Plant Metabolomics: Tomato Metabolite Profiling and Identification Employing High Resolution LC-MS Strategies

Helen Welchman[1], David Portwood[2], Mark Earll[2], Mark Seymour[2], Charles Baxter[2], Zsuzsanna Ament[2], Gary Woffendin[1], Martin Hornshaw[1], Michael Athanas[3], Madalina Oppermann[4]

Thermo Fisher Scientific, [1]Hemel Hempstead, UK, [2]Syngenta, Jealotts Hill Int. Research Centre, UK, [3]VAST Scientific, USA, [4]Kungens Kurva, Sweden

Overview

Purpose: Metabolite profiling of a wild-type (WT) and of a ripening inhibited (RI) tomato variety in tomato fruits selected at 4 development stages.

Methods: Metabolite profiling and fragmentation were both performed on a high resolution, high mass accuracy platform. The mass spectrometer was mass calibrated prior to starting the sequence of injections. All data was acquired using external calibration and positive ionization mode.

Results: Masses were measured with high, sub-ppm to max 2 ppm accuracy, leading to identifications based on elemental composition analysis. Metabolite profiles were acquired at high resolution (30 000, 60 000, 100 000) and following chromatographic alignment and peak detection, statistical analysis was performed to discriminate metabolites, which may serve as markers for tomato species or tomato fruit development. Further identifications were carried out via resonance excitation CID experiments and HCD fragmentation. Results obtained on a hybrid system from metabolite profiling and identification experiments provide evidence that the strategies selected can be successfully applied in the LC-MS based detection and identification of metabolites in plant extracts.

Introduction

Food nutritional value, quality, resistance to pathogens, and flavor are among the traits monitored by governments and the food industry alike, in an attempt to promote the creation of robust, healthy, nutrition-rich cultivars that contribute to a sustained agro development. Metabolomics has been identified as a key mass spectrometry-based approach in the analysis of such characteristics. Syngenta is a world-leading agribusiness with a particular interest in seeds and crop protection. Herewith, results from metabolite profiling and identification experiments on tomato samples will be presented.

Methods

Sample preparation: Tomato (*Solanum lycopersicum*) fruits of Ailsa craig (wild-type) and Ailsa craig with the ripening inhibitor mutation (Rin) genotypes were selected at 4 development stages (day 15 or “green”, day 40 or “dark green”, day 47R or “pink” ripe fruit and day 54R or “red” ripe fruit). Whole fruit pericarp was ground in liquid nitrogen and the resulting powdered fruit tissue used for subsequent extractions.
Triplicate biological replicates of both tomato varieties at each of the 4 development stages were analyzed by a minimum of three technical replicates each (30 000 resolving power), with nonaplex sample analysis at high resolving power (60 000). Comparative analyses were carried out using the publications of Gillaspy et al (1993), Carrari and Fernie (2006), Carrari et al (2006), Bovy et al (2007).
Chromatography: Separations were performed on an Acquity® HSS T3 column, 150 x 2.1 mm, 1.8 µm particles (Waters Corporation, USA). A Thermo Scientific Accela pump was used. The mobile phase separation employed isocratic conditions (100% A for 2.5 min) at a flow rate of 220 µL/min, followed by a gradient from 100-90% A for 5 min with the flow rate increasing from 220-350 µL/min, then 90-0% A during 2.5 min, where it was kept isocratically for 2 min, then a fast return, 0.1 min, to 100% A where the column was equilibrated for 3.4 min, and still at 100% A for another 2.5 min during which the flow rate was reduced from 350-220 µL/min prior to the initiation of a new chromatographic separation. Solvent A (water, 0.2% formic acid) and solvent B (98% acetonitrile, 2% water with 0.2% formic acid) were used. The pump was operated at 220-350 µL/min, as described.
Mass spectrometry: MS detection was carried out using a Thermo Scientific LTQ Orbitrap Velos hybrid mass spectrometer (Figure 1) operated in positive mode using high resolution, full MS (*m/z* 85-900) at 30 000 and 60 000 resolving power, with acquisitions carried out at resolution 100 000 for a random, limited selection of samples. For samples measured in full scan at 30 000 resolution, a top 2 DDA method was employed, whereas samples measured at 60 000 resolution employed a top 8 DDA method. Normalized collision energy of 25% for CID was used. In further experiments, an inclusion list was added to the method for validation of some of the identifications. HCD MS^2 data was acquired on a limited sub-set of samples.
Data processing: Alignment, peak detection and metabolite identification based on elemental composition (Thermo Scientific SIEVE software version 1.3 beta with ChemSpider™ interface) were followed by statistical analysis (SIMCA-P™ software, version 12.0. Profiling was performed both in a pair-wise fashion, in a direct comparison of the metabolite profiles of the two tomato varieties, as well as in an assignment-free total cultivar analysis, thus supporting both supervised, as well as non-supervised statistical analysis approaches. CID was particularly useful for metabolite identification based on spectral library searches (Mass Frontier™ software, version 6.0), with DDA experiments and inclusion list-driven MS/MS for metabolite structural determination.

FIGURE 1. Instrument schematic of the LTQ Orbitrap Velos™.

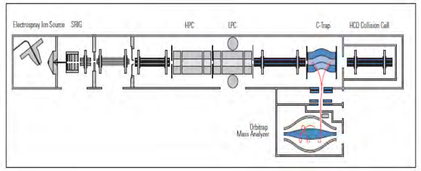
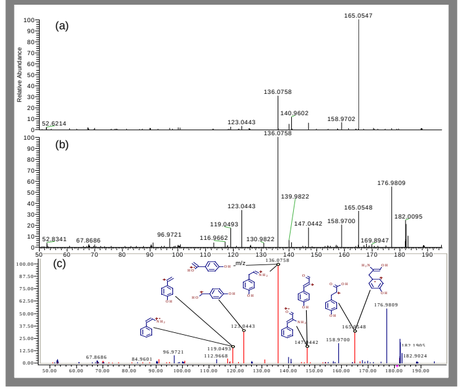


Figure 1 shows the design of the new LTQ Orbitrap Velos. The S-lens design increases the ion flux leading to higher overall sensitivity, and the dual cell linear trap has a high pressure region where ions are isolated, fragmented and scanned out at higher speed towards the lower pressure cell, which promotes a higher resolution detection of the ions at faster duty cycles. The combined C-trap and HCD cell have an increased ion transmission, enabling lower target settings for the fragmentation reaction. The overall improvement in speed and sensitivity leads to a powerful platform that offers efficient profiling, as well as comprehensive structural identification capabilities.

Metabolite profiling was performed with the SIEVE™ software, statistical analysis employed SIMCA-P™ software, while metabolite identification was based on a two-pronged approach: first, accurate mass was used to infer elemental composition, leading to preliminary metabolite identification; secondly, CID and HCD MS^2 product ion data were compared against theoretical fragmentation patterns derived with Mass Frontier™ software. Strong chromatographic and mass spectrometric performance led to reproducible generation of raw data files for processing and mining (Figures 2-6).

FIGURE 2. Fragmentation of tyrosine via different analytical modalities: CID (a); HCD (b); in-silico fragmentation and labeling in Mass Frontier software (c).



Results

Good chromatographic performance was obtained employing short separations, whereby hundreds of components were profiled. In conjunction with robust calibration values obtained via external calibration, metabolites were measured with small retention time drifts at high mass accuracy (Figure 3), leading to strongly suggestive identifications made by elemental composition software. Where theoretical matches were made to more than one elemental composition, additional validation of metabolite identification was provided by isotopic pattern matching in the high resolution full scan data (Figure 4) and by matching the in-silico MS^2 spectrum of the putatively identified metabolite to the observed spectrum (Figures 2, 5). No standards were available at the time of the experiment for confirmatory identification.
PCA along with OPLS-DA analyses were used to distinguished components of interest and potential markers, which were identified employing a mix of accurate mass and MS^2 strategies, with searches in spectral libraries.
However, further analyses and systematic studies are needed to establish direct correlations between composition (fingerprint), individual compounds, and variation attributes such as cultivar or growth stage dependencies. Mass errors registered were well within instrument specifications, with data acquired on the LTQ Orbitrap Velos instrument providing highly accurate mass data along batch acquisitions.

FIGURE 3. Extracted Ion Chromatograms (XIC) using +/- 5 ppm mass tolerance window across 3 batches of 24 LC-MS/MS runs, 72 runs in total. Retention time drifts are negligible, and mass accuracies show a deviation of +/- 1.5 ppm for the compound displayed below.

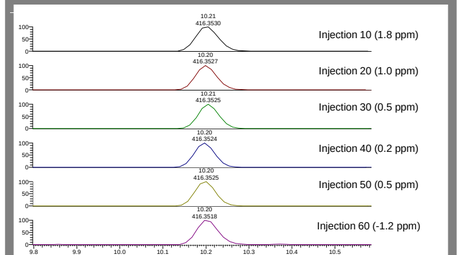


FIGURE 4. SIEVE-ChemSpider profiles and putative identification of metabolite leading to 2 proposed elemental compositions (a). High resolution (100 000) isotopic pattern verification of the 2 matches with zoom on A2 isotope of both proposed structures further validates the correct identification of glutathione.

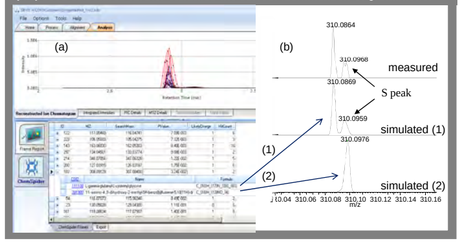


FIGURE 5. Non-differential analysis profiles of tomatidine in SIEVE software across the 72 samples acquired at 60 000 resolution reveals variations associated with a particular stage of fruit ripening (a); this stage can be distinguished by PCA analysis (b), while a comparison of the two varieties when the fruits are at day 54R (c), based on the good chromatographic performance obtained, reveals an interesting area around RT 10.2 min in the non-aligned chromatograms diagram. Bar chart of tomatidine concentrations (non-normalized, (d) and volcano plot with ratio calculation of relative abundance for aforementioned component (e). Built-in link to ChemSpider for putative compound identification (f), further verified by spectral matching of in-silico to measured spectrum and annotation in Mass Frontier software (g).

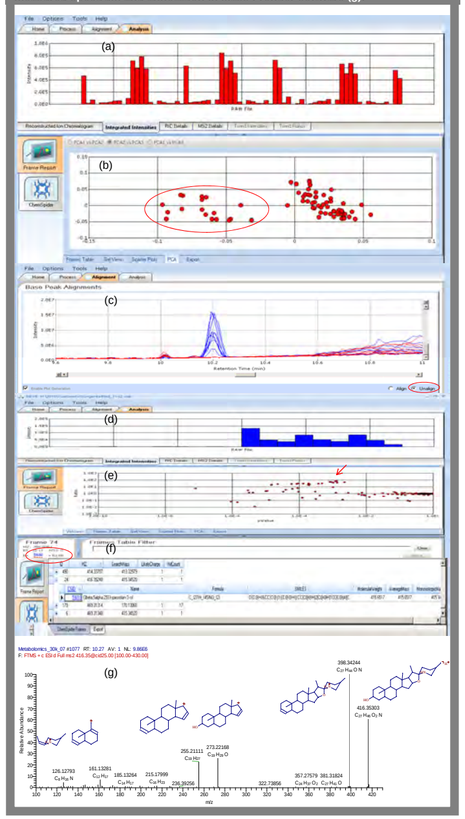
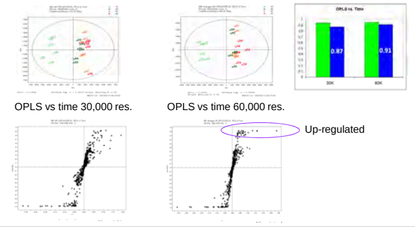


FIGURE 6. A comparison of OPLS versus time shows an improvement in predictive variance with the 60 000 resolution data. The high resolution data detects more variables reliably increasing with ripening than the data set acquired at lower resolution settings.



Conclusions

•Sample preparation for global metabolite discovery can be performed efficiently, with hundreds of features profiled using SIEVE software.

• Chromatographic conditions employed a short gradient, resulting in peak widths <0.3 min; such conditions were useful for both metabolite profiling and for metabolite identification experiments.

• Resolutions of 30 000 and of 60 000 were used for metabolite fingerprinting, enhancing discrimination of components, as well as metabolite identification based on accurate mass measurements, for components identified in publicly available databases. The higher resolution data provided a model with slightly enhanced predictive variance (Q2) and additional upregulated markers.

• Statistical analysis and modeling were carried out using SIMCA-P software. Orthogonal Partial Least Squares (OPLS) models were built with ripening time as the Y vector. The predictive merit of each model was verified by cross-validation, permutation testing and CV-Anova.

• Metabolite identification, a crucial component in metabolomics experiments, was performed using a two-stage approach:

1. Accurate mass determination generating elemental composition within a narrow mass tolerance window for identification based on accurate precursor masses.
2. CID mainly, and HCD product ion data, matching against theoretical fragmentation patterns derived with Mass Frontier software.

• The methods employed in this project led to the distinction of features which differentiate the tomato cultivars employed, as well as the developmental stages of the tomato fruits.

References

(1) Gillaspy, G. et al(1993) *The Plant Cell* **5**: 1439-51.
(2) Carrari, F. and Fernie, A.R. (2006), *Journal of Experimental Botany*, **57**: 1883-97.
(3) Carrari, F. et al (2006), *Plant Physiology*, **142**: 1380-96.
(4) Bovy, A. et al. (2007), *Metabolomics*, **3**: 399-412.

Acknowledgements

We would like to thank Bernard Delanghe, Christoph Heinrich and the software team from Thermo Fisher Scientific in Bremen, Germany, for their support.

The authors wish to acknowledge Graham Seymour and Charlie Hodgman from Nottingham University for the provision of tomato samples used in this study.