

Metabolomic Strategies Based on High-Resolution Mass Spectrometry as a Tool for Recognition of GMO (MON 89788 Variety) and Non-GMO Soybean: a Critical Assessment of Two Complementary Methods

Vojtech Hrbek¹ · Veronika Krtkova¹ · Josep Rubert¹ · Hana Chmelarova¹ · Katerina Demnerova² · Jaroslava Ovesna³ · Jana Hajslova¹

Received: 2 February 2016 / Accepted: 30 April 2017 / Published online: 30 May 2017
© Springer Science+Business Media New York 2017

Abstract This study was focused on distinguishing of genetically modified organism (Monsanto 89788 variety) and conventional soybeans by employing high-resolution mass spectrometry (HRMS) techniques for non-target screening of sample extracts. Two hyphenated instrumental platforms represented by (i) ultrahigh-performance liquid chromatography (U-HPLC) coupled to quadrupole/time of flight and (ii) ambient mass spectrometry with direct analysis in real time (DART) ion source-coupled OrbitrapMS were used. The statistical processing of generated data (metabolomic fingerprints) was performed by multivariate data analysis; principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) showed that both employed techniques enabled correct classification of genetically modified and conventional soybeans. In addition, some phosphatidylcholines and sugars were identified as the most significant markers.

Keywords Soybean · Genetically modified organism · Metabolomics · High-resolution mass spectrometry · Fingerprinting

Introduction

Various genetically modified organism (GMO) food crops (e.g. corn, potato or rice) have been grown commercially since 1994, when the first GMO plant (tomato) was introduced on the market (James 2011; Nap et al. 2003). In general terms, GMO means an organism, with the exception of human beings, in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination. Specific traits in GMO food crops often include resistance to various harmful/adverse factors—pests, diseases, cold, drought, etc., or tolerance to treatment by non-selective herbicides that kill all other undesirable plants such as weeds (Phipps and Park 2002; Senior and Dale 2002).

Soybean represents one of the most common genetically modified food crops; the estimated share of the total world production of GMO soybean was about 83% in (James 2015). According to United States Department of Agriculture (USDA, National Agricultural Statistics Service, June Agricultural Survey for the years 2000-16), GMO soybeans represented 94% of total soybean production of the USA in 2015 and also in 2016. Very common GMO soybean variety, Monsanto (MON 89788), is resistant against glyphosate, an active ingredient of the widely used non-selective herbicide known under the trade name Roundup. Glyphosate specifically inhibits enzyme, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase, causing important chemical

Electronic supplementary material The online version of this article (doi:10.1007/s12161-017-0929-8) contains supplementary material, which is available to authorized users.

✉ Jana Hajslova
jana.hajslova@vscht.cz

- ¹ Faculty of Food and Biochemical Technology, Department of Food Analysis and Nutrition, University of Chemistry and Technology, Prague, Technicka 3, 166 28 Prague 6, Czech Republic
- ² Faculty of Food and Biochemical Technology, Department of Biochemistry and Microbiology, University of Chemistry and Technology, Prague, Technicka 3, 166 28 Prague 6, Czech Republic
- ³ Crop Research Institute, Prague, Dmowska 507/73, 161 06 Prague 6, Czech Republic

intermediates during aromatic amino acid synthesis to be unavailable and, as a result, the plant dies. The resistance in GMO soybean is commonly achieved by insertion of different variants of the gene for EPSP synthase from soil bacteria *Agrobacterium tumefaciens* that is not inhibited by glyphosate (Arun et al. 2013; Dinon et al. 2010; Ujhelyi et al. 2008).

Although genetic modification typically results in higher yields and improved quality of respective crops, this approach still remains a subject of intensive debates regarding not only possible health risks for consumers but also due to other aspects such as environmental hazards. As documented in many published studies dealing with these issues (Devos et al. 2014; Hilbeck et al. 2014; Rajan and Letourneau 2012), GMO crops are matter of concerns mainly in the European Union (EU). With regards to all existing controversies, GMO crops are forbidden in eight EU member states (Austria, Bulgaria, Germany, Greece, Hungary, Italy, Luxembourg and Poland); in other EU countries, including Czech Republic, only GMO maize MON 810 can be grown under strictly specified conditions (Czarnak-Klos and Rodríguez-Cerezo 2010).

With regards to a diversity of attitudes to GMO crops, their labelling alike products thereof are a globally important issue (Tutelyan 2013). In the EU, it is mandatory to label any food as GMO, supposing that it contains more than 0.9% GMO ingredients according to Directive 2001/18/EC of the European Parliament and Council Regulation (EC) No. 1829/2003 and 1830/2003 (EC 2003a, b). In the recent years, several rapid, highly specific and sensitive methods based on polymerase chain reaction (PCR) have been standardised and applied as a useful tool for identifying GMOs. For quantification, real-time PCR is nowadays available in specialised laboratories; simultaneous detection of DNA fragments and their quantification is enabled (Del Gaudio et al. 2012; Mavropoulou et al. 2005; Zhang and Guo 2011).

In the recent years, metabolomic-based strategies have been introduced as a challenging tool for food authenticity (Rubert et al. 2015; Cubero-Leon et al. 2014; Simó et al. 2014). For this purpose, a broad set of metabolites contained in a tested sample (fingerprint) was assessed against those occurring in a 'reference' material. Advanced chemometric tools are combined to find possible differences and identify quality/authenticity markers. As regards GM crops, it is assumed that some, at least minor changes in the metabolic pathways as compared conventional ones may have occurred due to genetic manipulation.

Various instrumental techniques can be employed to obtain sample 'fingerprints'; amongst them, high-resolution mass spectrometry (HRMS) appears to be a dominating role (Rubert et al. 2015; Theodoridis et al. 2012). In most cases, HRMS is hyphenated with some separation technique, mostly (ultra)high-performance liquid chromatography ((U)HPLC). However, in recent years, ambient mass spectrometry

(AMS), which is a technique that omits chromatographic separation and provides mass spectral fingerprint of the entire sample, was also used. This technique has been shown to be a very promising tool, high-throughput authentication approach. Of the ion sources employed in AMS, direct analysis in real time (DART) is one (together with desorption electrospray ionisation (DESI)) of the most widely used (Cody et al. 2005).

The greatest advantage of U-HPLC-MS technology over DART-MS is that in addition to spectral separation of sample components, their chromatographic pre-separation is also involved. Moreover, significantly higher number of features obtained by this technique is because greater amount of analytes that enters into the MS source (in case of DART-MS, only those compounds that can be transferred into the gas phase and ionised) can be detected. U-HPLC-MS method allows a greater variability in analysis/separation of various substances by selecting different mobile and stationary phases; it is not possible in case of DART-MS. Another advantage of U-HPLC-MS is to achieve better performance characteristics (e.g. better repeatability) than DART-MS. The disadvantage of U-HPLC-MS is higher duration of analysis of one sample, usually from 10 to 20 min compared to analysis of one sample by DART-MS, only 5–10 s. The main advantages of DART ionisation compared to conventional separation and ionisation techniques include direct sample analysis under ambient conditions, minimal or no sample preparation and remarkably high sample throughput. One of the main drawbacks in DART-MS techniques is represented by matrix affects. Similar to electrospray ionisation (ESI) and/or atmospheric pressure chemical ionisation (APCI), mainly matrix-dependent signal suppression is encountered when examining real-life samples. The impact of matrix effects might be fairly severe, considering the absence of chromatographic (or any other) separation prior to sample ionisation. The most important method performance parameters such as linearity, precision and accuracy could be influenced by interfering matrix components (Vaclavik et al. 2010). Specificity of detection is another key assumption in ambient ionisation techniques since isobaric (and isomeric) interferences, potentially present in examined samples, can cause problems during both qualitative and quantitative analyses. Whilst tandem mass spectrometry (MS/MS) or mass spectrometry to the n th power (MS^n) measurements may overcome this problem in target analysis, the use of instruments with a high or ultrahigh resolving power is the unavoidable condition for reliable applications, both in target and profiling (fingerprinting) measurements of complex samples (Vaclavik et al. 2010).

The DART technique investigated in this study represents one of the APCI-related ionisation techniques employing a glow discharge for the ionisation. Metastable helium atoms, created from the glow discharge, react with ambient water,

oxygen or other atmospheric components to produce reactive ionizing species (Cody et al. 2005). The DART ion source was shown to be efficient for soft ionisation of a wide range of both polar and non-polar compounds and coupled with mass spectrometry represents a powerful analytical tool for metabolomic studies, especially in combination with chemometric methods (Stewart et al. 2014; Bylesjö et al. 2006). Until now, direct analysis in real time coupled with high-resolution mass spectrometry (DART-HRMS) has been most notably used in olive oil authentication (Vaclavik et al. 2009), as a powerful tool for beer origin recognition (Cajka et al. 2011), and for authentication of tomatoes and peppers from organic and conventional farming (Novotna et al. 2012).

Several studies based on metabolomic approaches for classification/discrimination of GM and conventional crops have been recently published (Kusano et al. 2015; Simó et al. 2014). The studies primarily deal with Fourier transform infrared spectroscopy (FT-IR) and nuclear magnetic resonance spectroscopy (NMR) analyses (Kim et al. 2009), and only one publication used mass spectrometry as a tool for metabolic fingerprinting (Vaclavik et al. 2013).

In this study, different analytical platforms were used in order to distinguish GMO and non-GMO soybean samples. For this purpose, polar extracts of reference materials of GMO and non-GMO soybeans were initially evaluated using LC-HRMS, providing characteristic soybean fingerprints. This technique provided more detail about the characterisation of the soybean samples, thanks to higher sensitivity and chromatographic separation. Subsequently, the same polar extracts were explored by DART-HRMS, which allowed very rapid analysis and GMO differentiation in real time.

Materials and Methods

Samples

In this study, the samples were provided by the Crop Research Institute (CRI) of Prague (Czech Republic), as a partner in the project NAZV-QI101B267, whereby this study was performed. A total of 49 samples of soybean were analysed, of which 30 were GMO (all samples were MON 89788 variety, collected from different growing areas) soybean and 19 conventional samples of different varieties were collected from different growing areas. The samples were in the form of seeds, which were homogenised into flour. The initial classification of the samples was performed by CRI using DNA analysis by PCR. This assay is based on the isolation of DNA, and the DNA amplification was verified using primers specific for the lectin gene. PCR was used to detect the presence of GMO elements for the 35S cauliflower mosaic virus (CaMV) promoter, nopaline synthase (NOS) terminator and EPSP synthase (Ovesna et al. 2010). The samples were stored

in darkness and dry conditions at room temperature (20 °C) until the time of laboratory processing.

To verify the suitability of the metabolomic approaches, commercially available ‘GMO soybean’ Certified Reference Materials (CRMs) with relevant modification, i.e. resistance against total herbicide glyphosate, were purchased from Sigma-Aldrich (Germany), ERMBF410DK ERM® Certified Reference Material, 1% Roundup Ready™, ERMBF410GK ERM® Certified Reference Material, 10% Roundup Ready™, ERMBF410AK Roundup Ready™ blank and ERM® Certified Reference Material.

Sample Preparation

The samples were placed into the freezer (−20 °C) and the next day immediately, after removing from the freezer were homogenised using the mill GRINDOMIX GM200 (Retsch, G). The samples immediately after homogenisation were weighed (2 g) into 15-mL polypropylene cuvettes. The extraction solvent, 10 mL of mixture, methanol:water (8:2, v/v), was added to the cuvette and manually shaken for 2 min. The mixture was then centrifuged (5 min, 20 °C, 10,000 rpm). The supernatant was collected using a 5-mL plastic syringe and filtered through a 0.22-µm filter. The filtrate was transferred to a 2-mL vial and analysed the same day.

Ultrahigh-Performance Liquid Chromatography Coupled with High-Resolution Mass Spectrometry Analysis

The chromatographic analysis was performed using an Acquity UPLC™ system (Waters Corp., Milford, MA, USA) using a Waters Acquity UPLC® BEH C18 column (100 × 2.1 mm i.d., 1.7 µm) at 60 °C and a flow rate of 0.4–0.5 mL/min. The mobile phase consisted of water-methanol (95:5, v/v) with 5 mM ammonium formate and 0.1% formic acid (A) and mixture of isopropanol:methanol:water (65:30:5, v/v/v) with 5 mM ammonium formate and 0.1% formic acid (B) with a gradient elution 0–1 min 90–50% (A) flow 0.4 mL/min, 1–5 min 50–20% (A) flow 0.4 mL/min, 5–11 min 20–0% (A) flow 0.4 mL/min, 11–12 min 0% (A) flow 0.5 mL/min and 12–14 min 90% (A) flow 0.5 mL/min for positive mode and gradient elution 0–4 min 90–50% (A) flow 0.4 mL/min, 4–6 min 50–0% (A) flow 0.4 mL/min, 6–11 min 0% (A) flow 0.5 mL/min and 11–14 min 90% (A) flow 0.5 mL/min for negative mode. The injection volume was 3 µL.

Mass spectrometric detection was performed on a Synapt G2 MS system (Waters Corp., Milford, MA, USA) equipped with an ESI source. Except MS¹, two types of data acquisition modes, MS^E and MS/MS, were carried out in order to investigate precursor ions and product ions. Nitrogen gas was used for nebulisation. The detection mode of the flight tube was selected to be a ‘W’ pattern. The usage of the W pattern in the

flight tube provides an advantage in comparison, for example with a 'V' pattern in the length of the flight tube. In case of the W pattern, the flight path is longer than for the V pattern and higher resolving power is achieved, up to 20,000 full width at half maximum (FWHM). Positive and negative ion spectra were recorded across the range of m/z 50–1200. The ion source conditions were as follows: capillary voltage, 1.0 kV in positive mode and -0.7 kV in negative mode; sampling cone voltage, ± 35 V; extraction cone voltage, ± 4.0 V; ESI source temperature, 120 °C; desolvation temperature, 350 °C; cone gas flow, 30 L/h; desolvation gas flow, 800 L/h; collision gas flow, 0.5 mL/min; and collision energy for MS^E acquisition mode, 4.0 eV for low-energy scans and 25–35 eV for high-energy scans. The instrument was tuned using leucine-enkephalin (2 ng/ μ L, water:methanol (50:50, v/v) with 0.1% formic acid) to provide a resolving power higher than 20,000 FWHM (m/z 556.2771 in ESI+ and m/z 554.2615 in ESI–), and the interval scan time was 0.02 s. The mass accuracy was maintained within the whole acquisition period by using a lock spray with the leucine-enkephalin as the reference compound to correct small mass drifts during the measurement. The mass calibration in both ionisation modes was performed by sodium formate solution (0.5%). Masslynx 4.1 (Waters Corp., Milford, MA, USA) was used to control the instrument.

DART-HRMS Analysis

For the analysis using ambient mass spectrometry, the DART ion source (DART-SVP) was fitted with a 12Dip-It™ tip scanner autosampler (IonSense, Saugus, MA, USA) coupled to an Exactive™ benchtop (Thermo Fisher Scientific, Bremen, Germany). A Vapur™ interface (IonSense, Saugus, MA, USA) was employed to couple the ion source to the mass spectrometer, and low vacuum in the interface chamber was maintained by a membrane pump (Vacuubrand, Wertheim, Germany). The distance between the exit of the DART gun and the ceramic transfer tube of the Vapur was set to 10 mm; the gap between the ceramic tube and the inlet to the heated capillary of the Exactive was 2 mm.

The DART and MS instruments were operated in both positive and negative ionisation modes, and the optimised settings were as follows: (i) DART positive ionisation helium flow 2.5 L min^{-1} , gas temperature 450 °C, discharge needle voltage 5000 V and grid electrode +350 V; (ii) DART negative ionisation helium flow 2.5 L min^{-1} , gas temperature 400 °C, discharge needle voltage 5000 V and grid electrode -350 V; and (iii) mass spectrometric detection capillary voltage ± 60 V, tube lens voltage ± 150 V and capillary temperature 250 °C. The sheath, auxiliary and sweep gases were disabled during DART-MS analyses.

The mass spectrometer was operated at mass resolving power 50,000 FWHM calculated for m/z 200. The mass

spectrum acquisition rate was 2 spectra s^{-1} . Liquid samples were delivered into the DART ionisation region with the use of 12 Dip-It tip scanner autosampler. Dip-It™ tips (IonSense, Saugus, MA, USA) were inserted into a holder and immersed in sample extracts placed in a 96-deep-well micro-plate (Life Systems Design, Merenschwand, Switzerland). The Dip-It holder was mounted onto the body of the autosampler, and the Dip-It tips were automatically moved at a constant speed of 0.5 mm s^{-1} through the helium gas between the exit of the DART gun and the inlet of the Vapur interface.

Standard external mass calibration of MS system in the range of 150–2000 m/z was performed both in positive and negative modes prior every measurement according to the manufacturer's instructions. Also, adjusted mass calibration in ESI(+) and ESI(–) in the mass range of m/z 50–750 using collisional induced dissociation (CID) at 25 eV was subsequently performed in order to cover the lower masses.

Quality Control

In order to check absence of carryover effects and to control the stability of fingerprints recorded, blank and quality control (QC) matrix samples were analysed within both DART-HRMS and U-HPLC-MS sequences. It should be noted that the in-batch sequence of tested samples was random (established based on random number generation) to avoid any possible time-dependent changes during DART-HRMS and U-HPLC-MS analyses, which could result in false clustering. To control the overall performance of instrumental system, QC samples were inserted into the sequence, always after a set of 10 test samples, and analysed under the same conditions. The QC sample was a pool of three (randomly selected) extracts of conventional soybean samples. In this way, the repeatability of sample fingerprints could be monitored. The good instrument performance was documented by a tight clustering of these QC samples (i.e. similarity of their fingerprints) in the principal component analysis (PCA) plot.

Data Analysis

Chemometric analysis included multivariate data analysis using unsupervised and supervised models. PCA and orthogonal partial least squares discriminant analysis (OPLS-DA) were employed based on SIMCA software (v. 13.0, 2011, Umetrics, Umea, Sweden; www.umetrics.com).

In the first stage, data processing and data pre-treatment must be carried out in order to capture the bulk of variation between different datasets. In this way, raw data generated by soybean analysis employing DART-HRMS technique (51 signals in positive ionisation mode, 57 signals in negative ionisation mode exceeding 2% abundance of the most intensive ion in average DART spectrum) in the form of absolute peak intensities were pre-processed using constant row sum;

that is, each variable was divided by the sum of all variables for each sample. This procedure transformed all the data to a uniform range of variability. DART-HRMS data were initially processed by Xcalibur 2.2 software and copied to MS Excel 2010. The macrofunction was used in a following step for creation the final tables which were exported to the SIMCA software. U-HPLC-HRMS technique produced 520 features in positive ionisation mode and 108 features in negative ionisation mode. In this case, data processing was performed by MassLynx 4.1 software subroutine MarkerLynx XS, and the resulting data file was exported to the SIMCA software.

Subsequently, Pareto scaling was applied prior PCA and OPLS-DA (Worley and Powers 2013). Then, PCA analysis enabled transformation of the original variables (normalised intensities of ions) to the new uncorrelated variables (principal components). In this way, the reduced dimensionality of the data was obtained whilst still preserving information from the original dataset. Additionally, OPLS-DA was subsequently applied to identify and reveal the most significant metabolites. The objective of OPLS-DA was to divide the systematic variation in the X block into two model parts, one part which models the co-variation between X and Y and another part which expresses the X variation that is not related (orthogonal) to Y . OPLS-DA was performed in order to provide a better distribution of samples and allow creation of a statistical model and validation.

The quality of the models was evaluated by the goodness-of-fit parameter (R^2X), the proportion of the variance of the response variable that is explained by the model (R^2Y) and the predictive ability parameter (Q^2), which was calculated by a k -fold internal cross validation of the data using a default option of the SIMCA software. In general terms, the value of R^2 must be higher than Q^2 and acceptable value of Q^2 is more than 0.5 (Blasco et al. 2015). In addition, the models were also evaluated in terms of recognition and prediction abilities. Recognition ability represents the percentage of samples in the training set, which were correctly classified. Prediction ability is the percentage of samples in the test set correctly classified by using the model developed during the training step. For this purpose, sevenfold internal cross validation was used (Berrueta et al. 2007). For the control of the Q^2 values, if they were stable and relevant (correctly calculated), the permutation test was used (Triba et al. 2015).

Selection of the Marker Ions (Markers) and Identification

The selection of marker ions, which have strong impact on the sample classification, can be done by several tools. One of these useful tools is an S-plot enabled by SIMCA software. The S-plot illustrating the distribution of detected features was employed for statistical evaluation. Features at the extremes of the S-plot, the outermost ions, can be considered as marker ions with the highest importance for sample separation. For

sorting the marker ions according to their importance, the variable importance in the projection (VIP) plots which explain X and show a correlation to Y can be used. There are 10 most important variables in a given model, i.e. those with VIP score >1 . For an explanation/confirmation of ions as markers also, a trend plot was used. In this way, the variability of the top ions across measurements of a set of different test samples could be illustrated.

In the case of U-HPLC-HRMS, tentative identification of compounds behind the marker ions was based on the estimation/calculation of elemental formula (accurate mass and mass error for respective m/z values in MS^1 was considered). To confirm suggested identification of marker ions, their product ions were investigated in MS/MS spectra. The identification of several other compounds occurring in sample fingerprints was supported by interpretation of MS^E spectra acquired in respective retention time. Both MS/MS and MS^E spectra were obtained by the using collision energy ramp ranging from 15 to 45 V. Online databases such as ChemSpider (www.chemspider.com) or Metlin (www.metlin.scripps.edu/index.php) were employed for compound identification. Regarding lipid identification, the information on their fragmentation as published by Krank et al. (2007) and Zhao et al. (2011) was used. In case of DART-HRMS, the tentative identification of marker ions, and other ions, was done by the similar way as described for U-HPLC-HRMS, except of using MS^E or MS/MS functions. These functions were not available in case of DART-HRMS.

Results and Discussion

As stated in the “Introduction” section, metabolomic studies generally aim at a comprehensive analysis of the metabolome, without a particular bias to specific groups of metabolites. In our study, we employed two fingerprinting strategies, U-HPLC-HMRS and DART-HRMS, for non-targeted analysis of both GMO and conventional soybean samples.

Selection of Extraction Solvent

The choice of extraction solvent was a critical decision to be made when planning this metabolomic-based study. We took into account that mainly qualitative aspects rather than quantitative ones are important for a sample characterisation; therefore, the optimisation of sample preparation step was aimed at isolation of a broadest possible representation of soybean metabolites. In the first step, the mixture of methanol:isopropanol (50:50, v/v) was examined as possible extraction solvent for non-polar metabolites (represented mainly by triacylglycerols). Nevertheless, these results did not show a statistically significant difference between the samples of conventional and GMO soybean. In the next step, we focused only on polar

and medium polar metabolome components which, in line with study by Marrelli et al. (2013), might be presumed to be important for a sample classification/distinguishing of GMO vs non-GMO soybean samples. Polar solvent mixture, methanol:water (8:2, v/v), was used in all follow-up experiments.

DART-HRMS Fingerprints

In the second phase of our experiments, the potential of high-throughput ambient ionisation technique, DART-HMRS, was investigated. Again, we started with an optimisation of DART-HRMS conditions for a detection of as broad as possible spectrum of metabolites occurring in soybean extracts. Helium beam temperature and desorption time were the major DART source operating parameters affecting the transfer of sample metabolites into a gas phase, their ionisation and transmission into MS system. When using 14-s desorption, of the tested temperatures (300, 350 and 400 °C), 350 °C enabled the highest number of ions and the highest responses both in DART ionisation modes ($[M + H]^+$ ions are typical in positive mode, whilst $[M - H]^-$ ions are mainly generated by negative ionisation).

Figure 1 shows the comparison of DART-HRMS fingerprints (i.e. averaged mass spectra across the entire desorption peak) of conventional soybean (A) and GMO soybean (B) extracts as obtained in a positive ionisation mode. Alike in the case of U-HPLC-HRMS, the fingerprints were similar, only slightly differing in relative intensities of individual ions. The only exception was relatively intensive m/z 163.0600, which occurred in conventional soybeans, whilst in GMO soybeans, it was either not detectable or occurred at very low intensity. The elemental composition of this marker ion was tentatively estimated to be $C_6H_{11}O_5$, which could correspond to protonated molecule of deoxyhexose (Δ ppm = 0.2) that might originate from precursor hexose through the loss of water in ion source.

The comparison of conventional (C) and GMO soybean (D) fingerprints obtained using the DART-HRMS technique in a negative ionisation mode is shown in Fig. 1. In this particular case, the difference between mass spectra was evident. Besides of differing relative intensities of some ions occurring in both mass spectra, also qualitative differences (some ions present/absent) could be observed. For example, the ion at m/z 293.2126 ($C_{18}H_{30}O_3$, Δ ppm = 1.8) was present only in the spectral fingerprint from conventional soybean, whilst m/z 295.2282 ($C_{18}H_{32}O_3$) vernolic acid (Δ ppm = 0.5), m/z 311.2232 ($C_{18}H_{32}O_4$, Δ ppm = 1.5), m/z 313.2387 ($C_{18}H_{34}O_4$, Δ ppm = 1.5), m/z 327.2187 ($C_{18}H_{32}O_5$, Δ ppm = 2.3) and m/z 329.2338 ($C_{18}H_{34}O_5$, Δ ppm = 1.2) were present in both fingerprints but they were more relatively intensive in conventional soybean compared to GMOs.

Ultrahigh-Performance Liquid Chromatography Coupled with High-Resolution Mass Spectrometry Fingerprints

In this phase of experiments concerned with obtaining a comprehensive characterisation of soybean extracts, optimal measurement conditions were searched. Generic settings (i.e. conditions found as optimal in previous, similar experiments) for U-HPLC separation and HRMS detection were employed. The testing of chromatography, including optimisation of the type of stationary/mobile phases, gradient optimisation or separation temperature, was done. In this study, reverse phase was chosen, providing chromatographic separation and peak resolution for a wide range of metabolites (with different molecular structures and polarities). In parallel, HILIC column was also tested; however, the number of molecular features was significantly reduced, reverse phase vs HILIC: 520 vs 135 in positive ionisation mode and 108 vs 63 in negative ionisation mode. In this way, contrary to an ambient MS, isobaric and isomeric metabolites could be identified. In addition, due to a lower matrix suppression, an improvement of detectability, based on the number of detected metabolites (U-HPLC-HRMS vs DART-HRMS: 520 vs 51 in positive ionisation mode, 108 vs 57 in negative ionisation mode), was observed.

The comparison of base peak chromatograms (BPCs) of aqueous methanolic extracts obtained by analysis of GMO and conventional soybeans in a positive ionisation mode did not show significant differences. The fingerprints of sample extracts contained phospholipids (PLs) as major components (retention time (tR) 7–9 min; see Fig. 2a) with m/z values in the range 710–790. The dominating PLs were phosphatidylcholines (PCs), partly separated according to the structure of bound fatty acids. The two most intensive peaks at tR 7.58 and 7.89 min were palmitoyl-arachidonoyl phosphatidylcholine (PC; 16:0/20:4, m/z 782.5686) and palmitoyl-linoleoyl PC (16:0/18:2, m/z 758.5705), the latter one partially co-eluted with palmitoyl-eicosatrienoyl PC peak (16:0/20:3, m/z 784.5858). Altogether, 17 PLs were initially identified (mass errors lower than 5 ppm).

Attention was also paid to the most polar metabolites eluting in the front part of chromatogram at tRs 1–2 min. Many of the peaks detected here corresponded to phytoestrogens, typical biologically active (estrogenic) secondary metabolites occurring in soybeans. In Fig. S1 (Supplementary data), both free and bound forms are shown: daidzein, genistein, glycitein and their conjugates (glycosides—daidzin, genistin, glycitin, and acetylated forms—acetyl daidzin, acetyl genistin, acetyl glycitin, malonyl derivatives: malonyl daidzin, malonyl genistin, malonyl glycitin). The profiles of these compounds were also assessed for both sample categories; however, again, no significant differences could be observed. The intensities of these

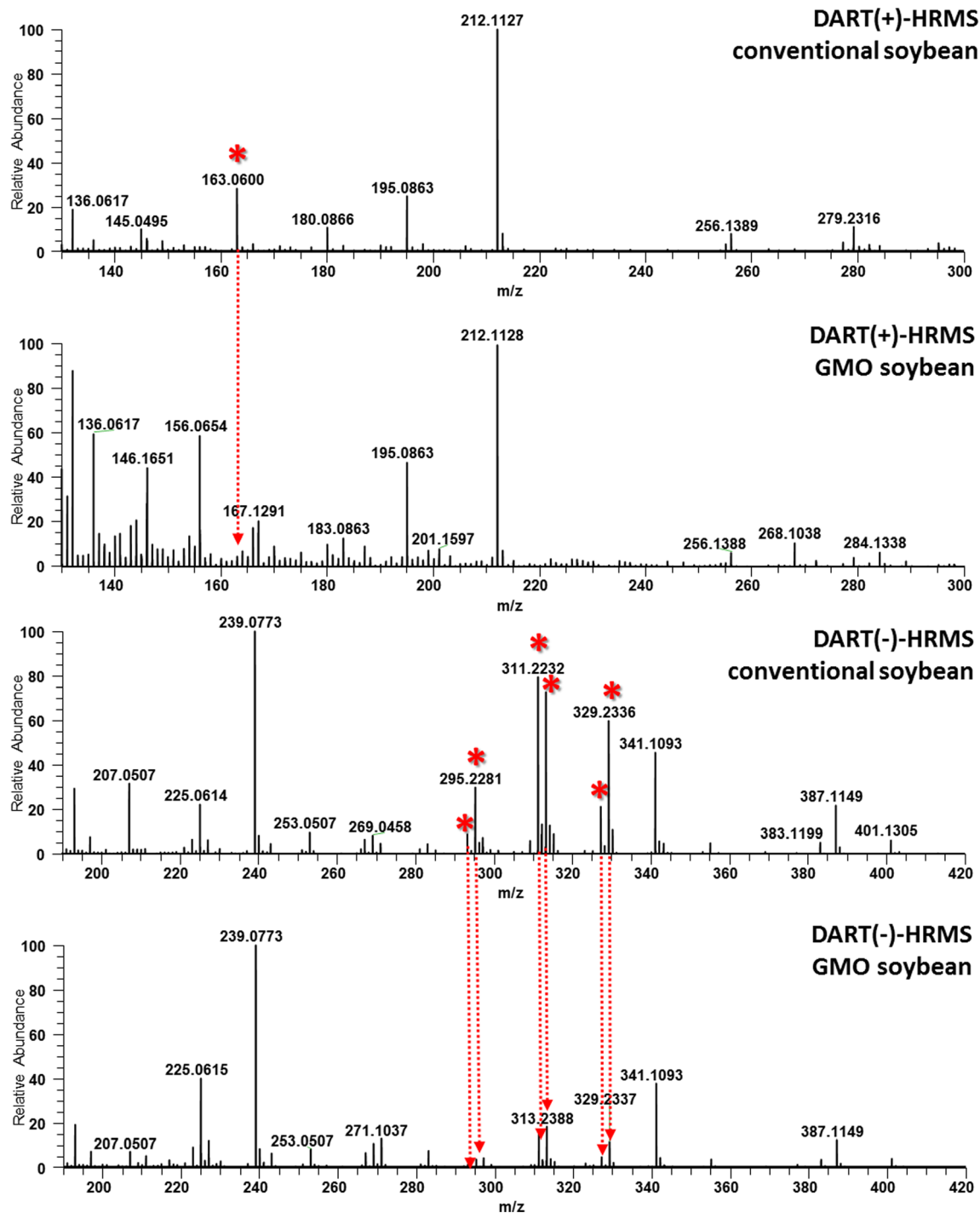


Fig. 1 DART(+)-MS mass spectra (m/z 130–300) obtained by analysis of 80% methanolic extracts of conventional soybean, GMO soybean and DART(-) MS mass spectrum (m/z 190–420) obtained by analysis of 80% methanolic extracts of conventional soybean, GMO soybean

compounds were similar in case of GMO and non-GMO soybean samples. Target analysis only for these compounds was also done, as an additional analysis, according to our in-house previously developed and validated methodology. The content of phytoestrogens was very similar; no trend for differentiation of soybean samples based on the content of individual phytoestrogens was observed.

Similar to the outcomes above (positive ionisation), the inspection of sample fingerprints recorded in a negative ionisation mode (see Fig. 2b) did not show unambiguous differences. The major metabolites identified here were lysophospholipids (tR 5.5 to 6.0), free fatty acids (tR 6.0 to 6.5 min) and phospholipids (tR 6.5–7.0 min). Most of phytoestrogens (eluted at tR 1–3 min) could be identified here as well.

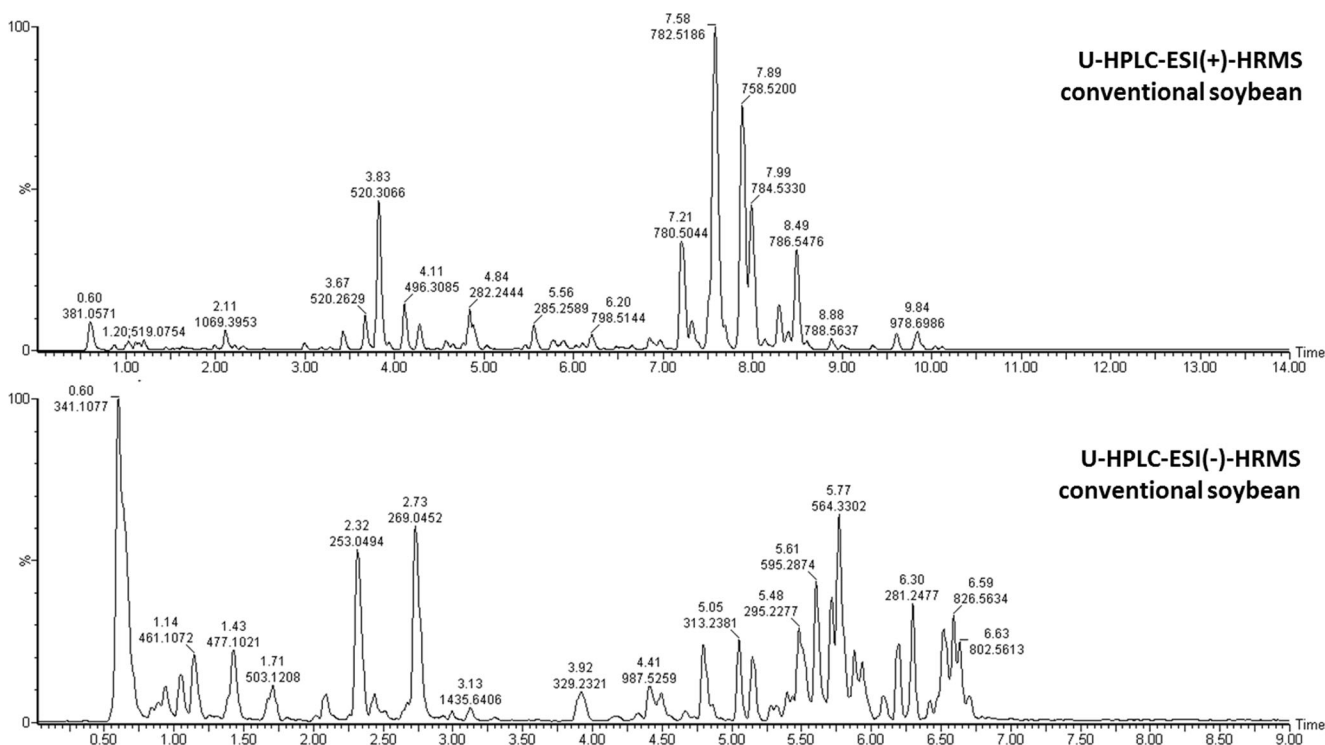


Fig. 2 U-HPLC-HRMS BPC chromatogram obtained by the analysis of 80% aqueous methanolic extract in positive ionisation and negative ionisation of conventional soybean sample

Methods Performance Characteristics

Within the validation protocol, repeatability of non-targeted measurements ($n = 6$) was calculated for both MS techniques employed in our study since this parameter plays an important role in classification of sample sets represented by normalised fingerprints, repeatability of each normalised feature intensities. Detected ion repeatabilities (after deconvolution when chromatographic separation was employed), expressed as relative standard deviations (RSDs), were in case of U-HPLC-HRMS in the range of 2.5–5.3% for positive and 2.7–6.2% for negative ionisation mode; in case of DART-HRMS, they ranged from 11.5 to 17.6% for positive mode and 12.3 to 18.4% for negative ionisation mode, as shown in Table 1. The range of RSD represents minimum and maximum values of RSD calculated for the normalised intensities of each individual ion (feature). RSDs obtained for QC samples were within respective ranges above. As expected, in both ionisation modes, lower values of RSD were obtained for U-HPLC-HRMS compared to DART-HRMS; since the latter case, the repeatability of a thermal desorption of analytes from the surface of sampling tips might be a limiting factor.

Chemometric Analysis

Generally, data acquisition with high-resolution mass spectrometry yields high volumes of raw data that need to be processed by advanced statistical tools. As far as compounds

most contributing to sample classification should be identified, then using a variety of bioinformatic tools aiming to derive information from measured data is necessary. The strategies employed in this study are briefly outlined below.

Multivariate Data Analysis

Initially, the data obtained by analysis of 49 soybean samples (conventional, $n = 19$; GMO soybean, $n = 30$) were processed by PCA both in LC(+/-) and DART(+/-) techniques. This unsupervised approach showed clustering behaviour related to the type of soybean (conventional vs GMO soybean; see Fig. 3). As discussed in detail in following paragraphs, the expected better class resolution in a discriminant problem was fulfilled when employing OPLS-DA. Further improved separation between classes was achieved regardless what type of technique was used (see Fig. 4).

As regards processing of U-HPLC-HRMS data, PCA clearly separated GMO and conventional samples obtained both positive and negative ionisation modes. In Fig. 3c presenting the positive ionisation data, PC1 and PC2 together described 52.8% of the sample set variability (33 and 19.8% for the PC1 and PC2, respectively); in case of negative ionisation data shown in Fig. 3d, this was even 68.3% (47.1 and 21.2% for the PC1 and PC2, respectively). Considering the fact that the first five PCs explain 87.7% (ESI+) and 87.1% (ESI-) of the total of variance, PC1/PC2 plot seemed to be a

Table 1 Detected ion repeatibilities (after deconvolution when chromatographic separation was employed), expressed as relative standard deviations

LC-HRMS		DART-HRMS	
Positive ionisation	Negative ionisation	Positive ionisation	Negative ionisation
2.5–5.3%	2.7–6.2%	11.5–17.6%	12.3–18.4%

RSD relative standard deviation

good starting point for sample clustering according to GMO or non-GMO.

In the next step (following PCA analysis), OPLS-DA was used. As it was expected, even more efficient separation of samples into groups was achieved, and the mathematical model obtained in this way reliably enabled correct classification of an unknown sample was created; recognition and prediction abilities of 100% in both ionisation modes were excellent.

Regarding the data obtained by DART-HRMS, PCA using positive ionisation mode showed much better sample clustering compared to negative ionisation mode. As shown in Fig. 3a presenting the positive ionisation data, PC1 and PC2 together described 58% of the sample set variability (38.1 and 19.9% for the PC1 and PC2, respectively); in case of negative ionisation data shown in Fig. 3b, this was even 63.7% (35.6 and 28.2% for the PC1 and PC2, respectively). Considering the fact that the first five PCs explain 83.3% (DART+) and 87.7% (DART-) of the total of variance, PC1/PC2 plot, also in this

case, promising clustering according to GMO or non-GMO could be obtained. This was reconfirmed by subsequent OPLS-DA analysis with a creation of mathematical models for sample classification. The quality of the models was evaluated by the goodness-of-fit parameter (R^2X), the proportion of the variance of the response variable that is explained by the model (R^2Y) and the predictive ability parameter (Q^2), which was calculated by a k -fold internal cross validation of the data using a default option of the SIMCA software. In general terms, the value of R^2 must be higher than Q^2 and acceptable value of Q^2 is more than 0.5 (Blasco et al. 2015). In addition, the models were also evaluated in terms of recognition and prediction abilities. Recognition ability represents the percentage of samples in the training set, which were correctly classified. Prediction ability is the percentage of samples in the test set correctly classified by using the model developed during the training step. For this purpose, sevenfold internal cross validation was used (Berrueta et al. 2007). For the control of

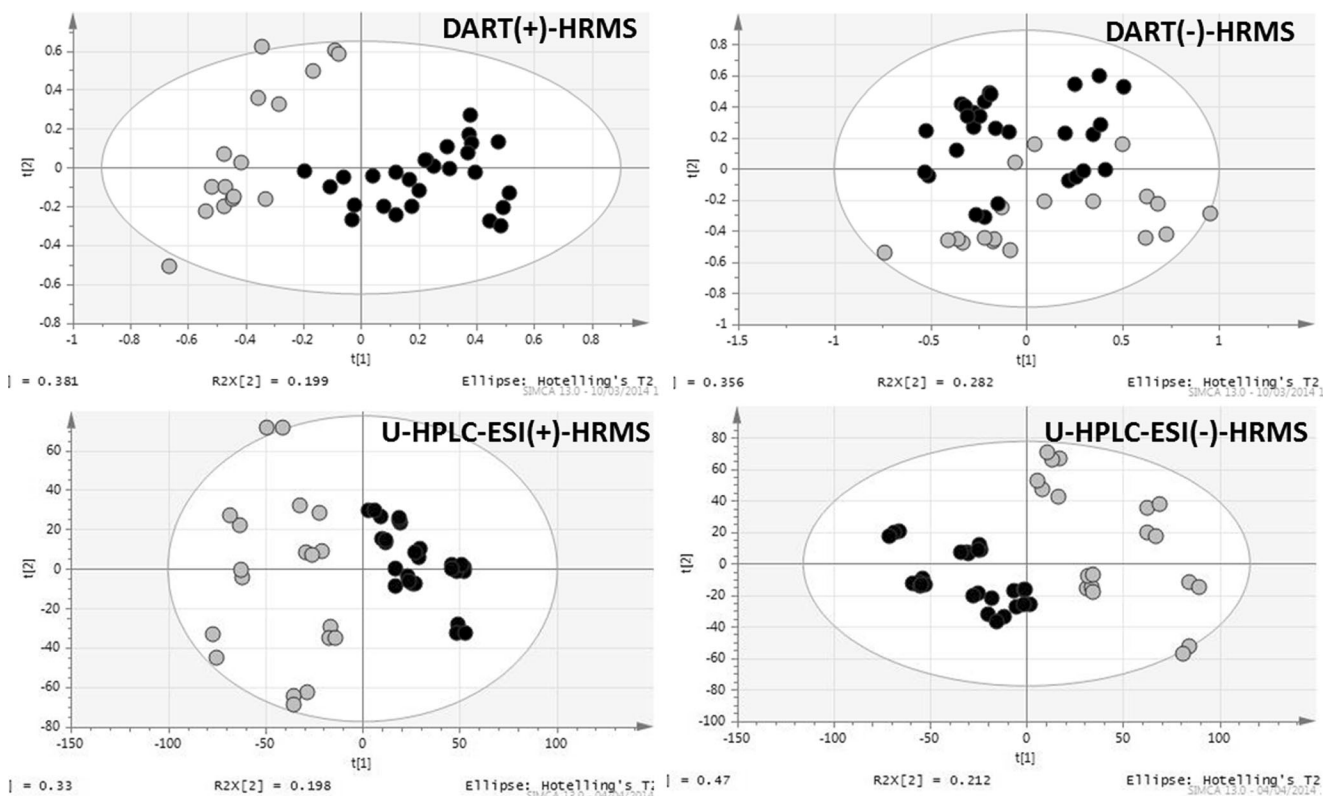


Fig. 3 PCA analysis of data generated by DART-HRMS analysis in positive mode, negative mode and by U-HPLC-HRMS analysis in positive mode, negative mode of conventional (grey) and GMO (black) soybean samples

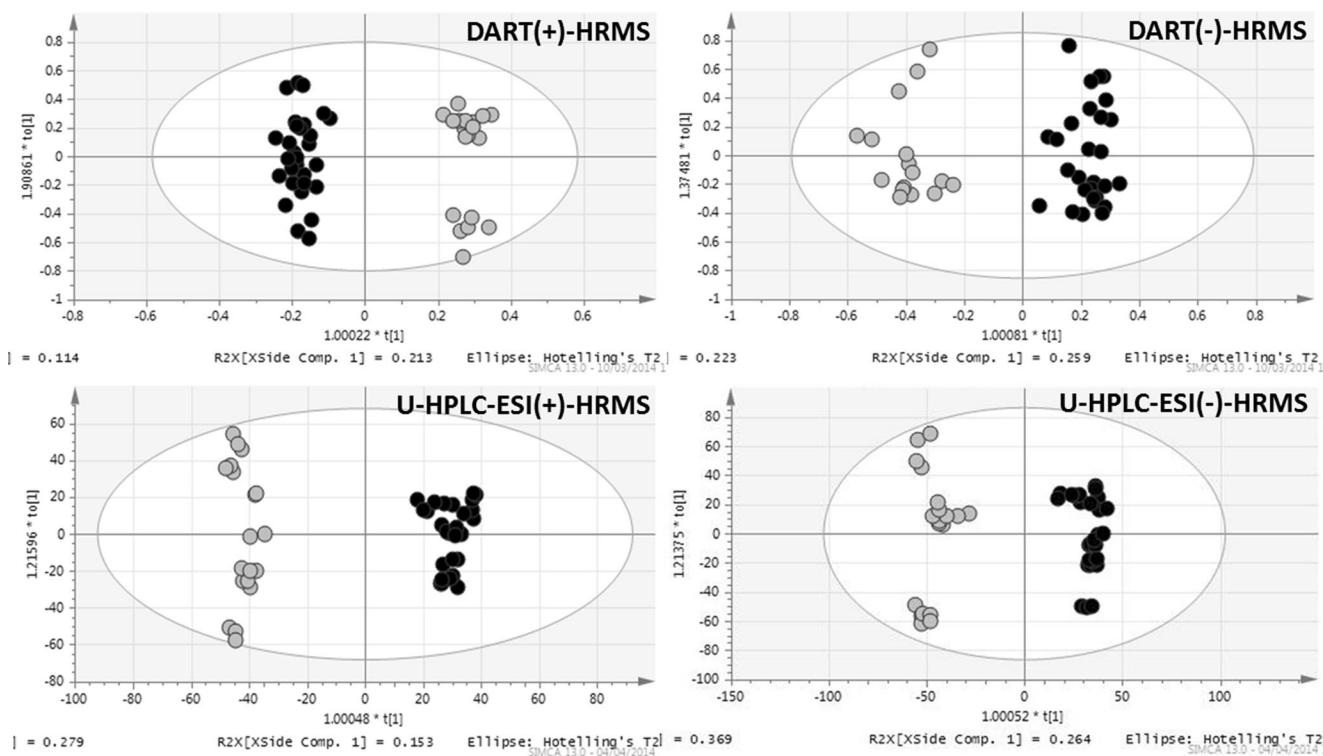


Fig. 4 OPLS-DA analysis of data generated by DART-HRMS analysis in positive mode, negative mode and by U-HPLC-HRMS analysis in positive mode, negative mode of conventional (*grey*) and GMO (*black*) soybean samples

the Q^2 values, if they were stable and relevant (correctly calculated), the permutation test was used (Triba et al. 2015).

Whilst for positive ionisation mode, both the recognition and prediction ability were 100%, for negative ionisation mode, these parameters were slightly lower, but still acceptable: recognition ability 94% and prediction ability 90%. The

values of R^2 and Q^2 summarised in Table 2 show that in all cases, good statistical models with Q^2 value which exceed acceptable limit 0.5 (Blasco et al. 2015), in all cases, enabling the classification of the samples correctly, were obtained. Also, the parameters from the permutation tests (number of permutations was 200) are included in Table 2 for both

Table 2 The quality parameters for the statistical models

Statistical model	Model diagnostics											
	LC-HRMS						DART-HRMS					
	Positive ionisation			Negative ionisation			Positive ionisation			Negative ionisation		
	R^2X	R^2Y	Q^2	R^2X	R^2Y	Q^2	R^2X	R^2Y	Q^2	R^2X	R^2Y	Q^2
PCA	0.717	–	0.573	0.910	–	0.816	0.870	–	0.506	0.962	–	0.719
OPLS-DA	0.666	0.983	0.953	0.739	0.974	0.966	0.889	0.979	0.949	0.810	0.945	0.902
Permutation test ^a												
GMO soybean samples												
R^2	(0.0, 0.536)			(0.0, 0.405)			(0.0, 0.072)			(0.0, 0.235)		
Q^2	(0.0, –0.480)			(0.0, –0.630)			(0.0, –0.379)			(0.0, –0.368)		
Non-GMO soybean samples												
R^2	(0.0, 0.538)			(0.0, 0.419)			(0.0, 0.074)			(0.0, 0.223)		
Q^2	(0.0, –0.441)			(0.0, –0.575)			(0.0, –0.375)			(0.0, –0.383)		

R^2 fit, Q^2 predictive ability

^a The number of permutations was 200

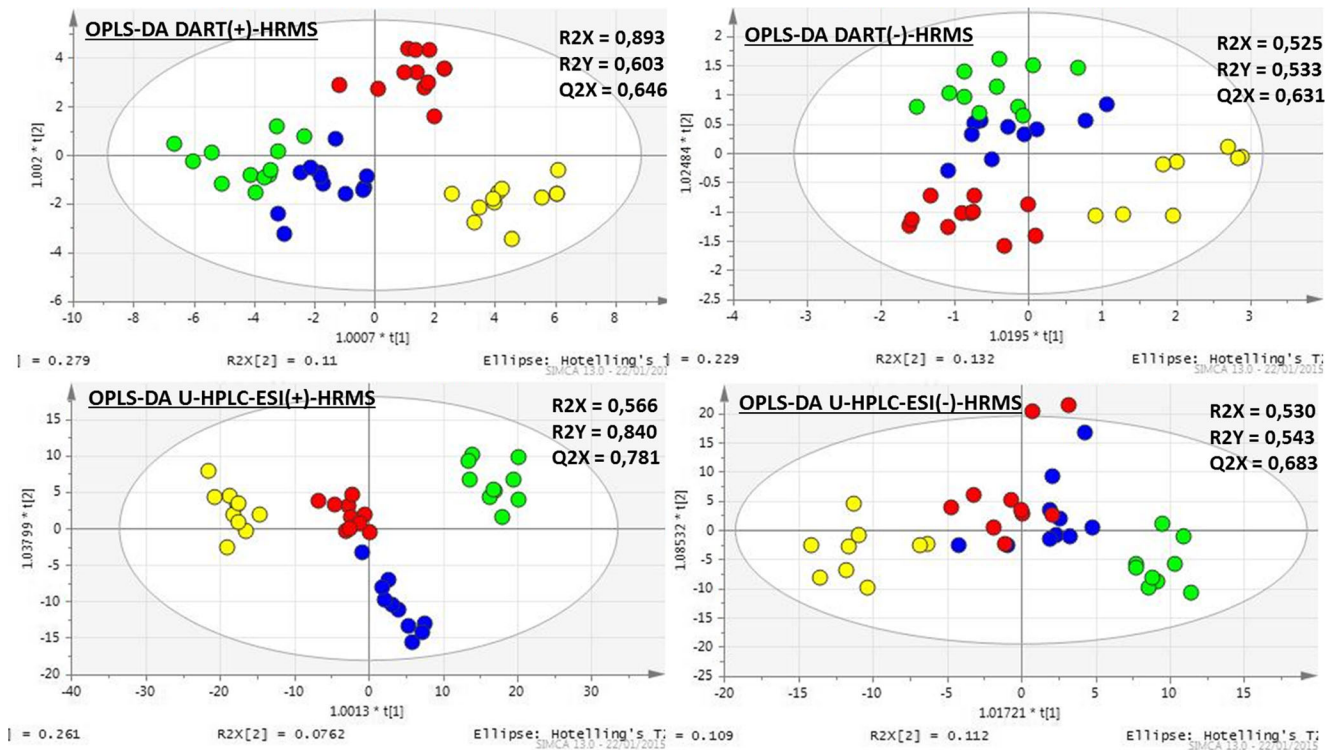


Fig. 5 Chemometric analysis of data generated by DART-HRMS and U-HPLC-HRMS for certified reference material analysis: *blank* = non-genetically modified soybean (*green*), soybean with 1% of GMO soybean

(*blue*), soybean with 5% of GMO soybean (*red*) and soybean with 10% of GMO soybean (*yellow*), OPLS-DA in positive and negative mode

technique, both group of samples and both ionisation. The permutation plots are listed in Fig. S2 (Supplementary data) for U-HPLC-HRMS data and Fig. S3 (Supplementary data) for DART-HRMS data.

As earlier emphasised, for verification of the entire analytical procedure (strategy) including statistical model, QC

samples were analysed. The QC samples were analysed after every 10 tested samples (for checking the consistency of obtaining results), and samples were randomly measured. The obtained fingerprints of analysed soybean samples and those of QC sample were continuously monitored throughout the sequence to avoid the problems with unstable

Table 3 Top 10 ‘markers’ enabling classification of soybean samples

LC-HRMS						DART-HRMS			
Markers sorted according to the importance									
Positive ionisation			Negative ionisation			Positive ionisation		Negative ionisation	
<i>m/z</i>	<i>tR</i> (min)	VIP score	<i>m/z</i>	<i>tR</i> (min)	VIP score	<i>m/z</i>	VIP score	<i>m/z</i>	VIP score
520.2995	3.82	3.39	341.1083	1.20	3.30	127.0390	3.79	128.0345	3.47
522.3109	4.29	2.95	564.3299	5.76	2.85	255.0651	2.60	269.0459	3.42
496.3399	4.11	2.09	595.2876	5.61	2.47	285.0756	2.11	181.0710	2.21
760.5295	8.29	2.04	695.4642	6.55	2.44	104.0708	1.90	105.0184	1.99
758.5705	7.89	1.94	295.2271	5.48	2.24	268.1039	1.59	253.0507	1.97
282.2510	4.83	1.92	571.2878	5.71	2.24	180.0866	1.53	329.2338	1.64
518.2797	3.43	1.79	269.0450	2.73	2.20	195.0863	1.51	119.0340	1.55
784.5858	7.98	1.77	387.1137	1.45	2.09	163.0600	1.43	295.2282	1.40
599.4453	7.33	1.72	683.2240	1.39	2.03	212.1128	1.11	327.2181	1.32
786.6012	8.50	1.70	566.3456	5.94	2.03	274.2124	1.12	293.2126	1.28

tR retention time, *VIP* variable importance in the projection

Table 4 Summary of significant markers obtained using the LC (+/-)-MS techniques

LC-MS—the most important markers for distinguishing samples—positive ionisation							
(<i>m/z</i>) [M + H] ⁺	Retention time (min)	Molecular formula	Identification	(<i>m/z</i>) [M + H] ⁺	Retention time (min)	Molecular formula	Identification
282.2510	4.83	C ₁₈ H ₃₆ NO	Unidentified	599.4453	7.33	C ₃₉ H ₆₆ O ₄	Unidentified
496.3399	4.11	C ₂₄ H ₅₀ NO ₇ P	LPC (C16:0)	758.5705	7.89	C ₄₂ H ₈₀ NO ₈ P	PC (16:0/18:2)
518.2797	3.43	C ₂₆ H ₄₈ NO ₇ P	LPC (C18:3)	760.5295	8.29	C ₄₂ H ₈₂ NO ₈ P	PC (16:0/18:1)
520.2995	3.82	C ₂₆ H ₅₀ NO ₇ P	LPC (C18:2)	784.5858	7.98	C ₄₄ H ₈₂ NO ₈ P	PC (16:0/20:3)
522.3109	4.29	C ₂₆ H ₅₂ NO ₇ P	LPC (C18:1)	786.6012	8.50	C ₄₄ H ₈₄ NO ₈ P	PC (18:0/18:2)
LC-MS—the most important markers for distinguishing samples—negative ionisation							
(<i>m/z</i>) [M – H] [–]	Retention time (min)	Molecular formula	Identification	(<i>m/z</i>) [M – H] [–]	Retention time (min)	Molecular formula	Identification
269.0450	2.73	C ₁₅ H ₁₀ O ₅	Genistein	566.3456	5.94	C ₃₄ H ₄₉ NO ₆	Unidentified
295.2271	5.48	C ₁₈ H ₃₂ O ₃	Unidentified	571.2878	5.71	C ₃₉ H ₄₀ O ₄	Unidentified
341.1083	1.20	C ₁₂ H ₂₂ O ₁₁	Disaccharide	595.2876	5.61	C ₄₁ H ₄₀ O ₄	Unidentified
387.1137	1.45	C ₁₈ H ₂₈ O ₉	Unidentified	683.2240	1.39	C ₄₉ H ₃₂ O ₄	Unidentified
564.3299	5.76	C ₃₄ H ₄₇ NO ₆	Unidentified	695.4642	6.55	C ₄₆ H ₆₄ O ₆	Unidentified

LPC lysophosphatidylcholine, PC phosphatidylcholine

measurement condition. Not only visual inspection but also statistical assessment of generated data provided relevant control. As illustrated by PCA statistical analysis shown in Fig. S4 (Supplementary data), repeatable fingerprints of QC samples were obtained by both techniques that are documented by their excellent clustering in both ionisation modes.

Analysis of CRM

CRMs of blank (non-GMO soybeans) and 10% of GMO soybeans were used for preparation of admixtures, which contain 5% of GMO soybeans. The sample preparation was the same as in “Sample Preparation” section. CRM material was prepared in six repetitions and analysed first day; after 1 week, new six repetitions of CRM were prepared and analysed (for better description of repeatability/reproducibility). The obtained data were processed in the same way as was described above. The

results of statistical analyses (OPLS-DA) are shown in Fig. 5. Figure 5 illustrates not only a good separation of blank soybeans and samples with 10% of GMO soybeans but also acceptable separation of 5 and 1% GMO soybean samples. OPLS-DA models enabled reliable classification of both GMO (10, 5, 1%) and non-GMO soybean samples. Moreover, detection of GMO soybeans admixed to non-GMO soybeans (blank) and vice versa was feasible at levels as low as 1%.

Selection and Identification of the Markers

As it was already described in “Selection of the Marker Ions (Markers) and Identification” section, S-plot, VIP-plot and trend plot were used. Figure S5 (Supplementary data) shows an example of an S-plot for features obtained by analysis of GMO soybean extract by U-HPLC-HRMS in positive ionisation mode. Ten of the most important ions, the most

Table 5 Summary of significant markers obtained using the DART (+/-)-MS techniques

DART-MS—the most important markers for distinguishing samples—positive ionisation					
(<i>m/z</i>) [M + H] ⁺	Molecular formula	Tentative identification	(<i>m/z</i>) [M + H] ⁺	Molecular formula	Tentative identification
104.0708	C ₄ H ₉ O ₂ N	Unidentified	212.1128	C ₇ H ₁₇ O ₆ N	Unidentified
127.0390	C ₆ H ₆ O ₃	Maltol, isomaltol	255.0651	C ₁₅ H ₁₀ O ₄	Daidzein
163.0600	C ₆ H ₁₁ O ₅	Deoxyhexose	268.1039	C ₉ H ₁₇ O ₈ N	Unidentified
180.0866	C ₆ H ₁₄ O ₅ N	Deoxyhexose [M + NH ₄] ⁺	274.2124	C ₁₅ H ₂₉ O ₄	Unidentified
195.0863	C ₇ H ₁₄ O ₆	Methylglucosid	285.0760	C ₁₆ H ₁₂ O ₅	Glycitein, biochanin A, 5-O-methylgenistein
DART-MS—the most important markers for distinguishing samples—negative ionisation					
(<i>m/z</i>) [M – H] [–]	Molecular formula	Tentative identification	(<i>m/z</i>) [M – H] [–]	Molecular formula	Tentative identification
105.0184	C ₃ H ₆ O ₄	Glyceric acid	269.0459	C ₁₅ H ₁₀ O ₅	Genistein
119.0340	C ₄ H ₈ O ₄	Tetrose	293.2126	C ₁₈ H ₃₀ O ₃	Unidentified
128.0345	C ₅ H ₇ O ₃ N	Pyroglutamic acid	295.2282	C ₁₈ H ₃₂ O ₃	Unidentified
181.0710	C ₆ H ₁₄ O ₆	Sugar alcohol	327.2181	C ₁₈ H ₃₂ O ₅	Unidentified
253.0507	C ₁₅ H ₁₀ O ₄	Daidzein	329.2338	C ₁₈ H ₃₄ O ₅	Unidentified

remotes, were selected and the attention was paid to them. The importance of these 10 ions was evaluated by VIP-plot and summarised in Table 3 and Fig. S6 (Supplementary data). This procedure was done for obtaining the most important 10 markers for both techniques and both ionisation modes.

The identification of markers usually represents the last step within metabolomic studies. This is crucial in order to understand the metabolite pathway, since they can be the interesting intermediates, or final secondary metabolites. The relevance of chosen markers is documented by trend plot in Fig. S7 (Supplementary data), where the most important marker for each technique in positive or negative ionisation can be seen. In these trend plots, the distribution of GMO and non-GMO soybean samples, in analysed sample set, according the abundance of relevant markers was evident.

Tables 4 and 5 summarise suggested top 10 markers. The identification procedure is based on the accurate mass measurements, a possible empirical formula as is described in “Selection of the Marker Ions (Markers) and Identification” section. The proposed list in Tables 4 and 5 should be considered only as tentative identification.

Conclusions

Based on the critical assessment of the two alternative fingerprinting techniques, U-HPLC-HRMS and DART-HRMS, employed for distinguishing of genetically modified (MON 89788 variety), we could conclude that both of them demonstrated sufficient flexibility and discrimination power enabling to built-up relevant statistical models, and subsequent marker compound identification. The benefits and limitation of these instrumental approaches can be characterised as follows:

U-HPLC-HRMS was superior over DART-HRMS in terms of higher number of features (520 vs 51 in positive ionisation mode, 108 vs 57 in negative ionisation mode), thus providing more data available for statistical processing; moreover, the repeatabilities (RSDs, %) of detected ion intensities were better (i.e. lower) than those obtained by DART-HRMS. On this account, the recognition/prediction abilities enabled by OPLS-DA models constructed on U-HPLC-HRMS data were better compared to those achieved by the ambient technique. On the other hand, DART-HRMS technique enabled significantly higher throughput of sample measurements (approximately 2 min, in both ionisation modes, per sample vs 14 min (ESI+) and 9 min (ESI-) in case of U-HPLC-HRMS usage). It is worth to notice that in particular case, processing of raw data obtained by DART-HRMS technique was rather complicated since their export data from Xcalibur 2.2 to MS Excel 2010 (required by SIMCA chemometric SW) had to be done manually.

As far as identification of marker compounds, the usage of U-HPLC-HRMS technique was clearly the preferred option. The main reason for that was obtaining high-resolution fragmentation mass spectra and interpretation of which can be supported by respective databases. In case of DART-HRMS, the availability of tandem mass analyser would enable improved identification of some metabolites (potential markers), although, because of the absence of chromatographic separation, isobaric and isomeric compounds would remain unresolved anyway.

In conclusion, DART-HRMS technique seemed to be a promising analytical technology for authentication of conventional vs GMO soybean samples. However, the usage of U-HPLC-HRMS technique was classified as more appropriate analytical strategy.

Compliance with Ethical Standards

Funding This work was supported by the Ministry of Agriculture of the Czech Republic (NAZV-QI101B267), Operational Programme Prague—Competitiveness (CZ.2.16/3.1.00/21537 and CZ.2.16/3.1.00/24503) and by the ‘National Program of Sustainability I’—NPU I (LO1601—No. MSM-43760/2015). Josep Rubert thanks the Generalitat Valenciana (Conselleria d’Educació, Cultura i Esport) for the VALi + d postdoctoral fellowship ‘Contractació de personal investigador en formació en fase postdoctoral 2014’ (APOSTD/2014/120).

Conflict of Interest Vojtech Hrbek declares that he/she has no conflict of interest. Veronika Krtkova declares that he/she has no conflict of interest. Josep Rubert declares that he/she has no conflict of interest. Hana Chmellarova declares that he/she has no conflict of interest. Katerina Demnerova declares that he/she has no conflict of interest. Jaroslava Ovesna declares that he/she has no conflict of interest. Jana Hajslova declares that he/she has no conflict of interest.

Compliance with Ethical Requirements This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Not applicable.

References

- Arun Ö, Yilmaz F, Muratoglu K (2013) PCR detection of genetically modified maize and soy in mildly and highly processed foods. *Food Control* 32:525–531
- Berrueta LA, Alonso-Salces RM, Heberger K (2007) Supervised pattern recognition in food analysis. *J Chromatogr A* 1158:196–214
- Blasco H, Blaszczyński J, Billaut J, Nadal-Desbarats L, Pradat P, Devos D, Moreau C, Andres C, Emond P, Corcia P, Slowinski (2015) Comparative analysis of targeted metabolomics: dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. *J Biomed Inform* 53:291–299
- Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J (2006) OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemom* 20:341–351
- Cajka T, Riddellova K, Tomaniova T, Hajslova J (2011) Ambient mass spectrometry employing a DART ion source for metabolomic

- fingerprinting/profiling: a powerful tool for beer origin recognition. *Metabolomics* 7:500–508
- Cody RB, Laramée JA, Durst HD (2005) Versatile new ion source for the analysis of materials in open air under ambient conditions. *Anal Chem* 77:2297–2302
- Cubero-Leon E, Peñalver R, Maquet A (2014) Review on metabolomics for food authentication. *Food Res Int* 60:95–107
- Czarnak-Klos M, Rodríguez-Cerezo E (2010) Best practice documents for coexistence of genetically modified crops with conventional and organic farming. 1. Maize crop production. European Coexistence Bureau (ECoB). Institute for Prospective Technological Studies. Publications Office of the European Union, Luxembourg, 10, 4619
- Del Gaudio S, Cirillo A, Di Bernardo G, Galderisi U, Cipollaro M (2012) Verification of real-time PCR methods for qualitative and quantitative testing of genetically modified organisms. *J Food Qual* 35:442–447
- Devos Y, Aguilera J, Diveki Z, Gomes A, Liu Y, Paoletti C, Jardin P, Herman L, Perry JN, Waigmann E (2014) EFSA's scientific activities and achievements on the risk assessment of genetically modified organisms (GMOs) during its first decade of existence: looking back and ahead. *Transgenic Res* 23:1–25
- Dinon AZ, Trembl D, Mello CS, Arisi ACM (2010) Monitoring of GMO in Brazilian processed meat and soy-based products from 2007 to 2008. *J Food Compos Anal* 23:226–229
- EC (2003a) Regulation (EC) 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. *Official Journal of the European Communities L* 268:1–23.36
- EC (2003b) Regulation (EC) 1830/2003 of the European Parliament and of the Council of 22 September concerning the traceability and labelling of genetically modified organisms and the traceability of food and feed products produced from genetically modified organisms and amending. Repealing Council Directive 2001/18/EC. *Official Journal of the European Communities L* 268:24–28.36
- Hilbeck A, Weiss G, Oehen B, Römbke J, Jänsch S, Teichmann H, Lang A, Otto M, Tappeser B (2014) Ranking matrices as operational tools for the environmental risk assessment of genetically modified crops on non-target organisms. *Ecol Indic* 36:367–381
- James C (2011) Executive summary of global status of commercialized biotech/GM crops: 2011. ISAAA Briefs No: 43, Ithaca
- James C (2015) 20th anniversary (1996 to 2015) of the global commercialization of biotech crops and biotech crop highlights in 2015. (ISAAA Brief No. 51) Ithaca, NY. (<http://www.isaaa.org/resources/publications/pocketk/16/>. Accessed 24 April 2017, <http://www.isaaa.org/resources/publications/briefs/51/executivesummary/default.asp>. Accessed 24 April 2017)
- Kim HS, Kim SK, Park YS, Kwon SY, Liu JR, Joung H, Jeon JH (2009) Metabolic profiles of genetically modified potatoes using a combination of metabolite fingerprinting and multivariate analysis. *Biotechnol Bioprocess Eng* 14:738–747
- Krank J, Murphy RC, Barkley RM, Duchoslav E, MsAnoy A (2007) Qualitative analysis and quantitative assessment of changes in neutral glycerol lipid molecular species within cells. *Methods Enzymol* 432:1–20
- Kusano M, Baxter I, Fukushima A, Oikawa Y, Nakabayashi R, Bouvrette DJ, Achard F, Jakubowski AR, Ballam JM et al (2015) Assessing metabolomic and chemical diversity of a soybean lineage representing 35 years of breeding. *Metabolomics* 11:261–270
- Marrelli M, Tudisco R, Mastellone V, Conforti F (2013) A comparative study of phytochemical composition of genetically and nongenetically modified soybean (*Glycine max* L.) and evaluation of antitumor activity. *Nat Prod Res* 27:574–578
- Mavropoulou AK, Koraki T, Ioannou PC, Christopoulos TK (2005) High-throughput double quantitative competitive polymerase chain reaction for determination of genetically modified organisms. *Anal Chem* 77:4785–4791
- Nap JP, Metz PLJ, Escaler M, Conner AJ (2003) The release of genetically modified crops into the environment part I. Overview of current status and regulations. *Plant J* 33:1–18
- Novotna H, Kmiecik O, Galazka M, Krtkova V, Hurajova A, Schulzova V, Hallmann E, Rembalkowska E, Hajslova J (2012) Metabolomic fingerprinting employing DART-TOFMS for authentication of tomatoes and peppers from organic and conventional farming. *Food Additives and Contaminants: Part A* 29:1335–1346
- Ovesna J, Kucera L, Hodek J, Demnerova K (2010) Reliability of PCR based screening for identification and quantification of GMOs. *Czech Journal of Food Science* 28:133–138
- Phipps RH, Park JR (2002) Environmental benefits of genetically modified crops: global and European perspectives on their ability to reduce pesticide use. *J Anim Feed Sci* 11:1–18
- Rajan SR, Letourneau GK (2012) What risk assessments of genetically modified organisms can learn from institutional analyses of public health risks. *J Biomed Biotechnol* 2012:1–8
- Rubert J, Zachariasova M, Hajslova J (2015) Advances in high-resolution mass spectrometry based on metabolomics studies for food—a review. *Food Additives & Contaminants: Part A* 32:1685–1708
- Senior IJ, Dale PJ (2002) Herbicide-tolerant crops in agriculture: oilseed rape as a case study. *Plant Breed* 121:97–107
- Simó C, Ibáñez C, Valdés A, Cifuentes A, García-Canas V (2014) Metabolomics of genetically modified crops. *Int J Mol Sci* 15:18941–18966
- Stewart S, Ivy MA, Anslyn EV (2014) The use of principal component analysis and discriminant analysis in differential sensing routines. *Chem Soc Rev* 43:70–84
- Theodoridis GA, Gika HG, Want EJ, Wilson ID (2012) Liquid chromatography-mass spectrometry based global metabolite profiling: a review. *Anal Chim Acta* 711:7–16
- Triba MN, Moyec LL, Amathieu R, Goossens C, Bouchemal N, Nahon P, Rutledge DN, Savarina P (2015) PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Mol Biosyst* 11:13–19
- Tutelyan VA (2013) Genetically modified food sources safety assessment and control introduction, 1st edn. Elsevier INC, San Diego, pp 18–27
- Ujhelyi G, Vajda B, Béki E, Neszlényi K, Jakab J, Jánosi A, Némédi E, Gelencsér É (2008) Surveying the RR soy content of commercially available food products in Hungary. *Food Control* 19:967–973
- USDA National Agricultural Statistics Service, *June Agricultural Survey for the years 2000–16*. (<https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/recent-trends-in-ge-adoption.aspx>. Accessed 24 April 2017)
- Vaclavik L, Cajka T, Hrbek V, Hajslova J (2009) Ambient mass spectrometry employing direct analysis in real time (DART) ion source for olive oil quality and authenticity assessment. *Anal Chim Acta* 645:56–63
- Vaclavik L, Rosmus J, Popping B, Hajslova J (2010) Rapid determination of melamine and cyanuric acid in milk powder using direct analysis in real time-time-of-flight mass spectrometry. *Journal of Chromatography A* 1217:4204–4211
- Vaclavik L, Zachariasova M, Hrbek V, Hajslova J (2010) Analysis of multiple mycotoxins in cereals under ambient conditions using direct analysis in real time (DART) ionization coupled to high resolution mass spectrometry. *Talanta* 82:1950–1957
- Vaclavik L, Ovesna J, Kucera L, Hodek J, Demnerova K, Hajslova J (2013) Application of ultra-high performance liquid chromatography-mass spectrometry (UHPLC-MS) metabolomic fingerprinting to characterise GM and conventional maize varieties. *Czech Journal of Food Science* 31:368–375

- Worley B, Powers R (2013) Multivariate analysis in metabolomics. *Current Metabolomics* 1:92–107
- Zhang D, Guo J (2011) The development and standardization of testing methods for genetically modified organisms and their derived products. *J Integr Plant Biol* 53:539–551
- Zhao YY, Xiong Y, Curtis JM (2011) Measurement of phospholipids by hydrophilic interaction liquid chromatography coupled to tandem mass spectrometry: the determination of choline containing compounds in foods. *J Chromatogr A* 1218:5470–5479