Advancing coffee plant metabolomics: A comparative QTOF study

Andrea Montis[1,2], Victoria Paredes-Orejudo[1,2], Axelle Bourez[1,2], Cedric Delporte[1,2], Florence Souard[1,2], Pierre Van Antwerpen[1,2], Jack Steed[3], Ernst Bouvin[4] and Jianru Stahl-Zeng[5]

[1]RD3 – Pharmacognosy, Bioanalysis and Drug Discovery, Belgium; [2]Université libre de Bruxelles, Belgium; [3]SCIEX, UK; [4]SCIEX, Belgium; [5]SCIEX, Germany

Metabolic profiling of plant tissues, such as leaves, flowers and phloem, can provide valuable insights into the diverse array of chemical compounds, or metabolites, that influence flavor, scent and metabolic processes. Here, we present a comprehensive characterization study that focuses on *Coffea arabica* L. (ARA) and *Coffea anthonyi* Stoff. (ANTH). Currently, information regarding the chemical composition of ANTH is limited. Two quadrupole time-of-flight (QTOF) mass spectrometry systems, the QTOF A system and the SCIEX ZenoTOF 7600 system, were employed for metabolite identification and molecular networking analysis. Our results demonstrate superior performance of the ZenoTOF 7600 system, highlighting its efficacy in untargeted metabolomics studies on plants. This study expands our understanding of the metabolic diversity within coffee species and underscores the utility of advanced mass spectrometry platforms to elucidate metabolite profiles essential for flavor development and metabolic pathway exploration.

Key benefits of the ZenoTOF 7600 system for coffee plant metabolomics

- **Single-injection workflow:** A single injection on the ZenoTOF 7600 system allows the user to collect both TOF MS and TOF MS/MS data. In comparison, the QTOF A system requires 2 separate injections.
- **Sensitivity gains:** The presence of the Zeno trap provides a 5–20x gain in MS/MS sensitivity, resulting in fewer losses of biorelevant molecules compared to the QTOF A system
- **Superior metabolite coverage:** In all the data processing methods employed, the ZenoTOF 7600 system provided additional insights into the samples analyzed compared to the QTOF A system (Figure 1)
- **Simple data conversion:** Quick and easy data conversion from *.wiff to *.mzXML format, allowing for third-party data processing to be utilized

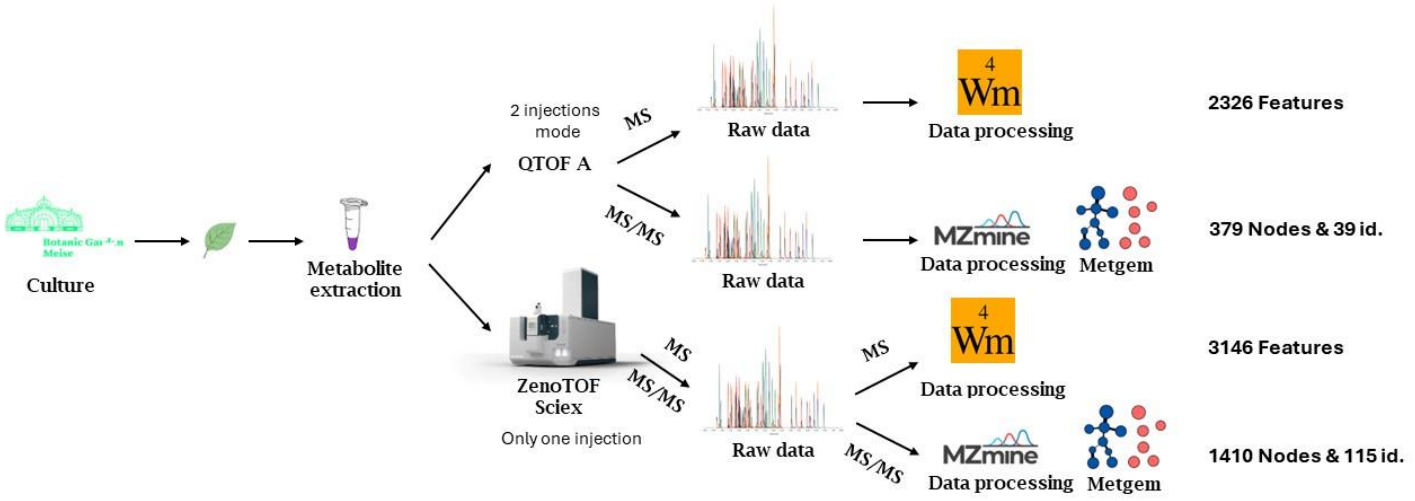

Figure 1. Step-by-step workflow for the metabolic characterization of coffee leaves by LC-MS and LC-MS/MS. In this analysis, the QTOF A system and the ZenoTOF 7600 system were compared. The ZenoTOF 7600 system allows for a single-injection workflow and has a higher scan speed compared to the QTOF A system. The QTOF A system has a slower scan speed and requires at least 2 injections to achieve comprehensive TOF MS and TOF MS/MS coverage.

Introduction

ANTH is a diploid species endemic to southeast Cameroon and northwest Republic of the Congo. ANTH exhibits morphological resemblances to *C. eugenioides*, the maternal progenitor of ARA.[1-2] ARA is a prominent contributor to the coffee market and predominantly inhabits mountainous and sub-mountainous forests ranging from 1000–2000 m in the eastern regions of the Democratic Republic of Congo, Uganda, Kenya, Rwanda, Burundi and southern Sudan. Whereas ANTH thrives in low-altitude forest habitats (360–650 m), ARA demonstrates optimal growth at high altitudes exceeding 1000 m.[3] Given the current trajectory of climatic changes, significant portions of high-altitude forests, including habitats of endemic coffee species such as ARA, might be at risk. ANTH can be considered a promising "lowland" alternative to the "highland" ARA for the production of coffee beverages or food supplements.

In this study, we analyze and compare coffee leaves from these 2 species using 2 QTOF platforms: the QTOF A system and the ZenoTOF 7600 system. We focus on the features and fingerprints obtained, which are representative of the chemical composition profile acquired.

Methods

Samples and reagents: Ultra-high purity water was extracted by filtration using a Milli-Q system from Millipore. LC-MS quality acetonitrile, formic acid (FA) and trifluoroacetic acid (TFA) were purchased from Fisher Scientific. Coffee samples (leaves) were collected in the tropical green house of Meise Botanic Garden. Samples were conserved in Meise Botanic Garden for ARA and ANTH plants.

Sample preparation: Four ARA and 4 ANTH samples were collected. Each sample was collected from a different plant and prepared as described below.

Collected leaf samples were dried by packing them in sealed plastic bags filled with silica gel. The silica gel was replaced every day for 1 week. After 1 week, 1.0 g of leaves and 100.0 mg of petals were powdered for each sample with an ULTRA-TURRAX Tube Drive control (Q-Lab, Vilvoorde, Belgium). The powdering time was 10 minutes at 6000 rotations per minute for the leaves. Fruits (1.0 g) and flower gynoeciums (100.0 mg) were ground for 20 minutes after briefly grinding with mortar and pestle. The duration of grinding was defined to obtain a homogeneous powder. The extraction was performed by suspending 15.0 mg of powdered samples in 1.5 mL of Milli-Q water for 5 min in a 55 kHz ultrasonic bath.[4] Between 4–6 technical replicates were prepared for the leaves and fruits of each species. Eight technical replicates were prepared for the flowers. Samples were filtered through a 0.2 µm cellulose acetate membrane and stored at −20°C until LC-MS analysis.

Chromatography: A Poroshell 120 EC-C18 column (2.7 µm, 100 mm × 2.1 mm) with guard column (2.7 µm, 5 mm × 2.1 mm) from Agilent was used to carry out the chromatographic separation. The column temperature was set to 55.0°C. The mobile phases were composed of 0.025% TFA and 0.075% FA in water (solvent A) and in acetonitrile (ACN, solvent B) for positive ion mode analysis. The applied gradient started at 0% solvent B and from 0–5 min, ramped to 10% solvent B. From 8–9 min, solvent B increased from 10% to 12.5% and from 9–11 min, solved B increased from 12.5% to 15%. From 11–17 min, solvent B increased from 15% to 80% and from 17–18 min increased to 100%. This 100% solvent B concentration was maintained from 18–19 min and then was reduced to 0% from 19–20 min. The post-run cycle lasted 8 min at 0.5 mL/min. The injection volume was 10 µL.

Mass spectrometry: Analysis was performed using the QTOF A system and the ZenoTOF 7600 system. The data were acquired in positive ionization mode. The source conditions used are outlined in Table 1. TOF MS and TOF MS/MS conditions used for data-dependent acquisition (DDA) experiments are listed in Tables 2 and 3, respectively. The conditions used for SWATH data-independent acquisition (DIA) experiments are shown in Table 4.

Table 1. Source conditions and gas parameters.

Parameter	Value
Curtain gas	*35 psi*
Collision gas	*7 psi*
Ion spray voltage	*5500 V*
Temperature	*350°C*
Gas 1	*50 psi*
Gas 2	*60 psi*

Table 2. TOF MS parameters for DDA experiments.

Parameter	Value
TOF MS start mass	*100 Da*
TOF MS stop mass	*1700 Da*
Accumulation time	*0.1 s*
Declustering potential	*80 V*
Collision energy	*10 V*

Table 3. TOF MS/MS conditions for DDA experiments.

Parameter	Value
TOF MS/MS start mass	*50 Da*
TOF MS/MS stop mass	*1000 Da*
Accumulation time	*0.08 s*
Declustering potential	*80 V*
Collision energy	*35 V*
Collision energy spread	*15*

Table 4. Conditions for variable window SWATH DIA experiments.

Parameter	Value
Variable Q1 windows	*10*
SWATH DIA start mass	*100 Da*
SWATH DIA stop mass	*1000 Da*
Declustering potential	*50 V*
Collision energy	*35 V*
Collision energy spread	*15*
TOF MS accumulation time	*0.1 s*
TOF MS/MS accumulation time	*0.04 s*

Data processing: ProteoWizard MSConvert tools (version 3.03.9393, 64-bit) were used to convert raw data to *.mzXML file format. Peak Picking was chosen as a filter option. All data pre-processing and data processing were performed on the Workflow4Metabolomics (W4M) infrastructure (https://workflow4metabolomics.org). Data pre-processing was performed in XCMS software, which allows the user to extract ions and align them across samples. This software also enables peak grouping, retention time correction and the annotation of isotope peaks, adducts and fragments. The Matched Filter algorithm was used with the parameters adapted for the QTOF A system.[5] Filtering and normalization were then applied. Noise signal was estimated using blanks (injection solvent) as a reference. Additional detailed steps and parameters that were used for data processing are publicly available on the W4M workflow repository (https://usegalaxy.fr/u/amontis/h/coffee-published-2022).

Unsupervised multivariate analysis was carried out on the W4M platform by performing a principal component analysis (PCA), which provided a prior overview of cluster separation. The parameters used included: 2 predictive components, pareto scaling and 4–7 cross-validation segments. The variables responsible for the separation between groups were initially provided by the loadings plots and then confirmed using the Biosigner tool with a seed value from 1 to 9. Biosigner is a tool that includes 3 different binary classifiers: PLS-DA (partial least square discriminant analysis), RF (random forest) and SVM (support vector machines). These classifiers can be used to select the features that were most discriminant for each sample.[6]

Comparative performance results

The goal of this study was to exclusively elucidate the distinctions between 2 analytical instruments by maintaining identical LC conditions and employing the same Poroshell 120 EC-C18 column. Coffee samples were subjected to analysis using both the ZenoTOF 7600 system and the QTOF A system. Specifically, 4 samples of ARA leaves and 4 samples of ANTH leaves were selected and analyzed using both instruments.

The datasets obtained from both analyses were processed utilizing the Galaxy W4M online platform. The data processing procedure included ion extraction, alignment, peak grouping, retention time correction and identification of isotopes, adducts and fragments. Data processing yielded 3277 extracted features for samples analyzed with the QTOF A system and 5668 extracted features for samples analyzed with the ZenoTOF 7600 system. Subsequent normalization procedures resulted in 2326 remaining metabolites for the QTOF A system and 3146 remaining metabolites for the ZenoTOF 7600 system (Table 5). These identified metabolites underwent multivariate and univariate statistical analyses for further evaluation.

The PCA revealed that the amount of variability described by PC1 and PC2 was comparable between the QTOF A system and ZenoTOF 7600 system datasets (Figure 2).

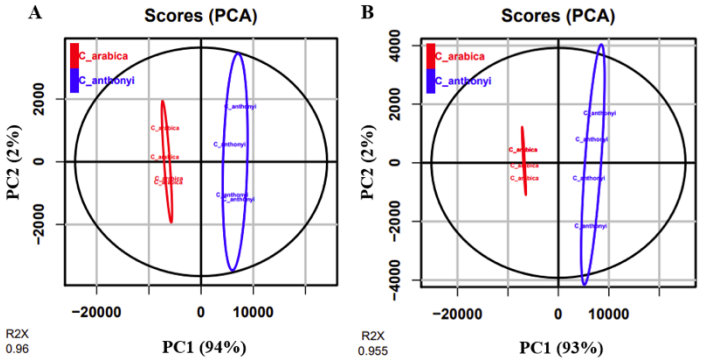

Figure 2. PCA score plots generated using 2 principal components to show the inter-species separation of the ARA and ANTH leaves. Samples were analyzed in positive electrospray ionization mode using the QTOF A system (A) or the ZenoTOF 7600 system (B). R^2X scores are shown in parentheses.

The loadings plots corresponding to the PCA plots revealed 11 metabolites for both datasets, albeit with some disparity in their composition (Figure 3). This observation indicates that while certain metabolites were identified as primarily contributing to the interspecies segregation, the specific composition of these metabolites varied between analyses conducted with the QTOF system A and the ZenoTOF 7600 system.

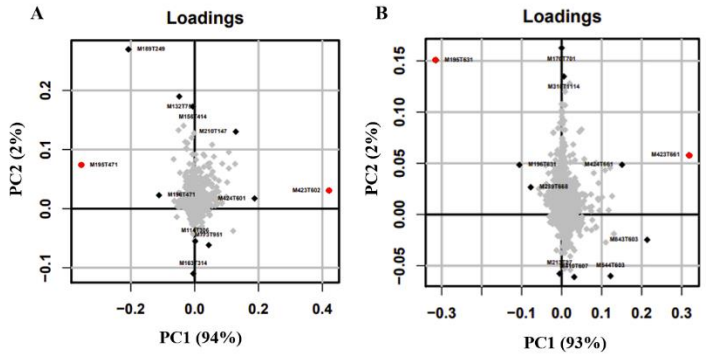

Figure 3. Loadings plots showing the main metabolites responsible for the separation of ARA and ANTH leaves. The samples were analyzed with the QTOF A system (A) or the ZenoTOF 7600 system (B) in positive electrospray ionization mode. The red data points indicate the metabolites that were annotated and confirmed by supplementary statistical analyses.

A PLS-DA was conducted and the metadata were filtered based on the most discriminant metabolites. Metabolites with a variable importance in the projection (VIP) value >2 were retained. Following filtration, 75 metabolites remained for the samples analyzed with the ZenoTOF 7600 system, whereas 47 metabolites remained for samples analyzed with the QTOF A system. Table 5 reveals a higher number of discriminant metabolites in the ZenoTOF 7600 system dataset. Notably, among the 75 metabolites identified with the ZenoTOF 7600 system, 33 were also detected with the QTOF A system, while the remaining 42 metabolites were exclusively detected with the ZenoTOF 7600 system.

Metabolic networking analyses were then performed to better visualize the ions detected with both instruments and to perform some annotation. Molecular networking is a bioinformatic approach that allows a global interpretation of data. Figures 4 and 5 show the molecular network outputs obtained using the QTOF A system and ZenoTOF 7600 system, respectively.

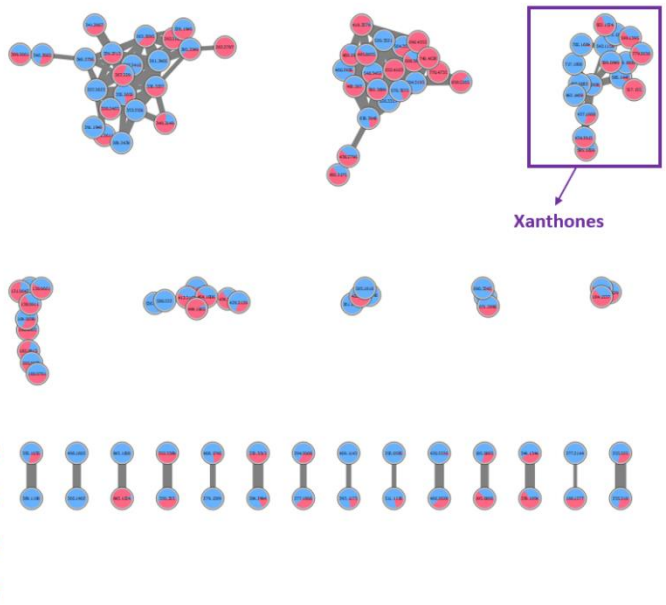

Figure 4. Molecular network obtained with the QTOF A system. Three major clusters and 5 minor clusters are shown. The isolated nodes have been filtered. Each circle indicates the relative abundance of the species in the ANTH (blue) and ARA (red) samples. The purple box indicates the cluster of xanthones.

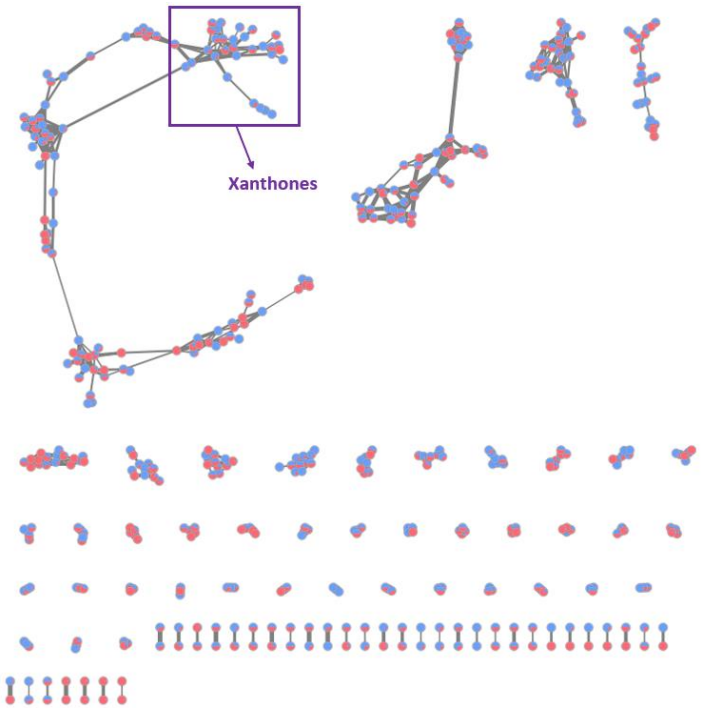

Figure 5. Molecular network obtained with the ZenoTOF 7600 system. Four major clusters and 39 minor clusters are shown. The isolated nodes have been filtered. Each circle indicates the relative abundance of the species in the ANTH (blue) and ARA (red) samples. The purple box indicates the cluster of xanthones.

In the constructed networks, 379 nodes were observed for the data collected on the QTOF A system, whereas 1410 nodes were identified for the data collected on the ZenoTOF 7600 system. The compound database from MetGem proposed 39 potential chemical structures for the molecular networks generated for the QTOF A system data and 115 for those generated with the ZenoTOF 7600 system data (Table 5). Annotation required further verification to confirm that standard compounds corresponded to the detected molecules. The larger number of compounds detected with ZenoTOF 7600 system data suggests that molecules previously isolated in the molecular networks generated from the QTOF A system might be linked within clusters in the molecular networks generated from the ZenoTOF 7600 system. For example, clusters containing xanthones were notably larger in the molecular networks generated from the ZenoTOF 7600 system data, indicating the availability of more spectral data for effective annotation. Notably, ion m/z = 239.1503, identified as primicarb, was linked to other nodes in the ZenoTOF 7600 system molecular network but appeared as an isolated node in the QTOF A system molecular network. Similarly, ion m/z = 391.1025, previously identified as a benzophenone derivative, and ion m/z = 289.0706, identified as dihydrokaempferol, exhibited similar differences in linkage between the molecular networks generated from the QTOF A system and ZenoTOF 7600 system data. These findings indicate the capability of the ZenoTOF 7600 system to detect metabolites even at low intensities, as demonstrated by the detection of dihydrokaempferol in ANTH samples, in which its presence was previously undetected by the QTOF A system. These differences are summarized and reported in Table 5.

Conclusions

The ZenoTOF 7600 system demonstrated superior suitability for untargeted metabolomics, as indicated by the higher number of metabolites detected compared to the QTOF A system (Table 5). This enhanced detection capability provides a more comprehensive overview of metabolite content in samples, aligning with the aim of untargeted metabolomics to capture the entire metabolome for maximal information on inter-group differences. Importantly, analysis with the ZenoTOF 7600 system revealed unique metabolites that were undetectable using the QTOF A system, enhancing the potential for annotation and identification of discriminant metabolites. While comparing results between the 2 instruments is valuable, focusing on metabolites identified as discriminant by both the QTOF A system and the ZenoTOF 7600 system is particularly insightful.

Notably, 3 metabolites (m/z 195.0875, m/z 355.1025 and m/z 423.0928) emerged consistently across both platforms, reinforcing the significance of compounds within this family. Specifically, these metabolites corresponded to caffeine, 5-CQA and mangiferin, which were previously identified as key differentiators between ARA and ANTH leaf samples using the QTOF A system.

The ZenoTOF 7600 system offers the advantage of conducting untargeted metabolomics studies and fragmentation for molecular networking in a single run, whereas the QTOF A system requires 2 separate injections. This feature reduces solvent usage and time taken in performing the analysis. Comparing metabolomic studies conducted with the QTOF A system and the ZenoTOF 7600 system revealed differences in post-processing results. While both systems provide general insights into interspecies differences in coffee, the ZenoTOF 7600 system enables a more comprehensive metabolite list

for further investigation of species-specific markers. Selection of the instrument should therefore be based on the desired level of information. The ZenoTOF 7600 system detects a greater number of metabolites both pre- and post-processing, leading to increased identification of discriminant metabolites and warranting targeted metabolomic studies.

Table 5. Comparison of LC-MS and LC-MS/MS outputs of coffee leaf samples.

Differential parameters	QTOF A system	ZenoTOF 7600 system
W4M LC-MS analysis		
Features after pre-processing	*3277*	*5668*
Features after signal drift correction	*2326*	*3146*
Features contributing to PLS-DA (VIP >2)	*47*	*75*
MetGem LC-MS/MS analysis (cosine score >0.5)		
Nodes	*379*	*1410*
Isolated nodes	*260*	*932*
Identified molecules	*39*	*115*
Xanthone cluster nodes	*17*	*31*

References

1. Stoffelen P, Noirot M, Couturon E, et al. Coffea anthonyi, a new self-compatible Central African coffee species, closely related to an ancestor of Coffea arabica. TAXON 2009; 58: 133–140. doi:10.1002/tax.581014

2. Vanden Abeele S, Janssens SB, Asimonyio Anio J, et al. Genetic diversity of wild and cultivated Coffea canephora in northeastern DR Congo and the implications for conservation. Am J Bot 2021; 108: 2425–2434. doi:10.1002/ajb2.1769

3. Nugroho D, Basunanda P, Mw S. Physical Bean Quality of Arabica Coffee (Coffea Arabica) Cultivated at High and Medium Altitude. Pelita Perkeb Coffee Cocoa Res J 2016; 32: 151–161. doi:10.22302/iccri.jur.pelitaperkebunan.v32i3.241

4. Alonso-Salces RM, Serra F, Reniero F, et al. Botanical and geographical characterization of green coffee (Coffea arabica and Coffea canephora): chemometric evaluation of phenolic and methylxanthine contents. J Agric Food Chem 2009; 57: 4224–4235. doi:10.1021/jf8037117

5. Smith C, Want E, O'Maille G, et al. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. Anal Chem 2006; 78. doi:10.1021/ac051437y

6. Rinaudo P, Boudah S, Junot C, et al. biosigner: A New Method for the Discovery of Significant Molecular Signatures from Omics Data. Front Mol Biosci 2016; 3: 26. doi:10.3389/fmolb.2016.00026

Headquarters
500 Old Connecticut Path | Framingham, MA 01701
USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please all the division headquarters or refer to our website at
sciex.com/offices