Robust Metabolite Profiling and Identification Employing High Resolution MS Strategies and Dedicated Software

Helen Welchman[1], David Portwood[2], Mark Earll[2], Mark Seymour[2], Charles Baxter[2], Martin Hornshaw[1], Madalina Oppermann[3], Graham Seymour[4], Charlie Hodgman[4]

Thermo Fisher Scientific, [1]Hemel Hempstead, UK and [3]Kungens Kurva, Sweden; [2]Syngenta, Bracknell, UK; [4]Nottingham University, Nottingham, UK

Overview

Purpose: The purpose of this study was to highlight metabolic differences between different cultivars of tomato through development and ripening, as part of a study to understand the metabolic and genetic basis of ripening. Metabolite profiling of a wild-type (WT) and of three ripening inhibited tomato varieties in tomato fruits selected at 13 development stages was performed.

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 30000 resolution 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 collision-induced dissociation (CID) experiments and higher energy collision-induced dissociation (HCD) experiments. Using multivariate analysis, differences in metabolic trajectories have been observed for the tomato cultivars, and some of the key metabolite differences have been identified. 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

Four varieties of tomato were grown and sampled over the period of flowering, fruiting and ripening as part of an **"Exploiting Systems Biology"** project. Polar and non-polar extracts were obtained from the tomatoes and subjected to an untargeted metabolomics study using LC-MS/MS. The study compared three non-ripening varieties with a normal (wild type) tomato during fruit development. Development stages were 15, 20, 25, 30 and 40 days after flowering, breaker stage (BR), when fruit starts to first turn color (approx 47 days +/- 1 day), BR 1, 2, 3, 4, 5, 6 and 7. Whole fruit pericarp was ground in liquid nitrogen, and the resulting powdered fruit tissue was used for subsequent extractions. Triplicate analytical replicates of all four tomato varieties at each of the 13 development stages were analyzed. 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 2.0 mm, 1.7 µm particles (WATERS®, USA). An ACQUITY UPLC® system was used (WATERS, USA). Solvents used were 0.2% formic acid in water (Solvent A) and 98/2/0.2 acetonitrile/water/formic acid (Solvent B). The gradient started at 100% A (hold 2.5 minutes) at 0.25 mL/min followed by a gentle ramp to 10% B after 7.5 minutes, increasing the flow rate of 0.4 mL/min; then to 100% B after 10 minutes, hold 2 minutes before equilibrating back to starting conditions after 18 minutes.

Mass Spectrometry

MS detection was carried out using a Thermo Scientific LTQ Orbitrap Velos hybrid mass spectrometer operated in positive mode, full MS (*m/z* 85-900) at 30000 resolving power. MS^2 measurements were data dependent analysis (DDA)-driven for metabolite MS^2 confirmatory purposes, top 3 selection, and HCD fragmentation with normalized collision energy of 50.

Data Analysis

SIEVE™ software version 1.3 with ChemSpider™ interface was used for alignment, peak detection and metabolite identification based on elemental composition.

Component Elucidator algorithm (Technology Preview, prototype version 1.0, build 19) was also employed for the generation of a full list of components quantified across the entire data set. SIMCA-P™ software, version 12.0, was used for statistical analysis. Profiling was performed both in a pair-wise fashion, in a direct comparison of the metabolite profiles of either two tomato varieties or for the same tomato variety, comparing two development stages. Mass Frontier™ software, version 6.0 was used for confirmatory MS/MS metabolite structural determination.

FIGURE 1. Workflow employed.

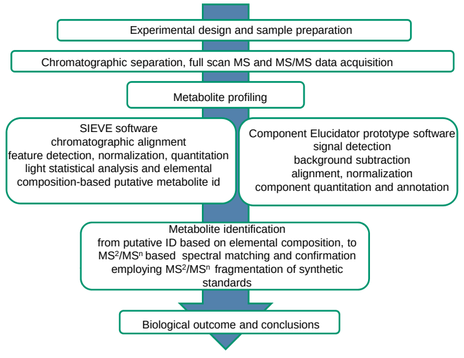
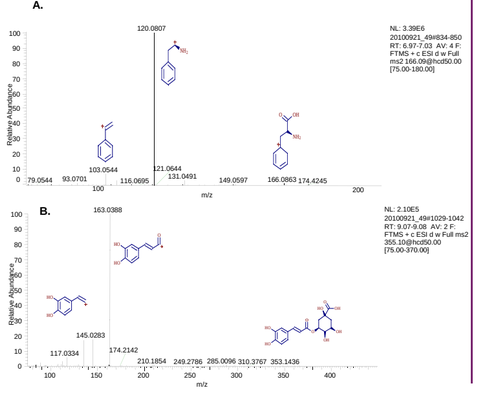


FIGURE 2. HCD fragmentation and labeling of (A.) phenylalanine and (B.) chlorogenic acid. Both compounds were fragmented in one of the wild-type tomato samples.



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, leading to strongly suggestive identifications made by elemental composition software (Figs. 3-5). Additional validation of metabolite identification was provided by matching the in-silico MS^2 spectrum of the putatively identified metabolite to the observed spectrum (Fig. 2). In some instances, standards were available for confirmatory identification of both MS^2 fragmentation patterns, as well as for matching of chromatographic retention times.

FIGURE 3. Glutamic acid profiling, quantification and identification.

(A) Its presence detected with Component Elucidator in a tomato QC sample, where the top panel is a TIC chromatographic profile, the middle panel displays the full MS acquired at the selected time (1.53 min), and the table in the lower panel shows *m/z*, intensity and S/N corresponding to glutamic acid.

(B) In SIEVE, in a comparison between wild-type samples at breaker vs breaker +7 days, where the top panel represents the integrated intensities measured across triplicates in the two groups, the middle panel represents a Volcano plot of ratio between groups (calculated at 3.67) vs p-value, and the lower panel represents putative metabolite identification in ChemSpider, specifically in the KEGG database.

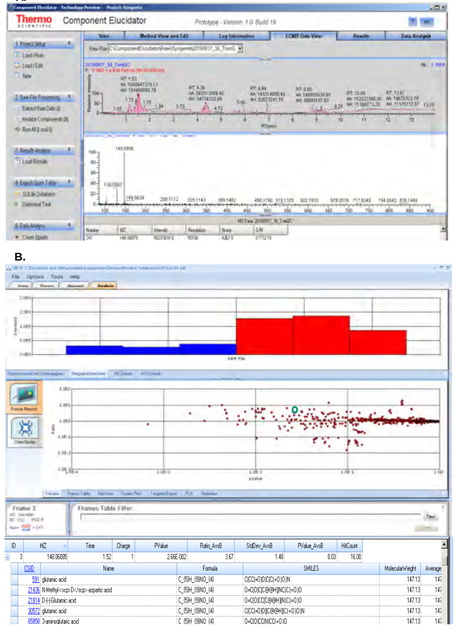


FIGURE 4. Metabolite profiling of wild-type tomatoes, breaker vs breaker +7 days.

A. PCA plot.

B. Table highlighting some of the differences detected between samples.

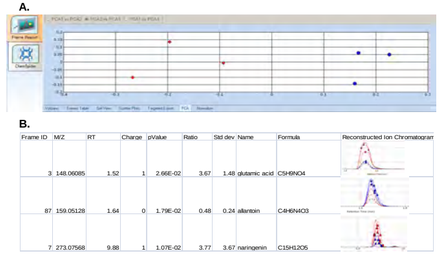


Figure 5. Metabolite profiling of wild-type vs ripening-inhibited tomato varieties at breaker stage.

A. Chromatographic alignment and scoring of breaker stage of two varieties.

B. Table highlighting some of the differences detected between samples at breaker stage, where triangles in the Reconstructed Ion Chromatogram reflect MS^2 acquisitions.

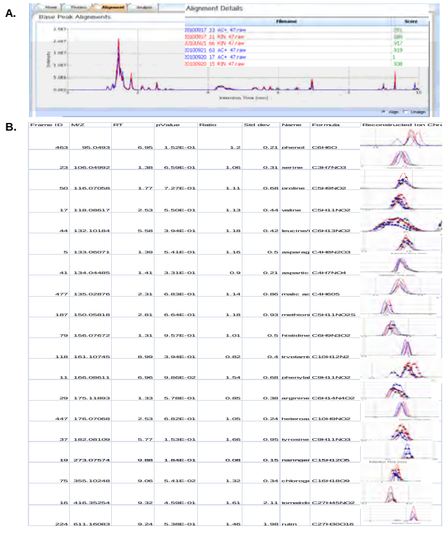
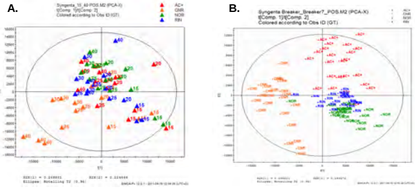


FIGURE 6. SIMCA-P results of tomato variety discrimination in (A.) pre-breaker and (B.) post-breaker stages.



Conclusion

- The workflow employed (Fig. 1) was efficient in the study of both developmental and cultivar-related metabolic differences associated with tomato fruit ripening.
- Background subtraction is a useful and desirable feature; employing Component Elucidator algorithm in the total analysis of 9 blank runs, 44 quality control (QC) runs, 10 runs containing tomato pericarp mix, and the 156 samples which represent the four tomato varieties at 13 different development stages, as triplicate runs, led to the generation of a component table listing 319 potential components. While data evaluation and validation of interesting metabolite identities is on-going, quantitation consistency between QC and tomato mix samples is a feature observed.
- Chromatographic conditions employed short gradients, resulting in peak widths of under 30 s; such conditions were useful for both metabolite profiling and for metabolite identification experiments.
- Resolution of 30000 was used for metabolite fingerprinting, with good discrimination of components, as well as metabolite identification based on accurate mass measurements, for components identified in publicly available databases.
- Statistical analysis and modelling was carried out using SIMCA-P software from Umetrics and hotelling ellipses of pre-breaker and post-breaker stages visually described feature discrimination between the four tomato varieties used.
- Metabolite identification, a crucial component in metabolomics experiments, was performed using two approaches:
  1. Accurate mass determination generating elemental composition within a narrow mass tolerance window (0.001 Da) for identification based on accurate precursor masses.
  2. MS^2 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), *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 colleagues at Thermo Fisher Scientific, including Mark Sanders, David Peake and Thomas McClure, as well as collaborators at Bristol-Myers Squibb, for their support and constructive proposals regarding prototype software employment.