Journal of Chromatography A, 1216 (2009) 7314–7323

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Metabolomics of transgenic maize combining Fourier transform-ion cyclotron resonance-mass spectrometry, capillary electrophoresis-mass spectrometry and pressurized liquid extraction

Carlos Leon[a], Irene Rodriguez-Meizoso[a], Marianna Lucio[b], Virginia Garcia-Cañas[a], Elena Ibañez[a], Philippe Schmitt-Kopplin[b,*], Alejandro Cifuentes[a,**]

[a] Institute of Industrial Fermentations (CSIC), Juan de la Cierva 3, Madrid, Spain
[b] Institut für Ökologische Chemie (Helmholtz Zentrum Muenchen), Ingolstaedter Landstrasse 1, Neuherberg, Germany

ARTICLE INFO

Article history:
Available online 3 May 2009

Keywords:
Metabolomics
Corn
Capillary electrophoresis-mass spectrometry (CE-MS)
FT-MS
Transgenic foods
GMOs
Bt maize
Metabolites

ABSTRACT

In this work, the potential of combining capillary electrophoresis-time-of-flight-mass spectrometry (CE-TOF-MS) and Fourier transform-ion cyclotron resonance-mass spectrometry (FT-ICR-MS) for metabolomics of genetically modified organisms (GMOs) is demonstrated. Thus, six different varieties of maize, three of them transgenic (PR33P66 Bt, Tietar Bt and Aristis Bt) and their corresponding isogenic lines (PR33P66, Tietar and Aristis) grown under the same field conditions, were analyzed. Based on the ultrahigh resolution and remarkable mass accuracy provided by the 12-T FT-ICR-MS it was possible to directly analyze a good number of metabolites whose identity could be proposed based on their specific isotopic pattern. For identification of metabolite isomers, CE-TOF-MS was also used combining the information on nominal mass with electrophoretic mobility corroborating in that way the identity of several new biomarkers. Furthermore, PLE extractions were evaluated in order to establish selective extraction as an additional criterion to obtain useful information in maize metabolomics. Differences in the metabolite levels were found between the three transgenic maize varieties compared with their wild isogenic lines in some specific metabolic pathways. To our knowledge, this is the first time that an approach as the one presented in this work (pressurized liquid extraction + FT-ICR-MS + CE-TOF-MS) is shown for a metabolomic study.

1. Introduction

Genetically modified organisms (GMOs) have brought about a great impact in agriculture and food industry. The most common genetic modifications implemented in plants so far are applied to improve resistance of crops to plagues or pesticides, while some of the newest modifications are being designed to provide better nutritional properties [1,2].

The most common example of genetically induced resistance to plagues in crops (mainly corn) is the inclusion of the *Cry1Ab* gene from *Bacillus thuringiensis* (Bt), an aerobic spore-forming bacterium commonly found in the environment. This gene regulates the synthesis of a Bt protein that acts as an insect toxin in the genetically modified plant [3,4]. Namely, the *Cry1Ab* gene is the prototype for the gene currently expressed in most commercial Bt maize events for the control of the European corn borer (*Ostrinia nubilalis*), a major insect pest of maize in North American and European agriculture [5,6]. These advantages over conventional pesticides have induced a rise in GM-crops expressing Bt toxins [7,8].

However, the use of GMOs in foods is facing numerous criticisms from consumers and ecological organizations that have led some countries to regulate the production, growth and commercialization of GMOs [9,10]. Moreover, investigations on the chemical composition of some GMOs have shown unexpected modifications in this type of organisms [11]. This corroborates the need of new and more powerful analytical developments able to face the complexity of this problem. The development of more powerful analytical tools is also highly required in order to address the forthcoming second generation of GMOs, in which significant changes in metabolites such as polyphenols, vitamins, fatty acids or amino acids will be introduced [12–16].

The development of metabolomics provides the necessary information about metabolite levels allowing the evaluation of their changes in response to different perturbations in the investigated organism [17]. This perturbation can be due to, e.g., pharmaceuticals [18,19], environmental stressors [20], diet related changes [21], genetic modifications [22,23] or food contaminants [24].

* Corresponding author. Fax: +49 089 31873358.
** Corresponding author. Fax: +34 91 5644853.
E-mail addresses: schmitt-kopplin@helmholtz-muenchen.de (P. Schmitt-Kopplin), acifuentes@ifi.csic.es (A. Cifuentes).

0021-9673/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2009.04.092

The main analytical tools used for metabolomics are so far based on NMR and MS techniques [25]. Due to sensitivity limitations in NMR, the application of MS in this field is increasingly growing [26]. It may be used either combined with liquid chromatography [27], gas chromatography [28] and capillary electrophoresis [28–30], or as a stand-alone technique. In this regard, the use of high-field Fourier transform-ion cyclotron resonance-mass spectrometry (FT-ICR-MS) enables the assignment of thousands of elemental compositions of metabolites in complex mixtures by virtue of its unique mass accuracy (<100 ppb) and resolution (>500,000 in full scan mode at m/z 500) allowing to detect fine-scale diversity in complex mixtures [31]. These properties allow the correspondence between accurate masses to elementary compositions and annotation to potential metabolites giving as a result, e.g., the generation of specific metabolic maps for each organism [32] or the detection and structure confirmation of specific biomarkers [33,34]. The great applications of MS and in particular FT-ICR-MS to metabolomics have been reviewed elsewhere [35–37].

Metabolomic studies have also been possible thanks to the great advance in statistical and data evaluation tools [38,39]. However, still some important challenges remain unsolved or need to be improved. For instance, unequivocal identification of biomarkers is yet one of the main analytical challenges as metabolite databases are still limited to the core metabolome, and information achieved by these techniques is not enough to distinguish isomers [26]. This problem can partially be overcome by gathering as much information as possible in the elucidation of the biomarker, for example, through the optimization of data evaluation [40–42] or the use of separation techniques (chromatographic or electro-driven) coupled online or atline to MS. In this way, complementary information can be achieved providing a more complete picture on the possible biomarkers. In this regard, the potential of capillary electrophoresis-mass spectrometry (CE-MS) coupling has already been demonstrated in several metabolomics studies [29,43,44]. The combination of different analytical technologies will be, therefore, highly desired in order to identify unknown metabolites in biofluids and natural samples allowing the detection of biomarkers.

Pressurized liquid extraction (PLE) is one of the so-called green technologies [45]. PLE combines high pressure and high temperature in order to modify the properties of the extraction solvents (boiling point, solubility, polarity, dielectric constant, etc.) allowing the selection of the type of extracted metabolites according to their polarity. Temperature is a critical parameter in PLE since it can modify the dielectric constant of the solvent and promote mass transfer rate while the high pressure allows keeping the solvent at the liquid state. The small solvent volume required in PLE makes this technique highly applicable to many analytical approaches including metabolomics, since it can provide representative extracts of samples in an automatic way while facilitates the subsequent compound identification based on the polarity of the solvent used as extractant.

The goal of this work is to demonstrate the potential of combining FT-ICR-MS, CE-MS and PLE for GMOs metabolomics. Thus, FT-ICR-MS results are confirmed (and complemented) by CE-MS data helping to identify new biomarkers. Moreover, PLE is used for sample extraction and fractionation, prior to FT-ICR-MS analysis, in order to include polarity of compounds as additional selectivity criterion for metabolomics.

2. Experimental

2.1. Chemicals

All chemicals were of analytical reagent grade and used as received. Methanol from Scharlau (Barcelona, Spain), hexane from Panreac Química (Barcelona, Spain) and water were used as extraction agents for maize. Water was deionized by using a Milli-Q system from Millipore (Bedford, MA, USA). The extracts were stored at 4 °C and warmed at room temperature before used. LC–MS grade methanol from Riedel-de Häen (Seelze, Germany), and deionized water were used as solvents for FT-ICR-MS injection. Formic acid from Riedel-de Haën (Seelze, Germany) and 2-propanol from Scharlau (Barcelona, Spain) were used for the CE buffers and sheath liquid respectively.

2.2. Maize samples

The investigated varieties of conventional and transgenic maize were obtained from a field assay carried out in Estación Experimental Agrícola Mas Badía in Tallada d'Empordá (Girona, Spain) using commercial varieties. Namely, in order to skip any influence from the growing conditions, Aristis maize (wild type and its Bt transgenic variety), Tietar maize (wild type and its Bt transgenic variety), and PR33P66 maize (wild type and its Bt transgenic variety) were grown at exactly the same time under the same field conditions. In all cases, the maize genetic modification consisted of introducing a new Cry-type gene able to synthesize a new protein (protein Bt) that acts as insecticide obtaining a maize resistant to some worm plagues. Each sample was a mixture of 8–12 kernels grained all together in order to minimize phenotypic variegation.

The transgenic and non-transgenic nature of all these maize samples was confirmed based on their DNA using an analytical procedure developed in our laboratory and described elsewhere [46–50].

2.3. PLE procedure

Three different solvents (water, methanol and hexane) and two extraction temperatures (100 and 175 °C) were applied for 20 min at 1500 psi for the extraction of metabolites from transgenic and conventional maize using PLE. Maize kernels were freshly grounded and stored in the fridge at 4 °C. Next, 1 g of sample was weighed and extracted by PLE using a Dionex (Sunnyvale, CA, USA) ASE 200 system. The cell volume was 11 ml, with a flush volume of 60%.

2.4. Ultrasonic extraction procedure

A metabolite extraction procedure using ultrasonic equipment based on a previous work from our group [44] was used for comparison. Namely, 2 g of sample was weighed and extracted with 10 ml of methanol:water (50:50) in ultrasonic bath for 10 min. After sonication, samples were centrifuged for 5 min at 3000 rpm and liquid phases were filtered through a 0.45 μm filter. Liquid phases were taken to dryness in a rotatory evaporator and redissolved in 0.5 ml of methanol.

2.5. FT-ICR-MS conditions

Samples were diluted 1:20 in methanol before flow injection in the FT-ICR-MS. High resolution mass spectra were carried out on a Bruker (Bremen, Germany) APEX Qe Fourier transform-ion cyclotron resonance-MS equipped with a 12 T superconducting magnet and an Apollo II electrospray (ESI) source operated with 1000 scan (1 MW) in the positive ion mode and 500 scan (4 MW) in the negative one. Scan range was 150–2000 m/z for both modes. Spectra were externally calibrated on clusters of arginine (10 mg/l in methanol) and internally systematically on fatty acids (ESI(−)) and solvent diesters (ESI(+)). Samples were flow-injected straight forward in the ESI using a Hamilton 250 μl syringe with a simple syringe injection pump at 0.12 ml/h.

2.6. CE-TOF-MS analysis

Analysis were carried out with a P/ACE 5010 CE apparatus from Beckman Instruments coupled to a time-of-flight micrOTOF MS detector from Bruker Daltonik (Bremen, Germany) using an orthogonal electrospray interface (ESI, model G1607A from Agilent Technologies, Palo Alto, CA, USA). CE-TOF-MS conditions applied in this part of the work were similar to those already described elsewhere [44]. Briefly, an uncoated fused-silica capillary with 50 μm I.D., 375 μm O.D. and 80 cm of total (and detection) length was used. CE buffer was 5% formic acid in water (pH 1.90), and other conditions were run voltage of 20 kV, samples were injected in triplicate hydrodynamically for 15 s at 0.5 psi. The nebulizing gas and voltage were stopped during the washing procedure. Electrical contact at the electrospray needle tip was established via a sheath liquid that consisted of isopropanol–water (50:50, v/v) delivered at a flow rate of 0.24 ml/h by a 74900-00-05 Cole Parmer syringe pump (Vernon Hills, IL, USA). The nebulizer and dry gas conditions were 0.4 bar N_2 and 4 l/min N_2 at 180 °C. For the metabolite search, DataAnalysis 3.4 from Bruker Daltonics was used. The comparison between CE-TOF-MS and FT-ICR-MS results was done using the 150–450 m/z range in ESI(+) ionization mode.

2.7. Data evaluation

FT-ICR-MS spectra were exported after calibration into peak lists at a signal-to-noise ratio (S/N) = 2. The preliminary structure of the files is assigned through software developed in Python that provides the binning at 1 ppm of all the spectra. The statistical analyses were done with SIMCA-P 11.5 (Umetrics, Umea, Sweden) and SAS version 9.1 (SAS Institute, Cary, NC, USA) in order to summarize and visualize the main properties of the samples. From the statistical models it is possible to extrapolate lists of masses characteristic for the different groups under study. From these lists the elemental formulas were assigned for each peak in batch mode by a software tool written in-house in Python and FORTRAN 95. The generated formulas were validated by setting sensible chemical constraints (nitrogen rule, atomic oxygen to carbon ratio O/C < 1, element counts: hydrogen H< (2 + C2), carbon C ≤ 100, oxygen O ≤ 80, nitrogen N ≤ 5 and sulfur S ≤ 1) and only the $^{12}C_n$ masses in conjunction with their automated generated theoretical isotope patterns (existence of the $^{12}C_{(n-1)}$ isotope) were taken into consideration for the van Krevelen data visualization [33,51]. In a different data evaluation approach the raw mass data were uploaded in the Mass Translator into Pathways [32] developed at Helmholtz Zentrum Muenchen. The MassTRIX approach enables a direct confrontation of the exact mass data with the elementary composition space of the KEGG database and the resulting hits are only based on identical elementary compositions and thus do not differentiate isomers. Despite the restrictions explained above, this approach provides a first screening of the FT-ICR-MS data.

3. Results and discussion

3.1. FT-ICR-MS for metabolomics of transgenic maize

The objective of these FT-ICR-MS analyses was to describe in a non-targeted way the metabolome as detected using ESI(−) and ESI(+) analyses of the ultrasonic extracts of maize, taking advantage of the high resolution, accuracy and sensitivity provided by the 12-T FT-ICR-MS instrument. To do this, maize samples were firstly extracted using an ultrasonic extraction procedure optimized previously [44]. Finely ground kernels of six different varieties of maize, three of them transgenic (PR33P66 Bt, Tietar Bt and Aristis Bt) and their corresponding isogenic lines (PR33P66, Tietar and Aristis) grown under the same field conditions were extracted and analyzed by FT-ICR-MS. Since the direct analysis of the extracts by FT-ICR-MS provided saturation of the MS signal, three different dilutions of these extracts in methanol (1–20, 1–200 and 1–1000) were tried before direct injection of the samples, giving the 1–20 dilution the best results. Therefore, all the samples were diluted 20 times before FT-ICR-MS analysis. Methanol samples (blank) were firstly injected in order to confirm that different ions appeared in the extracts. The diluted extracts were directly infused in the FT-ICR-MS in both the positive and the negative ESI mode. Analysis of these samples resulted in average in 5500 mass signals depending on the sample and the ionization mode, from which less than 1000 signals fulfilled the above mentioned validation constraints for the assignment of elemental compositions. The raw data were uploaded into MassTRIX [32], a server that aims to annotate high precision mass spectrometry data and to display the results on *Zea mays* specific annotated metabolites in the Kyoto Encyclopedia Gene and Genome (KEGG) and their related pathway maps. The number of compounds definitely assigned to a valid elemental composition found in different metabolic pathways of *Z. mays* for each sample in both ionization modes can be seen in Table 1. From these data it can be observed that in general more compounds were detected in the genetically modified (GM) varieties compared to their conventional isogenic lines. These differences were investigated using a multivariate approach, moreover, considering that the classification of samples based on their metabolic profile is one of the main issues in this type of research. The classical method used for this purpose is the partial least squares-discriminant analysis (PLS-DA). It is a multivariate method used to classify and it is suitable when the number of experiments (in this occasion spectra) is small compared to the amount of variables (m/z) and when it is present multicollinearity.

Table 1
Number of compounds from the metabolic pathways of *Z. mays* identified by FT-ICR-MS in the six varieties of maize (3 GMOs and their respective non-trasgenic isogenic lines). All samples were analyzed by FT-ICR-MS operated with 1000 scan (1 MW) in the positive mode and with 500 scan (4 MW) in the negative mode. KEGG database, available through the Mass Translator into Pathways, MassTRIX software [32] was used for the determination of the compounds.

	Positive FT-ICR-MS	Negative FT-ICR-MS
PR33P66	153	307
PR33P66 Bt	156	513
TIETAR	232	297
TIETAR Bt	111	496
ARISTIS	143	429
ARISTIS Bt	221	482

As can be seen in Fig. 1, the metabolomes can be differentiated in a two component model with a good level of prediction (PLS-DA model with $R^2(Y) = 0.99$ and $Q^2(\text{cum}) = 0.52$). This supervised analysis reflects the natural grouping that was possible to see also with a principal component analysis (PCA) (data not shown). The dimensionality of the original data space was reduced through the calculus of these principal components in order to get an overview of the variation, so that groups, trends and outliers can be identified among the observations. The reason why such a compression was possible is that variables are correlated with each other. Due to the low number of samples, however, this method could not be externally validated. The score scatter plot (Fig. 1A) showed a considerable separation between the two groups, the blue one for the GMO spectra and red for the isogenic lines demonstrating the good classification capability of this approach.

The loadings were constructed by projecting the variables (m/z) on to these new axes. The whole points, in Fig. 1B, are all the loading points. Mathematically they describe the correlation between variables (m/z values) and also between the scores. The extreme

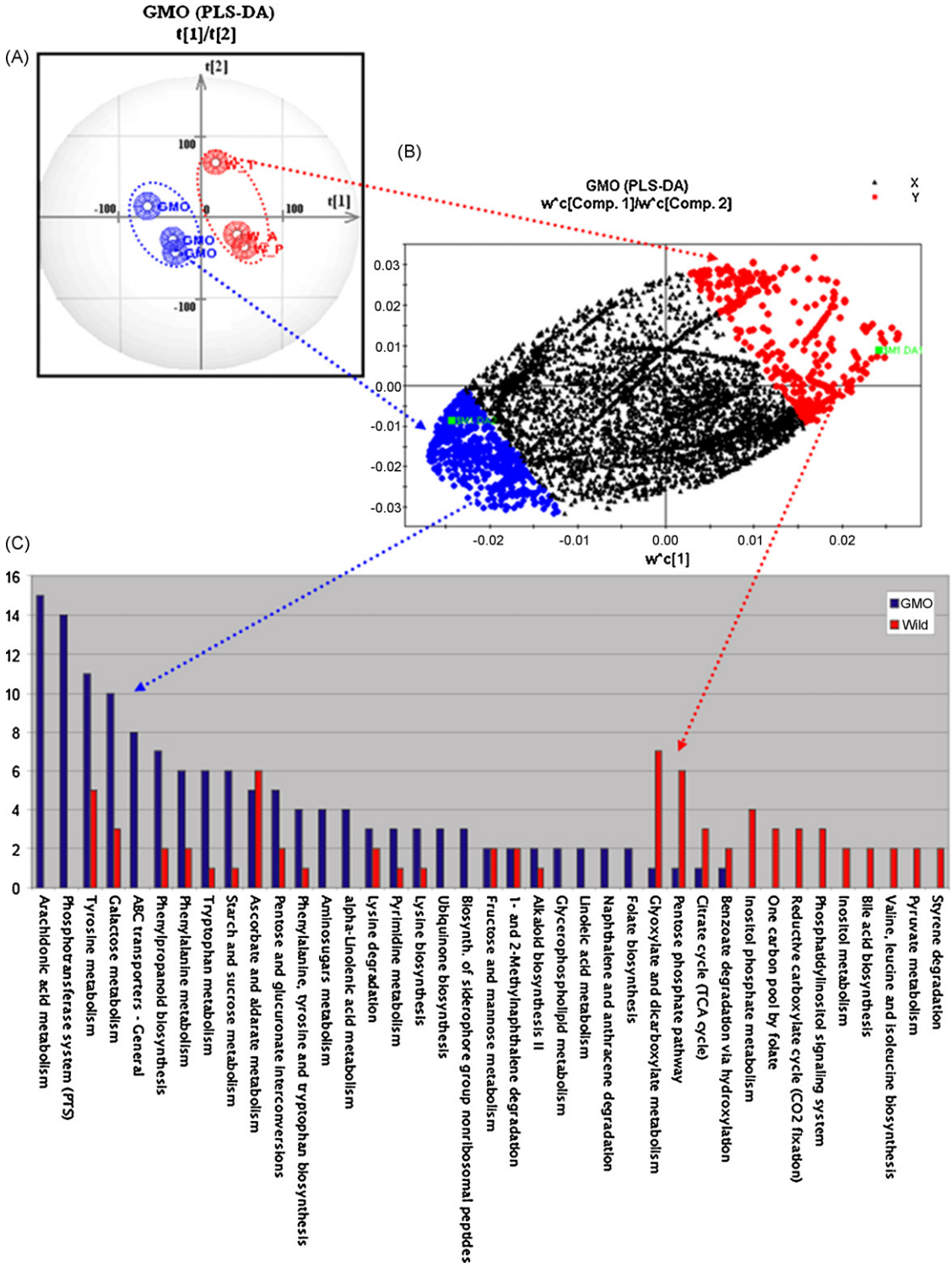


Fig. 1. PLS-DA model (Q^2(cum) = 0.52 and $R^2(Y)$ = 0.99) with six different maize varieties analyzed by FT-ICR-MS. Samples (A) PR33P66; (B) PR33P66 Bt; (C) Tietar; (D) Tietar Bt; (E) Aristis; and (F) Aristis Bt. See text for other experimental conditions. The score scatter plot underlines a different pattern for the transgenic (they are represented in blue color) and isogenic lines (red color). The different properties of the discriminative masses (represented in blue and red in the loading plot) are investigated with MassTRIX. The model was built up with the data measured in negative mode.

groups, colored in blue and red (Fig. 1B), are the loadings most correlated with the two different groups (blue for the GMOs and red for the wild samples). These loadings were the most discriminative masses that contribute to differentiate the two groups. The evaluation of the importance was done based on the regression coefficient values. The two lists of discriminant m/z were uploaded in MassTRIX showing the total number of compounds identified and present in *Z. mays* (Fig. 1C).

The blue bars and the red ones in Fig. 1C are the number of isomeric hits within a pathway different in the GMO group and isogenic lines. Mainly compounds from arachidonic acid metabolism, phosphotransferase system (PTS), tyrosine and galac-

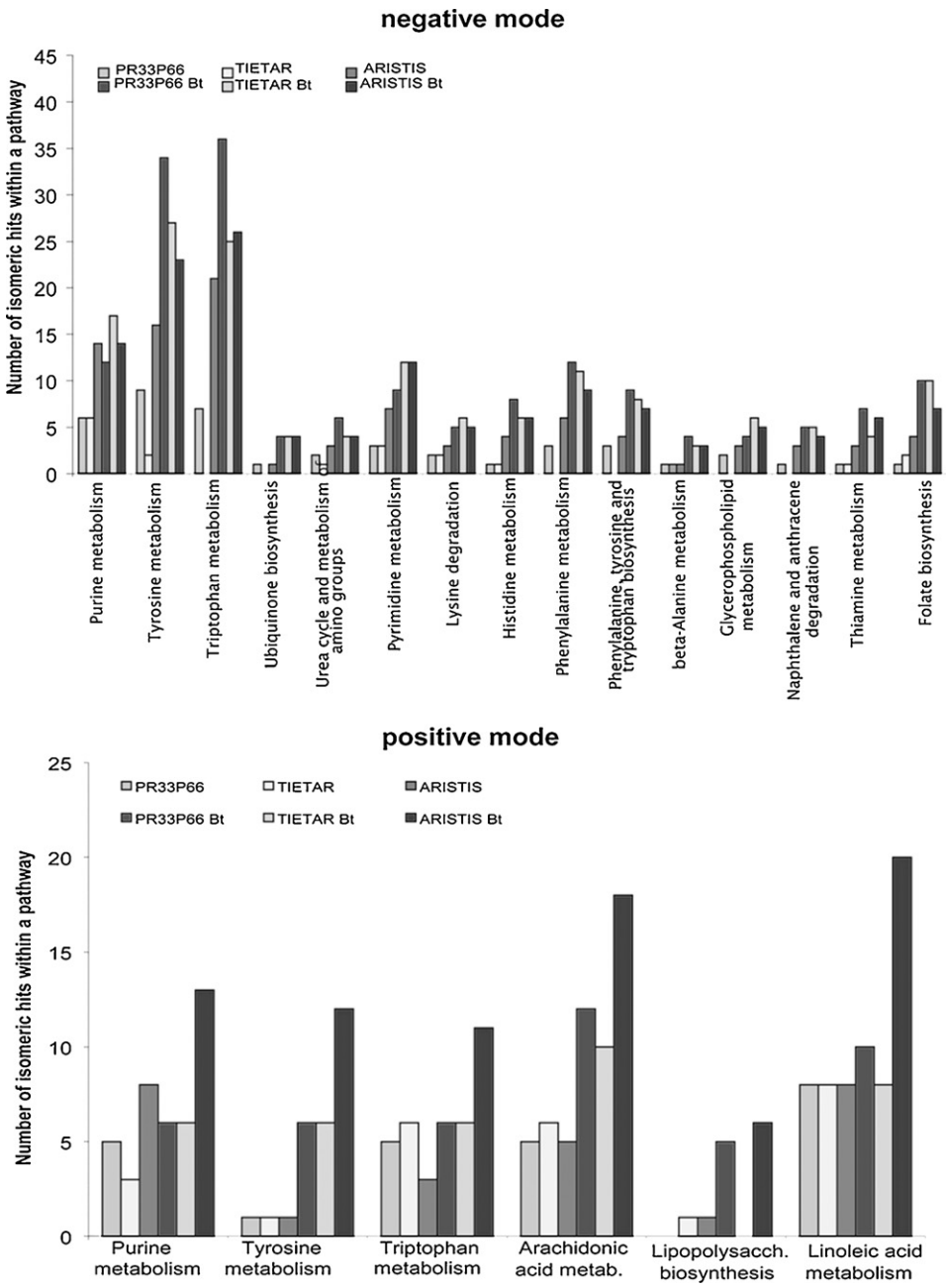


Fig. 2. Number of compounds identified in some metabolic pathways in six different varieties of maize, three transgenic (PR33P66 Bt, Tietar Bt and Aristis Bt) and their corresponding isogenic non-transgenic lines (PR33P66, Tietar and Aristis) by FT-ICR-MS using both positive and negative mode. Positive ion mode was operated at 1000 scan (1 MW) and negative one at 500 scan (4 MW). Scan range: 150–2000 *m*/*z*.

tose metabolism characterized GMO group (see the bars-graphic in Fig. 1C). In particular in all the list of masses selected, as representative for isogenic lines group, none was assigned to the arachidonic acid metabolism or phosphotransferase system (PTS) as can be deduced from the red bars in Fig. 1C.

As mentioned, except for Tietar in Table 1 in the positive mode, in general a higher number of compounds were detected in the GM samples compared to their conventional isogenic lines showing an over-expression in some metabolic pathways of the transgenic varieties compared to the wild ones. These results are in accordance with those found in the literature, highlighting differences in the metabolism of GMOs with respect to their parental wild lines [11,43,44].

Interestingly, most of these variations are due to specific metabolic pathways, with a higher number of metabolites found, particularly in the GM-varieties. In Fig. 2, the number of compounds identified in some of these particular pathways derived from the information obtained in the MassTRIX are indicated (wild samples correspond to the first three bars of each metabolic pathway while GM samples correspond to the last three bars). As can be seen, the amino acid metabolic pathways (or amino acid related) have clearly a bigger number of metabolites, as for example tyrosine and tryptophan, detected in both ionization modes, or the lysine, phenylalanine, histidine or β-alanine metabolic pathways in the negative mode. Additionally, it is shown that other metabolic pathways, such as the purine metabolism or the folate biosynthesis are clearly altered.

Based on the mentioned FT-ICR-MS results, differences between the metabolic profiles of transgenic and wild maize can be highlighted and elementary compositions of specific GMO biomarkers

Table 2
Comparison among the experimental relative molecular masses (M_r) obtained by CE-TOF-MS, the experimental M_r values obtained by FT-ICR-MS and the theoretical M_r values of tentatively identified compounds. Experimental conditions: FT-ICR-MS operated with 1000 scan (1 MW) in the positive mode. CE buffer was 5% formic acid in water (pH 1.90), other conditions were run voltage of 20 kV, samples injected hydrodynamically for 15 s at 0.5 psi. Sheath liquid in the CE-TOF-MS consisted of isopropanol–water (50:50, v/v).

Compound number	Formula	Theoretical m/z	FT-ICR-MS M_r	FT-ICR-MS error (mDa)	FT-ICR-MS error (ppm)	CE-TOF-MS M_r	CE-TOF-MS error (mDa)	CE-TOF-MS error (ppm)	Tentative compound
1	$[C_7H_{16}NO_3]^+$	162.1124	162.1126	0.2	0.12	162.106	6.4	3.9	L-Carnitine
2	$[C_6H_{15}N_4O_2]^+$	175.1189	175.1191	0.2	0.11	175.112	6.9	3.9	Arginine
3	$[C_9H_{12}NO_3]^+$	182.0811	182.0804	0.7	0.38	182.073	8.1	4.4	Tyrosine
4	$[C_9H_{20}N_5O_3]^+$	246.1561	246.1544	1.7	0.69	246.234	77.9	31.6	Beta-Alanyn-L-arginine
5	$[C_{14}H_{21}N_2O_3]^+$	265.1546	265.1550	0.4	0.15	265.100	54.6	20.5	Subaphyllin
6	$[C_{17}H_{14}NO_3]^+$	280.0968	280.0924	4.4	1.57	280.084	12.8	4.6	Graveoline
7	$[C_{18}H_{30}N_3O_2]^+$	320.2332	320.2355	2.3	0.7	320.224	9.2	2.8	
8	$[C_{15}H_{19}O_9]^+$	343.1023	343.1224	19.7	5.74	343.100	2.3	0.6	L-Caffeoyl-beta-D-glucose
9	$[C_{25}H_{32}N_3O_4]^+$	438.2387	438.2394	0.7	0.15	438.232	6.7	1.5	Lunarine
10	$[C_{15}H_{12}O_4]^+$	256.0730	256.0725	0.5	0.2	255.978	95	37.1	Apigeninidin
11	$[C_8H_8NO_2]^+$	150.0549	150.0528	2.1	1.39	150.065	10.1	6.7	5,6-Dihydroxyindole
12	$[C_{20}H_{35}O_4]^+$	339.2529	339.2535	0.6	0.1	339.235	17.9	5.2	

investigated. Special attention was paid to compounds assigned to the metabolic pathways by the MassTRIX software. Following these criteria a list with 33 highly probable compounds could be obtained. When considering all FT-ICR-MS mass data without any filtering by the software, a much higher number of MS signals (hundreds of m/z ions) could be found but without any structure information except their elementary composition.

Moreover, FT-ICR-MS cannot differentiate among structures between isomers and an additional analytical dimension is many times needed to determine/confirm the possible identity of the several biomarkers found having the same molecular formula [26]. CE-TOF-MS was applied here in order to gain additional and complementary information on the metabolic profile.

3.2. FT-ICR-MS and CE-TOF-MS for metabolomics of transgenic maize

The same six varieties of maize as above were subjected to CE-TOF-MS under the conditions specified under Section 2. Information on electrophoretic mobility of the solutes at the separation pH was used as additional criterion to elucidate their nature. In Table 2, the molecular mass values and the tentative compounds identified by both CE-TOF-MS and FT-ICR-MS analysis are compared. Ten compounds could be clearly confirmed combining the electrophoretic mobility as additional criterion together with the mass data provided by FT-ICR-MS and TOF-MS. The structure of these compounds was determined by identification in the KEGG database. However, MS/MS experiments might give more accurate structural information about these compounds. Moreover, their relative molecular mass values determined by CE-TOF-MS and FT-ICR-MS showed a very good agreement as shown in Table 2. However, it has to be clarified that the list of compounds given in Table 2 shows only those compounds that were found by both FT-ICR-MS and CE-TOF-MS. FT-ICR-MS alone detects a much higher number of metabolites but this technique is not able to find directly any structural information when these compounds are not found in databases or not analyzed by CE-TOF-MS. On the other hand, CE-TOF-MS also gives more peaks that cannot be assigned to structures directly. From this list only four compounds (numbers 1, 10, 11 and 12) would correspond to possible biomarkers, i.e., they are found in all the transgenic maizes but not (or at very low concentration) in their respective non-transgenic parental lines. A more detailed explanation on how the identification of these possible biomarkers was carried out is exemplified below.

In Fig. 3 we can see as example the CE-TOF-MS extracted ion electropherogram of one of these tentative biomarkers (compound 10

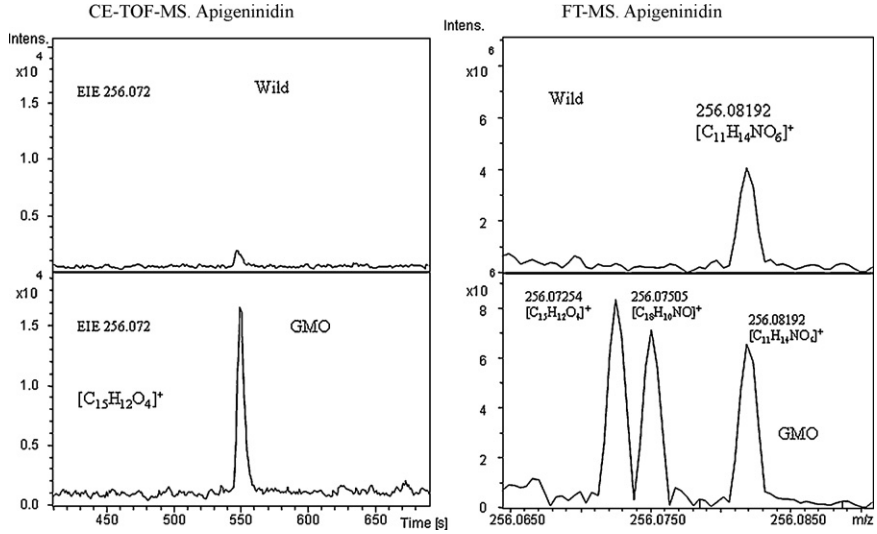


Fig. 3. Comparison between (A) a wild variety of maize (PR33P66) and (B) a transgenic one (PR33P66 Bt) analyzed by CE-TOF-MS (EIE 256.072 m/z, identified as $[C_{15}H_{12}O_4]^+$) and FT-ICR-MS. CE-MS conditions: bare silica capillary with 80 cm and 50 μm I.D.; BGE composed of 5% formic acid at pH 1.90; run voltage 20 kV; sample injection 50 s at 0.5 psi; sheath liquid: 2-propanol:H_2O (50:50, v/v) at 0.24 ml/h; nebulizer gas at 0.4 bar; dry gas flowing at 4 l/min and 180 °C. FT-ICR-MS in the positive mode all the other conditions as in Fig. 1.

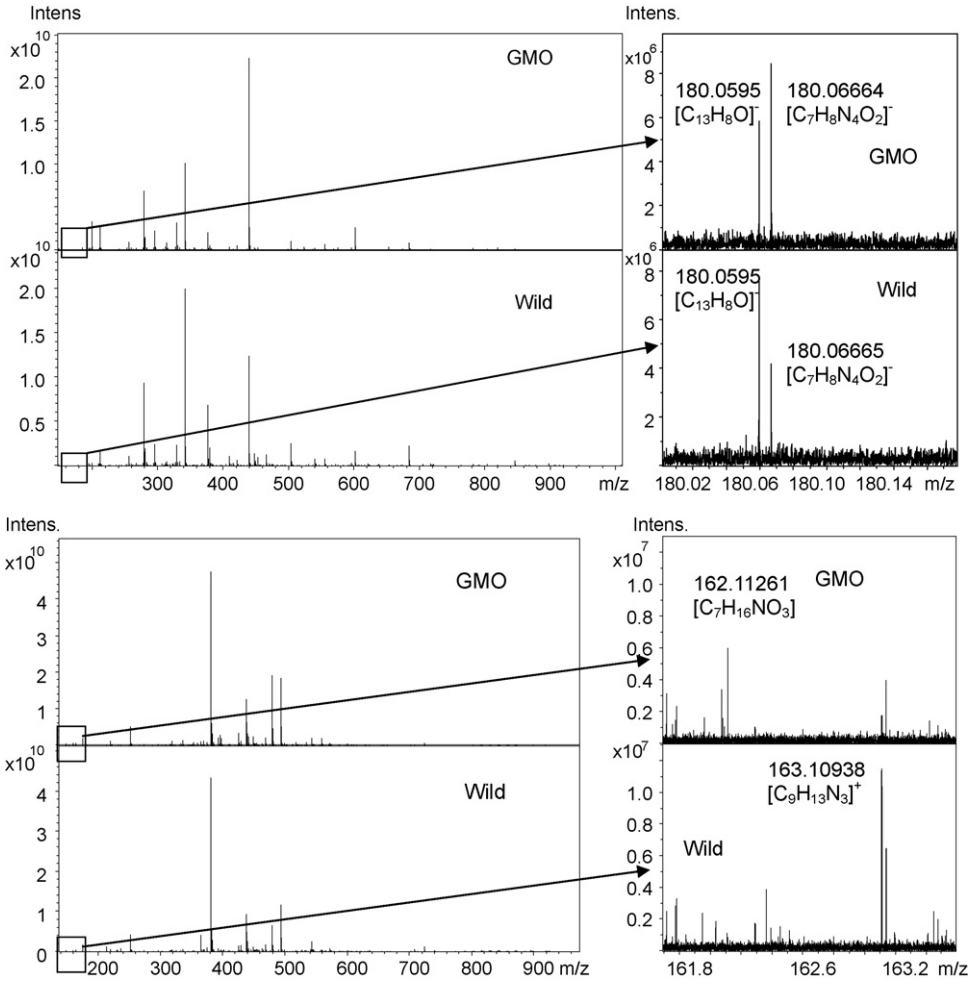
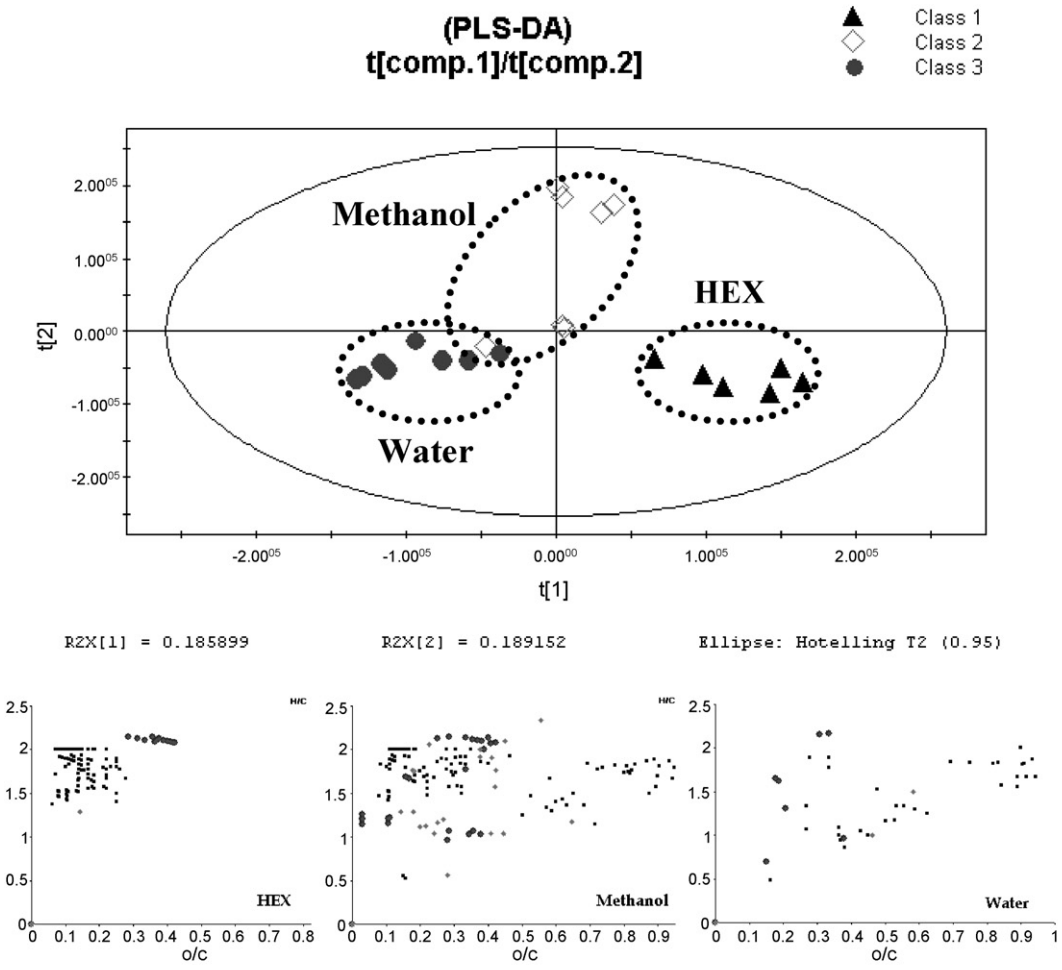


Fig. 4. FT-ICR-MS mass spectra of tentatively identified compounds L-carnitine and tyrosine for Aristis and Aristis Bt maize. Positive ionization mode for L-carnitine (162.1126 *m*/*z*) and negative for tyrosine (180.0666 *m*/*z*). Other experimental conditions as in Fig. 1.

in Table 2, 256.072 *m*/*z*), and the spectra obtained by FT-ICR-MS for both a transgenic and a wild maize. As can be seen, the compound found by FT-ICR-MS was confirmed by CE-TOF-MS. The tentative molecular formulae related to this mass (and isotopic pattern) correspond to $[C_{15}H_{12}O_4]^+$ (theoretical monoisotopical *m*/*z* value is 256.07301) that according to the KEGG database corresponds to apigeninidin, a compound present in the flavonoid biosynthesis metabolism of maize. As can be seen in Fig. 3, the ion 256.07301 *m*/*z* is present in the GMO sample as confirmed by FT-ICR-MS and CE-TOF-MS, while it does not appear in the wild one. Moreover, the electrophoretic mobility is in agreement with this assignment because apigeninidin is expected to bear positive charge at pH 1.90. Therefore, the information obtained by CE-TOF-MS on migration behavior confirms the tentative assignment made by FT-ICR-MS. Other ions that appear in the FT-ICR-MS spectra in Fig. 3 and are not detected in CE-TOF-MS could be different metabolites and even transgenic biomarkers. For example *m*/*z* 256.08192 could be a benzoxazinon derivative present in maize [52], but only molecular mass is not definite criteria to undoubtedly confirm this.

In Fig. 4 two additional examples of the analytical power provided by combining FT-ICR-MS and CE-TOF-MS are shown. Thus, compound 1 corresponding to the ion 162.11261 *m*/*z* together with its isotopic pattern was detected by FT-ICR-MS in the transgenic maize but not in the wild line as can be seen in Fig. 4. The use of CE-TOF-MS corroborated the existence of this compound only in the GM samples and, moreover, it could be identified as L-carnitine ($[C_7H_{16}NO_3]^+$) based on its electrophoretic mobility at pH 1.90. This compound can be considered as a possible biomarker since it systematically appears with higher intensity in all the transgenic lines compared to the wild ones. For comparison in the negative mode, spectra of *m*/*z* 180.06664 tentatively identified in both CE-TOF-MS and FT-ICR-MS experiments as tyrosine ($[C_9H_{10}NO_3]^-$) is also shown.

An interesting case in combining FT-ICR-MS and CE-TOF-MS is shown considering the ion 150.05494 *m*/*z* (compound 11 in Table 2). Thus, using FT-ICR-MS it was possible to assign to this ion the molecular formula $[C_8H_8NO_2]^+$. However, in this case, three different compounds were obtained from the metabolic map of maize for this single mass and formula, namely, 4-hydroxymandelonitrile, 2-formylaminobenzaldehyde and 5,6-dihydroxyindole. CE-TOF-MS extracted ion electropherogram of this *m*/*z* gave one peak in all the transgenic varieties and no peaks in the wild ones, thus confirming the FT-ICR-MS results and showing that this compound could be used as a biomarker. Moreover, based on their electrophoretic mobilities 4-hydroxymandelonitrile, and 2-formylaminobenzaldehyde are compounds that at the running pH are expected to be neutral, so they will not move into the capillary and could not be detected by CE-TOF-MS. Only the third possible compound, 5,6-dihydroxyindole, is the most probable one since it has two aromatic rings and a secondary amine that can bear a positive charge at the separation pH being, therefore, the one detected by CE-TOF-MS. This compound is present in maize in the tyrosine metabolism, which seems to be altered in these GMOs. Thus, CE-TOF-MS is shown to be an useful tool for the con-



KEGG Pathway	HEX	MET	WATER	KEGG Pathway	HEX	MET	WATER	KEGG Pathway	HEX	MET	WATER
Arachidonic acid met.	67	2	0	Tyrosine metabolism	0	2	22	Terpenoid biosynthesis	1	4	3
C21-Steroid hormon met.	36	0	0	Ascorbate, alderate met	0	0	19	Biosynthesis of steroids	5	4	0
Androgen, estrogen met.	32	0	0	Tryptophan metabolism	0	0	16	Val, Leu-Ile metabolism	0	2	0
Biosynt unsat fatty acids	12	6	0	Pentose-glucuronate interc	0	0	14	Ubiquinone biosynt.	5	2	6
Bile acid biosynthesis	7	1	0	Starch and sucrose met	0	0	14	Tyrosine metabolism	0	2	22
Biosynthesis of steroids	5	4	0	Galactose metabolism	0	1	12	Pyruvate metabolism	0	2	2
Sphingolipid metabolism	2	0	0	Fructose, mannose met	0	0	11	Phenylprop biosynt.	1	2	23
Fatty acid elongation	1	0	0	Phenylalanine metabolism	0	0	10	Insect hormone biosynt	1	2	0
Fatty acid metabolism	1	0	0	Phe, Tyr-Trp biosynth.	0	1	7	Fatty acid biosynthesis	7	2	0
				Val, Leu-Ile biosynth.	0	2	0	Brassinosteroid biosynt	5	2	0

Fig. 5. PLS-DA model with one of the maize extracts obtained at two different temperatures (100 and 175 °C) and using three different solvents (hexane, methanol and water). For the three different groups (different solvents) the masses with the highest coefficient of regression are considered. These masses are represented in van Krevelen diagram and submitted in the MassTRIX revealing that the main differences between hexane and water consist in the massive presence of carbohydrate metabolism in the water and fatty acid in the hexane.

firmation of metabolite isomers determined by FT-ICR-MS analysis. Clearly, although the other two compounds were not confirmed as metabolites by FT-ICR-MS + CE-TOF-MS, their existence cannot be excluded.

Based on these examples it is possible to assess the limitations of approaches leading to the characterization of biomarkers. Although the high resolution and sensitivity provided by FT-ICR-MS allows the detection of an impressive number of masses and isotopically confirmed elementary compositions, this technique shows limitations for distinguishing among multiple assignations to the same molecular formula (elucidation of isomers). CE-TOF-MS provides m/z information and adds the electrophoretic mobility data as additional criterion for isomers elucidation. However, CE-TOF-MS can only confirm the ones that will have an electrophoretic mobility in the chosen mobility window governed by the buffer pH and experimental conditions. On the other hand, the use of a database such as MassTRIX (KEEG) to tentatively assign metabolite isomers to elementary compositions found with FT-ICR-MS are mostly limited to the core metabolome of the chosen organism; particular secondary metabolites are not present in the database yet.

Therefore, additional restrictive criteria during the extraction step can be very necessary in order to make a correct assignment when working on metabolomics and/or biomarker discovery. Following this idea, the combination of PLE with FT-ICR-MS was next investigated in order to consider the polarity of the compounds as an additional criterion to find biomarkers.

3.3. PLE and FT-ICR-MS for transgenic maize metabolomics analysis

Extracts from Aristis and Aristis Bt maize seeds were obtained using PLE with three solvents of different polarity (namely, hexane, methanol and water) at two different temperatures, 100 °C and 175 °C. Extracts obtained were injected and analyzed by direct infusion in ESI(−)–FT-ICR-MS as described above. The data were statistically evaluated with a PLS-DA model, as can be seen in Fig. 5. Model validation was carried out via permutation and with the cross-validation (CV). The goodness of prediction of the CV is assessed by the Q^2 measure. Model validation with the number of permutations equaling 200 generated intercepts of $R^2 = 0.399$ and $Q^2 = -0.307$, which means the occurrence of no overfittings and the model validity [53] and it gives confidence that there is a useful linear discriminant relationship present in the data being analyzed [54]. The scoring plot of the model exhibited a clear clustering of three distinct groups. The value of Q^2 and $R^2(Y)$ is respectively 0.62 and 0.76. The analysis and the visualization reveal that the main differences between hexane group and water group consist in the relevant presence of carbohydrate metabolism in the water and methanol, and fatty acid in the hexane. This is determined by the position in the van Krevelen diagrams of the compounds [50] (Fig. 5). The main properties of the group are also presented visually with the van Krevelen mass diagrams [31,33,51,55]. As can be seen in Fig. 5, there are mainly three groups corresponding to the three different extraction solvents used in PLE, water, methanol and hexane. This means that the primary clustering parameter is the solvent type and not the temperature (samples with hexane are clustered together, same with methanol and with water) and it is explained along the first component. The properties that make the water group homogeneous are related with the presence of carbohydrate metabolism and amino acid as determined with the MassTRIX approach. This is clear also through the inspection of van Krevelen diagrams and also from the compound coming from the pathways. The main characteristics of hexane group are linked with the strong presence of fatty acid biosynthesis and the complete absence of the metabolites related to the carbohydrate. The methanol group is between these groups, with the tendency to be more similar to the hexane group, exception for one sample that is internal to the water group, although this group does not appear to be very homogenous. Internal to the cluster the samples tend to group by the type (GM or wild) rather than the temperature exception for the methanol in which the temperature factor seems to have a dominant relevance. This tendency is also corroborated by the van Krevelen diagrams of Fig. 6. Thus, in this Fig. 6 extracted compounds have been classified depending on the type of atoms that they contain, i.e., compounds that contain CHO, CHON, CHOS or CHONS (Table 3). Water is the solvent that extracts by far the highest number of compounds at 175 °C and hexane is shown to be a more selective solvent in this particular case. However, water extracts were strongly affected at higher extraction temperatures with a typical "caramel/Popcorn smell" of the corn seed extracts, slight yellowish color and significant increase in CHO-components with higher double bond equivalences at lower H/C (oxidized phenolic range); the significant increase in N-containing compounds (CHON) may correspond to Maillard-type of products in the visualized H/C and O/C ranges. The methanol and hexane extracts were not as affected by the temperature; in methanol only S-containing components slightly increased in number (CHOS

Table 3
Number of metabolites with their corresponding class of elementary composition found in various extracts. Extracts obtained using PLE with water, methanol and hexane at 100 °C and 175 °C for 20 min (data plotted on van Krevelen diagrams in Fig. 6).

	Water		MetOH		Hexane	
	100 °C	175 °C	100 °C	175 °C	100 °C	175 °C
CHO	279	703	303	314	103	155
CHOS	113	249	133	199	26	23
CHON	46	284	50	48	8	12
CHONS	54	81	51	122	8	5

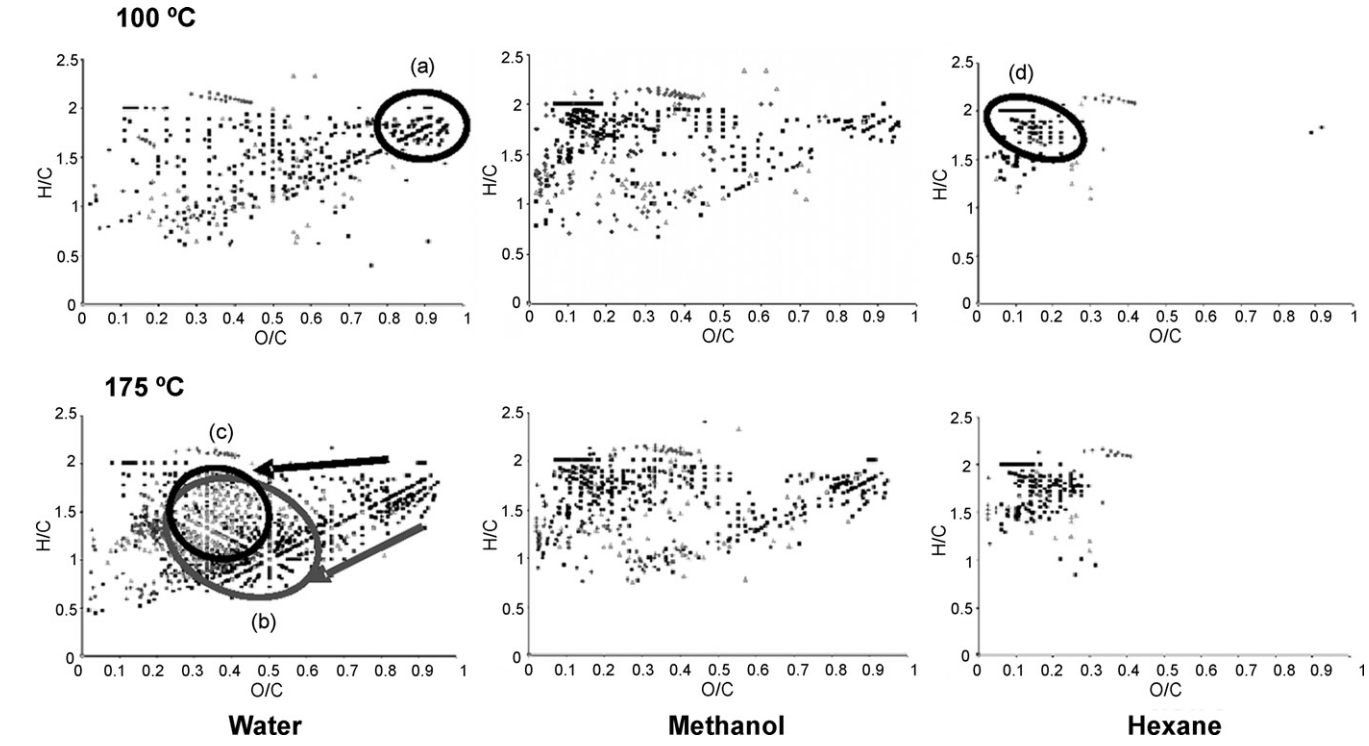


Fig. 6. van Krevelen diagrams (atomic H/C versus O/C) of the maize seed extraction with various solvents and temperatures; (a) carbohydrate type, (b) condensed carbohydrates as CHO, (c) Maillard and Amadori reaction products, and (d) fatty acid type.

and CHONS), but no other variations were found in the different regions.

The extraction methods using pressurized liquids are demonstrated to be a good starting point for metabolomics and biomarker studies. Thus, PLE can provide interesting information based on the polarity of the extracted compounds because it seems to be also a feasible combination to gain additional information on the samples under study in this type of shot-gun metabolomics approaches.

4. Conclusions

Six different maize samples (three GMOs and their respective isogenic wild lines) were directly analyzed by FT-ICR-MS in order to study their metabolic profiles. The study of the compounds found and their metabolic pathways indicates that in transgenic maize more compounds were identified in comparison with the wild maize, particularly in some amino acid related metabolic pathways. Thus, it has been found that some metabolic pathways are altered in the transgenic varieties compared with their respective isogenic lines. Moreover, it is demonstrated that the use of PLE prior to FT-ICR-MS can add additional information based on the polarity of the compounds providing an additional criteria for metabolomic studies.

The ultrahigh resolution and mass accuracy of FT-ICR-MS enable a very large screening of metabolites based on their ionization potential in ESI(+) or ESI(−) and thus a non-targeted approach with assignment of thousands of elementary compositions. Statistical analysis of the data reduces these lists to a lower number of possible m/z biomarkers whose structure cannot directly be given from elementary composition alone. Structural restrictions can be obtained with organism-specific databases (KEGG, MassTRIX) leading to possible biomarker chemical structures but still without possible differentiations in isomers. Results from CE-TOF-MS help to overcome some limitations of FT-ICR-MS in the identification of metabolite isomers and highlight the necessity of having complementary analytical information (e.g., exact mass versus nominal mass, chemical structure versus electrophoretic mobility of compounds) to help to identify possible biomarkers apart of their exact mass value.

CE-TOF-MS filters the observable compounds based on the potential electrophoretic mobility and can be easily involved in routine analysis for screening samples automatically whether FT-ICR-MS is a less accessible and more expensive technology but with higher accuracy and direct elementary composition assignment. Direct coupling of CE with FT-ICR-MS is certainly possible but only with a loss of sensitivity leading to lower biomarker hits and with a lower flexibility in routine use for screening purposes.

Acknowledgements

This work was supported by Projects AGL2005-05320-C02-01 and CONSOLIDER INGENIO 2010 CSD2007-00063 FUN-C-FOOD (Ministerio de Educación y Ciencia), S-505/AGR-0153 (ALIBIRD, Comunidad de Madrid) and HA2006-0057 (Ministerio de Educación y Ciencia). C.L. wants to thank the Comunidad Autonoma de Madrid for a grant. We would like to thank Dr. Pedro J. Martinez for fruitful discussions.

References

[1] M.G. Koziel, G.L. Beland, C. Bowman, N.B. Carozzi, R. Crenshaw, L. Crossland, J. Dawson, N. Desai, M. Hill, S. Kadwell, K. Launis, K. Lewis, D. Maddox, K. McPherson, M.R. Meghji, E. Merlin, R. Rhodes, G.W. Warren, M. Wright, S.V. Evola, Nat. Biotechnol. 11 (1993) 194.
[2] A.G. Gao, S.M. Hakimi, C.A. Mittanck, Y. Wu, B.M. Woerner, D.M. Stark, D.M. Shah, J. Liang, C.M.T. Rommens, Nat. Biotechnol. 18 (2000) 1307.
[3] N.Z. Shaban, M.H. Helmy, M.A.R. El-Kersh, B.F. Mahmoud, Comp. Biochem. Physiol. C 135 (2003) 405.
[4] R. Karlova, M. Weemen-Hendriks, S. Naimov, J. Ceron, S. Dukiandjiev, R.A. De Maagd, J. Invertebr. Pathol. 88 (2005) 169.
[5] M. Geiser, S. Schweitzer, C. Grimm, Gene 48 (1986) 109.
[6] F.S. Betz, B.G. Hammond, R.L. Fuchs, Regul. Toxicol. Pharmacol. 32 (2000) 156.
[7] C. Manetti, C. Bianchetti, L. Casciani, C. Castro, M.E. Di Cocco, A. Miccheli, M. Motto, F. Conti, J. Exp. Bot. 57 (2006) 2613.
[8] C. James, Int. Service Acquisit. Agric. Appl. 32 (2004) 1.
[9] Regulation (EC) 1829/2003, Parliament, European Union, Brussels, 2003.
[10] Regulation (EC) 1830/2003, Parliament, European Union, Brussels, 2003.
[11] Safety of Genetically Engineered Foods: Approaches to Assessing Unintended Health Effects, Institute of Medicine and National Research Council of the National Academies, 2004, p. 39.
[12] G. Flachowsky, A. Chesson, K. Aulrich, Arch. Anim. Nutr. 59 (2005) 1.
[13] C. Simo, A. Rizzi, C. Barbas, A. Cifuentes, Electrophoresis 26 (2005) 1432.
[14] C. Simo, C. Barbas, A. Cifuentes, Electrophoresis 26 (2005) 1306.
[15] E. Ibañez, A. Cifuentes, Crit. Rev. Food Sci. 41 (2001) 413.
[16] A. Cifuentes, B. Bartolome, C. Gomez-Cordoves, Electrophoresis 22 (2001) 1561.
[17] C. Castro, C. Manetti, Anal. Biochem. 371 (2007) 194.
[18] T.A. Clayton, J.C. Lindon, O. Cloarec, H. Antti, C. Charuel, G. Hanton, J.P. Provost, J.L.L. Net, D. Baker, R.J. Walley, J.R. Everett, J.K. Nicholson, Nature 440 (2006) 1073.
[19] J.K. Nicholson, J. Connelly, J.C. Lindon, E. Holmes, Nat. Rev. Drug Discov. 1 (2002) 153.
[20] M.R. Viant, E.S. Rosenblum, R.S. Tjeerdema, Environ. Sci. Technol. 37 (2003) 4982.
[21] Y. Wang, H. Tang, J.K. Nicholson, P.J. Hylands, J. Sampson, E. Holmes, J. Agric. Food Chem. 53 (2005) 191.
[22] G. Le Gall, I.J. Colquhoun, A.L. Davis, G.J. Collins, M.E. Verhoeyen, J. Agric. Food Chem. 51 (2003) 2447.
[23] C. Manetti, C. Bianchetti, M. Bizzarri, L. Casciani, C. Castro, G. D'Ascenzo, M. Delfini, M.E. Di Cocco, A. Lagana, A. Miccheli, M. Motto, F. Conti, Phytochemistry 65 (2004) 3187.
[24] Y. Picó, D. Barceló, Trends Anal. Chem. 27 (2008) 821.
[25] K. Hollywood, D.R. Brison, R. Goodacre, Proteomics 6 (2006) 4716.
[26] J. Chen, X. Zhao, J. Fritsche, P. Yin, P. Schmitt-Kopplin, W. Wang, X. Lu, H.U. Haring, E.D. Schleicher, R. Lehmann, G. Xu, Anal. Chem. 80 (2008) 1280.
[27] E.J. Want, A. Nordstrom, H. Morita, G. Siuzdak, J. Proteome Res. 6 (2007) 459.
[28] K. Yuan, H. Kong, Y. Guan, J. Yang, G. Xu, J. Chromatogr. B 850 (2007) 236.
[29] E.E.K. Baidoo, P.I. Benke, C. Neususs, M. Pelzing, G. Kruppa, J.A. Leary, J.D. Keasling, Anal. Chem. 80 (2008) 3112.
[30] F. Benavente, R. van der Heijden, U.R. Tjaden, J. van der Greef, T. Hankemeier, Electrophoresis 27 (2006) 4570.
[31] N. Hertkorn, M. Meringer, R. Gugisch, C. Ruecker, M. Frommberger, E.M. Perdue, M. Witt, P. Schmitt-Kopplin, Anal. Bioanal. Chem. 389 (2007) 1311.
[32] K. Suhre, P. Schmitt-Kopplin, Nucleic Acids Res. 36 (2008) 481.
[33] R. Rossello-Mora, M. Lucio, A. Peña, J. Brito-Echeverria, A. Lopez-Lopez, M. Valens-Vadell, M. Frommberger, J. Anton, P. Schmitt-Kopplin, ISME J. 2 (2008) 242.
[34] Y. Wang, J. Wang, M. Yao, X. Zhao, J. Fritsche, P. Schmitt-Kopplin, Z. Cai, D. Wan, X. Lu, S. Yang, J. Gu, H.U. Haring, E.D. Schleicher, R. Lehmann, G. Xu, Anal. Chem. 80 (2008) 4680.
[35] S.C. Brown, G. Kruppa, J.L. Dasseux, Mass Spectrom. Rev. 24 (2005) 223.
[36] S.G. Villas-Bôas, S. Mas, M. Åkesson, J. Smedsgaard, J. Nielsen, Mass Spectrom. Rev. 24 (2005) 613.
[37] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass Spectrom. Rev. 26 (2007) 51.
[38] L. Eriksson, E. Johansson, S. Wold, in: C.F.A.S.G. (Ed.), Quantitative Structure–Activity Relationships in Environmental Sciences—VII, SETAC Press, Pensacola, FL, 1997, p. 381.
[39] D.J. Crockford, J.C. Lindon, O. Cloarec, R.S. Plumb, S.J. Bruce, S. Zirah, P. Rainville, C.L. Stumpf, K. Johnson, E. Holmes, J.K. Nicholson, Anal. Chem. 78 (2006) 4398.
[40] S. Bijlsma, I. Bobeldijk, E.R. Verheij, R. Ramaker, S. Kochhar, I.A. Macdonald, B. Van Ommen, A.K. Smilde, Anal. Chem. 78 (2006) 567.
[41] T.M.D. Ebbels, H.C. Keun, O.P. Beckonert, M.E. Bollard, J.C. Lindon, E. Holmes, J.K. Nicholson, J. Proteome Res. 6 (2007) 4407.
[42] A. Nordstrom, G. O'Maille, C. Qin, G. Siuzdak, Anal. Chem. 78 (2006) 3289.
[43] R. García-Villalba, C. León, G. Dinelli, A. Segura-Carretero, A. Fernandez-Gutierrez, V. García-Cañas, A. Cifuentes, J. Chromatogr. A 1195 (2008) 164.
[44] T. Levandi, C. Leon, V. Garcia-Cañas, M. Kaljurand, A. Cifuentes, Anal. Chem. 80 (2008) 6329.
[45] J.A. Mendiola, M. Herrero, A. Cifuentes, E. Ibañez, J. Chromatogr. A 1152 (2007) 234.
[46] V. Garcia-Cañas, A. Cifuentes, R. Gonzalez, Anal. Chem. 76 (2004) 2306.
[47] V. Garcia-Cañas, R. Gonzalez, A. Cifuentes, J. Agric. Food Chem. 50 (2002) 4497.
[48] V. Garcia-Cañas, R. Gonzalez, A. Cifuentes, J. Sep. Sci. 25 (2002) 577.
[49] V. Garcia-Cañas, R. Gonzalez, A. Cifuentes, J. Agric. Food Chem. 50 (2002) 1016.
[50] V. Garcia-Cañas, R. Gonzalez, A. Cifuentes, Electrophoresis 25 (2004) 2219.
[51] N. Hertkorn, M. Frommberger, P. Schmitt-Kopplin, M. Witt, B. Koch, E.M. Perdue, Anal. Chem. 80 (2008) 8908.
[52] V. Cambier, T. Hance, E. De Hoffmann, Phytochem. Anal. 10 (1999) 119.
[53] L. Eriksson, P. Hagberg, E. Johansson, S. Rannar, O. Whelehan, A. Astrom, T. Lindgren, J. Chemometr. 15 (2001) 337.
[54] P. Jonathan, W.V. McCarthy, A.M.I. Roberts, J. Chemometr. 10 (1996) 189.
[55] D.W. van Krevelen, Fuel 29 (1950) 269.