. As a result of the separation feature and the re-concentration effect of the  
63 modulation step, peak capacity over 1000 is easily obtained [18, 19]. The merits of improved  
64 peak capacity and structure-based grouping of 2D GC over 1D GC has been extensively reported  
65 [20] (and references cited therein). One may expect a growing interest for GCxGC/MS in various  
66 fields, and one may quote metabolomics where huge set of data are compiled as an example of  
67 its likely application.

68 Although 2D GC has advantages, 1D GC is currently the predominant and effective analytical  
69 method within laboratories all over the world. Improved knowledge about the composition of co-  
70 eluting compounds or low concentration compounds reserved in the testing sample demands  
71 enhanced resolution. And conversely, the enhanced resolution can give insight for mass spectra  
72 analysis reserved in a lower resolution data [21]. A fact is that, until recently, routine analysis in  
73 many fields was relying on low resolution mass spectrometry, the larger market in volume for  
74 manufacturers: the aforementioned approach of reintroducing optimized data obtained through  
75 2D GC/MS into 1D GC/MS may present interesting applications for the optimized use of low  
76 resolution mass spectrometers in large workflow analysis. For these reasons, the evaluation of  
77 the chromatographic and mass spectra information revealed by 2D GC/MS profile that can be  
78 favorably used for 1D GC/MS data mining and biomarker discovery within AMDIS deserves to  
79 be investigated.

80 The aim of the present study is to scrutinize how the improved knowledge obtained by Pyrolysis-  
81 GCxGC/MS profile can be beneficial for compound identification in a 1D GC/MS data  
82 collection process like the routine procedure of using AMDIS. The analysis was focused on low  
83 concentration and co-eluting compounds analysis, and focuses on the identification of the  
84 pyrolysis fingerprint of East Asian handmade paper samples. The method was firstly carried out  
85 with the examination of Py-GCxGC/MS data to generate a 2D profile for manual identification  
86 of low concentrated and co-eluting compounds in 1D GC/MS data. Then, an in-house AMDIS  
87 library was constructed using the mass spectra gained from a GGxGC/MS analysis and  
88 corresponding 1D GC/MS profile and their performance in the AMDIS automatic identification  
89 process was evaluated.

90

## 91 **Experimental**

### 92 *Material*

93 Material used in this study consisted of handmade paper made from kozo (*Broussonetia kazinoki*  
94 Sieb.) prepared following the Japanese traditional processes. Kozo is a plant native to Asia and  
95 its inner bark fiber is commonly used in East Asian hand papermaking which accounted for a  
96 significant fiber resources in the history of hand papermaking.

97

### 98 *Py-GC/MS Analysis*

99 Py-GC/MS analyses were performed using a vertical micro-furnace-type pyrolyzer PY-2020iD  
100 (Frontier Lab, Fukushima, Japan) directly connected to the injection port of a Shimadzu QP2010  
101 gas chromatograph linked with a Shimadzu QP2010Plus quadrupole mass spectrometer  
102 (Shimadzu, Champs-sur-Marne, France). The sample was placed in a stainless steel sample cup.  
103 The paper samples were weighted with an XP2U Ultra Micro Balance (Mettler Toledo, Viroflay,  
104 France). Typical quantities of handmade paper used in this work were ca. 100 µg weighted with  
105 the microbalance. The sample cup was placed on top of the pyrolyzer at near ambient  
106 temperature. The sample cup was introduced into the furnace at 500 °C, and then the temperature  
107 program of the gas chromatograph oven was started. The Py-GC interface was held at 320 °C.  
108 Chromatographic separation was carried out on a Frontier Lab DB5-ms 5% phenyl–95%  
109 dimethyl polysiloxane fused silica capillary column (30 m length, 0.25 mm inner diameter and  
110 coated with a 0.25 µm film thickness). The oven temperature was initially held 3 min at 40 °C,  
111 and then ramped at 5 °C min<sup>-1</sup> to 325 °C, where it was held for 10 min. The total duration of GC

112 analysis was 70 min. The helium carrier gas Alphagaz 1 (Air Liquide, France) was used in the  
113 linear velocity mode ( $1 \text{ mL min}^{-1}$ ). The injector was held at  $280^\circ\text{C}$  and used in split mode (1:10  
114 of the total flow). The mass spectrometer was operated at  $5000 \text{ u.s}^{-1}$ , with a scan range from 50  
115 to 500 u, using electron ionization at 70 eV. The interface was kept at  $300^\circ\text{C}$  and the MS source  
116 at  $200^\circ\text{C}$ . The mass spectra of the pyrolysis products were recorded using the average spectrum  
117 + average and subtract spectrum modes in the Shimadzu GCMS Solution software.  
118 Identifications were previously achieved on the basis of EI mass spectra by interpretation of the  
119 main fragmentations and using the NIST MS library (2011) [22].

120

### 121 *Py-GCxGC/MS Analysis*

122 Analysis was conducted with a Shimadzu QP 2010-Ultra gas chromatograph mass spectrometer  
123 (Shimadzu, Champs-sur-Marne, France) equipped with a two-stage thermal modulator ZX 2  
124 (Zoex, Houston, USA). Pyrolysis was performed using a vertical micro-furnace-type pyrolyzer  
125 PY-3030iD (Frontier Lab, Fukushima, Japan) directly connected to the injection port of the gas  
126 chromatograph. The sample was placed in a stainless steel sample cup. Typical quantities of  
127 sample paper used in this work were ca.  $50 \mu\text{g}$  weighted with an XP2U Ultra Micro Balance  
128 (Mettler Toledo, Viroflay, France). The pyrolysis temperature was defined to be  $500^\circ\text{C}$ . The  
129 pyrolyzer interface was held at  $320^\circ\text{C}$ . An OPTIMA-5HT column ( $30\text{m} \times 0.25\text{mm}$  I.D.,  $0.25\mu\text{m}$   
130 film thickness, Macherey-Nagel, Hoerd, France) was used as first dimension column and a  
131 Zebron ZB-50 ( $2.8\text{m} \times 0.1\text{mm}$  I.D.,  $0.1\mu\text{m}$  film thickness, Phenomenex, Le Pecq, France) was  
132 used as a second-dimension column and for the loop modulator system. The separation was  
133 carried out at a constant pressure of 300 kPa using Helium Alphagaz 1 as carrier gas (Air

134 Liquide, Bagneux, France). The ZX 2 two-stage thermal modulator was utilized with a  
135 modulation period of 10 s with a programmed hot pulse of 0.350 s. A two-step temperature  
136 program was used for the hot jet set at 200 °C for 30 min and subsequently raised to 280 °C. The  
137 oven temperature was initially held 1 min at 100 °C, and then ramped at 2 °C min<sup>-1</sup> to 325 °C,  
138 where it was held for 25 min. The injector was held at 280°C and used in split mode (1:30 of the  
139 total flow). The mass spectrometer was operated at 20,000 u.s<sup>-1</sup>, with a scan range from 50 to  
140 500 u, using electron ionization at 70 eV. The interface was kept at 280 °C and the ion source at  
141 200 °C.

142

#### 143 *Data treatment*

144 Data processing of the Py-GCxGC/MS raw data was achieved using GC Image software, version  
145 2.4 (Lincoln, Nebraska). Identification of paper components was performed by comparing the  
146 mass spectra of unknown components with the ones of the NIST mass spectra library, 2011  
147 edition. More information about the identification refers to [22, 23]. The mass spectra from  
148 GCxGC profile and 1D profile were extracted using GC Image software, version 2.4 (Lincoln,  
149 Nebraska), GCMS Postrun Analysis (a software build within Shimadzu analysis platform) and  
150 AMDIS. The procedure is provided in Supplementary Figure S3. The deconvolution of mass  
151 spectra was carried out triplicate within AMDIS for the evaluation of the quality of the mass  
152 spectra from GCxGC profile and 1D profile. **In order to evaluate the efficiency of the Py-  
153 GCxGC/MS approach, specific characteristics of the AMDIS software like filtering procedure  
154 were set at default value (“filtering off”) during the treatment of data.** As AMDIS cannot handle

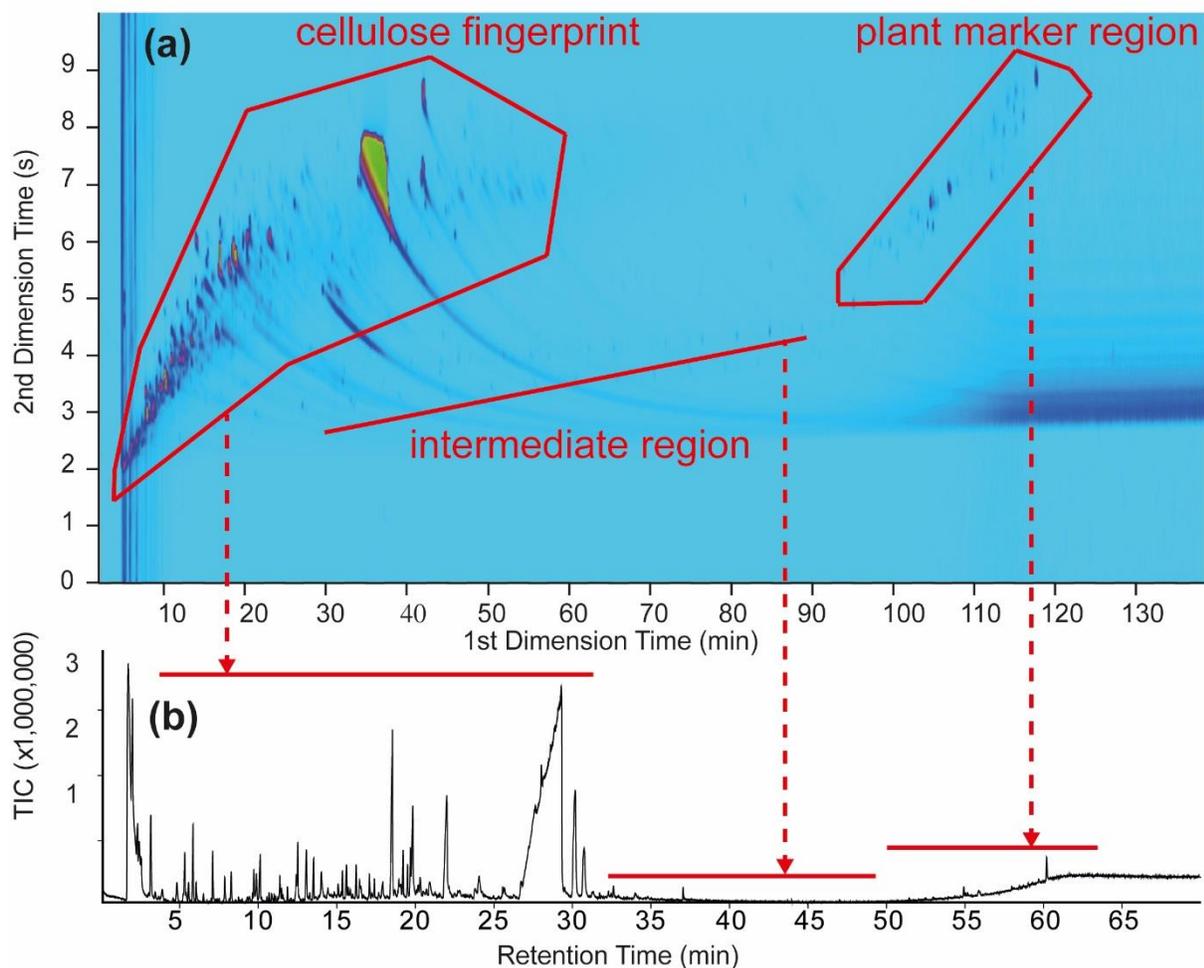
155 Shimadzu data format (.qgd), the data acquired by the Shimadzu platform was previously saved  
156 as CDF format (.cdf) by the GCMS Postrun Analysis software.

157

## 158 **Results and discussion**

159 *The pros and cons of AMDIS deconvolution for Py-GC/MS treatment of handmade*  
160 *paper pyrolysis products*

161 In our study for characterization of East Asian handmade papers using pyrolysis-gas  
162 chromatography and mass spectrometry, the identification of pyrolysis product of handmade  
163 papers were performed with both Py-GC/MS and Py-GCxGC/MS analysis and the  
164 chromatograms were shown in Figure 1. It can be seen that the GCxGC/MS and 1D GC/MS  
165 profiles correspond to each other according to their eluting order in the first dimension of  
166 separation. The part of the chromatogram associated with the cellulose pyrolysis products  
167 presents high similarities and the latter part (plant marker region) in both chromatograms present  
168 composition differences which can be used for tracing fiber origins [22-24].

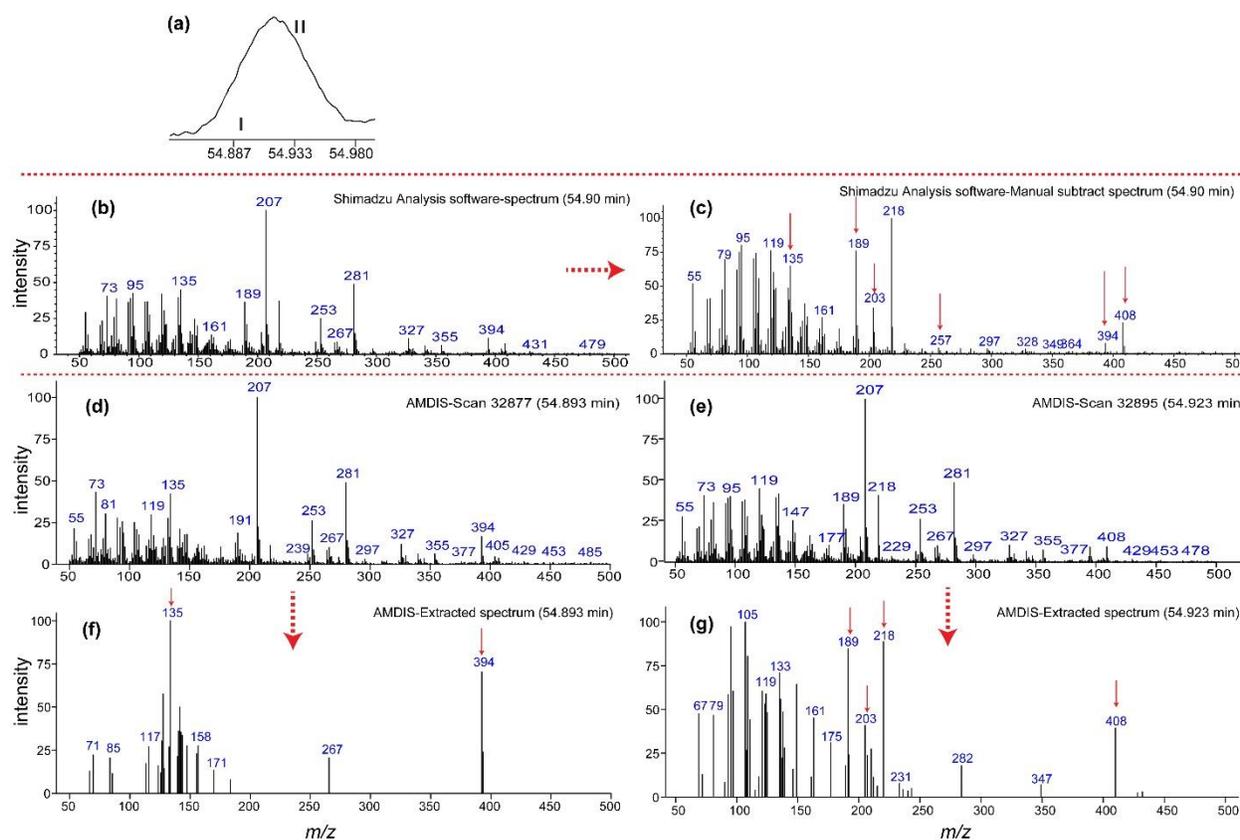


169

170 **Figure 1.** Py-GCxGC/MS (a) and Py-GC/MS (b) chromatograms obtained from the pyrolysis of  
 171 a kozo handmade paper sample

172 The manual examination of the mass spectra of the plant marker region in Figure 1 was first  
 173 carried out treating Py-GC/MS data with the commercially dedicated software GCMS Postrun  
 174 Analysis (GC/MS Postrun Analysis is the commercial software build within the Shimadzu  
 175 analysis platform) using subtraction of background from nearby baseline for peak investigation.  
 176 For example, the base peak mass spectra,  $m/z$  218 (Figure 2c), generated by a retro Diels Alder  
 177 fission of the triterpenic ring structures was considered here as a characteristic fragment ion in  
 178 the mass spectra of such structures [25-27]. In the mass spectra of triterpenes, comparison

179 between fragment peak intensities at  $m/z$  203 and  $m/z$  189 allows the distinction between the  
 180 oleanane and ursane skeleton [26]: the fragment ion at  $m/z$  203 is more intense than the peak at  
 181  $m/z$  189 for the 12-oleanene derivatives series, while usually the reverse occurs in the mass  
 182 spectra of an 12-ursene derivative (fragment peaks at  $m/z$  203 and  $m/z$  189 may have similar  
 183 intensities in that case). Variation of the ratios between the fragment ions at  $m/z$  203 and  $m/z$  189  
 184 may be considered as a second criterion for the identification of ursane and oleanane structures.  
 185 Mass spectra analysis following this procedure permits to assign the peak (compound II in Figure  
 186 2a) as an ursane structure (urs-2,12-diene, MW 408) with a base peak of  $m/z$  218 and the  
 187 corresponding ratio between  $m/z$  189 and  $m/z$  203 in favor of  $m/z$  189 (Figure 2b and Figure 2c).  
 188 However, the presence of peaks at  $m/z$  135, 394 was noticed. This latter peak might be assigned  
 189 to co-eluting compound(s) presenting another skeleton structure than the previous triterpene.



190

191 **Figure 2.** The examined peak issued from Py-GC/MS analysis (a) with its corresponding  
192 scanned mass spectrum (b) and manually subtracted spectrum (c) using the GCMS Postrun  
193 Analysis software (Shimadzu) for peak identification. AMDIS automatic mass spectra  
194 deconvolution was conducted and resulted in the scanned spectra (d and e) and automatic  
195 extracted spectra (f and g)

196 Besides the commercially dedicated software, AMDIS provide a way of deconvolution for mass  
197 examination. To illustrate this point, AMDIS deconvolution was carried out for examination of  
198 co-eluting compounds around compound II in Figure 2a. As is known, the AMDIS  
199 deconvolution ability largely rely on the analyze settings. The following deconvolution settings  
200 in AMDIS with shape requirements set as ‘medium’, resolution set as ‘high’, and sensitivity set  
201 as ‘very high’ were applied here. Under these conditions, AMDIS generated ca. 50 peaks within  
202 the retention time range 54.8 min–55.1 min. Among these compounds, most of them revealed to  
203 be false peaks, but a compound assigned to a stigmastane structure was located and identified  
204 (compound I, Figure 2d, Figure 2f) together with the previously identified peak in GCMS  
205 Postrun Analysis software (Shimadzu) (Figure 2e, Figure 2g). Actually, the number of detected  
206 compounds with AMDIS can be found dependent on the chosen analysis settings. With the  
207 parameters: resolution, sensitivity and shape requirement all set as ‘medium’ for deconvolution,  
208 AMDIS deconvolution revealed only five peaks in the same region. Of the five automated  
209 identified peaks, only the compound urs-2,12-diene was detected and other extracted spectra  
210 were not conclusive for compound identification.

211 In the above example from the pyrolysis product of East Asian handmade papers, AMDIS  
212 demonstrated its ability to locate the co-eluting compounds but the drawbacks of AMDIS lies in  
213 the fact that: i) the number of returned compounds by AMDIS largely depends on the analysis

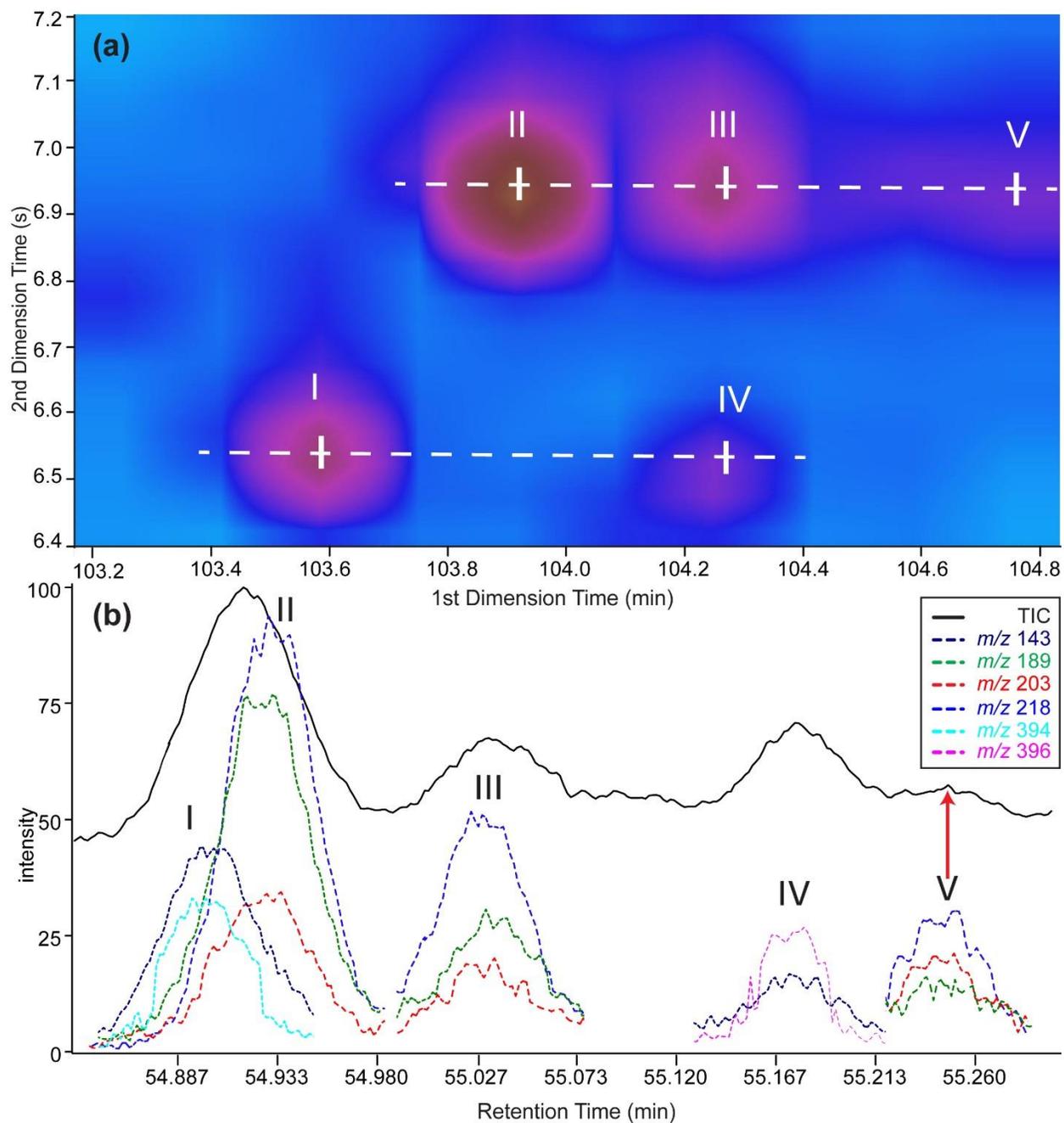
214 settings; ii) when using lower resolution and lower sensitivity settings for analysis of GC/MS  
215 data, AMDIS produces less peaks to check but poses the risk of losing the co-eluting peaks; iii)  
216 while when using optimized settings, AMDIS produces many false peaks and checking the  
217 AMDIS identified peaks is tedious and time-consuming. The observed false peaks were in  
218 accordance with the previous report that AMDIS may produce artifact components  
219 corresponding to noise or give multiple peak assignments that correspond to a single  
220 chromatographic peak [9].

221 The separation of compounds in mixtures is the main aim of GC/MS, meanwhile, an efficient  
222 separation is a guarantee for a reliable mass spectral identification. However, as observed in  
223 routine 1D workflow analysis, it is not always easy to obtain an ideally 'pure' mass spectrum in  
224 consideration of mutual interference from other mass spectra (chromatographic peaks  
225 overlapping) as well as interference with the background. Even by adding another dimension for  
226 further separation, compound co-elution is still predicted to occur and can be observed in real  
227 applications [28, 29]. Commercial mass spectrometer based platform software usually propose  
228 solutions for identifying and integrating compounds based on peak area or height using total ion  
229 count (TIC), or peaks from the extracted ion chromatograms (EICs) by selecting  $m/z$   
230 characteristic values for given compounds. Meanwhile, most of the currently applied manual  
231 identification approaches use simple subtraction of background from nearby baseline or a  
232 shoulder of a given peak of interest: an approach that can reveal insufficient to handle  
233 overlapping and/or co-eluting peaks. For this reason, the increased resolution capabilities of the  
234 analytical method could clearly constitute a plus for the AMDIS automatic data treatment as will  
235 be discussed below.

236

237 *The benefit of GCxGC profile use for 1D data treatment*

238 Our previous study revealed that by using Py-GCxGC/MS, a factor of gain in sensitivity of ca. 6  
239 can be expected for the analysis of East Asian handmade papers [23]. The aspect of how this  
240 benefit can be used for 1D data mining is explored here. Since the tested samples are identical,  
241 the presence of compounds in 2D profile may lead to quick location of co-eluting compounds or  
242 new compounds difficult to attribute in the corresponding 1D GC/MS profile. Also, a well-  
243 known feature of 2D analysis is the structure-based (or structure grouping) compound  
244 distribution in the 2D chromatogram that may be helpful for analysis of a series of compounds  
245 with similar structures in 1D GC/MS data.



246

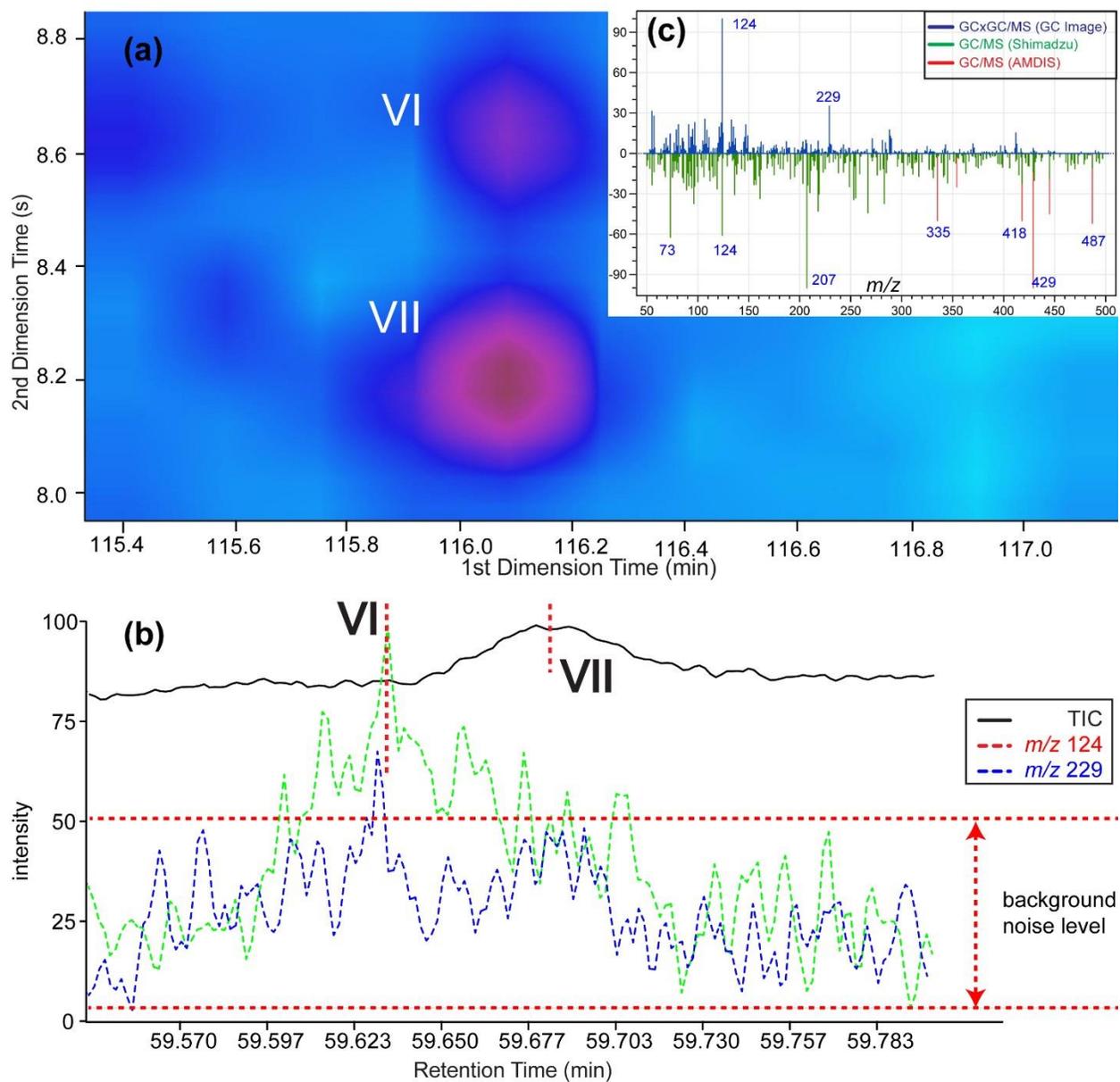
247 **Figure 3.** Py-GCxGC/MS profile for five pyrolysis products within the plant marker region  
 248 distribution (a) and examination of the corresponding 1D data (b) revealed with an extracted ion  
 249 chromatograms (EICs) procedure

250 The region of elution of five pyrolysis products presented as an example and analyzed in the Py-  
 251 GCxGC/MS profile and the corresponding 1D Py-GC/MS profile (Figure 3). The compounds I-

252 V showed close volatilities as indicated by their retention time  $^1t_R$ . However, due to their  
253 difference in structure, they presented different polarities and by the benefit of a GCxGC  
254 separation, they were separated into two different groups (Figure 3a). The first group consists of  
255 phytosterol structure compounds eluting first around  $^2t_R=6.5$ s (compound I and IV) followed by  
256 the second group mainly consisting of terpenic structure compounds eluting at  $^2t_R=6.95$ s  
257 (compound II, III and V). From the structure layout in the 2D analysis, we gained the knowledge  
258 that within the analyzed region, there were two different groups of eluting compounds. For a  
259 quick location of these compounds in the 1D data, fragment ions selected from their 2D mass  
260 spectra were used (namely,  $m/z$  at 143, 189, 203, 218, 394 and 396), which is convenient and a  
261 real time-saving method compared to a manual check for the identification of peaks  
262 deconvoluted by AMDIS. The order of the selected ions in Figure 3b clearly depicts all the  
263 compounds in the 1D data. It can be seen that compounds I and II are closely co-eluting within a  
264 retention time difference of less than 2 s and the other three compounds are also easily located by  
265 the selected ions within AMDIS after the help of 2D separation for a reliable assignment.

266 The procedure involving the acquired compound distribution knowledge seems quick and  
267 simple, but without the improved separation obtained by GCxGC, AMDIS alone would be  
268 inefficient or even unsuccessful in targeting all these compounds. In particular, the direct  
269 deconvolution of compounds III, IV and V using AMDIS with the previously reported parameter  
270 settings during the 1D Py-GC/MS analysis was not conclusive. For the terpenic compounds III  
271 and V, AMDIS did not extract the  $m/z$  218, 203 from the mass spectra, which adversely were  
272 showing a high level of noise (Supplementary Figure S1-1 and Supplementary Figure S1-3), a  
273 critical point for their identification (only  $m/z$  189 was extracted for compound V). For  
274 compound IV, although the  $m/z$  396 is extracted, ions such as  $m/z$  147 and 95 were missing

275 (Supplementary Figure S1-2). On the contrary, the GCxGC separation even with less sample  
276 quantity clearly presents better mass spectra (Supplementary Figure S1). In considering  
277 interference from background and co-eluting compounds, reliability of AMDIS for detecting  
278 unknown and low concentration of compounds shall be taken into account for biomarker  
279 discovery.

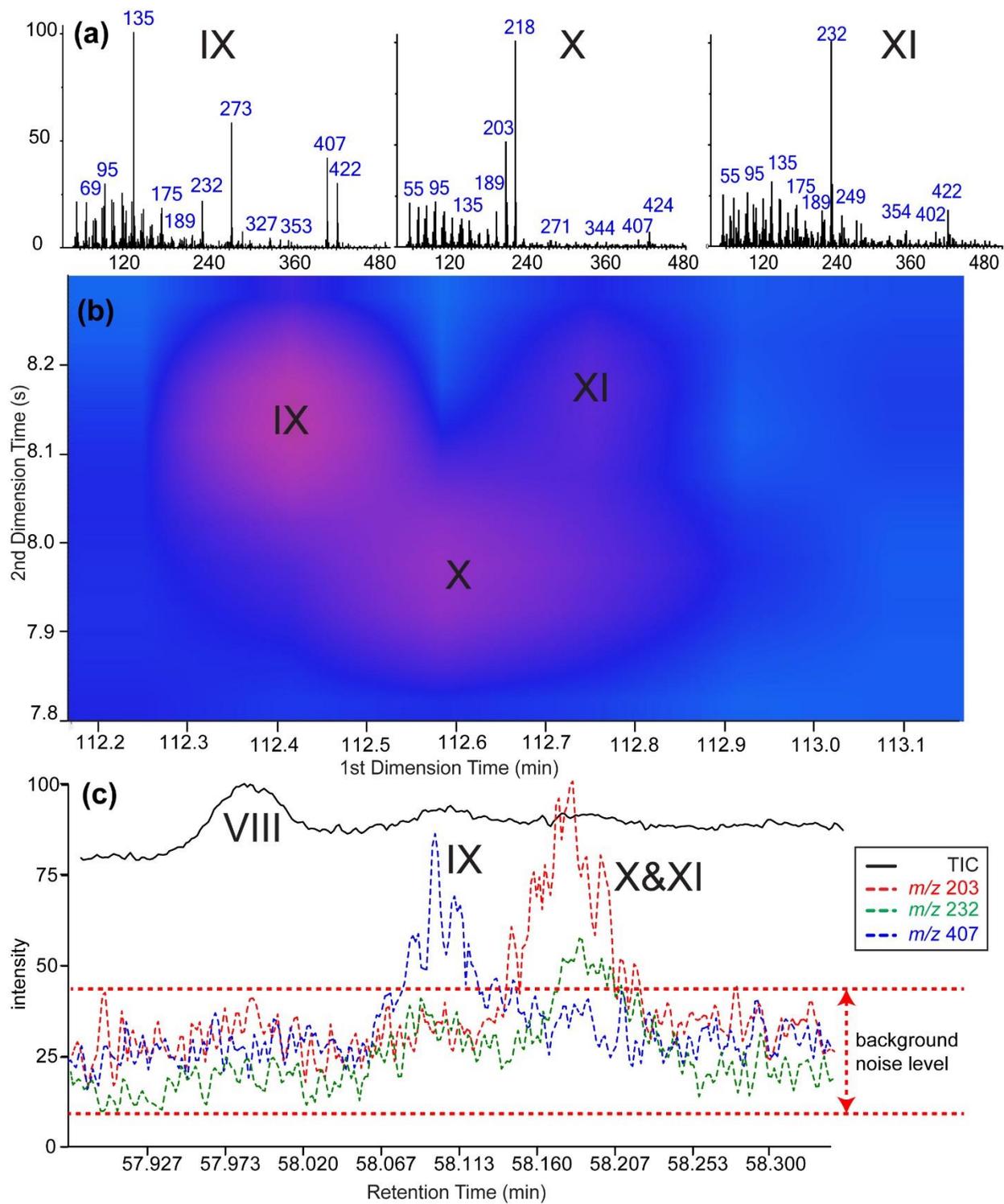


280

281 **Figure 4.** Py-GCxGC/MS profile (a) distribution for examination of the corresponding 1D data  
282 revealed with an extracted ion chromatograms (EICs) procedure (b). Insert reports (c) were  
283 compound VI's mass spectra from the Py-GCxGC/MS profile (blue), while GCMS Postrun  
284 Analysis (Shimadzu) manual check (green) and AMDSIS automatic extraction (red)

285 Due to the detection limit of 1D GC/MS and the interference from the background, using  
286 AMDIS to detect signals barely above background is not easy or even impossible in some cases.  
287 For example, in Figure 4a, compounds VI and VII presented almost the same retention times in  
288  $^1t_r$  and their differentiation relied largely on their separation in the second dimension  $^2t_r$  for their  
289 detection. After their identification in the 2D profile, the 1D profile is re-examined using  
290 AMDIS. Previously the compound VI (Figure 4b) was not taken into account as it does not have  
291 a sufficient signal/noise ratio to emerge from the background. Even after knowing the existence  
292 of compound VI in the chromatogram, it is not easy to locate it in the GCMS Postrun Analysis  
293 software (Shimadzu) using subtraction of background from nearby baseline as its intensity is too  
294 low and the background mass spectra (with predominant  $m/z$  at 73, 207, 281) were still influent  
295 after the subtraction procedure (Figure 4c). Also, AMDIS deconvolution cannot satisfactorily  
296 extract the characteristic ions such as  $m/z$  124 and  $m/z$  229 as revealed in Figure 4c. The co-  
297 elution of compound VI and compound VII in the 2D profile considering their retention time in  
298 the first dimension  $^1t_r$  confirmed that compound VI is eluting near compound VII or even within  
299 compound VII in the corresponding 1D profile. From the mass spectrum of compound VI  
300 extracted from the 2D analysis (Figure 4c), the intensive ions at  $m/z$  124 and  $m/z$  229 were  
301 selected to reinspect the 1D profile. From the AMDIS display of  $m/z$  124 and  $m/z$  229, the  
302 presence of VI around compound VII was confirmed (Figure 4b). In 1D data, the compound VI  
303 firstly eluted straight before the elution of compound VII and continued to elute within the peak

304 of compound VII. However, the intensity of VII is comparatively higher and compound VI was  
305 hidden in the background as was illustrated by Fig.4b. In this typical case where a compound at  
306 low concentration elutes closely with a compound at a higher concentration, AMDIS cannot  
307 work properly. Only after the insight of compounds eluting from the 2D profile, AMDIS using  
308 the selected EIC procedure can locate the compound of lower intensity. The success of  
309 separating compounds VI and VII in Figure 4 is largely due to their difference in structure  
310 resulting in a difference in polarity allowing their efficient separation in the second dimension.  
311 Identification of compounds VI and VII is finally accomplished after reinjection of 2D fragment  
312 ion identification. This example illustrated the expected contribution of a GCxGC analysis for  
313 resolution of unfavorable cases in a monodimensional separation.



314

315 **Figure 5.** Mass spectra for compounds IX, X and XI (a) and their Py-GCxGC/MS profile (b) for  
316 examination of the corresponding 1D data (c) revealed with an extracted ion chromatograms  
317 (EICs) procedure

318 To illustrate the benefit of use of Py-GCxGC/MS profiles for other pyrolysis products detected in  
319 the plant marker region, Figure 5 reports the identification process of another three compounds.

320 In this example, AMDIS deconvolution worked well only with compound VIII, which is  
321 characterized by a satisfying signal to noise ratio in the 1D profile, while it cannot extract  
322 informative ions for compounds IX, X and XI since their signals were very low and submerged  
323 in the background (Figure 5c). From the 2D profile, it is clear that three peaks were clearly  
324 detected in the region  $^1t_r = 112.3 \text{ min} - 112.8 \text{ min}$ ,  $^2t_r = 7.8 \text{ s} - 8.3 \text{ s}$  (Figure 5b). And the  
325 corresponding mass spectra are reported in Figure 5a. The main fragments  $m/z$  at 203, 407 and  
326 232 were selected for their display in AMDIS for the treatment of 1D GC/MS data (for  
327 compound XI, the base peak  $m/z$  232 was chosen while for compound IX, X the  $m/z$  of lower  
328 intensity 203, 407 was selected to compensate the difference of concentration between  
329 compounds IX, X vs compound XI). It is shown that the three compounds were present in the 1D  
330 profile with comparatively low intensity (especially for compound XI) with compound IX  
331 eluting first and compounds X, XI co-eluting together.

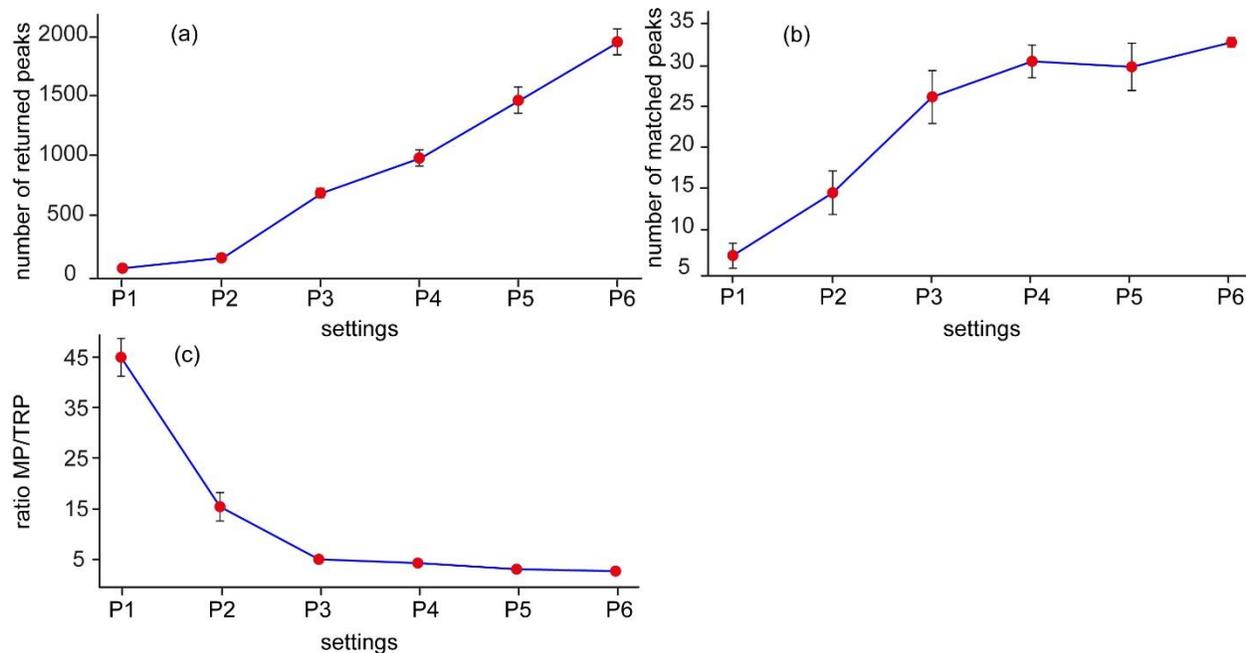
332 The three above examples illustrated that by using the 2D profile and reinjecting the mass  
333 spectra information for an EIC identification procedure, co-eluting compounds in 1D analysis  
334 can be efficiently detected and located within AMDIS. Instead of spending time in choosing the  
335 right ions (EICs) for the optimization of the AMDIS identification procedure (sometimes in vain  
336 due to the low resolution and low mass spectra quality in 1D profile), the 2D profile can act as  
337 'peak finder' for the 1D GC/MS data analysis. Also, considering that the required quantity of

338 sample for 2D analysis is at least half than that of the 1D analysis (ca. 50  $\mu\text{g}$  vs ca. 100  $\mu\text{g}$ ), it  
339 can be seen from their mass spectra that the 2D improved separation offering more reliable mass  
340 spectra and consequently, the identification of products during 1D analysis can be confirmed by  
341 their corresponding 2D spectra. On our opinion, an important feature reported for AMDIS, which  
342 lays in its suitability at detecting very low levels of metabolite compounds in biological samples  
343 [6], may be noticeably increased by the combination of 1D analysis with mass spectral  
344 information derived from a 2D analysis. The 2D profile, when analyzing complex samples, can  
345 also contribute to improve the reliability of the mass spectra for comparison with a  
346 comprehensive library or even in the case of unidentified compounds [30].

347

348 *GCxGC/MS profiles combined to AMDIS automatic identification process in 1D*  
349 *GC/MS data analysis*

350 The advantage of AMDIS is indicated by its name: Automated Mass Spectral Deconvolution and  
351 Identification System. In routine analysis, AMDIS is a powerful tool for quick peak  
352 identification. One important part of AMDIS is its library search function which was frequently  
353 used for match and identification of recognized compounds. Although several library has been  
354 embedded in AMDIS, such as NISTFF (for flavours and fragrances analysis), NISTTOX (for  
355 toxicity analysis) and NISTCW (for chemical weapon analysis) [31], the main drawback is still  
356 evident as the library provided may not contain enough compounds, enough mass spectra  
357 information and even, in some specific domains, standard libraries may be missing. In practice,  
358 lots of AMDIS users are intend to build in-house library which proved to be effective and  
359 convenient for their specific studies [32, 33].



360  
 361 **Figure 6.** Total number of peaks returned (a) and number of matched peaks with target library  
 362 (b) by AMDIS deconvolution process with different settings applied to the parameters  
 363 “resolution” (RE), “sensitivity” (SE), and “peak shape requirement” (PE) (P1:all Low; P2:all  
 364 Medium; P3: all High; P4: RE,SE High, PE Medium; P5: all Very high; P6: RE,SE Very high, PE  
 365 Medium). (c) ratio of matched peak (MP)/total returned peaks (TRP)

366 As has been mentioned above, the number of peaks returned by AMDIS deconvolution process  
 367 largely depends on the analysis settings. To define an optimized setting, an in-house target  
 368 compound library was built using the assigned compounds in the plant marker regions (for  
 369 compound identification, refer to [23]) and tested with different parameter settings (Figure 6).  
 370 Even considering optimized parameter settings in AMDIS, one must be aware of the influence of  
 371 these settings during the AMDIS deconvolution procedure. It can be seen from Figure 6 that the  
 372 setting of parameters for the parameters “resolution”, “sensitivity” and “shape requirement”,  
 373 greatly influences the final results. As an illustration, when changing parameter values from low  
 374 to medium (P1 to P2), the total number of peaks returned increases slowly, but when using the

375 ‘high’ settings (P3), the number of generated peaks rise noticeably rendering time consuming the  
 376 treatment of data after deconvolution (Figure 6a). The deconvolution returned peaks continued to  
 377 increase when using ‘very high’ settings (P3 to P5). The influence of peak shape is also depicted  
 378 in Figure 6. The number of returned peaks also increased when set the peak shape requirement as  
 379 medium (P3 to P4, P5 to P6). The number of peaks matched with the library search also varies  
 380 with these different settings (Figure 6b): the number of matched peaks with library search using  
 381 high settings increased by 2~3 times compared with those using low and medium settings (P3  
 382 compared to P1 and P2) while continue to use ‘very high’ settings did not have much influence  
 383 in the number of matched peaks. The problem raised when using the high settings for AMDIS  
 384 deconvolution process is that it will generate a large number of false peaks, to be eliminated  
 385 during the post-processing of data. It can be seen from Figure 6c that the number of matched  
 386 peaks was less than 5% of the total peaks returned from the deconvolution when use ‘high’  
 387 settings. In the following, deconvolution settings with “high resolution”, “high sensitivity” and  
 388 “medium shape requirement” were used as optimized settings for 1D GC/MS data analysis.

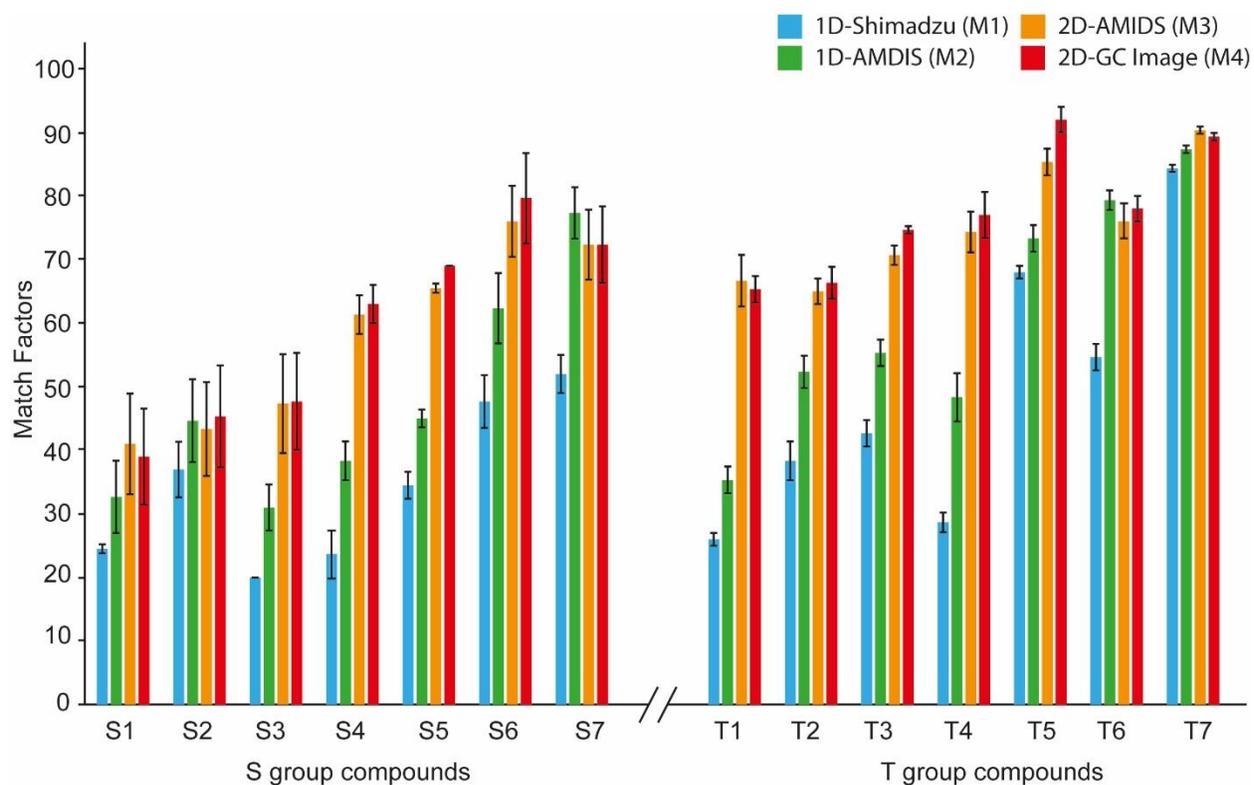
389 **Table 1.** List of compounds used for building up an in-house target library for the evaluation the  
 390 quality of the GCxGC/MS and 1D GC/MS mass spectra.

Retention Time	<i>m/z</i>	Assigned formula	Molecular mass	Attribution	TIC response	Label*
53.023	365, 121, 135, 93	C30H48	408	A-neooleana-3(5),12-diene	72305	S2
53.394	145, 81, 121, 95	C29H48	396	stigmasta-3,5-diene	48551	S3
53.953	406, 255, 95, 123	C30H46	406	oleanane-2,9,12-triene	118881	S4
54.343	218, 203,189, 95	C30H48	408	oleanane-2,12-diene	117283	T4
54.712	105, 57, 145, 396	C29H50O	414	NA stigmastan	91922	S1
54.917	135, 143, 394, 81	C29H48O	412	stigmasta-4,6-dien-3-ol	631803	S5*
54.917	218, 189, 93, 95	C30H48	408	urs-2,12-diene	631803	T5*
55.178	147, 81, 145, 396	C29H50O	414	stigmasta-3-ol	130403	S6
58.465	218, 203, 189, 95	C30H50O	426	beta amyirin (Olean-12-en-3-ol)	119246	T3
58.726	218, 95, 107, 219	C30H48O	424	urs-12-en-3-one	82013	T1
59.023	218, 95, 81, 189	C30H50O	426	alpha-amyirin (urs-12-en-3-ol)	151263	T2

59.277	466, 255, 121, 95	C32H50O2	466	oleanane-12,14-diene-3-acetate	62445	S7
59.678	218, 203, 189, 69	C32H52O2	468	oleanane-12-ene-3-acetate	249653	T6
60.196	218, 189, 95, 69	C32H52O2	468	alpha-amyrin acetate (urs-12-ene-3-acetate)	1438017	T7

391 Label\* is corresponding to the match factors illustrated in Figure 7  
392 S5 and T5\* are co-eluting, the same TIC area was used for both.

393 As a benefit of the increased resolution and sensitivity of the GCxGC system, both the  
394 chromatographic separation and resulting pyrolysis product identifications were improved. It  
395 also resulted in an increased mass spectra quality even with a lower quantity of sample.  
396 However, the question “how these characteristics of GCxGC analysis are able to improve the  
397 mass spectra collection and to benefit 1D data treatment” has not yet been investigated. To test  
398 the performance of reintroducing the GCxGC mass spectra in the 1D mass spectra data treatment  
399 procedure of AMDIS, two categories of compounds in the plant marker region were selected:  
400 one category constituted mainly of stigmasta compounds (reported as S type) and another  
401 category constituted of terpenic compounds (reported as T type). The compound information is  
402 listed in Table 1 and their distribution in the 1D Py-GC/MS chromatogram is presented in  
403 Supplementary Figure S2. Their respective mass spectra were obtained following the procedure  
404 described in Supplementary Figure S3 which involved the GCxGC data collection of mass  
405 spectra using GCxGC Image, and the 1D GC/MS data collection of mass spectra using the  
406 GCMS postrun analysis software as well as AMDIS. Finally, compilation of this set of mass  
407 spectra formed an AMDIS in-house target library assembling 56 mass spectra (Supplementary  
408 Figure S4). For the purpose of mass spectra comparison, the AMDIS analysis was conducted  
409 with a simple type of analysis model (a minimum match factor was set at 20% and multiple  
410 identifications per compound was allowed).



411

412 **Figure 7.** Report of the match factors for each products with the in-house target library  
 413 consisting of mass spectra derived from GCxGC/MS and 1D GC/MS data collection. For the  
 414 collecting of mass spectra using M1, M2, M3 and M4, refers to methods in Supplementary  
 415 Figure S2

416 The match factors for each category of compounds with mass spectra collected from Py-  
 417 GCxGC/MS and 1D Py-GC/MS analysis were shown in Figure 7. As all the applied parameters  
 418 were the same during the procedure of deconvolution, the differences observed in the match  
 419 factor was mainly caused by the quality of the mass spectra (the ability to extract characteristic  
 420 ions from the test chromatograms is also a matter of interest). As can be seen in Figure 7, the  
 421 mass spectra collected from the GCxGC profile generally provided a 10% to 20% higher match  
 422 factor than the mass spectra collected from the 1D GC/MS data. For S1~S7, The GCxGC mass

423 spectra have less interference from the background (see Supplementary Figure S4-1 to  
424 Supplementary Figure S4-28) and presented a higher match factor except for S1 and S2. This  
425 could be correlated to the fact that the AMDIS deconvolution missed to extract enough  
426 characteristic ions of S1 and S2 for comparison with the mass spectra in the in-house target  
427 library, which can be rationalized considering the low concentration and background interference  
428 of these two compounds. The T type compounds, which all have characteristic fragment ions  $m/z$   
429 at 218, 203 and 189 of high intensities were more illustrative for a series of compounds with  
430 similar structure as they are belonging to the same family with high similarity observed in their  
431 mass spectra patterns (Supplementary Figure S4-29 to Supplementary Figure S4-56). The  
432 chromatographic response for T1~T4 in 1D analysis is comparatively lower compared to the  
433 corresponding GCxGC/MS profile, which resulted in mass spectra containing high intensity of  
434 background ions (column bleeding) (Supplementary Figure S4-29 to Supplementary Figure S4-  
435 44). On the contrary, the GCxGC/MS profile provided a higher detection level associated to a  
436 increased resolution of chromatographic peaks which resulted in an improved quality of the mass  
437 spectra and a higher match factor in the AMDIS automatic identification process when using  
438 GCxGC reintroduced mass spectra. As can be exemplified from the mass spectrum of  
439 compound T5, the intensity of  $m/z$  218 and 189 were comparatively higher in the collected  
440 GCxGC mass spectrum than the 1D mass spectrum and their difference were mainly residing in  
441 the content of low  $m/z$  fragment ions from 50 to 180 u (Supplementary Figure S5). This can be  
442 explained by the co-elution of compound T5 with compound S5: in the 1D mass spectrum, the  
443  $m/z$  50-180 fragment ions result from the fragmentation of both T5 and S5, which cannot be fully  
444 separated by AMDIS deconvolution, while in the GCxGC profile, as T5 and S5 have been  
445 separated in the two dimensions,  $m/z$  218 and 203 ions have comparatively higher intensities

446 with the decrease of  $m/z$  50-180. As for compound T6, using the GCMS Postrun Analysis  
447 software (Shimadzu) for background subtraction can still contain high background and column  
448 bleeding ions like  $m/z$  at 207 (Supplementary Figure S4-49) which may cause its comparatively  
449 lower match factor, while background subtraction using AMDIS can have the same match factor  
450 with mass spectra from the GCxGC profile. It is also observed that for the compounds present in  
451 comparatively high concentration, like compound T7, both the 1D and GCxGC method can  
452 produce mass spectra with lower background interference (Supplementary Figure S4-53 to  
453 Supplementary Figure S4-56, Supplementary Figure S5), which demonstrated that in such case,  
454 the 1D chromatographic detection level and resolution are satisfactory for proper identification.

455 Among the four way of getting mass spectra for target library construction, the manual  
456 background subtraction using the GCMS Postrun Analysis software did not produce high match  
457 factors as there were still high level of background interference (Supplementary Figure S4). The  
458 algorithm for background subtraction in AMDIS reduced the level of background interference  
459 compared to the use of GCMS Postrun Analysis software but has the limitation of detection and  
460 resolution reported above during 1D data collection. The mass spectra produced from a  
461 GCxGC/MS profile has fewer contaminations from the background (column bleeding) and are  
462 observed with higher match factors in the AMDIS automatic identification procedure. This result  
463 constitutes a plus for the AMDIS automatic data treatment to provide a complementary way of  
464 analysis for AMDIS use. Moreover, to our opinion, the approach adopted here demonstrated its  
465 potential to circumvent the inability to reliably extract weak signals in AMDIS automatic  
466 deconvolution procedure due to the inherent difficulty of effectively establishing thresholds for  
467 distinguishing significant signal from noise.

468

## 469 **Conclusions**

470 The potential use of GCxGC/MS profiles to conduct an improved AMDIS procedure for  
471 compound identification in 1D GC/MS data was reported in this study with application to the  
472 identification of pyrolysis products of East Asian handmade papers. The benefits of GCxGC  
473 profiles use for manual and automatic 1D GC/MS data mining with AMDIS was explored. The  
474 result showed that the presence of compounds in 2D profile may lead to the quick location of co-  
475 eluting compounds or the quick identification of new compounds in 1D GC/MS chromatograms  
476 eluding background contamination in the mass spectra and avoiding mass spectra mixing due to  
477 ineffective chromatographic separation. Clearly, the structured layout of products in a 2D  
478 chromatogram may help for 1D analysis of a series of compounds with similar structures. As a  
479 consequence, the compound distribution in GCxGC/MS profiles can be used as ‘peak finder’ for  
480 manual check of low concentration and co-eluting compound identification in 1D GC data.  
481 GCxGC/MS profiles improved the identification of unresolved patterns in 1D GC/MS analysis,  
482 as well as permitted to improve the EICs identification procedure for low signal peak detection  
483 and unresolved peak patterns showing a high dynamic in their corresponding chromatographic  
484 signals.

485 The performance of an in-house library constructed with mass spectra collected from  
486 GCxGC/MS profiles and corresponding 1D profiles for AMDIS automatic identification process  
487 is also evaluated. The result showed that the mass spectra from the GCxGC profile generally  
488 have 10% to 20% higher match factors than the mass spectra obtained from the 1D profile. Also  
489 the way of collecting mass spectra was shown to affect the quality of the mass spectra with the  
490 influence of background ions interference. The algorithm for background subtraction in AMDIS  
491 reduced the level of background interference but with a limitation in its effectiveness for

492 identification in 1D analysis. The mass spectra produced from a GCxGC/MS profile have higher  
493 matches observed in the AMDIS automatic identification process which may further contribute to  
494 the 1D data analysis when combined with parameters such as retention time and retention  
495 indexes.

496 The present study demonstrated how 2D profiles provide a complementary way of analysis for  
497 AMDIS use: a future application may consist for a given sample system of interest, firstly, in a  
498 scrutinizing application of 2D GCxGC/MS to appropriately identify the components of this  
499 system, and, second, in applying a routine use of AMDIS in a subsequent 1D GC/MS analysis.  
500 Extension of the application of the present method may be done through the use of an in-house  
501 library, which can be easily shared by scientists working in the same domain of interest.  
502 Similarly, a  $m/z$  table of interest build from the EIC issued from the 2D GCxGC analysis can be  
503 provided for routine 1D GC/MS analysis when 2D GCxGC/MS is not at disposal.

504 The ability of AMDIS for identifying spectra with very low levels of signals once high-quality  
505 spectrum is provided and operated with optimized parameters (filtering procedure in the post-  
506 processing of data, retention index use) was previously reported, and is currently effective in  
507 routine AMDIS uses. The further investigation of the optimized AMDIS parameters with the  
508 described 2D method (the automated in-house library search together with an EIC examination  
509 procedure), may contribute to the identification of similar plant species that have common major  
510 compounds and may only be differentiated by their minor compounds. By definition, there must  
511 be a potential usage of the proposed method and in-house libraries collected from 2D analysis in  
512 an extended field of applications like cultural heritage, botany or material science.

513

514 **Acknowledgement**

515 Pr Masamitsu Inaba (Conservation Science Laboratory, Graduate School of Conservation Tokyo  
516 Geijutsu Daigaku, Japan) and Mr Masazumi Seki (Kochi prefectural paper technology center,  
517 Kochi, Japan) are gratefully acknowledged for providing the kozo reference paper samples. B. H.  
518 is grateful to the China Scholarship Council for their financial support through a grant in the  
519 course of the PhD program of Sorbonne Universités/China Scholarship Council.

520

## 521 **References**

- 522 1. Choe, S., Woo, S.H., Kim, D.W., Park, Y., Choi, H., Hwang, B.Y., Lee, D., Kim, S.:  
523 Development of a target component extraction method from GC-MS data with an in-house  
524 program for metabolite profiling. *Anal. Biochem.* 426, 94-102 (2012)
- 525 2. Aggio, R., Villas-Boas, S.G., Ruggiero, K.: Metab: an R package for high-throughput  
526 analysis of metabolomics data generated by GC-MS. *Bioinformatics.* 27, 2316-2318 (2011)
- 527 3. Wehrens, R., Weingart, G., Mattivi, F.: metaMS: An open-source pipeline for GC-MS-  
528 based untargeted metabolomics. *J. Chromatogr. B.* 966, 109-116 (2014)
- 529 4. Behrends, V., Tredwell, G.D., Bundy, J.G.: A software complement to AMDIS for  
530 processing GC-MS metabolomic data. *Anal. Biochem.* 415, 206-208 (2011)
- 531 5. Stein, S.E.: An integrated method for spectrum extraction and compound identification  
532 from gas chromatography/mass spectrometry data. *J. Am. Soc. Mass Spectrom.* 10, 770-781  
533 (1999)
- 534 6. Halket, J.M., Przyborowska, A., Stein, S.E., Mallard, W.G., Down, S., Chalmers, R.A.:  
535 Deconvolution gas chromatography/mass spectrometry of urinary organic acids—potential for  
536 pattern recognition and automated identification of metabolic disorders. *Rapid Commun. Mass*  
537 *Spectrom.* 13, 279-284 (1999)
- 538 7. Mallard, W.G., Andriamiharavo, N.R., Mirokhin, Y.A., Halket, J.M., Stein, S.E.:  
539 Creation of libraries of recurring mass spectra from large data sets assisted by a dual-column  
540 workflow. *Anal. Chem.* 86, 10231-10238 (2014)

- 541 8. Stein, S.E. and Scott, D.R.: Optimization and testing of mass spectral library search  
542 algorithms for compound identification. *J. Am. Soc. Mass Spectrom.* 5, 859-866 (1994)
- 543 9. Lu, H.M., Dunn, W.B., Shen, H.L., Kell, D.B., Liang, Y.Z.: Comparative evaluation of  
544 software for deconvolution of metabolomics data based on GC-TOF-MS. *Trac-Trends Anal.*  
545 *Chem.* 27, 215-227 (2008)
- 546 10. Cerdan-Calero, M., Sendra, J.M., Sentandreu, E.: Gas chromatography coupled to mass  
547 spectrometry analysis of volatiles, sugars, organic acids and aminoacids in Valencia Late orange  
548 juice and reliability of the Automated Mass Spectral Deconvolution and Identification System  
549 for their automatic identification and quantification. *J. Chromatogr. A.* 1241, 84-95 (2012)
- 550 11. Smits, M.M., Carleer, R., Colpaert, J.V.: PYQUAN: A rapid workflow around the  
551 AMDIS deconvolution software for high throughput analysis of pyrolysis GC/MS data. *J. Anal.*  
552 *Appl. Pyrolysis.* 118, 335-342 (2016)
- 553 12. Wu, L.J., Liu, W., Cao, J.L., Li, Q.Q., Huang, Y., Min, S.G.: Analysis of the aroma  
554 components in tobacco using combined GC-MS and AMDIS. *Anal. Methods.* 5, 1259-1263  
555 (2013)
- 556 13. Buszewska-Forajta, M., Bujak, R., Yumba-Mpanga, A., Siluk, D., Kaliszan, R.: GC/MS  
557 technique and AMDIS software application in identification of hydrophobic compounds of  
558 grasshoppers' abdominal secretion (*Chorthippus* spp.). *J. Pharm. Biomed. Anal.* 102, 331-339  
559 (2015)
- 560 14. Smart, K.F., Aggio, R.B., Van Houtte, J.R., Villas-Bôas, S.G.: Analytical platform for  
561 metabolome analysis of microbial cells using methyl chloroformate derivatization followed by  
562 gas chromatography-mass spectrometry. *Nat. Protoc.* 5, 1709-1729 (2010)

- 563 15. Liu, Z. and Phillips, J.B.: Comprehensive two-dimensional gas chromatography using an  
564 on-column thermal modulator interface. *J. Chromatogr. Sci.* 29, 227-231 (1991)
- 565 16. Adahchour, M., Beens, J., Vreuls, R., Brinkman, U.T.: Recent developments in  
566 comprehensive two-dimensional gas chromatography (GC× GC): I. Introduction and  
567 instrumental set-up. *Trac-Trends Anal. Chem.* 25, 438-454 (2006)
- 568 17. Mondello, L., Tranchida, P.Q., Dugo, P., Dugo, G.: Comprehensive two - dimensional  
569 gas chromatography - mass spectrometry: A review. *Mass Spectrom. Rev.* 27, 101-124 (2008)
- 570 18. Gilar, M., Olivova, P., Daly, A.E., Gebler, J.C.: Orthogonality of separation in two-  
571 dimensional liquid chromatography. *Anal. Chem.* 77, 6426-6434 (2005)
- 572 19. Davis, J.M., Stoll, D.R., Carr, P.W.: Dependence of Effective Peak Capacity in  
573 Comprehensive Two-Dimensional Separations on the Distribution of Peak Capacity between the  
574 Two Dimensions. *Anal. Chem.* 80, 8122-8134 (2008)
- 575 20. Tranchida, P.Q., Franchina, F.A., Dugo, P., Mondello, L.: Comprehensive two -  
576 dimensional gas chromatography - mass spectrometry: Recent evolution and current trends.  
577 *Mass Spectrom. Rev.* 35, 524-534 (2016)
- 578 21. Pedroso, M.P., Ferreira, E.C., Hantao, L.W., Bogusz, S., Augusto, F.: Identification of  
579 volatiles from pineapple (*Ananas comosus* L.) pulp by comprehensive two - dimensional gas  
580 chromatography and gas chromatography/mass spectrometry. *J. Sep. Sci.* 34, 1547-1554 (2011)
- 581 22. Avataneo, C. and Sablier, M.: New criteria for the characterization of traditional East  
582 Asian papers. *Environ. Sci. Pollut. Res.* 24, 2166-2181 (2017)

- 583 23. Han, B., Daheur, G., Sablier, M.: Py-GCxGC/MS in cultural heritage studies: An  
584 illustration through analytical characterization of traditional East Asian handmade papers. J.  
585 Anal. Appl. Pyrolysis. 122, 458-467 (2016)
- 586 24. Han, B., Vial, J., Inaba, M., Sablier, M.: Analytical characterization of East Asian  
587 handmade papers: A combined approach using Py-GCxGC/MS and multivariate analysis. J.  
588 Anal. Appl. Pyrolysis. (2017)
- 589 25. Budzikiewicz, H., Wilson, J.M., Djerassi, C.: Mass spectrometry in structural and  
590 stereochemical problems. XXXII. 1 Pentacyclic triterpenes. J. Am. Chem. Soc. 85, 3688-&  
591 (1963)
- 592 26. Karliner, J. and Djerassi, C.: Terpenoids. LVII. Mass Spectral and Nuclear Magnetic  
593 Resonance Studies of Pentacyclic Triterpene Hydrocarbons. J. Org. Chem. 31, 1945-& (1966)
- 594 27. Ogunkoya, L.: Application of mass spectrometry in structural problems in triterpenes.  
595 Phytochemistry. 20, 121-126 (1981)
- 596 28. Davis, J.M.: Statistical-overlap theory for elliptical zones of high aspect ratio in  
597 comprehensive two-dimensional separations. J. Sep. Sci. 28, 347-359 (2005)
- 598 29. Liu, S.Y. and Davis, J.M.: Dependence on saturation of average minimum resolution in  
599 two-dimensional statistical-overlap theory: Peak overlap in saturated two-dimensional  
600 separations. J. Chromatogr. A. 1126, 244-256 (2006)
- 601 30. Stein, S.: Mass Spectral Reference Libraries: An Ever-Expanding Resource for Chemical  
602 Identification. Anal. Chem. 84, 7274-7282 (2012)

603 31. Davies, T.: The new automated mass spectrometry deconvolution and identification  
604 system (AMDIS). *Spectrosc. Eur.* 10, 24-27 (1998)

605 32. Schilling, M.R., Heginbotham, A., van Keulen, H., Szelewski, M.: Beyond the basics: A  
606 systematic approach for comprehensive analysis of organic materials in Asian lacquers. *Stud.*  
607 *Conserv.* 61, 3-27 (2016)

608 33. Härtig, C.: Rapid identification of fatty acid methyl esters using a multidimensional gas  
609 chromatography–mass spectrometry database. *J. Chromatogr. A.* 1177, 159-169 (2008)

610

611 **Figure captions**

612 **Figure 1.** Py-GCxGC/MS (a) and Py-GC/MS (b) chromatograms obtained from the pyrolysis of  
613 a kozo handmade paper sample

614 **Figure 2.** The examined peak issued from Py-GC/MS analysis (a) with its corresponding  
615 scanned mass spectrum (b) and manually subtracted spectrum (c) using the GCMS Postrun  
616 Analysis software (Shimadzu) for peak identification. AMDIS automatic mass spectra  
617 deconvolution was conducted and resulted in the scanned spectra (d and e) and automatic  
618 extracted spectra (f and g)

619 **Figure 3.** Py-GCxGC/MS profile for five pyrolysis products within the plant marker region  
620 distribution (a) and examination of the corresponding 1D data (b) revealed with an extracted ion  
621 chromatograms (EICs) procedure

622 **Figure 4.** Py-GCxGC/MS profile (a) distribution for examination of the corresponding 1D data  
623 revealed with an extracted ion chromatograms (EIC) procedure (b). Insert reports (c) were  
624 compound VI's mass spectra from the Py-GCxGC/MS profile (blue), while GCMS Postrun  
625 Analysis (Shimadzu) manual check (green) and AMDSIS automatic extraction (red)

626 **Figure 5.** Mass spectra for compounds IX, X and XI (a) and their Py-GCxGC/MS profile (b) for  
627 examination of the corresponding 1D data (c) revealed with an extracted ion chromatograms  
628 (EIC) procedure

629 **Figure 6.** Total number of peaks returned (a) and number of matched peaks with target library  
630 (b) by AMDIS deconvolution process with different settings applied to the parameters  
631 "resolution" (RE), "sensitivity" (SE), and "peak shape requirement" (PE) (P1:all Low; P2:all  
632 Medium; P3: all High; P4: RE,SE High, PE Medium; P5: all Very high; P6: RE,SE Very high, PE  
633 Medium). (c) ratio of matched peak (MP)/total returned peaks (TRP)

634 **Figure 7.** Report of the match factors for each products with the in-house target library  
635 consisting of mass spectra derived from GCxGC/MS and 1D GC/MS data collection. For the  
636 collecting of mass spectra using M1, M2, M3 and M4, refers to Supplementary Figure S2

637 **Table 1.** List of compounds used for building up an in-house target library for the evaluation the  
638 quality of the GCxGC/MS and 1D GC/MS mass spectra.

639

Figure 1.

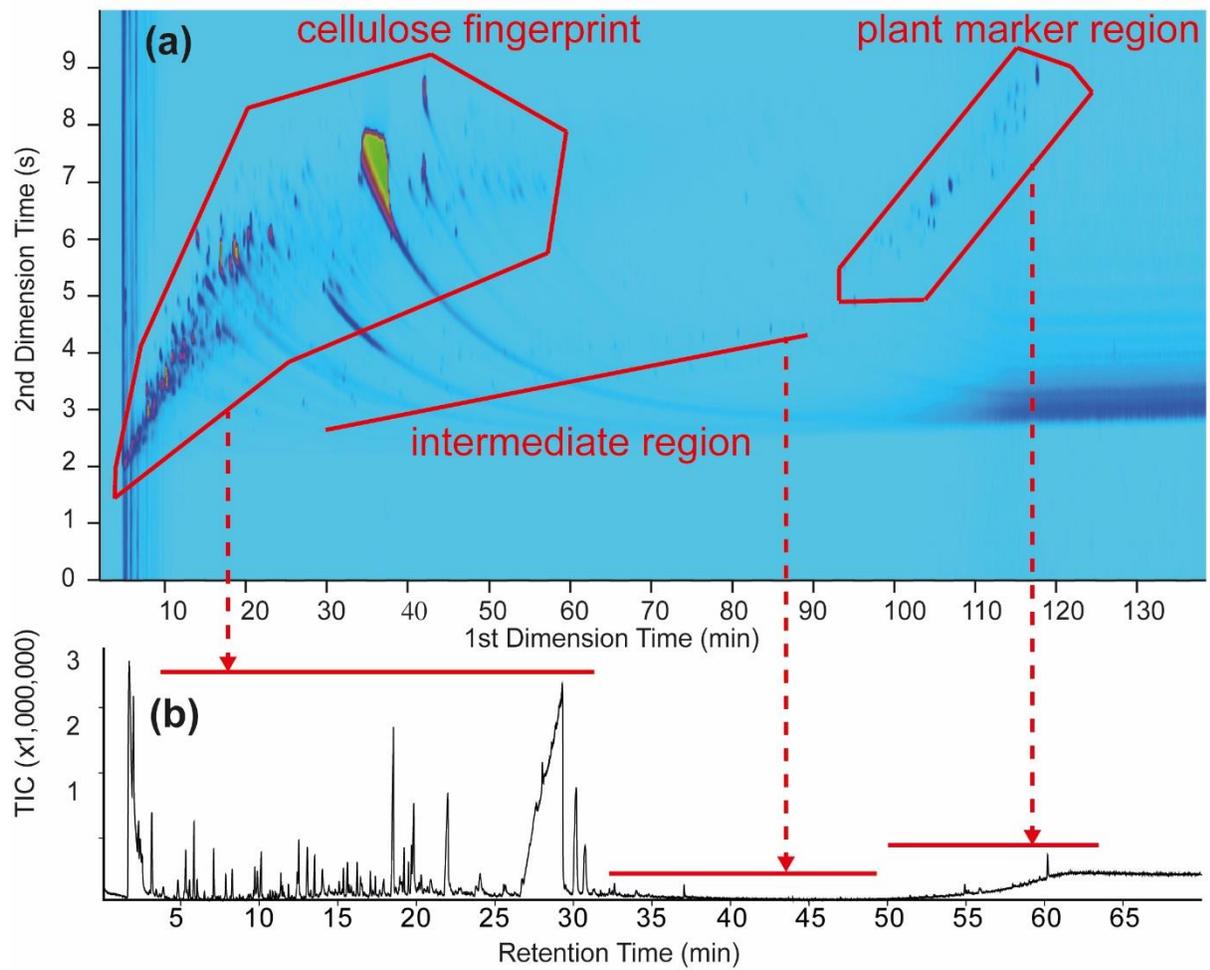


Figure 2.

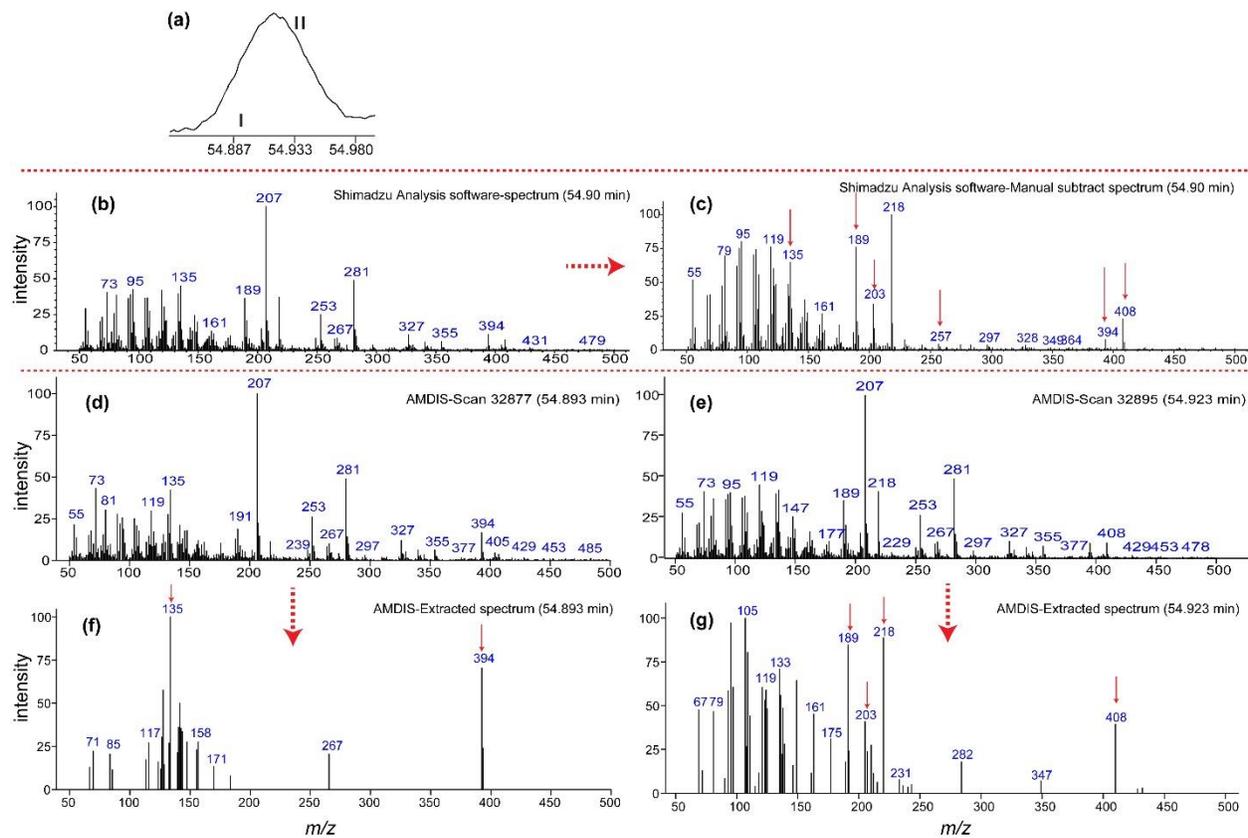


Figure 3.

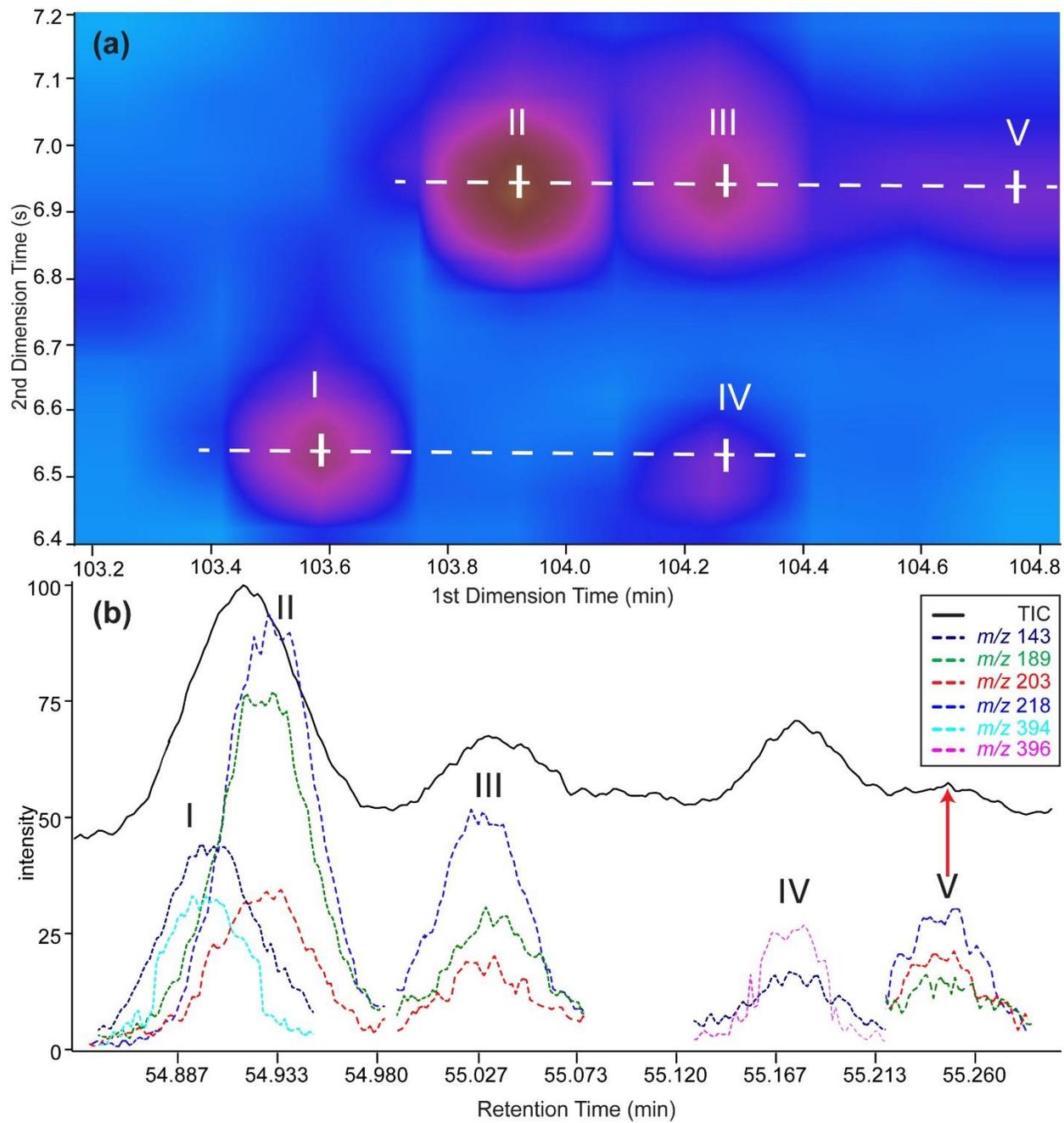


Figure 4.

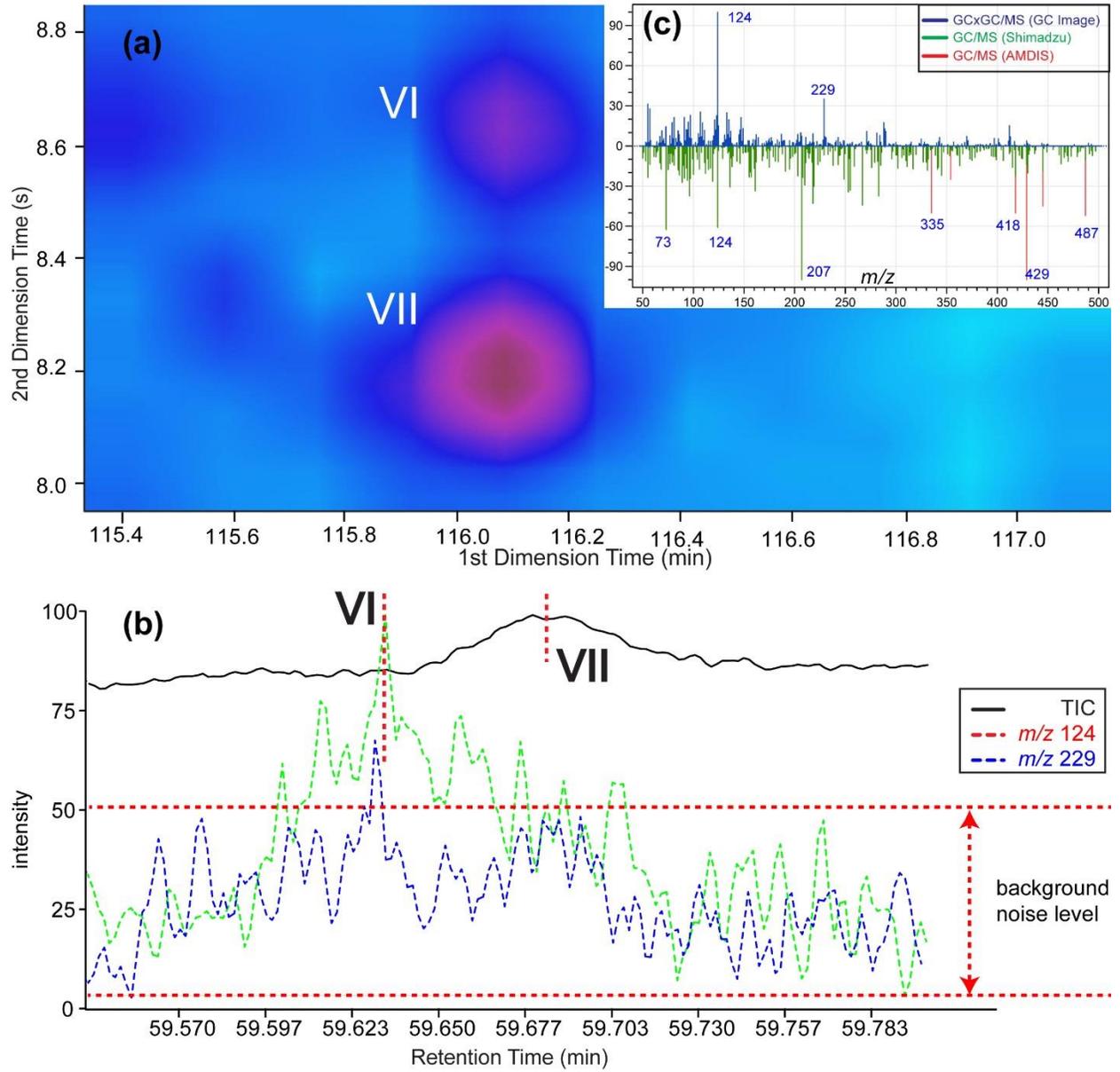


Figure 5.

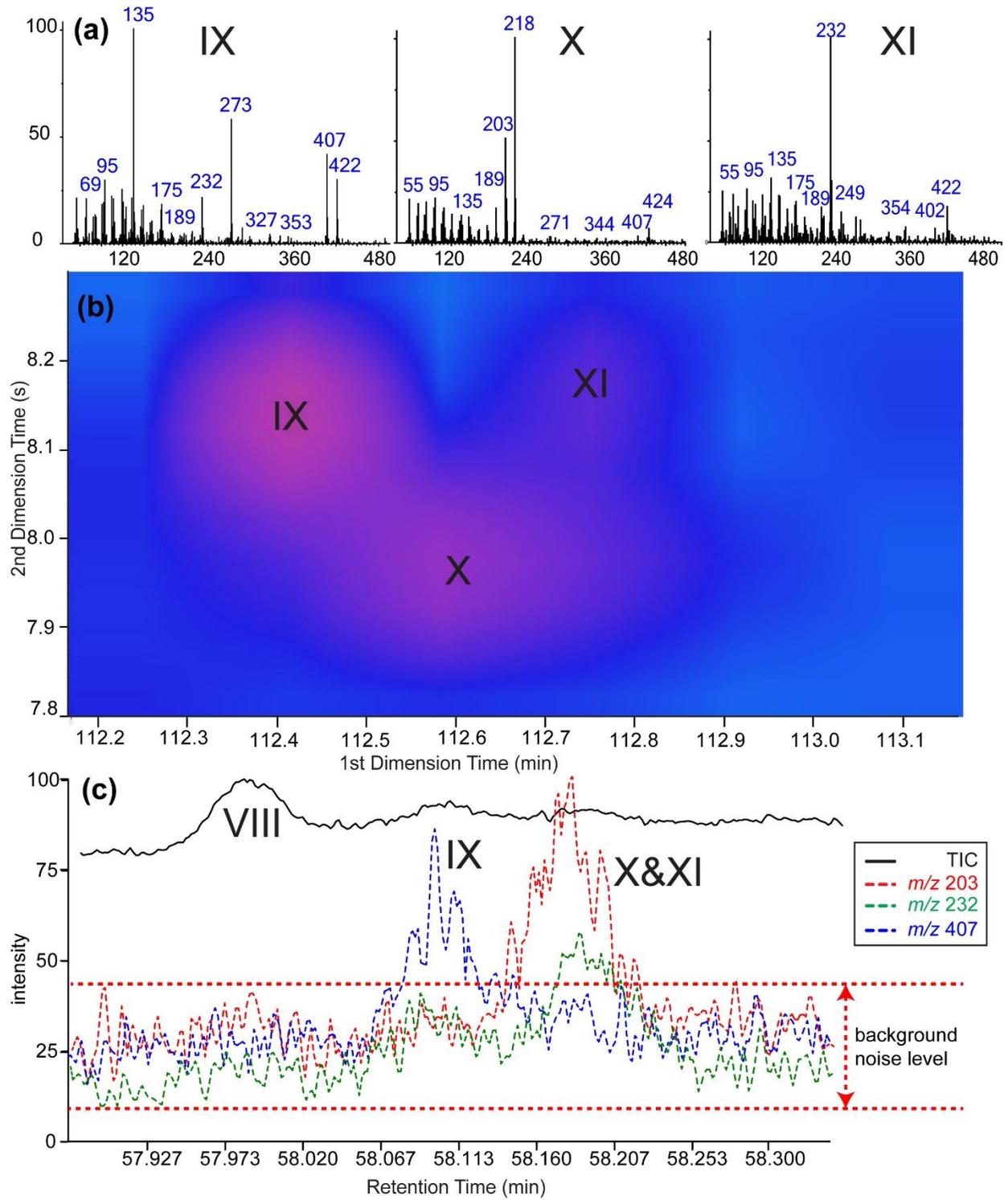


Figure 6.

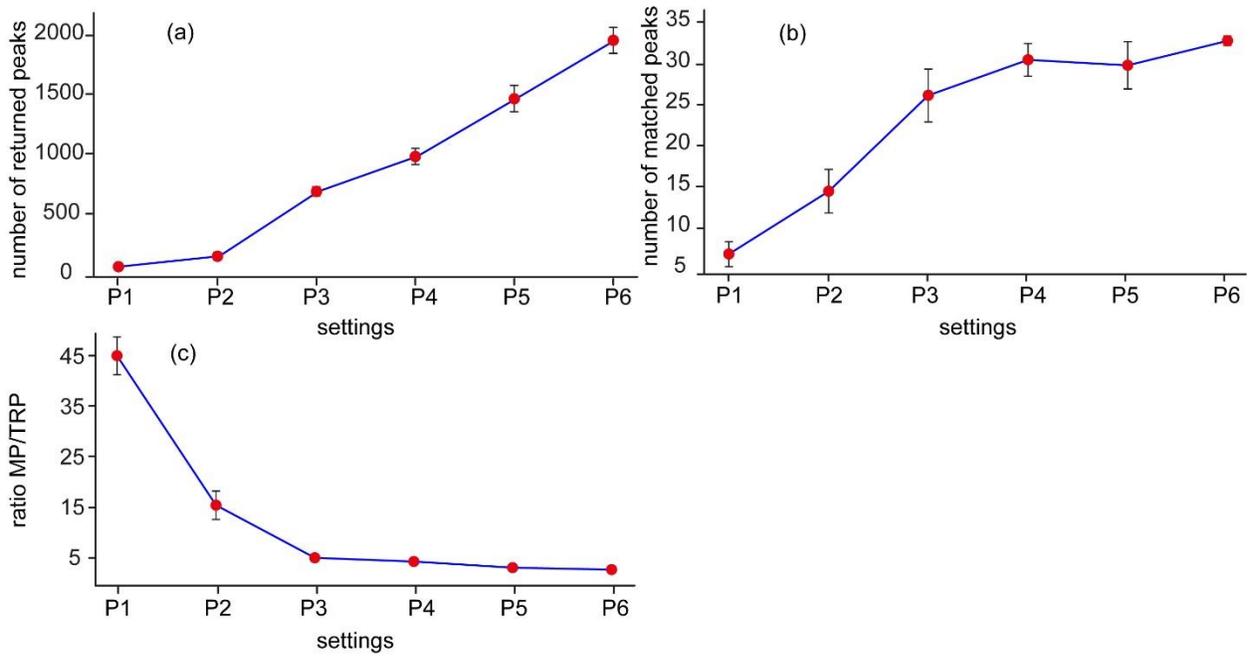


Figure 7.

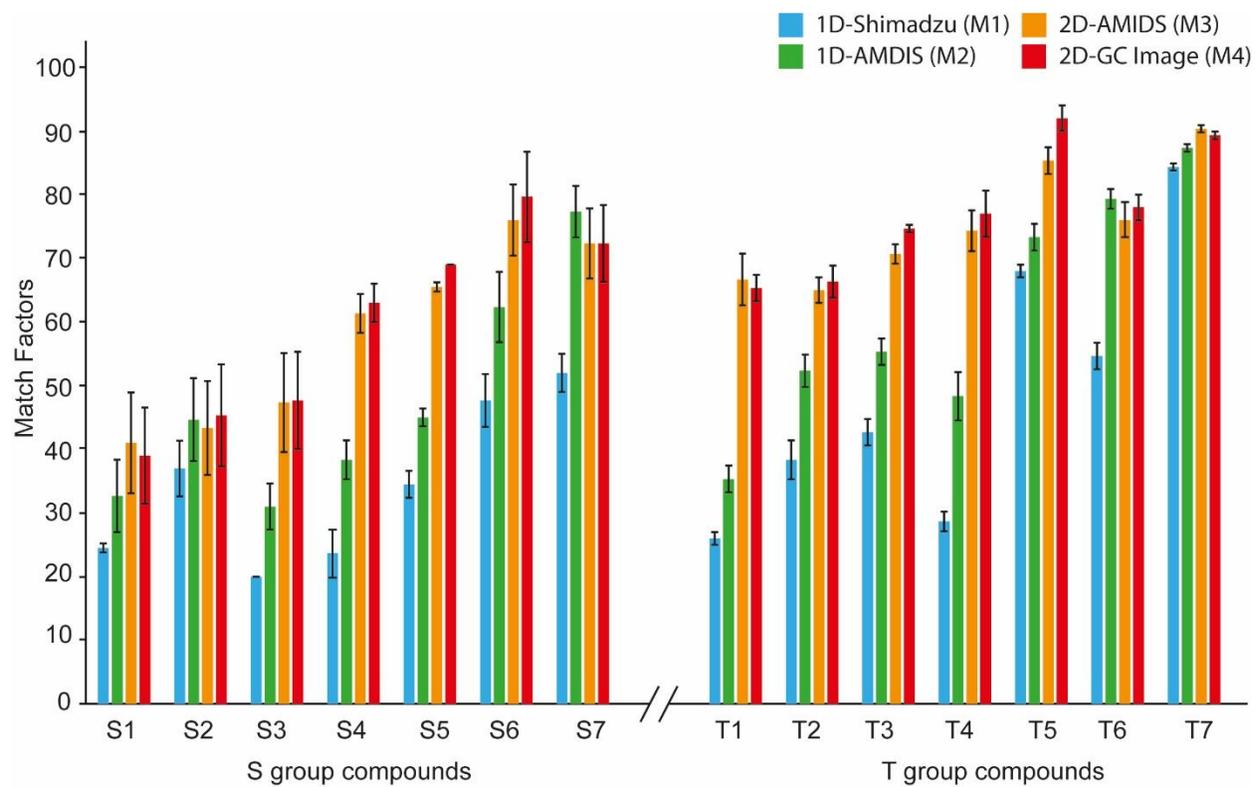


Table 1.

Retention Time	<i>m/z</i>	Assigned formula	Molecular mass	Attribution	TIC response	Label*
53.023	365, 121, 135, 93	C30H48	408	A-neooleana-3(5),12-diene	72305	S2
53.394	145, 81, 121, 95	C29H48	396	stigmasta-3,5-diene	48551	S3
53.953	406, 255, 95, 123	C30H46	406	oleanane-2,9,12-triene	118881	S4
54.343	218, 203,189, 95	C30H48	408	oleanane-2,12-diene	117283	T4
54.712	105, 57, 145, 396	C29H50O	414	NA stigmastan	91922	S1
54.917	135, 143, 394, 81	C29H48O	412	stigmasta-4,6-dien-3-ol	631803	S5*
54.917	218, 189, 93, 95	C30H48	408	urs-2,12-diene	631803	T5*
55.178	147, 81, 145, 396	C29H50O	414	stigmasta-3-ol	130403	S6
58.465	218, 203, 189, 95	C30H50O	426	beta amyrin (Olean-12-en-3-ol)	119246	T3
58.726	218, 95, 107, 219	C30H48O	424	urs-12-en-3-one	82013	T1
59.023	218, 95, 81, 189	C30H50O	426	alpha-amyrin (urs-12-en-3-ol)	151263	T2
59.277	466, 255, 121, 95	C32H50O2	466	oleanane-12,14-diene-3-acetate	62445	S7
59.678	218, 203, 189, 69	C32H52O2	468	oleanane-12-ene-3-acetate	249653	T6
60.196	218, 189, 95, 69	C32H52O2	468	alpha-amyrin acetate (urs-12-ene-3-acetate)	1438017	T7

Label\* is corresponding to the match factors illustrated in Figure 7  
 S5 and T5\* are co-eluting, the same TIC area was used for both.

## Graphic

