



**HAL**  
open science

# Combined Nuclear Magnetic Resonance Spectroscopy and Mass Spectrometry Approaches for Metabolomics

Marine P M Letertre, Gaud Dervilly, Patrick Giraudeau

► **To cite this version:**

Marine P M Letertre, Gaud Dervilly, Patrick Giraudeau. Combined Nuclear Magnetic Resonance Spectroscopy and Mass Spectrometry Approaches for Metabolomics. *Analytical Chemistry*, 2021, 93 (1), pp.500-518. 10.1021/acs.analchem.0c04371 . hal-03001696

**HAL Id: hal-03001696**

**<https://hal.science/hal-03001696>**

Submitted on 12 Nov 2020

**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.

1 **Combined Nuclear Magnetic Resonance Spectroscopy and Mass**  
2 **Spectrometry Approaches for Metabolomics**

3 Marine P .M. Letertre<sup>a</sup>, Gaud Dervilly<sup>b</sup>, Patrick Giraudeau<sup>a,\*</sup>

4 a Université de Nantes, CNRS, CEISAM UMR 6230, F-44000 Nantes, France

5 b LABERCA, Oniris, INRAE, 44307 Nantes, France

6 \* Email: [patrick.giraudeau@univ-nantes.fr](mailto:patrick.giraudeau@univ-nantes.fr) .

7 Fundamental and Applied Reviews in Analytical Chemistry 2021.

8

9

10 Metabolomics is based on cutting-edge analytical methods that provide a “snapshot” of all the  
11 detectable metabolites, small molecules generally weighing under 1200 Da, present in complex  
12 biological samples. It presents a wide scope of applications and has allowed considerable progress to  
13 be made in health and disease research,<sup>1,2</sup> pharmaceutical sciences,<sup>3-5</sup> personalized medicine,<sup>6-10</sup>  
14 microbiome research,<sup>11,12</sup> but also in food and nutrition,<sup>13-16</sup> agriculture,<sup>17,18</sup> marine environmental  
15 research<sup>19,20</sup> or exposome research.<sup>21-23</sup> Metabolomics can be undertaken through two different  
16 approaches. Untargeted methods aim at capturing a broad view of all the metabolites present in a  
17 biological sample, without any *a priori*, up to the limit of the accessible metabolites which depends on  
18 the detection limit of the applied analytical method, the physicochemical properties of the analytes,  
19 as well as the sample handling or preparation applied. An untargeted assay often aims to identify one  
20 or several new biomarkers of a particular phenotype, which can be for instance markers of interest of  
21 a specific disease, or markers of effect following an exposition to a physical or a chemical stress or a  
22 therapeutic treatment, and to elucidate their structures. It can also help to build a model capable to  
23 predict a specific condition, such as in foodomics. Once untargeted analyses have been performed and  
24 that effect biomarkers have been discovered, or if there are known exposure biomarkers of interest  
25 (e.g., xenobiotic such a specific drug or a chemical contaminant and their metabolites), a targeted  
26 approach will make it possible to perform a quantitative analysis of those compounds. In the case of  
27 biomarker discovery, quantitative insights through targeted analysis are often needed to validate that  
28 a metabolite is indeed a real biomarker. Halfway between untargeted and targeted approaches stands  
29 a slightly different strategy, where compounds from the same class of metabolites or a particular  
30 biochemical pathway (e.g., bile acids or amino acid metabolism) need to be broadly captured and if  
31 possible in a quantitative way but when the rest of the metabolome is not relevant to the research  
32 problem.

33 Currently, the two main analytical techniques used to apply those approaches are nuclear  
34 magnetic resonance (NMR) spectroscopy, based most of the time on the detection of <sup>1</sup>H or <sup>13</sup>C nuclei,  
35 and mass spectrometry (MS), often coupled with separation techniques such as liquid or gas  
36 chromatography (LC or GC), capillary electrophoresis (CE), or ion mobility (IM). NMR spectroscopy is a  
37 noninvasive technique, as the sample can be recovered and used in a following experiment (providing  
38 that the sample preparation needed for NMR analysis such as D<sub>2</sub>O addition, does not interfere with  
39 the following experiment). In contrast, MS analysis, due to the nature of the technique, is destructive  
40 but, as only relatively small volumes are required (as low as few microliters of an often very diluted  
41 sample), this does not necessarily cause problems. NMR spectroscopy also presents the advantage of  
42 providing accurate quantitative results, making possible the quantitation of multiple analytes with a  
43 single internal or external reference.<sup>24</sup> Furthermore, <sup>1</sup>H NMR spectroscopy has the strong advantage

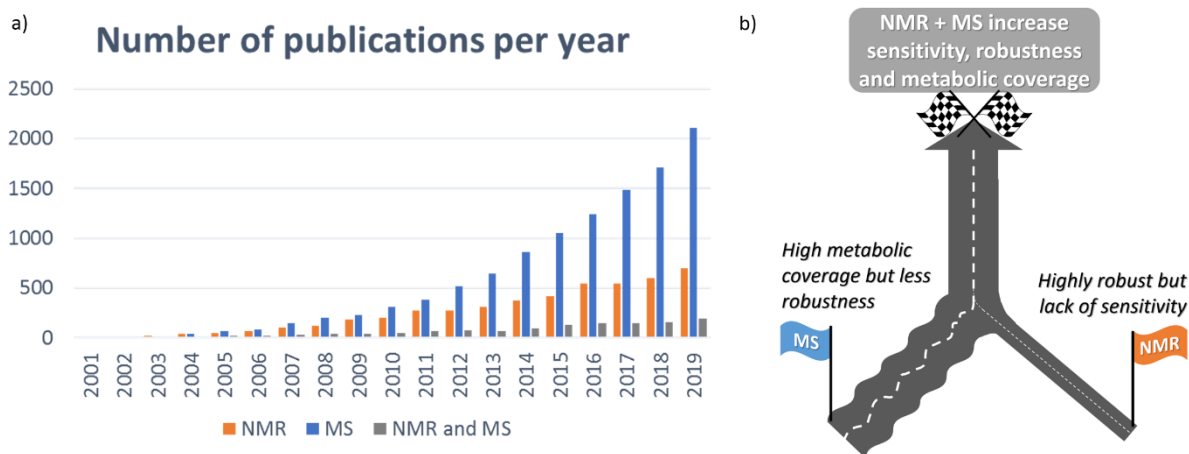
44 of being a robust and reproducible technique, both through time and between laboratories.<sup>25</sup>  
45 However, the main drawback of NMR spectroscopy is its relative lack of sensitivity, together with  
46 ubiquitous signal overlap in the <sup>1</sup>H NMR spectrum of biological samples, which limits the identification  
47 of metabolites and the discovery of significant biomolecular changes and biomarkers. Recent progress  
48 has been made to overcome such limitations. The overlap issue has been solved by including the  
49 acquisition of 2D NMR data sets in metabolomics workflows to reduce signal overlap while providing  
50 crucial information to elucidate the structure of metabolites.<sup>26,27</sup> While 2D NMR suffers from long  
51 acquisition times, numerous methodological developments made it possible to reduce acquisition  
52 times to a few minutes for biological samples, and the quantitative issues associated with 2D NMR  
53 have also been addressed through pulse sequence developments or calibration strategies.<sup>28-30</sup> More  
54 recently, proof-of-concept developments have been addressing the sensitivity issues of NMR  
55 metabolomics, notably through hyperpolarization strategies such as dissolution dynamic nuclear  
56 polarization (d-DNP)<sup>31-33</sup> or para-hydrogen induced polarization (PHIP).<sup>34,35</sup> These techniques offer an  
57 unprecedented boost in NMR sensitivity by several orders of magnitude; however, they are very recent  
58 and not yet implemented in routine for metabolomics.

59 MS-based techniques are the most widely implemented strategies for metabolomics purposes  
60 (Figure 1), especially UPLC-MS with electrospray ionization (ESI), thanks to the greater sensitivity that  
61 this technique offers. This fact has been reinforced in recent years thanks to the development of high-  
62 resolution (HR)-MS techniques and the possibility to determine the accurate mass of a compound.  
63 However, these techniques are somewhat less robust than NMR ones, and although targeted assays  
64 are generally comparable between laboratories,<sup>36,37</sup> untargeted methods require careful quality  
65 control (QC, biological pool sample) procedures to assess robustness and repeatability over time.<sup>38-41</sup>  
66 Also, the important sensitivity that MS and especially HRMS techniques offer comes with its drawback,  
67 such as ion suppression. However, using multiple ionization modes (positive and negative) for ESI and  
68 several chromatographic systems (reversedpPhase [RP] and hydrophilic interaction chromatography  
69 [HILIC]), as well as any other strategy enabling signals deconvolution such as IM,<sup>42</sup> are clear advantages  
70 to increase metabolite coverage.

71 Metabolomics greatly benefits from the tremendous progress made in both MS and NMR in the  
72 past couple of decades, in terms of sensitivity, resolution, and rapidity, but also from advances in  
73 statistical analysis and bioinformatics methods. However, most of these recent advances are costly,  
74 time-consuming, and require advanced technical skills, which makes them not easily accessible.  
75 Furthermore, none of the analytical methods existing today allow a full capture of the metabolome.  
76 This is due to multiple factors, such as sensitivity limitations, loss of metabolites during sample  
77 collection, handling, preparation, and analysis (e.g., nonretention/nonelution, ionization efficiency,

78 signal overlapping...). In consequence, the use of both routine NMR spectroscopy and MS-based  
79 techniques through an integrated platform is a sensible and powerful option to maximize metabolome  
80 coverage, facilitate metabolite identification and biomarkers discovery, and build more robust models  
81 through the use of multiple data set integrations.<sup>43,44</sup>

82 The combined use of NMR and MS has been long exploited for the structural characterization of  
83 new metabolites of interest, especially in the natural products field.<sup>45</sup> Analysis of 1D and 2D NMR  
84 spectra to extract chemical shift and coupling information, as well as direct infusion MS to obtain an  
85 exact  $m/z$  ratio and fragmentation patterns, help to identify and structurally define new metabolites  
86 following several steps of purification. In metabolomics, many published studies relied on only one of  
87 the two techniques, mainly for opportunistic reasons, i.e., researchers focusing on metabolomics  
88 studies were using the closest available technique in their laboratory or institute. Figure 1 shows that  
89 the proportion of MS versus NMR in metabolomics has been increasing over time, mainly for cost and  
90 sensitivity reasons. However, Figure 1 also sheds light on a still modest but substantial increase of the  
91 number of studies that make use of both techniques, suggesting that the combination of MS and NMR  
92 for metabolomics could be highly valuable. This review focuses on the description of such powerful  
93 combination, which can be done through various ways. NMR and MS can be combined at the hardware  
94 level through physical association of the two techniques. However, in most cases, it relies on  
95 integrating the respective data sets in a common chemometric software and multivariate statistical  
96 analysis pipeline. Such integration can be performed at different levels, through cross-comparison,  
97 correlation, or multiblock integration. In the review herein, we describe the principles of such  
98 combination, highlighting how it has provided a considerable paradigm shift in metabolomics in the  
99 past few years. The benefits of gathering these techniques in postmetabolomics workflow through  
100 more targeted approaches to improve metabolite identification, quantitative assays, and fluxomic  
101 analysis are also exposed, before discussing the perspectives of integrating several metabolomics and  
102 omics methods in general. Individual methodological advances made in NMR or MS metabolomics are  
103 excluded from the scope of this review, but they have been thoroughly reviewed recently<sup>46,47</sup> and will  
104 certainly benefit to combined MS and NMR approaches.



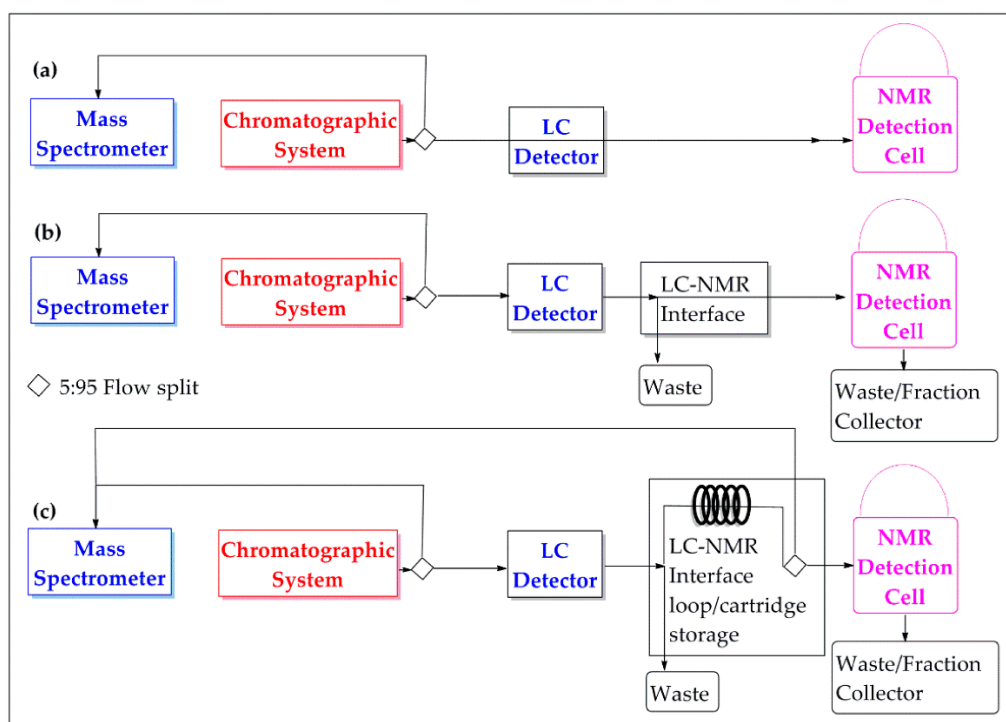
105

106 Figure 1. (a) Results obtained by searching through *pubmed.ncbi.nlm.nih.gov* in [Title/Abstract], the following  
 107 terms: “(nuclear magnetic resonance OR NMR) AND (metabolomics OR metabonomics OR metabolic profiling)”,  
 108 shown in orange; “(mass spectrometry OR MS) AND (metabolomics OR metabonomics OR metabolic profiling)”,  
 109 shown in blue; “(nuclear magnetic resonance OR NMR) AND (mass spectrometry OR MS) AND (metabolomics OR  
 110 metabonomics OR metabolic profiling)”, shown in gray. Research was done on September 18, 2020. (b)  
 111 Schematic representation of the advantages and the drawbacks of NMR and MS-based analytical methods and  
 112 the benefits of combining them.

### 113 NMR and MS Hardware Combination

114 As stated above, the combination of several NMR (e.g., 1D/2D,  $^1\text{H}/^{13}\text{C}$ ) or MS platforms (e.g.,  
 115 LC/GC, ESI+/-, RP/HILIC) increases the metabolic coverage, and the combined use of both NMR and MS  
 116 platforms can also help toward that same objective. Such a combination was initially applied in the  
 117 natural products research field, in order to help with structural elucidation, through an off-line  
 118 platform based on the comparison of NMR chemical shifts and coupling constants as well as HRMS to  
 119 obtain exact  $m/z$  and fragmentation patterns. This hyphenation led to the development of online  
 120 platforms including both NMR and MS hardware, often preceded by an LC system. This type of system  
 121 found applications in drug metabolism research or drug discovery from natural products, especially to  
 122 help with dereplication, consisting of identifying known natural compounds from active fractions to  
 123 avoid spending time on compounds which had already previously been discovered.<sup>48</sup> The hyphenated  
 124 use of LC-MS-NMR was achieved thanks to postcolumn splitters, which send 10% of the outgoing flow  
 125 from the LC column to the MS system and the remaining 90% to the NMR system.<sup>44</sup> A commercial  
 126 NMR-MS interface was also developed, composed of a splitter controlled by the operator and a double  
 127 dilutor.<sup>44</sup> The latter allows, on the one hand, the prevention of an extensive use of deuterated solvent  
 128 as the sample is mixed in  $\text{D}_2\text{O}$  just before entering the NMR system rather than before entering the LC  
 129 system and, on the other hand, to dilute the sample in the appropriate solvent for ionization and MS  
 130 detection.<sup>44</sup>

131 Three different ways exist to set up such a combined platform, namely, through a continuous-flow  
132 mode, a stop-flow mode or a storage mode (Figure 2).<sup>49</sup> By using the dynamic continuous-flow mode,  
133 the sample already mixed with deuterated solvent is sent to the LC system before flowing separately  
134 and continuously in the MS and the NMR systems. One of the first examples of this continuous-flow  
135 mode was made by Shockcor and co-workers in 1996 to analyze a urine sample from an individual  
136 administrated with paracetamol, in order to identify with more confidence paracetamol metabolites  
137 and urinary endogenous compounds.<sup>50</sup> Phenylacetylglutamine, not previously detected by <sup>1</sup>H NMR  
138 spectroscopy alone due to spectral overlapping but usually detected by HPLC-MS and confirmed by  
139 the use of a standard, was well identified thanks to the good resolution obtained by this HPLC-NMR-  
140 MS system.<sup>50</sup> The NMR data set made it possible to clearly define which paracetamol-glucuronide  
141 isomer was observed, and this would not have been possible by the unique use of HPLC-MS or HPLC-  
142 MS/MS alone without comparison to an internal standard. However, this technique is limited by the  
143 time-evolving LC gradient composition which induces a bias in the NMR measurement. Indeed, this  
144 leads to an evolution of the position of the solvent peaks which makes it difficult to maintain an  
145 efficient solvent signal suppression over time.<sup>49</sup> Most importantly, the short residence time of nuclear  
146 spins in the NMR detection cell strongly limits signal averaging, which in turns impacts the resulting  
147 NMR signal-to-noise ratio (SNR). This issue can be addressed by using a static analysis such as the stop-  
148 flow mode, during which a valve pauses the LC flow when a peak is detected or selected and that the  
149 corresponding analyte has reached the NMR detection cell. This approach provides enough time for  
150 the NMR measurement to be performed with a satisfactory SNR. However, stop-flow broadens the LC  
151 peaks thus limiting the chromatographic resolution. As a consequence, a storage mode has often been  
152 preferred, either performed through an online or off-line setup. For this mode, the different fractions  
153 coming out from the LC system can be collected and stored in a loop while the NMR analysis is running.  
154 The sample collection can also be done in a cartridge, most of the time a solid-phase extraction (SPE)  
155 cartridge, which receives 95% of the LC eluent (the other 5% being sent toward the MS system) and  
156 which efficiently retains and concentrates analytes before NMR characterization, preventing an  
157 extended use of deuterated solvent.<sup>51</sup>



158

159 Figure 2. Schematic setups of the different LC–NMR working modes with parallel mass spectrometer (MS)  
 160 detection: (a) online/continuous-flow mode, (b) stop-flow mode and (c) loop/cartridge storage mode. Figure  
 161 reproduced from ref. <sup>49</sup> under Creative Commons Attribution 4.0 International License  
 162 (<http://creativecommons.org/licenses/by/4.0/>).

163 LC-MS-SPE-NMR has been elegantly applied to help with the structural elucidation of urinary  
 164 phenolic compounds in humans following tea consumption, once classical off-line LC-Orbitrap Fourier  
 165 transform MS (LC-FTMS) and 1D <sup>1</sup>H NMR analysis were performed separately to select features of  
 166 interest for further characterization.<sup>52</sup> This approach efficiently provides comprehensive structural  
 167 confirmation of the fragmentation patterns of the selected feature, while simultaneously providing  
 168 quantitative data based on the <sup>1</sup>H NMR spectroscopy part of the system. This hyphenated system has  
 169 found successful applications in natural product research, especially when it comes to structural  
 170 elucidation or to differentiate isomeric or isobaric compounds.<sup>49</sup> However, several drawbacks still limit  
 171 its widespread use, such as the bulky and expensive equipment, associated with a limited sensitivity.  
 172 The latter drawback has pushed toward the development of integrated microprobe technologies and  
 173 capillary separation.<sup>49,53</sup> Lin and co-workers reported an LC-MS-NMR platform with microscale  
 174 instruments, namely, a nanoSplitter LC-MS and a microdroplet NMR, for increased sensitivity.<sup>54</sup> The  
 175 use of a highly sensitive instrument such as a nanoelectrospray MS, requested only 2% of the LC  
 176 column eluent, driving away the rest to a UV-guided collection for concentration prior to be stored  
 177 waiting for an off-line NMR analysis. This setup limited the use of deuterated solvent as it was added  
 178 just before the NMR analysis, which can thus be done retrospectively once the LC-MS analysis have  
 179 been performed. However, it reintroduced an additional step of sample handling prior to NMR



180 analysis. And although the sensitivity was further optimized in this application by using a microcoil  
181 NMR probe,<sup>54</sup> the robustness of such equipment is questionable,<sup>53</sup> which is a major bottleneck to  
182 create high-quality databases aiming to accelerate structural characterization of low concentration  
183 analytes in complex matrices.<sup>55</sup>

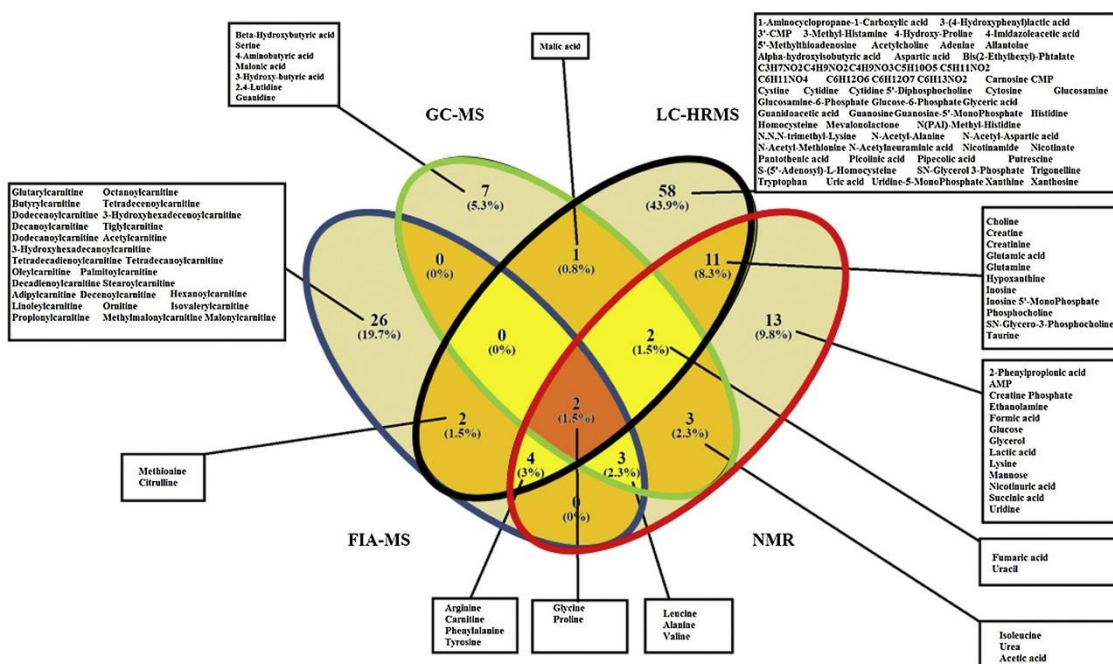
184 Beyond its ability to help with structural elucidation in natural product research and the efforts  
185 made to minimize sample handling and preparation while increasing its sensitivity, the hyphenated LC-  
186 MS-NMR system did not receive the success expected by some in the past decade. The known  
187 limitations of such a platform, namely, the incompatibility of the solvents for MS and NMR or the  
188 extended use of expensive deuterated solvent, the low flow rate for efficient ionization and MS  
189 detection, and the long acquisition time requested for a sufficient NMR sensitivity,<sup>56</sup> were not  
190 overcome. As such, it seems unlikely that such a system will be further encouraged by the  
191 metabolomics community, especially when promising computed-based combination solutions are  
192 being developed.

## 193 **NMR and MS Data Set Combination for Metabolomics Analysis**

### 194 **Cross-Comparison of NMR and MS Data Sets to Increase Metabolic Coverage**

195 Although good convergence was found between different techniques in interlaboratory studies,<sup>57</sup>  
196 individual analytical techniques do not necessarily cover the same types of metabolites. The parallel  
197 use of NMR and MS methods can highly improve the quality of metabolomics studies in a variety of  
198 ways. The most obvious case where the combination of the two techniques can be beneficial is the  
199 increase of metabolic coverage.<sup>58</sup> This was illustrated, for instance, in a study aiming to investigate the  
200 biomolecular processes behind the mycotoxins production of cereals infected by the plant pathogen  
201 *Fusarium graminearum*.<sup>59</sup> In this work, NMR spectroscopy and LC-QTOF-MS (based on a reversed-phase  
202 stationary phase) analyses were performed to measure polar and semipolar compounds, respectively.  
203 In total, 15 amino acids or derivatives, 3 sugars and polyols, 4 tricarboxylic acid (TCA) organic acid  
204 derivatives and 4 nucleosides and nucleotides were identified or putatively annotated by NMR or 2D  
205 NMR, while 55 sesquiterpenes and 10 polyketides were highlighted by MS or MS/MS.<sup>59</sup> None of the  
206 metabolites identified or annotated by one technique were claimed to be identified by the other. More  
207 studies with key numbers highlight the advantage of combining NMR and MS method to increase the  
208 metabolic coverage. Goulitquer et al. combined one <sup>1</sup>H NMR, one GC-MS, and seven LC-MS data sets  
209 to explore the changes induced in the metabolome and the lipidome of human gastric cancer cells  
210 following treatment with anticancer drugs.<sup>60</sup> The LC-MS data sets were acquired on three different  
211 instruments (UHPLC-LTQ-Orbitrap, UHPLC-Exactive, UPLC-HRMSe Q-TOF) and provided analyses in  
212 both positive and negative modes. Out of the 111 metabolites and lipids annotated, only 9 were

213 common to the LC-MS and the GC-MS data sets, 4 between the LC-MS and the NMR, 2 between the  
 214 GC-MS and the NMR, and 6 were concordant between the three platforms. A recent study successfully  
 215 attempted to capture the broadest picture of the human serum metabolome.<sup>61</sup> To do so, five analytical  
 216 platforms were used, namely, <sup>1</sup>H NMR, GC-MS, LC-ESI-MS/MS, TLC/GC-FID-MS and DI-MS. Over 3500  
 217 distinct metabolites were identified, and from those, only 29 were commonly identified by NMR and  
 218 GC-MS, 13 by NMR and DI-MS, 14 between GC-MS and DFI-MS, 8 between the three analytical  
 219 methods just cited, and 53 between DFI-MS and TLC/GC-FID-MS.<sup>61</sup> It should be noted that the  
 220 spectacular effort made through this study was complemented with an extensive literature research,  
 221 called “bibliomic”, which found 665 other serum metabolites already reported in the literature but not  
 222 detected by the five analytical platforms applied. Quantitative data were also reported for a portion  
 223 of the over 4000 metabolites, showing acceptable agreement between the concentrations obtained  
 224 from the different analytical methods but still with some exceptions.<sup>61</sup> In a similar way, the combined  
 225 use of NMR, FIA-MS/MS, GC-MS, and LC-HRMS was applied to explore the skeletal muscle  
 226 metabolome,<sup>62</sup> in order to assess their performance as well as different sample extraction protocols.  
 227 Here again, only 2 metabolites were commonly detected by the four analytical methods, 2 metabolites  
 228 common to LC-HRMS, GC-MS, and NMR, 3 common ones between FIA-MS, GC-MS, and NMR, and 4  
 229 common ones between LC-HRMS, FIA-MS, and NMR (Figure 3).<sup>62</sup> The GC-MS covered 7 unique  
 230 metabolites, against 13 ones for the NMR, 26 for the FIA-MS and 58 for the LC-HRMS. Although these  
 231 numbers suggest that GC-MS was the less efficient tool to study such samples,<sup>62</sup> some of the 7  
 232 metabolites detected by GC-MS could be of crucial importance to understand key metabolic pathway  
 233 alterations or biomarker discovery.



235 Figure 3. Venn diagram representing specificity and overlap of metabolites reliably detected by each analytical  
236 method. Red, blue, green, and black circles represent metabolites analyzed respectively, by NMR, FIA-MS, GC-  
237 MS, and LC-HRMS. Crossed zone represents overlaps between methods. Figure reprinted from J. Pharm.  
238 Biomed. Anal., Vol. 148, Bruno, C.; Patin, F.; Bocca, C.; Nadal-Desbarats, L.; Bonnier, F.; Reynier, P.;  
239 Emond, P.; Vourc'h, P.; Joseph-Delafont, K.; Corcia, P.; Andres, C. R.; Blasco, H. The Combination of  
240 Four Analytical Methods to Explore Skeletal Muscle Metabolomics: Better Coverage of Metabolic  
241 Pathways or a Marketing Argument?, p 273-279 (ref <sup>62</sup>). Copyright 2018, with permission from Elsevier.

242 Several examples have shown how increasing the metabolic coverage with a second metabolomics  
243 approach can help gaining a deeper understanding of a biological process. For instance, Allwood et al.  
244 used this strategy to understand fragrance and maturity development in five melon cultivars,<sup>63</sup> mainly  
245 depending on volatile organic compounds (VOCs). As it is often the case to measure VOCs, GC-MS was  
246 used, with thermal desorption. Optimization of the sampling method based on the use of a  
247 polydimethylsiloxane membrane allowed the detection of 58 VOCs.<sup>63</sup> Principal component analysis  
248 (PCA), heatmap, and relative changes highlighted clear differences between five melon cultivars, which  
249 were further supplemented with quantitative measures of amino-acids, known precursors of certain  
250 VOCs, by <sup>1</sup>H NMR spectroscopy. Reverse correlations between amino acids concentrations and VOCs  
251 levels were observed, clearly showing that consumption of amino acids to produce VOCs directly  
252 influence melon fragrance and maturity.

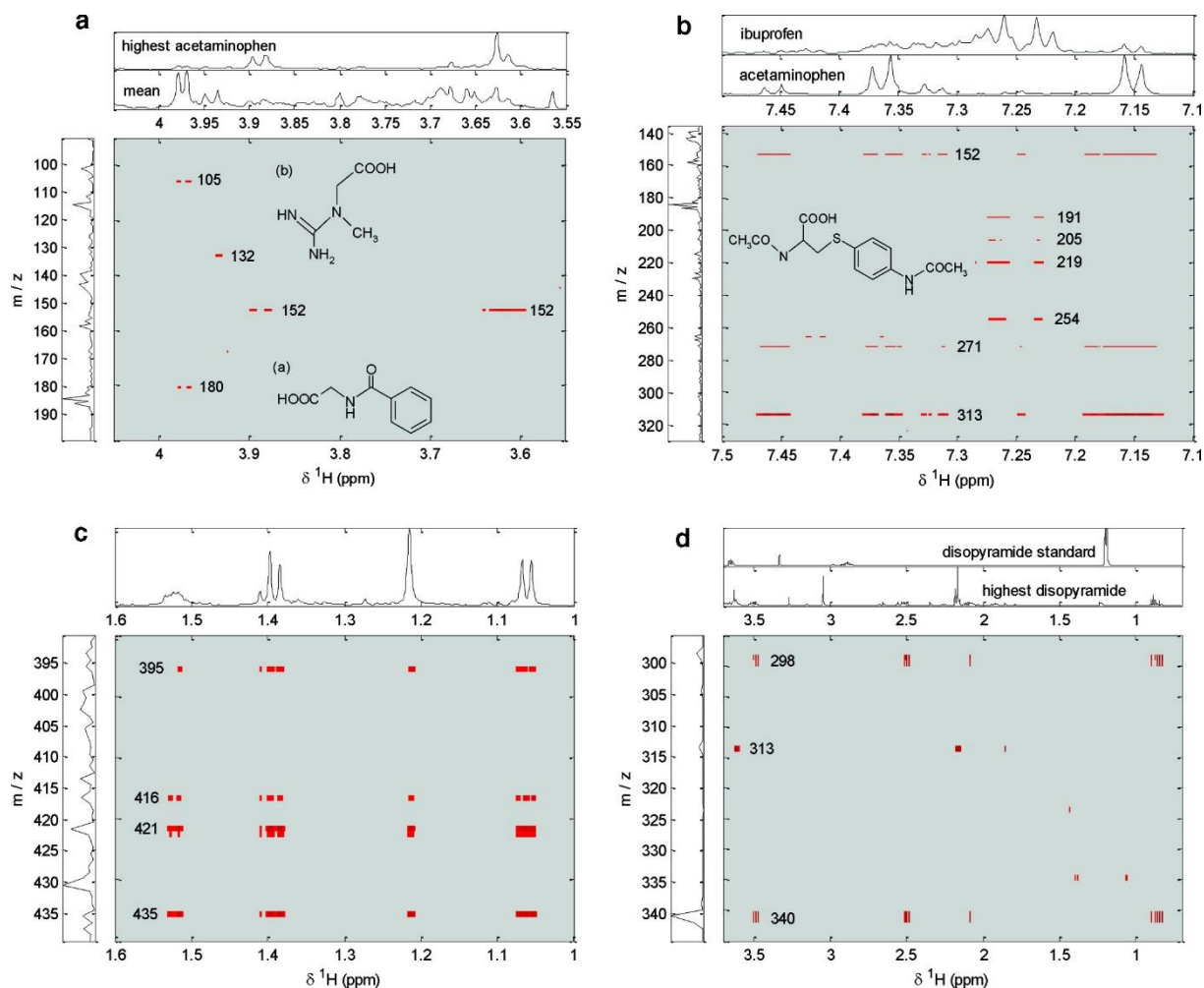
253 Still, contradictory results are sometimes encountered when both techniques are applied to the  
254 same matrix. For instance, different concentrations can be measured in the NMR and the MS data set,  
255 or a compound can be detected by one technique and not by the other, which is actually the reason  
256 why using both in combination increases the metabolic coverage. A typical example of this problematic  
257 has been reported by Atherton et al. in 2006,<sup>64</sup> where metabolic profiling of several tissues from  
258 control or peroxisome proliferator-activated receptor-alpha (PPAR-alpha) null mice were explored by  
259 using <sup>1</sup>H NMR spectroscopy, high-resolution magic angle spinning (HR-MAS) <sup>1</sup>H NMR spectroscopy, GC-  
260 MS, and LC-MS. In this study, significant changes were observed in the cardiac metabolic profile of the  
261 control vs muted mice by <sup>1</sup>H NMR, HR-MAS <sup>1</sup>H NMR, and GC-MS, but the order of magnitude of these  
262 changes were different from one method to the other. Indeed, as <sup>1</sup>H NMR spectroscopy has limited  
263 sensitivity, only the most concentrated compounds appeared significantly different in the PCA model,  
264 while the number of metabolites significantly impacting the GC-MS partial least-square discriminant  
265 analysis (PLS-DA) model were 5 times more numerous.<sup>64</sup> This was explained by the fact that  
266 metabolites with hydroxyl and amine functional groups were easily detected by GC-MS, even though  
267 they might not be the most concentrated metabolites. This constitutes a perfect representation of how  
268 difficult it is to catch, within a complex biological sample, metabolites present in a broad variety of  
269 concentration, polarity, and mass range.<sup>64</sup> On a side note, in order to increase the number of  
270 metabolites detected by both NMR and MS techniques, an elegant approach is to use a smart isotope

271 tag, such as  $^{15}\text{N}$ -cholamine, which present the advantage of being a sensitive isotope for NMR and of  
272 having a permanent charge for MS efficient detection.<sup>65</sup> Chemical derivatization with this smart tag  
273 allows the simultaneous detection of carboxylic acid derivatives without ambiguity. However, this  
274 approach is restricted to metabolites containing a carboxyl group and thus covers a limited part of the  
275 metabolome.

## 276 **Correlation of NMR and MS Data Sets**

277 Rather than taking into consideration the different variables within a sample, correlation analysis  
278 is based on taking into consideration the intensities of the same variable across different samples.<sup>66</sup>  
279 The first correlation tools that appeared in metabolomics focused on the 2D correlation of vibrational  
280 spectroscopic data, such as IR or Raman.<sup>67</sup> Numerous statistical tools for spectroscopic correlation  
281 arose from it but mainly for the interpretation of NMR data sets.<sup>68</sup> One of the most common tools is  
282 based on statistical total correlation spectroscopy (STOCSY), which correlates signals showing similar  
283 variations across samples within 1D  $^1\text{H}$  NMR spectra, in order to better extract individual metabolite  
284 spectral patterns and facilitate the identification of biomarkers.<sup>69</sup> It was further adapted to several  
285 other statistical tools,<sup>68</sup> such as Het-STOCSY, to correlate heteronuclear NMR signals, STOCSY-editing,  
286 which aimed to correlate only the endogenous compounds without taking into consideration the  
287 exogenous ones, or also statistical heterospectroscopy (SHY), which aimed to correlate signals from  $^1\text{H}$   
288 NMR with LC-MS ones. Through this tool, the intrinsic covariance of the NMR chemical shifts and the  
289  $m/z$  signal intensities of the same features is analyzed to help biomarker discovery and achieve a  
290 deeper understanding of the biological alterations due to a specific drug treatment or disease. The  
291 efficiency of the SHY method was illustrated through a proof-of-concept study measuring the effect of  
292 hydrazine treatment in rat urine samples.<sup>70</sup> Prior to the correlation, the data sets need to be formatted.  
293 As such, a cubic spline was used to smooth the NMR spectra, and MS spectra were binned to produce  
294 2D histograms which were further summed by a specific retention time window to create pseudo  
295 direct infusion spectra and prevent the loss of the LC-MS signals eluting closely to the chromatographic  
296 dead volume.<sup>70</sup> Subsequent normalization was required to take into consideration the different  
297 dilution factors of the urinary metabolites. Correlation coefficients were calculated by using a Pearson  
298 correlation and visualized according to a specified cutoff.<sup>70</sup> This powerful tool was then applied to  
299 human urine samples within an epidemiological study where the studied population was not  
300 controlled or selected.<sup>71</sup> Even so, the use of therapeutic treatment was easily detected in the samples,  
301 and the additional use of  $\text{MS}^E$  (the combined use of low and high collision energy to simultaneously  
302 detect the precursor and the fragments  $m/z$ ) data allowed the annotation of unreported drug  
303 metabolites (Figure 4). Together with the detection of more common endogenous metabolites, it

304 showed that SHY is a useful statistical tool to explore the xenometabolome and its effect on metabolic  
 305 phenotypes.<sup>71</sup>



306  
 307 Figure 4. SHY plots: (a) Correlation (cutoff 0.7) of a hippurate NMR signal (doublet at 3.97 ppm) with the hippurate  
 308 parent ion (m/z 180, neutral molecule shown as inset a) and a fragment due to cleavage of glycine in-source (m/z  
 309 105). Also shown are correlations due to creatine (singlet at 3.93 ppm, m/z 132, neutral molecule shown as inset  
 310 b) and acetaminophen-related signals (doublet at 3.89 ppm, multiplet at 3.62 ppm, m/z 152). The inset NMR  
 311 spectra are the mean spectrum and the spectrum for the sample with the highest concentration of  
 312 acetaminophen. The inset MS spectrum is the mean. (b) Correlation (cutoff 0.7) between aromatic NMR signals  
 313 for acetaminophen and ibuprofen metabolites and various nominal m/z values. Those for acetaminophen can be  
 314 immediately identified as belonging to the unmodified drug (m/z 152), its cysteinyl conjugate (m/z 271), and its  
 315 acetylcysteinyl conjugate (m/z 313, neutral molecule shown as inset). Those for ibuprofen require investigation  
 316 of the MSE spectra. The inset NMR spectra are for those samples having the highest respective NMR intensities,  
 317 and the inset mass spectrum is the sum of the corresponding mass spectra. Some trace of previous  
 318 acetaminophen usage is present in the NMR spectrum illustrating ibuprofen. (c) Correlation (cutoff 0.8) between  
 319 some aliphatic NMR signals for ibuprofen metabolites and m/z values in a higher range than given in Figure 1b.  
 320 The inset spectrum is for the sample having the highest ibuprofen intensity. Correlations due to the first  $^{13}\text{C}$   
 321 isotope can be distinguished for the strong signal at m/z 421. (d) Correlation (cutoff 0.8) of disopyramide-related  
 322 NMR signals with m/z values for disopyramide (m/z 340) and its known metabolite N-dealkyldisopyramide (m/z  
 323 298). The first  $^{13}\text{C}$  isotope correlations are visible. The inset NMR spectra are for a disopyramide standard (top)  
 324 and for the urine sample giving the highest relevant signal intensities (bottom). The inset mass spectrum is for  
 325 the same urine sample. It is clear that the urine NMR signals do not match the NMR signals for the disopyramide

326 standard (although in UPLC-MS an exact match to the standard was obtained). The NMR signals must therefore  
327 be due to an unidentified disopyramide metabolite, a covarying endogenous metabolite, or an additional  
328 (unknown) drug that was being taken by the single subject who was taking disopyramide. Figure reproduced  
329 from Crockford, D. J.; Maher, A. D.; Ahmadi, K. R.; Barrett, A.; Plumb, R. S.; Wilson, I. D.; Nicholson, J. K.  
330 *Anal. Chem.* 2008, 80 (18), 6835–6844 (ref <sup>71</sup>). Copyright 2008 American Chemical Society.

331 SHY also found applications in other fields than toxicology, as illustrated by Marti et al., who used  
332 this statistical tool to assess the authenticity and geographical metabolic differences of cold-pressed  
333 lemon oil.<sup>72</sup> In this study, the NMR/MS correlation complemented by the use of geranial and neral  
334 isomer standards, helped to determine the presence of both isomers but showed that the geranial one  
335 was present in the samples at a higher concentration.<sup>72</sup> SHY was also used in the natural product field  
336 to identify bioactive compounds while avoiding a time-consuming isolation process.<sup>73</sup> To do so,  
337 microfractionation of samples prior to LC-MS (positive and negative ionization mode) and NMR  
338 analyses were performed, as well as bioactivity assays. The selection of active NMR compounds, often  
339 made difficult by the important overlap with inactive compounds, was facilitated by the correlation  
340 with LC-MS signals, making SHY a powerful tool for the deconvolution of natural products.<sup>73</sup>  
341 Correlation analyses between NMR and MS data sets can also be used to confirm the annotations of  
342 the discriminative features previously identified by PCA applied to one of the methods, as it was done  
343 to study biological processes underlying the urine samples of patients with inborn errors of  
344 metabolism.<sup>74</sup> In this particular example, unsupervised PCA analysis was first performed on a NMR  
345 data set and the features that significantly impacted the distribution of the samples were annotated.  
346 Second, NMR spectra and MS spectra, from a DESI-MS data set, were bucketed into the same number  
347 of bins (594) to obtain a square matrix that was subjected to a Pearson correlation.<sup>74</sup> With this  
348 approach, a common list of discriminant features can be annotated with more confidence when a MS  
349 feature is positively correlated to an NMR annotation. Furthermore, it can also help to gain a better  
350 understanding of the biochemical reactions lying behind a specific condition, as a negative correlation  
351 can be explained by the consumption of the precursor compound and the production of another  
352 compound.<sup>74</sup> However, since the metabolic coverage of two analytical methods is different, it makes  
353 sense to assume that a list of common features between the NMR and the DESI-MS data sets will  
354 capture only a restricted portion of the urinary metabolome and that important biomarkers could be  
355 missed. Still, statistical analysis based on correlation coefficients can be easily implemented while  
356 allowing a straightforward interpretation of the results.

357 Although SHY seems to be the most widely correlation tool used in metabolomics to correlate  
358 NMR and MS data sets, another statistical correlation tool based on networks was also reported. In a  
359 first example, correlation networks were used to visualize the relationships between melon fruit  
360 analytes that were identified and for some quantified by various analytical techniques (namely <sup>1</sup>H NMR

361 spectroscopy and GC-MS analyses of polar compounds, HPLC analysis of lipophilic isoprenoids,  
362 untargeted LC-MS analysis of semipolar compounds, untargeted GC-MS analysis of volatile compounds  
363 and elemental profiling for mineral elements).<sup>75</sup> This extensive metabolic profiling allowed the  
364 identification or annotation of about 1932 features and 15 mineral elements. Within these features,  
365 only a small proportion were detected by several analytical methods, which proves the essential need  
366 of using several complementary analytical techniques to increase metabolic coverage. Following  
367 feature selection based on a two-way ANOVA, Spearman correlation coefficients were calculated and  
368 a cutoff ( $> 0.90$ ) was applied, which set the number of features or mineral elements in the correlation  
369 network at 715.<sup>75</sup> A network cartography was then created based on the Fruchterman-Reingold  
370 algorithm, where analytes were represented by the nodes and the Spearman correlation coefficients  
371 by the distance of the link between the nodes. This correlation network analysis made it possible to  
372 identify clusters of metabolites which were coregulated, to establish global changes in metabolic  
373 composition, and to highlight the association between primary and secondary metabolites with  
374 minerals or volatile compounds.<sup>75</sup> This approach was not developed to help with structural elucidation  
375 but rather to obtain a broader picture of the biological process and better understand metabolic  
376 interactions. It also applies to other kind of biological interactions; indeed, correlation networks were  
377 used to explore gene-metabolite association in tomato fruit, for instance.<sup>76</sup>

## 378 **Multiblock Fusion**

379 The use of multiblock data integration, or data fusion, has been increasing in omics sciences for a  
380 couple of decades and this approach can be applied at different levels.<sup>66</sup> Low-level data fusion consists  
381 of combining the preprocessed individual blocks at the data level without performing any variable  
382 selection prior to modeling the resulting block and in interpreting the global outcome. Although this is  
383 a straightforward way to approach data integration, careful consideration must be taken regarding  
384 scaling and normalizing the individual data sets. Indeed, because of the sensitivity and robustness  
385 differences existing between analytical methods (e.g., NMR or HRMS), the analytical response for a  
386 single compound will vary from one data set to another. Data set manipulation to overcome this  
387 obstacle can give too much weight to similar variables (e.g., isotopes and fragments from the same  
388 metabolite or metabolites from the same pathway), as it is the case when the individual data sets are  
389 autoscaled. Scaling can thus also be performed by considering sub-blocks of similar variables, which  
390 weight the influence of sub-blocks according to their size. Either way, low-level data fusion has been  
391 described as an approach which provides only limited useful information in metabolomics.<sup>66</sup>  
392 Furthermore, because the totality of each individual data set is being integrated as is, the size of the  
393 resulting data matrices before modeling is considerable. As such, mid-level data fusion considers only  
394 the most discriminant features highlighted by block-wise statistical analyses of the individual blocks,

395 which can be complemented by applying an additional technique to further reduce the dimensionality  
396 of the integrated matrix.<sup>66</sup> Also, assessing separately each data set through mid-level data fusion  
397 highlighted metabolites detected by several methods, which could be over-represented in the global  
398 model and which may introduce bias in biomarker discovery. Finally, data fusion can also be applied  
399 through a high-level approach, where individual blocks are preprocessed and modeled separately, as  
400 for a mid-level approach, but where the global output (e.g., the predictive algorithms) of each of the  
401 individual model is integrated, rather than their discriminant features.<sup>66</sup> It is important to mention that  
402 statistical model validation is crucial in metabolomics in order to properly assess the performance of  
403 the model without overfitting it,<sup>77,78</sup> but although several validation tools exist, no common agreement  
404 has been found regarding which one is the most suitable.<sup>79-81</sup> The predictive power of the combined  
405 output obtained following a high-level approach is highly expected to be equal or higher than the  
406 predictive power of the best performing individual model, and thus the error rate is also expected to  
407 be reduced.<sup>82</sup> However, the predictive performance of the global output will increase more  
408 importantly if the classifiers used present similar discriminative performance, which is often the case  
409 in metabolomics.<sup>82</sup> In such cases, integration of the individual output through correlation networks, as  
410 presented in the Correlation of NMR and MS Data Sets section, might be a solution to jointly interpret  
411 individual results while preserving the predictive performance of the individual analysis.<sup>66</sup> In a way,  
412 correlation networks can be considered as a high-level data fusion approach but from which the  
413 biological interpretation can often be complicated by an extensive visualization output.

414 Following the selection of the data integration approach, several modeling techniques are  
415 available and can be applied for individual data set analysis (for mid- or high-level approaches, before  
416 data fusion of the selected features or the individual global outputs, respectively) or by global analysis  
417 (for low- or mid-level approaches once data fusion has been made).<sup>66</sup> The resulting models naturally  
418 highlight possible association between variables from different data sets to improve biological  
419 interpretation but also serve to assess the contribution of each individual data set to the global model.  
420 Sequential multiblock analysis, as unsupervised single block multivariate statistical analysis, aims to  
421 describe the general trend of the matrix and is based on the calculation of one single component at a  
422 time followed by a deflation procedure to calculate the next one. Several sequential multiblock  
423 methods exist, from the simplest which are SUM-PCA or consensus PCA (CPCA) to more complex such  
424 as hierarchical PCA (HPCA), generalized PCA (GPCA), multivariate component models or multiple factor  
425 analysis.<sup>66,83</sup> Similarly, several modeling tools exist to apply predictive supervised analysis, such as PLS  
426 regression or discriminant analysis, orthogonal-PLS (O-PLS or O2-PLS), hierarchical PLS (HPLS) or  
427 multiblock PLS (MBPLS).<sup>66,84</sup> Those statistical methods allow the combination of data sets from  
428 different analytical methods, whatever the size of the different blocks. In metabolomics, however,



429 even though the numbers of variables can change from one NMR to one MS block for instance, it is  
430 preferred to have the same number of objects (or samples) for all blocks. The following paragraphs  
431 illustrate how such data fusion and data modeling approaches have maximized the potential of  
432 combining NMR and MS data in metabolomics.

433 In a first example, HPCA was applied to the study of three melon cultivars by  $^1\text{H}$  NMR spectroscopy  
434 and by GC-MS, but not only to fuse both data sets.<sup>85</sup> Indeed, in this example, classical PCA on  $^1\text{H}$  NMR  
435 managed to discriminate the samples coming from different spatial positions in the melon fruit, but it  
436 failed when the PCA was based on the GC-MS data set. As such, the authors assigned each of the three  
437 melon cultivars as an individual block and applied HPCA on them, which successfully highlighted  
438 metabolic differences linked to the spatial positions of the samples.<sup>85</sup> Furthermore, a HPCA model was  
439 also built on the combination of both analytical techniques and showed the robustness of this  
440 statistical tool as the compounds detected by both NMR and GC-MS were located at similar positions  
441 of the loading plots.<sup>85</sup> In another plant study exploring tomato fruits and leaves,<sup>86</sup> the integration of  $^1\text{H}$   
442 NMR, LC-MS, and GC-MS data was directly done through a low-data level fusion approach, regardless  
443 of their individual performance but further association was made with correlation networks to  
444 facilitate the interpretation of the biological pathway regulations.

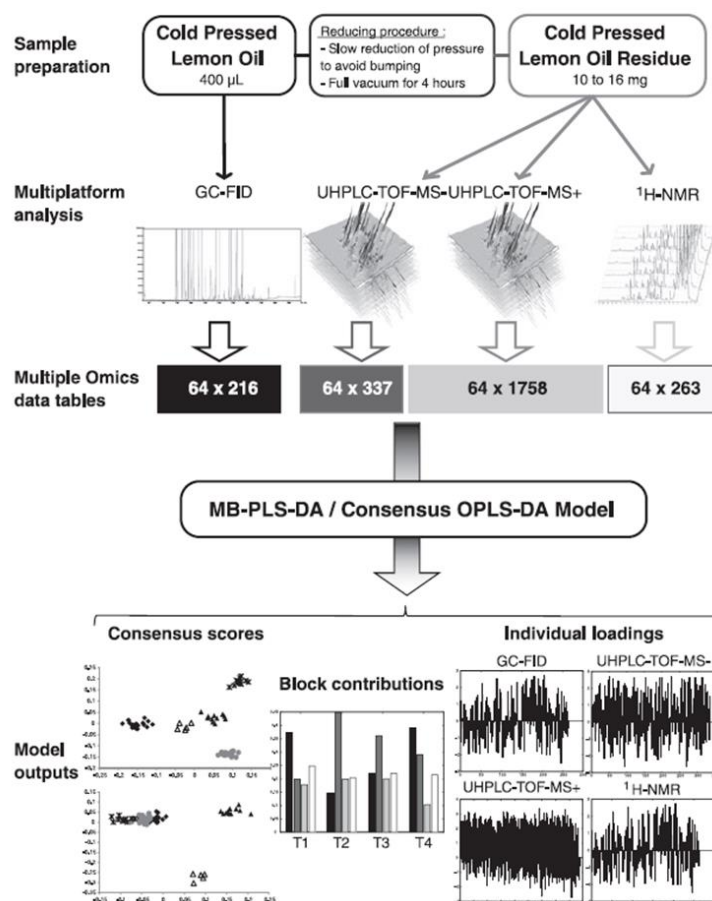
445 In a second example,  $^1\text{H}$  NMR spectroscopy and two HRMS instruments (TOF and Orbitrap) were  
446 used to analyze honey samples from different botanical origins.<sup>87</sup> PCA and PLS-DA were applied on  
447 each of the individual data sets, before PCA modeling based on mid-level data fusion was performed  
448 from two different angles. The first one was based on the fusion of the PCA scores of each of the data  
449 sets, to prevent any loss of information, and the second one was based on the fusion of selected  
450 variables from the individual PLS-DA models, to remove any irrelevant information. Both data fusion  
451 approaches performed better than the individual models in term of discriminative power and sample  
452 misclassification.<sup>87</sup> The mid-level data fusion between NMR and HRMS-Orbitrap with variable  
453 selections had the best discrimination of all the models reported, without misclassification, while the  
454 mid-level data fusion between NMR and HRMS-Orbitrap without variable selections led to  
455 misclassification. The fusion of NMR with HRMS-TOF data with or without variable selection did not  
456 misclassify the samples but underperformed the discrimination of the samples with variable selection  
457 compared to the fusion of the NMR and HRMS-Orbitrap data sets.<sup>87</sup>

458 Another study focused on the metabolic profiles of plasma samples from patients with stable  
459 carotid atherosclerosis versus healthy subjects using GC-MS and  $^1\text{H}$  NMR spectroscopy.<sup>88</sup> The individual  
460 PCA models showed good separation of the samples but only along the third principal component,  
461 which proved the presence of discriminative variables irrelevant to the pathology characterization. As

462 in single block multivariate analyses, PLS, or orthogonal signal correction (OSC), also called OPLS,  
463 provided a better sample separation since the sample classification is included in the model.  
464 Supervised analyses were thus performed, and the PLS-DA and OPLS-DA models of both individual data  
465 sets did present a higher discriminative power compared to PCA models. Low-level data fusion was  
466 then performed, and the supervised analysis of the resulting data set performed as well as the  
467 individual supervised models.<sup>88</sup> Here, the application of the combined NMR and MS supervised analysis  
468 could thus be questioned, but a Pearson correlation between the metabolites scores obtained from  
469 the predictive component of the combined OPLS-DA model allowed a broader understanding of the  
470 metabolic pathway alteration than if only one technique would have been used for statistical analyses.

471 Even if supervised methods often enable a better separation than unsupervised analyses, they can  
472 sometimes fail, and combining several analytical technologies might enhance the discriminative  
473 performance of a model and help to highlight specific biomarkers. For instance, Gu et al. used <sup>1</sup>H NMR  
474 spectroscopy and direct analysis in real time (DART)-MS to discriminate serum samples from patients  
475 with breast cancer from healthy controls.<sup>89</sup> No distinct separation between the samples were observed  
476 in the PCA models of each of the analytical data sets, although a slight grouping along the first principal  
477 component was observed for the <sup>1</sup>H NMR PCA model. Furthermore, both PLS-DA and OSC-PLS-DA  
478 models based on each of the individual data sets did misclassify an important number of samples,  
479 which was clearly beyond the acceptance rate when it comes to health applications.<sup>89</sup> As such, PLS-DA  
480 and OSC-PLS-DA models were rebuilt by setting the Y dummy matrix, the classification variable, to the  
481 first principal component of the <sup>1</sup>H NMR PCA model, which performed slightly better than the DART-  
482 MS PCA model, and the X matrix to the DART-MS data set, which was more sensitive. These models  
483 both performed better than the individual supervised models, with a major preference for the OSC-  
484 PLS-DA which resulted in a lowest misclassification score due to the removal of confounding factors  
485 following the orthogonal signal correction.<sup>89</sup> Another study, which aimed to differentiate the  
486 extraction protocols of cold-pressed lemon oil, supervised modeling analyses, namely, MB-PLS-DA and  
487 consensus (C)-OPLS-DA, elegantly showed the benefits of using orthogonal projection to improve the  
488 separation between samples.<sup>90</sup> In this illustration, a low-data level fusion of untargeted data sets  
489 obtained by <sup>1</sup>H NMR, GC-FID and LC-MS in positive and negative ionization modes was used (Figure  
490 5).<sup>90</sup> The supervised analyses then showed a much better separation, and thus interpretability, when  
491 it came to the C-OPLS-DA compared to the MB-PLS-DA, although their predictive performance was  
492 similar. It is important to note that even if the data matrix resulting from the fusion of these four data  
493 sets was extensive, high-level data fusion successfully discriminated the same samples according to  
494 their geographical origins in another study,<sup>91</sup> but not according to their extraction processes.  
495 Therefore, low-level data fusion can be a useful alternative when other modelling tools failed, at the

496 condition to take particular care for the scaling of the different datasets. As illustrated in another study  
 497 applying low-level data fusion, both of the  $^1\text{H}$  NMR and direct infusion (DI)-ESI-MS data sets were  
 498 scaled first to unit variance and second by the square root of the block variable count, in order to  
 499 ensure fairness in the consideration of each block.<sup>92</sup> Other key elements were to note in this report, as  
 500 a thorough optimization of the sample preparation in order to prevent important sample handling.  
 501 This allowed the analysis of one single sample by both  $^1\text{H}$  NMR and DI-ESI-MS, and optimization of the  
 502 DI-ESI-MS protocol was undertaken to limit the matrix effect.<sup>92</sup> Also, the backscaled  $^1\text{H}$  NMR and DI-  
 503 ESI-MS loading plots obtained from the MB-PLSDA, which outperformed the single block PLS, were  
 504 complemented by additional MS accurate mass and MS/MS experiments to compare with the NMR  
 505 signals and facilitate metabolite identification.<sup>92</sup> Overall, this study provides a complete illustration of  
 506 how to combine NMR and MS data sets from sample preparation, data set acquisition, multivariate  
 507 analyses and metabolite identification.



508  
 509 Figure 5. Methodology used to integrate metabolomic data from multiple analytical platforms for a  
 510 comprehensive characterization of lemon essential oils. Figure reproduced from Integrating  
 511 Metabolomic Data from Multiple Analytical Platforms for a Comprehensive Characterization of Lemon  
 512 Essential Oils, Mehl, F.; Marti, G.; Merle, P.; Delort, E.; Baroux, L.; Sommer, H.; Wolfender, J.-L.; Rudaz,  
 513 S.; Boccard, J. Flavor Fragr., Vol. 30, Issue 2 (ref<sup>90</sup>). Copyright 2015 Wiley.

515           Multiblock data fusion is thus starting to be anchored in the metabolomics landscape, and some  
516 attempts are made to refine and bring some originality into those methods and to increase their  
517 application scope. For instance, C-PLS-DA and C-O-PLS-DA have been applied to integrate MS with two  
518 NMR data sets, one  $^1\text{H}$  data set and one 2D J-resolved (J-res) NMR.<sup>93</sup> Also, data fusion can be based on  
519 multiple kernel learning (MKL). This approach was applied to explore plasma metabolic alterations in  
520 three different chronic diseases for example, namely, acute coronary syndrome, breast and colon  
521 cancers.<sup>94</sup> Serum samples were analyzed by NMR and LC-MS, and the resulting data sets were fused  
522 with the metadata of the patients, which aimed to represent their lifestyle. The MKL fusion model  
523 increased the performance of the individual models when it came to the acute coronary syndrome,  
524 but slightly underperformed compared to the individual NMR model for the breast cancer condition,  
525 and none of the individual or fusion model performed well regarding the colon cancer conditions.<sup>94</sup>  
526 This could be due to the increase presence of confounding variables in the matrix, which complicated  
527 the selection of discriminative variables. Several methods were actually developed to optimize variable  
528 selection. Deng and co-workers proposed one based on backward variable elimination from PLS-DA  
529 models and combined with Monte Carlo cross validation (MCCV-BVE-PLSDA).<sup>95</sup> This method could be  
530 considered as half way between low-level and mid-level data-fusion, as all the variables were  
531 considered during the first iteration and since only the most predictive variables were kept during the  
532 last iteration. A similar approach had been already proposed, called a Sparse multiblock PLS regression  
533 (Sparse MBPLSR), implemented with a cross model validation in order to ensure the reliable and stable  
534 variable selection for biomarker discovery.<sup>96</sup> In a more recent report, the variable selection before  
535 obtaining the final PLS-DA model was made in three steps. First, a curation step was applied to remove  
536 all the redundant variables. Second, PCA and PLS-DA were performed and the resulting variables were  
537 further filtered by performing different kind of statistical analyses, namely one-way ANOVA Sparse  
538 PLS, or least absolute shrinkage and selection operator (LASSO). Third, PLS-DA was performed on each  
539 of the resulting new subset of selected variables.<sup>97</sup> All the PLS-DA following variable selection  
540 performed better than the PLS-DA before variable selection and the models which had the best  
541 performance were the one based on the variables selected by the Sparse PLS and the LASSO  
542 techniques.<sup>97</sup> Further matrix reduction can be performed but it depends on operator willingness to  
543 discard information without affecting biomarker discovery or biological pathway understanding. In any  
544 case, although a 2014 review mentioned that the multiblock fusion of NMR and MS techniques  
545 followed by supervised analyses was not common in metabolomics studies,<sup>66</sup> we hope to have shown  
546 that since it has gained great interest as it maximizes the complementarity between both analytical  
547 techniques. Although multiblock analyses do not always end up giving better results than single block  
548 analysis and that it is important to keep a critical eye on the usefulness of this hyphenation, it has  
549 already found various application fields.

## 550 **Postmetabolomics Analysis: Reaching the Aim and Going Beyond**

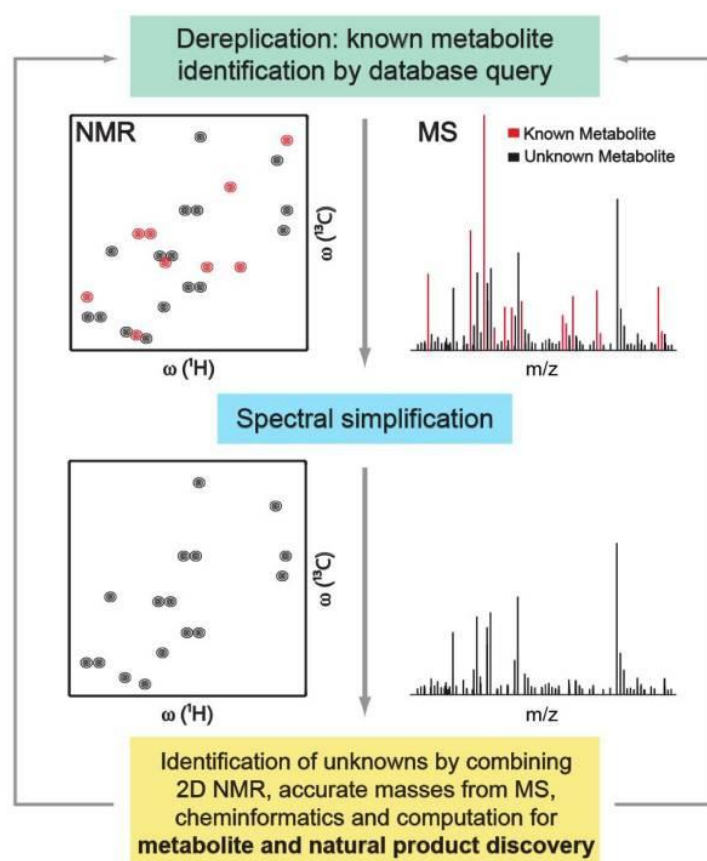
### 551 **Identification**

552 While increasing the metabolic coverage is of the utmost importance in metabolomics, it might be  
553 even more crucial to identify the metabolites covered. Indeed, metabolite identification is clearly seen  
554 as a major bottleneck in the field of metabolomics, especially in MS-based techniques. Thanks to the  
555 robustness of <sup>1</sup>H NMR spectroscopic analysis, NMR databases are easier to produce and more  
556 trustworthy than MS databases. Generally, signals are compared and matched to experimental or  
557 theoretical spectra registered in in-house or online databases. If a compound of interest cannot be  
558 identified, it is an unknown compound annotated at a level 4 of confidence, the lowest according to  
559 the criteria used by the Metabolomics Standards Initiative (MSI).<sup>98,99</sup> If the chemical class of the  
560 compound can be determined, the annotation level is 3. If a compound can be determined by  
561 comparison to a database, it is a level 2 of annotation. Finally, if the compound matches with at least  
562 two orthogonal parameters (e.g., the *m/z* and the retention time) compared to an authentic standard  
563 which has been spiked into a sample, then the level of confidence is 1 and the term “identification”  
564 can be used (although care still needs to be taken in the case of isomeric compounds, but  
565 stereochemistry should soon be taken into consideration through a new and revised reporting  
566 standards which are being discussed in the community with the Metabolite Identification Task group  
567 of the Metabolomics Society leading on this initiative [Personal communication, Prof. Warwick Dunn,  
568 cochair Metabolite Identification task group]), in contrast to the term “annotation” for levels 2, 3, and  
569 4.<sup>99</sup> Also sometimes, subconfidence groups in the level 2 annotation emerge, 2b being given if the  
570 annotation to a specific metabolite is based on one orthogonal parameter (e.g., *m/z* or  $\delta$  values  
571 matching to database) and 2a if the annotation is based on two orthogonal parameters (e.g., *m/z*  
572 values and retention time or *m/z* and  $\delta$  values matching to databases) without spiking the  
573 corresponding authentic standard. Hence, annotating a compound by using both NMR and MS  
574 analytical techniques provides more confidence in the annotation level and facilitates structural  
575 elucidation of unknowns.

576 Identification of unknowns is particularly difficult in plant metabolomics due to the lack of chemical  
577 standards to confirm the identification of a new metabolite. Initial strategies developed were thus  
578 based on accurate mass measurements by HRMS, allowing to obtain chemical formula and matching  
579 them to possible chemical structures available in databases, before comparing their fragmentation  
580 pattern.<sup>100</sup> Successful candidates were confirmed following purification and NMR analysis for  
581 structural characterization. This method could be limited by the fact that the lack of sensitivity of NMR  
582 measurements could restrict the annotation of new metabolites or biomarkers. The use of capillary

583 NMR to overcome this limitation was seen as an alternative,<sup>101</sup> but the process was not less time-  
584 consuming and labor-intensive. More recent strategies have thus been proposed to increase the  
585 identification of unknowns by relying on the complementarity between NMR and MS. The one that  
586 attracted the most attention in metabolomics is the so-called SUMMIT, for Structures of Unknown  
587 Metabolomics Mixture components by MS/NMR.<sup>102</sup> Its principle relies on HRMS measurements of the  
588 accurate masses of the different analytes present within a complex sample in order to determine their  
589 molecular formulas. From those, all the possible scaffolds which can correspond to these formulas are  
590 predicted, and the list can be extensive. A 1D or 2D NMR spectra is then predicted for each of the  
591 predicted scaffolds. These predicted NMR spectra are then compared to experimental HSQC NMR  
592 spectra, previously deconvoluted into <sup>13</sup>C-<sup>1</sup>H HSQC chemical shifts of each metabolite by combining  
593 information from 2D NMR experiments. The possible scaffold hits are finally ranked according the level  
594 of concordance between the predicted and the experimental spectra.<sup>102</sup> This method was initially  
595 developed to prevent time-demanding steps such as sample separation/purification or interrogating  
596 metabolic databases and was successfully applied to the identification of previously known compound  
597 from *E. coli* as a proof of principle.<sup>102</sup> However, subsequent publications showed that the approach  
598 could be further improved by the interrogation of databases, making the method more general and  
599 efficient.<sup>103,104</sup> In the future, using the SUMMIT strategy could also help guiding retrospectively to a  
600 specific MS platform *ad hoc* or encourage the use of multiple MS platform to increase the chance to  
601 detect metabolites by both MS and NMR. Recently after the introduction of SUMMIT MS/NMR,  
602 another strategy called NMR/MS Translator was proposed by the same group, as a tool which could  
603 be used prior to SUMMIT MS/NMR.<sup>105</sup> Following the 1D or 2D NMR spectral acquisition, NMR/MS  
604 Translator questions NMR databases and from the obtained hits, it calculates isotopes, adducts, and  
605 fragments. From those, MS spectra are reconstructed and compared to the upstream acquired MS<sup>1</sup>  
606 experimental spectra.<sup>105</sup> Interrogating databases is limited by the fact that the hits will depends on  
607 how well a database is furnished, and yet there have always been a lot of disparities from one database  
608 to another.<sup>58</sup> Therefore, authors highlighted that the eye confirmation of the resulting annotations by  
609 an operator should always be applied to prevent false identification, as reported with NMR/MS  
610 Translator for 11 urinary metabolites. However once known metabolites are identified, unknown  
611 metabolites are easily distinguished and SUMMIT MS/NMR can come to help (Figure 6),<sup>103</sup> for which  
612 automation efforts have been pursued.<sup>104</sup> While these approaches are very elegant, they imply that  
613 metabolites need to be detected by both NMR and MS measurements. However, as already explained  
614 in Cross-Comparison of NMR and MS Data Sets to Increase Metabolic Coverage the metabolic coverage  
615 between NMR and MS measurements is limited by different parameters such as the low sensitivity of  
616 NMR or for MS techniques the ionization efficiency, the choice of the chromatographic phase, and the  
617 ionization mode. Therefore, methods such as SUMMIT or NMR/MS translator are restricted to

618 compounds that can be detected by both the NMR and MS techniques. As such, physical or chemical  
619 derivatization<sup>103</sup> prior to analysis might give access to an increased number of unknowns that would  
620 not have been commonly detected by NMR and MS otherwise. Although the associated sample  
621 handling could be more time-demanding, it would lead to promising perspectives to uncover more  
622 unknown metabolites.



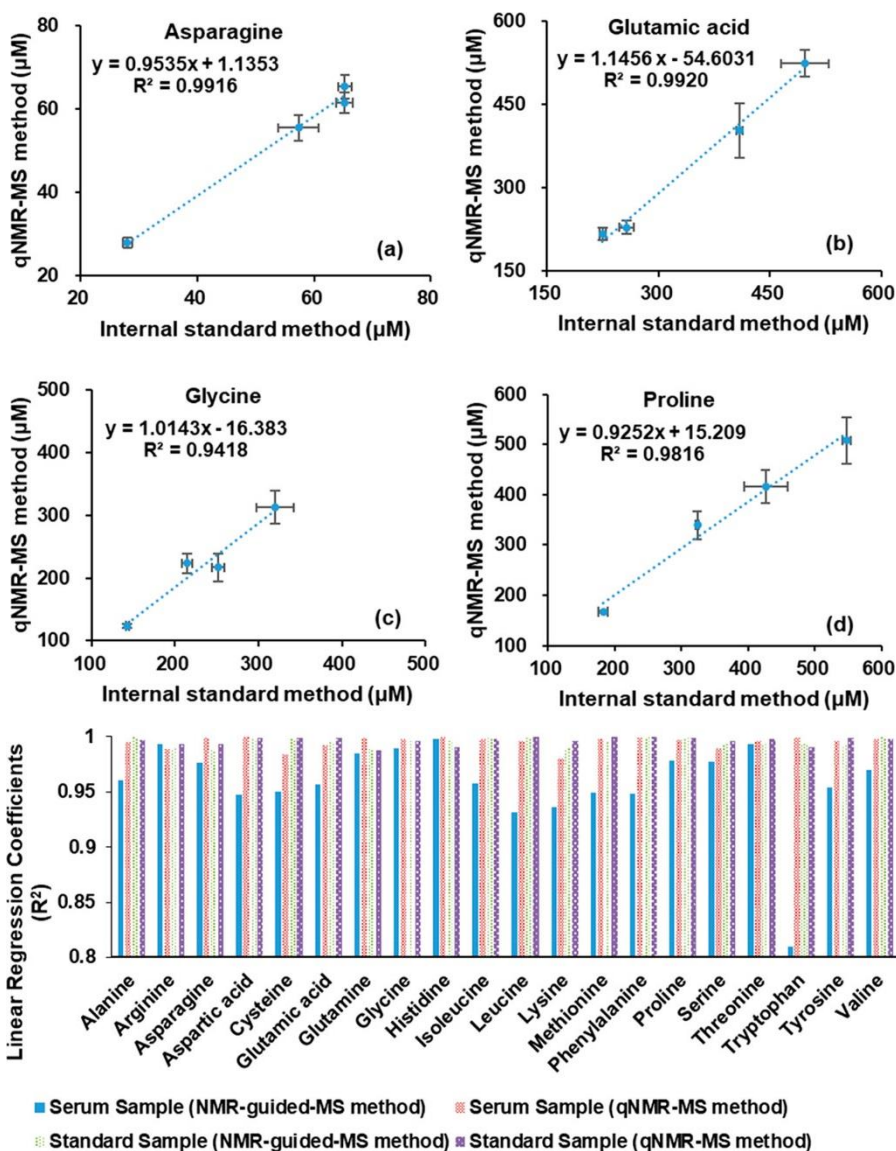
623  
624 Figure 6. Integrated metabolomics workflow for the identification of known and unknown metabolites in  
625 complex mixtures. Combined use of metabolomics databases with experimental NMR and MS spectra (e.g.,  
626 NMR/MS Translator<sup>105</sup>) allows the rapid identification of a maximal number of known metabolites present in the  
627 mixture, while unidentified signals are used as fingerprints of unknowns. Next, structures of unknown  
628 metabolites can be elucidated or vastly narrowed down by the combined use of multidimensional NMR, MS,  
629 cheminformatics, and computation (e.g., SUMMIT MS/NMR<sup>102</sup>). Figure reprinted from Curr. Opin. Biotechnol.,  
630 Vol. 43, Bingol, K.; Brüsweiler, R. Knowns and Unknowns in Metabolomics Identified by  
631 Multidimensional NMR and Hybrid MS/NMR Methods, p 17-24 (ref <sup>103</sup>). Copyright 2017, with permission  
632 from Elsevier.

### 633 Quantitation

634 Quantitative information about metabolites constitutes an invaluable asset to obtain a deeper  
635 understanding of the biological reactions and processes occurring in an organism, as changes in  
636 metabolite concentrations reflect changes in protein concentrations or gene expressions. For this  
637 reason, efforts are being made to collect quantitative data and incorporate them in databases, as it

638 was done for the human serum metabolome.<sup>61</sup> Quantitative data also helps toward more robust,  
639 discriminant, and predictive statistical models.<sup>106</sup> NMR- and MS-based quantitation methods are well  
640 established.<sup>106,107</sup> NMR quantitation is relatively straightforward and lies on the use of a single internal  
641 standard such as TSP (sodium d<sub>4</sub>-3-(trimethylsilyl)-propionate), provided that data are acquired and  
642 processed under proper conditions known as qNMR.<sup>108</sup> For complex mixtures where peak overlaps  
643 prevent accurate quantitation, quantitative procedures based on 2D NMR have also been  
644 developed.<sup>109</sup> MS quantitation is more time-demanding as it requires further development to target  
645 the metabolites of interest and free from ion suppression phenomenon. Most robust targeted assays  
646 lie on either the combined use of stable isotope labeled (SIL) internal standards and calibration curves,  
647 especially in regulatory science,<sup>110</sup> or on differential <sup>12</sup>C/<sup>13</sup>C isotope labeling using for instance isotope  
648 reagent such as <sup>13</sup>C-dansyl chloride to label metabolites.<sup>111</sup> However, SIL internal standards are not  
649 always available and costly. Chemical analogues of the metabolites of interest can be used as internal  
650 standards but this is a less reliable approach and results must be interpreted with care. The idea of  
651 combining both NMR and MS techniques to obtain more robust quantitative assays and overcome the  
652 difficulties encountered through MS-based assays is relatively recent.<sup>106</sup> A first approach has been  
653 described as an “NMR-guided-MS quantitation”.<sup>112</sup> In this approach, NMR absolute quantitation of the  
654 metabolites of interest was performed for a serum sample which was randomly selected to be the  
655 reference sample. The concentrations obtained were then set as references for the multiple reaction  
656 monitoring transitions of the metabolites in the MS experiment, during which the rest of the samples  
657 were analyzed.<sup>112</sup> Without the use of internal standards, 30 serum metabolites were successfully  
658 quantified and the correlation between the concentration obtained by the NMR-guided-MS approach  
659 with the one obtained by NMR were above 0.92 for most metabolites. Metabolite concentrations  
660 which showed poor correlation between the NMR-guided-MS approach and NMR draw special  
661 attention to the fact that ion formation is not always stable during MS measurements (which might be  
662 due to ion suppression or ionization efficiency due to e.g., source clogging) and that matrix effect can  
663 skew the results.<sup>112</sup> To alleviate this problem, the same group proposed to combine NMR and MS  
664 techniques with chemical derivatization through a so-called qNMR-MS method.<sup>113</sup> As for the NMR-  
665 guided-MS method, NMR was first used to obtain metabolite concentrations of a randomly assigned  
666 reference sample. This sample was then derivatized with SIL internal standards and mixed with the  
667 remaining study samples, which had been derivatized with unlabeled internal standards. The  
668 comparison between the labeled and the unlabeled signals allowed absolute quantitation by MS,  
669 making it possible to account for matrix effects.<sup>113</sup> Results showed excellent agreement between a  
670 classical internal standards methods and proved that the matrix effect, which can be important in  
671 complex biological samples such as serum, is well corrected in the qNMR-MS method compared to the  
672 NMR-guided-MS one (Figure 7).<sup>113</sup>





673

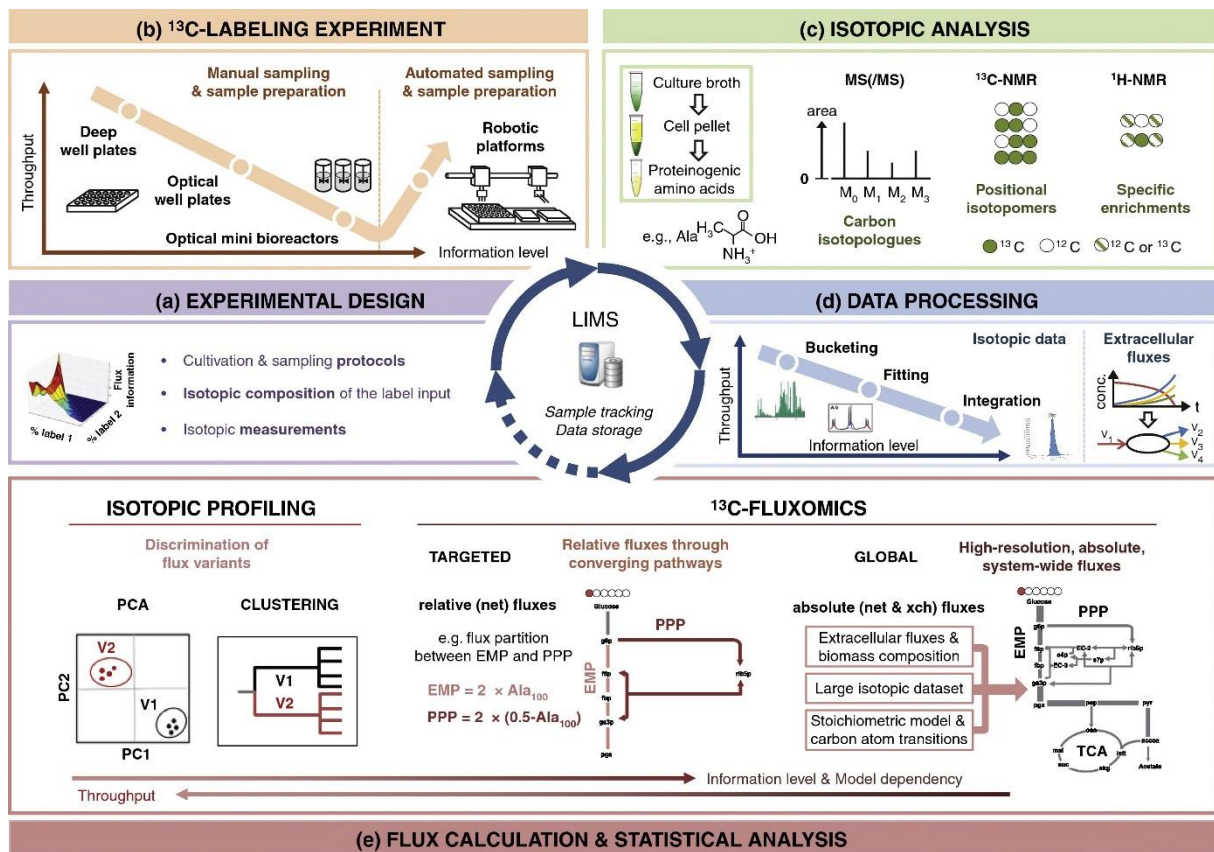
674

675 Figure 7. Linear regression equations for four amino acids in human serum samples: (a) asparagine, (b) glutamic  
 676 Acid, (c) glycine, and (d) proline. Coefficients of determination ( $R^2$ ) values show excellent agreement between  
 677 qNMR-MS and internal standard methods. Coefficients of determination ( $R^2$ ) for quantitating amino acids in  
 678 human serum and standard samples. Each sample was tested using two quantitation calculation methods: NMR-  
 679 guided-MS and qNMR-MS. NMR-guided-MS uses the peak area of each amino acid derivative to quantify, while  
 680 qNMR-MS uses the peak area ratio between labeled and unlabeled MS peaks of each amino acid derivative to  
 681 quantify metabolites. Figure reproduced from Fei, Q.; Wang, D.; Jasbi, P.; Zhang, P.; Nagana Gowda, G. A.;  
 682 Raftery, D.; Gu, H. Combining NMR and MS with Chemical Derivatization for Absolute Quantification  
 683 with Reduced Matrix Effects. *Anal. Chem.* 2019, 91 (6), 4055–4062 (ref <sup>113</sup>). Copyright 2019 American  
 684 Chemical Society.

### 685 Stable-Isotope Resolved Metabolomics (SIRM)

686 Stable-isotope resolved metabolomics (SIRM) is an atom-based approach which aims to measure,  
 687 following a stimulus, metabolic pathways alterations, and fluxes, which explains that it is also called  
 688 fluxomics. This field is based on stable isotope tracers which are used to label precursor molecules,  
 689 making it possible to measure metabolic reactions and quantify the metabolite byproducts based on

690 the measurement of the tracer atoms. The combined use of NMR and MS analytical techniques for  
691 SIRM is an evidence, as NMR spectroscopy can distinguish isotopomers, compounds which have the  
692 same number of heavy isotopes but at different positions, while MS can quantify the isotopologues,  
693 isomers that differ by their isotope composition. The most commonly used stable isotope-labeled  
694 precursor in fluxomics is U-<sup>13</sup>C-glucose, but several others are available and can be selected according  
695 the metabolic pathway of interest.<sup>114</sup> Some studies actually used multiple labeled precursors in order  
696 to cover more metabolic pathways.<sup>115</sup> SIRM is thus a very promising tool for clinical applications such  
697 as understating metabolic disturbances in diseases, pharmaceutical<sup>116</sup> and toxicological research, and  
698 highlighting novel drug targets, especially because experiments can be performed *in vitro* through cell  
699 cultures, on *ex vivo* tissue models, or *in vivo* human or animal models.<sup>117</sup> It helped for instance to  
700 understand the mechanism of action of a lithium therapeutic treatment for bipolar disorders.<sup>115</sup> It is  
701 also of foremost interest for cancer research,<sup>118,119</sup> such as this *in vivo* application for lung cancer  
702 patients infused with <sup>13</sup>C-glucose which highlighted an upregulation of the glycolysis and Krebs cycle  
703 activity, as well as the unexpected pyruvate carboxylation, compared to the noncancerous tissues  
704 surrounding the tumor.<sup>120</sup> The most used MS technologies in SIRM are GC-MS and FT-ICR-MS, the first  
705 one being more affordable but the second being more sensitive, having higher resolution, and allowing  
706 high-throughput workflows to be applied more easily.<sup>121</sup> In NMR, fast 2D or 3D methods are generally  
707 used to accurately quantify the populations of all isotopomers, given the complex spectra resulting  
708 from complex metabolite mixtures with multiple isotopic patterns.<sup>122–124</sup> The combination of 2D NMR  
709 with FT-ICR-MS highlighted for instance precursor metabolites for the synthesis of specific  
710 glycerophospholipids while obtaining accurate m/z which facilitate identification, within a relatively  
711 rapid workflow.<sup>125</sup> However, SIRM is a rapidly growing field and LC-MS has also been applied in  
712 combination with NMR to develop a workflow allowing the flux analysis of isoprenoids in yeast, which  
713 could be further applied to other organisms.<sup>126</sup> Upon further progresses made in terms of  
714 computational tools and high-throughput workflow (Figure 8)<sup>127</sup> for fluxomics analysis, the  
715 combination of NMR and MS within this field will for sure unravel the understanding of numerous  
716 metabolic pathways within a large scope of applications.



717

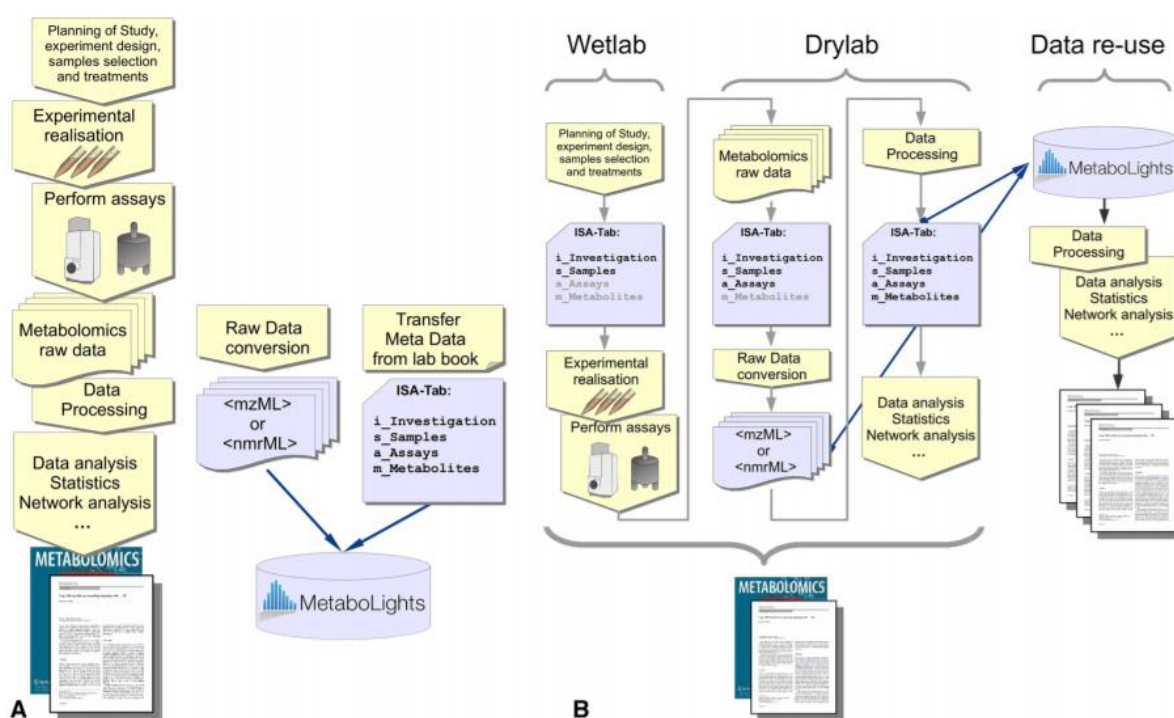
718 Figure 8. Overview of a high-throughput <sup>13</sup>C-fluxomics workflow. (a) Experimental design allows maximizing the  
 719 flux information depending on the biological question addressed by the study. Besides defining appropriate  
 720 cultivation conditions (medium composition, temperature, pH, among others), the label input must be optimized  
 721 according to the isotopic data that can be measured, the metabolic systems investigated, and the experimental  
 722 costs. (b) Robotic and manual systems are available to perform <sup>13</sup>C-labeling experiments and sample preparation.  
 723 A trade-off between throughput and fine control of growth is determined by the cell cultivation system used and  
 724 by the degree of workflow automation. (c) NMR-based and/or MS-based isotopic analyses are then carried out  
 725 to measure the <sup>13</sup>C-incorporation into metabolites. (d) Data processing is necessary to extract meaningful  
 726 isotopic information from the raw data. In contrast to isotopic profiling or targeted <sup>13</sup>C-fluxomics, global <sup>13</sup>C-  
 727 fluxomics also requires measuring extracellular (production and consumption) fluxes from the time-course  
 728 variations of extracellular concentrations. (e) Finally, different computational approaches for flux calculation and  
 729 statistical analysis are applied according to the purpose of the investigations (i.e., isotopic profiling, targeted  
 730 fluxomics or global fluxomics) and the level of biological knowledge required to extract the flux information.  
 731 Figure reprinted from Curr. Opin. Biotechnol., Vol. 43, Heux, S.; Bergès, C.; Millard, P.; Portais, J.-C.;  
 732 Létisse, F. Recent Advances in High-Throughput <sup>13</sup>C-Fluxomics, p 104-109 (ref <sup>127</sup>). Copyright 2017, with  
 733 permission from Elsevier.

## 734 Future Perspectives

### 735 Public Databases

736 The previous sections clearly showed that metabolite identification relies on comparing  
 737 experimental spectral data to those present in various databases. With Metabolomics being a rapidly  
 738 growing field, databases have been expanding as well. Common examples are the Human Metabolome

739 Database,<sup>128</sup> the human serum metabolome,<sup>61</sup> METLIN,<sup>129</sup> the Biological Magnetic Resonance Data  
 740 Bank,<sup>130</sup> LipidMaps,<sup>131</sup> or the Yeast Metabolomics Database.<sup>132</sup> However, as indicated by their names,  
 741 these databases are either instrument-, organism-, or compound class-specific. Furthermore, high-  
 742 quality spectra are sometimes missing, which limits annotation confidence, or experimental conditions  
 743 are not reported, which makes the spectral comparison more difficult and less trustworthy. Therefore,  
 744 such databases do not always make good use of the MS/NMR complementarity. To overcome this  
 745 issue, MetaboLights, the first metabolomics open-access repository regardless of the species or the  
 746 analytical technique used, was released in 2012.<sup>133,134</sup> It aimed to offer a repository where compound  
 747 structure, spectral reference, biological concentration, location, and role as well as raw data could be  
 748 submitted, stored, shared with the metabolomics community, and reused. However, automatic  
 749 reporting of metabolomics data is not always performed by the community, although it has improved  
 750 in the past couple of years.<sup>135</sup> Undoubtedly, this has been helped by initiatives such as COSMOS  
 751 (Coordination of Standards in MetabOmicS), which aimed to promote data reporting according to  
 752 specific standards, in order to facilitate data exchange and dissemination.<sup>136</sup> An experimental workflow  
 753 has been proposed to guide members of the metabolomics community in their standardization efforts  
 754 (Figure 9) and which ensure FAIR (Findable, Accessible, Interoperable and Reusable) data sharing.<sup>137</sup>  
 755 This is of paramount importance for the future of metabolomics and even if it goes without saying that  
 756 this should be applied for both NMR and MS experiments; even taken separately, it is obvious that it  
 757 will also help toward further combined use of NMR and MS analysis by gathering both spectral data  
 758 for a same compound.

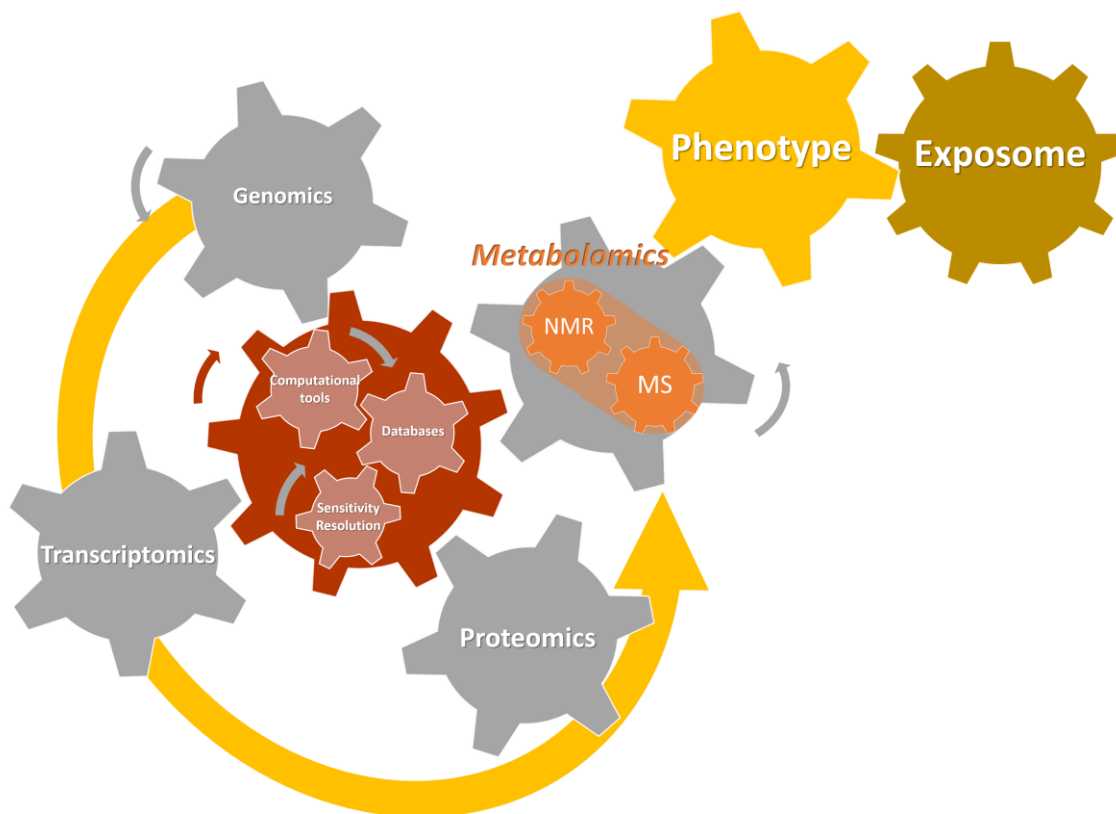


759

760 Figure 9. Experimental workflows in metabolomics. Shown in light blue are the relevant parts where data  
761 standards come into play. Annotated data deposition in open repositories allow for data reanalysis and reuse.  
762 (a) Traditional workflow using tools which do not depend on data standards, and where data annotation and  
763 data publication happen together with manuscript submission. (b) Fully standard embedded workflow, where  
764 data annotation is part of the standard operational procedures, data processing can use open software, and data  
765 publication is an integral part of the dissemination. Figure reproduced from ref. <sup>137</sup> under Creative Commons  
766 Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

## 767 **Integration to Other Platforms (Outside Metabolomics)**

768 The new bioinformatic tools allowing the combination of different analytical methods will certainly  
769 not stop their boundaries to the metabolomics field. A growing interest has been noticed for the  
770 integration of untargeted metabolomics data with DNA sequencing techniques for instance, to  
771 investigate the microbiome and its role in health and disease or in pharmaceutical studies to  
772 investigate impact that drugs can have on the human metabolic profile.<sup>138</sup> Integrating multiple omics  
773 data from genomics, transcriptomics, proteomics, and metabolomics as well as epigenomics or  
774 pharmacogenomics provides a deeper understanding of complex biological mechanisms linked to a  
775 specific condition. Although computational challenges persist with respect to the integration of large  
776 omics data sets together and their biological interpretation, there is a will to go toward that goal in the  
777 omics community,<sup>139</sup> and some tools, such as the R package mixOmics, have already been  
778 developed.<sup>140</sup> Omics integration is thus already attracting attention in a number of different fields,  
779 either to have more insight into the human health risks linked to environmental chemical exposition,<sup>141</sup>  
780 ecological interactions through chemical signals,<sup>142</sup> plant biology through integration of transcriptome  
781 and metabolome data sets,<sup>143</sup> or many others. Some examples cited in this review did actually combine  
782 transcriptomics and NMR- and MS-based metabolomics to study tomato fruit composition and used  
783 correlation analyses to highlight associations between genes and metabolites, to finally visualize them  
784 in a common network.<sup>76</sup> Proteome and metabolome data sets were also combined in the same aim of  
785 studying tomato fruit composition.<sup>144</sup> This combination successfully promoted biomarkers discovery in  
786 diseases with a case study based on encephalomyelitis<sup>145</sup> and promises further applications to  
787 toxicological studies following a recent published workflow.<sup>146</sup> Parallel developments in computational  
788 tools and databases, together with the pursuit of better sensitivity and resolution for NMR and MS  
789 methods, will undoubtedly increase the level of biological information that can be accessed (Figure  
790 10).



791

792 Figure 10. Schematic representation of systems biology, from which the heart relies on current progress being  
 793 made in computational tools for multiomics integration, expanding databases, and increased sensitivity and  
 794 resolution of current analytical instruments, which will allow new discovery in genomics, transcriptomics,  
 795 proteomics, and finally metabolomics, which is the closest field which can explain an organism phenotype and  
 796 which the exposome has constant impact.

797 **Conclusion**

798 In this review, various ways to combine NMR spectroscopy and MS technologies for metabolomics  
 799 have been discussed, from hardware combination to computational tools, highlighting how it made it  
 800 possible to access crucial information at the postmetabolomics level. It seems clear that metabolomics  
 801 is a growing and dynamic field which will greatly benefit from NMR and MS combination, in particular  
 802 from the computational point of view. Indeed, hardware combination seems a bit outdated when it  
 803 comes to metabolomics but will probably remain relevant in natural product research. In terms of  
 804 computational combination of NMR and MS-based metabolomics, the solutions to do so are clearly  
 805 multiplying, and the massive improvements which have been made in the recent years guarantee that  
 806 the use of correlation and data fusion approaches will spread. This will allow one to obtain more robust  
 807 statistical models and will definitely help biomarkers discoveries and a more profound understanding  
 808 of biological systems. This will be even truer if efforts in integrating metabolomics with other omics  
 809 are maintained to obtain deeper capture of the biological processes lying under an organism  
 810 phenotype and of its constant interaction with its environment, the so-called exposome.<sup>147</sup>



## 811 **Acknowledgements**

812 P.G. acknowledges support of the European Research Council under the European Union's Horizon  
813 2020 Research and Innovation Program (ERC Grant Agreement No. 814747/SUMMIT). The Corsaire  
814 metabolomics core facility is also acknowledged as well as the national infrastructure MetaboHUB.

## 815 **Biographies**

816 **Dr. Marine Letertre** studied chemistry at the Universities of Rennes and Nantes, where she received  
817 her Master Diploma in 2015. During her studies, she acquired a rich international experience with  
818 internships in Ireland and New-Zealand. Then she did a Ph.D at Imperial College London, where she  
819 applied NMR and LC-MS methods to investigate the two-way interactions existing between the gut  
820 microbiota and host drug metabolism. After obtaining her Ph.D degree in 2020, she joined the Corsaire  
821 metabolomics platform as a facilitator. In 2021, she is joining the CEISAM research institute in Nantes  
822 as a post-doctoral fellow to work on the complementarity of NMR and MS for metabolomic  
823 applications in collaboration with the National Veterinary College of Nantes. In 2020 she was elected  
824 Chair of the Early-career Member Network of the Metabolomics Society.

825 **Dr. Gaud Dervilly** studied biochemistry and food science at AgroParisTech (former ENSIA), where she  
826 received her engineer Diploma and Ph.D. degree in 1997 and 2001, respectively. She further holds an  
827 habilitation degree in Analytical Chemistry from Nantes University since 2007. She is head deputy of  
828 INRAE research unit within the National Veterinary College of Nantes, France. For 20 years, her  
829 research activity has been devoted to Chemical Food Safety issues. She is responsible for the  
830 management of research projects related to the modelling of contaminants transfer along the food  
831 chain and the evaluation of consumer's chemical exposure. Her competences range from targeted  
832 mass spectrometric approaches to more global and non-targeted strategies, such as metabolomics, to  
833 study the effects of chemical exposure and related biomarkers discovery, in a risk assessment  
834 perspective.

835 **Prof. Patrick Giraudeau** studied physics and chemistry at the University of Nantes, where he received  
836 his Ph.D. degree in 2008. He worked for one year each as a postdoctoral researcher in the Department  
837 of Chemical Physics at the Weizmann Institute of Science (Israel). In 2009, he became an associate  
838 professor at the University of Nantes, where he became a full professor in 2017. In 2014, he became a  
839 fellow of the *Institut Universitaire de France*, and received a consolidator grant from the European  
840 Research Council in 2018. His research activities at the CEISAM research institute are focused on the  
841 development of quantitative NMR methods for the analysis of complex mixtures, including  
842 applications to metabolomics and fluxomics. Research highlights include the development of fast

843 multi-dimensional quantitative experiments at high fields and also on benchtop spectrometers, as well  
844 as recent investigations in dissolution dynamic nuclear polarization.

845 The authors declare no competing financial interest.

## 846 References

- 847 (1) German, J. B.; Hammock, B. D.; Watkins, S. M. Metabolomics: Building on a Century of  
848 Biochemistry to Guide Human Health. *Metabolomics* **2005**, *1* (1), 3–9.  
849 <https://doi.org/10.1007/s11306-005-1102-8>.
- 850 (2) Holmes, E.; Wilson, I. D.; Nicholson, J. K. Metabolic Phenotyping in Health and Disease. *Cell*  
851 **2008**, *134* (5), 714–717. <https://doi.org/10.1016/j.cell.2008.08.026>.
- 852 (3) Kell, D. B.; Goodacre, R. Metabolomics and Systems Pharmacology: Why and How to Model  
853 the Human Metabolic Network for Drug Discovery. *Drug Discov. Today* **2014**, *19* (2), 171–182.  
854 <https://doi.org/10.1016/j.drudis.2013.07.014>.
- 855 (4) Tolstikov, V. Metabolomics: Bridging the Gap between Pharmaceutical Development and  
856 Population Health. *Metabolites* **2016**, *6* (3). <https://doi.org/10.3390/metabo6030020>.
- 857 (5) Ufer, M.; Juif, P.-E.; Boof, M.-L.; Muehlan, C.; Dingemans, J. Metabolite Profiling in Early  
858 Clinical Drug Development: Current Status and Future Prospects. *Expert Opin. Drug Metab.*  
859 *Toxicol.* **2017**, *13* (8), 803–806. <https://doi.org/10.1080/17425255.2017.1351944>.
- 860 (6) Nicholson, J. K. Global Systems Biology, Personalized Medicine and Molecular Epidemiology.  
861 *Mol. Syst. Biol.* **2006**, *2* (1), 52. <https://doi.org/10.1038/msb4100095>.
- 862 (7) Kamleh, M. A.; Spagou, K.; Want, E. J. Metabolic Profiling in Disease Diagnosis, Toxicology and  
863 Personalized Healthcare. *Curr. Pharm. Biotechnol.* **2011**, *12* (7), 976–995.  
864 <https://doi.org/10.2174/138920111795909069>.
- 865 (8) Koen, N.; Du Preez, I.; Loots, D. T. Chapter Three - Metabolomics and Personalized Medicine.  
866 In *Advances in Protein Chemistry and Structural Biology*; Donev, R., Ed.; Personalized Medicine;  
867 Academic Press, 2016; Vol. 102, pp 53–78. <https://doi.org/10.1016/bs.apcsb.2015.09.003>.
- 868 (9) Li, B.; He, X.; Jia, W.; Li, H. Novel Applications of Metabolomics in Personalized Medicine: A  
869 Mini-Review. *Mol. J. Synth. Chem. Nat. Prod. Chem.* **2017**, *22* (7).  
870 <https://doi.org/10.3390/molecules22071173>.
- 871 (10) Jacob, M.; Lopata, A. L.; Dasouki, M.; Abdel Rahman, A. M. Metabolomics toward Personalized  
872 Medicine. *Mass Spectrom. Rev.* **2019**, *38* (3), 221–238. <https://doi.org/10.1002/mas.21548>.
- 873 (11) Li, M.; Wang, B.; Zhang, M.; Rantalainen, M.; Wang, S.; Zhou, H.; Zhang, Y.; Shen, J.; Pang, X.;  
874 Zhang, M.; Wei, H.; Chen, Y.; Lu, H.; Zuo, J.; Su, M.; Qiu, Y.; Jia, W.; Xiao, C.; Smith, L. M.; Yang,  
875 S.; Holmes, E.; Tang, H.; Zhao, G.; Nicholson, J. K.; Li, L.; Zhao, L. Symbiotic Gut Microbes  
876 Modulate Human Metabolic Phenotypes. *Proc. Natl. Acad. Sci.* **2008**, *105* (6), 2117–2122.  
877 <https://doi.org/10.1073/pnas.0712038105>.
- 878 (12) Noecker, C.; Chiu, H.-C.; McNally, C. P.; Borenstein, E. Defining and Evaluating Microbial  
879 Contributions to Metabolite Variation in Microbiome-Metabolome Association Studies.  
880 *mSystems* **2019**, *4* (6). <https://doi.org/10.1128/mSystems.00579-19>.
- 881 (13) Koulman, A.; Volmer, D. A. Perspectives for Metabolomics in Human Nutrition: An Overview.  
882 *Nutr. Bull. BNF* **2008**, *33* (4), 324–330. <https://doi.org/10.1111/j.1467-3010.2008.00733.x>.
- 883 (14) Sébédio, J.-L. Metabolomics, Nutrition, and Potential Biomarkers of Food Quality, Intake, and  
884 Health Status. *Adv. Food Nutr. Res.* **2017**, *82*, 83–116.  
885 <https://doi.org/10.1016/bs.afnr.2017.01.001>.
- 886 (15) Garcia-Perez, I.; Poma, J. M.; Gibson, R.; Chambers, E. S.; Hansen, T. H.; Vestergaard, H.;  
887 Hansen, T.; Beckmann, M.; Pedersen, O.; Elliott, P.; Stamler, J.; Nicholson, J. K.; Draper, J.;  
888 Mathers, J. C.; Holmes, E.; Frost, G. Objective Assessment of Dietary Patterns by Use of  
889 Metabolic Phenotyping: A Randomised, Controlled, Crossover Trial. *Lancet Diabetes*  
890 *Endocrinol.* **2017**, *5* (3), 184–195. [https://doi.org/10.1016/S2213-8587\(16\)30419-3](https://doi.org/10.1016/S2213-8587(16)30419-3).



- 891 (16) Tebani, A.; Bekri, S. Paving the Way to Precision Nutrition Through Metabolomics. *Front. Nutr.*  
892 **2019**, *6*. <https://doi.org/10.3389/fnut.2019.00041>.
- 893 (17) Dixon, R. A.; Gang, D. R.; Charlton, A. J.; Fiehn, O.; Kuiper, H. A.; Reynolds, T. L.; Tjeerdema, R.  
894 S.; Jeffery, E. H.; German, J. B.; Ridley, W. P.; Seiber, J. N. Applications of Metabolomics in  
895 Agriculture. *J. Agric. Food Chem.* **2006**, *54* (24), 8984–8994.  
896 <https://doi.org/10.1021/jf061218t>.
- 897 (18) Prado, R. M. do; Porto, C.; Nunes, E.; Aguiar, C. L. de; Pilau, E. J. Metabolomics and Agriculture:  
898 What Can Be Done? *mSystems* **2018**, *3* (2). <https://doi.org/10.1128/mSystems.00156-17>.
- 899 (19) Motti, C. Environmental Marine Metabolomics: From Whole Organism System Biology to  
900 Ecosystem Management. *J. Mar. Sci. Res. Dev.* **2012**, *2* (3), 1–2. <https://doi.org/10.4172/2155-9910.1000e110>.
- 902 (20) Sogin, E. M.; Puskás, E.; Dubilier, N.; Liebeke, M. Marine Metabolomics: A Method for  
903 Nontargeted Measurement of Metabolites in Seawater by Gas Chromatography–Mass  
904 Spectrometry. *mSystems* **2019**, *4* (6). <https://doi.org/10.1128/mSystems.00638-19>.
- 905 (21) Bloszies, C. S.; Fiehn, O. Using Untargeted Metabolomics for Detecting Exposome Compounds.  
906 *Curr. Opin. Toxicol.* **2018**, *8*, 87–92. <https://doi.org/10.1016/j.cotox.2018.03.002>.
- 907 (22) Walker, D. I.; Valvi, D.; Rothman, N.; Lan, Q.; Miller, G. W.; Jones, D. P. The Metabolome: A Key  
908 Measure for Exposome Research in Epidemiology. *Curr. Epidemiol. Rep.* **2019**, *6*, 93–103.
- 909 (23) Labine, L. M.; Simpson, M. J. The Use of Nuclear Magnetic Resonance (NMR) and Mass  
910 Spectrometry (MS)–Based Metabolomics in Environmental Exposure Assessment. *Curr. Opin.*  
911 *Environ. Sci. Health* **2020**, *15*, 7–15. <https://doi.org/10.1016/j.coesh.2020.01.008>.
- 912 (24) Holzgrabe, U.; Deubner, R.; Schollmayer, C.; Waibel, B. Quantitative NMR Spectroscopy—  
913 Applications in Drug Analysis. *J. Pharm. Biomed. Anal.* **2005**, *38* (5), 806–812.  
914 <https://doi.org/10.1016/j.jpba.2005.01.050>.
- 915 (25) Ward, J. L.; Baker, J. M.; Miller, S. J.; Deborde, C.; Maucourt, M.; Biais, B.; Rolin, D.; Moing, A.;  
916 Moco, S.; Vervoort, J.; Lommen, A.; Schäfer, H.; Humpfer, E.; Beale, M. H. An Inter-Laboratory  
917 Comparison Demonstrates That <sup>1</sup>H-NMR Metabolite Fingerprinting Is a Robust Technique  
918 for Collaborative Plant Metabolomic Data Collection. *Metabolomics* **2010**, *6* (2), 263–273.  
919 <https://doi.org/10.1007/s11306-010-0200-4>.
- 920 (26) Van, Q. N.; Issaq, H. J.; Jiang, Q.; Li, Q.; Muschik, G. M.; Waybright, T. J.; Lou, H.; Dean, M.;  
921 Uitto, J.; Veenstra, T. D. Comparison of 1D and 2D NMR Spectroscopy for Metabolic Profiling.  
922 *J. Proteome Res.* **2008**, *7* (2), 630–639. <https://doi.org/10.1021/pr700594s>.
- 923 (27) Robinette, S. L.; Ajredini, R.; Rasheed, H.; Zeinomar, A.; Schroeder, F. C.; Dossey, A. T.; Edison,  
924 A. S. Hierarchical Alignment and Full Resolution Pattern Recognition of 2D NMR Spectra:  
925 Application to Nematode Chemical Ecology. *Anal. Chem.* **2011**, *83* (5), 1649–1657.  
926 <https://doi.org/10.1021/ac102724x>.
- 927 (28) Jézéquel, T.; Deborde, C.; Maucourt, M.; Zhendre, V.; Moing, A.; Giraudeau, P. Absolute  
928 Quantification of Metabolites in Tomato Fruit Extracts by Fast 2D NMR. *Metabolomics* **2015**,  
929 *11* (5), 1231–1242. <https://doi.org/10.1007/s11306-015-0780-0>.
- 930 (29) Marchand, J.; Martineau, E.; Guitton, Y.; Dervilly-Pinel, G.; Giraudeau, P. Multidimensional  
931 NMR Approaches towards Highly Resolved, Sensitive and High-Throughput Quantitative  
932 Metabolomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 49–55.  
933 <https://doi.org/10.1016/j.copbio.2016.08.004>.
- 934 (30) Marchand, J.; Martineau, E.; Guitton, Y.; Le Bizec, B.; Dervilly-Pinel, G.; Giraudeau, P. A  
935 Multidimensional <sup>1</sup>H NMR Lipidomics Workflow to Address Chemical Food Safety Issues.  
936 *Metabolomics Off. J. Metabolomic Soc.* **2018**, *14* (5), 60. <https://doi.org/10.1007/s11306-018-1360-x>.
- 938 (31) Bornet, A.; Maucourt, M.; Deborde, C.; Jacob, D.; Milani, J.; Vuichoud, B.; Ji, X.; Dumez, J.-N.;  
939 Moing, A.; Bodenhausen, G.; Jannin, S.; Giraudeau, P. Highly Repeatable Dissolution Dynamic  
940 Nuclear Polarization for Heteronuclear NMR Metabolomics. *Anal. Chem.* **2016**, *88* (12), 6179–  
941 6183. <https://doi.org/10.1021/acs.analchem.6b01094>.

- 942 (32) Plainchont, B.; Berruyer, P.; Dumez, J.-N.; Jannin, S.; Giraudeau, P. Dynamic Nuclear  
943 Polarization Opens New Perspectives for NMR Spectroscopy in Analytical Chemistry. *Anal.*  
944 *Chem.* **2018**, *90* (6), 3639–3650. <https://doi.org/10.1021/acs.analchem.7b05236>.
- 945 (33) Lerche, M. H.; Yigit, D.; Frahm, A. B.; Ardenkjær-Larsen, J. H.; Malinowski, R. M.; Jensen, P. R.  
946 Stable Isotope-Resolved Analysis with Quantitative Dissolution Dynamic Nuclear Polarization.  
947 *Anal. Chem.* **2018**, *90* (1). <https://doi.org/10.1021/acs.analchem.7b02779>.
- 948 (34) Eshuis, N.; Aspers, R. L. E. G.; van Weerdenburg, B. J. A.; Feiters, M. C.; Rutjes, F. P. J. T.;  
949 Wijmenga, S. S.; Tessari, M. 2D NMR Trace Analysis by Continuous Hyperpolarization at High  
950 Magnetic Field. *Angew. Chem. Int. Ed.* **2015**, *54* (48), 14527–14530.  
951 <https://doi.org/10.1002/anie.201507831>.
- 952 (35) Eshuis, N.; van Weerdenburg, B. J. A.; Feiters, M. C.; Rutjes, F. P. J. T.; Wijmenga, S. S.; Tessari,  
953 M. Quantitative Trace Analysis of Complex Mixtures Using SABRE Hyperpolarization. *Angew.*  
954 *Chem. Int. Ed.* **2015**, *54* (5), 1481–1484. <https://doi.org/10.1002/anie.201409795>.
- 955 (36) Orellana, G.; Vanden Bussche, J.; Van Meulebroek, L.; Vandegheuchte, M.; Janssen, C.;  
956 Vanhaecke, L. Validation of a Confirmatory Method for Lipophilic Marine Toxins in Shellfish  
957 Using UHPLC-HR-Orbitrap MS. *Anal. Bioanal. Chem.* **2014**, *406* (22), 5303–5312.  
958 <https://doi.org/10.1007/s00216-014-7958-6>.
- 959 (37) Nye, L. C.; Williams, J. P.; Munjoma, N. C.; Letertre, M. P. M.; Coen, M.; Bouwmeester, R.;  
960 Martens, L.; Swann, J. R.; Nicholson, J. K.; Plumb, R. S.; McCullagh, M.; Gethings, L. A.; Lai, S.;  
961 Langridge, J. I.; Vissers, J. P. C.; Wilson, I. D. A Comparison of Collision Cross Section Values  
962 Obtained via Travelling Wave Ion Mobility-Mass Spectrometry and Ultra High Performance  
963 Liquid Chromatography-Ion Mobility-Mass Spectrometry: Application to the Characterisation  
964 of Metabolites in Rat Urine. *J. Chromatogr. A* **2019**, *1602*, 386–396.  
965 <https://doi.org/10.1016/j.chroma.2019.06.056>.
- 966 (38) Zelena, E.; Dunn, W. B.; Broadhurst, D.; Francis-McIntyre, S.; Carroll, K. M.; Begley, P.; O’Hagan,  
967 S.; Knowles, J. D.; Halsall, A.; Wilson, I. D.; Kell, D. B. Development of a Robust and Repeatable  
968 UPLC–MS Method for the Long-Term Metabolomic Study of Human Serum. *Anal. Chem.* **2009**,  
969 *81* (4), 1357–1364. <https://doi.org/10.1021/ac8019366>.
- 970 (39) Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown,  
971 M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Kell, D. B.; Goodacre,  
972 R. Procedures for Large-Scale Metabolic Profiling of Serum and Plasma Using Gas  
973 Chromatography and Liquid Chromatography Coupled to Mass Spectrometry. *Nat. Protoc.*  
974 **2011**, *6* (7), 1060–1083. <https://doi.org/10.1038/nprot.2011.335>.
- 975 (40) Dunn, W. B.; Lin, W.; Broadhurst, D.; Begley, P.; Brown, M.; Zelena, E.; Vaughan, A. A.; Halsall,  
976 A.; Harding, N.; Knowles, J. D.; Francis-McIntyre, S.; Tseng, A.; Ellis, D. I.; O’Hagan, S.; Aarons,  
977 G.; Benjamin, B.; Chew-Graham, S.; Moseley, C.; Potter, P.; Winder, C. L.; Potts, C.; Thornton,  
978 P.; McWhirter, C.; Zubair, M.; Pan, M.; Burns, A.; Cruickshank, J. K.; Jayson, G. C.; Purandare,  
979 N.; Wu, F. C. W.; Finn, J. D.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Goodacre, R.; Kell, D.  
980 B. Molecular Phenotyping of a UK Population: Defining the Human Serum Metabolome.  
981 *Metabolomics* **2015**, *11* (1), 9–26. <https://doi.org/10.1007/s11306-014-0707-1>.
- 982 (41) Gika, H. G.; Zisi, C.; Theodoridis, G.; Wilson, I. D. Protocol for Quality Control in Metabolic  
983 Profiling of Biological Fluids by U(H)PLC-MS. *J. Chromatogr. B* **2016**, *1008*, 15–25.  
984 <https://doi.org/10.1016/j.jchromb.2015.10.045>.
- 985 (42) Hernández-Mesa, M.; Monteau, F.; Le Bizec, B.; Dervilly-Pinel, G. Potential of Ion Mobility-  
986 Mass Spectrometry for Both Targeted and Non-Targeted Analysis of Phase II Steroid  
987 Metabolites in Urine. *Anal. Chim. Acta X* **2019**, *1*, 100006.  
988 <https://doi.org/10.1016/j.acax.2019.100006>.
- 989 (43) Pan, Z.; Raftery, D. Comparing and Combining NMR Spectroscopy and Mass Spectrometry in  
990 Metabolomics. *Anal. Bioanal. Chem.* **2007**, *387* (2), 525–527. <https://doi.org/10.1007/s00216-006-0687-8>.
- 991

- 992 (44) Marshall, D. D.; Powers, R. Beyond the Paradigm: Combining Mass Spectrometry and Nuclear  
993 Magnetic Resonance for Metabolomics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *100*, 1–16.  
994 <https://doi.org/10.1016/j.pnmrs.2017.01.001>.
- 995 (45) Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The Re-Emergence of Natural Products for Drug  
996 Discovery in the Genomics Era. *Nat. Rev. Drug Discov.* **2015**, *14* (2), 111–129.  
997 <https://doi.org/10.1038/nrd4510>.
- 998 (46) Giraudeau, P. NMR-Based Metabolomics and Fluxomics: Developments and Future Prospects.  
999 *Analyst* **2020**, *145* (7), 2457–2472. <https://doi.org/10.1039/D0AN00142B>.
- 1000 (47) Ren, J.-L.; Zhang, A.-H.; Kong, L.; Wang, X.-J. Advances in Mass Spectrometry-Based  
1001 Metabolomics for Investigation of Metabolites. *RSC Adv.* **2018**, *8* (40), 22335–22350.  
1002 <https://doi.org/10.1039/C8RA01574K>.
- 1003 (48) Corcoran, O.; Spraul, M. LC–NMR–MS in Drug Discovery. *Drug Discov. Today* **2003**, *8* (14), 624–  
1004 631. [https://doi.org/10.1016/S1359-6446\(03\)02749-1](https://doi.org/10.1016/S1359-6446(03)02749-1).
- 1005 (49) Gebretsadik, T.; Linert, W.; Thomas, M.; Berhanu, T.; Frew, R. LC-NMR for Natural Products  
1006 Analysis: A Journey from an Academic Curiosity to a Robust Analytical Tool. *Sci* **2019**, *1* (1), 31.  
1007 <https://doi.org/10.3390/sci1010031>.
- 1008 (50) Shockcor, J. P.; Unger, S. E.; Wilson, I. D.; Foxall, P. J. D.; Nicholson, J. K.; Lindon, J. C. Combined  
1009 HPLC, NMR Spectroscopy, and Ion-Trap Mass Spectrometry with Application to the Detection  
1010 and Characterization of Xenobiotic and Endogenous Metabolites in Human Urine. *Anal. Chem.*  
1011 **1996**, *68* (24), 4431–4435. <https://doi.org/10.1021/ac9606463>.
- 1012 (51) Bhatia, A.; Sarma, S. J.; Lei, Z.; Sumner, L. W. UHPLC-QTOF-MS/MS-SPE-NMR: A Solution to the  
1013 Metabolomics Grand Challenge of Higher-Throughput, Confident Metabolite Identifications.  
1014 In *NMR-Based Metabolomics: Methods and Protocols*; Gowda, G. A. N., Raftery, D., Eds.;  
1015 Methods in Molecular Biology; Springer: New York, NY, 2019; pp 113–133.  
1016 [https://doi.org/10.1007/978-1-4939-9690-2\\_7](https://doi.org/10.1007/978-1-4939-9690-2_7).
- 1017 (52) van der Hooft, J. J. J.; de Vos, R. C. H.; Mihaleva, V.; Bino, R. J.; Ridder, L.; de Roo, N.; Jacobs,  
1018 D. M.; van Duynhoven, J. P. M.; Vervoort, J. Structural Elucidation and Quantification of  
1019 Phenolic Conjugates Present in Human Urine after Tea Intake. *Anal. Chem.* **2012**, *84* (16),  
1020 7263–7271. <https://doi.org/10.1021/ac3017339>.
- 1021 (53) Wolfender, J.-L.; Ndjoko, K.; Hostettmann, K. The Potential of LC-NMR in Phytochemical  
1022 Analysis. *Phytochem. Anal.* **2001**, *12* (1), 2–22. [https://doi.org/10.1002/1099-1565\(200101/02\)12:1<2::AID-PCA552>3.0.CO;2-K](https://doi.org/10.1002/1099-1565(200101/02)12:1<2::AID-PCA552>3.0.CO;2-K).
- 1024 (54) Lin, Y.; Schiavo, S.; Orjala, J.; Vouros, P.; Kautz, R. Microscale LC-MS-NMR Platform Applied to  
1025 the Identification of Active Cyanobacterial Metabolites. *Anal. Chem.* **2008**, *80* (21), 8045–8054.  
1026 <https://doi.org/10.1021/ac801049k>.
- 1027 (55) Hooft, J. J. J.; De Vos, R.; Ridder, L.; Vervoort, J. J. M.; Bino, R. Structural Elucidation of Low  
1028 Abundant Metabolites in Complex Sample Matrices. *Metabolomics* **2013**, *9*,  
1029 <https://doi.org/10.1007/s11306-013-0519-8>.
- 1030 (56) Silva Elipe, M. V. Advantages and Disadvantages of Nuclear Magnetic Resonance Spectroscopy  
1031 as a Hyphenated Technique. *Anal. Chim. Acta* **2003**, *497* (1), 1–25.  
1032 <https://doi.org/10.1016/j.aca.2003.08.048>.
- 1033 (57) Martin, J.-C.; Maillot, M.; Mazerolles, G.; Verdu, A.; Lyan, B.; Migné, C.; Defoort, C.; Canlet, C.;  
1034 Junot, C.; Guillou, C.; Manach, C.; Jacob, D.; Bouveresse, D. J.-R.; Paris, E.; Pujos-Guillot, E.;  
1035 Jourdan, F.; Giacomoni, F.; Courant, F.; Favé, G.; Le Gall, G.; Chassaigne, H.; Tabet, J.-C.; Martin,  
1036 J.-F.; Antignac, J.-P.; Shintu, L.; Defernez, M.; Philo, M.; Alexandre-Gouaubau, M.-C.; Amiot-  
1037 Carlin, M.-J.; Bossis, M.; Triba, M. N.; Stojilkovic, N.; Banzet, N.; Molinié, R.; Bott, R.;  
1038 Goulitquer, S.; Caldarelli, S.; Rutledge, D. N. Can We Trust Untargeted Metabolomics? Results  
1039 of the Metabo-Ring Initiative, a Large-Scale, Multi-Instrument Inter-Laboratory Study.  
1040 *Metabolomics* **2015**, *11* (4), 807–821. <https://doi.org/10.1007/s11306-014-0740-0>.
- 1041 (58) Moco, S.; Vervoort, J.; Moco, S.; Bino, R. J.; De Vos, R. C. H.; Bino, R. Metabolomics  
1042 Technologies and Metabolite Identification. *TrAC Trends Anal. Chem.* **2007**, *26* (9), 855–866.  
1043 <https://doi.org/10.1016/j.trac.2007.08.003>.

- 1044 (59) Atanasova-Penichon, V.; Legoahec, L.; Bernillon, S.; Deborde, C.; Maucourt, M.; Verdal-Bonnin,  
1045 M.-N.; Pinson-Gadais, L.; Ponts, N.; Moing, A.; Richard-Forget, F. Mycotoxin Biosynthesis and  
1046 Central Metabolism Are Two Interlinked Pathways in *Fusarium Graminearum*, as  
1047 Demonstrated by the Extensive Metabolic Changes Induced by Caffeic Acid Exposure. *Appl.*  
1048 *Environ. Microbiol.* **2018**, *84* (8). <https://doi.org/10.1128/AEM.01705-17>.
- 1049 (60) Goulitquer, S.; Croyal, M.; Lalande, J.; Royer, A.-L.; Guitton, Y.; Arzur, D.; Durand, S.; Le Jossic-  
1050 Corcos, C.; Bouchereau, A.; Potin, P.; Akoka, S.; Antignac, J.-P.; Krempf, M.; Ferchaud-Roucher,  
1051 V.; Giraudeau, P.; Corcos, L. Consequences of Blunting the Mevalonate Pathway in Cancer  
1052 Identified by a Pluri-Omics Approach. *Cell Death Dis.* **2018**, *9* (7), 745.  
1053 <https://doi.org/10.1038/s41419-018-0761-0>.
- 1054 (61) Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.;  
1055 Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang, P.;  
1056 Hollander, Z.; Pedersen, T. L.; Smith, S. R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J.  
1057 W.; Goodfriend, T.; Wishart, D. S. The Human Serum Metabolome. *PLOS ONE* **2011**, *6* (2),  
1058 e16957. <https://doi.org/10.1371/journal.pone.0016957>.
- 1059 (62) Bruno, C.; Patin, F.; Bocca, C.; Nadal-Desbarats, L.; Bonnier, F.; Reynier, P.; Emond, P.; Vourc'h,  
1060 P.; Joseph-Delafont, K.; Corcia, P.; Andres, C. R.; Blasco, H. The Combination of Four Analytical  
1061 Methods to Explore Skeletal Muscle Metabolomics: Better Coverage of Metabolic Pathways  
1062 or a Marketing Argument? *J. Pharm. Biomed. Anal.* **2018**, *148*, 273–279.  
1063 <https://doi.org/10.1016/j.jpba.2017.10.013>.
- 1064 (63) Allwood, J. W.; Cheung, W.; Xu, Y.; Mumm, R.; De Vos, R. C. H.; Deborde, C.; Biais, B.; Maucourt,  
1065 M.; Berger, Y.; Schaffer, A. A.; Rolin, D.; Moing, A.; Hall, R. D.; Goodacre, R. Metabolomics in  
1066 Melon: A New Opportunity for Aroma Analysis. *Phytochemistry* **2014**, *99*, 61–72.  
1067 <https://doi.org/10.1016/j.phytochem.2013.12.010>.
- 1068 (64) Atherton, H. J.; Bailey, N. J.; Zhang, W.; Taylor, J.; Major, H.; Shockcor, J.; Clarke, K.; Griffin, J.  
1069 L. A Combined <sup>1</sup>H-NMR Spectroscopy- and Mass Spectrometry-Based Metabolomic Study of  
1070 the PPAR- $\alpha$  Null Mutant Mouse Defines Profound Systemic Changes in Metabolism Linked to  
1071 the Metabolic Syndrome. *Physiol. Genomics* **2006**, *27* (2), 178–186.  
1072 <https://doi.org/10.1152/physiolgenomics.00060.2006>.
- 1073 (65) Tayyari, F.; Gowda, G. A. N.; Gu, H.; Raftery, D. <sup>15</sup>N-Cholamine—A Smart Isotope Tag for  
1074 Combining NMR- and MS-Based Metabolite Profiling. *Anal. Chem.* **2013**, *85* (18), 8715–8721.  
1075 <https://doi.org/10.1021/ac401712a>.
- 1076 (66) Boccard, J.; Rudaz, S. Harnessing the Complexity of Metabolomic Data with Chemometrics. *J.*  
1077 *Chemom.* **2014**, *28* (1), 1–9. <https://doi.org/10.1002/cem.2567>.
- 1078 (67) Noda, I. Generalized Two-Dimensional Correlation Method Applicable to Infrared, Raman, and  
1079 Other Types of Spectroscopy. *Appl. Spectrosc.* **1993**, *47* (9), 1329–1336.
- 1080 (68) Robinette, S. L.; Lindon, J. C.; Nicholson, J. K. Statistical Spectroscopic Tools for Biomarker  
1081 Discovery and Systems Medicine. *Anal. Chem.* **2013**, *85* (11), 5297–5303.  
1082 <https://doi.org/10.1021/ac4007254>.
- 1083 (69) Cloarec, O.; Dumas, M.-E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier,  
1084 D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical Total Correlation Spectroscopy: An  
1085 Exploratory Approach for Latent Biomarker Identification from Metabolic <sup>1</sup>H NMR Data Sets.  
1086 *Anal. Chem.* **2005**, *77* (5), 1282–1289. <https://doi.org/10.1021/ac048630x>.
- 1087 (70) Crockford, D. J.; Holmes, E.; Lindon, J. C.; Plumb, R. S.; Zirah, S.; Bruce, S. J.; Rainville, P.;  
1088 Stumpf, C. L.; Nicholson, J. K. Statistical Heterospectroscopy, an Approach to the Integrated  
1089 Analysis of NMR and UPLC-MS Data Sets: Application in Metabonomic Toxicology Studies.  
1090 *Anal. Chem.* **2006**, *78* (2), 363–371. <https://doi.org/10.1021/ac051444m>.
- 1091 (71) Crockford, D. J.; Maher, A. D.; Ahmadi, K. R.; Barrett, A.; Plumb, R. S.; Wilson, I. D.; Nicholson,  
1092 J. K. <sup>1</sup>H NMR and UPLC-MSE Statistical Heterospectroscopy: Characterization of Drug  
1093 Metabolites (Xenometabolome) in Epidemiological Studies. *Anal. Chem.* **2008**, *80* (18), 6835–  
1094 6844. <https://doi.org/10.1021/ac801075m>.

- 1095 (72) Marti, G.; Boccard, J.; Mehl, F.; Debrus, B.; Marcourt, L.; Merle, P.; Delort, E.; Baroux, L.;  
1096 Sommer, H.; Rudaz, S.; Wolfender, J.-L. Comprehensive Profiling and Marker Identification in  
1097 Non-Volatile Citrus Oil Residues by Mass Spectrometry and Nuclear Magnetic Resonance. *Food*  
1098 *Chem.* **2014**, *150*, 235–245. <https://doi.org/10.1016/j.foodchem.2013.10.103>.
- 1099 (73) Bertrand, S.; Azzollini, A.; Nievergelt, A.; Boccard, J.; Rudaz, S.; Cuendet, M.; Wolfender, J.-L.  
1100 Statistical Correlations between HPLC Activity-Based Profiling Results and NMR/MS  
1101 Microfraction Data to Deconvolute Bioactive Compounds in Mixtures. *Molecules* **2016**, *21* (3),  
1102 259. <https://doi.org/10.3390/molecules21030259>.
- 1103 (74) Pan, Z.; Gu, H.; Talaty, N.; Chen, H.; Shanaiah, N.; Hainline, B. E.; Cooks, R. G.; Raftery, D.  
1104 Principal Component Analysis of Urine Metabolites Detected by NMR and DESI-MS in Patients  
1105 with Inborn Errors of Metabolism. *Anal. Bioanal. Chem.* **2007**, *387* (2), 539–549.  
1106 <https://doi.org/10.1007/s00216-006-0546-7>.
- 1107 (75) Moing, A.; Aharoni, A.; Biais, B.; Rogachev, I.; Meir, S.; Brodsky, L.; Allwood, J. W.; Erban, A.;  
1108 Dunn, W. B.; Kay, L.; Koning, S. de; Vos, R. C. H. de; Jonker, H.; Mumm, R.; Deborde, C.;  
1109 Maucourt, M.; Bernillon, S.; Gibon, Y.; Hansen, T. H.; Husted, S.; Goodacre, R.; Kopka, J.;  
1110 Schjoerring, J. K.; Rolin, D.; Hall, R. D. Extensive Metabolic Cross-Talk in Melon Fruit Revealed  
1111 by Spatial and Developmental Combinatorial Metabolomics. *New Phytol.* **2011**, *190* (3), 683–  
1112 696. <https://doi.org/10.1111/j.1469-8137.2010.03626.x>.
- 1113 (76) Mounet, F.; Moing, A.; Garcia, V.; Petit, J.; Maucourt, M.; Deborde, C.; Bernillon, S.; Gall, G. L.;  
1114 Colquhoun, I.; Defernez, M.; Giraudel, J.-L.; Rolin, D.; Rothan, C.; Lemaire-Chamley, M. Gene  
1115 and Metabolite Regulatory Network Analysis of Early Developing Fruit Tissues Highlights New  
1116 Candidate Genes for the Control of Tomato Fruit Composition and Development. *Plant Physiol.*  
1117 **2009**, *149* (3), 1505–1528. <https://doi.org/10.1104/pp.108.133967>.
- 1118 (77) Rubingh, C. M.; Bijlsma, S.; Derks, E. P. P. A.; Bobeldijk, I.; Verheij, E. R.; Kochhar, S.; Smilde, A.  
1119 K. Assessing the Performance of Statistical Validation Tools for Megavariate Metabolomics  
1120 Data. *Metabolomics* **2006**, *2* (2), 53–61. <https://doi.org/10.1007/s11306-006-0022-6>.
- 1121 (78) Xi, B.; Gu, H.; Baniyadi, H.; Raftery, D. Statistical Analysis and Modeling of Mass Spectrometry-  
1122 Based Metabolomics Data. *Methods Mol. Biol. Clifton NJ* **2014**, *1198*, 333–353.  
1123 [https://doi.org/10.1007/978-1-4939-1258-2\\_22](https://doi.org/10.1007/978-1-4939-1258-2_22).
- 1124 (79) Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van  
1125 Duijnhoven, J. P. M.; van Dorsten, F. A. Assessment of PLS-DA Cross Validation. *Metabolomics*  
1126 **2008**, *4* (1), 81–89. <https://doi.org/10.1007/s11306-007-0099-6>.
- 1127 (80) Esbensen, K. H.; Geladi, P. Principles of Proper Validation: Use and Abuse of Re-Sampling for  
1128 Validation. *J. Chemom.* **2010**, *24* (3–4), 168–187. <https://doi.org/10.1002/cem.1310>.
- 1129 (81) Szymańska, E.; Saccenti, E.; Smilde, A. K.; Westerhuis, J. A. Double-Check: Validation of  
1130 Diagnostic Statistics for PLS-DA Models in Metabolomics Studies. *Metabolomics Off. J.*  
1131 *Metabolomic Soc.* **2012**, *8* (Suppl 1), 3–16. <https://doi.org/10.1007/s11306-011-0330-3>.
- 1132 (82) Doeswijk, T. G.; Smilde, A. K.; Hageman, J. A.; Westerhuis, J. A.; van Eeuwijk, F. A. On the  
1133 Increase of Predictive Performance with High-Level Data Fusion. *Anal. Chim. Acta* **2011**, *705*  
1134 (1), 41–47. <https://doi.org/10.1016/j.aca.2011.03.025>.
- 1135 (83) Smilde, A. K.; Westerhuis, J. A.; Jong, S. de. A Framework for Sequential Multiblock Component  
1136 Methods. *J. Chemom.* **2003**, *17* (6), 323–337. <https://doi.org/10.1002/cem.811>.
- 1137 (84) Westerhuis, J. A.; Kourti, T.; MacGregor, J. F. Analysis of Multiblock and Hierarchical PCA and  
1138 PLS Models. *J. Chemom.* **1998**, *12* (5), 301–321. [https://doi.org/10.1002/\(SICI\)1099-128X\(199809/10\)12:5<301::AID-CEM515>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-128X(199809/10)12:5<301::AID-CEM515>3.0.CO;2-S).
- 1140 (85) Biais, B.; Allwood, J. W.; Deborde, C.; Xu, Y.; Maucourt, M.; Beauvoit, B.; Dunn, W. B.; Jacob,  
1141 D.; Goodacre, R.; Rolin, D.; Moing, A. 1H NMR, GC-EI-TOFMS, and Data Set Correlation for Fruit  
1142 Metabolomics: Application to Spatial Metabolite Analysis in Melon. *Anal. Chem.* **2009**, *81* (8),  
1143 2884–2894. <https://doi.org/10.1021/ac9001996>.
- 1144 (86) Bénard, C.; Bernillon, S.; Biais, B.; Osorio, S.; Maucourt, M.; Ballias, P.; Deborde, C.; Colombié,  
1145 S.; Cabasson, C.; Jacob, D.; Vercambre, G.; Gautier, H.; Rolin, D.; Génard, M.; Fernie, A. R.;  
1146 Gibon, Y.; Moing, A. Metabolomic Profiling in Tomato Reveals Diel Compositional Changes in

- 1147 Fruit Affected by Source–Sink Relationships. *J. Exp. Bot.* **2015**, *66* (11), 3391–3404.  
 1148 <https://doi.org/10.1093/jxb/erv151>.
- 1149 (87) Spiteri, M.; Dubin, E.; Cotton, J.; Poirel, M.; Corman, B.; Jamin, E.; Lees, M.; Rutledge, D. Data  
 1150 Fusion between High Resolution (1)H-NMR and Mass Spectrometry: A Synergetic Approach to  
 1151 Honey Botanical Origin Characterization. *Anal. Bioanal. Chem.* **2016**, *408* (16), 4389–4401.  
 1152 <https://doi.org/10.1007/s00216-016-9538-4>.
- 1153 (88) Teul, J.; Rupérez, F. J.; Garcia, A.; Vaysse, J.; Balayssac, S.; Gilard, V.; Malet-Martino, M.;  
 1154 Martin-Ventura, J. L.; Blanco-Colio, L. M.; Tuñón, J.; Egido, J.; Barbas, C. Improving Metabolite  
 1155 Knowledge in Stable Atherosclerosis Patients by Association and Correlation of GC-MS and 1H  
 1156 NMR Fingerprints. *J. Proteome Res.* **2009**, *8* (12), 5580–5589.  
 1157 <https://doi.org/10.1021/pr900668v>.
- 1158 (89) Gu, H.; Pan, Z.; Xi, B.; Asiago, V.; Musselman, B.; Raftery, D. Principal Component Directed  
 1159 Partial Least Squares Analysis for Combining Nuclear Magnetic Resonance and Mass  
 1160 Spectrometry Data in Metabolomics: Application to the Detection of Breast Cancer. *Anal.*  
 1161 *Chim. Acta* **2011**, *686* (1), 57–63. <https://doi.org/10.1016/j.aca.2010.11.040>.
- 1162 (90) Mehl, F.; Marti, G.; Merle, P.; Delort, E.; Baroux, L.; Sommer, H.; Wolfender, J.-L.; Rudaz, S.;  
 1163 Boccard, J. Integrating Metabolomic Data from Multiple Analytical Platforms for a  
 1164 Comprehensive Characterisation of Lemon Essential Oils. *Flavour Fragr. J.* **2015**, *30* (2), 131–  
 1165 138. <https://doi.org/10.1002/ffj.3230>.
- 1166 (91) Mehl, F.; Marti, G.; Boccard, J.; Debrus, B.; Merle, P.; Delort, E.; Baroux, L.; Raymo, V.; Velazco,  
 1167 M. I.; Sommer, H.; Wolfender, J.-L.; Rudaz, S. Differentiation of Lemon Essential Oil Based on  
 1168 Volatile and Non-Volatile Fractions with Various Analytical Techniques: A Metabolomic  
 1169 Approach. *Food Chem.* **2014**, *143*, 325–335.  
 1170 <https://doi.org/10.1016/j.foodchem.2013.07.125>.
- 1171 (92) Marshall, D. D.; Lei, S.; Worley, B.; Huang, Y.; Garcia-Garcia, A.; Franco, R.; Dodds, E. D.;  
 1172 Powers, R. Combining DI-ESI–MS and NMR Datasets for Metabolic Profiling. *Metabolomics*  
 1173 **2015**, *11* (2), 391–402. <https://doi.org/10.1007/s11306-014-0704-4>.
- 1174 (93) Palaric, C.; Pilard, S.; Fontaine, J.-X.; Boccard, J.; Mathiron, D.; Rigaud, S.; Cailleu, D.; Mesnard,  
 1175 F.; Gut, Y.; Renaud, T.; Petit, A.; Beaumal, J.-Y.; Molinié, R. Processing of NMR and MS  
 1176 Metabolomics Data Using Chemometrics Methods: A Global Tool for Fungi Biotransformation  
 1177 Reactions Monitoring. *Metabolomics* **2019**, *15* (8), 107. <https://doi.org/10.1007/s11306-019-1567-5>.
- 1179 (94) Acar, E.; Gürdeniz, G.; Savorani, F.; Hansen, L.; Olsen, A.; Tjønneland, A.; Dragsted, L. O.; Bro,  
 1180 R. Forecasting Chronic Diseases Using Data Fusion. *J. Proteome Res.* **2017**, *16* (7), 2435–2444.  
 1181 <https://doi.org/10.1021/acs.jproteome.7b00039>.
- 1182 (95) Deng, L.; Gu, H.; Zhu, J.; Nagana Gowda, G. A.; Djukovic, D.; Chiorean, E. G.; Raftery, D.  
 1183 Combining NMR and LC/MS Using Backward Variable Elimination: Metabolomics Analysis of  
 1184 Colorectal Cancer, Polyps, and Healthy Controls. *Anal. Chem.* **2016**, *88* (16), 7975–7983.  
 1185 <https://doi.org/10.1021/acs.analchem.6b00885>.
- 1186 (96) Karaman, İ.; Nørskov, N. P.; Yde, C. C.; Hedemann, M. S.; Bach Knudsen, K. E.; Kohler, A. Sparse  
 1187 Multi-Block PLSR for Biomarker Discovery When Integrating Data from LC–MS and NMR  
 1188 Metabolomics. *Metabolomics* **2015**, *11* (2), 367–379. <https://doi.org/10.1007/s11306-014-0698-y>.
- 1190 (97) Fernandez, O.; Urrutia, M.; Berton, T.; Bernillon, S.; Deborde, C.; Jacob, D.; Maucourt, M.;  
 1191 Maury, P.; Duruflé, H.; Gibon, Y.; Langlade, N. B.; Moing, A. Metabolomic Characterization of  
 1192 Sunflower Leaf Allows Discriminating Genotype Groups or Stress Levels with a Minimal Set of  
 1193 Metabolic Markers. *Metabolomics* **2019**, *15* (4). <https://doi.org/10.1007/s11306-019-1515-4>.
- 1194 (98) Fiehn, O.; Robertson, D.; Griffin, J.; van der Werf, M.; Nikolau, B.; Morrison, N.; Sumner, L. W.;  
 1195 Goodacre, R.; Hardy, N. W.; Taylor, C.; Fostel, J.; Kristal, B.; Kaddurah-Daouk, R.; Mendes, P.;  
 1196 van Ommen, B.; Lindon, J. C.; Sansone, S.-A. The Metabolomics Standards Initiative (MSI).  
 1197 *Metabolomics* **2007**, *3* (3), 175–178. <https://doi.org/10.1007/s11306-007-0070-6>.

- 1198 (99) Sumner, L.; Amberg, A.; Barrett, D.; Beale, M.; Beger, R.; Daykin, C.; Fan, T.; Fiehn, O.;  
1199 Goodacre, R.; Griffin, J.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A.;  
1200 Lindon, J.; Marriott, P.; Nicholls, A.; Viant, M. Proposed Minimum Reporting Standards for  
1201 Chemical Analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards  
1202 Initiative (MSI). *Metabolomics Off. J. Metabolomic Soc.* **2007**, *3*, 211–221.  
1203 <https://doi.org/10.1007/s11306-007-0082-2>.
- 1204 (100) Marti, G.; Erb, M.; Boccard, J.; Glauser, G.; Doyen, G. R.; Villard, N.; Robert, C. a M.; Turlings,  
1205 T. C. J.; Rudaz, S.; Wolfender, J.-L. Metabolomics Reveals Herbivore-Induced Metabolites of  
1206 Resistance and Susceptibility in Maize Leaves and Roots. *Plant Cell Environ.* **2013**, *36* (3), 621–  
1207 639. <https://doi.org/10.1111/pce.12002>.
- 1208 (101) Glauser, G.; Guillarme, D.; Grata, E.; Boccard, J.; Thiocone, A.; Carrupt, P.-A.; Veuthey, J.-L.;  
1209 Rudaz, S.; Wolfender, J.-L. Optimized Liquid Chromatography–Mass Spectrometry Approach  
1210 for the Isolation of Minor Stress Biomarkers in Plant Extracts and Their Identification by  
1211 Capillary Nuclear Magnetic Resonance. *J. Chromatogr. A* **2008**, *1180* (1), 90–98.  
1212 <https://doi.org/10.1016/j.chroma.2007.12.021>.
- 1213 (102) Bingol, K.; Bruschweiler-Li, L.; Yu, C.; Somogyi, A.; Zhang, F.; Bruschweiler, R. Metabolomics  
1214 Beyond Spectroscopic Databases: A Combined MS/NMR Strategy for the Rapid Identification  
1215 of New Metabolites in Complex Mixtures. *Anal. Chem.* **2015**, *87* (7), 3864–3870.  
1216 <https://doi.org/10.1021/ac504633z>.
- 1217 (103) Bingol, K.; Bruschweiler, R. Knowns and Unknowns in Metabolomics Identified by  
1218 Multidimensional NMR and Hybrid MS/NMR Methods. *Curr. Opin. Biotechnol.* **2017**, *43*, 17–  
1219 24. <https://doi.org/10.1016/j.copbio.2016.07.006>.
- 1220 (104) Leggett, A.; Wang, C.; Li, D.-W.; Somogyi, A.; Bruschweiler-Li, L.; Bruschweiler, R. Chapter  
1221 Eleven - Identification of Unknown Metabolomics Mixture Compounds by Combining NMR,  
1222 MS, and Cheminformatics. In *Methods in Enzymology*; Wand, A. J., Ed.; Biological NMR Part B;  
1223 Academic Press, 2019; Vol. 615, pp 407–422. <https://doi.org/10.1016/bs.mie.2018.09.003>.
- 1224 (105) Bingol, K.; Bruschweiler, R. NMR/MS Translator for the Enhanced Simultaneous Analysis of  
1225 Metabolomics Mixtures by NMR Spectroscopy and Mass Spectrometry: Application to Human  
1226 Urine. *J. Proteome Res.* **2015**, *14* (6), 2642–2648.  
1227 <https://doi.org/10.1021/acs.jproteome.5b00184>.
- 1228 (106) Djukovic, D.; Raftery, D.; Gowda, N. Chapter 16 - Mass Spectrometry and NMR Spectroscopy  
1229 Based Quantitative Metabolomics. In *Proteomic and Metabolomic Approaches to Biomarker*  
1230 *Discovery (Second Edition)*; Issaq, H. J., Veenstra, T. D., Eds.; Academic Press: Boston, 2020; pp  
1231 289–311. <https://doi.org/10.1016/B978-0-12-818607-7.00016-5>.
- 1232 (107) Djukovic, D.; Nagana Gowda, G. A.; Raftery, D. Chapter 18 - Mass Spectrometry and NMR  
1233 Spectroscopy–Based Quantitative Metabolomics. In *Proteomic and Metabolomic Approaches*  
1234 *to Biomarker Discovery*; Issaq, H. J., Veenstra, T. D., Eds.; Academic Press: Boston, 2013; pp  
1235 279–297. <https://doi.org/10.1016/B978-0-12-394446-7.00018-2>.
- 1236 (108) Simmler, C.; Napolitano, J. G.; McAlpine, J. B.; Chen, S.-N.; Pauli, G. F. Universal Quantitative  
1237 NMR Analysis of Complex Natural Samples. *Curr. Opin. Biotechnol.* **2014**, *25*, 51–59.  
1238 <https://doi.org/10.1016/j.copbio.2013.08.004>.
- 1239 (109) Marchand, J.; Martineau, E.; Guitton, Y.; Dervilly-Pinel, G.; Giraudeau, P. Multidimensional  
1240 NMR Approaches towards Highly Resolved, Sensitive and High-Throughput Quantitative  
1241 Metabolomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 49–55.  
1242 <https://doi.org/10.1016/j.copbio.2016.08.004>.
- 1243 (110) 2002/657/EC: Commission Decision of 12 August 2002 Implementing Council Directive  
1244 96/23/EC Concerning the Performance of Analytical Methods and the Interpretation of Results  
1245 (Text with EEA Relevance) (Notified under Document Number C(2002) 3044); 2002; Vol. 221.
- 1246 (111) Guo, K.; Li, L. Differential 12C-/13C-Isotope Dansylation Labeling and Fast Liquid  
1247 Chromatography/Mass Spectrometry for Absolute and Relative Quantification of the  
1248 Metabolome. *Anal. Chem.* **2009**, *81* (10), 3919–3932. <https://doi.org/10.1021/ac900166a>.

- 1249 (112) Nagana Gowda, G. A.; Djukovic, D.; Bettcher, L. F.; Gu, H.; Raftery, D. NMR-Guided Mass  
1250 Spectrometry for Absolute Quantitation of Human Metabolites. *Anal. Chem.* **2018**, *90* (3),  
1251 2001–2009. <https://doi.org/10.1021/acs.analchem.7b04089>.
- 1252 (113) Fei, Q.; Wang, D.; Jasbi, P.; Zhang, P.; Nagana Gowda, G. A.; Raftery, D.; Gu, H. Combining NMR  
1253 and MS with Chemical Derivatization for Absolute Quantification with Reduced Matrix Effects.  
1254 *Anal. Chem.* **2019**, *91* (6), 4055–4062. <https://doi.org/10.1021/acs.analchem.8b05611>.
- 1255 (114) Lane, A. N.; Fan, T. W. -M.; Higashi, R. M. Isotopomer-Based Metabolomic Analysis by NMR  
1256 and Mass Spectrometry. In *Methods in Cell Biology*; Biophysical Tools for Biologists, Volume  
1257 One: In Vitro Techniques; Academic Press, 2008; Vol. 84, pp 541–588.  
1258 [https://doi.org/10.1016/S0091-679X\(07\)84018-0](https://doi.org/10.1016/S0091-679X(07)84018-0).
- 1259 (115) Fan, T. W.-M.; Yuan, P.; Lane, A. N.; Higashi, R. M.; Wang, Y.; Hamidi, A. B.; Zhou, R.; Guitart,  
1260 X.; Chen, G.; Manji, H. K.; Kaddurah-Daouk, R. Stable Isotope-Resolved Metabolomic Analysis  
1261 of Lithium Effects on Glial-Neuronal Metabolism and Interactions. *Metabolomics* **2010**, *6* (2),  
1262 165–179. <https://doi.org/10.1007/s11306-010-0208-9>.
- 1263 (116) Lane, A. N.; Higashi, R. M.; Fan, T. W.-M. Preclinical Models for Interrogating Drug Action in  
1264 Human Cancers Using Stable Isotope Resolved Metabolomics (SIRM). *Metabolomics* **2016**, *12*  
1265 (7), 118. <https://doi.org/10.1007/s11306-016-1065-y>.
- 1266 (117) Fan, T. W.-M.; Lorkiewicz, P.; Sellers, K.; Moseley, H. N. B.; Higashi, R. M.; Lane, A. N. Stable  
1267 Isotope-Resolved Metabolomics and Applications for Drug Development. *Pharmacol. Ther.*  
1268 **2012**, *133* (3), 366–391. <https://doi.org/10.1016/j.pharmthera.2011.12.007>.
- 1269 (118) Lane, A. N.; Fan, T. W.-M.; Bousamra, M.; Higashi, R. M.; Yan, J.; Miller, D. M. Stable Isotope-  
1270 Resolved Metabolomics (SIRM) in Cancer Research with Clinical Application to NonSmall Cell  
1271 Lung Cancer. *OMICS J. Integr. Biol.* **2011**, *15* (3), 173–182.  
1272 <https://doi.org/10.1089/omi.2010.0088>.
- 1273 (119) Lane, A. N.; Higashi, R. M.; Fan, T. W.-M. NMR and MS-Based Stable Isotope-Resolved  
1274 Metabolomics and Applications in Cancer Metabolism. *TrAC Trends Anal. Chem.* **2019**, *120*,  
1275 115322. <https://doi.org/10.1016/j.trac.2018.11.020>.
- 1276 (120) Fan, T. W.; Lane, A. N.; Higashi, R. M.; Farag, M. A.; Gao, H.; Bousamra, M.; Miller, D. M. Altered  
1277 Regulation of Metabolic Pathways in Human Lung Cancer Discerned by <sup>13</sup>C Stable Isotope-  
1278 Resolved Metabolomics (SIRM). *Mol. Cancer* **2009**, *8* (1), 1–19. <https://doi.org/10.1186/1476-4598-8-41>.
- 1280 (121) Higashi, R. M.; Fan, T. W.-M.; Lorkiewicz, P. K.; Moseley, H. N. B.; Lane, A. N. Stable Isotope-  
1281 Labeled Tracers for Metabolic Pathway Elucidation by GC-MS and FT-MS. *Mass Spectrom.*  
1282 *Metabolomics* **2014**, 147–167. [https://doi.org/10.1007/978-1-4939-1258-2\\_11](https://doi.org/10.1007/978-1-4939-1258-2_11).
- 1283 (122) Giraudeau, P.; Massou, S.; Robin, Y.; Cahoreau, E.; Portais, J.-C.; Akoka, S. Ultrafast  
1284 Quantitative 2D NMR: An Efficient Tool for the Measurement of Specific Isotopic Enrichments  
1285 in Complex Biological Mixtures. *Anal. Chem.* **2011**, *83* (8), 3112–3119.  
1286 <https://doi.org/10.1021/ac200007p>.
- 1287 (123) Reardon, P. N.; Marean-Reardon, C. L.; Bukovec, M. A.; Coggins, B. E.; Isern, N. G. 3D TOCSY-  
1288 HSQC NMR for Metabolic Flux Analysis Using Non-Uniform Sampling. *Anal. Chem.* **2016**, *88* (5),  
1289 2825–2831. <https://doi.org/10.1021/acs.analchem.5b04535>.
- 1290 (124) Sinnaeve, D.; Dinclaux, M.; Cahoreau, E.; Millard, P.; Portais, J.-C.; Létisse, F.; Lippens, G.  
1291 Improved Isotopic Profiling by Pure Shift Heteronuclear 2D J-Resolved NMR Spectroscopy.  
1292 *Anal. Chem.* **2018**, *90* (6), 4025–4031. <https://doi.org/10.1021/acs.analchem.7b05206>.
- 1293 (125) Lane, A. N.; Fan, T. W.-M.; Xie, Z.; Moseley, H. N. B.; Higashi, R. M. Isotopomer Analysis of Lipid  
1294 Biosynthesis by High Resolution Mass Spectrometry and NMR. *Anal. Chim. Acta* **2009**, *651* (2),  
1295 201–208. <https://doi.org/10.1016/j.jaca.2009.08.032>.
- 1296 (126) Castaño-Cerezo, S.; Kulyk-Barbier, H.; Millard, P.; Portais, J.-C.; Heux, S.; Truan, G.; Bellvert, F.  
1297 Functional Analysis of Isoprenoid Precursors Biosynthesis by Quantitative Metabolomics and  
1298 Isotopologue Profiling. *Metabolomics* **2019**, *15* (9). <https://doi.org/10.1007/s11306-019-1580-8>.
- 1299



- 1300 (127) Heux, S.; Bergès, C.; Millard, P.; Portais, J.-C.; Létisse, F. Recent Advances in High-Throughput  
1301 <sup>13C</sup>-Fluxomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 104–109.  
1302 <https://doi.org/10.1016/j.copbio.2016.10.010>.
- 1303 (128) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt,  
1304 D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.;  
1305 Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.;  
1306 Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; Macinnis, G. D.;  
1307 Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B.  
1308 D.; Vogel, H. J.; Querengesser, L. HMDB: The Human Metabolome Database. *Nucleic Acids Res.*  
1309 **2007**, *35* (Database issue), D521–526. <https://doi.org/10.1093/nar/gkl923>.
- 1310 (129) Smith, C. A.; O’Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.;  
1311 Abagyan, R.; Siuzdak, G. METLIN: A Metabolite Mass Spectral Database. *Ther. Drug Monit.*  
1312 **2005**, *27* (6), 747–751. <https://doi.org/10.1097/01.ftd.0000179845.53213.39>.
- 1313 (130) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading,  
1314 S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.;  
1315 Markley, J. L. BioMagResBank. *Nucleic Acids Res.* **2008**, *36* (Database issue), D402–408.  
1316 <https://doi.org/10.1093/nar/gkm957>.
- 1317 (131) Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. LIPID MAPS Online Tools for Lipid Research.  
1318 *Nucleic Acids Res.* **2007**, *35* (Web Server issue), W606–612.  
1319 <https://doi.org/10.1093/nar/gkm324>.
- 1320 (132) Jewison, T.; Knox, C.; Neveu, V.; Djoumbou, Y.; Guo, A. C.; Lee, J.; Liu, P.; Mandal, R.;  
1321 Krishnamurthy, R.; Sinelnikov, I.; Wilson, M.; Wishart, D. S. YMDB: The Yeast Metabolome  
1322 Database. *Nucleic Acids Res.* **2012**, *40* (Database issue), D815–D820.  
1323 <https://doi.org/10.1093/nar/gkr916>.
- 1324 (133) Salek, R. M.; Haug, K.; Conesa, P.; Hastings, J.; Williams, M.; Mahendraker, T.; Maguire, E.;  
1325 González-Beltrán, A. N.; Rocca-Serra, P.; Sansone, S.-A.; Steinbeck, C. The MetaboLights  
1326 Repository: Curation Challenges in Metabolomics. *Database J. Biol. Databases Curation* **2013**,  
1327 *2013*, bat029. <https://doi.org/10.1093/database/bat029>.
- 1328 (134) Haug, K.; Salek, R. M.; Conesa, P.; Hastings, J.; de Matos, P.; Rijnbeek, M.; Mahendraker, T.;  
1329 Williams, M.; Neumann, S.; Rocca-Serra, P.; Maguire, E.; González-Beltrán, A.; Sansone, S.-A.;  
1330 Griffin, J. L.; Steinbeck, C. MetaboLights—an Open-Access General-Purpose Repository for  
1331 Metabolomics Studies and Associated Meta-Data. *Nucleic Acids Res.* **2013**, *41* (D1), D781–  
1332 D786. <https://doi.org/10.1093/nar/gks1004>.
- 1333 (135) Haug, K.; Cochrane, K.; Nainala, V. C.; Williams, M.; Chang, J.; Jayaseelan, K. V.; O’Donovan, C.  
1334 MetaboLights: A Resource Evolving in Response to the Needs of Its Scientific Community.  
1335 *Nucleic Acids Res.* **2020**, *48* (D1), D440–D444. <https://doi.org/10.1093/nar/gkz1019>.
- 1336 (136) Salek, R. M.; Neumann, S.; Schober, D.; Hummel, J.; Billiau, K.; Kopka, J.; Correa, E.; Reijmers,  
1337 T.; Rosato, A.; Tenori, L.; Turano, P.; Marin, S.; Deborde, C.; Jacob, D.; Rolin, D.; Dartigues, B.;  
1338 Conesa, P.; Haug, K.; Rocca-Serra, P.; O’Hagan, S.; Hao, J.; van Vliet, M.; Sysi-Aho, M.; Ludwig,  
1339 C.; Bouwman, J.; Cascante, M.; Ebbels, T.; Griffin, J. L.; Moing, A.; Nikolski, M.; Oresic, M.;  
1340 Sansone, S.-A.; Viant, M. R.; Goodacre, R.; Günther, U. L.; Hankemeier, T.; Luchinat, C.;  
1341 Walther, D.; Steinbeck, C. COordination of Standards in MetabOlogicS (COSMOS): Facilitating  
1342 Integrated Metabolomics Data Access. *Metabolomics* **2015**, *11* (6), 1587–1597.  
1343 <https://doi.org/10.1007/s11306-015-0810-y>.
- 1344 (137) Rocca-Serra, P.; Salek, R. M.; Arita, M.; Correa, E.; Dayalan, S.; Gonzalez-Beltran, A.; Ebbels, T.;  
1345 Goodacre, R.; Hastings, J.; Haug, K.; Koulman, A.; Nikolski, M.; Oresic, M.; Sansone, S.-A.;  
1346 Schober, D.; Smith, J.; Steinbeck, C.; Viant, M. R.; Neumann, S. Data Standards Can Boost  
1347 Metabolomics Research, and If There Is a Will, There Is a Way. *Metabolomics* **2016**, *12*.  
1348 <https://doi.org/10.1007/s11306-015-0879-3>.
- 1349 (138) Zimmermann, M.; Zimmermann-Kogadeeva, M.; Wegmann, R.; Goodman, A. L. Mapping  
1350 Human Microbiome Drug Metabolism by Gut Bacteria and Their Genes. *Nature* **2019**, *570*  
1351 (7762), 462–467. <https://doi.org/10.1038/s41586-019-1291-3>.

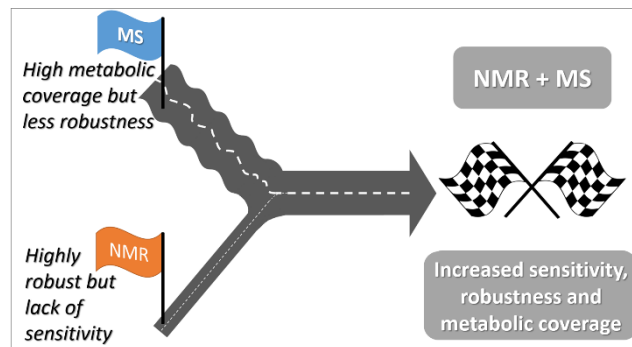
- 1352 (139) Manzoni, C.; Kia, D. A.; Vandrovцова, J.; Hardy, J.; Wood, N. W.; Lewis, P. A.; Ferrari, R.  
1353 Genome, Transcriptome and Proteome: The Rise of Omics Data and Their Integration in  
1354 Biomedical Sciences. *Brief. Bioinform.* **2018**, *19* (2), 286–302.  
1355 <https://doi.org/10.1093/bib/bbw114>.
- 1356 (140) Rohart, F.; Gautier, B.; Singh, A.; Cao, K.-A. L. MixOmics: An R Package for ‘omics Feature  
1357 Selection and Multiple Data Integration. *PLOS Comput. Biol.* **2017**, *13* (11), e1005752.  
1358 <https://doi.org/10.1371/journal.pcbi.1005752>.
- 1359 (141) Yao, X.; Cao, D.; Wang, F.; Zhang, W.; Ma, C.; Song, M. An Overview of Omics Approaches to  
1360 Characterize the Effect of Perfluoroalkyl Substances in Environmental Health. *TrAC Trends*  
1361 *Anal. Chem.* **2019**, *121*, 115367. <https://doi.org/10.1016/j.trac.2018.12.021>.
- 1362 (142) Brunetti, A. E.; Neto, F. C.; Vera, M. C.; Taboada, C.; Pavarini, D. P.; Bauermeister, A.; Lopes, N.  
1363 P. An Integrative Omics Perspective for the Analysis of Chemical Signals in Ecological  
1364 Interactions. *Chem. Soc. Rev.* **2018**, *47* (5), 1574–1591. <https://doi.org/10.1039/C7CS00368D>.
- 1365 (143) Saito, K.; Hirai, M. Y.; Yonekura-Sakakibara, K. Decoding Genes with Coexpression Networks  
1366 and Metabolomics – ‘Majority Report by Precogs.’ *Trends Plant Sci.* **2008**, *13* (1), 36–43.  
1367 <https://doi.org/10.1016/j.tplants.2007.10.006>.
- 1368 (144) Pascual, L.; Xu, J.; Biais, B.; Maucourt, M.; Ballias, P.; Bernillon, S.; Deborde, C.; Jacob, D.;  
1369 Desgroux, A.; Faurobert, M.; Bouchet, J.-P.; Gibon, Y.; Moing, A.; Causse, M. Deciphering  
1370 Genetic Diversity and Inheritance of Tomato Fruit Weight and Composition through a Systems  
1371 Biology Approach. *J. Exp. Bot.* **2013**, *64* (18), 5737–5752. <https://doi.org/10.1093/jxb/ert349>.
- 1372 (145) Blanchet, L.; Smolinska, A.; Attali, A.; Stoop, M. P.; Ampt, K. A.; van Aken, H.; Suidgeest, E.;  
1373 Tuinstra, T.; Wijmenga, S. S.; Luider, T.; Buydens, L. M. Fusion of Metabolomics and Proteomics  
1374 Data for Biomarkers Discovery: Case Study on the Experimental Autoimmune  
1375 Encephalomyelitis. *BMC Bioinformatics* **2011**, *12* (1), 254. [https://doi.org/10.1186/1471-2105-](https://doi.org/10.1186/1471-2105-12-254)  
1376 [12-254](https://doi.org/10.1186/1471-2105-12-254).
- 1377 (146) González-Ruiz, V.; Schvartz, D.; Sandström, J.; Pezzatti, J.; Jeanneret, F.; Tonoli, D.; Boccard, J.;  
1378 Monnet-Tschudi, F.; Sanchez, J.-C.; Rudaz, S. An Integrative Multi-Omics Workflow to Address  
1379 Multifactorial Toxicology Experiments. *Metabolites* **2019**, *9* (4), 79.  
1380 <https://doi.org/10.3390/metabo9040079>.
- 1381 (147) Wild, C. P. The Exposome: From Concept to Utility. *Int. J. Epidemiol.* **2012**, *41* (1), 24–32.  
1382 <https://doi.org/10.1093/ije/dyr236>.
- 1383

1384

1385

For table of contents only

1386



1387