

American University in Cairo

## AUC Knowledge Fountain

---

Theses and Dissertations

Student Research

---

Summer 6-15-2021

### Comparative metabolites profiling and fingerprinting via a multiplex approach of UV and LC/MS of coffee seeds in the Middle East and in relation to their antioxidant effects

Enas Elhawary  
enaselhawary17@aucegypt.edu

Follow this and additional works at: <https://fount.aucegypt.edu/etds>

 Part of the [Food Science Commons](#)

---

#### Recommended Citation

##### APA Citation

Elhawary, E. (2021). *Comparative metabolites profiling and fingerprinting via a multiplex approach of UV and LC/MS of coffee seeds in the Middle East and in relation to their antioxidant effects* [Master's Thesis, the American University in Cairo]. AUC Knowledge Fountain.

<https://fount.aucegypt.edu/etds/1651>

##### MLA Citation

Elhawary, Enas. *Comparative metabolites profiling and fingerprinting via a multiplex approach of UV and LC/MS of coffee seeds in the Middle East and in relation to their antioxidant effects*. 2021. American University in Cairo, Master's Thesis. *AUC Knowledge Fountain*.

<https://fount.aucegypt.edu/etds/1651>

This Master's Thesis is brought to you for free and open access by the Student Research at AUC Knowledge Fountain. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of AUC Knowledge Fountain. For more information, please contact [thesisadmin@aucegypt.edu](mailto:thesisadmin@aucegypt.edu).

**Comparative metabolites profiling and fingerprinting via a  
multiplex approach of UV and LC/MS of coffee seeds in the  
Middle East and in relation to their antioxidant effects**

A thesis submitted by

**Enas Ali Elhawary**

For the master's degree in chemistry with a concentration in food chemistry

Under the supervision of

**Prof. Dr. Mohamed Ali Farag**

Chemistry Department, AUC  
Pharmacognosy Department  
Faculty of Pharmacy  
Cairo University

**Comparative metabolites profiling and fingerprinting via a  
multiplex approach of UV and LC/MS of coffee seeds in the Middle  
East and in relation to their antioxidant effects**

**A Thesis Submitted by**  
**Enas Ali Elhawary**

**to the**  
**School of Sciences and Engineering**  
**Master program**  
**Has been approved by**

**Prof. Dr. Mohamed Ali Farag (Supervisor)**

Professor, Department of Chemistry, AUC

\_\_\_\_\_

**Dr. Tamer El-Idreesy (Internal Examiner)**

Associate Professor, Department of Chemistry, AUC

\_\_\_\_\_

**Dr. Sherweit H. El-Ahmady (External Examiner)**

Professor, Department of Pharmacognosy, Ain shams university

\_\_\_\_\_

**Dr. Ehab Elsawy**

\_\_\_\_\_

Associate Professor, Department of Chemistry, AUC

**Graduate Program Director**

**Date**

**School Dean**

**Date**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## **ACKNOWLEDGEMENT**

First of all, thanks to Allah by the grace of whom this work was accomplished.

I would like to express my deepest gratitude to Professor Dr. Mohamed Ali Farag for supervising this work and valuable guidance during this study; without Dr Farag's supervision and persistence help, this research would not have been possible. I am thankful to all my professors at AUC's chemistry department, whom I have learnt a lot from throughout the course of my study.

No words could express my gratitude and deep appreciation to my family, especially my husband, Walid Shehata, to whom words are not enough to describe his care and support. I would like to extend my gratitude to my dear beloved father, Dr. Ali Elhawary and mother, Dr. Fatma Omar, for their continuous support and encouragement. I am deeply thankful for my beloved sisters, who spread no effort to give me all the assistance and love. Moreover, I would like to express my sincere appreciation to my lovely children Laila, Karma and Adham for their understanding during the periods I was busy with my research.

## Abstract

Coffee is a worldwide beverage of increasing consumption owing to its unique flavor and several health benefits. The current research aimed at investigating multiple bioactive compounds in coffee belonging to the two major species of *Coffea*, including *C. arabica* and *C. canephora* or *C. Robusta*, either green or roasted, along with 15 commercial specimens collected from the Middle East. Two platforms were employed, including ultra-high-performance liquid chromatography coupled to mass spectrometry (UHPLC/MS) and UV for profiling and fingerprinting, respectively. Identification of the UPLC/MS dataset was aided by using the global natural product social molecular networking (GNPS). A total of 91 peaks were annotated belonging to different metabolites classes, including 23 hydroxycinnamates (*i.e.*, chlorogenic acid), 5 hydroxycinnamic amides (*i.e.*, Coumaroyl tryptophan), 17 diterpenes (*i.e.*, Cafestol), 15 fatty acids (*i.e.*, Trihydroxy-octadecanoic acid), 7 sphingolipids, 6 nitrogenous compounds (*i.e.*, *N*-tricosanoyl-hydroxytryptamine), 2 alkaloids (*i.e.*, Caffeine), 2 sugars (*i.e.*, Acetyl-di-feruloyl sucrose), 1 coumarin (dihydroxypsoralen-*O*-hexoside), and 1 fatty acid amides (docosenamide). Several metabolites are reported for the first time in coffee, including novel hydroxycinnamates (5), sphingolipids (4) and ceramides (4), diterpenes (2), coumarin (1), fatty acid amide (1). Spectral datasets from both UV and UPLC/MS were subjected to multivariate data analysis for discrimination between the 19 coffee accessions, including principal component analysis (PCA) and orthogonal projection to latent structures-discriminate analysis (OPLS-DA). PCA model of the UV dataset (200-450 nm) provided effective discrimination of the unroasted from the roasted ones and further to classify closely related samples based on their origin. Moreover, UPLC/MS models could further distinguish between green and roasted and also roasted and instant samples. The results of both platforms were comparable, suggesting the UV technique as an alternative tool for UPLC/MS. Moreover, total phenolics assay (TPC) and antioxidant assays (DPPH and FRAP) were employed to correlate between both assays and UPLC/MS semi-quantitative dataset using partial least squares-discriminant (PLS), with a positive correlation between metabolites and assays. Potential markers that were revealed from the PLS model included caffeoylquinic acid, dihydroxypsoralen-*O*-hexoside, dicaffeoyl-quinolactone and caffeine. Moreover, mineral analysis was employed in different coffee specimens for major elements in coffee using

inductively coupled plasma atomic emission spectrometry (ICP-AES). To the best of our knowledge, this is the first comprehensive comparative metabolomics approach targeting a large number of coffee specimens and to be used for future quality control determination in coffee revealing for seeds roasting indices or origin.

**Keywords:** Antioxidant; Chemometrics; Chlorogenic acid; Coffee; UPLC/MS; UV

# Table of Contents

<b>Introduction.....</b>	<b>1</b>
<b>Goal of the study .....</b>	<b>4</b>
<b>Review of Literature.....</b>	<b>5</b>
<b>Materials, Apparatus, and Techniques.....</b>	<b>30</b>
<b>Chapter 1: .....</b>	<b>39</b>
<b>1.1 Metabolites profiling of green and roasted coffee beans of arabica and canephora species.....</b>	<b>39</b>
<b>1.2 UPLC/MS metabolites identification using global natural products social networking (GNPS).....</b>	<b>41</b>
<b>1.3 UPLC/MS metabolites profiling of coffee seed extracts .....</b>	<b>55</b>
<b>1.3.1 Alkaloids .....</b>	<b>55</b>
<b>1.3.2 Organic acids.....</b>	<b>56</b>
<b>1.3.3 Hydroxy cinnamates .....</b>	<b>58</b>
<b>1.3.4 Diterpenes .....</b>	<b>72</b>
<b>1.3.5 Fatty acids.....</b>	<b>80</b>
<b>1.3.6 Nitrogenous compounds .....</b>	<b>87</b>
<b>1.3.7 Miscellaneous.....</b>	<b>91</b>
<b>1.4 Multivariate data analysis and fingerprinting of coffee samples analyzed using UPLC/MS.....</b>	<b>92</b>
<b>1.4.2. Instant versus roasted samples .....</b>	<b>94</b>
<b>1.4.3. Cardamom versus plain samples .....</b>	<b>96</b>
<b>2 Chapter II: Fingerprinting of coffee seeds <i>via</i> UV spectroscopy coupled to chemometrics.....</b>	<b>98</b>
<b>2.1 UV fingerprinting of coffee seeds.....</b>	<b>99</b>
<b>2.1.1 Roasted versus non-roasted .....</b>	<b>99</b>
<b>2.1.2 Instant versus roasted samples .....</b>	<b>100</b>
<b>2.1.3 Cardamom versus plain samples.....</b>	<b>101</b>
<b>2.1.4 Geographical origin and species .....</b>	<b>103</b>

2.2	Comparison between UPLC-MS and UV fingerprinting multivariate data analysis models.....	104
3	Chapter III: Determination of total phenolics in coffee specimens and in relation to its in vitro antioxidant assays.....	106
3.1	Determination of total phenolics (TPC) of coffee species .....	107
3.2	Antioxidant activity.....	107
3.2.1	DPPH assay.....	107
3.2.2	FRAP assay .....	108
3.3	Correlation between biological assays and UPLC/MS metabolite profile .....	110
	Chapter IV: Mineral's analysis .....	112
	Conclusion .....	116
	Recommendation.....	117
	References .....	118



## List of tables

<b>Table 1:</b> Structures of chlorogenic acids in <i>C. arabica</i> and <i>C. canephora</i> .....	8
<b>Table 2:</b> Structures of alkaloids isolated from coffee seeds.....	14
<b>Table 3:</b> Structures of diterpenes reported in coffee. ....	16
<b>Table 4:</b> Structures of fatty acids and fatty acid amides isolated from coffee species. ....	18
<b>Table 5:</b> A list of coffee specimens analyzed by UPLC/MS and UV, including origin, degree of roasting, and sample code used in the text.....	30
<b>Table 6:</b> Metabolites identified in methanol extracts of authenticated green <i>C.canephora</i> (GCC), green <i>C. arabica</i> (GCA), roasted <i>C. canephora</i> (RCC), and roasted <i>C. arabica</i> (RCA) via UHPLC-PDA-ESI-MS in both negative and positive ionization modes. ....	44
<b>Table 7:</b> Summarized results of TPC, DPPH, FRAP $\pm$ SD (n=3). The corresponding sample codes are listed in <b>Table 5</b> .....	109
<b>Table 8:</b> Results of mineral analyses in 10 coffee samples. All values in the table are expressed as Mean $\pm$ SD expressed in ppm. ....	114

## List of figures

<b>Figure 1:</b> Illustrative figure for the study plan. ....	4
<b>Figure 2:</b> Structures of the main hydroxycinnamic acids in coffee. Compound (1) dimethoxycinnamic acid and compound (2) trimethoxycinnamic acid. ....	8
<b>Figure 3:</b> Chemical structures of hydroxycinnamic amides reported in coffee. ....	13
<b>Figure 4:</b> Chemical structures and names of 5-hydroxytryptamides (C-5HTs) carboxylic acids derivatives. ....	21
<b>Figure 5:</b> Representative UPLC/MS base peak chromatograms of green <i>Coffea arabica</i> (A), green <i>Coffea canephora</i> (B), roasted <i>Coffea arabica</i> (C), roasted <i>Coffea canephora</i> (D) in positive ionization mode. In negative ionization mode, green <i>Coffea arabica</i> (E), green <i>Coffea canephora</i> (F), roasted <i>Coffea arabica</i> (G), roasted <i>Coffea canephora</i> (H). ....	40
<b>Figure 6:</b> Full molecular networking from UPLC/MS-MS data in negative ionization mode from green <i>Coffea canephora</i> (dark green color), green <i>Coffea arabica</i> (light green), roasted <i>Coffea canephora</i> (dark brown color), roasted <i>Coffea arabica</i> (light brown). The network exhibited a pie chart reflecting parent ions and their abundance in each sample. ....	42
<b>Figure 7:</b> Molecular network of major clusters created from coffee samples (GCA, GCC, RCA, RCC). For all the networks, nodes are color-coded based on the roasting and species present and labelled by their parent ions. Light and dark green corresponds to green <i>Coffea canephora</i> and green <i>Coffea arabica</i> , respectively, while light and dark brown correspond to roasted <i>Coffea canephora</i> and roasted <i>Coffea arabica</i> , respectively. (A): Molecular network of hydroxycinnamates, (B): Hydroxycinnamic lactones, (C): Hydroxycinnamic amides, (D): Diterpenes glycosides (E), Sphingolipids. ....	43
<b>Figure 8:</b> MS/MS spectrum of peak (P6) in the positive ion mode. ....	55
<b>Figure 9:</b> MS/MS spectrum of peak (P5) in positive ion mode. ....	56
<b>Figure 10:</b> MS/MS spectrum of peak (P3) in the negative ionization mode. ....	57
<b>Figure 11:</b> MS/MS spectrum of peak (P2) in the negative ionization mode. ....	57
<b>Figure 12:</b> MS/MS spectrum of peak (P7) in the negative ion mode. ....	61
<b>Figure 13:</b> Fragmentation pattern for caffeoylquinic acids (P7, P8, P9) ....	61
<b>Figure 14:</b> MS/MS spectrum of peak (P19) in negative ion mode ....	62
<b>Figure 15:</b> MS/MS spectrum of peak (P20) in negative ion mode. ....	62
<b>Figure 16:</b> MS/MS spectrum of peak (P23) in negative ion mode. ....	63
<b>Figure 17:</b> MS/MS spectrum of peak (P24) in negative ionization mode ....	63
<b>Figure 18:</b> MS/MS spectrum of peak (p16) in negative ion mode ....	64
<b>Figure 19:</b> MS/MS spectrum of peak (P22) in negative ion mode ....	64
<b>Figure 20:</b> MS/MS spectrum of peak (P28) in negative ion mode. ....	65
<b>Figure 21:</b> MS/MS spectrum of peak (P27) in negative ion mode ....	65
<b>Figure 22:</b> MS/MS spectrum of peak (P4) in negative ion mode ....	66
<b>Figure 23:</b> MS/MS spectrum of peak (P17) in negative ion mode ....	66

<b>Figure 24:</b> MS/MS spectrum of peak (P13) in negative ion mode .....	67
<b>Figure 25:</b> MS/MS spectrum of peak (P10) in negative ion mode .....	67
<b>Figure 26:</b> MS/MS spectrum of peak (P18) in negative ion mode .....	68
<b>Figure 27:</b> MS/MS spectrum of (P21) in negative ion mode.....	68
<b>Figure 28:</b> MS/MS spectrum of peak (P25) in negative ion mode .....	69
<b>Figure 29:</b> MS/MS spectrum of peak (P30) in negative ion mode .....	69
<b>Figure 30:</b> MS/MS spectrum of peak (P82) in negative ion mode .....	70
<b>Figure 31:</b> MS/MS spectrum of peak (P84) in negative ion mode .....	70
<b>Figure 32:</b> MS/MS spectrum of peak (P83) in negative ion mode .....	71
<b>Figure 33:</b> MS/MS spectrum of peak (P85) in negative ion mode .....	71
<b>Figure 34:</b> MS/MS spectrum of peak (P86) in negative ion mode .....	72
<b>Figure 35:</b> MS/MS spectrum of peak (P35) in positive ion mode .....	74
<b>Figure 36:</b> MS/MS spectrum of peak (P37) in positive ion mode .....	74
<b>Figure 37:</b> MS/MS spectrum of peak (P45) in positive ion mode.....	75
<b>Figure 38:</b> MS/MS spectrum of peak (P43) in positive ion mode .....	75
<b>Figure 39:</b> MS/MS spectrum of peak (P38) in positive ion mode .....	76
<b>Figure 40:</b> MS/MS spectrum of peak (P39) in positive ion mode .....	76
<b>Figure 41:</b> MS/MS spectrum of peak (P41) in positive ion mode .....	77
<b>Figure 42:</b> MS/MS spectrum of peak (P36) in positive ion mode .....	77
<b>Figure 43:</b> MS/MS spectrum of peak (P47) in negative ion mode .....	78
<b>Figure 44:</b> MS/MS spectrum of peak (P49) in negative ion mode .....	78
<b>Figure 45:</b> MS/MS spectrum of peak (P48) in negative ion mode .....	79
<b>Figure 46:</b> MS/MS spectrum of peak (P50) in negative ion mode .....	79
<b>Figure 47:</b> MS/MS spectrum of peak (P51) in negative ion mode .....	80
<b>Figure 48:</b> MS/MS spectrum of peak (P71) in negative ion mode .....	82
<b>Figure 49:</b> MS/MS spectrum of peak (P52) in negative ion mode .....	82
<b>Figure 50:</b> MS/MS spectrum of peak (P72) in negative ion mode .....	83
<b>Figure 51:</b> MS/MS spectrum of peak (P75) in positive ion mode .....	83
<b>Figure 52:</b> MS/MS spectrum of peak (P56) in negative ion mode .....	84
<b>Figure 53:</b> MS/MS spectrum of peak (P57) in negative ion mode .....	84
<b>Figure 54:</b> MS/MS spectrum of peak (P60) in negative ion mode .....	85
<b>Figure 55:</b> MS/MS spectrum of peak (P62) in negative ion mode .....	85
<b>Figure 56:</b> MS/MS spectrum of peak (P65) in positive ion mode .....	86
<b>Figure 57:</b> MS/MS spectrum peak (P63) in positive ion mode .....	86
<b>Figure 58:</b> MS/MS spectrum peak (P59) in positive ion mode .....	87
<b>Figure 59:</b> MS/MS spectrum peak (P76) in positive ion mode .....	88
<b>Figure 60:</b> MS/MS spectrum peak (P77) in positive ion mode .....	88
<b>Figure 61:</b> MS/MS spectrum peak (P78) in positive ion mode .....	89

<b>Figure 62:</b> MS/MS spectrum peak (P79) in positive ion mode.....	89
<b>Figure 63:</b> MS/MS spectrum peak (P80) in positive ion mode .....	90
<b>Figure 64:</b> MS/MS spectrum peak (P81) in positive ion mode .....	90
<b>Figure 65:</b> MS/MS spectrum peak (P1) in positive ion mode .....	91
<b>Figure 66:</b> MS/MS spectrum peak (P33) in negative ion mode .....	92
<b>Figure 67:</b> Multivariate PCA and OPLS analyses of coffee specimens labelled as (A) and (C) for green and roasted specimens, respectively. (A) Principal component analysis (PCA) score plot, (B) PCA loading plot, (C) Orthogonal projections to latent structures discriminant analysis (OPLS-DA) score plot, and (D) OPLS-DA S-plot show covariance p [1] against the correlation p(cor) [1] of the variables of the discriminating component of the OPLS-DA models. Selected variables are highlighted in the S-plot and are discussed in the text. ....	94
<b>Figure 68:</b> Multivariate PCA and OPLS analyses of coffee specimens labelled as (A) and (C) for Instant and roasted specimens, respectively (A) Principal component analysis (PCA) score plot, (B) PCA loading plot, (C) Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot, and (D) OPLS-DA S-plot, show covariance p [1] against the correlation p(cor) [1] of the variables of the discriminating component of the OPLS-DA models. Selected variables are highlighted in the S-plot and are discussed in the text. ....	96
<b>Figure 69:</b> Multivariate PCA and OPLS analyses of coffee specimens labelled as (cardamom) and (no cardamom) for roasted cardamom and plain roasted specimens, respectively (A) Principal component analysis (PCA) score plot, (B) PCA loading plot, (C) Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot, and (D) OPLS-DA S-plot. Show covariance p [1] against the correlation p(cor) [1] of the variables of the discriminating component of the OPLS-DA models. Selected variables are highlighted in the S-plot and are discussed in the text. ....	97
<b>Figure 70:</b> PCA score plot (A) of roasted and green, OPLS based on roasting (B), S-line based on roasting (C) PCA of instant and roasted samples (D), OPLS for roasted and instant samples (E), S-plot for instant and roasted samples (F), PCA for plain roasted and instant samples (G). The S-plot shows the correlation(cor) and covariance p [1] between variables(wavelengths) and the predictive score of the discriminating component of OPLS-DA. The discriminant wavelengths of the important variables list are highlighted and discussed in the text. ....	102
<b>Figure71:</b> PCA score plot derived from modelling roasted arabica coffee (RCA), roasted canephore (RCC), and 7 roasted commercial samples (BRK, LRCS, LRS, LRCK, LRCM, LRCQ, BRA (A). PCA score plot derived from modelling roasted commercial samples against each other (B).....	104
<b>Figure 72:</b> Total phenolic content (TPC) of investigated coffee specimen and values are expressed as mg gallic acid equivalent/g extract (mg GAE/g extract), DPPH (2,2-diphenyl-1-picrylhydrazyl) scavenging activity, and values are in IC50±SD (µg/mL), and ferric reducing antioxidant power (FRAP) and values are expressed in mg Trolox equivalent per mg extract (mg	

TE/mg extract). Each bar represents mean $\pm$ SE (n=3); the corresponding coffee codes are listed in <b>Table 5</b> .....	110
<b>Figure 73:</b> Radar charts created from VIP score of the main metabolites which contributed in antioxidant activity. A: Dicafeoyl quinolactone showed maximum concentration in roasted samples and light roasted (RCA, RCC, LRCM), respectively. B: Caffeoylquinic acid showed maximum concentration green, and light roasted (BRA, LRCM), respectively. C: Feruloylquinic acid showed maximum concentration in green, and light roasted (GCS, GCE, BRA), respectively. Samples codes are listed in <b>Table 5</b> .....	111
<b>Figure 74:</b> Mineral's level expressed as ppm in coffee samples, A: K and Mg levels in coffee samples B: Other minerals level in coffee samples (Mn, Pb, Zn, Se, Fe, Cu, Na, and Ca).....	115

## List of abbreviations

amu	Atomic mass unit
Anova	Analysis of variance
AO	Antioxidant
BRA	Lab roasted green coffee (Aswan)
BRK	Lab roasted green coffee (Kuwait)
CC	Column chromatography
CETP	cholesterol ester transfer protein
CGA	Chlorogenic acids
CNS	Central nervous system
CQA	Caffeoylquinic acid
CQL	caffeoylquinic lactone
CSAs	caffeoyl shikimic acid
CYP	Cytochrome P-450
DM	Diabetes Mellitus
DPPH	2,2 -Diphenyl-1-picrylhydrazyl reagent
EPR	Electron paramagnetic resonance
F-C	Folin–Ciocalteu
FDA	Food and drug administration
FFAs	Free fatty acids
Fig.	Figure
FQA	Feruoylquinic acid
FRAP	Ferric reducing antioxidant power
GNPS	Global natural product social molecular networking
HCA	Hydroxycinnamic acid
HO <sup>•</sup>	Hydroxyl radical
HPLC-DAD-MS	High-performance liquid chromatography-diode array detector

	mass spectrometry
HRKC	Heavily roasted coffee blended with cardamom
IA	Iso-citric acid
IC	Inhibition coefficient
ICP-AES	Inductively coupled plasma atomic emission spectrometry
LC-MS	Liquid chromatography-mass spectrometry
LDL	Low-density lipoprotein
M	Mean
mg	Milligram
µg	Microgram
ml	Milliliter
µl	Microliter
m/m	Millions
MN	Molecular networking
MVA	Multivariate data analysis
$O_2^{\bullet-}$	Superoxide radical
OPLS-DA	Orthogonal projection to latent structures-discriminate analysis
P	covariance
Pcor	Correlation
PCA	Principal component analysis
<i>P</i> -CoQA	p-coumaroylquinic acid

PDA	Photodiode array
PDA	photodiode array detector
PPM	Parts per million
QA	Quinic acid
RCA	Roasted <i>C. Arabica</i>
RCC	Roasted <i>C. canephora</i> or <i>C. Robusta</i>
ROO <sup>•</sup>	Peroxyl radical
R <sub>t</sub>	Retention time
SD	Standard Deviation
SE	Standard error
SIMCA	Soft independent modelling class analogy
SPSS	Social packaging for statistical study software
TE	Trolox equivalent
TFC	Total flavonoid content
TPC	Total phenolic content
UV	Ultraviolet
UPLC	Ultra-performance liquid chromatography
VIP	Variable influence on projection
W	weight



## Introduction

Beverages-containing caffeine, including coffee, are consumed daily worldwide to improve consciousness and cognitive functions. About 3 billion cups of coffee are consumed daily, expressed economically at the cost of *ca.* US \$ 200 billion annually [1, 2]. Among more than 120 species of *Coffea*, coffee is brewed mainly from the seeds of *Coffea arabica* L. and *Coffea canephora* L. var. *robusta* or *C. robusta*. Therefore, both species are the most important commercially in coffee production [3]. The arabica coffee is much favored by most of the consumers having a more sensorial aroma and flavor compared to robusta [4]. In 2017, 160 million coffee packs were exported, 63.3% derived from the arabica coffee, while the remainder was mainly robusta coffee [5]. Consumption of coffee is also attributed to its characteristic flavor and aroma besides from its CNS stimulant action, which is acceptable by consumers. As a beverage, it gathers various flavors like sour, bitter, and astringent taste attributed to the presence of carboxylic acids, alkaloids, and chlorogenic acids, respectively [1]. The addition of herbs and or spices to coffee is one of the most common traditions in the Arab and Gulf area. On top of spices, cardamom (*Elettaria cardamomum* Maton) is considered a natural antioxidant source and used as spices, food flavoring agent, and additive. In addition to its significant antioxidant effect, strong antimicrobial properties have been reported. Moreover, it has been used for the treatment of anorexia, bronchitis, cold, and cough [6].

Particularly, coffee products, *i.e.*, instant, espresso, coffee brews, coffee substitutes, have long been reported as a major source of dietary antioxidants due to their richness in phytochemicals. Besides caffeine (1% to 4%), coffee seeds are rich in other secondary metabolites of several health benefits. Major secondary metabolites in coffee include phenolic acids, *i.e.*, chlorogenic acids that have various pharmacological effects as anti-inflammatory, hepatoprotective, anti-cancer, and anti-diabetic [7]. The high abundance of chlorogenic acid and its derivatives (7% to 12%), *i.e.*, caffeoylquinic acids, dicaffeoylquinic acids, feruloyl quinic acids, and *p*-coumaroylquinic acids, results in additional potential bioactivities, including alleviation of the cellular oxidative and inflammatory stress modulation [7]. Moreover, coffee was found to be enriched in diterpenes (Cafestol and kahweol 0.7-3.5%), tryptophan alkaloids, and other secondary metabolites[8]. Post-harvesting of coffee seeds, especially roasting to

improve their sensory characters, are associated with the production of roasting products, *i.e.*, melanoidins with additional antioxidant and anti-inflammatory effects. The melanoidins account for 29 % *w/w* of the dry coffee weight and responsible for the browning of coffee seeds as well [9, 10] [11].[12] Also, 850 volatile compounds constitute the aroma of roasted coffee seeds, *i.e.*, smokey, nutty, roasty, and sulfury aroma. They belong to different chemical classes, *i.e.*, pyrazines, pyridines, pyrroles, furans, and sulfur-containing compounds [13]. These phytochemicals, *i.e.*, polyphenols and melanoidins, have the capacity for free radical scavenging and metal chelation [14, 15]. The *in-vitro* spectrophotometric antioxidant assays, *i.e.*, 2,2-diphenyl-1-picrylhydrazyl (DPPH), 2,2-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), and ferric reducing antioxidant power (FRAP), are frequently carried out to assess antioxidant effects of biological extracts and are based on the detection of reactive oxygen species, *i.e.*, hydroxyl, hydrogen peroxides, and superoxide radicals. However, their sensitivity varies based on chemicals' nature and PH. Consequently, it has been recommended to perform multiple assays for the better evaluation of antioxidant properties [14].

Synergism among several metabolites has been reported for natural products bioactivity, including the antioxidant activity as in the case of coffee chemicals [16, 17]. Hence, in-depth phytochemical analysis of natural metabolites with advanced analytical techniques, *i.e.*, metabolomics, is warranted. Further, variation within the different species, the complexity of plants' chemistry, and low levels of most secondary metabolites in plants, in addition to diversity in chemical dimensions, warrant for development of sensitive analytical techniques to monitor all these differences [18]. Besides, one of the main concerns in coffee production is the adulteration of roasted coffee to gain economic benefit either by blending low-quality coffee beans or adding other ingredients, *i.e.*, brown sugar, coffee husks, maize, soybean, etc. [19]. Toci *et al.* have recently reviewed the adulteration practices that necessitate an authenticity quality control protocol to determine coffee origin [20].

For the reliable analysis of coffee metabolites in various specimens, an untargeted detection method allowing for the identification of different classes should be followed. A typical platform for plant extracts profiling includes chromatographic techniques such as ultra-high-performance liquid chromatography coupled to mass spectrometry (UPLC/MS), which presents an excellent combination of selectivity and sensitivity with different modes of ionization to increase the identified metabolites scores[21].

UV-VIS technique has been previously applied for the identification of metabolome components in coffee, where UV spectral bands provided semi qualitative and quantitative analytical information about the bioactive components, *i.e.*, phenolic acids, methylxanthines, chlorogenic acids, and pentacyclic alcohol. Moreover, as a cheaper, simpler, and non-destructive technique, UV fingerprinting can be considered an alternative tool for UPLC/MS for coffee samples QC analysis [22, 23].

In both analytical methods, experimentalists are challenged with too many variables such as huge datasets, geographical origin, harvesting time, and chemotypes. Consequently, coupling chemometrics and multivariate analysis to UV fingerprinting allow combining all the challenging variables with data[19].

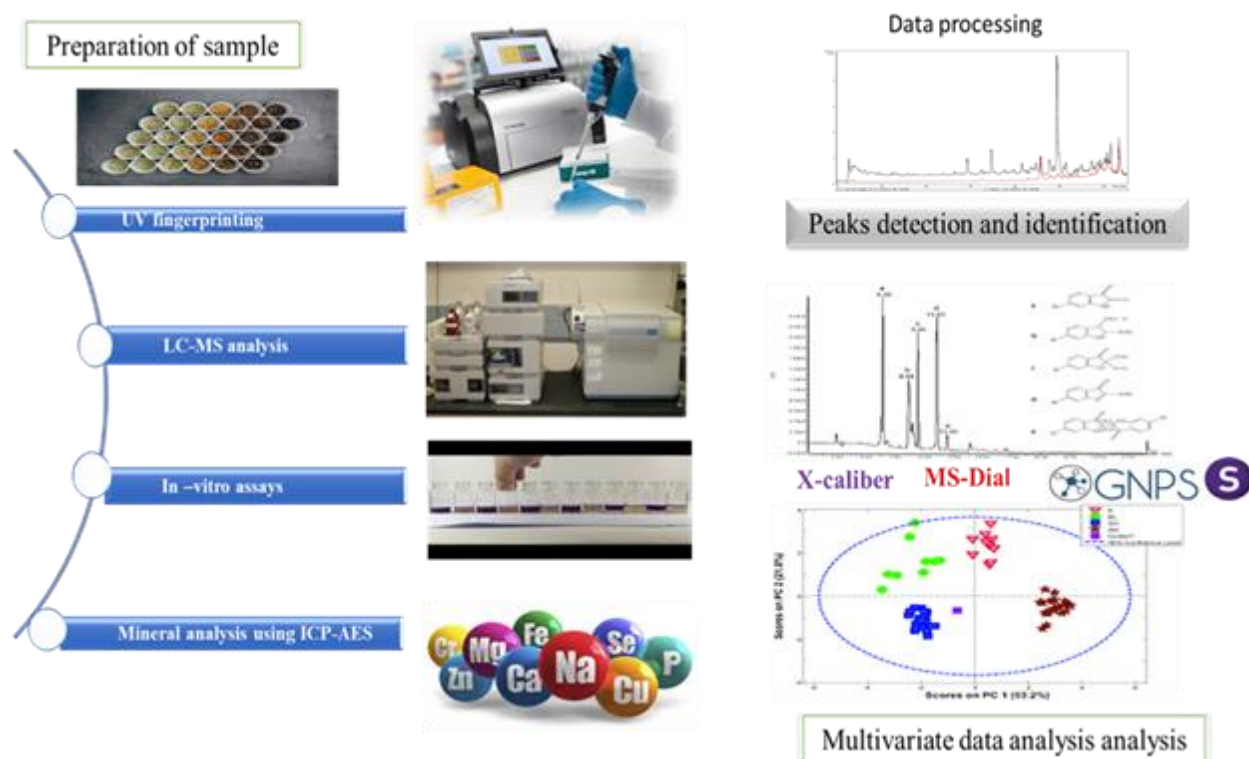
The current research presents a comprehensive analysis of 15 arabica and robusta coffee commercial products collected from the Middle East region in addition to 4 authenticated green and roasted arabica and Robusta coffee. The workflow of this study employed a metabolite profiling using RP-liquid chromatography for coffee samples to obtain the maximum number of metabolites' peaks, and in comparison, to UV fingerprinting of all samples as a less costly alternative method for liquid chromatography. For UPLC/MS metabolites' visualization, global natural product molecular networking (GNPS) was employed for detailed analysis of acquired UPLC-tandem mass for annotation and comparison of roasted and unroasted samples of both species. GNPS has been applied in many clinical fields, discovering natural phytoconstituents and drug metabolism [24]. Therefore, in our study, it was helpful for the comparison and annotation of UPLC/MS peaks based on their MS/MS fragmentation patterns. Aside, a PLS model combining the UPLC/MS dataset and the antioxidant assays' data was established to correlate the major metabolites contributing to the antioxidant activity, and radar charts were made to identify these metabolites in the different samples. Finally, the determination of elemental composition in coffee specimens for major elements in coffee using inductively coupled plasma atomic emission spectrometry (ICP-AES) was employed to determine elements level in different types of coffee samples and to study the effect of roasting on their levels.

## Goal of the study

The main goals of this study are:

1. Metabolite's identification using UPLC/MS in coffee samples.
2. Comparison between UV fingerprinting versus UPLC/MS for coffee specimens' classification.
3. Identification of roasting process impact on coffee metabolites composition as analyzed using LCMS and UV.
4. Correlation between metabolite profile and antioxidant effect using PLS modelling.
5. Determination of elemental composition in coffee specimens using ICP-AES.

An illustrative diagram explaining the study theme is shown in **Figure 1**



**Figure 1:** Illustrative figure for the study plan.

## Review of Literature

Phenylpropanoids are the dominant class that exists in natural drugs and particularly coffee beans such as flavonoids, coumarins, and hydroxy-cinnamic acid derivatives (HCA). An extensive survey of coffee literature revealed that HCA was the main secondary metabolite represented by chlorogenic acids and their derivatives that mostly contributed to the taste and the sensory properties of coffee beverage[7]. Caffeine is a major alkaloid in coffee that has strong stimulant properties, whereas chlorogenic acids impart an antioxidant effect. The roasting process is a crucial step in the coffee industry which affects its organoleptic properties and nutritional value. Green raw coffee is mainly formed of 60% Carbohydrates, 15% lipids, 10% proteins, phenolic compounds up to 10%. Besides lower levels of caffeine and trigonelline, approximately 1% and mineral content do not exceed 4%. During roasting, some of the percentages of some content are decreased, such as carbohydrates (40%), proteins (6%), chlorogenic acids (4%), while lipids, caffeine, trigonelline, and minerals possess the same relative percentage. [12]. Upon roasting, the chemical composition of raw coffee beans undergoes massive transformation processes such as esterification, thermal isomerization, acyl migration, dehydration, and lactonization (epimerization) [25]. Furthermore, roasting parts of carbohydrates get degraded to mono and oligosaccharides (low molecular weight compounds) and interact with amino acids yielding Maillard reaction products. Pyrolysis reaction products like pyrazines produced during roasting significantly affect the aroma of coffee seed [26]. Flavor and aroma, as well as phytonutrient levels, are highly affecting the quality of coffee products [27]. This review aims to provide up-to-date data information on the chemical composition of commercial coffee products in the Middle East and further determine the effect of roasting on the coffee chemical profile, as detailed in the next subsections.

### A. Phytoconstituents in coffee

#### I. Hydroxycinnamates derivatives

Esterification reaction between hydroxycinnamic acids (HCA) and organic acids that are widely distributed in coffee seeds leads to the formation of hydroxycinnamoyl esters, i.e., chlorogenic acids (CGAs) as the major class of phenolics in coffee [7]. Quinic acid (QA) and shikimic acid are the most common organic acids, while caffeic, ferulic, *p*-coumaric, sinapic,

dimethoxy, and trimethoxy cinnamic acids are the main hydroxycinnamic acids found in coffee. Representative structures of hydroxycinnamic acids are shown in **Figure 2** [12].

Chlorogenic acids are subdivided into different classes, caffeoylquinic acids, feruloylquinic acids, *p*-coumaroylquinic acids, dicaffeoylquinic acids, di-feruloyl quinic acids, di-*p*-coumaroylquinic acids, tricaffeoylquinic acids, dimethoxy cinnamoyl quinic acids, feruloyl-dimethoxy-cinnamoyl quinic acids, caffeoyl-dimethoxy-cinnamoyl quinic acids, dimethoxycinnamoyl-feruloyl caffeoylquinic acids, dimethoxy-feruloyl-caffeoylquinic acids, *p*-coumaroyl caffeoylquinic acids, *p*-coumaroyldimethoxycinnamoylquinic acids, *p*-coumaroyl-, sinapoylquinic acids, sinapoyl feruloyl quinic acids, dicaffeoyl-sinapoylquinic acids, feruloylquinic sinapoyl-caffeoylquinic acids, trimethoxycinnamoyl-feruloylquinic acids, trimethoxycinnamoyl-caffeoylquinic acids, dicaffeoyl-feruloylquinic acids, 3-(3,5-dihydroxy-4-methoxy)cinnamoyl-4- feruloylquinic acids, diferuloyl-caffeoylquinic acids, **Table 1**.

They have several pharmacological effects, mainly a strong antioxidant, anti-diabetic, and they are the main contributor to astringent and bitterness of coffee beverage [28]. Earlier studies on mono-acyl chlorogenic acids revealed that three positional isomers commonly exist naturally with the hydroxy group on C3, C4, and C5, respectively, to form three Regio isomers[7]. Other HCA derivatives identified from the esterification reaction between isocitric acid (IA) and caffeoyl or feruloyl and coumaroyl have been reported in coffee[27]. The differentiation between IA and QA isomers is not well reported due to the similarity of the molecular mass of HCA, caffeoyl, i.e., 354 Da, coumaroyl, i.e., 338 Da, feruloyl, i.e., 368 Da [27].

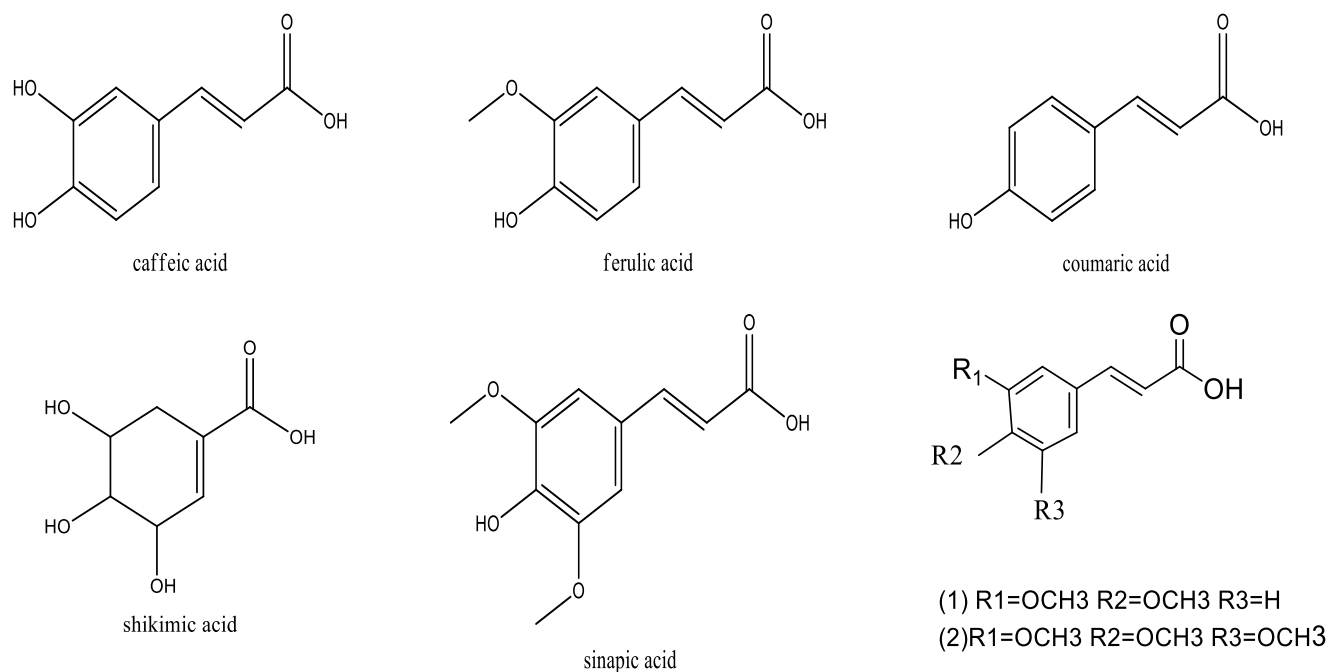
A recent study identified more than 100 different hydroxycinnamoyl esters in green and roasted samples from *C. arabica* and *C. robusta*, including 8 caffeoyl-feruloyl quinic acids, with 5 isomers of *p*-coumaroyl caffeoylquinic acids detected in green beans of arabica and canephora [28]. Another series of isomers that are associated with methyl residue that is quinic acid carrying one (*E*)-cinnamic acid and one 3,4 dimethoxy cinnamic acid were assigned as 4 - dimethoxy-cinnamoyl quinic acid, 3- caffeoyl-dimethoxy cinnamic, and 3-feruloyl- dimethoxy cinnamic acid isomers [28].

Recent research revealed minor chlorogenic acid in green canephora coffee seeds using LC-MS, 15 chlorogenic acid conjugates, trimethoxycinnamoylquinic acids, and a series of triacyl quinic acids were detected, identified as a phytochemical marker for differentiation between arabica and canephora species[29].

Roasting changes the coffee chemical profile and increases its complexity due to the formation of quinides, chlorogenic derivatives, and shikimates. Many lactones had been identified in arabica and canephora roasted coffee via UHPLC-ESI-qTOF-MS/MS, i.e., caffeoyl, feruloyl, and diferuloyl quinides described as chlorogenic lactones (CGLs). They are formed due to roasting through esterification reaction on all hydroxyl groups, including C1, followed by dehydration and formation of lactones which contribute to coffee flavor [30].

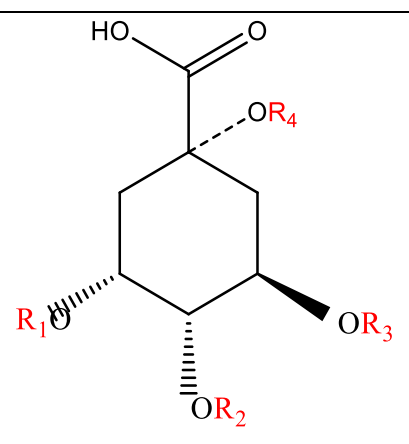
Another important hydroxycinnamates derivative, hydroxycinnamoyl amides which are formed between hydroxycinnamic acids and amino acids reported in both green and roasted coffee beans of arabica and canephora[31].

For example, caffeoyl tyrosine, caffeoyl tryptophan, feruloyl tryptophan, and coumaroyl tryptophan are detected in both arabica and canephora species, and their role as an anti-inflammatory, anticarcinogenic, and anti-diabetic have been reported, but the mechanism is not known, **Figure 3** [32].



**Figure 2:** Structures of the main hydroxycinnamic acids in coffee. Compound **(1)** dimethoxycinnamic acid and compound **(2)** trimethoxycinnamic acid.

**Table 1:** Structures of chlorogenic acids in *C. arabica* and *C. canephora*

Compound name	Chemical Structure				species
					<i>C. arabica</i> and <i>C. canephora</i>
	R1	R2	R3	R4	
Quinic acid	H	H	H	H	
3-Caffeoyl-quinic acid	caffeoyl	H	H	H	<i>C. arabica</i> and <i>C. canephora</i>
4- Caffeoyl-quinic acid	H	caffeoyl	H	H	<i>C. arabica</i> and <i>C.</i>

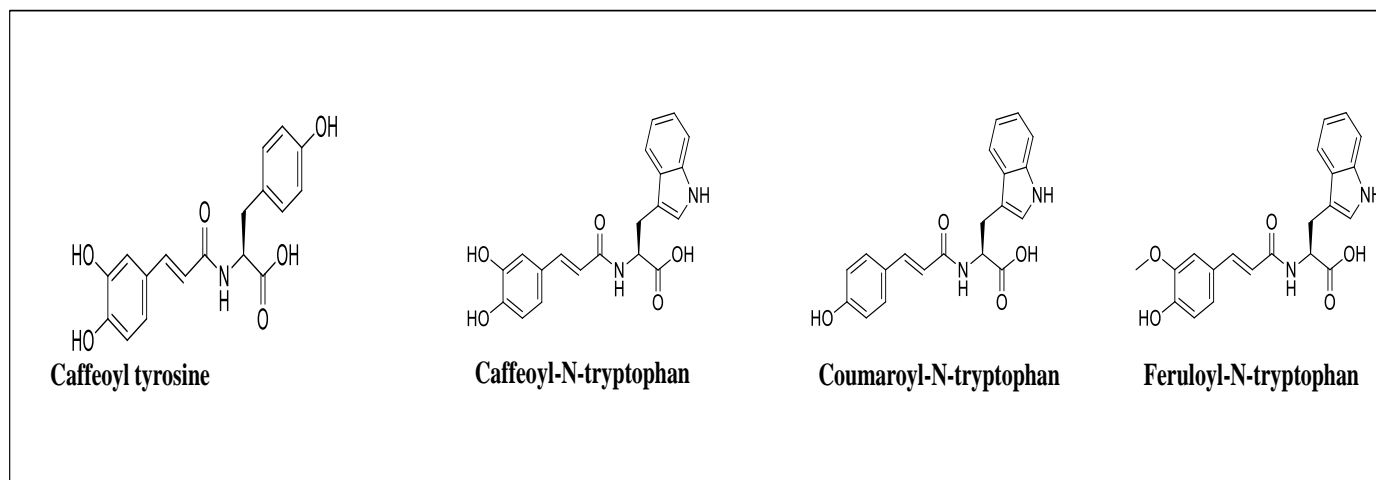


					<i>canephora</i>
5-Caffeoyl-quiunic acid	H	H	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3- <i>p</i> -Coumaroyl-quinic acid	<i>p</i> -coumaroyl	H	H	H	<i>C. arabica</i> and <i>C. canephora</i>
4- <i>p</i> -Coumaroyl-quinic acid	H	<i>p</i> -coumaroyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
5- <i>P</i> -coumaroyl-quinic acid	H	H	<i>p</i> -coumaroyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Feruoyl-quinic acid	Feruloyl	H	H	H	<i>C. arabica</i> and <i>C. canephora</i>
4-Feruoyl-quinic acid	H	Feruloyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
5-Feruloyl-quinic acid	H	H	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3,4-Dicaffeoylquinic acid	caffeoyl	caffeoyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
4,5-Dicaffeoylquinic acid	H	caffeoyl	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3,5-Dicaffeoylquinic acid	caffeoyl	H	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3,5-Diferuoylquinic acid	Feruloyl	H	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3,4- Diferuoylquinic acid	Feruloyl	Feruloyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
4,5- Diferuoylquinic acid	H	Feruloyl	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
4- <i>O</i> -Dimethoxycinnamoylquinic acid	H	Dimethoxy-cinnamoyl quinic acid	H	H	<i>C. arabica</i> and <i>C. canephora</i>

5-O-Dimethoxycinnamoylquinic acid	H	H	Dimethoxycinnamoylquinic acid	H	<i>C. arabica</i> and <i>C. canephora</i>
3-O-Dimethoxycinnamoylquinic acid	Dimethoxycinnamoylquinic acid	H	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3,4-Dimethoxycinnamoylquinic acid	Dimethoxycinnamoylquinic acid	Dimethoxycinnamoylquinic acid	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Caffeoyl-4-dimethoxycinnamoylquinic acid	Caffeoyl	Dimethoxycinnamoylquinic acid	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Caffeoyl-5-dimethoxycinnamoylquinic acid	Caffeoyl	H	Dimethoxycinnamoylquinic acid	H	<i>C. arabica</i> and <i>C. canephora</i>
4-Caffeoyl-5-dimethoxycinnamoylquinic acid	H	Caffeoyl	Dimethoxycinnamoylquinic acid	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Feruloyl-4-dimethoxycinnamoylquinic acid	Feruloyl	Dimethoxycinnamoylquinic acid	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Feruloyl-5-dimethoxycinnamoylquinic acid	Feruloyl	H	Dimethoxycinnamoylquinic acid	H	<i>C. arabica</i> and <i>C. canephora</i>
4-Feruloyl-5-dimethoxycinnamoylquinic acid	H	Feruloyl	Dimethoxycinnamoylquinic acid	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Feruloyl,4-caffeoylquinic acid	Feruloyl	Caffeoyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-feruloyl, 5-caffeoylquinic acid	Feruloyl	H	Caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>

4-feruoyl, 5-caffeoylquinic acid	H	Feruloyl	Caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-caffeoyl,4-feruoylquinic acid	Caffeoyl	Feruloyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-caffeoyl, 5-feruoylquinic acid	Caffeoyl	H	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
4-caffeoyl, 5-feruoylquinic acid	H	Caffeoyl	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-Caffeoyl -4 <i>p</i> -Coumaroylquinic acid	Caffeoyl	<i>p</i> -Coumaroyl	H	H	Green <i>C. arabica</i> and <i>C. canephora</i>
3-Caffeoyl -5 <i>p</i> -Coumaroylquinic acid	C	H	<i>p</i> -Coumaroyl	H	Green <i>C. arabica</i> and <i>C. canephora</i>
4-Caffeoyl-5- <i>p</i> -Coumaroylquinic acid	H	Caffeoyl	<i>p</i> -Coumaroyl	H	Green <i>C. arabica</i> and <i>C. canephora</i>
3- <i>p</i> -Coumaroyl-4-caffeoylquinic acid	<i>p</i> -Coumaroyl	caffeoyl	H	H	Green <i>C. arabica</i> and <i>C. canephora</i>
3- <i>p</i> -Coumaroyl-5-caffeoylquinic acid	caffeoyl	<i>p</i> -Coumaroyl	H	H	Green <i>C. arabica</i> and <i>C. canephora</i>
4- <i>p</i> -Coumaroyl-5-caffeoylquinic acid	H	<i>p</i> -Coumaroyl	Caffeoyl	H	Green <i>C. arabica</i> and <i>C. canephora</i>
3-Sinapoylquinic acid	Sinapoyl	H	H	H	Green <i>C. canephora</i>
4- Sinapoylquinic acid	H	Sinapoyl	H	H	Green <i>C. canephora</i>
5-Sinapoylquinic acid	H	H	Sinapoyl	H	Green <i>C. canephora</i>
3-sinapoyl-4-caffeoylquinic acid	Sinapoyl	caffeoyl	H	H	Green <i>C. canephora</i>
3-sinapoyl-5-caffeoylquinic acid	Sinapoyl	H	caffeoyl	H	Green <i>C. canephora</i>
4-sinapoyl-3-caffeoylquinic acid	caffeoyl	Sinapoyl	H	H	Green <i>C. canephora</i>
3-feruloyl-4-sinapoylquinic acid	Feruloyl	Sinapoyl	H	H	Green <i>C. canephora</i>

3-feruloyl-5-sinapoylquinic acid	Feruloyl	H	Sinapoyl	H	Green <i>C. canephora</i>
4-feruloyl-5-sinapoylquinic acid	H	Feruloyl	Sinapoyl	H	Green <i>C. canephora</i>
3-sinapoyl-4-feruloylquinic acid	Sinapoyl	Feruloyl	H	H	Green <i>C. canephora</i>
3-sinapoyl-5-feruloylquinic acid	Sinapoyl	H	Feruloyl	H	Green <i>C. canephora</i>
4-sinapoyl-5-feruloylquinic acid	H	Sinapoyl	Feruloyl	H	Green <i>C. canephora</i>
3-trimethoxycinnamoyl-4-feruloylquinic acid	Trimethoxy cinnamoyl	Feruloyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-trimethoxycinnamoyl-5-caffeoylquinic acid	Trimethoxy cinnamoyl	H	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
4-trimethoxycinnamoyl-5-caffeoylquinic acid	H	Trimethoxy cinnamoyl	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-trimethoxycinnamoyl-4-feruloylquinic acid	Trimethoxy cinnamoyl	Feruloyl	H	H	<i>C. arabica</i> and <i>C. canephora</i>
3-trimethoxycinnamoyl-5-feruloylquinic acid	Trimethoxy cinnamoyl	H	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
4-trimethoxycinnamoyl-5-feruloylquinic acid	H	Trimethoxy cinnamoyl	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-caffeoyl-4,5-diferuloylquinic acid	caffeoyl	Feruloyl	Feruloyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-sinapoyl-4,5-dicaffeoylquinic acid	Sinapoyl	caffeoyl	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3,4 dicaffeoyl-5-sinapoylquinic acid	caffeoyl	caffeoyl	Sinapoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3,4diferuoyl-5-caffeoylquinic acid	Feruloyl	Feruloyl	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>
3-dimethoxycinnamoyl-4-feruloyl-5-caffeoylquinic acid	Dimethoxy-cinnamoyl	Feruloyl	caffeoyl	H	<i>C. arabica</i> and <i>C. canephora</i>



**Figure 3:** Chemical structures of hydroxycinnamic amides reported in coffee.

## II. Alkaloids

Caffeine is a purine alkaloid that is one of the most stimulants consumed chemicals found naturally in many plants and manufactured products. Tea, coffee, chocolates, energy drinks are examples of known caffeine sources, chemically known as 1,3,7-trimethylxanthine,

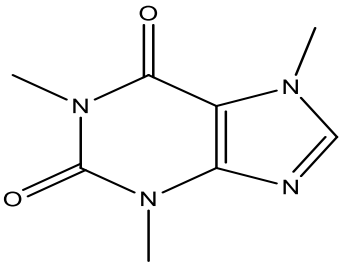
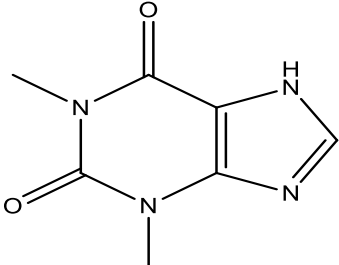
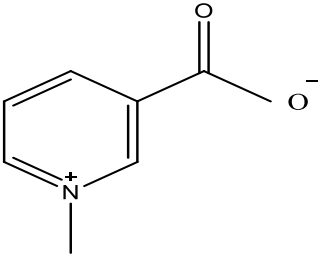
**Table 2**, [30]. Typically, one cup of coffee is estimated to contain about 80-150 mg caffeine which differs according to species and method of preparation. Robusta coffee had been reported to have double the caffeine in arabica, approximately 2.7%. Roasting has been reported to significantly change the chemical composition of coffee seeds, with an increase in the concentration of caffeine as the temperature increase through prolonged heating as in heavily roasted coffee significantly increases the caffeine content [33-35].

Caffeine is a diuretic that is shortly excreted from the body and doesn't accumulate in the body. It is metabolized in the liver by cytochrome P450(CYP1A2) into methylxanthine and methyl urea [36]. Many biological effects had been reported, including treatment of bronchopulmonary dysplasia, parkinsonism, apnea, coronary heart diseases, liver cirrhosis, hypotension, headache, fatigue, language delay, and antioxidant. On the other hand, adverse effects are detected upon excess consumption (>400 mg per day), hypertension, hyperstimulation of CNS, insomnia, alertness, panic attacks, and caffeine addiction. Extreme sub-lethal doses (6–10 mg/kg) for adults may lead to chills, nausea, palpitations, headache, and flushing [33, 36, 37].

Theobromine and theophylline also belong to the purine family that is formed during the degradation of caffeine and are reported for their role mainly as a smooth muscle relaxant in bronchospasm.[33-35]. It should be noted that coffee contains no theobromine, with only trace levels of theophylline [38].

Trigonelline is a pyridine alkaloid formed among methylation of nicotinic acid, called *N*-methyl nicotinic acid, that is abundant in coffee seeds and to mediate for its health benefits [33]. A study in diabetic mice confirmed its anti-diabetic role in diabetes mellitus patients [33] [52].

**Table 2:** Structures of alkaloids isolated from coffee seeds.

Compound name	Chemical Structure	Species
Caffeine		<i>C. arabica</i> and <i>C. canephora</i>
Theophylline		<i>C. arabica</i> and <i>C. canephora</i>
Trigonelline		<i>C. arabica</i> and <i>C. canephora</i>

### III. Diterpenes

The coffee seed oil contains alcoholic diterpenes belonging to the kauaurane family, i.e., Cafestol and kahweol present in esterified form by fatty acids and has different UV absorption

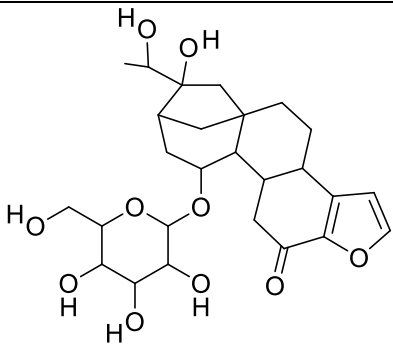
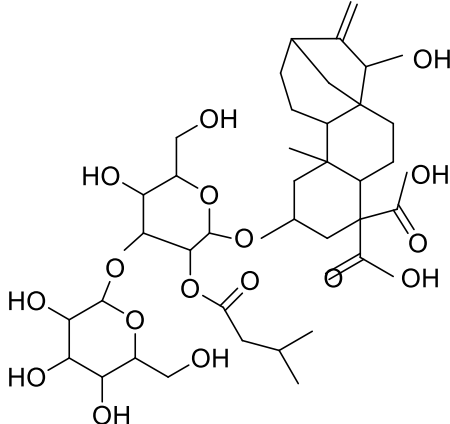
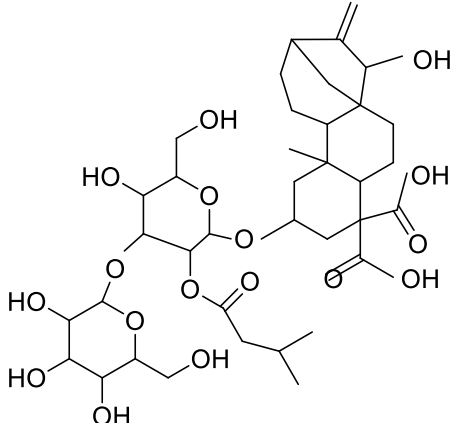
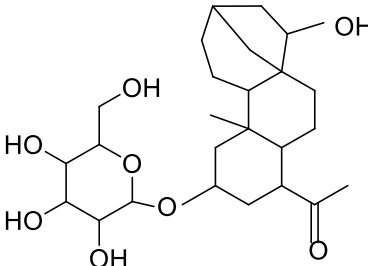
maxima due to extra double bond at 224, 290 nm, respectively[39]. Also, their content varies from species and roasting process as their percentage in green robusta(0.3-1.5%) is less than green arabica(1.2-1.9%)[8]. Moreover, 16-O- methyl cafestol(16-OMC) and 16-O- methyl kahweol was found more in canephora species than in arabica[8]. Roasting has been reported to change the chemical composition of diterpenes as they undergo dehydration reactions producing dehydrated compounds such as dehydrocafestol, dehydrokahweol, isokahweol, and secokahweol[35]. However, a study reported twelve kahweol and fourteen cafestol derivatives which are not affected by the thermal process. Prolonged roasting has been found in some studies to reduce the diterpene content with a 70% reduction in kahweol, and Cafestol in arabica, whereas a 60% decrease was reported for Cafestol in canephora **Table 3** [35,36].

Several studies studied the biological effect of coffee diterpenes, decreased risk for colorectal cancer, antioxidant, hepatoprotective, anticarcinogens, and anti-inflammatory effects were reported [8]. Negative effects of coffee diterpenes included hypercholesteremia, cardiovascular diseases, and thrombosis development in over-dosed consumers[35]. In one study, diterpene content was reported to vary based on the preparation method, i.e., diterpenes in Espresso 5 -10 times the filtered coffee. Besides, the content of diterpenes can vary in different parts of plants(endosperm, pericarp, perisperm) [35, 36].

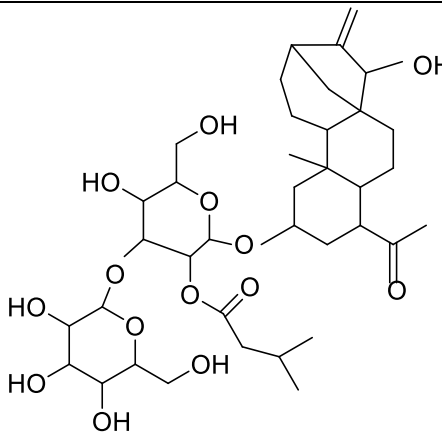
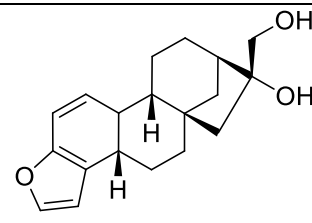
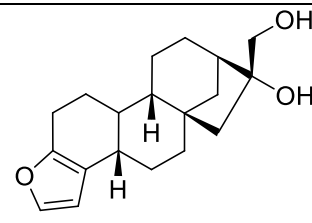
Another type of furokaurane that exist as glucosides known as mozambioside, a bitter taste furokaurane glucoside that is a potential marker for green and roasted arabica species with a UV max at 281nm[40].

Moreover, ent-kauranes diterpenoid known as atractyligenins was previously reported in the seventies in both arabica and canephora but with more abundance in arabica species[31]. They are structurally related to atractyloside and carboxyatractyloside that are close to phytotoxin atractyloside structure and reported to be degraded during roasting. However, it has been reported in coffee spent of both species and has been investigated by LCMS, with several atractyligenin and their carboxy derivatives identified suggesting that roasting temperature was not high enough for complete degradation **Table 3** [31].

**Table 3:** Structures of diterpenes reported in coffee.

Compound name	Chemical structure	Species
Mozambioside		<i>C. arabica</i>
3'-O-β-D-glucopyranosyl-2'-O-isovaleryl-2β-(2-deseoxyattractyligenin)		Green <i>C. arabica</i> and green <i>C. canephora</i>
2-O-β-glucopyranosyl-attractyligenin		Green <i>C. arabica</i> and green <i>C. canephora</i>
3'-O-β-D-glucopyranosyl-2'-O-isovaleryl-2β-(2-desoxyattractyligenin)		Green <i>C. arabica</i> and green <i>C. canephora</i>



3'-O-β-D-glucopyranosyl - 2'-O-isovaleryl- 2β-(2-carboxyatractylig enin		G and green C. <i>canephora</i> reen <i>C. arabica</i>
Kahweol		<i>C. canephora</i>
Cafestol		<i>C. arabica</i>

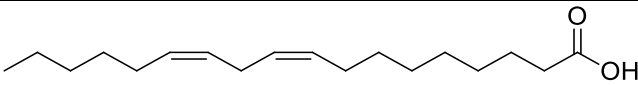
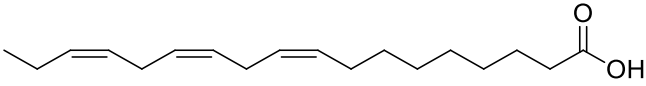
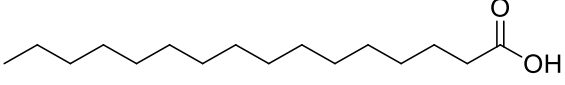
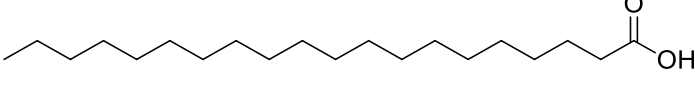
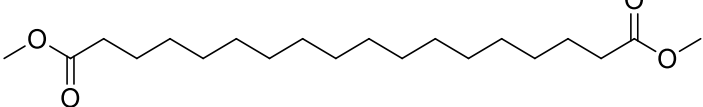
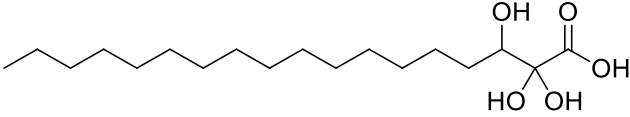
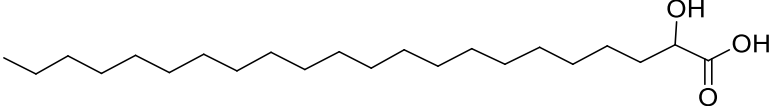
#### IV. Fatty acids and sphingolipids.

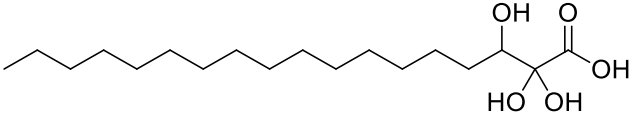
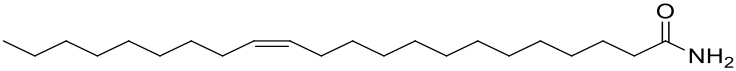
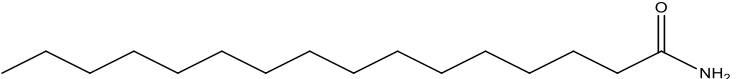
Coffee oil lipid profile is composed mainly of triacylglycerols as in most vegetable oils (75% unesterified form), diterpene esters (7-20%), and sterols (2-5%w/w) [26]. Fatty acids are found mainly in the endosperm of coffee seeds and are considered as a flavour carrier contributing to organoleptic properties. It is subdivided into saponifiable and unsaponifiable parts [41, 42]. The Saponifiable part consists mainly of triglycerides, phospholipids, kauarane diterpenes, while the unsaponifiable part includes free fatty acids, tocopherols, and diterpenes. Moreover, the percentage of oil in arabica green beans is approximately 15%, while for canephora reaches only 10% [42]. However, coffee is not considered a source of fat as lipids have low solubility in boiled water, but in some types of instant coffee, they can be added after extraction to improve the taste and flavor. Only a few studies on the effect of roasting on lipid profile had been conducted, and trans fatty acid production was detected in traces(0.1%) [41, 42].

Much work has been done towards the identification of fatty acids in coffee and analysis of its profile in different species and under different conditions like storage, roasting, etc. Fatty acids in coffee exist mostly as triglycerides, and trace amounts as free fatty acid, i.e., palmitic acid, are the major saturated fatty acid (30%), while linoleic acid and linolenic are the main unsaturated fatty acid (40%). Other fatty acids such as myristic, stearic, eicosenoic, and oleic are present in coffee oil [42].

In addition to free fatty acids, acylated fatty acid amides were detected by LCMS, including docosenoic acid amide, eicosenoic acid, and hexadecenoic acid amides, **Table 4** [26]. Several diseases such as cancer, hypercholesterolemia, obesity, atherosclerosis, and Alzheimer's had been correlated to sphingolipids and their metabolites [43]. Few studies were done on the identification of sphingolipids in natural plants, and no reports on coffee have been conducted.

**Table 4:** Structures of fatty acids and fatty acid amides isolated from coffee species.

Compound name	Chemical structure	Species
Linoleic acid		<i>C. arabica</i> and <i>C. canephora</i>
Linolenic acid		<i>C. arabica</i> and <i>C. canephora</i>
Palmitic acid		<i>C. arabica</i> and <i>C. canephora</i>
Arachidic acid		<i>C. arabica</i> and <i>C. canephora</i>
Dimethyl octadecanedioate		<i>C. arabica</i> and <i>C. canephora</i>
Hydroxy-tetracosanoic acid		<i>C. arabica</i> and <i>C. canephora</i>
Hydroxy-docosanoic acid		<i>C. arabica</i> and <i>C. canephora</i>

Trihydroxy-octadecanoic acid		<i>C. arabica</i> and <i>C. canephora</i>
Eicosenoic acyl amide		<i>C. arabica</i> and <i>C. canephora</i>
Hexadecanoic acyl amide		<i>C. arabica</i> and <i>C. canephora</i>

## V. Carbohydrates

Carbohydrates constitute more than half the weight of dry raw coffee beans (~ 60% dry matter). The roasting process affects the sugar content of green coffee beans for both low and high molecular weight sugars. Low molecular weight carbohydrates, mainly sucrose that are present at small levels, with higher levels in arabica (5-8%) than robusta (2-5%) [44]. Nevertheless, sucrose is degraded during roasting to drop to trace levels in roasted coffee beans (0.23-.033%) [45]. Moreover, reducing sugars that were produced during the hydrolysis of sucrose like glucose and fructose were reported to be rapidly degraded upon roasting into furan compounds [45].

In contrast, polysaccharides showed higher stability upon roasting, i.e., arabinogalactan, galactomannan, and mannan, which gave the characteristic thickness of coffee beans [46]. Besides, cellulose and hemicellulose are present in green coffee beans [44]. Polysaccharides level was directly related to the quality of foam formed in *espresso* coffee, and they are further not metabolized by digestive enzymes being a source of dietary fibers that contribute to coffee anti-colon cancer effect [47].

## VI. Protein

Protein and amino acids are present in green coffee and, to account for 15% of its dry weight, found more abundant in robusta coffee than in arabica. Free amino acids are lower than protein ( 0.15–2.5%) in *C. arabica* and *C. canephora* species [48].

During roasting, they play an essential role in Maillard's reaction by interacting with sugars. Accordingly, they are the main contributors to the flavor, color, and aroma of coffee [13].

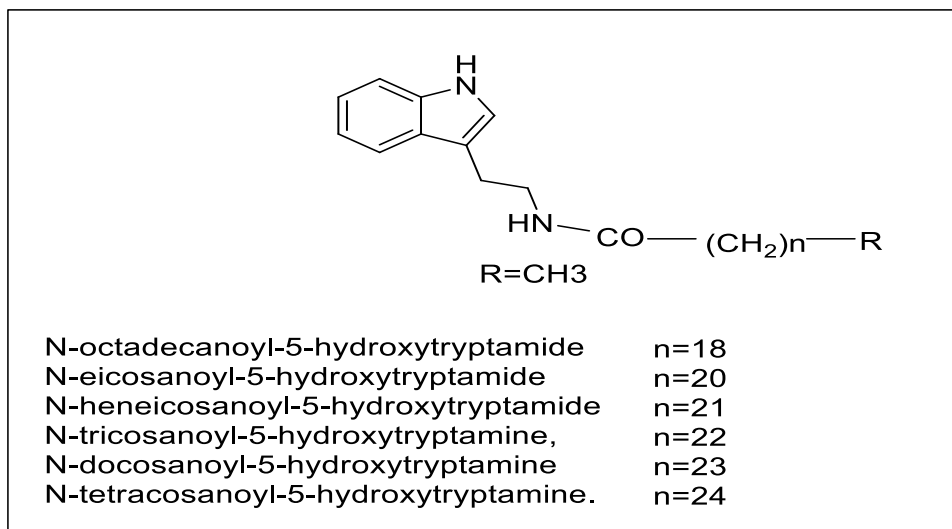
Amino acids such as tyrosine, phenylalanine, and tryptophan are present in green coffee and highly degraded with roasting resulting in negligible amounts in roasted beans [48].

Other factors that were reported to affect the amino acid content include decaffeination, fermentation, drying, and storage [12, 49]. Likewise, storage of green coffee beans at high temperatures increases the amino acid concentration due to nonenzymatic reaction and pyrolysis [48]. Several reports were conducted on protein content in coffee seeds on their important role in flavour production. However, the exact proportion and concentration had not been confirmed and varied according to the used methods for estimation [48]. Moreover, bound protein to arabinogalactans was reported that are most probably enzymes. Earlier studies reported that the polyphenol oxidase enzyme oxidizes phenolic acids to quinones which react with protein and consequently changing the quality of coffee beans [48].

## VII. Nitrogenous compounds

Coffee wax has been found to contain carboxylic acid-5-hydroxytryptamine (C-5HTs), where the amino group of 5-hydroxytryptamine (5-HT) is combined with different chain lengths of fatty acids. For example, octadecanoyl C20, eicosanoyl C20, heneicosanoyl C21, docosanoyl C22, tricosanoyl C23, and tetracosanoyl C24 producing *N*-octadecanoyl-5-hydroxytryptamine and *N*-eicosanoyl-5-hydroxytryptamine, *N*-heneicosanoyl-5-hydroxytryptamine, *N*-tricosanoyl-5-hydroxytryptamine, *N*-docosanoyl-5-hydroxytryptamine, and *N*-tetracosanoyl-5-hydroxytryptamine, respectively, **Figure 4** [50].

Additionally, these compounds were reported to have an irritant effect on the stomach in high doses, so the interest in waxy layer removal has been increased. Several methods have conducted an improvement of gastric discomforts, such as the steaming method, dewaxing, and decaffeination [32, 50].



**Figure 4:** Chemical structures and names of 5-hydroxytryptamides (C-5HTs) carboxylic acids derivatives.

### VIII. Vitamins and minerals.

Coffee seeds represent a good source for minerals such as Na, Mg, K, Mn, Zn, Ca, Ba, Ni, Cu, and Fe. The USDA Nutrition database reported that 250 ml of coffee brew and 30 ml espresso can provide up to 8 mg and 25 mg of magnesium, respectively, while they can provide 115 mg and 34 mg of potassium, respectively [33].

In addition to vitamins, i.e., nicotinic acid, vitamin D,  $\alpha$ -tocopherol, thiamine (B1), riboflavin(B2), niacin (B3), and pantothenic acid (B5) with a concentration range of 5-11% of the RDA [37]. The relation between methods of preparation and mineral content had been investigated and showed that the highest level of Ca and Cu were found in Espresso coffee, while the lowest in the aero press. Moreover, Fe was very low in French coffee while its level increased significantly in the brewing coffee as along with phosphorous (P) [34].

## B. Coffee and metabolomic studies

Recently, MS-based metabolomics techniques have delivered an extraordinary combination of selectivity and sensitivity, which can be employed as an effective platform for metabolites profiling [7]. Different MS techniques provide different ionization modes to gain a wide range of identified metabolites [19]. Moreover, UPLC/MS or fluorescence detection (HPLC/FLD) and UV (HPLC/UV) are well suited for secondary metabolites fingerprinting as a powerful analytical technique for natural plants' characterization and classification [21]. Besides, to provide a fingerprint of the extract, direct spectroscopic devices are typically used, including UV, IR, and NMR [51]. Among previous fingerprinting approaches in coffee include the discrimination between defective and non-defective roasted coffees by IR and authentication of commercial Brazilian arabica blends composed of roasted coffee with  $^1\text{H}$ -NMR [52]. The common factor in all previous techniques lies in the generation of huge datasets which warrant the application of statistical modelling tools, including principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA). The conventional analytical methods are challenged with several variables such as huge data sets, geographical location, harvesting time, and chemotypes. These models provide an indicator for methods reliability and insights into separations between the investigated sample groups [53]. Compared to the extensive application of metabolomics in coffee analysis, much less has been applied for coffee analysis in the context of its pharmacological activities.

In one study, the application of metabolomics in the coffee analysis included profiling of phenolic compounds and melanoidin, including the Indonesia Arabica and Robusta coffee extracts based on the FTIR platform. The results revealed that upon roasting, an increase in melanoidins content and a decrease in phenolics. However, there was a significant decrease in antioxidant activity that demonstrated the weak antioxidant effect of melanoidins in coffee compared to phenolic compounds. These results were based on a PLS metabolomic fingerprint that demonstrated a positive correlation of hydroxyl group on antioxidant activity [54].

Pérez-Míguez *et al.* has applied a non-targeted metabolomic approach based on LC-MS for the identification of metabolites in arabica coffee seeds roasted at different degrees. They found chemical compounds that could distinguish between different degrees of roasting, such as caffeoylquinic acid, chlorogenic acid lactones, *N*-caffeoyl tryptophan which showed a marked

increase. However, compounds that showed an opposite trend like caffeoyl feruloyl quinic acid, mozambioside, coumaroyl quinic acid, and dicaffeoylquinic acid showed a marked decrease [55].

Another study used ion trap mass spectrometry coupled to LCMS as a stronger analytical tool for further explanation of the mechanism for fragmentation of diacyl chlorogenic acids in green canephora coffee. The results revealed several fragmentation patterns which are diagnostic for diacyl chlorogenic acid, such as  $m/z$  299, 255 for caffeoyl derivatives, and  $m/z$  313, 269 for feruloyl derivatives corresponded to the loss of cinnamoyl moiety [7], **Table 1**.

A comprehensive review on green and decaffeinated coffee revealed 9 major and 17 minor chlorogenic acids identified by HPLC, which are considered as phytochemical markers to differentiate between *C. arabica* and *C. canephora*. Additionally, providing a full comprehensive profile for coffee to gain more knowledge for biological coffee value [35].

An earlier metabolomic study reported that caffeoylquinic acid isomers are the major chlorogenic acids mainly, 5-O-caffeoylquinic acid isomer, and it showed a 33% decline upon roasting concurrently with an increase in 3- and 4-caffeoyl isomers, though with still 5-O-caffeoylquinic acid as the major isomer. In this study, HPLC analysis along with antioxidant activity was performed to evaluate the effect of roasting on antioxidant activity of caffeoylquinic acid, reporting that antioxidant activity of medium roasted samples was double the green coffee. During roasting, a portion of chlorogenic acids undergoes hydrolysis to form lactones and Maillard reaction products which contribute to the seeds antioxidant activity [35].

During the roasting of coffee, shikimate and lactone derivatives of feruloylquinic acids and caffeoylquinic acid were also formed [56]. Jaiswal *et al.* investigated commercial roasted canephora coffee samples by LC/MS and could discriminate between them based on their chlorogenic acid lactones and hydroxycinnamoyl shikimates composition using tandem mass spectrometry [56].

Research using UPLC-DAD-QTOF/MS on 19 commercial coffee revealed nine new methyl hydroxycinnamoyl quinnats annotated as feruloyl quinic acid methyl quinate and mono

and diacyl hydroxycinnamoyl derivatives were also detected in arabica and canephora coffee samples as shown in **Table 1** [57].

Metabolomic analysis using ion mobility spectrometry-mass spectrometry (IMS-MS), NMR, ambient sonic spray ionization mass spectrometry was done to evaluate the roasting process effect on coffee [55]. Also, a non-targeted approach based on HPLC/UV has been effective for the authentication of various commercial coffee products having a different geographical origin, variety, and roasting degree [19].

A comprehensive metabolomics study employed LCMS to characterize minor triacyl hydroxycinnamic acids and sinapic acids in green canephora and reported new series of trimethoxy cinnamoylquinic acid and sinapoyl quinic acid derivatives based on their molecular weight and fragmentation pattern. Some structures of trimethoxy cinnamoyl chlorogenic acids and sinapoylquinic acids are depicted in **Table 1**.

Several analytical techniques were employed for the characterization of phenolic acid glycosides, such as GC-MS, LCMS. The fragmentation pathways usually involve cleavage of aglycon and intact sugar, and aglycon ion represents the base peak.

Asides, a study on methylxanthines class in coffee was done by HPLC and UV spectroscopy at 273 nm and reported three main compounds: 1,3,7-trimethylxanthine (caffeine), 3-dimethylxanthine (theobromine), and 1,3-dimethylxanthine (theophylline). Recently, HPLC was employed for the quantitative determination of caffeine in instant coffee, black tea, decaffeinated coffee, and roasted coffee detected at 122, 217, 26 and 350  $\mu\text{g mL}^{-1}$ , respectively [35].

Moreover, several methods have been reported for the analysis of diterpenes in coffee. HPLC was reported to be the most efficient method as it provides direct analysis without derivatization and provides the highest number of identified compounds[35]. Later, the change in the composition of lipids due to roasting, origin, or type was analyzed by RP-HPLC and with no major difference reported in triglycerides profile [42]. A recent study on coffee arabica and robusta used gel chromatographic and capillary gas techniques to assess free fatty acids without diterpenes, and triglycerides and 9 acids were detected mainly stearic and oleic, which were more abundant in robusta coffee [42].



Sugars content was quantified by ion-exchange chromatography; saccharose was the major sugar in green coffee and detected at higher levels in arabica (5-10%), while approximately 3-5% was reported for canephora coffee [35].

A newly published study investigated amino acids in coffee by HPLC, ion exchange, and direct method with amperemeter. The resulting data revealed several amino acids in green coffee powder with concentrations (4.4-1000 mg /100g), i.e., aspartic acid, glycine, tryptophan, histidine, glutamic acid, serine, arginine, tyrosine, alanine, valine, etc. [32].

Although processed coffee is a natural source of nutritional compounds and is believed to be a major source of antioxidants, it contains acrylamide that is considered a carcinogen (class 2A) according to the international agency for cancer research. Acrylamide is formed during the thermal process of food, such as baking, frying, and roasting, *via* the Maillard reaction that involves the reaction between amino acids and sugar during the heat. A previous study done on different kinds of coffee using LCMS quantified acrylamide in an instant, Turkish, and brewed coffee detected at 16.5 to 79.5 ng mL<sup>-1</sup>, 5.3–54.8 ng mL<sup>-1</sup>, and 5.9 to 38.8 ng mL<sup>-1</sup>, respectively [58]. However, the determination of acrylamide in coffee by LCMS is a challenging process due to the coextraction of coffee extract and difficulties to identify low molecular weight compounds [58].

### **C. Coffee seeds health benefits**

- **Antioxidant activity**

The enrichment of coffee in phenolic compounds, mainly cinnamic acids and chlorogenic acids, was studied and to attribute strong antioxidant properties; a 200 ml cup of coffee encompasses 50–350 mg chlorogenic acids and 35–175 mg of cinnamic acids [59].

The antioxidant activity (AO) of coffee seeds of different species and roasting degrees has been extensively studied. Coffee phytochemical elements were found to be able to scavenge free radicals, i.e., reactive oxygen species accumulated from oxidative stress reactions (Heiberweiss or Fenton reactions ) [15,60].

Antioxidant assay of crude caffeine revealed that it exerted hydrophilic and lipophilic antioxidant activity (146  $\mu\text{mol TE/g}$  and 65  $\mu\text{mol TE/g}$  respectively) [33].

Recent research assessed the antioxidant properties of caffeine to scavenge hydroxyl radical ( $\text{HO}^{\bullet}$ ) and superoxide ( $\text{O}_2^{\bullet-}$ ) radicals by electron paramagnetic resonance (EPR) that showed high bimolecular levels while coffee metabolites (methyl uric acid and methylxanthine) was more capable to scavenge peroxy radicals ( $\text{ROO}^{\bullet}$ ) at a micromolar level [60].

The same study identified chlorogenic acids (CGAs) as potential antioxidants, mainly 5-CQA as a radical scavenger for hydroxyl ( $\text{HO}^{\bullet}$ ), superoxide ( $\text{O}_2^{\bullet-}$ ) radical, and peroxy radicals. Furthermore, the di-chlorogenic acid isomer exhibited a more powerful scavenging ability than mono-isomers, likely attributed to higher numbers of hydroxyl groups [60].

The antioxidative potential of bioactive components in coffee revealed that caffeine and ferulic acids exerted significant scavenging activity in experimental mice as well as diterpenes, i.e., Cafestol and kahweol, and suggestive that these chemicals act synergistically to exert an antioxidant action [36].

The effect of roasting degree (light, medium, heavy) on the antioxidant activity of *C. arabica* and *C. robusta* using 4 in-vitro assays (TPC, DPPH, FRAP, TEAC) was determined, with considerable differences in effects among coffee species more pronounced than that due to roasting degree [60]. Among the assays, TPC and FRAP were found more sensitive than ABTS and DPPH assays for discriminating between different species of coffee and the roasting process [61].

The effect of the roasting process on the antioxidant activity of green coffee beans versus medium roasted coffee analyzed using HPLC-MS revealed a double effect in roasted coffee likely attributed to Maillard reaction products [30].

Melanoidins produced during the roasting process were also studied, high and low molecular weight products were isolated using different chromatographic techniques for further antioxidant *in-vitro* assays. The results indicated higher antioxidant activity for low molecular than high molecular weight compounds due to more scavenging of peroxy radicals [60].

The antioxidant properties of various bioactive in coffee and their metabolites in different species has been investigated using ORAC and ABTS assays revealed that caffeic and chlorogenic acids in addition to their metabolized products, i.e., as dihydro ferulic and meta coumaric acid possessed strong antioxidant properties [62].

Comparison of the effect of melanoidins and phenolic compounds was conducted on the antioxidant capacity in arabica and Indonesia coffee using *in-vitro* antioxidant assays (DPPH and FRAP). Reduced polyphenolic content was associated with increasing roasting degree and concurrent with increased melanoidins concentration [63].

Similar research using LDL oxidation assay revealed that the antioxidant activity of phenolic compounds in coffee differed according to species and roasting degree. Results identified polyphenols in green robusta coffee at double that of arabica and no significant difference after roasting [64].

- **Anti-inflammatory Activity**

Inflammation is a defensive reaction mediated by chemicals released from neutrophils and macrophages, such as cytokines, oxidants, and chemokines. Several studies reported that caffeine in coffee could prevent the action of cytokines in rodents, regulate phosphorylation, and delete the cell inflammatory response and thus has a significant role in the treatment of neuroinflammation [65-67].

Likewise, the anti-inflammatory effect of chlorogenic acid was determined using COX enzymes assays in human and mice models and exhibited high activity compared to other components [68]. Diterpenes, i.e., Cafestol and kahweol, along with ferulic acid, have been reported for anti-inflammatory potential via inhibiting the production of COXs, NO, and PGE2 [69, 70].

- **Antimicrobial activity**

The different components of coffee exhibited antibacterial and anti-fungal effects, with chlorogenic acids found to inhibit *Escherichia coli* and to have a significant effect against hepatitis B and herpes simplex virus. Similarly, ferulic acid showed significant anti-viral activity against influenza virus A & B and antifungal activity against *Microcystis aeruginosa* [71].

- **Stimulation of central nervous system**

The stimulatory effect of coffee is mainly associated with its richness of caffeine through the antagonism of the adenosine receptors A1 and A2A subtypes. Adenosine is a natural nucleoside in all body cells that has mainly inhibitory action and caffeine antagonism that leads to the stimulation of CNS, hypertension, and diuresis [37]. Caffeine is quickly absorbed in the stomach and intestine and is distributed to all the body organs, mainly the brain and metabolized by cytochrome P450 in the liver. Cafestol and kahweol have been reported to possess a neuroprotective function and increase noradrenaline secretions [36]. Ferulic acid has also been found to enhance cerebral ischemia, protect against nerve injury in neuroblastoma, and act as an anti-depressant upon continuous consumption [72]. Several studies showed the neuro-protection capacity of caffeine in the treatment of many CNS disorders such as apnea, depression, neuroinflammation, seizures, oxidative stress. Evidence for the anxiolytic effects of caffeine on anxiety-related behaviour in male and female rats was also tested [53].

- **Treatment of Hypercholesterolemia**

An increase in blood serum LDL cholesterol levels in heavy coffee consumers from Scandinavia was reported [37]. Recently, a meta-analysis demonstrated the effect of the brewing method of coffee on cholesterol levels, with filtered coffee found associated with lower serum cholesterol levels [73].

Diterpenes, mainly Cafestol and kahweol, were the cholesterol-raising chemicals found abundant in Turkish, boiled, espresso, and French coffee (6–14 mg/cup) [74].

An investigation of the mechanism of diterpenes in cholesterol elevation was done in French press coffee, and the results revealed a significant increase in cholesterol ester transfer protein (CETP) in the blood that concluded to be the cause of the elevation in LDL and VLDL cholesterol [75].

- **Prevention of Type 2 Diabetes Mellitus (DM)**

A study in Germany suggested that the risk of developing type2 DM decreased to 50% with the increasing consumption of coffee, similar to a large cohort study in the US, Finland, Netherlands, and Sweden, which are known for their high coffee consumption [37] [76]. Clinical

studies in the short term had reported a caffeine dose-dependent decrease in insulin levels and glucose tolerance [77]. However, an earlier study reported no significant relationship between diabetes mellitus and coffee consumption that was attributed to the brewing method [77]. Moreover, clinical studies used oral glucose tolerance testing (OGTT) as an additional tool for confirming the non-significant association between coffee consumption and the prevention of DM [78].

Trigonelline abundance in coffee seeds is likely to mediate for the anti-diabetic effect on mice leading to a reduced level of glucose and improvement in the peripheral neuropathy associated with DM [33].

In experimental studies on rats, chlorogenic acids and caffeine were reported to lower insulin and blood glucose levels by exhibiting insulin resistance in some rats [79, 80]. Besides, ferulic acid showed a hypoglycemic effect by decreasing insulin resistance [81].

- **Prevention of Colorectal Cancer**

Epidemiological studies revealed that coffee consumption had been inversely linked with the incidence of colon cancer, a 25% reduction in coffee drinkers than non-drinkers was found in a meta-analysis [82]. In contrast, an American study on people taking decaffeinated coffee revealed up to a 50% lower risk for colorectal cancer [83]. Several studies investigated the metabolites in coffee that could prevent colon cancer. Diterpenes were suggested to function as a chemopreventive agent in unfiltered coffee found to decrease the synthesis of bile acids, decrease DNA adducts, and enhance glutathione synthesis [8, 84, 85].

- **Effects on heart and blood pressure**

Several studies investigated coffee active components effects against heart diseases, with chlorogenic acids found to manage blood pressure and protect against heart hypertrophy, cardiovascular deficiency and decrease hypoxia in hypertensive patients [86-88].

Likewise, ferulic acid has been reported to exhibit anti-thrombotic effects on platelet aggregation [89] and vasorelaxant action in rats [81]. Diterpenes, i.e., Cafestol, have also been found to exhibit anti-atherosclerosis impact in humans and suggestive that different chemical classes in coffee mediate for effect on the cardiovascular system [90].

## Materials, Apparatus, and Techniques.

### 1. Materials

- **Plant material**

Two authenticated species of economic value were included, namely *Coffee Arabica* and *Coffee canephora* var. Robusta collected from Brazil presented as green and roasted specimens along with 17 commercial coffee specimens collected from the Arab world of different roasting stages and blends. Some of the commercial samples have cardamom as a blend. Details of coffee specimens are presented in **Table 5**.

**Table 5:** A list of coffee specimens analyzed by UPLC/MS and UV, including origin, degree of roasting, and sample code used in the text.

Sample name		Producer and Origin	Degree of roasting*	Sample code
Authentic roasted samples	Roasted <i>C. Arabica</i>	Mina Gerais, Brazil	12.9	RCA
	Roasted <i>C. canephora</i> or <i>C. Robusta</i>		6.6	RCC
Authentic green samples	Green <i>C. arabica</i>		7.3	GCA
	Green <i>C. canephora</i> or <i>C. Robusta</i>		1.6	GCC
Commercial samples	Lightly roasted blended with cardamom	Maatouk, Saudi Arabia	1.8	LRCM
	Lightly roasted coffee	Shahi, Saudi Arabia	2.8	LRS
	Heavily roasted blended with cardamom	Alameed coffee, Kuwait	6.0	HRKC
	Lab.-roasted green coffee		1.6	BRK
	Lightly roasted blended with cardamom		1.0	LRCK
	Lab.-roasted green coffee	Aswan, Egypt	1.0	BRA
	Lightly roasted blended with Qassim blend	Saudi-Arabia	---**	LRSQ
	Lightly roasted blended with cardamom	Shahi, Saudi Arabia	4.1	LRCS
		Qatar	4.2	LRCQ
	Instant <i>C. arabica</i>	Maxima coffee	32.3	ICA
	Instant Arabian coffee blended with cardamom	NESCAFE Arabiana	7.1	ICC
	Green coffee	Bayara, Emirates	4.4	GCU

	Aswan, Egypt	3.5	GCE
	Saudi-Arabia	1.3	GCS
	Alameed coffee, Kuwait	5.1	GCK

\*: Relative to green coffee seed based on melanoidins content measured using UV spectrophotometer at 200-450 nm.

\*\*: Unavailable

- **Materials for the chromatographic study**

Hexamethyldisiloxane (HMDS) and methanol-d4 (99.80% D) were obtained from Deutero GmbH (Kastellaun, Germany).

Acetonitrile (HPLC grade) and formic acid (LC-MS grade) were purchased from J. T. Baker (Deventer, The Netherlands).

MilliQ water was used for UPLC-MS analysis.

All other chemicals were supplied by Sigma Aldrich (St. Louis, MO, USA)

- **Materials for UV fingerprinting**

100% MeOH was purchased from Sigma- Aldrich (St. Louis, MO, USA)

- **Materials for determination of total phenolic content**

Gallic acid, Folin-ciocalteu reagent (FCR) were purchased from Sigma –Aldrich (St. Louis, MO, USA).

- **Materials for in-vitro antioxidant assays (DPPH radical and FRAP assays)**

DPPH (2,2-diphenyl -1-picrylhydrazyl), Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid), acetate buffer with PH=3.6 (Sodium acetate, glacial acetic acid, and distilled water

Ferric chloride (FeCl<sub>3</sub>) and concentrated hydrochloric acid (conc HCl),

TPTZ (tripyrityltriiazine) were purchased from Sigma – Aldrich (St. Louis, MO, USA)

- **Materials for mineral analysis**

HNO<sub>3</sub> (65%), H<sub>2</sub>O<sub>2</sub> (30%)

## 2. Equipment

- Shaker incubator BJPX-S20 China

- Rotary evaporator (Buchi, Geneva, Switzerland)
- Electric balance (Sartorius, Germany)
- Hot plate shaker (Edmund Bühler GmbH, Germany)
- Gen5 UV-vis microplate reader (Biotek, version 2.0)
- Inductively coupled plasma atomic emission spectrometry (ICP-AE), Thermo sci, model: iCAP 6000 series, made in England.

- **Equipment for High-resolution UPLC-Orbitrap-MS-Analysis**

UPLC (Dionex Ultimate 3000, Thermo fisher scientific, Bremen, Germany) with a photodiode array(200-600nm) and Water Acquity HSS T3, RP-18 column (150\*1 mm,1.8µm particle size, a pore size of 100Å). The eluted compounds were detected in the range of  $m/z$  100 to 1000 using a MicroTOF-Q hybrid quadrupole time-of-flight mass spectrometer (Bruker Daltonics) equipped with an Apollo II electrospray ion source in positive and negative ion modes.

### 3. Softwares:

- Thermo Xcaliber Roadmap 2.2SP1, Therm Fisher Scientific Inc.
- Soft independent modelling class analogy (SIMCA14.1)MKS Umetrics, Umea, Sweden
- Global Natural Products Social Molecular Networking (GNPS), ccmsweb@cs.ucsd.edu
- Social package for statistical study software (SPSS17, SPSSInc., Chicago, USA)
- MS dial software, version 4.16 Hiroshi Tsugawa (RIKEN)
- Graph prism 7(GraphPad Software, Inc)

### 4. Techniques

- **Roasting of coffee**

Lab roasting was performed inside a glass sealed vial using a Hot Plate Shaker SM 30 AT control (Edmund Bühler GmbH, Germany) set at 120 °C and a shaking speed of 150 rpm for 6 h.

- **Sample preparation for determination of total phenolic content (TPC) and antioxidant assays (DPPH, FRAP)**



3 g of the different coffee samples were extracted with 30 ml methanol (100%) by sonication for 30 min at room temperature and further filtered using Whatman filter paper. Further extraction using the same amount of methanol for 30 min and filtration, transfer to rotavap with filtrates combined and under reduced pressure at a temperature not exceeding 45°C till dryness to yield concentrated extracts. Dried extracts were kept at -5°C for stock solution preparations until further analysis.

- **Samples preparation for UPLC-MS analysis**

The coffee powder (150 mg) of each specimen was homogenized with 5 mL 100% MeOH containing 10 µg/mL umbelliferone as an internal standard for the relative quantification using a Turrax mixer (11,000 rpm) for five 20 s periods. A period of 1 min between each mixing period was made to prevent heating. Extracts were then vortexed vigorously and centrifuged at 3,000 x g for 30 min to remove plant debris. Then, 1 mL was aliquoted and placed on a (500 mg) C<sub>18</sub> cartridge preconditioned with methanol and Milli-Q water. Samples were then eluted twice using 3 mL methanol, the eluent was evaporated under a nitrogen stream, and the obtained dry residue was resuspended in 1 mL methanol. Two µL were afterwards subjected to LC-MS analysis in triplicates (n=3). Chromatographic separations were performed using an Acquity UPLC system (Waters) equipped with an HSS T3 column (100 × 1.0 mm, particle size 1.8 µm; Waters), where elution started at 10% acetonitrile (LC-MS grade, J.T. Baker, Deventer, Netherlands) till 100% acetonitrile over 10 minutes and kept at 100% for extra 2 min. The eluted compounds were detected in the range of *m/z* 100 to 1000 using a MicroTOF-Q hybrid quadrupole time-of-flight mass spectrometer (Bruker Daltonics) equipped with an Apollo II electrospray ion source in positive and negative ion modes.

- **Identification of metabolites analyzed using UPLC-MS**

Metabolites were identified by their retention time, accurate mass, fragments, isotopic distribution, UV–vis spectra, and error. High-resolution raw files were imported into X-caliber software qual browser (<https://www.thermofisher.com/>). The analysis was performed in both positive and negative modes, and the ion mass spectra derived from cations (M+H) or anions (M–H) accompanied by many fragmentation patterns. Relative comparison of spectral data to the

literature references, in-house data, natural product database of the phytochemical dictionary (CRC, Wiley), and GNPS aided in metabolites annotation.

- **UPLC-MS data processing for multivariate data analysis**

The original UPLC-MS low-resolution files of all authenticated and commercial samples were converted into mzML files using MS Convert. GUI(<http://proteowizard.sourceforge.net/download.html>) then to abf files using ABF converter (<https://www.reifycs.com/AbfConverter/>). Accordingly, MS dial software (<http://prime.psc.riken.jp/compms/msdial/main.html>) was utilized for the analysis of data with the following parameters: retention time (0-20min), mass range (0-1000 Da), accurate mass tolerance (MS1) 0.01, MS2, 0.025 Da, minimum peak height (1000), sigma (0.7), maximum charge number (2), MS1 tolerance for alignment (0.015 Da), and peak height 1000. Peak abundance mass list was then exported for multivariate data analysis where final ID and metabolites were Pareto scaled using SIMCA 14.1(Umetrics, Umea, Sweden). Unsupervised principal component analysis (PCA) models were validated based on  $R_2$  and  $Q_2$ . Supervised OPLS -DA that analyze the pre-classified groups and identify markers via S -a plot that was validated with P-value, covariance (p), and correlation (pcor). The regression extension analysis of PLS connects the model in two matrices of X and Y variables, with VIP values from the resulting model used to show the concentration of metabolites in different samples via radar charts.

- **Molecular based networking of coffee specimens using GNPS**

Raw data files of both authenticated arabica samples (GCA, RCA) and canephora samples (GCC, RCC) were converted to mzML format using Proteo Wizard as a converting tool to be uploaded to GNPS via WinSCP (<https://winscp.net>). In GNPS, molecular networks were generated for both negative and positive files separately, where the following parameters were adjusted was adjusted: 0.02 Da parent mass tolerance, 0.01 Da fragment ion tolerance, cosine score was 0.7 or above, and a minimum of four matching peaks were used in network building. The spectral library (Respect) was used for searching for the new spectra. For network visualization, cystoscope open-source software (version 3.8.2) where each spectrum is represented by the nodes and alignment of the spectrum to spectrum signified by the edges [24].

- **Samples preparation for UV fingerprinting**

3 gm of each sample was extracted with 30 ml 100% methanol and macerated for 2 hours, then centrifuged. Aliquot of 200  $\mu$ L was used prepared four replicates in 96-well for the Gen 5 Greener UV microplate reader (Gen 5 kitted with 96-well quartz cell with 1 nm spectral resolution in the UV region), and the absorption spectra were recorded from 200-450 nm.

- **Multivariate data analysis for UV fingerprinting**

The UV models were developed using the four genuine samples of green and roasted arabica and canephora specimens alongside the 15 commercial samples from different regions of the Middle East. Firstly, spectral data were converted to a data matrix using excel (Microsoft office 365 proplus), columns represent variables (spectral points), and rows represent the corresponding wavelength. In the excel sheet, the matrix was constructed of all samples with their biological replicates for samples against 250 variables spanning the readings between 200-450 nm. The dataset was first subjected to unsupervised techniques, including principal component analysis (PCA) using SIMCA software. All variables were mean-centred and pareto scaled before data analysis, where the centred data were divided by the square root of standard deviation.

- **Determination of total phenolics content by Follin -Ciocalteu reagent.**

Total phenolics content (TPC) of the coffee specimen extracts prepared as explained (section 3.1 )was determined calorimetrically using Folin–Ciocâlteu reagent as described by Zhang *et al.*, with slight modifications, with gallic acid being used as a standard for quantification [31]. Briefly, 20  $\mu$ L aliquots of samples were mixed with 100  $\mu$ L of 10% Folin–Ciocâlteu reagent left for 5 min in the dark followed by the addition of 80  $\mu$ L 7.5 mg sodium bicarbonate and incubated in the dark for 30 min. The absorbance of all samples was measured at 765 nm. Besides, a standard curve of gallic acid was established in the concentration range of 1-100  $\mu$ g/mL. All measurements were made in triplicates with the TPC expressed as mg gallic acid equivalent/mg extract (mg GAE/mg extract).

- **DPPH radical scavenging assay**

The DPPH radical scavenging assay was carried out as described by Hidalgo *et al.*, with slight modification [33]. Briefly, 30  $\mu\text{L}$  of each extract was mixed with 270  $\mu\text{L}$   $6 \times 10^{-5}$  M DPPH. The mixture was then left in the dark for 30 min, and absorbance was recorded afterwards at 517 nm. A negative control sample was made of 30  $\mu\text{L}$  100% methanol instead of samples aliquot. All measurements were made in triplicates using a microplate reader in different concentrations, *i.e.*, 0.01, 0.1, and 0.5  $\mu\text{g}/\text{mL}$ ; results are expressed as mean  $\pm$  SD.

The radical scavenging activity was measured for each specimen as % inhibition of DPPH =  $(1 - A_s/A_c) \times 100$ . The results were expressed in  $\text{IC}_{50} \pm \text{SD}$  ( $\mu\text{g}/\text{mL}$ ), representing the percentage required by the samples to decrease DPPH absorption by 50% and therefore, higher  $\text{IC}_{50}$  values indicate the lower antioxidant activity of coffee samples.

- **Ferric reducing antioxidant power (FRAP) assay.**

The FRAP assay is also a colorimetric assay that measures the ferric reducing power of samples [14, 15]. According to Fernández-Poyatos, *et al.*'s method, FRAP was conducted. Briefly, the stock solution of FRAP reagent consists of 10 mM TPTZ (2,4,6-tripyridyl-S-triazine) in 40 mM in HCl (10 Mm), acetate buffer (300 mM, pH 3.6), 20 mM  $\text{FeCl}_3$ . 175  $\mu\text{L}$  of freshly prepared FRAP solution was mixed with 25  $\mu\text{L}$  of extract and incubated in the dark for 30 min till being recorded at 593 nm. A Trolox calibration curve (0.01-0.1 mg/mL) was constructed and results were expressed as mg Trolox equivalents per mg extract (mg TE/g) [91]. All the measurements were performed in triplicates and expressed as mean  $\pm$  SD.

- **Minerals analysis**

10 Coffee samples were analyzed ( $n=3$ ) for their major minerals reported in coffee in addition to toxic elements, *i.e.*, Zn, Cu, Mg, Fe, Ca, Se, Na, Pb, Mn, and K were determined by ICP-AES based on Santos *et al.*'s metod [92]. ICP-AES, i CAP 600 series, and microwave equipment (milestone ETHOS lab station with easy Control software HRP1000/105 high-pressure segmented rotor) were used for analysis. The advanced microwave digestion method was used first for the digestions of all samples using the microwave drying technique with  $\text{HNO}_3$  and  $\text{H}_2\text{O}_2$ . Argon gas was used for the excitation of all elements' atoms, and blank values of each

element were deducted from the samples' values. Analyzed coffee samples included LRCM, GCS, RCA, GCA, HRKC, LRCK, BRK, GCK, GCE, and BRA, **Table 5**.

- **Statistical analysis**

Data were analyzed using SPSS,1,2 for windows (Boston, USA). Data were expressed as mean  $\pm$  standard deviation. Differences between different samples were statistically significant when  $P \leq 0.05$ .  $IC_{50}$  was calculated from Graph prism 8 software.

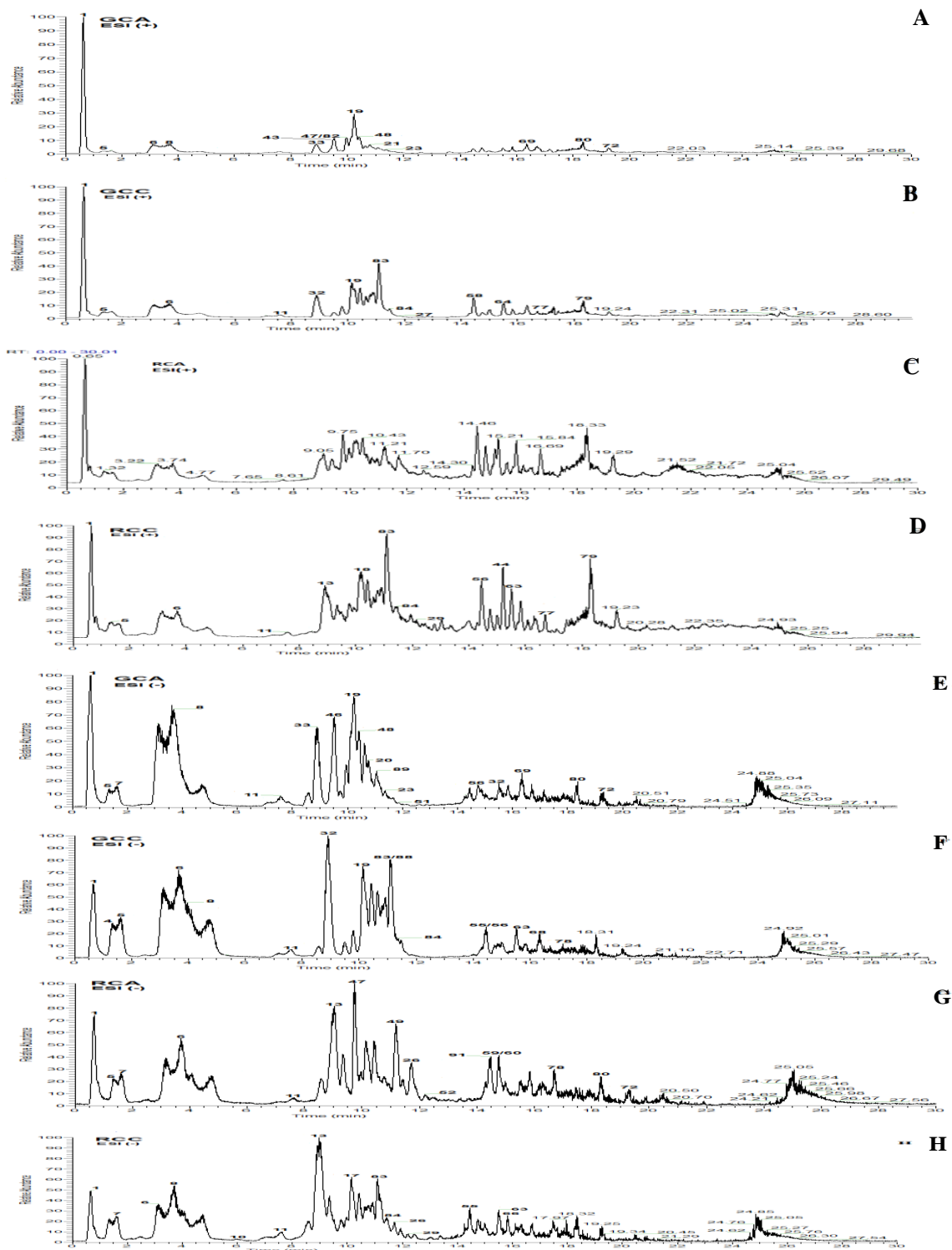
**Part I**

**Phytochemical study**

## **Chapter 1: Metabolite's profiling of coffee *via* UPLC / ESI/ MS coupled to chemometrics.**

### **1.1 Metabolites profiling of green and roasted coffee beans of arabica and canephora species**

The main objective of this study was to identify heterogeneity in coffee secondary metabolites in the context of its geographical origin, roasting process, different blends in the Middle East. To achieve such an objective, methanol extracts prepared from coffee specimens were analyzed using UPLC/MS. The metabolite identification included time, accurate mass, fragments, isotopic distribution, UV–vis spectra, and error) reported from literature and natural product database of the phytochemical dictionary. The analysis was performed in both positive and negative ionization modes, and the ion mass spectra derived from cations (M+H) or anions (M–H) accompanied by many fragmentation patterns aided in metabolites annotation. The metabolites eluted in the order of alkaloids, hydroxycinnamic acids, diterpenoids, *N*-alkonyl fatty acids, sphingolipids, and fatty acids. A total of 91 peaks are annotated belonging to different metabolites classes 23 hydroxycinnamates (*i.e.*, chlorogenic acid), 5 hydroxycinnamic amides, 17 diterpenes (*i.e.*, Cafestol), 15 fatty acids (*i.e.*, Trihydroxy-octadecanoic acid), 7 sphingolipids, 6 nitrogenous compounds (*i.e.*, *N*-tricosanoyl-hydroxytryptamine), 2 alkaloids (*i.e.*, caffeine), 2 sugars (*i.e.*, Acetyl-di-feruloyl sucrose), 1 coumarin (dihydroxypsoralen-*O*-hexoside) and 1 fatty acid amide (docosenamide), **Figure 5 & Table 6.**

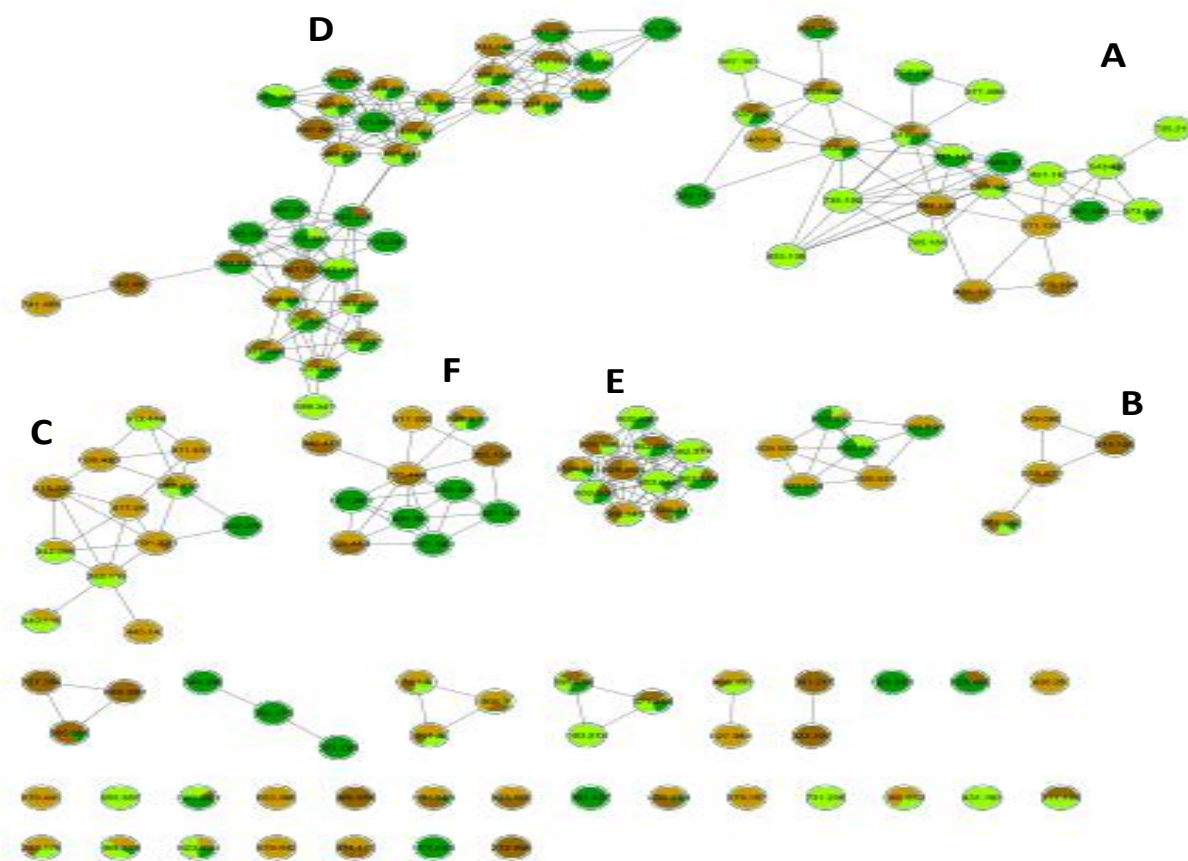


**Figure 5:** Representative UPLC/MS base peak chromatograms of green *Coffea arabica* (A), green *Coffea canephora* (B), roasted *Coffea arabica* (C), roasted *Coffea canephora* (D) in positive ionization mode. In negative ionization mode, green *Coffea arabica* (E), green *Coffea canephora* (F), roasted *Coffea arabica* (G), roasted *Coffea canephora* (H).

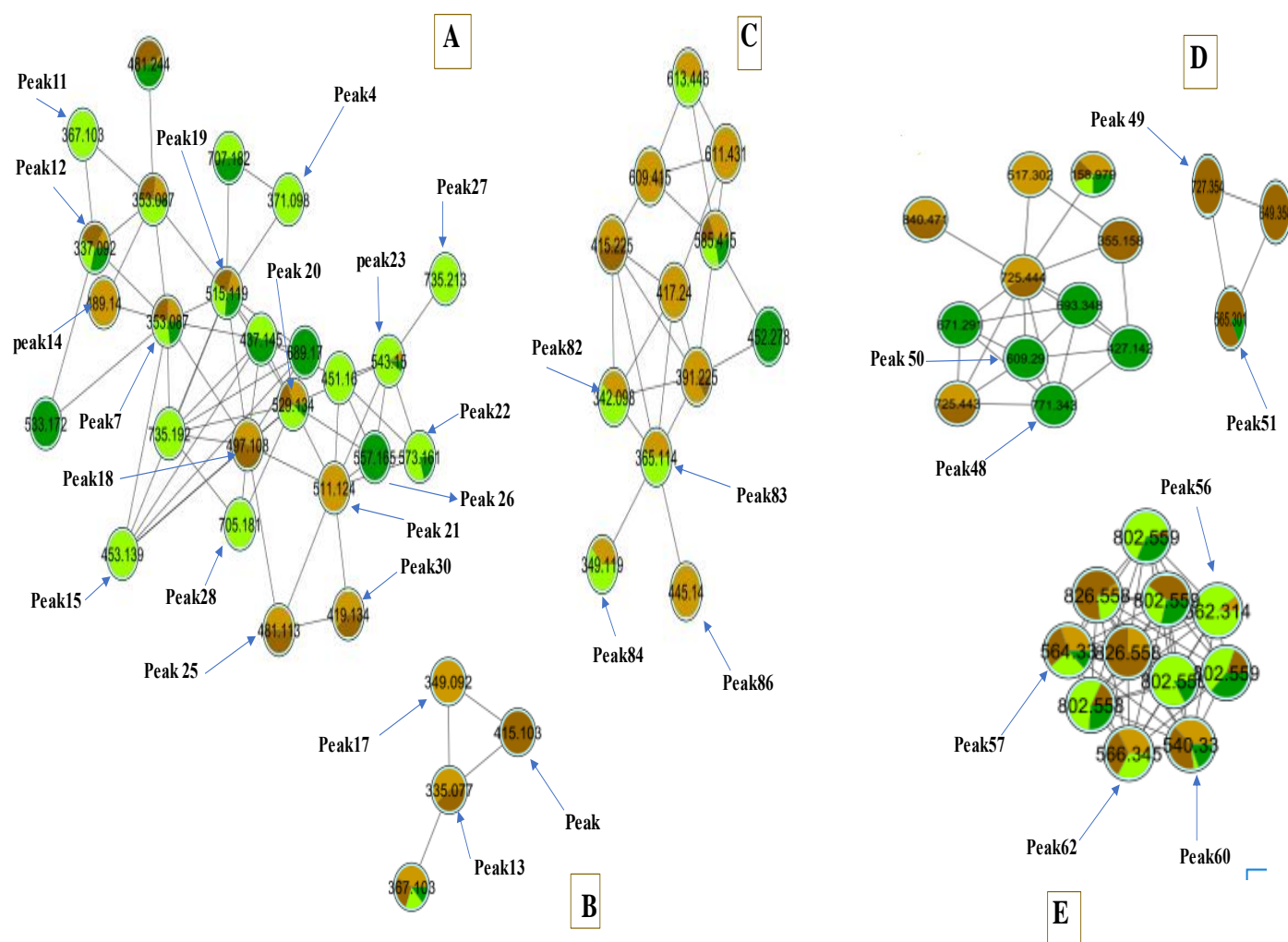


## 1.2 UPLC/MS metabolites identification using global natural products social networking (GNPS)

GNPS (Global natural products social networking) has been successfully applied for visualization of coffee metabolites obtained from the UP LC/MS platform [24]. The graphical display has aided in the annotation and dereplication of the metabolites obtained from the UPLC/MS datasets. The created molecular networking (MN) encompassed 145 connected nodes consisted of 11 clusters where the nodes of this network represent the compounds' parent ion, and the colors of the node represent the roasting and species attributes provided from the metadata file, **Figure 6**. The identified metabolites grouped in classes are listed in **Table 6** are detected from UPLC/MS data and GNPS.



**Figure 6:** Full molecular networking from UPLC/MS-MS data in negative ionization mode from green *Coffea canephora* (dark green color), green *Coffea arabica* (light green), roasted *Coffea canephora* (dark brown color), roasted *Coffea arabica* (light brown). The network exhibited a pie chart reflecting parent ions and their abundance in each sample.



**Figure 7:** Molecular network of major clusters created from coffee samples (GCA, GCC, RCA, RCC). For all the networks, nodes are color-coded based on the roasting and species present and labelled by their parent ions. Light and dark green corresponds to green *Coffea canephora* and green *Coffea arabica*, respectively, while light and dark brown correspond to roasted *Coffea canephora* and roasted *Coffea arabica*, respectively. **(A):** Molecular network of hydroxycinnamates, **(B):** Hydroxycinnamic lactones, **(C):** Hydroxycinnamic amides, **(D):** Diterpenes glycosides **(E),** Sphingolipids.

**Table 6:** Metabolites identified in methanol extracts of authenticated green *C.canephora* (GCC), green *C. arabica* (GCA), roasted *C. canephora* (RCC), and roasted *C. arabica* (RCA) via UHPLC-PDA-ESI-MS in both negative and positive ionization modes.

Peak no	Metabolite	R <sub>t</sub> (min)	UV max (nm)	Mass error (ppm)	Mol. formula	[M-H]	[M+H]	MS <sup>n</sup> ions ( <i>m/z</i> )	Coffee speci men
<b>Coumarins</b>									
P1	Dihydroxypsoral en- <i>O</i> -hexoside	0.66	265	7.92	C <sub>17</sub> H <sub>17</sub> O <sub>10</sub> <sup>+</sup>		381.07895	219, 201	RCA RCC GCA GCC
<b>Organic acids</b>									
P2	Quinic acid	0.67	265	-0.11	C <sub>7</sub> H <sub>11</sub> O <sub>6</sub> <sup>-</sup>	191.0 5589		111, 173	RCA RCC GCA GCC
P3	Isocitric acid	0.89	372	-0.13	C <sub>6</sub> H <sub>7</sub> O <sub>7</sub> <sup>-</sup>	191.0 1965		111, 173, 155, 127	RCA RCC GCA GCC
<b>Phenolic acid glycoside</b>									
P4	Dihydroferulic acid 4- <i>O</i> - glucuronide	0.94	---*	-1.11	C <sub>16</sub> H <sub>19</sub> O <sub>10</sub> <sup>-</sup>	371.0 9756		353, 191, 135	GCC
<b>Alkaloids</b>									
P5	Trigonelline	1.37	---	-1.0	C <sub>7</sub> H <sub>7</sub> NO <sub>2</sub> <sup>+</sup>		138.08978	120, 110, 69, 90	RCA RCC GCA GCC
P6	Caffeine	3.54	---	-5.4	C <sub>8</sub> H <sub>10</sub> N <sub>4</sub> O <sub>2</sub>		195.19037	137	RCA

									+		RCC GCA GCC
<b>Hydroxycinnamates derivatives</b>											RCA RCC GCA GCC
P7	3- <i>O</i> -Caffeoyl quinic acid	1.9	221,325	-0.1	C <sub>16</sub> H <sub>17</sub> O <sub>9</sub> <sup>-</sup>	353.0 8777	191, 179, 135				
P8	5- <i>O</i> -Caffeoyl quinic acid	3.69	221,325	-1.06	C <sub>16</sub> H <sub>17</sub> O <sub>9</sub> <sup>-</sup>	353.0 8771	191, 179				RCA RCC GCA GCC
P9	4- <i>O</i> -Caffeoyl quinic acid	5.09	221,325	-1.06	C <sub>16</sub> H <sub>17</sub> O <sub>9</sub> <sup>-</sup>	353.0 8844	173, 179, 191				RCA RCC GCA GCC
P10	Caffeoyl shikimic acid	6.15	301,284	-2.18	C <sub>16</sub> H <sub>15</sub> O <sub>8</sub> <sup>-</sup>	335.0 7687	179, 161, 135				RCA RCC
P11	Feruloyl quinic acid isomer	7.66	221,325	-1.07	C <sub>17</sub> H <sub>19</sub> O <sub>9</sub> <sup>-</sup>	367.1 0324	161, 193, 135				RCA RCC GCA GCC
P12	<i>p</i> -Coumaroyl quinic acid	7.86	221,325	-1.69	C <sub>16</sub> H <sub>17</sub> O <sub>8</sub> <sup>-</sup>	337.0 9256	191, 163				RCA RCC GCA GCC
P13	Caffeoyl-quinolactone	8.99	221,325	-1.55	C <sub>16</sub> H <sub>15</sub> O <sub>8</sub> <sup>-</sup>	335.0 7648	161, 135, 179				RCA RCC
P14	Unknown chlorogenic acid derivative **	9.66	218,322	-0.76	C <sub>24</sub> H <sub>25</sub> O <sub>11</sub> <sup>-</sup>	489.1 3986	353, 315, 255, 191, 297				RCC
P15	Unknown diacyl chlorogenic acid	9.7	220,320	-2.3	C <sub>21</sub> H <sub>25</sub> O <sub>11</sub> <sup>-</sup>	453.1 3919	353, 335, 291				GCC

derivative**								
P16	Methyl - <i>O</i> -feruloylquininate	10.01	218,32 6	-2.06	$C_{18}H_{21}O_9^-$	381.1 1981	175, 160, 193	RCC RCA
P17	Feruloyl-quinolactone	10.15	218,32 6	-0.51	$C_{17}H_{17}O_8^-$	349.0 9271	175, 193, 149, 134	RCC RCA
P18	Dicaffeoyl-quinolactone**	10.19	325	-0.27	$C_{25}H_{23}O_{11}^-$	497.1 075	335	RCA RCC
P19	Dicaffeoyl quinic acid isomer	10.23	220,32 5	-1.76	$C_{25}H_{23}O_{12}^-$	515.1 1859	353, 335	RCA RCC GCA GCC
P20	Caffeoyl-feruloylquinic acid	10.82	325	-1.75	$C_{26}H_{25}O_{12}^-$	529.1 3428	367, 353	RCA RCC GCA GCC
P21	Caffeoyl-feruloyl quinolactone**	11.74	220,32 5	-1.8	$C_{26}H_{23}O_{11}^-$	511.1 2366	335, 179.161	RCC
P22	Sinapoyl-feruoylquinic acid	11.28	221,32 4	-1.57	$C_{28}H_{29}O_{13}$	573.1 6046	349, 397	GCA GCC
P23	Diferuloylquinic acid.	11.31	221,32 4		$C_{27}H_{27}O_{12}^-$	543.1 499	367, 349	RCC GCC
P24	Caffeoyl-	11.30	222,32	-1.57	$C_{27}H_{27}O_{12}$	543.1	381.367,	GCA

	dimethoxycinnamoylquinic acid		4			500		335	
P25	Unknown quinolactone derivative **	11.6	221	-2.37	$C_{25}H_{21}O_{10}^-$	481.1 1288		335, 179, 161	RCC RCA
P26	Feruloyl - dimethoxycinnamoylquinic acid	11.75	222	1.62	$C_{28}H_{29}O_{12}^-$	557.1 684		381, 349	GCA GCC
P27	Triacyl- <i>O</i> -caffeoyl- <i>O</i> -feruloyl- <i>O</i> -sinapoylquinic acid (new)	11.9	222	-1.91	$C_{37}H_{35}O_{16}^-$	735.1 9165		573, 529	GCC
P28	Di- <i>O</i> -feruloyl- <i>O</i> -caffeoylquinic acid	11.98	222	-3.9	$C_{36}H_{33}O_{15}^-$	705.1 7969		543, 529	GCC
P29	Caffeoyl-feruloylquinic acid lactone	12.1	222	-2.13	$C_{26}H_{23}O_{11}^-$	437.1 4474		335, 193, 179	RCC
P30	Unknown quinolactone derivative	12.23	223	-2.8	$C_{21}H_{23}O_9^-$	419.1 3358		335, 317, 255, 179	RCA RCC
P31	Unknown chlorogenic acid	13.15	222	-1.23	$C_{21}H_{25}O_{10}^-$	437.1 444		173, 275	GCA GCC
<b>Sugars</b>									
P32	Di-o-hexoside	0.7	*	-0.3	$C_{12}H_{21}O_{11}^-$	341.1 0883		*	GCA GCC
P33	Acetyl-diferuloylsucrose	8.85	221,325	-1.01	$C_{34}H_{39}O_{18}^-$	735.2 1344		367	GCC GCA

P34	Acylsucroses dihydroxycinna moyl	9.73	220,32 7	-0.34	$C_{29}H_{36}O_{18}^-$	671.2 9065	627	GCA
<b>Diterpenes</b>								
P35	Cafestol	9.39	222	-1.54	$C_{20}H_{29}O_3^+$	317.21063	299, 271, 253	GCA
P36	Trihydroxy- kauradien- olide**	9.45	217	-2.1	$C_{20}H_{27}O_5^+$	347.18457	329, 285	GCA
P37	Dehydrocafestol	9.49	220	-1.32	$C_{20}H_{27}O_2^+$	299.20016	145, 191, 281, 253	GCA RCA
P38	Mozambioside	10.03	298	-1.38	$C_{26}H_{37}O_{10}^+$	509.224	347, 329, 311	RCAG CA
P39	Bengalensol- <i>O</i> - hexoside	11.75	221	-1.38	$C_{26}H_{35}O_9^+$	491.2417	329, 311	RCA
P40	Trihydroxy- kauranoic acid	10.67	220	-2.21	$C_{20}H_{31}O_5^-$	351.2 171	289, 321	RCA
P41	Bengalensol	11.63	221	-1.35	$C_{20}H_{25}O_4^+$	329.17429	293, 311, 237	RCA
P42	Dihydroxy- kauren-oic acid	12.64	223	-0.218	$C_{20}H_{29}O_4^-$	333.2 0706	303	RCA
P43	16- methylKahweol	12.78	222	-1.79	$C_{20}H_{27}O_4^+$	331.19006	314, 296, 145, 279	RCC RCA
P44	Dehydro- kahweol	13.6	222	-1.33	$C_{20}H_{25}O_2^+$	297.18451	279, 145	GCA
P45	Dehydrocafestol derivative**	15.28	225	-1.62	$C_{20}H_{25}O^-$	281.18954	263, 173, 131	RCA RCC
<b>Diterpene glycosides</b>								
P46	Carboxyatractyli genin- <i>O</i> - hexoside	9.57	324,22 1	-1.74	$C_{26}H_{37}O_{11}^-$	525.2 3322	396, 203	GCA GCC



P47	Atracyligenin-O-hexoside	9.82	219,311	-1.1	$C_{25}H_{37}O_9^-$	481.2 4377	301	GCA GCC
P48	Desoxycarboxyatractyligenin-O-hexoside	10.5	220	0.05	$C_{37}H_{55}O_{17}^-$	771.3 4338	727	GCA
P49	Desoxyatractyligenin-O-hexoside	11.21	---	0.45	$C_{36}H_{55}O_{15}^-$	727.3 5376	643, 625	RCA GCA
P50	Carboxyatractyligenin-O-hexoside	11.72	221	-1.64	$C_{31}H_{45}O_{12}^-$	609.2 9065	565	GCA
P51	Isovaleryl-atractyligenin-O-hexoside derivative	11.82	221	-1.23	$C_{30}H_{45}O_{10}^-$	565.3 009	481, 463, 303	RCA GCA
<b>Fatty acids and sphingolipids</b>								
P52	Trihydroxy-octadecaenoic acid	12.2	223	-1.87	$C_{18}H_{33}O_5^-$	329.2 3273	311, 293, 229, 171	RCA GCA
P53	hexosyl)-2-(pentanoyloxy) propyl dodecenoate	13.03	223	-1.29	$C_{26}H_{45}O_{10}^-$	517.3 02	473, 367	RCC
P54	Linoleic acid methyl ester**	13.19	223	-1.47	$C_{17}H_{25}O_4^-$	293.1 7523	236, 221	RCC GCC
P55	Unknown fatty acid	13.6	222	4.18	$C_{14}H_{29}O_8^-$	325.1 8228	183	RCA RCC GCA GCC
P56	Sphingolipid conjugate I**	14.00	222	1.39	$C_{27}H_{49}NO_9P^-$	562.3 1610	502	RCA RCC

								GCA GCC
P57	Sphingolipid conjugate II**	14.55	222	-2.18	$C_{27}H_{51}NO_9$ $P^-$	564.3 2947	504	RCA RCC GCA GCC
P58	Phosphatidyl inositol hexanoic acid derivative	14.68	224	-3.21	$C_{25}H_{48}O_{12}$ $P^-$	571.2 8705	391, 315, 255, 241	RCA RCC GCA GCC
P59	Ceramide conjugate I**	14.75	224	-0.45	$C_{22}H_{49}O_6N$ $_4P$		496.33881 478, 184	RCA RCC GCA GCC
P60	Sphingolipid conjugate III**	14.78	222	-1.9	$C_{25}H_{51}NO_9$ $P^-$	540.3 2941	480	RCA RCC GCA GCC
P61	Unknown fatty acid	14.85	224	-7.34	$C_{13}H_{27}O_8^-$	311.1 6809	183	RCA RCC GCA GCC
P62	Sphingolipid conjugate IV**	14.97	224	-1.45	$C_{27}H_{53}NO_9$ $P^-$	566.3 4552	506	RCA RCC GCA GCC
P63	Ceramide conjugate II**	15.08	221	-0.98	$C_{21}H_{45}O_2N$ $_9P_2^+$		522.35461 504, 184	RCA RCC GCA GCC
P64	Decanoic acid, propanediol	15.48	222	8.46	$C_{24}H_{51}O_{10}^-$	499.3 5391	481, 455, 322, 279	RCA RCC GCA

								GCC
P65	Ceramide conjugate III**	15.7	222	-1.42	$C_{23}H_{46}O_8N_5$	520.37042	502, 184	RCC GCC GCC GCA
P66	Unknown fatty acid	15.77	224	-0.06	$C_{20}H_{38}O_{2N}$	324.28943	307, 263, 245	RCC RCA
P67	Hydroxy-octacosadodecenoic acid	16.22	225	-8.13	$C_{28}H_{31}O_3^-$	415.2 2482	279	RCC
P68	Unknown fatty acid	16.3	225	-8.08	$C_{26}H_{55}O_{10}^-$	527.3 8416	509, 350, 307	RCC RCA
P69	Hexanedioic acid, propanediol ester	16.36	224	10.47	$C_{24}H_{51}O_9^-$	483.3 5849	465, 324, 158	GCA GCC
P70	Unknown fatty acid	16.49	225	-9.23	$C_{15}H_{31}O_8^-$	339.1 9919	183	RCA RCC GCA GCC
P71	Dimethyl octadecanedioate	19.18	226	0.83	$C_{20}H_{37}O_4^-$	341.2 6892	313, 269	RCA RCC GCA GCC
P72	Hydroxy-docosanoic acid	19.35	224	-2.97	$C_{22}H_{43}O_3^-$	355.3 2135	309	RCA RCC GCA GCC
P73	Hydroxy-tetracosanoic acid	20.58	222	-2.5	$C_{24}H_{47}O_3^-$	383.3 5208	337	RCA RCC GCA GCC
P74	Unknown fatty	20.97	227	-0.03	$C_{38}H_{55}O_3^+$	559.41382	541, 279,	RCC

	acid ester							223, 183	GCC
<b>Fatty acyl amides</b>								RCA RCC GCC GCA	
P75	docosenamide) **	16.69	226	-1.06	$C_{22}H_{44}NO$	338.3 4137	321, 303		
<b>Nitrogenous compounds</b>									RCA GCA RCC GCC
P76	<i>N</i> -Heneicosanoyl-hydroxytryptamine.	15.92	226	-1.4	$C_{30}H_{51}N_2O_3^+$	487.39	469, 177, 160		
P77	<i>N</i> -tricosanoyl-hydroxytryptamine	16.69	226	-0.65	$C_{23}H_{55}N_2O_3^+$	515.42	497, 177, 160		RCA RCC GCA GCC
P78	<i>N</i> -docosanoyl-hydroxytryptamide	19.3	226	-1.09	$C_{32}H_{55}N_2O_5^+$	499.42	482, 177, 160		RCA RCC GCA GCC
P79	<i>N</i> -octadecanoyl-hydroxytryptamide	17.46	226	-1.81	$C_{28}H_{46}N_2O_2^+$	443.36185	426, 177, 160		RCA RCC GCA GCC
P80	<i>N</i> -Eicosanoyl-hydroxytryptamide	18.32	226	-1.68	$C_{30}H_{51}N_2O_2^+$	471.39349	454, 177, 160		RCA RCC GCA GCC

P81	<i>N</i> -tetracosanoyl-hydroxytryptamide.	20.32	226	-0.95	$C_{34}H_{59}N_2O_{2+}$	527.46	510.177.1 60	RCA RCC GCA GCC
<b>Hydroxycinnamates amide</b>								
P82	Caffeoyl tyrosine	9.8	220	9.75	$C_{18}H_{16}NO_6^-$	342.0 9775	206	GCA GCC
P83	Caffeoyl- <i>N</i> -tryptophan	11.08	221	-0.61	$C_{20}H_{17}N_2O_5^-$	365.1 1395	135,229	GCC RCC
P84	<i>p</i> -coumaroyl- <i>N</i> -tryptophan	11.49	221	-2.0	$C_{20}H_{17}N_2O_4^-$	349.1 1868	229	RCC GCC
P85	Feruloyl- <i>N</i> -tryptophan	11.66	222	9.14	$C_{21}H_{19}N_2O_5^-$	379.1 2805	335, 229	GCC
P86	Unknown hydroxy cinnamic acid amide**	12.1	229	-6.59	$C_{25}H_{21}O_6N_2^-$	445.1 405	309, 161, 229	RCA RCC
<b>Unknowns</b>								
P87	Unknown	9.69	222	-1.07	$C_{19}H_{25}O_2^+$	285.18472	267, 239, 229	RCA
P88	Unknown	15.3	222	-2.3	$C_{28}H_{33}O_4^-$	433.23483	153	GCA
P89	Unknown	14.41	---	-14.16	$C_{27}H_{47}O_{14}^-$	595.2 8729	415, 315, 279, 241	RCA RCC
P90	Unknown	11.81	222	-0.64	$C_{22}H_{27}O_{10}$	451.1 5973	349, 275, 173	GCA GCC
P91	Unknown	17.09	266	-9.88	$C_{26}H_{31}O_3$	391.2 2531	255, 153	

---

\*: not detected.

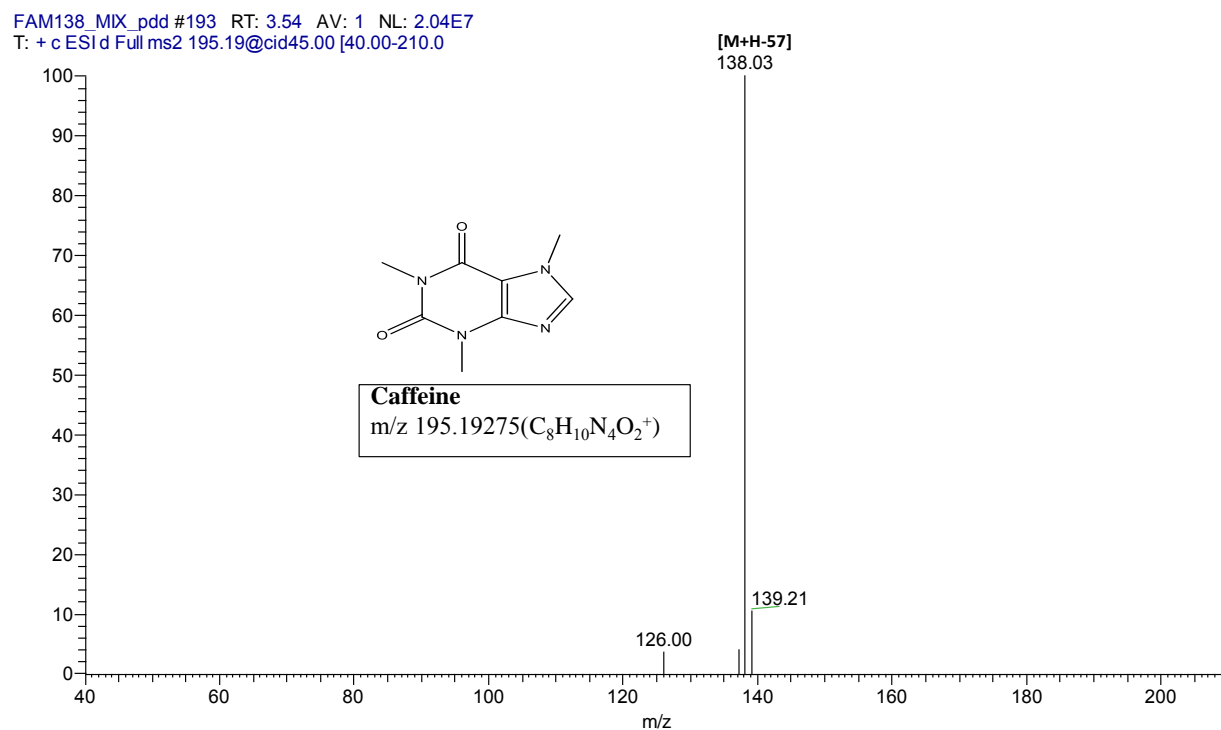
\*\*: reported for the first time.

### 1.3 UPLC/MS metabolites profiling of coffee seed extracts

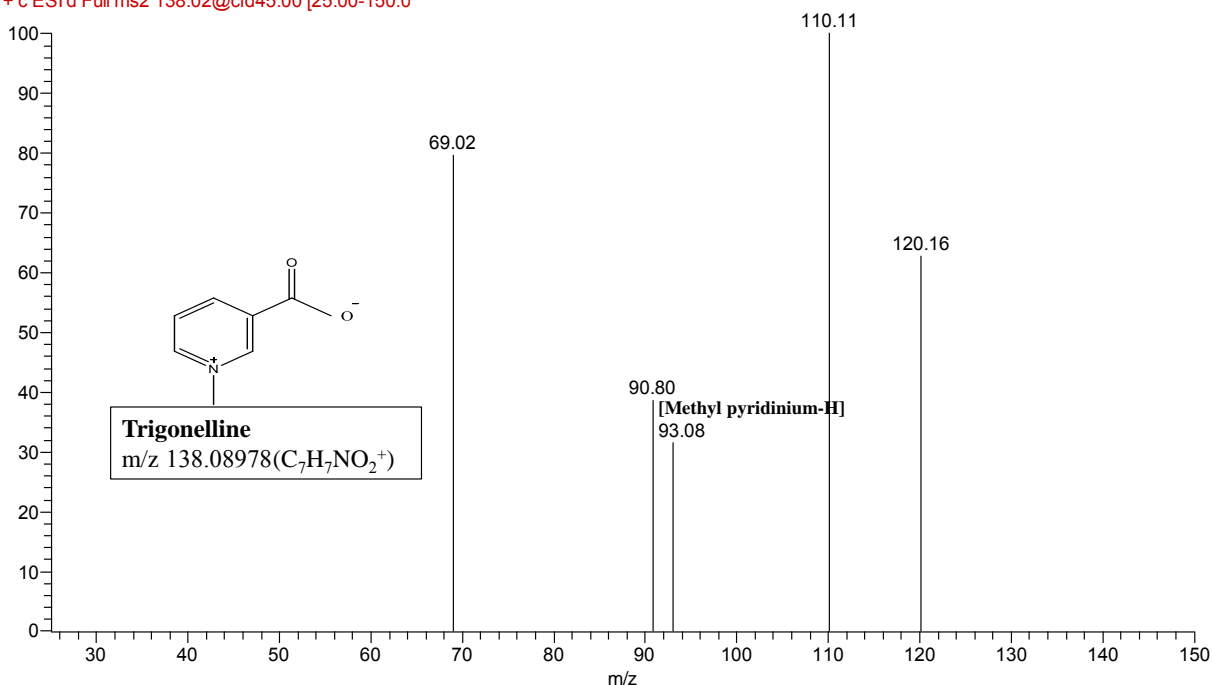
#### 1.3.1 Alkaloids

Caffeine, a characteristic component of coffee beans, belongs to xanthine alkaloids (1,3,7-trimethylxanthine) and employs several biological values, especially as a central nervous system stimulant[93]. It was detected in positive ionization mode owing to the presence of a nitrogen atom in peak (P6) at  $m/z$  195 and yielding fragment ion at  $m/z$  138 for the loss of methyl isocyanate (57 Da),

**Figure 8.** Trigonelline is a bitter alkaloid that contributes to coffee flavor, its level in roasted beans decreases upon the increasing degree of roasting. Moreover, trigonelline was also detected at  $m/z$  138 (P5) and produces fragment ion at  $m/z$  93 (methyl pyridinium), which is common upon degradation of trigonelline, **Figure 9, Table 6**, [94].



**Figure 8:** MS/MS spectrum of peak (P6) in the positive ion mode.



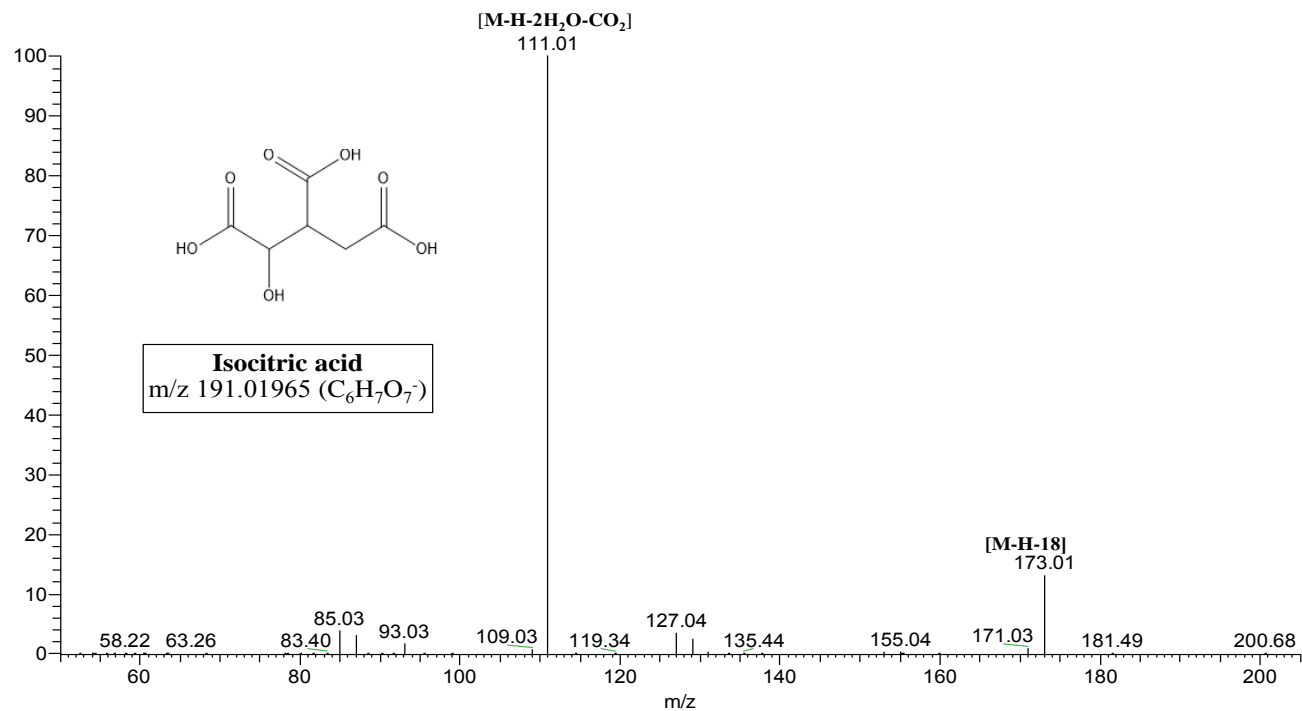
**Figure 9:** MS/MS spectrum of peak (P5) in positive ion mode.

### 1.3.2 Organic acids

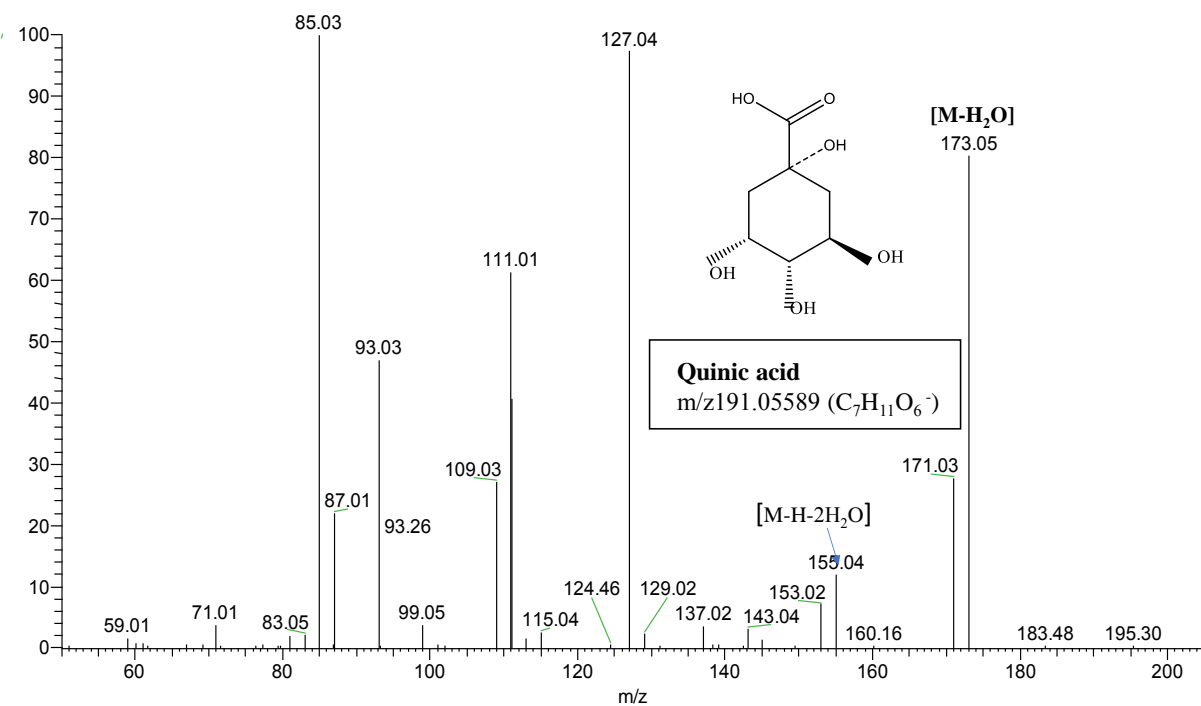
Quinic acid (QA) and (iso) citric acid (IA) are common organic acids that can be esterified by hydroxycinnamic acid (HCA) to form hydroxycinnamoylquinic acids or hydroxycinnamoyl-(iso)citric acids, respectively, [27]. Herein, the similarities and differences between QA and IA fragmentation patterns are highlighted using UPLC/MS in negative ionization mode.

Diagnostic fragments of QA and IA were reported for the detection of IA in pineapple juice and QA in coffee [27]. In **Table 6**, QA (P2) and IA (P3) showed the same molecular ion at  $m/z$  191, albeit base peak at  $m/z$  173 (M-H-H<sub>2</sub>O) for QA and  $m/z$  111 (M-H-2H<sub>2</sub>O-CO<sub>2</sub>) for IA could distinguished between the two acids. Moreover, diagnostic fragment peak at  $m/z$  155 (M-H-2H<sub>2</sub>O) was detected for QA and absent for IA (**Figure 10 & 11**).





**Figure 10:** MS/MS spectrum of peak (P3) in the negative ionization mode.



**Figure 11:** MS/MS spectrum of peak (P2) in the negative ionization mode.

### 1.3.3 Hydroxy cinnamates

Previous studies have reported several HCA derivatives as major chlorogenic acids in coffee [28]. These compounds included mono-acylquinic acids such as mono-caffeoyl, mono-*p*-coumaroyl, and mono-feruloyl or diacyl as di-caffeoylquinic acid isomers, and triacyl as *O*-caffeoyl-*O*-feruloyl-*O*-sinapoylquinic acid. In the negative ionization mode, the most abundant peaks were attributed for caffeoylquinic acid isomers, *i.e.*, 3-*O*-caffeoylquinic acid (P7), 5-*O*-caffeoylquinic acid (P8), and 4-*O*-caffeoylquinic acid (P9) and constituted cluster **A** in the molecular network obtained from GNPS, Error! Reference source not found.. All isomers showed a deprotonated parent ion peak (M-H)<sup>-</sup> at *m/z* 353, with further fragment ions at *m/z* 191 (M-H-caffeoyl) as a base peak, 179 (M-H-quinic), and 135 (M-H-quinic-CO<sub>2</sub>). However, the elution order could differentiate between different isomers and assigned as 3CQA, 5CQA, and 4CQA, respectively (Figures 12 &13) [35].

Later in elution time, diacyl CGAs, *i.e.*, dicaffeoyl quinic acid isomers (P19), showed (M-H)<sup>-</sup> at *m/z* 515 annotated as dicaffeoylquinic acid with fragment ions at *m/z* 353 (M-162) and 335 (M-180), corresponding to (caffeic acid-H<sub>2</sub>O), and caffeic acid loss, respectively, **Figure 14** [95].

Another peak (P20) at *m/z* 529 for (M-H)<sup>-</sup> and fragment ions at *m/z* 367 and *m/z* 353 correspond to the respective losses of caffeoyl and feruloyl moieties was annotated as caffeoyl-feruloylquinic acid **Figure 15** [96].

In agreement with the literature, several minor chlorogenic acid derivatives were also characterized in both green and roasted samples of both coffee species. For example, peak (P23) was detected at *m/z* 543 with fragment ions at *m/z* 367 (feruloylquinic acid-H), and 349 (feruloylquinic acid-H<sub>2</sub>O-H) corresponding to di-feruloyl quinic acid detected mainly in green canephora and to a less extent in roasted canephora **Figure 16**. Intriguingly, another peak at *m/z* 543 (P24) at the same retention time only in green arabica (GCA) produced fragment peaks at *m/z* 381 [dimethoxycinnamoyl-H], 367 [M-caffeoyl], and 335 [caffeoylquinic acid-H<sub>2</sub>O] and annotated as caffeoyl -dimethoxy cinnamoyl quinic acid **Figure 17** [29]. Besides, a peak of the deprotonated ion at *m/z* 381 was more abundant in RCC than RCA annotated as methyl-*O*-feruloyl quinate (P16) and confirmed by its fragment ions at *m/z* 175 (ferulic-H<sub>2</sub>O-H<sup>+</sup>), *m/z* 193 (ferulic-H<sup>+</sup>), and *m/z* 349 (feruloyl-quinolactone) **Figure 18**. Furthermore, a peak derived from sinapic acid was assigned to sinapoyl-feruloyl quinic acid (P22) at *m/z* 573 and fragment ions at *m/z* 397

(sinapoylquinic acid-H) and 349 (feruloylquinic acid-H<sub>2</sub>O-H) and previously reported in canephora but reported for the first time in green Arabica coffee (GCA) **Figure 19** [57].

Likewise, triacyl quinic acid was identified as deprotonated molecular ion at  $m/z$  705 and fragment ions at  $m/z$  543 and 529, with loss of caffeoyl moiety indicating the presence of di-*O*-feruloyl-*O*-caffeoylquinic acid (P28) and considered as a marker for canephora **Figure 20** [29]. Another triacyl quinic acid detected for the first time in canephora coffee (GCC) at  $m/z$  735 and tentatively identified as triacyl *O*-caffeoyl-*O*-feruloyl-*O*-sinapoylquinic acid (P27). The produced base peak at  $m/z$  573 and fragment ion  $m/z$  529 supported the annotation of these peaks as triacyl quinic acid **Figure 21**. Moreover, its identification was further supported by the obtained MN to display structural similarity to di-*O*-feruloyl-*O*-caffeoylquinic acid (P28). Furthermore, another compound tentatively assigned in negative ion mode is of great interest as a marker for green canephora (GCC). It yielded (M-H)<sup>-</sup> ion at  $m/z$  371 and MS2 at  $m/z$  353, 191, 173 and 135. The loss of 162 (hexose) directly connected to dicaffeoylquinic acid in the MN suggested this compound as phenolic acid glycoside and assigned as dihydro ferulic acid-*O*-glucuronide (P4), **Table 6 & Figure 22** [95].

Aside from hydroxycinnamic esters, lactones (quinides) were identified in both RCC and RCA samples and confirmed by clustering together in the MN **Figure 7B**, peak (P17) showed deprotonated ion (M-H) peak at  $m/z$  349 and MS2 fragments at  $m/z$  175 (ferulic acid-H-H<sub>2</sub>O) and  $m/z$  193 (ferulic acid-H) annotated as feruloyl-quinolactone, **Figure 23** [97].

Furthermore, in both arabica and canephora roasted samples, caffeoyl shikimic acid (CSA) and caffeoyl-quinolactone (CQL) displayed a similar (M-H) at  $m/z$  335 and fragment ions at  $m/z$  179(M-H-caffeoyl), 161(M-H-caffeoyl-H<sub>2</sub>O), and 135 (M-H-caffeoyl-CO<sub>2</sub>) in P10 and P14. However, shikimates showed a base peak at  $m/z$  179, while lactones at  $m/z$  161, **Figure 24 & Table 5**, [56]. Another lactone was identified as dicaffeoylquinolactone (P18) M+H at  $m/z$  497, and fragment ion at  $m/z$  335 (caffeoyl-quinolactone) revealed from MN, **Figure 7A & Figure 26**.

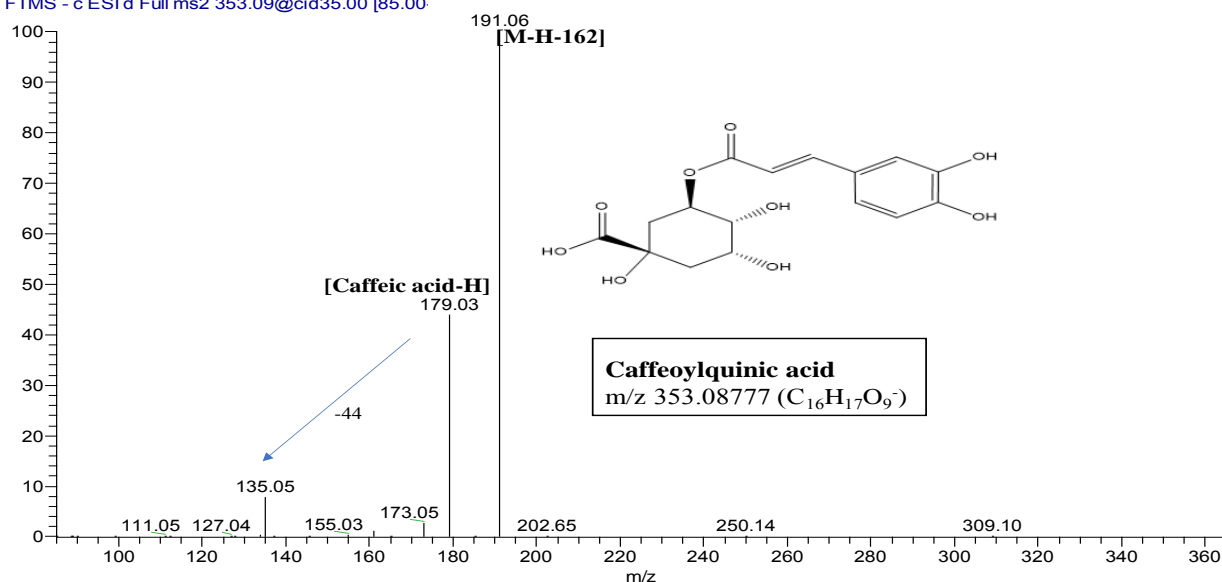
New lactone detected only in RCC was predicted with the aid of GNPS, **Figure 7A** (P21), (M-H) at  $m/z$  511, which is a dehydrated form of peak (P20) detected at  $m/z$  529, and fragment ions at  $m/z$  335 (caffeoyl lactone), 179 (caffeic-H) and 161(caffeic-H-H<sub>2</sub>O), suggesting for the

presence of caffeoyl-feruloylquinolactone reported for the first time, **Figure 27**. Besides, two novel peaks (P25) and (P30) were detected in both roasted specimens (M-H) at  $m/z$  481 and 419 respectively, sharing fragment ions at  $m/z$  335, 179, 161 suggesting for a chlorogenic lactone derivative, **Figure 28 & Figure 29**.

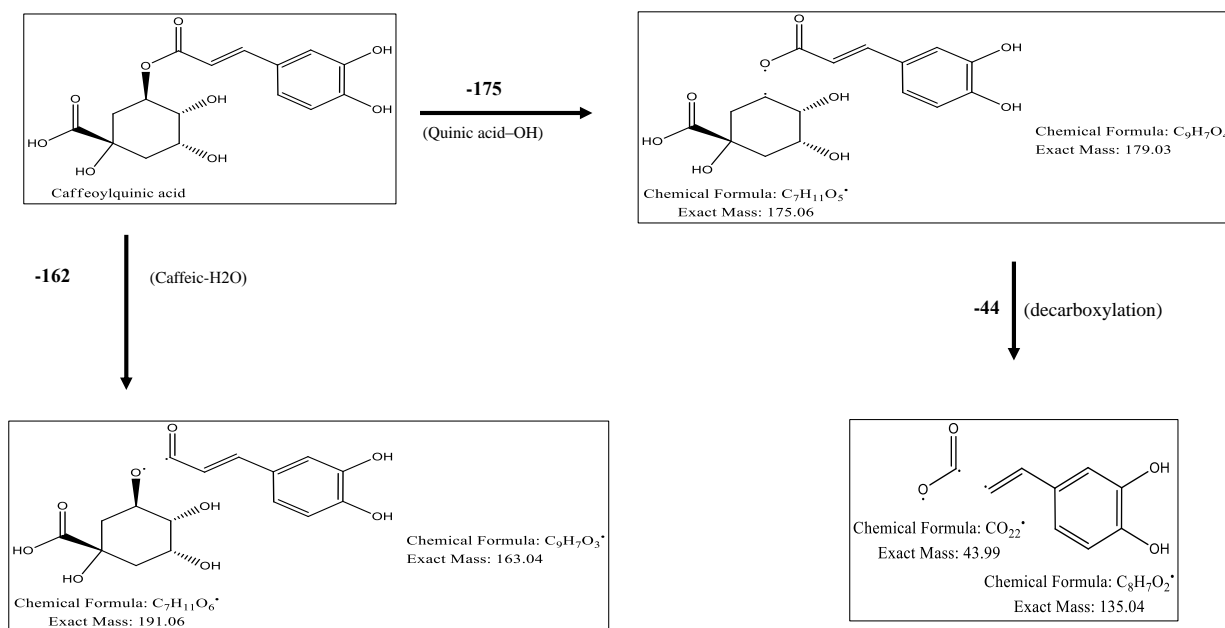
The presence of nitrogen atom was also revealed in hydroxycinnamate amides detected in both arabica and canephora specimens. Two peaks were detected at  $m/z$  349 and 365 for *p*-coumaroyl tryptophan (P84) and caffeoyl tryptophan (P83), respectively. Both compounds were previously reported as markers for green canephora species [96], with a loss of 120 Da corresponding to (*p*-coumaric-CO<sub>2</sub>), and 136 Da (caffeic acid-CO<sub>2</sub>) for coumaroyl tryptophan and caffeoyl tryptophan respectively, **Table 5**, **Figure 31** and **Figure 32**.

In negative mode, caffeoyl tyrosine (P82) was detected in GCA and GCC at  $m/z$  342, with MS/MS fragment at  $m/z$  206 corresponding to (caffeic-CO<sub>2</sub>) [96]. The UV band at 290 nm contributed to tryptophan and tyrosine, in addition to the bands of chlorogenic acid, *i.e.*, 220 and 321 nm. Finally, a novel tryptophan-derived hydroxycinnamic acid (P86) at  $m/z$  445 and fragment ion  $m/z$  at 229, which is characteristic for tryptophan identified for the first time in coffee. The molecular network **C** showed most of the identified hydroxycinnamic amides, as identified in **Table 5**, indicating their structural similarity (**Figure 30 & Figure 34**).

FAM139\_GCAn#358 RT: 1.33 AV: 1 NL: 1.49E5  
T: FTMS - c ESI d Full ms2 353.09@cid35.00 [85.00]

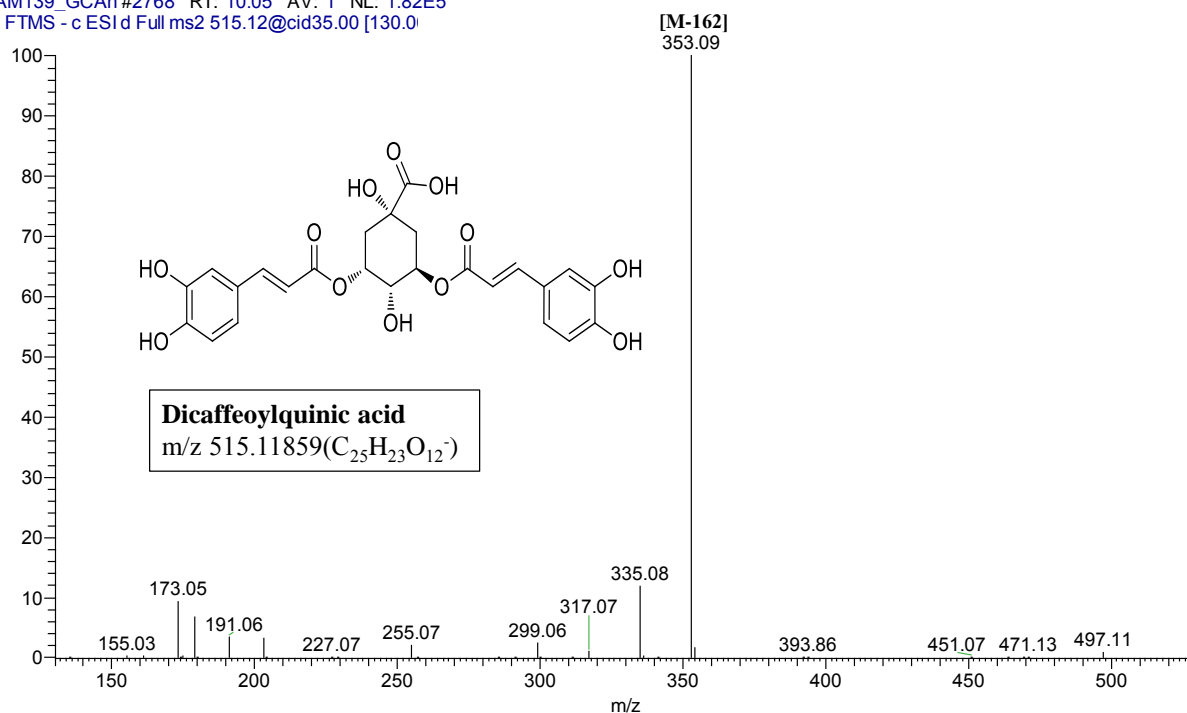


**Figure 12:** MS/MS spectrum of peak (P7) in the negative ion mode.



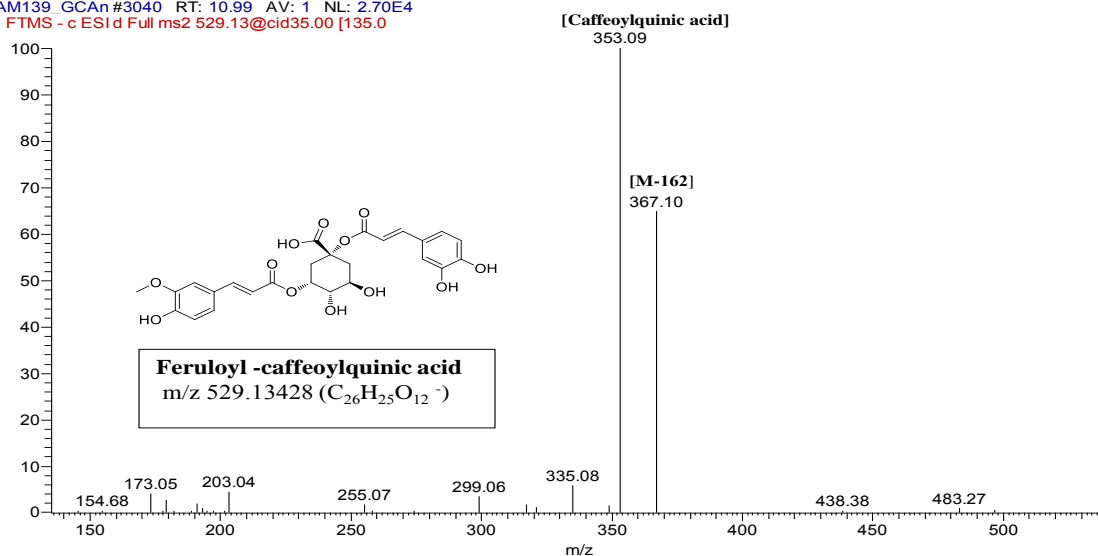
**Figure 13:** Fragmentation pattern for caffeoylquinic acids (P7, P8, P9)

FAM139\_GCAn#2768 RT: 10.05 AV: 1 NL: 1.82E5  
T: FTMS - c ESI d Full ms2 515.12@cid35.00 [130.0]



**Figure 14:** MS/MS spectrum of peak (P19) in negative ion mode

FAM139\_GCAn#3040 RT: 10.99 AV: 1 NL: 2.70E4  
F: FTMS - c ESI d Full ms2 529.13@cid35.00 [135.0]



**Figure 15:** MS/MS spectrum of peak (P20) in negative ion mode.

FAM139\_RCCn #3106 RT: 11.35 AV: 1 NL: 5.57E4  
 F: FTMS - c ESI d Full ms2 543.15@cid35.00 [135.0]

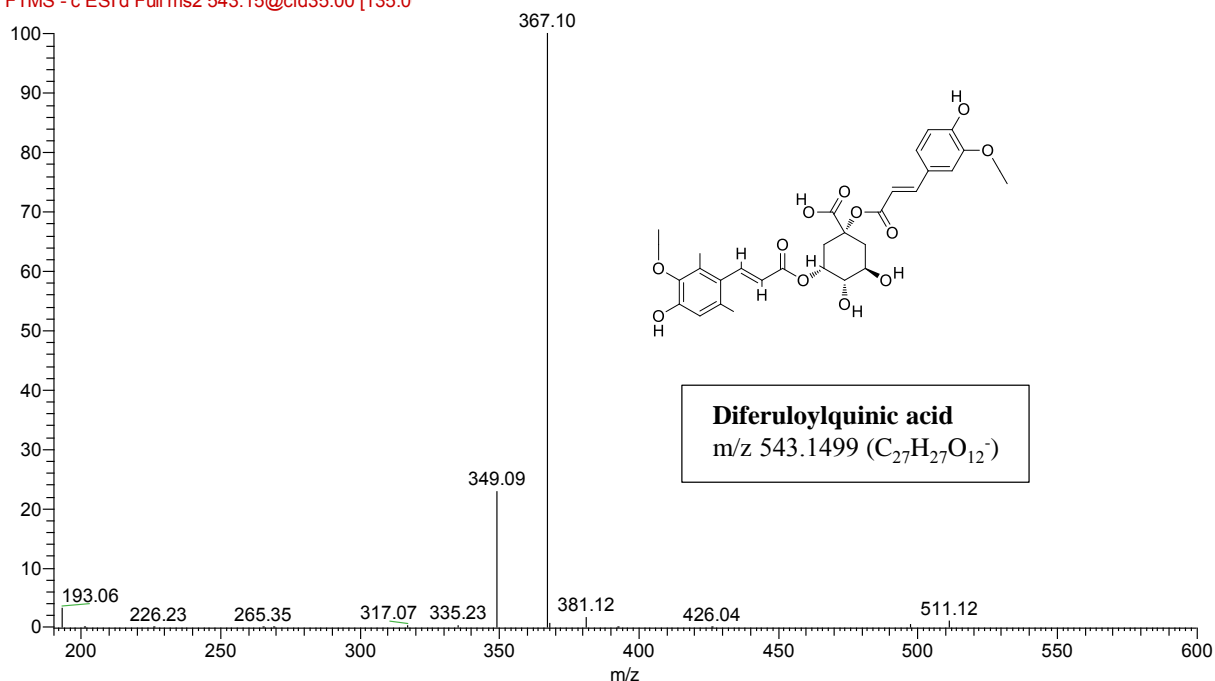


Figure 16: MS/MS spectrum of peak (P23) in negative ion mode.

FAM139\_GCA#3126 RT: 11.32 AV: 1 NL: 3.35E4  
 F: FTMS - c ESI d Full ms2 543.15@cid35.00 [135.0]

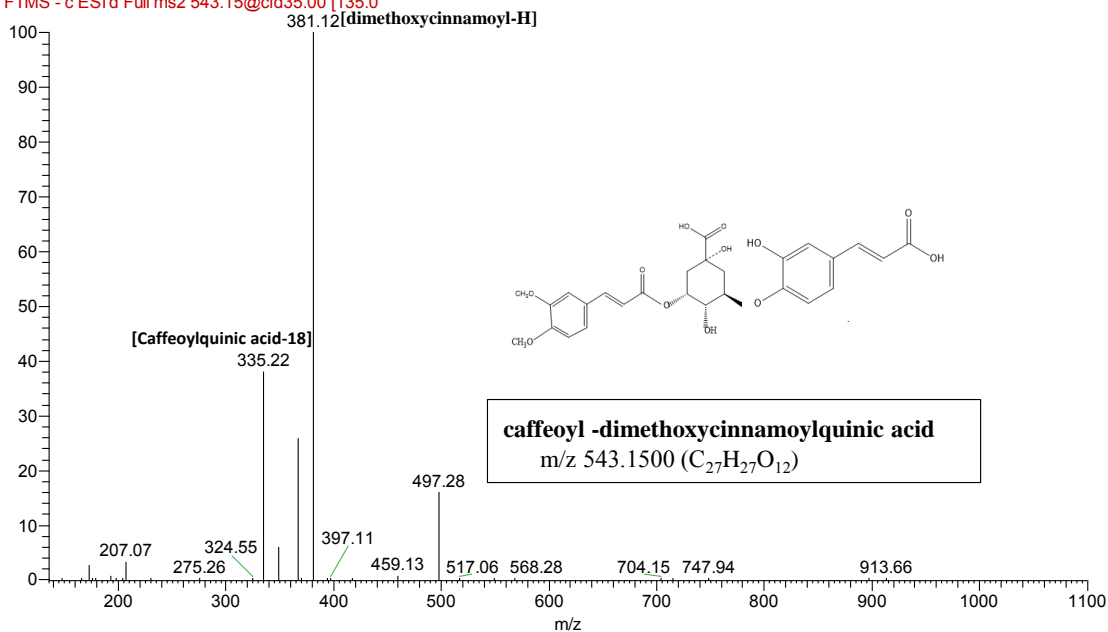
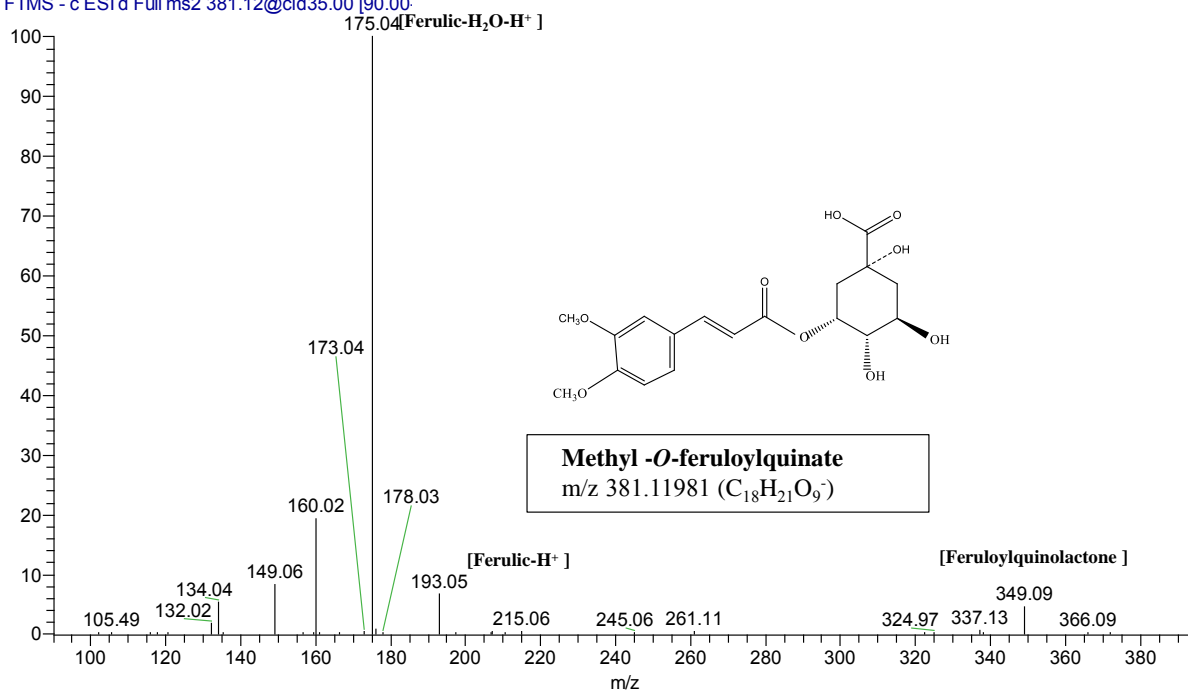


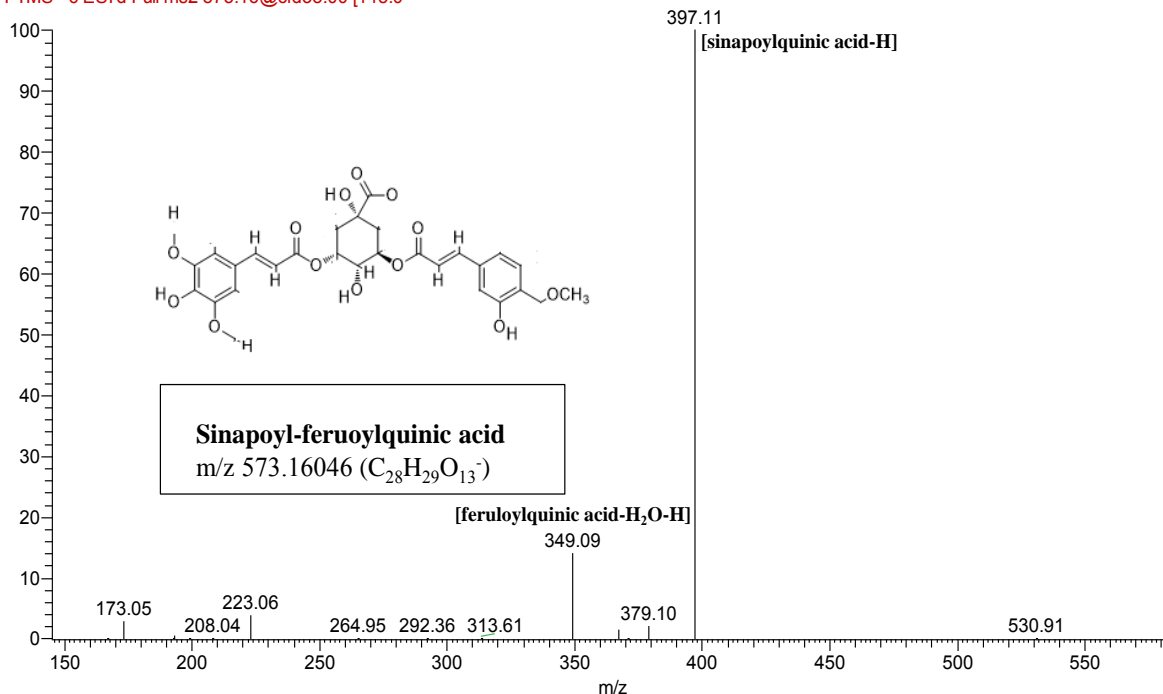
Figure 17: MS/MS spectrum of peak (P24) in negative ionization mode

FAM139\_RCCn #2738 RT: 10.02 AV: 1 NL: 3.11E4  
T: FTMS - c ESI d Full ms2 381.12@cid35.00 [90.00]



**Figure 18:** MS/MS spectrum of peak (p16) in negative ion mode

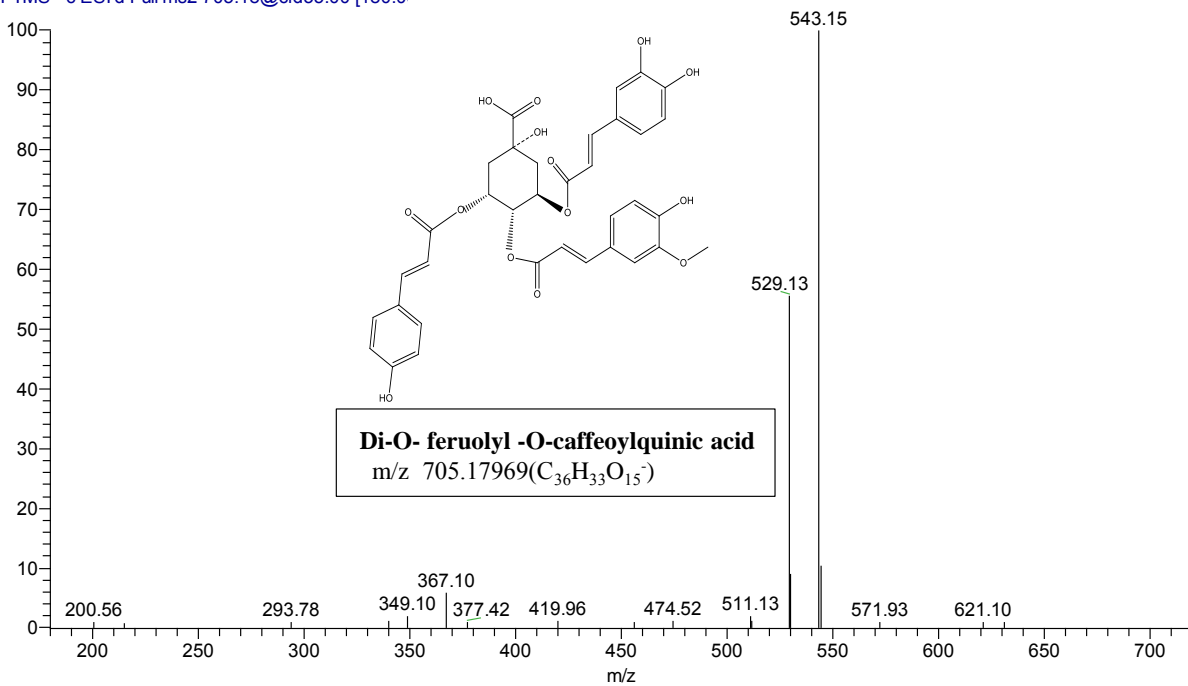
FAM139\_GCCn #3212 RT: 11.30 AV: 1 NL: 9.90E4  
F: FTMS - c ESI d Full ms2 573.16@cid35.00 [145.0]



**Figure 19:** MS/MS spectrum of peak (P22) in negative ion mode

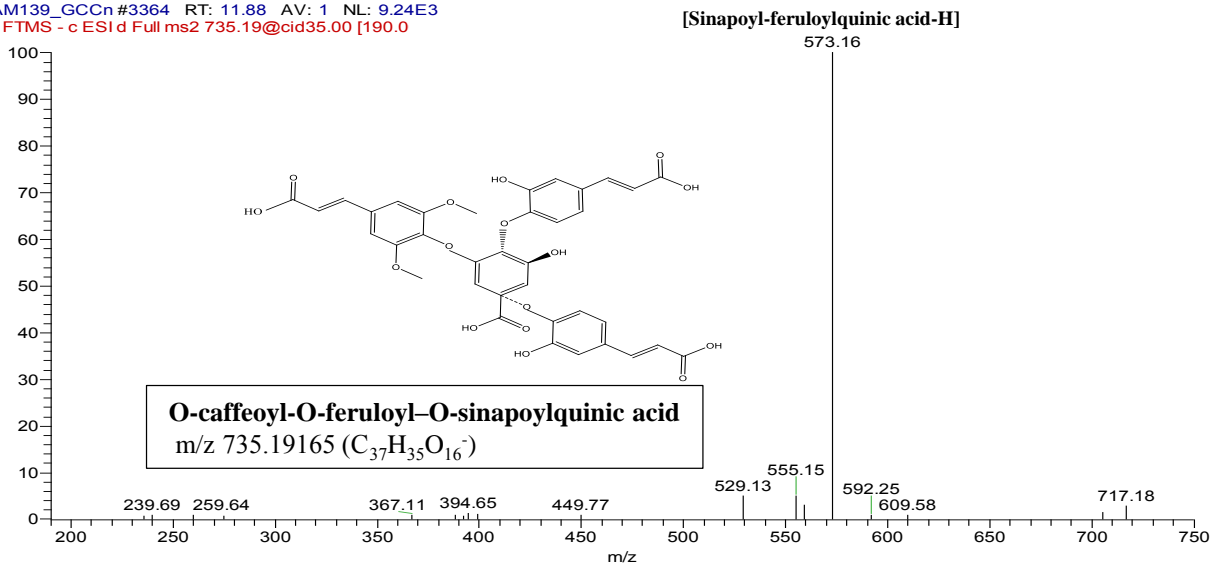


FAM139\_GCCn #3396 RT: 12.00 AV: 1 NL: 8.60E3  
T: FTMS - c ESI d Full ms2 705.18@cid35.00 [180.0]



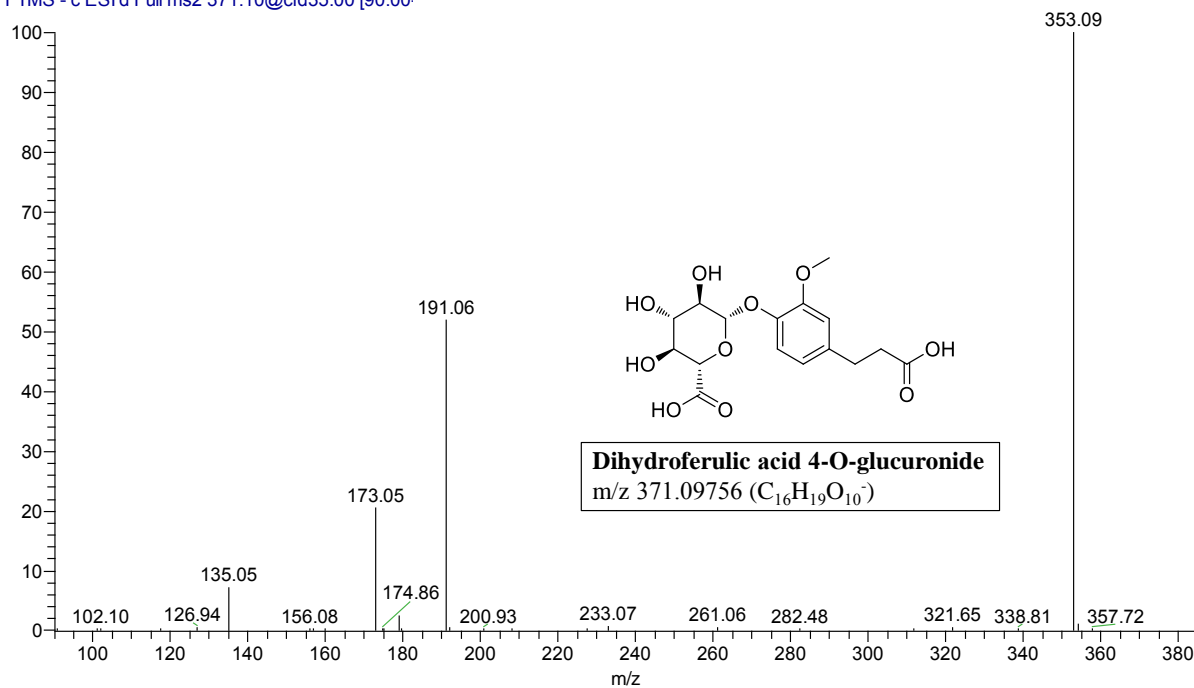
**Figure 20:** MS/MS spectrum of peak (P28) in negative ion mode.

FAM139\_GCCn #3364 RT: 11.88 AV: 1 NL: 9.24E3  
F: FTMS - c ESI d Full ms2 735.19@cid35.00 [190.0]



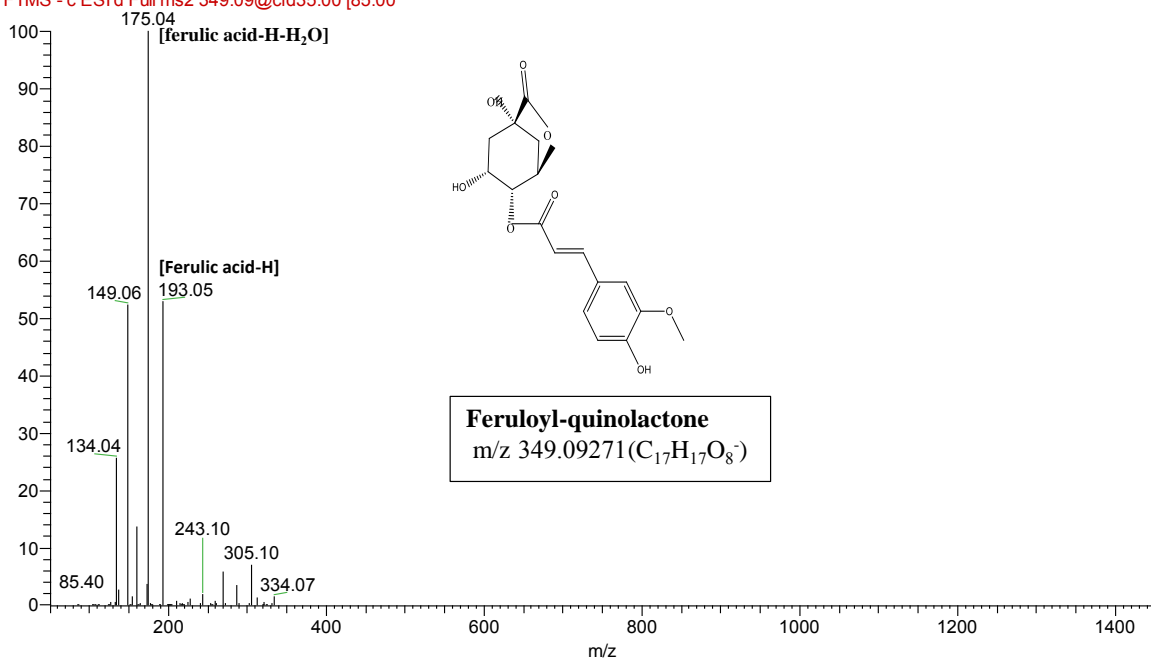
**Figure 21:** MS/MS spectrum of peak (P27) in negative ion mode

FAM139\_GCCn#248 RT: 0.94 AV: 1 NL: 2.97E4  
T: FTMS - c ESI d Full ms2 371.10@cid35.00 [90.00]



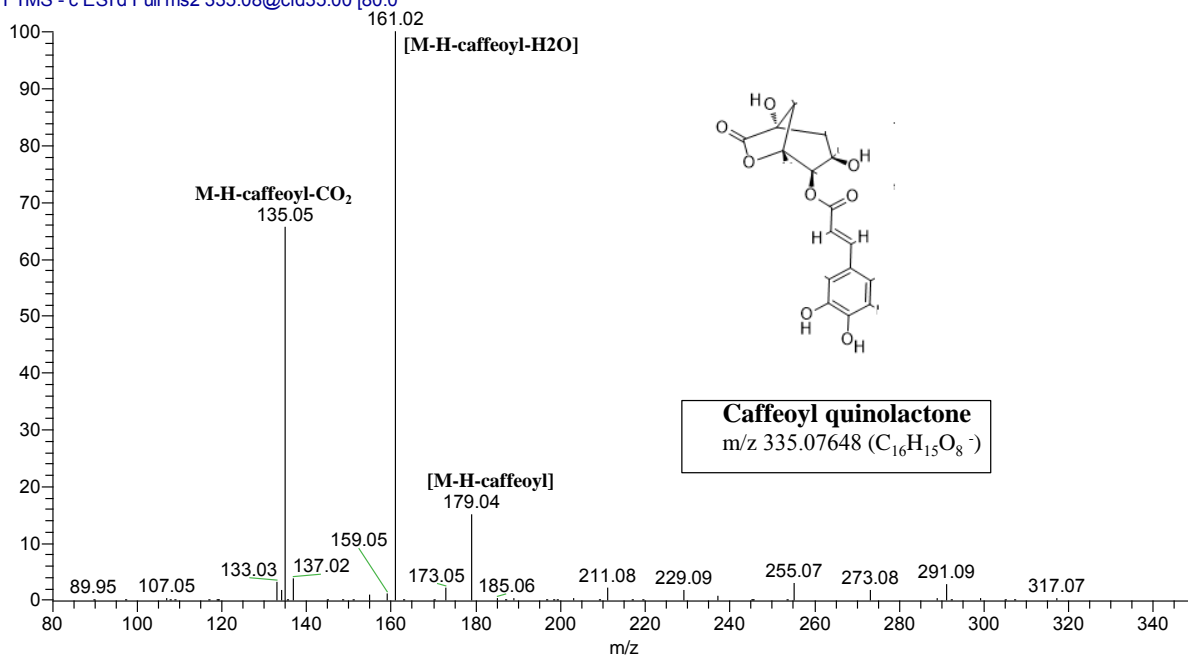
**Figure 22:** MS/MS spectrum of peak (P4) in negative ion mode

FAM139\_RCCn#2774 RT: 10.15 AV: 1 NL: 1.45E5  
F: FTMS - c ESI d Full ms2 349.09@cid35.00 [85.00]



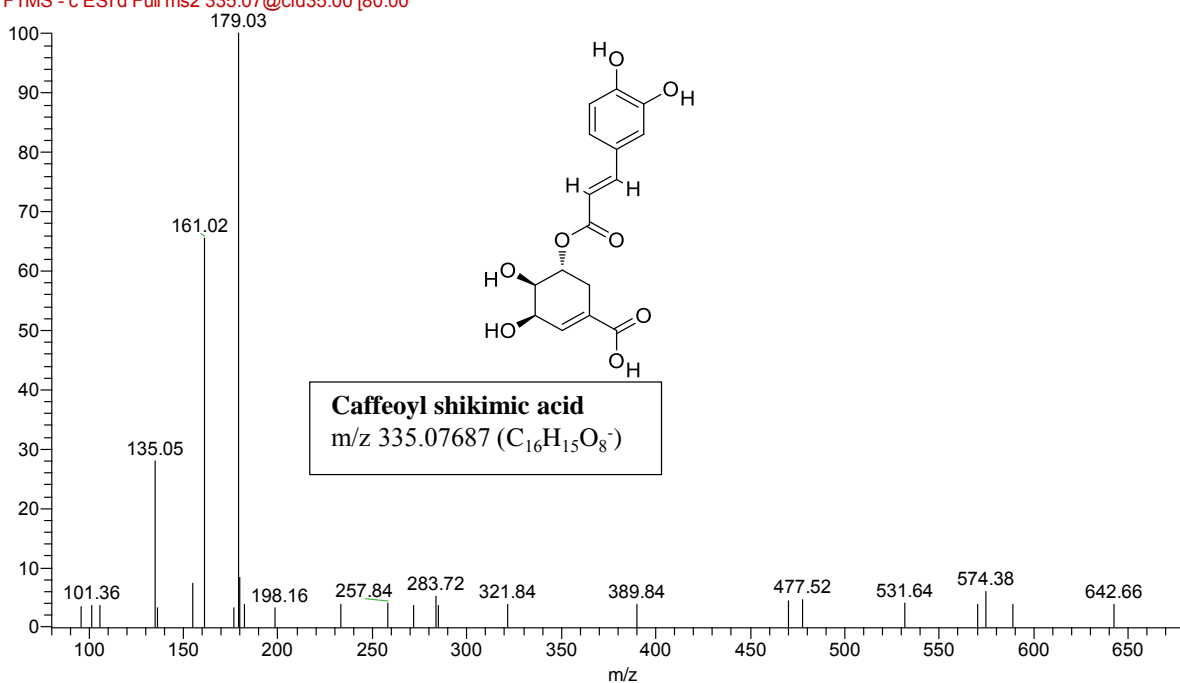
**Figure 23:** MS/MS spectrum of peak (P17) in negative ion mode

FAM139\_RCA#2440 RT: 8.97 AV: 1 NL: 1.34E5  
T: FTMS - c ESI d Full ms2 335.08@cid35.00 [80.0]



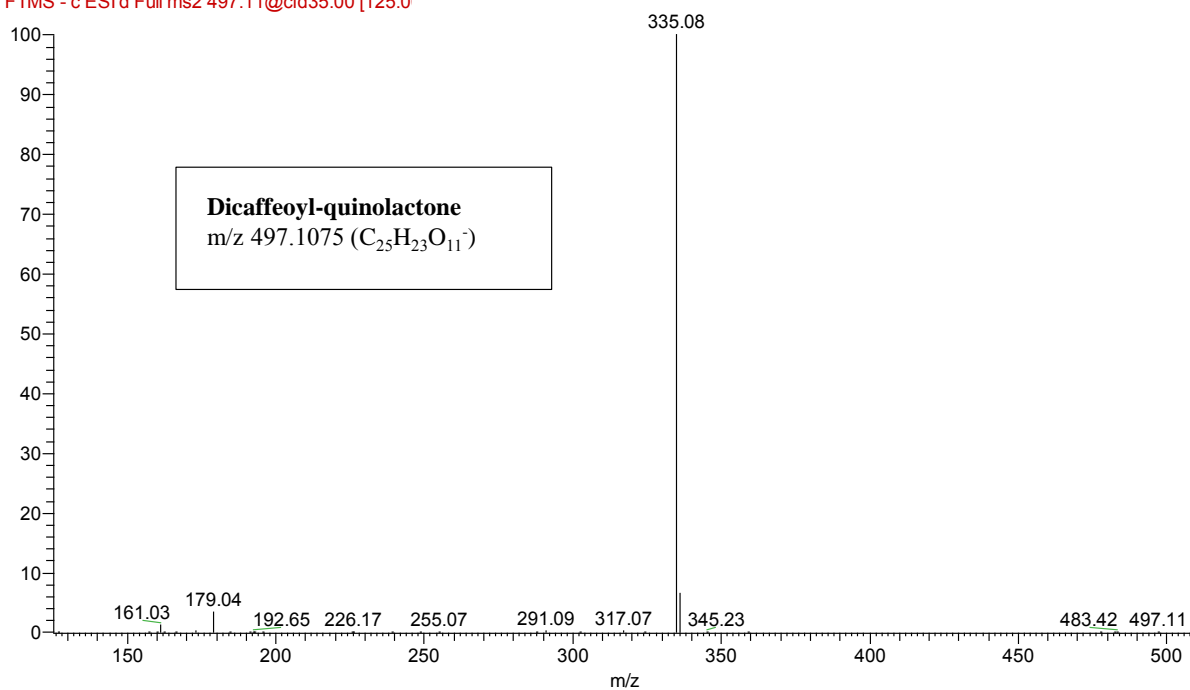
**Figure 24: MS/MS spectrum of peak (P13) in negative ion mode**

FAM139\_RCC#1720 RT: 6.27 AV: 1 NL: 2.08E3  
F: FTMS - c ESI d Full ms2 335.07@cid35.00 [80.00]



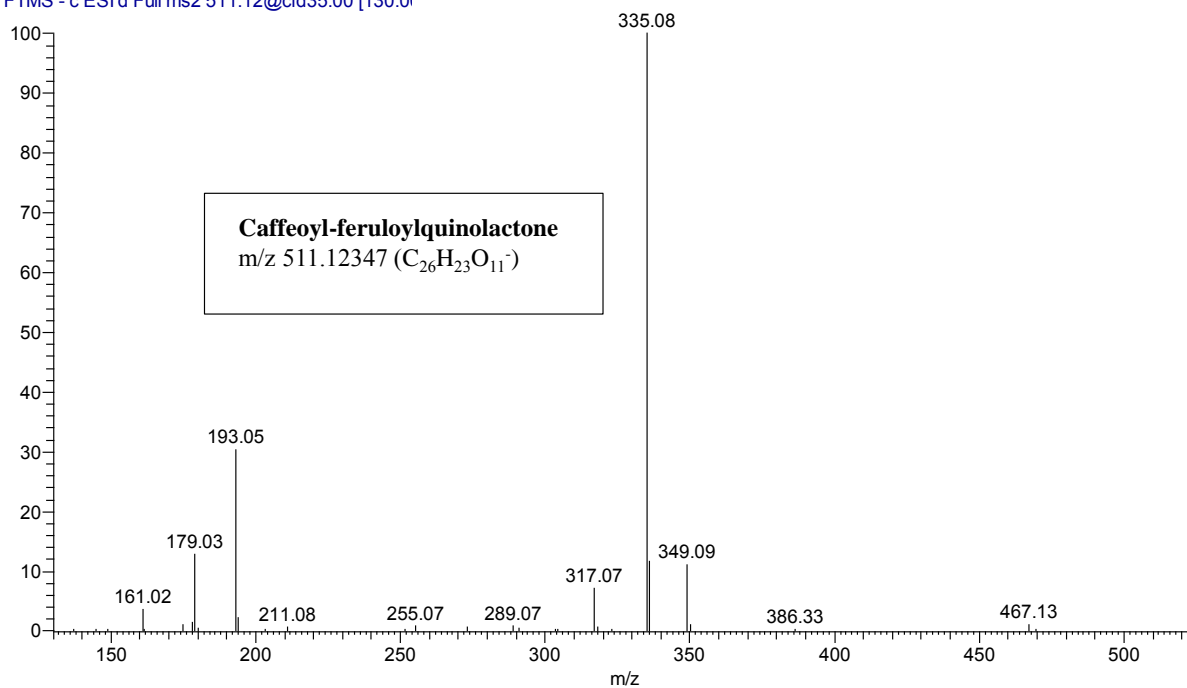
**Figure 25: MS/MS spectrum of peak (P10) in negative ion mode**

FAM139\_RCCn#3142 RT: 11.48 AV: 1 NL: 6.85E4  
F: FTMS - c ESI d Full ms2 497.11@cid35.00 [125.0]



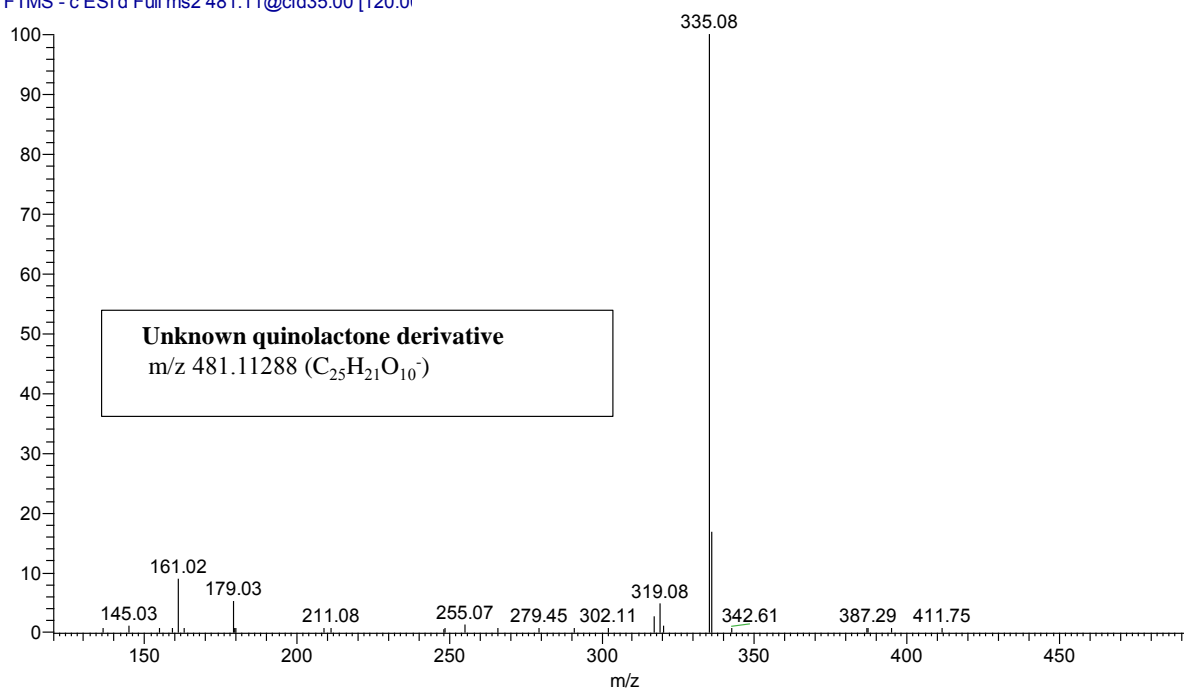
**Figure 26:** MS/MS spectrum of peak (P18) in negative ion mode

FAM139\_RCCn#3282 RT: 12.00 AV: 1 NL: 2.41E4  
T: FTMS - c ESI d Full ms2 511.12@cid35.00 [130.0]



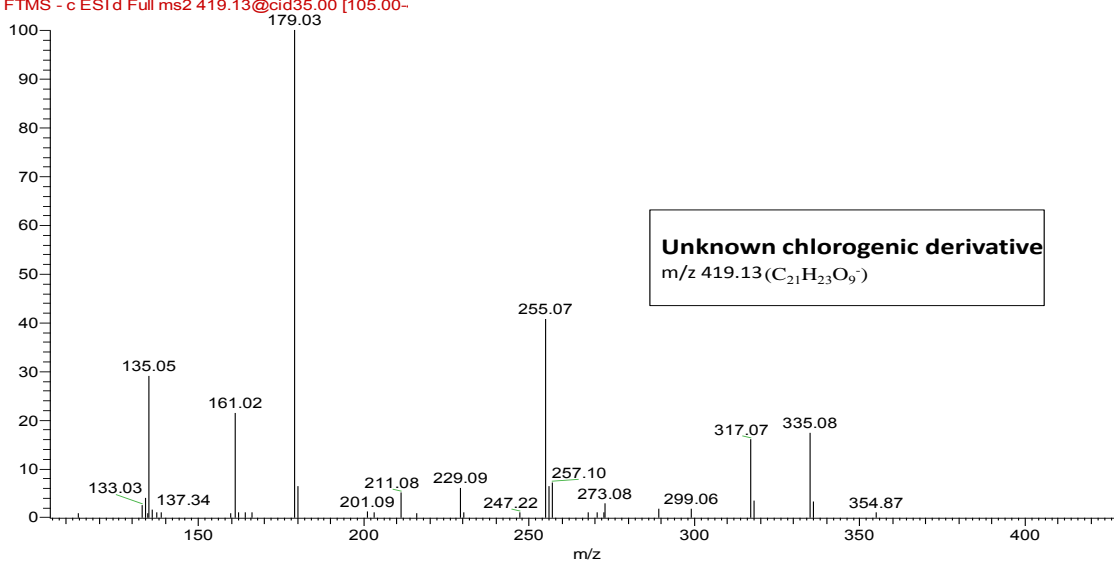
**Figure 27:** MS/MS spectrum of (P21) in negative ion mode

FAM139\_RCCn#3178 RT: 11.61 AV: 1 NL: 1.19E4  
T: FTMS - c ESI d Full ms2 481.11@cid35.00 [120.0]



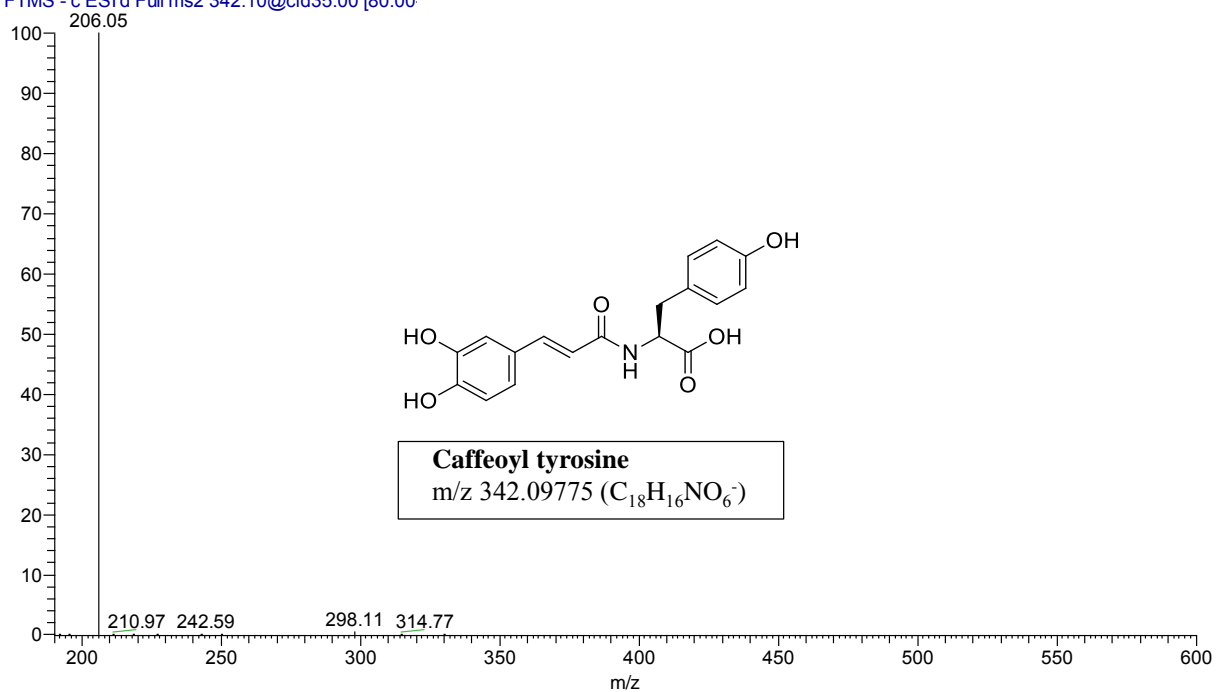
**Figure 28:** MS/MS spectrum of peak (P25) in negative ion mode

FAM139\_RCA#3314 RT: 12.19 AV: 1 NL: 8.21E3  
F: FTMS - c ESI d Full ms2 419.13@cid35.00 [105.00-]



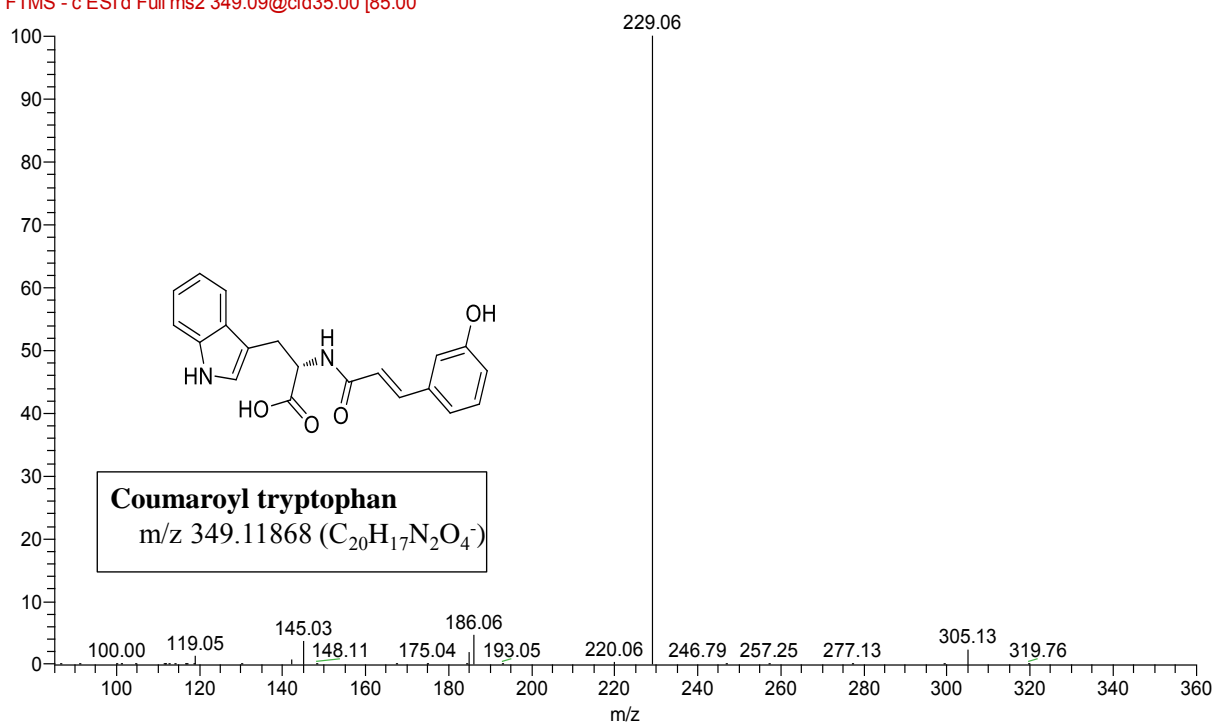
**Figure 29:** MS/MS spectrum of peak (P30) in negative ion mode

FAM139\_GCCn #2754 RT: 9.81 AV: 1 NL: 3.09E5  
T: FTMS - c ESI d Full ms2 342.10@cid35.00 [80.00]



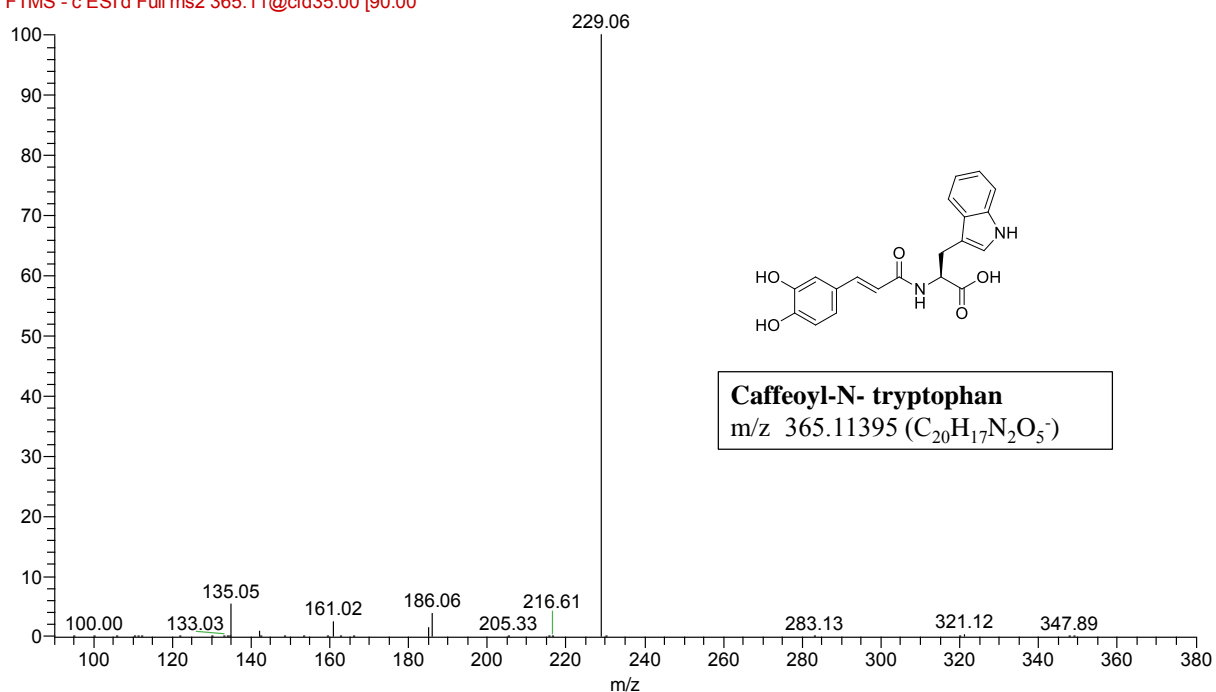
**Figure 30:** MS/MS spectrum of peak (P82) in negative ion mode

FAM139\_RCCn #3144 RT: 11.49 AV: 1 NL: 9.09E4  
F: FTMS - c ESI d Full ms2 349.09@cid35.00 [85.00]



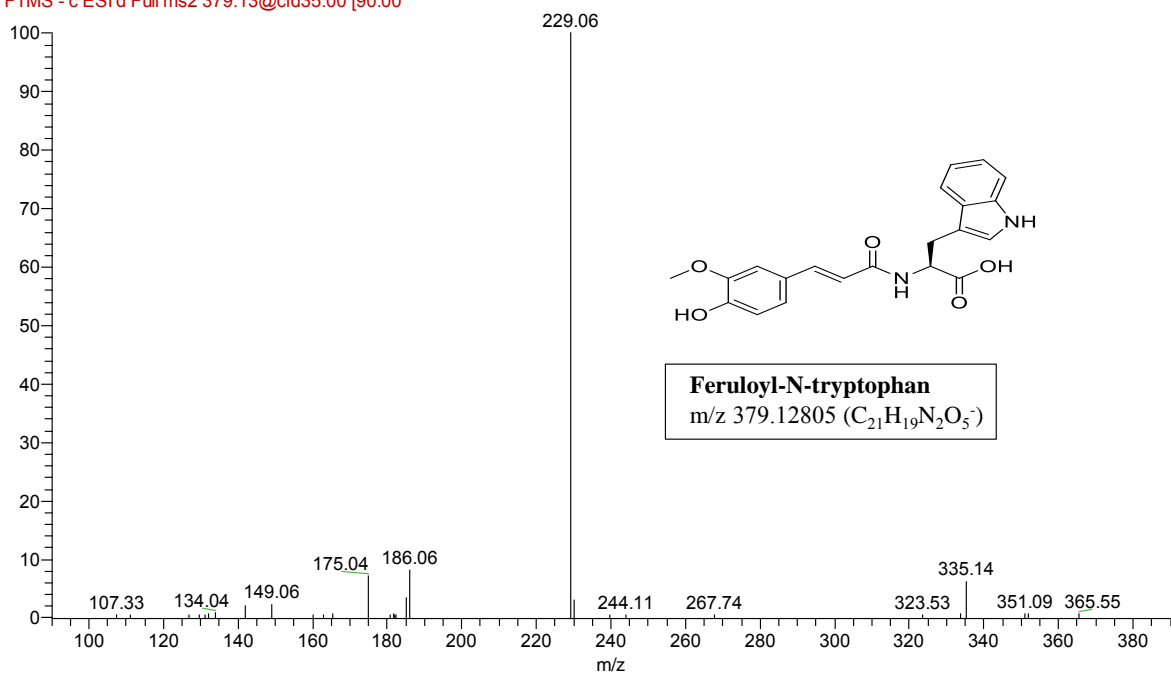
**Figure 31:** MS/MS spectrum of peak (P84) in negative ion mode

FAM139\_RCCn#3048 RT: 11.14 AV: 1 NL: 3.80E5  
 F: FTMS - c ESI d Full ms2 365.11@cid35.00 [90.00]

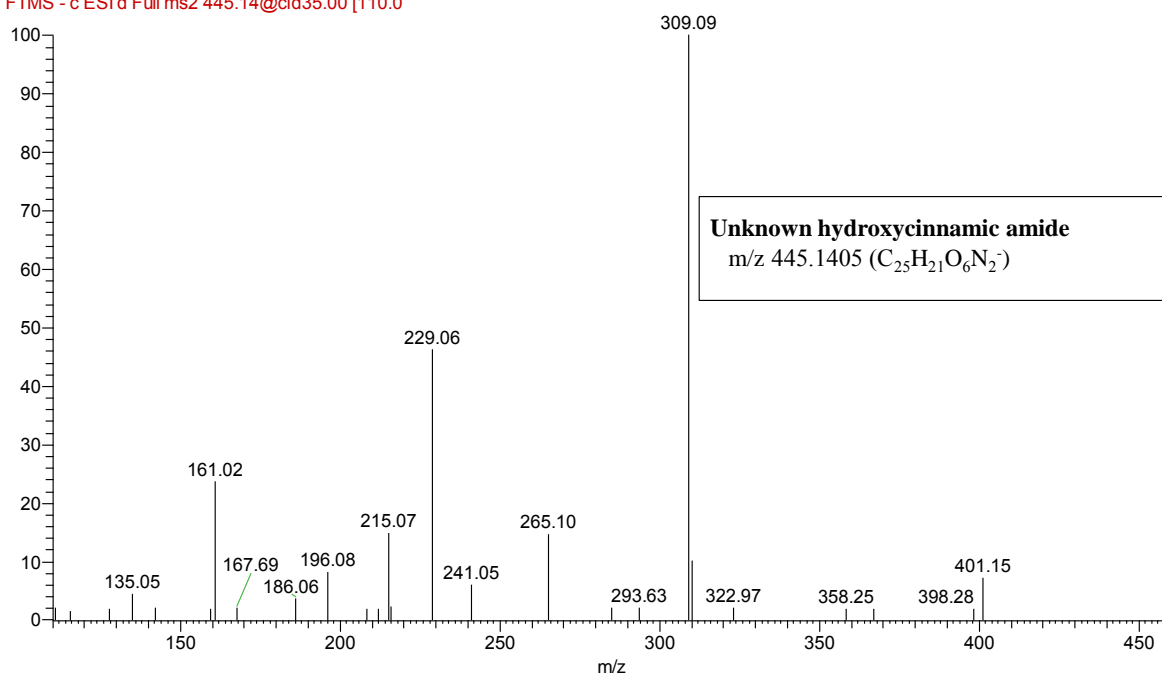


**Figure 32:** MS/MS spectrum of peak (P83) in negative ion mode

FAM139\_RCCn#3172 RT: 11.59 AV: 1 NL: 1.34E4  
 F: FTMS - c ESI d Full ms2 379.13@cid35.00 [90.00]



**Figure 33:** MS/MS spectrum of peak (P85) in negative ion mode



**Figure 34:** MS/MS spectrum of peak (P86) in negative ion mode

### 1.3.4 Diterpenes

The major diterpenes in coffee are cafestol and kahweol, which have been reported in both roasted arabica and canephora species [39]. In the current study, cafestol (P35) was detected in the positive ionization mode at  $m/z$  317, and its dehydrated analogue (P37) assigned as dehydrocafestol  $m/z$  299, **Figure 35** & **Figure 37**. Additionally, new diterpenes reported for the first time in coffee derived from dehydrocafestol (P45), producing a peak at  $m/z$  281 with fragmentation ion at  $m/z$  263 ( $M+H-H_2O$ ), 147 ( $M+H-C_{10}H_{18}O_2$ ), and 131 confirming the identity of cafestol entity, **Figure 37**. Although kahweol was not detected in our roasted canephora samples, 16-O-methylkahweol (P43) was identified at  $m/z$  331 with a base peak at  $m/z$  314 ( $M-H_2O$ ) and  $m/z$  296 corresponding to kahweol and dehydrokahweol, respectively (**Table 6** & **Figure 38**), [39].

Mozambioside, a diterpenoid glycoside of furokaurane identified as a marker for arabica species [98], was detected in both green and roasted arabica (P38) at  $m/z$  ( $M+H$ ) 509 with fragment ions at  $m/z$  347 ( $M\text{-hexose}+H$ ) and 329 ( $M\text{-hexose-H}_2O+H$ ) with a characteristic UV band at  $\lambda_{max}$  298 nm [55, 99]. Another peak at  $m/z$  491 indicates loss of water molecule from



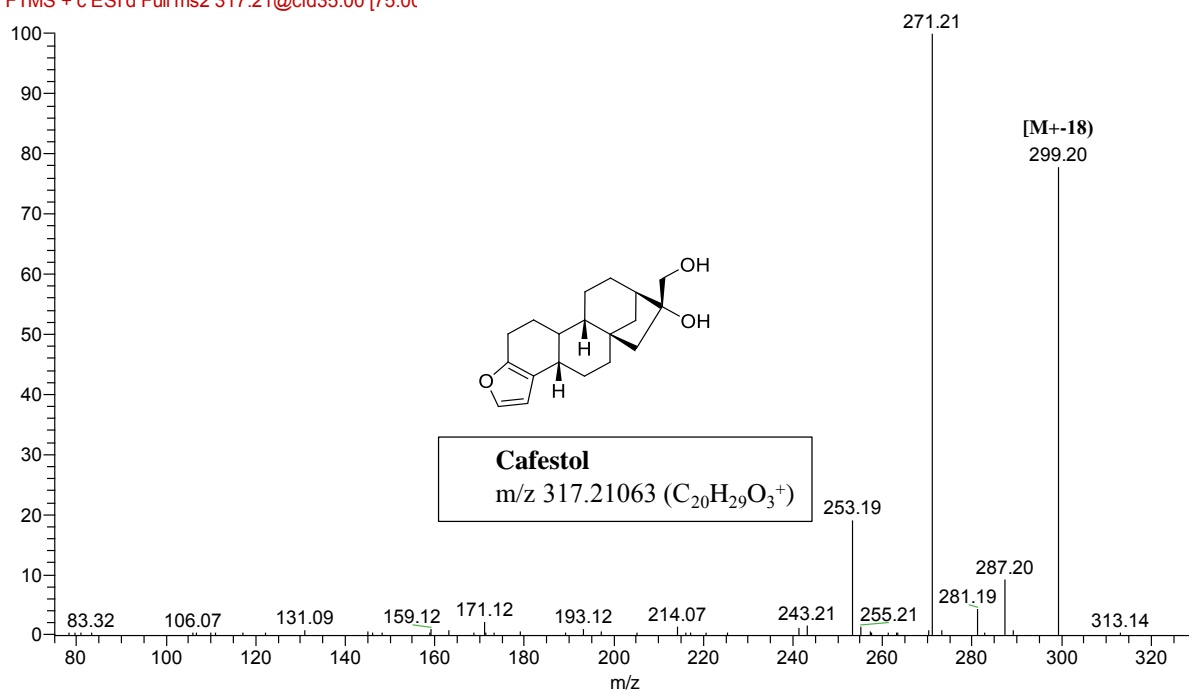
mozambioside and yields fragment ions at  $m/z$  329 (bengalensol) and  $m/z$  311(M-hexose) and annotated as bengalensol-*O*-hexoside (P39). Furthermore, bengalensol was detected at  $m/z$  329 (P41) with fragment ions at  $m/z$  311 (M-H<sub>2</sub>O) and  $m/z$  293 (M-2H<sub>2</sub>O) identified only in roasted arabica, **Figure 39, 40 and 41**.

Interestingly,  $m/z$  347 (M+H)<sup>+</sup> corresponded to a new diterpenoid annotated as trihydroxy-kauradien-olide (P36) based on fragment ions at  $m/z$  329 (M-18) and  $m/z$  285 (M-44) [100], posed as a new marker for green arabica species (GCA), found completely absent in RCA suggesting its degradation upon roasting, it was previously identified in *Isodon* species, **Figure 42**, [100].

Aside from diterpene furokaurane, Atractyligenins that are ent-kaurane diterpenoids detected to a lesser extent than diterpenes furokaurane. They lack a characteristic UV band at  $\lambda_{\max}$  298 and found only in the GCA, suggesting that they were degraded during the thermal process [44]. Interestingly, they are structurally related to atractyloside phytotoxin and inhibit adenine nucleotide in mitochondria exhibiting toxic effects in raw coffee[98]. An intense peak at  $m/z$  481 was detected in negative mode with a base peak at  $m/z$  301 indicating the loss of hexose and annotated as atracyligenin-*O*-hexoside (P47), **Figure 43**, [98]. Another analogue of atracyligenin was observed in P49 assigned as desoxy-atracyligenin-*O*-hexoside, **Table 6**,.

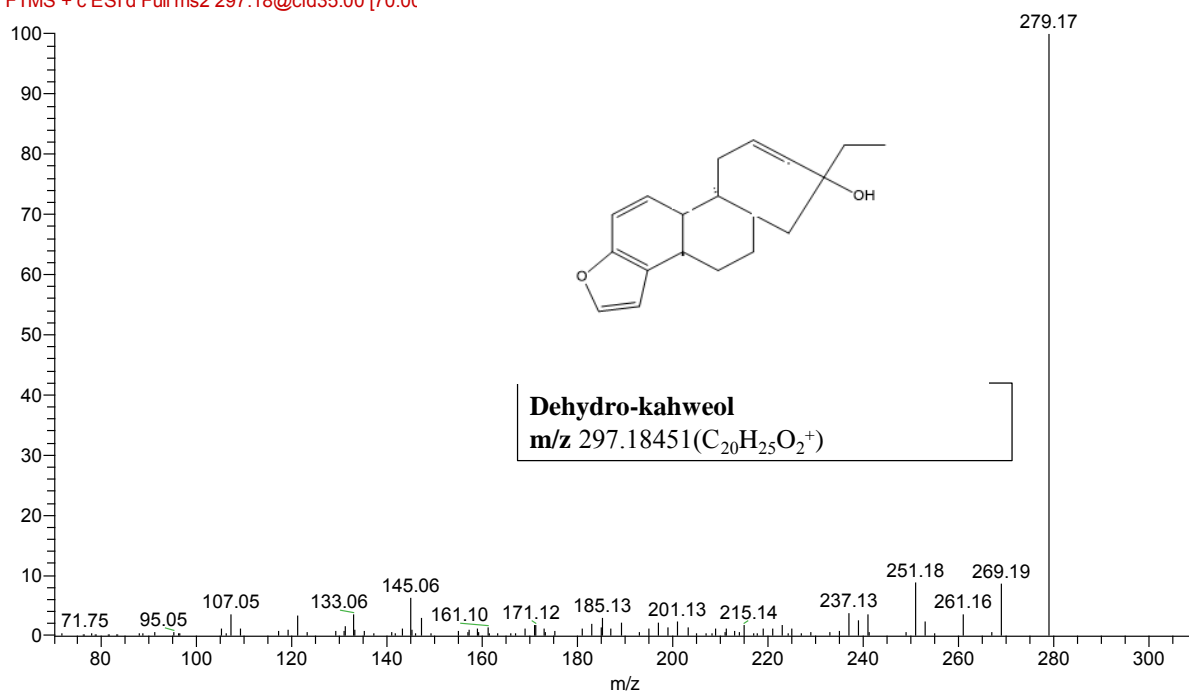
Besides, a peak at  $m/z$  771 with a base peak at  $m/z$  727 attributed to the loss of 44 Da confirming its assigning desoxycarboxyatracyligenin-*O*-hexoside (P48), **Figure 45**. Furthermore, an intense peak at  $m/z$  609 and its fragment ion at  $m/z$  565 (loss of 44 Da) suggested the presence of carboxyatracyligenin-*O*-hexoside in (P50) **Figure 46** [44,98].

FAM139\_GCA #1880 RT: 9.39 AV: 1 NL: 1.00E5  
 F: FTMS + c ESI d Full ms2 317.21@cid35.00 [75.0]



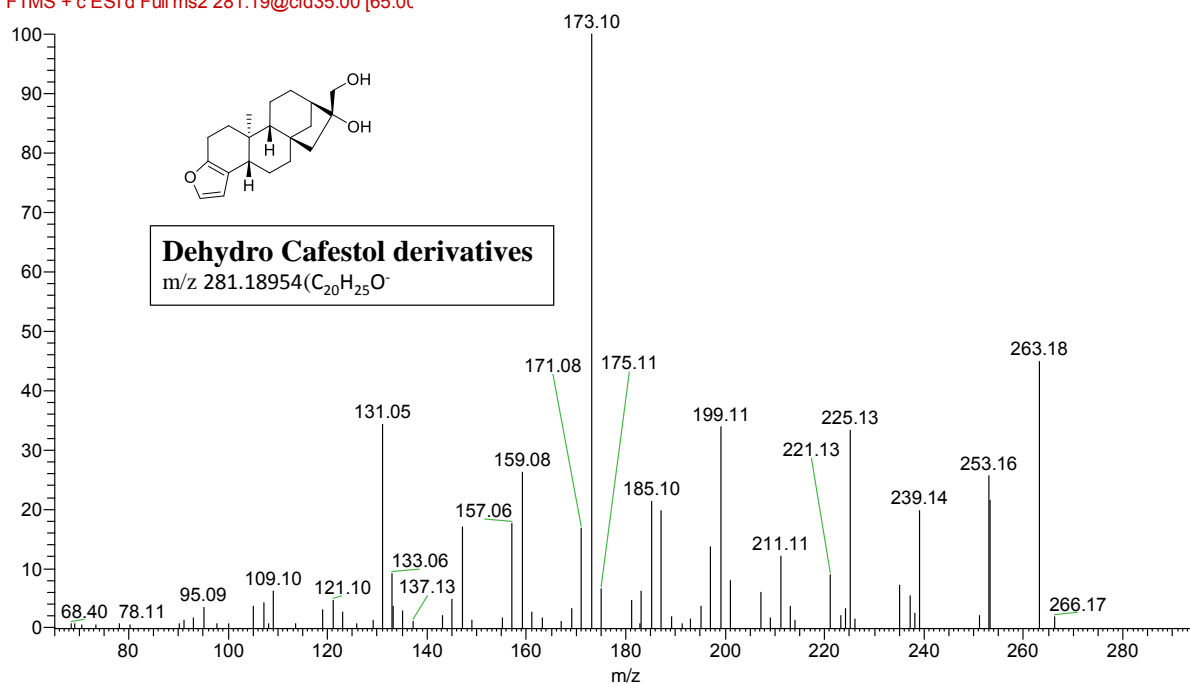
**Figure 35:** MS/MS spectrum of peak (P35) in positive ion mode

FAM139\_GCA #2766 RT: 13.60 AV: 1 NL: 1.24E5  
 F: FTMS + c ESI d Full ms2 297.18@cid35.00 [70.0]



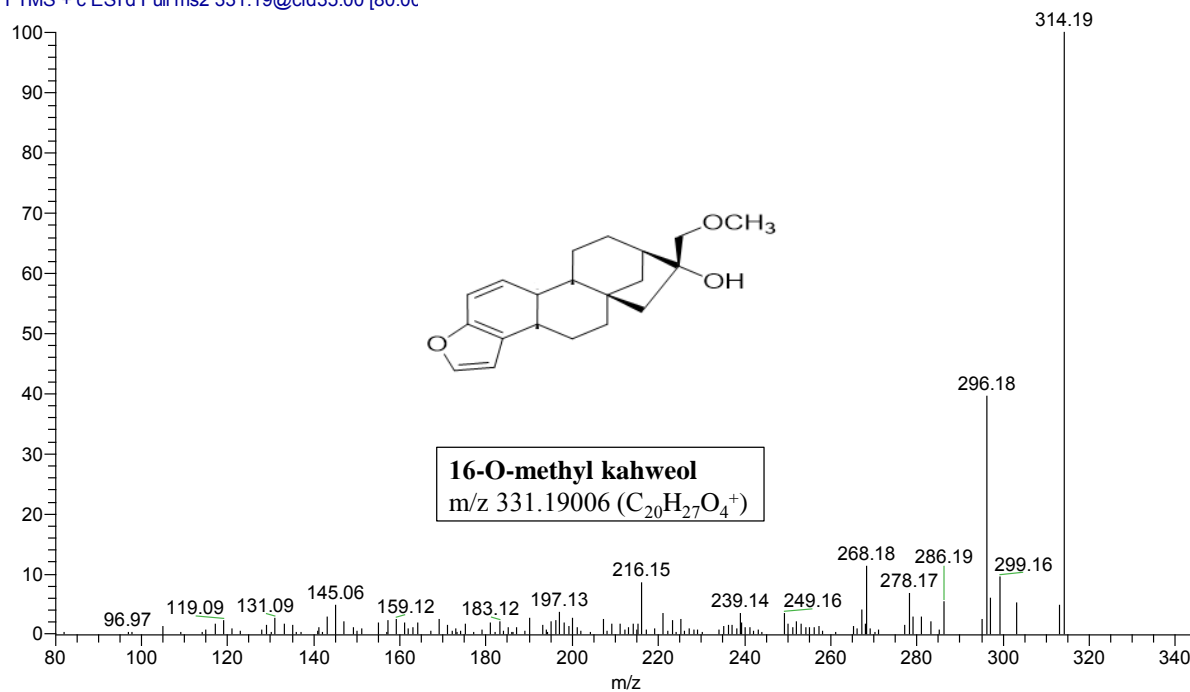
**Figure 36:** MS/MS spectrum of peak (P37) in positive ion mode

FAM139\_RCA #3216 RT: 15.31 AV: 1 NL: 4.14E4  
 F: FTMS + c ESI d Full ms2 281.19@cid35.00 [65.0C]



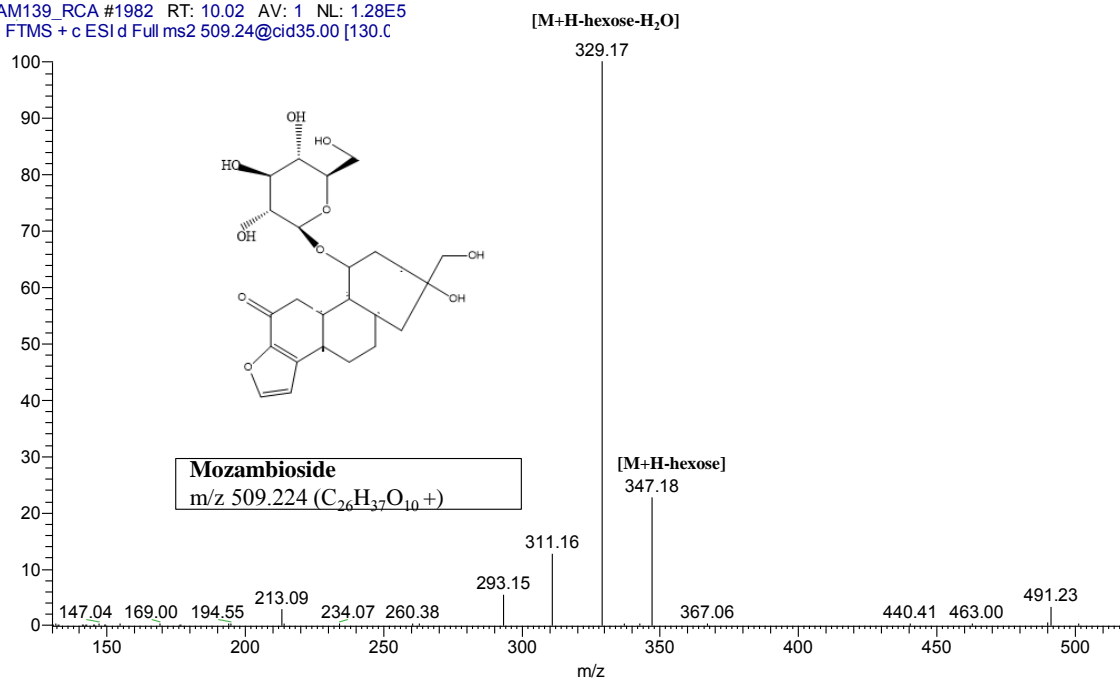
**Figure 37:** MS/MS spectrum of peak (P45) in positive ion mode.

FAM139\_RCC #2684 RT: 12.78 AV: 1 NL: 9.93E4  
 T: FTMS + c ESI d Full ms2 331.19@cid35.00 [80.0C]



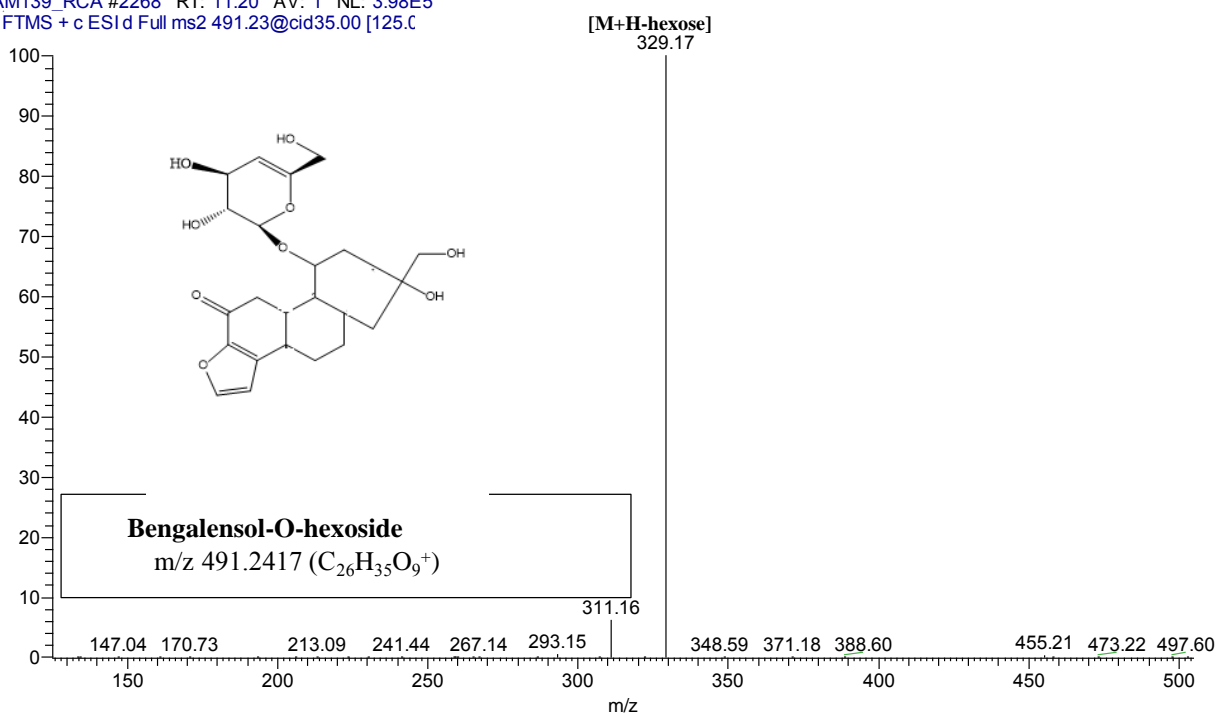
**Figure 38:** MS/MS spectrum of peak (P43) in positive ion mode

FAM139\_RCA #1982 RT: 10.02 AV: 1 NL: 1.28E5  
T: FTMS + c ESI d Full ms2 509.24@cid35.00 [130.0]



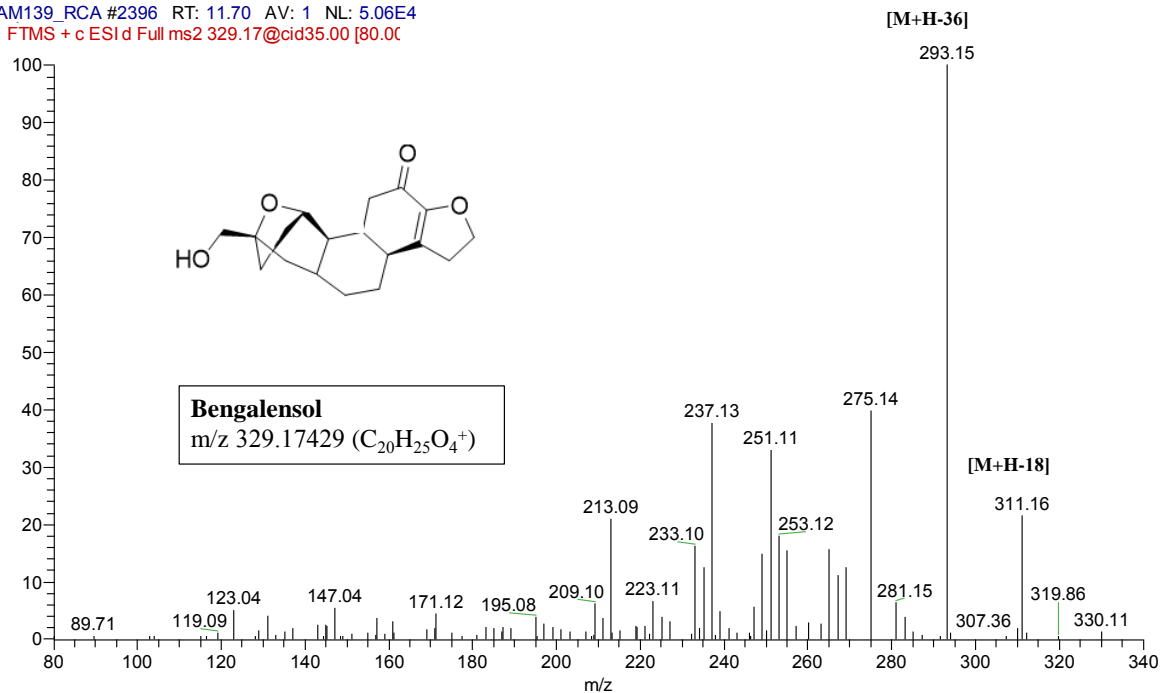
**Figure 39:** MS/MS spectrum of peak (P38) in positive ion mode

FAM139\_RCA #2268 RT: 11.20 AV: 1 NL: 3.98E5  
T: FTMS + c ESI d Full ms2 491.23@cid35.00 [125.0]



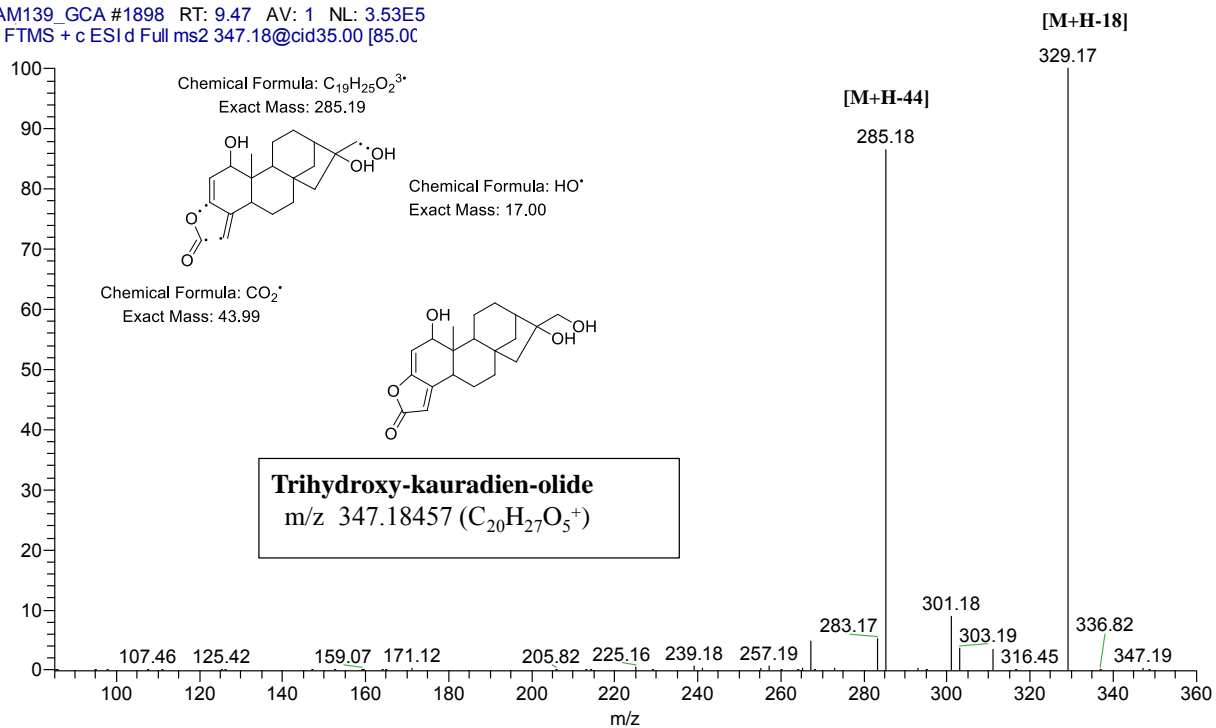
**Figure 40:** MS/MS spectrum of peak (P39) in positive ion mode

FAM139\_RCA #2396 RT: 11.70 AV: 1 NL: 5.06E4  
F: FTMS + c ESI d Full ms2 329.17@cid35.00 [80.0]



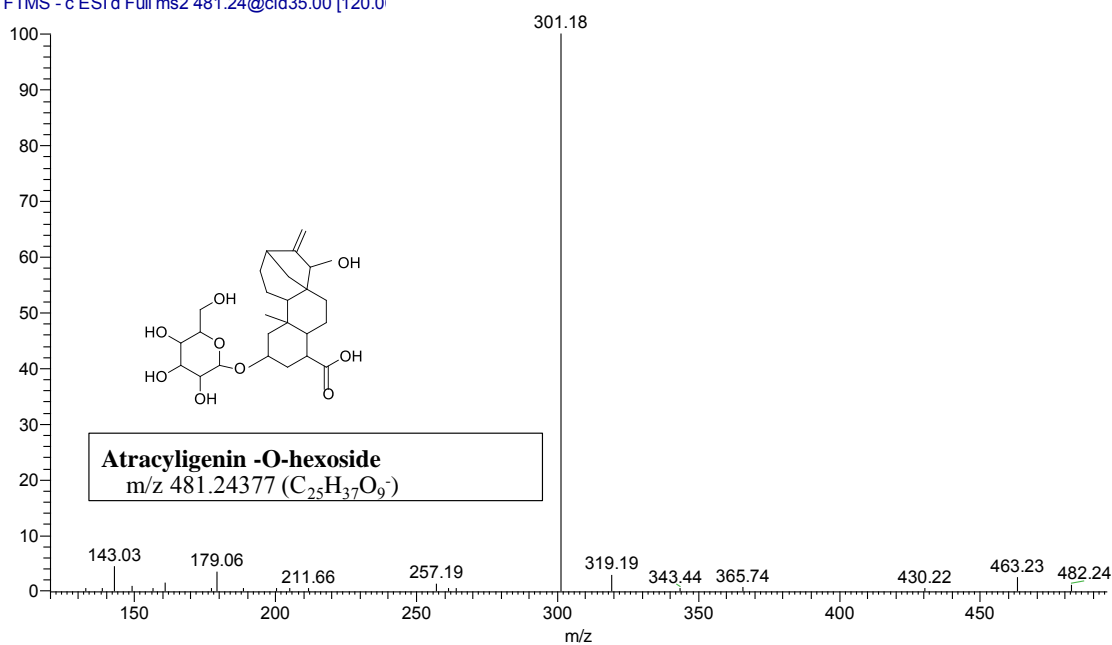
**Figure 41:** MS/MS spectrum of peak (P41) in positive ion mode

FAM139\_GCA #1898 RT: 9.47 AV: 1 NL: 3.53E5  
T: FTMS + c ESI d Full ms2 347.18@cid35.00 [85.0]



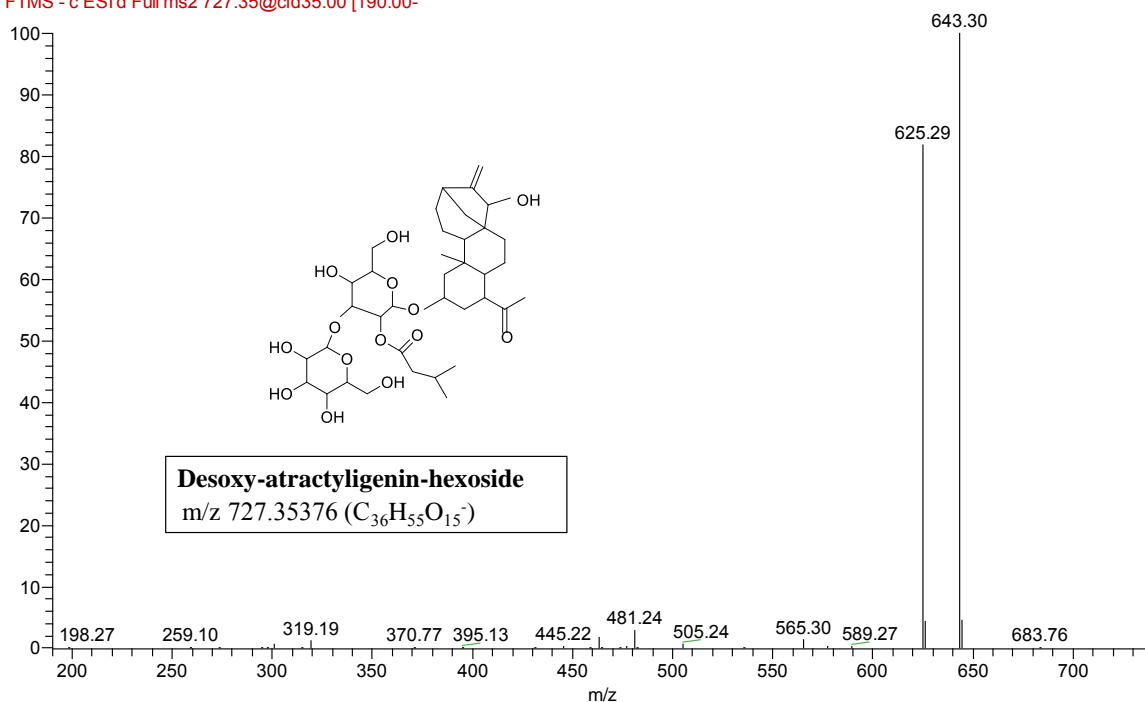
**Figure 42:** MS/MS spectrum of peak (P36) in positive ion mode

FAM139\_GCA#2630 RT: 9.53 AV: 1 NL: 4.01E4  
T: FTMS - c ESI d Full ms2 481.24@cid35.00 [120.0]



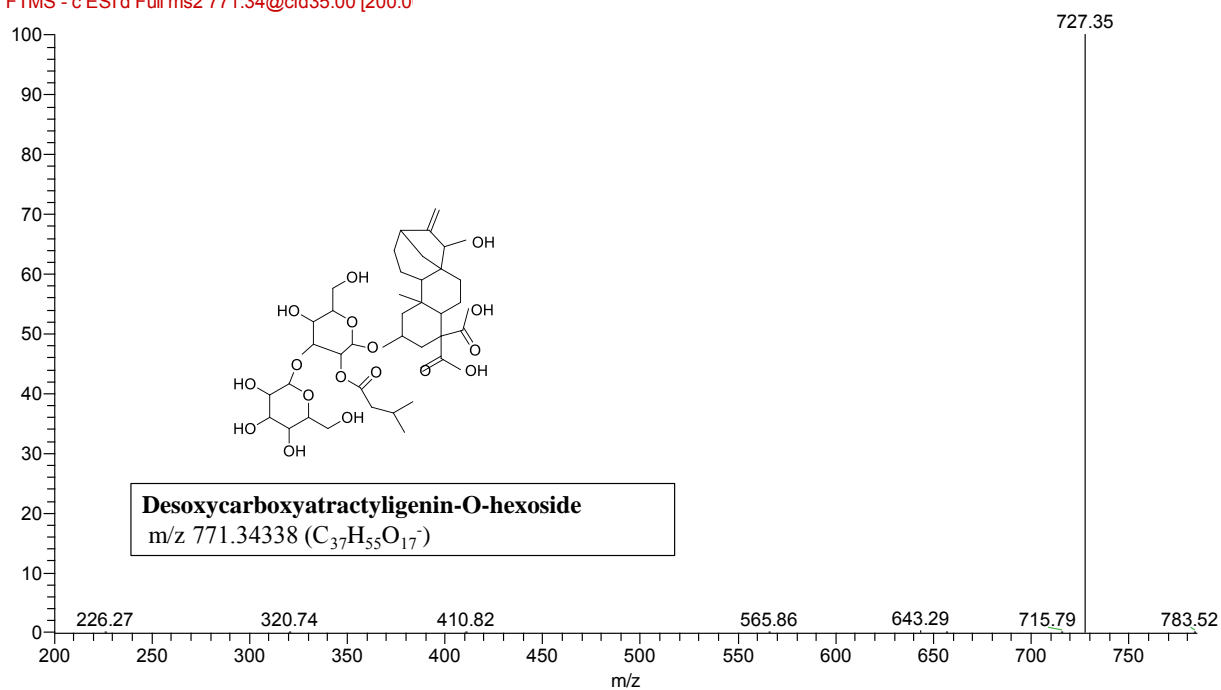
**Figure 43:** MS/MS spectrum of peak (P47) in negative ion mode

FAM139\_RCA#3056 RT: 11.21 AV: 1 NL: 1.23E5  
F: FTMS - c ESI d Full ms2 727.35@cid35.00 [190.00-]



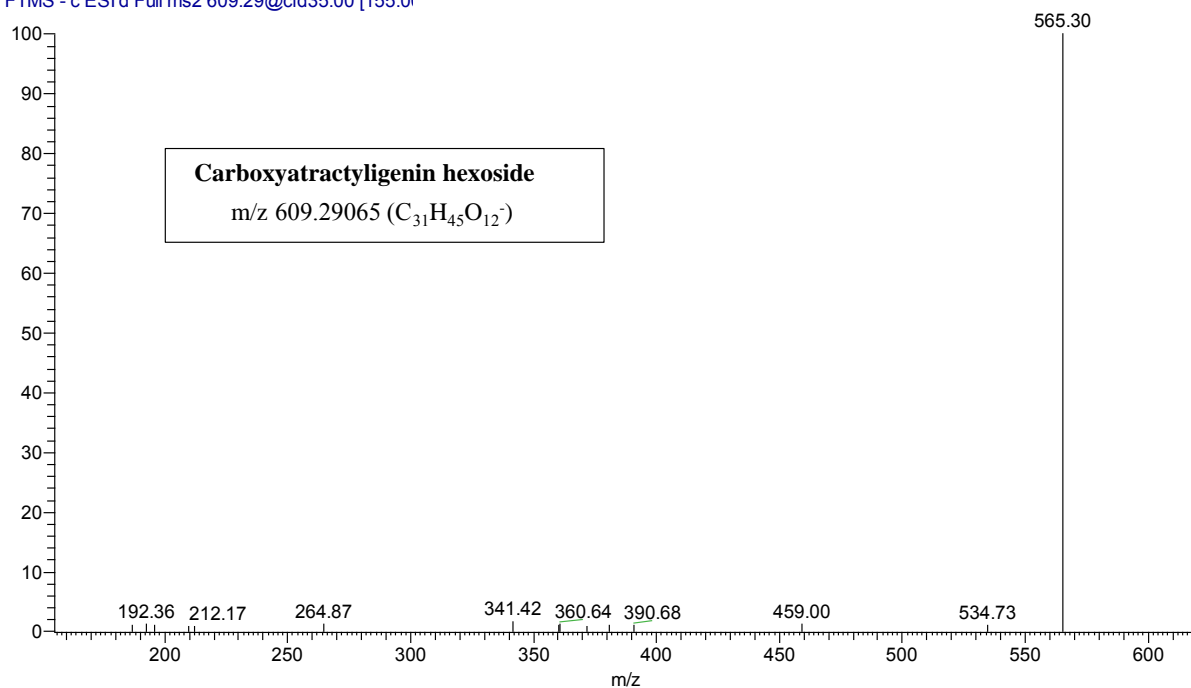
**Figure 44:** MS/MS spectrum of peak (P49) in negative ion mode

FAM139\_GcAn#2972 RT: 10.72 AV: 1 NL: 8.78E4  
 F: FTMS - c ESI d Full ms2 771.34@cid35.00 [200.0]

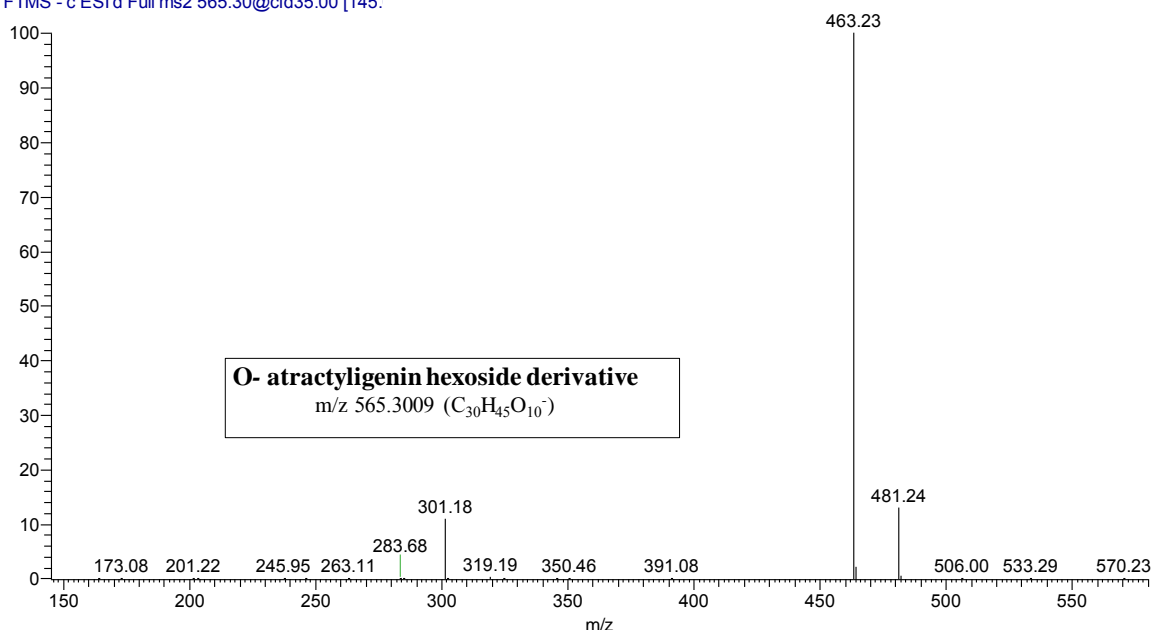


**Figure 45:** MS/MS spectrum of peak (P48) in negative ion mode

FAM139\_GcAn#3244 RT: 11.80 AV: 1 NL: 7.23E3  
 T: FTMS - c ESI d Full ms2 609.29@cid35.00 [155.0]



**Figure 46:** MS/MS spectrum of peak (P50) in negative ion mode



**Figure 47:** MS/MS spectrum of peak (P51) in negative ion mode

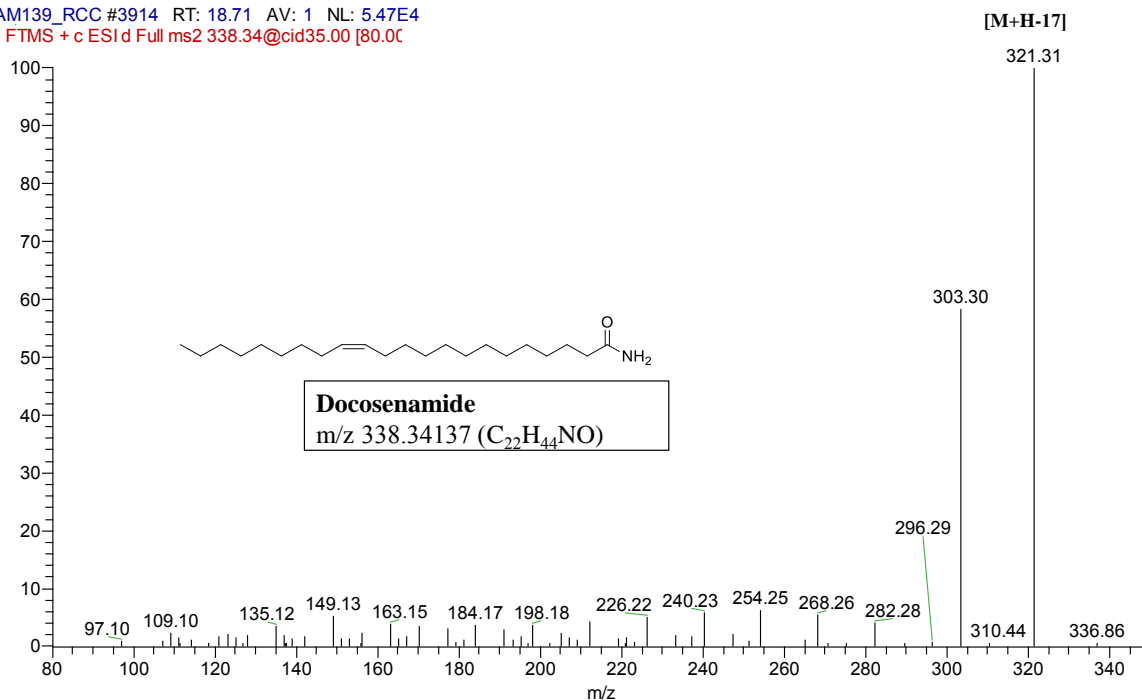
### 1.3.5 Fatty acids

Fatty acids and sphingolipids appeared lately at  $R_t > 12.00$  min of the UPLC chromatogram, **Table 6**. Negative ionization mode revealed several hydroxylated fatty acid peaks, For example, (P71) at  $m/z$  341 ( $M-H$ )<sup>-</sup> assigned as dimethyl octadecanedioate with fragment ions at  $m/z$  313( $M-2CH_3$ ) and  $m/z$  269 (hexadecanoic acid- $H$ ), **Figure 48**. Likewise, other hydroxylated fatty acids were detected in peaks (P 52), (P69), and (P72) that showed ( $M-H$ )<sup>-</sup> at  $m/z$  329,  $m/z$  355, and  $m/z$  483 respectively, and annotated with regard to literature. The successive loss of  $CH_2$  (14 Da) had confirmed their identity as fatty acid and was identified in both species, **Figure 49 & Figure 50, Table 5** [18,101].

Nitrogen-containing lipids were detected in (P75) at  $m/z$  338 ( $M+H$ ) annotated as docosenamide reported for the first time in coffee [104]. It showed MS/MS fragments at  $m/z$  321( $M-17$ ) corresponding to a loss of ammonia detected in both arabica and canephora species **Figure 51**.



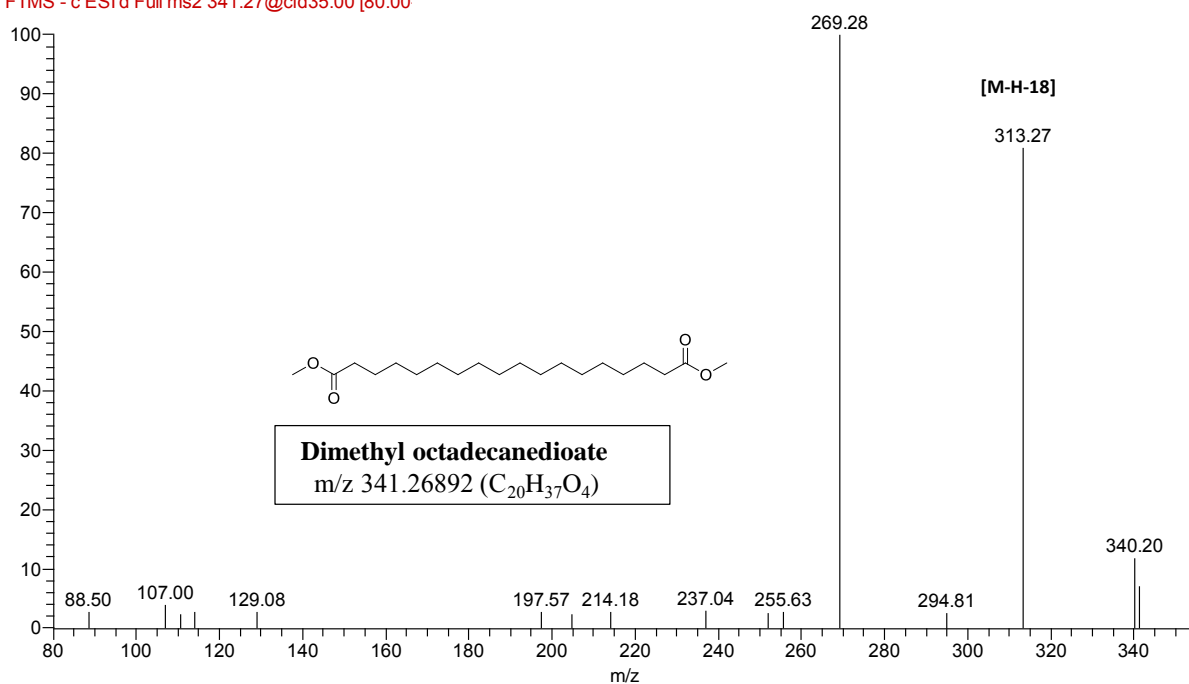
FAM139\_RCC #3914 RT: 18.71 AV: 1 NL: 5.47E4  
 F: FTMS + c ESI d Full ms2 338.34@cid35.00 [80.0C]



Sphingolipids are complex molecules of lipids and were abundant in all the samples (GCA, GCC, RCA, RCC), and several peaks were detected in the negative and positive ionization modes. In negative ion mode, four peaks with an even mass weight at  $m/z$  540, 564, 566, and 568 eluted and showed a loss of 60 AMU corresponding to trimethylammonium ( $C_3H_{10}N^+$ ) characteristic for sphingolipids annotated as sphingolipids **I**, **II**, **III**, **IV** (P56, P57, P60, P62), respectively as in **Figure 52-55**. The molecular network created from GNPS connected these conjugates confirming their structural similarity (cluster **E**), **Figure 7**. Another sphingolipid was observed in the negative mode at  $m/z$  571 with major fragments ions at  $m/z$  391 (M-hexose), 315 (M-palmitic acid) annotated as phosphatidylinositol hexanoic acid derivative (P58), **Figure 55** [18].

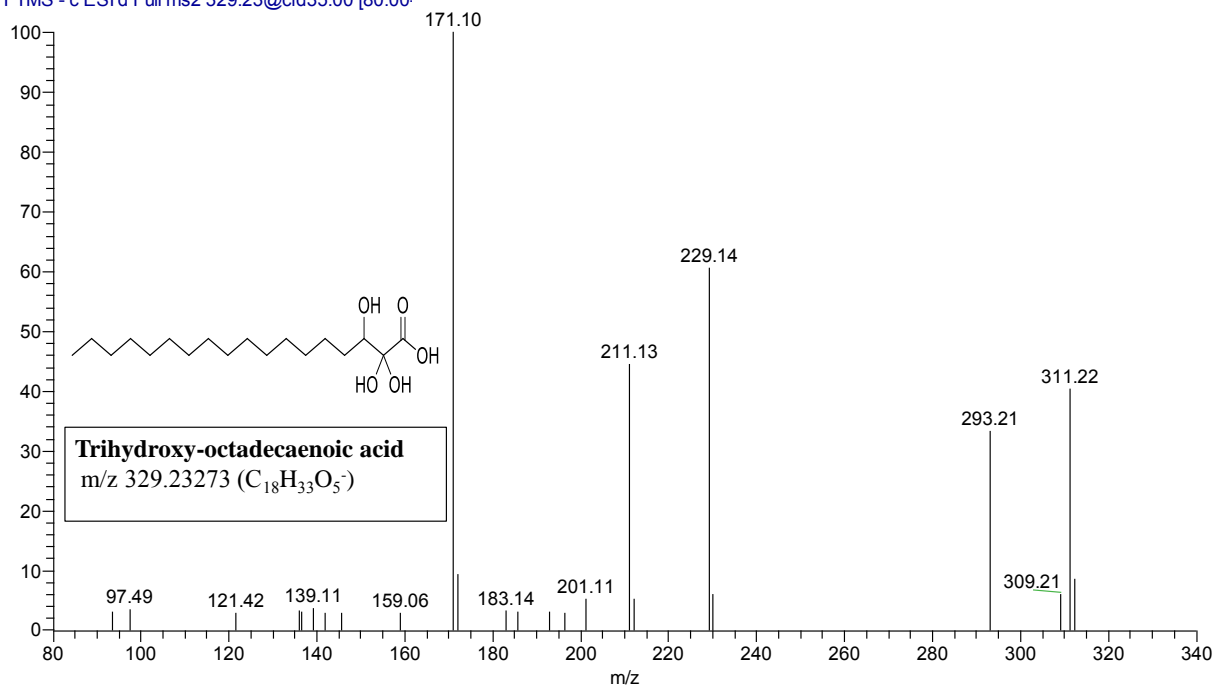
In positive mode, another class of sphingolipids known as glycosylated ceramides characterized by fragment ion at  $m/z$  184 due to the loss of phosphocholine were annotated in P59, P63, P65 at  $m/z$  496, 522, 520, respectively, **Figure 56-Figure 58**.

FAM139\_GCCn#5256 RT: 19.14 AV: 1 NL: 2.90E3  
 F: FTMS - c ESI d Full ms2 341.27@cid35.00 [80.00]



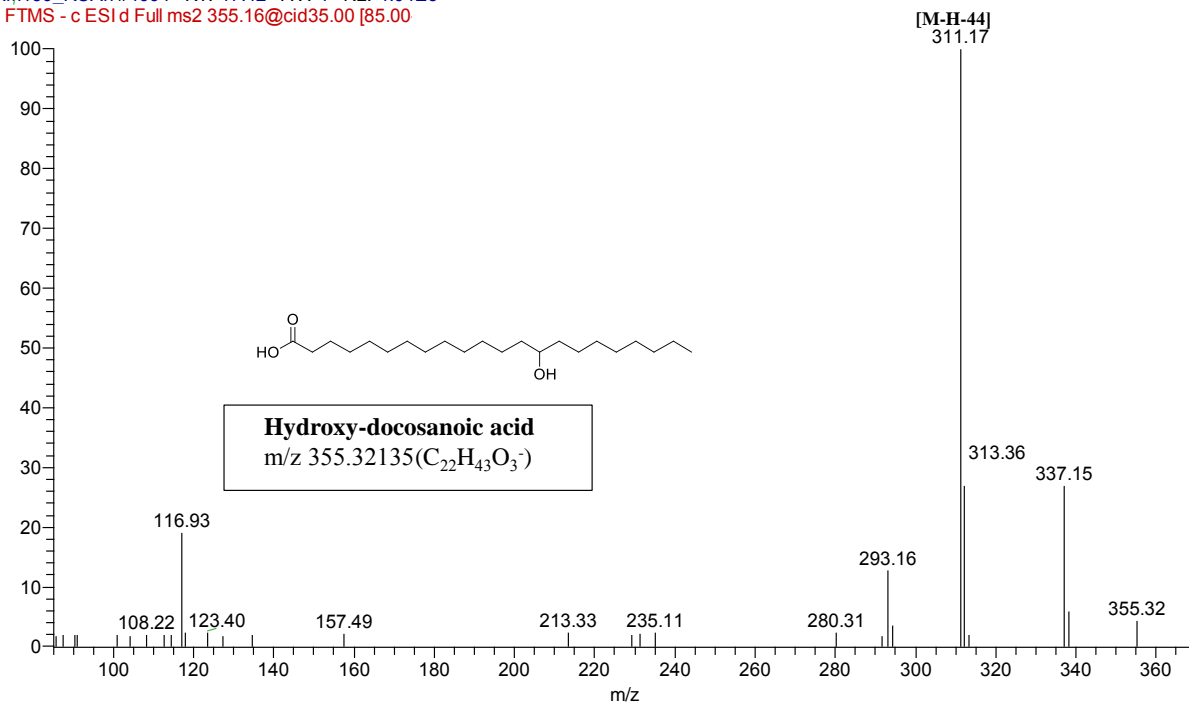
**Figure 48:** MS/MS spectrum of peak (P71) in negative ion mode

FAM139\_GCA#3362 RT: 12.24 AV: 1 NL: 2.61E3  
 T: FTMS - c ESI d Full ms2 329.23@cid35.00 [80.00]



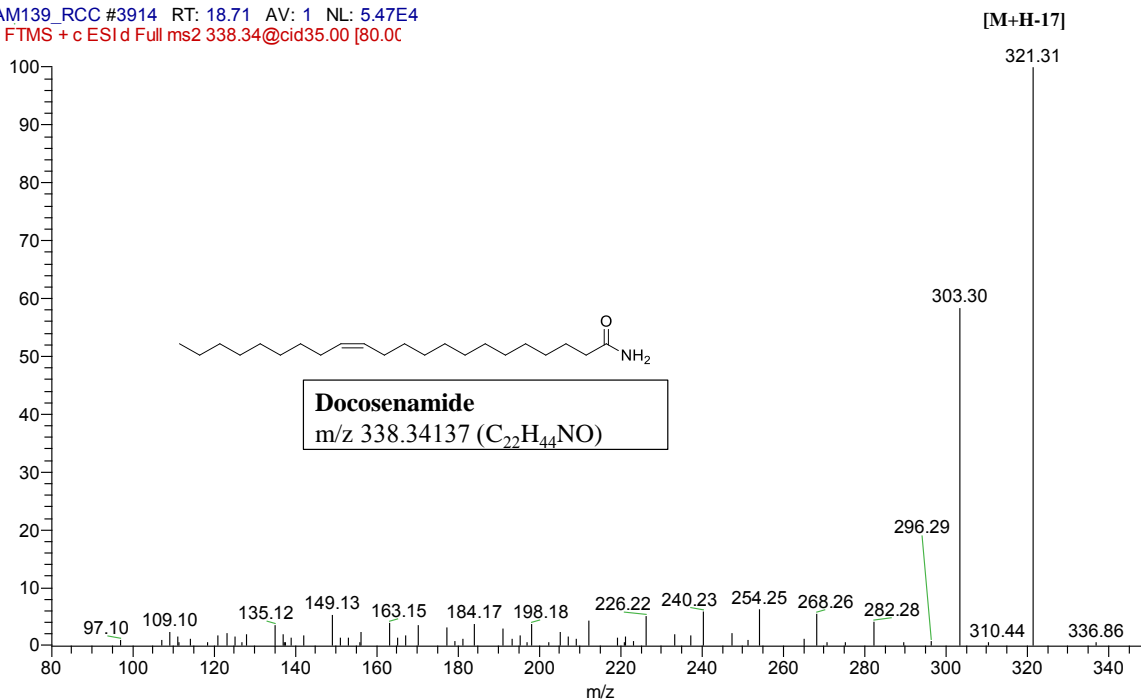
**Figure 49:** MS/MS spectrum of peak (P52) in negative ion mode

FAM139\_RCA#4694 RT: 17.42 AV: 1 NL: 4.04E3  
 F: FTMS - c ESI d Full ms2 355.16@cid35.00 [85.00]



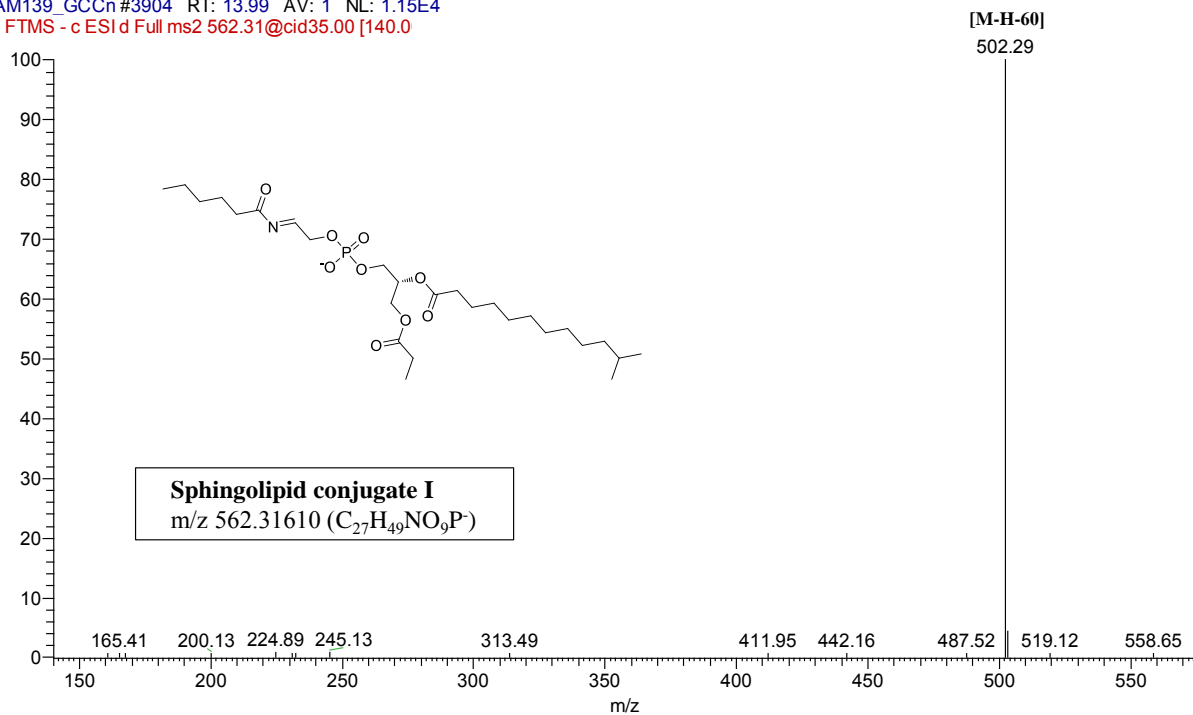
**Figure 50:** MS/MS spectrum of peak (P72) in negative ion mode

FAM139\_RCC #3914 RT: 18.71 AV: 1 NL: 5.47E4  
 F: FTMS + c ESI d Full ms2 338.34@cid35.00 [80.00]



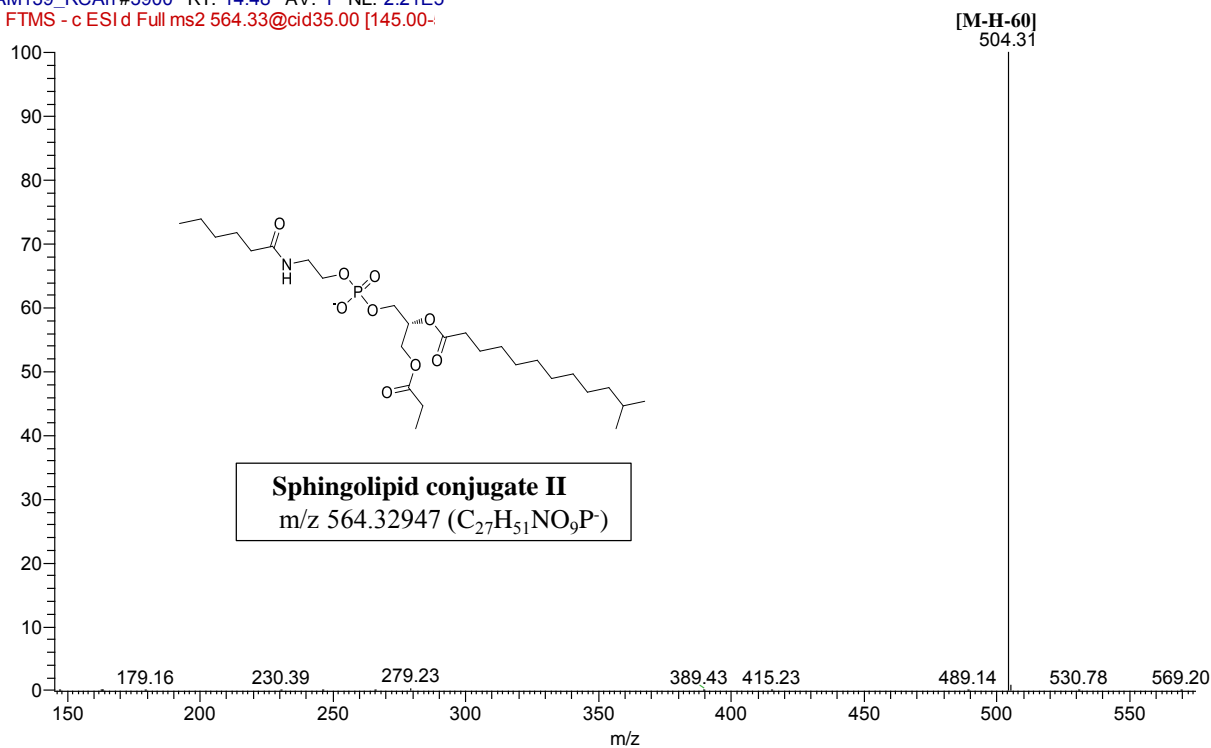
**Figure 51:** MS/MS spectrum of peak (P75) in positive ion mode

FAM139\_GCCn#3904 RT: 13.99 AV: 1 NL: 1.15E4  
 F: FTMS - c ESI d Full ms2 562.31@cid35.00 [140.0



**Figure 52:** MS/MS spectrum of peak (P56) in negative ion mode

FAM139\_RCA#3900 RT: 14.48 AV: 1 NL: 2.21E5  
 F: FTMS - c ESI d Full ms2 564.33@cid35.00 [145.00-



**Figure 53:** MS/MS spectrum of peak (P57) in negative ion mode

FAM139\_RCA#3978 RT: 14.78 AV: 1 NL: 1.38E5  
F: FTMS - c ESI d Full ms2 540.33@cid35.00 [135.0]

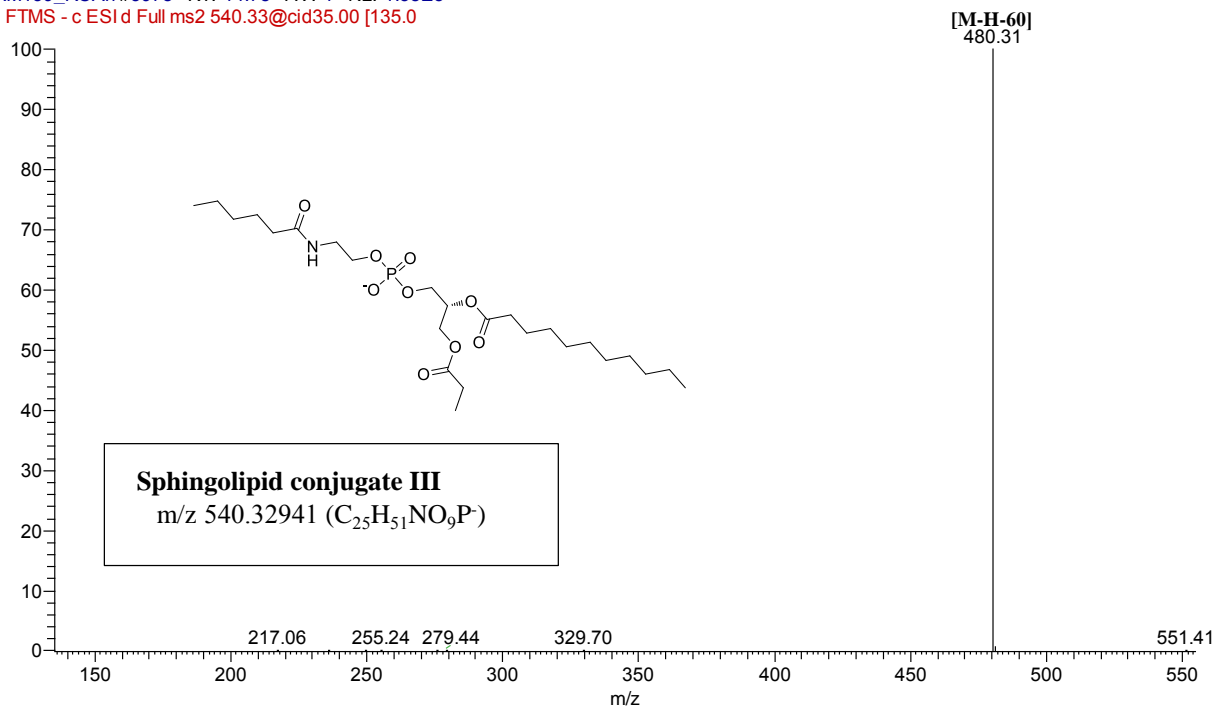


Figure 54: MS/MS spectrum of peak (P60) in negative ion mode

FAM139\_RCA#4052 RT: 15.06 AV: 1 NL: 5.26E4  
F: FTMS - c ESI d Full ms2 566.34@cid35.00 [145.0]

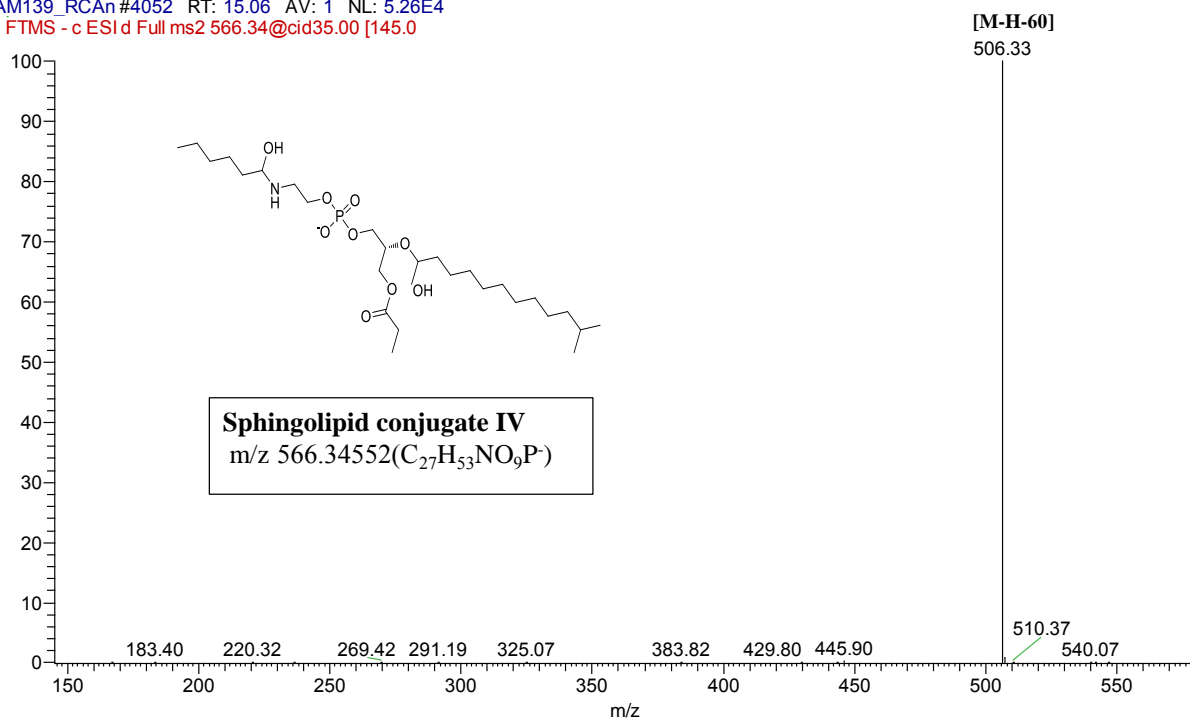
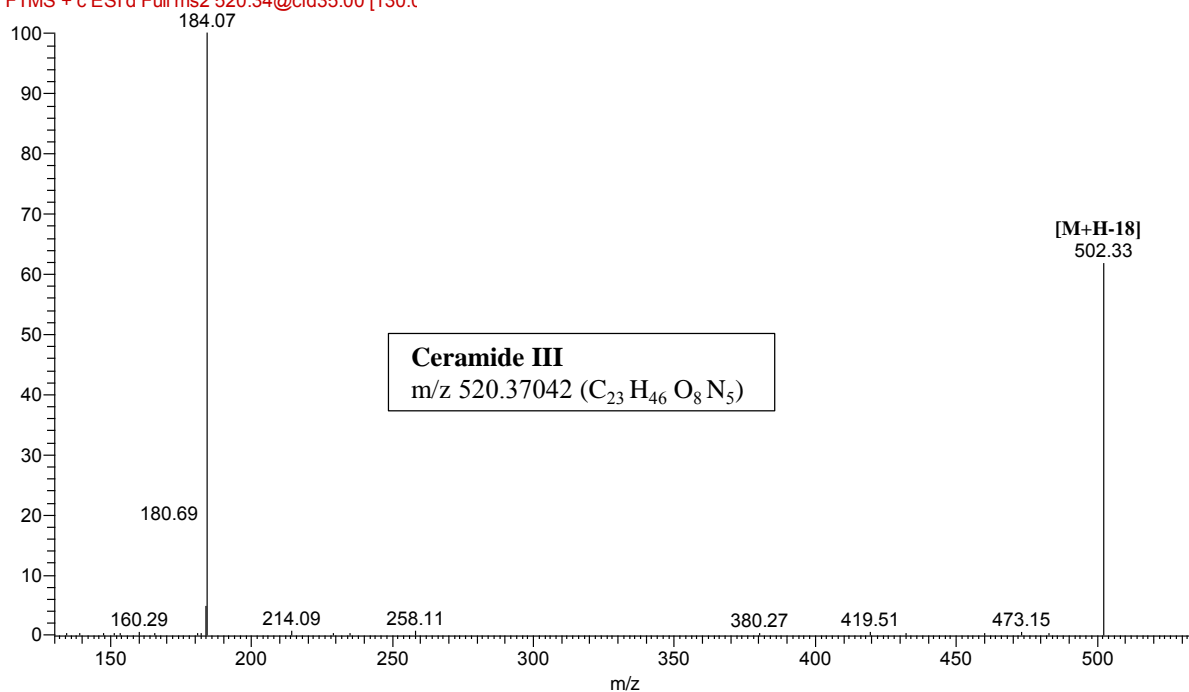


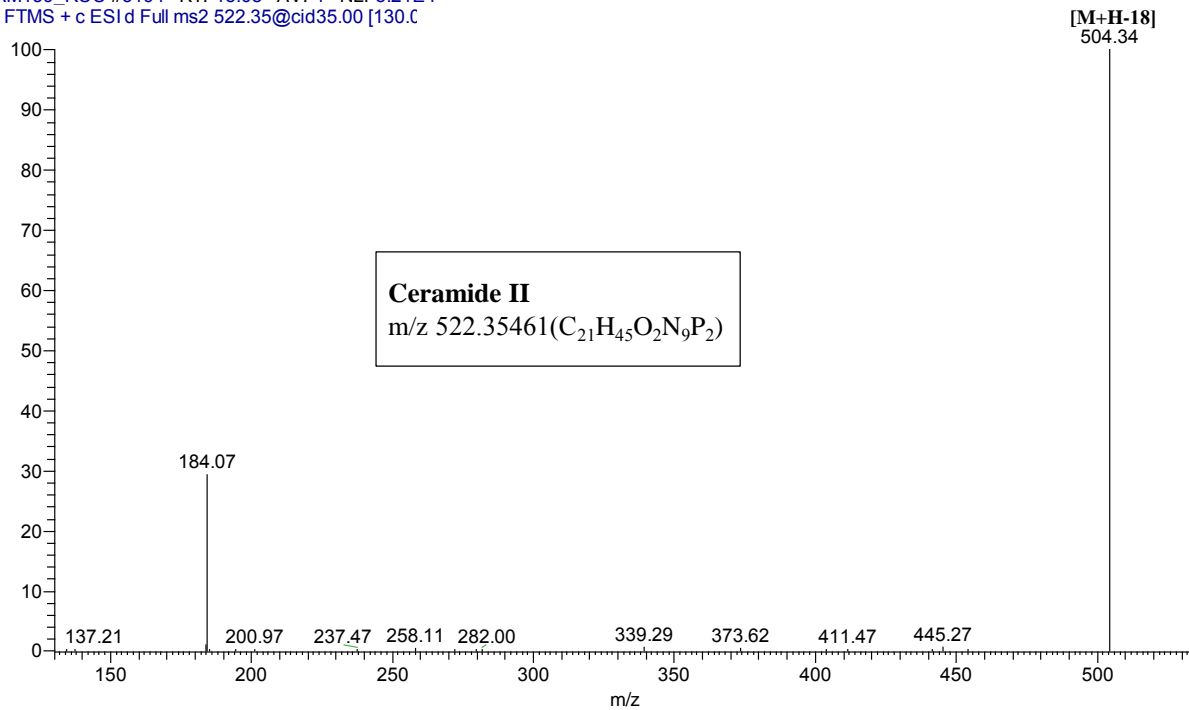
Figure 55: MS/MS spectrum of peak (P62) in negative ion mode

FAM139\_RCC #3022 RT: 14.29 AV: 1 NL: 7.50E4  
 F: FTMS + c ESI d Full ms2 520.34@cid35.00 [130.0]

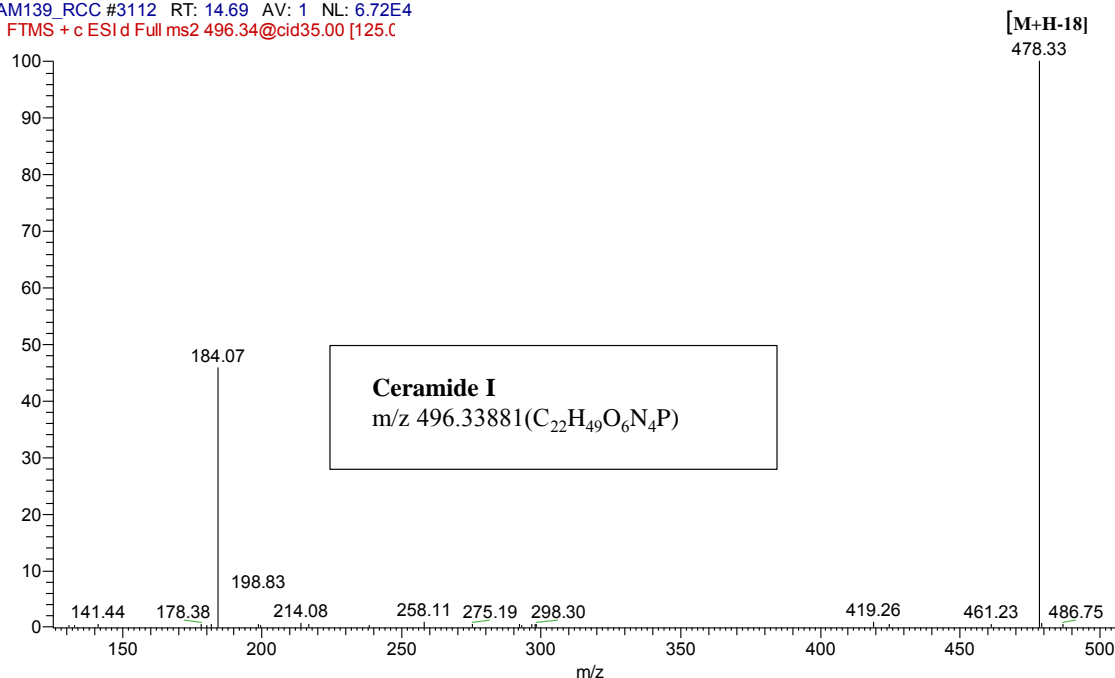


**Figure 56:** MS/MS spectrum of peak (P65) in positive ion mode

FAM139\_RCC #3194 RT: 15.08 AV: 1 NL: 8.21E4  
 T: FTMS + c ESI d Full ms2 522.35@cid35.00 [130.0]



**Figure 57:** MS/MS spectrum peak (P63) in positive ion mode

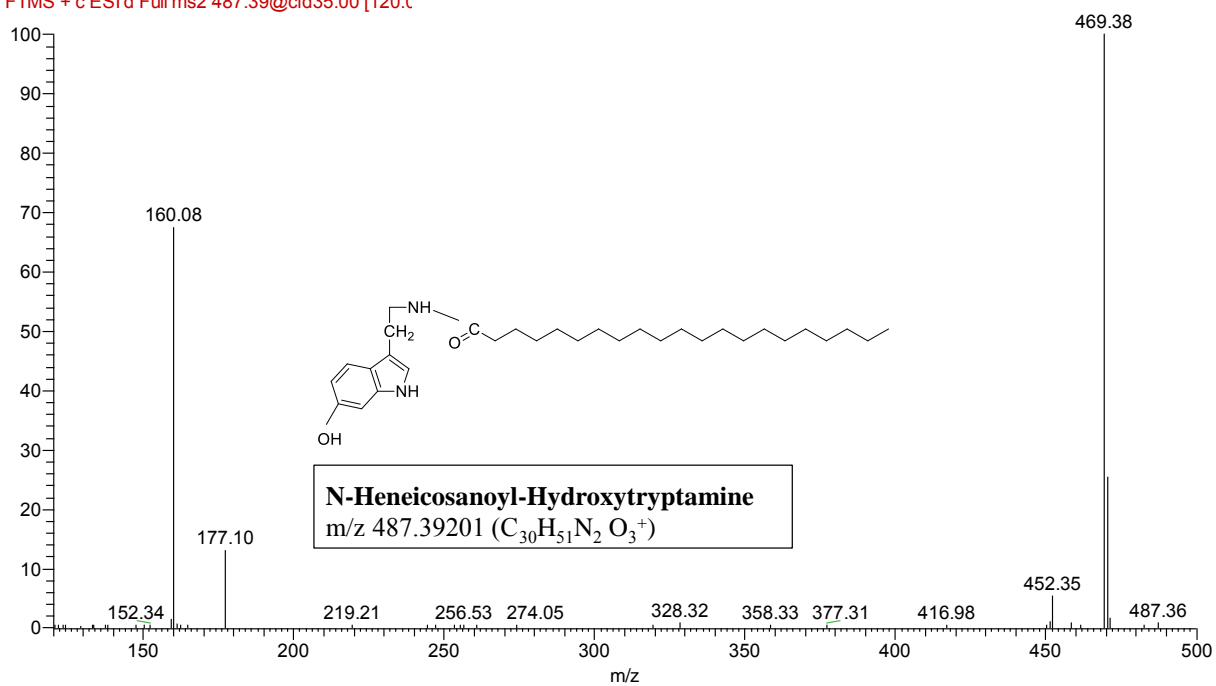


**Figure 58:** MS/MS spectrum peak (P59) in positive ion mode

### 1.3.6 Nitrogenous compounds

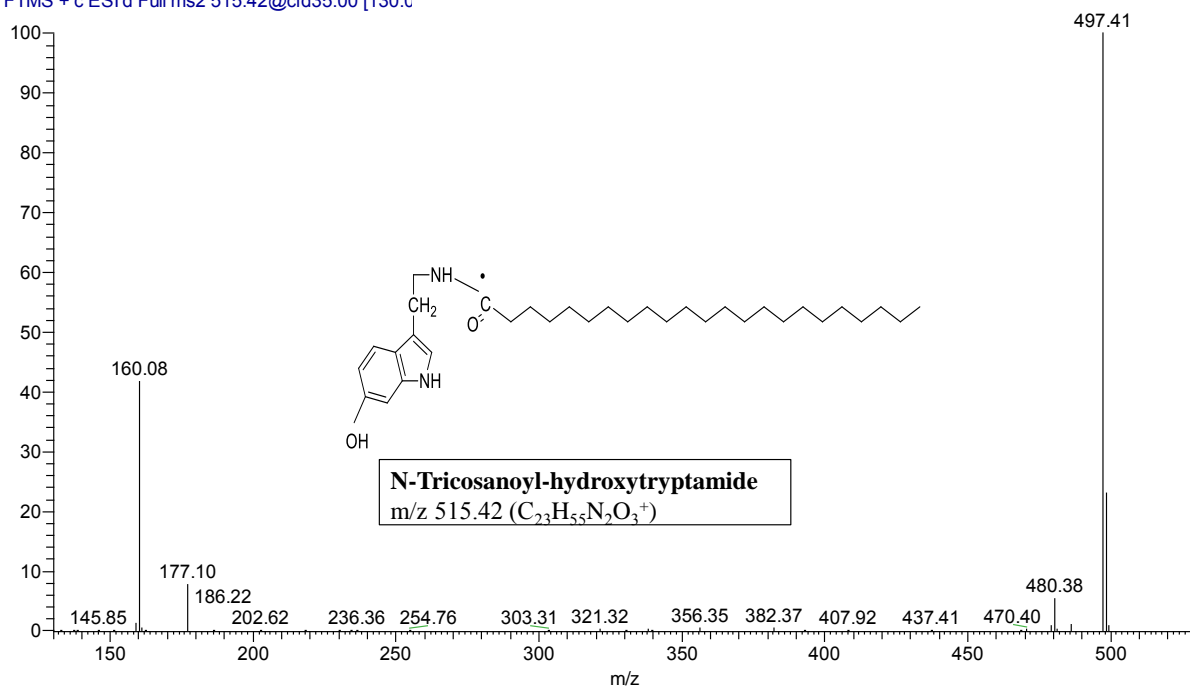
Several peaks corresponded to *N*-alkanoyl-5-hydroxytryptamides (C-5HTs) were detected in positive mode, where the amino group of 5-hydroxytryptamine (5-HT) is combined with an octadecanoyl (C18), eicosanoyl C20, heneicosanoyl C21 producing *N*-octadecanoyl-5-hydroxytryptamide (C18-5HT, P79), and eicosanoyl-5-hydroxytryptamide (EHT, P80), and heneicosanoyl-5-hydroxytryptamide (C21-5HT, P76) respectively. These compounds were detected at  $m/z$  443 for C18 5-HT,  $m/z$  471 for EHT, and  $m/z$  487 for C21-5HT in the positive ion mode, **Table 6**. Moreover, *N*-tricosanoyl-5-hydroxytryptamine (C23-5HT, peak 77), *N*-docosanoyl-5-hydroxytryptamine [C22-5HT, P78], and *N*-tetracosanoyl-5-hydroxytryptamine [C24-5HT, P81] were detected at  $m/z$  515, 499, and 527 respectively. All compounds shared the same fragment ion at  $m/z$  160, which corresponds to 5-hydroxy tryptamine and confirmed from UV  $\lambda_{\max}$  at 226 nm detected in both green and roasted samples of both canephora and arabica coffee seeds in agreement with previous reports, **Figure 59-64** [102, 103].

FAM139\_GCA #3228 RT: 15.92 AV: 1 NL: 5.91E4  
 F: FTMS + c ESI d Full ms2 487.39@cid35.00 [120.0]



**Figure 59:** MS/MS spectrum peak (P76) in positive ion mode

FAM139\_GCA #3398 RT: 16.69 AV: 1 NL: 2.28E5  
 T: FTMS + c ESI d Full ms2 515.42@cid35.00 [130.0]



**Figure 60:** MS/MS spectrum peak (P77) in positive ion mode



FAM139\_GCA #3916 RT: 19.31 AV: 1 NL: 1.16E5  
T: FTMS + c ESI d Full ms2 499.42@cid35.00 [125.0

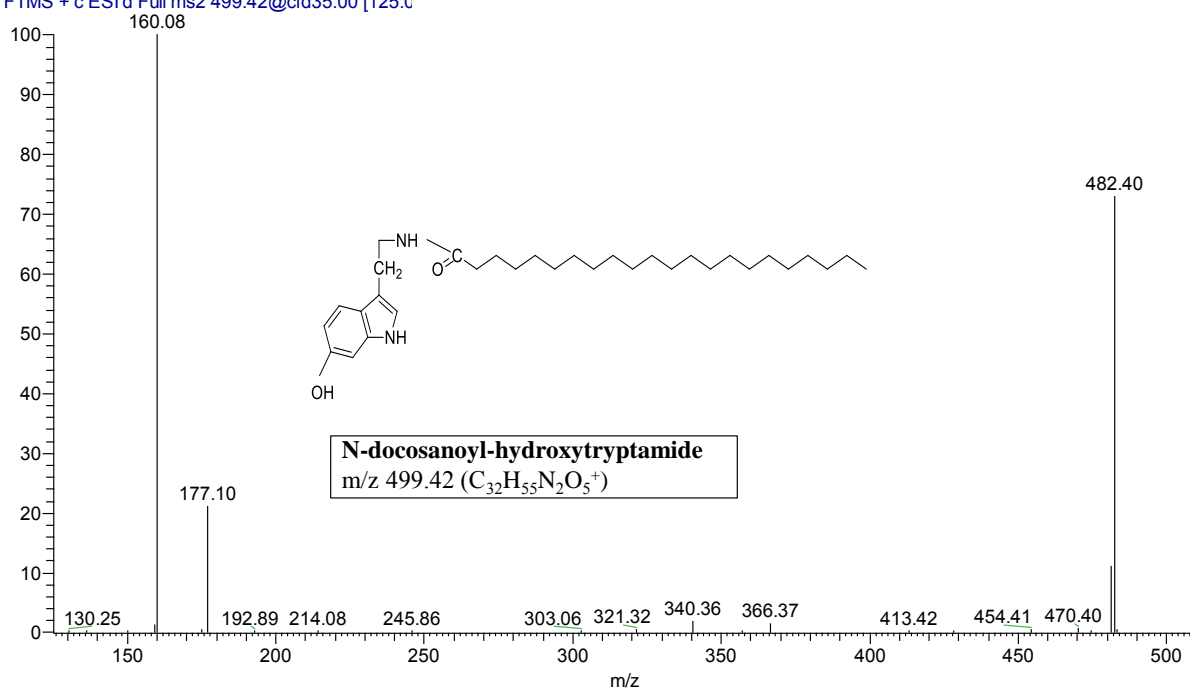


Figure 61: MS/MS spectrum peak (P78) in positive ion mode

FAM139\_GCA #3554 RT: 17.47 AV: 1 NL: 9.96E4  
F: FTMS + c ESI d Full ms2 443.36@cid35.00 [110.0

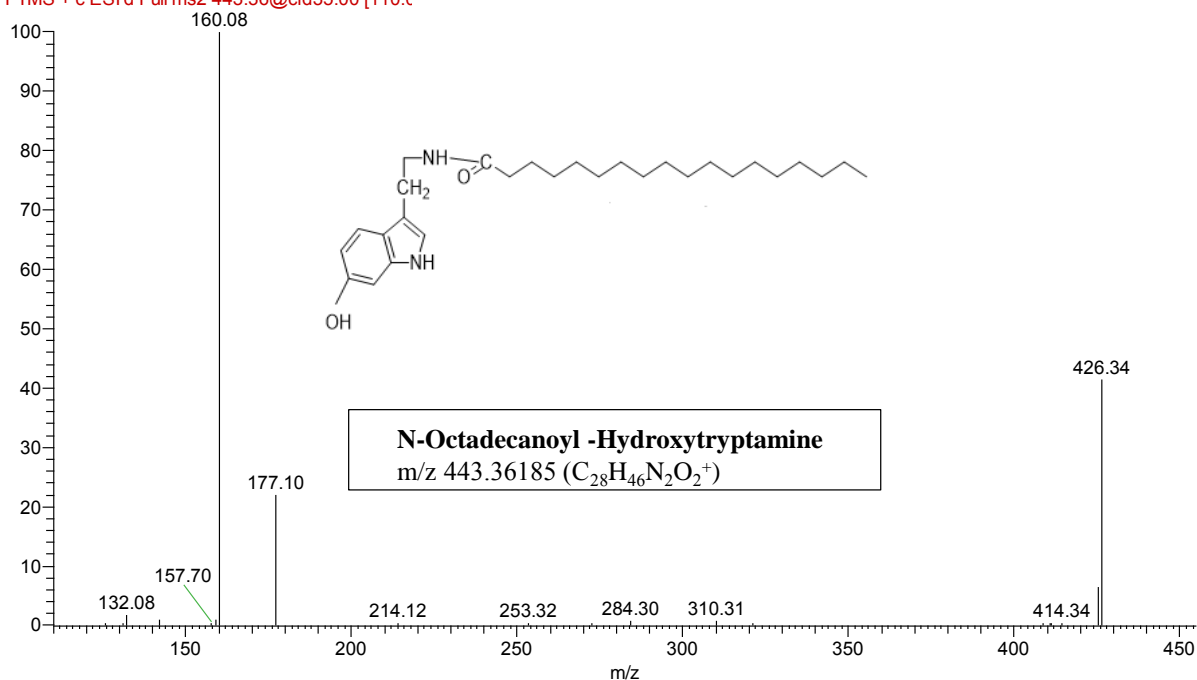
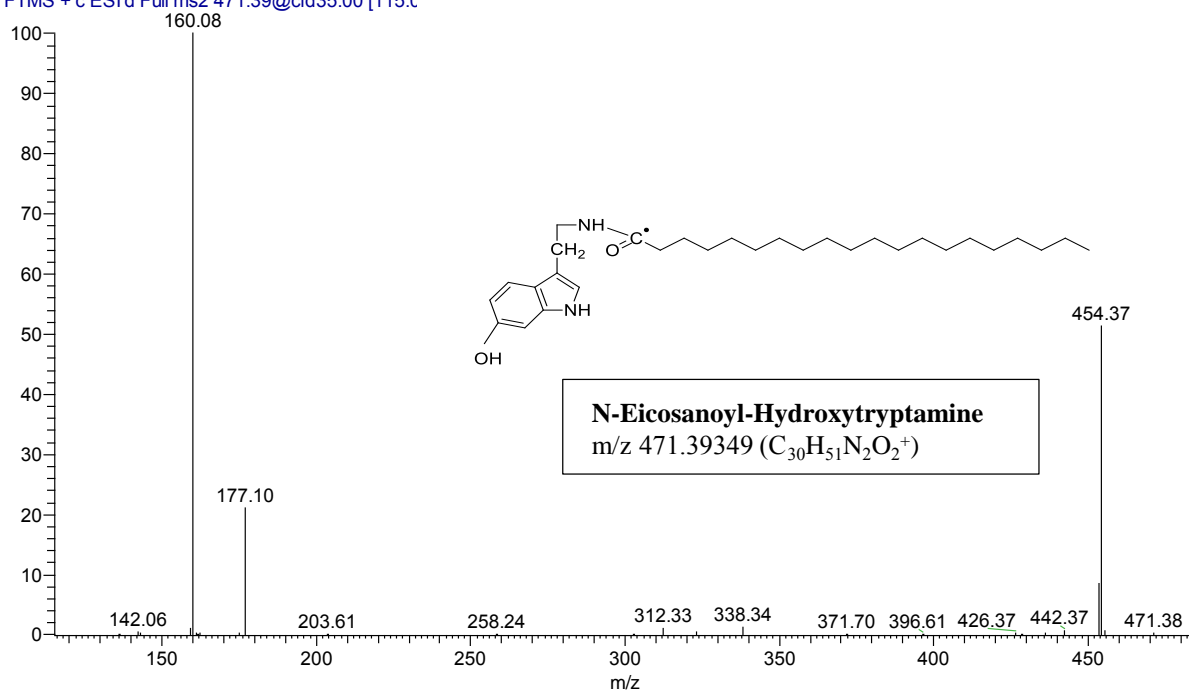


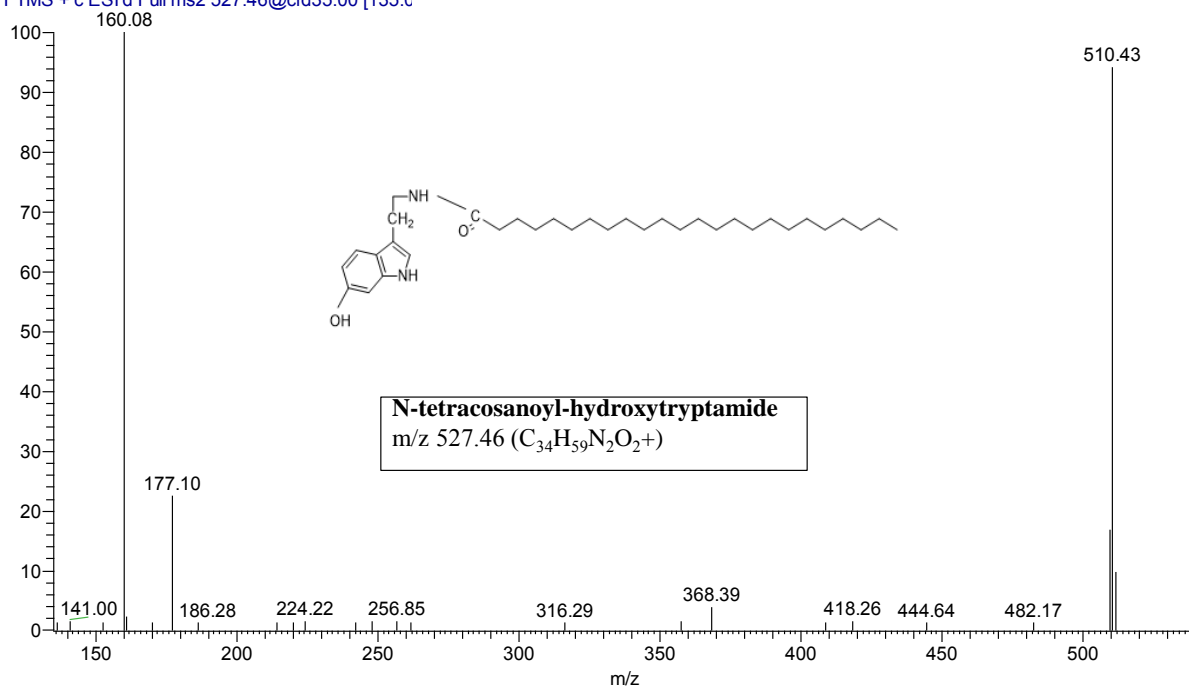
Figure 62: MS/MS spectrum peak (P79) in positive ion mode

FAM139\_GCA #3738 RT: 18.39 AV: 1 NL: 2.37E5  
T: FTMS + c ESI d Full ms2 471.39@cid35.00 [115.0]



**Figure 63:** MS/MS spectrum peak (P80) in positive ion mode

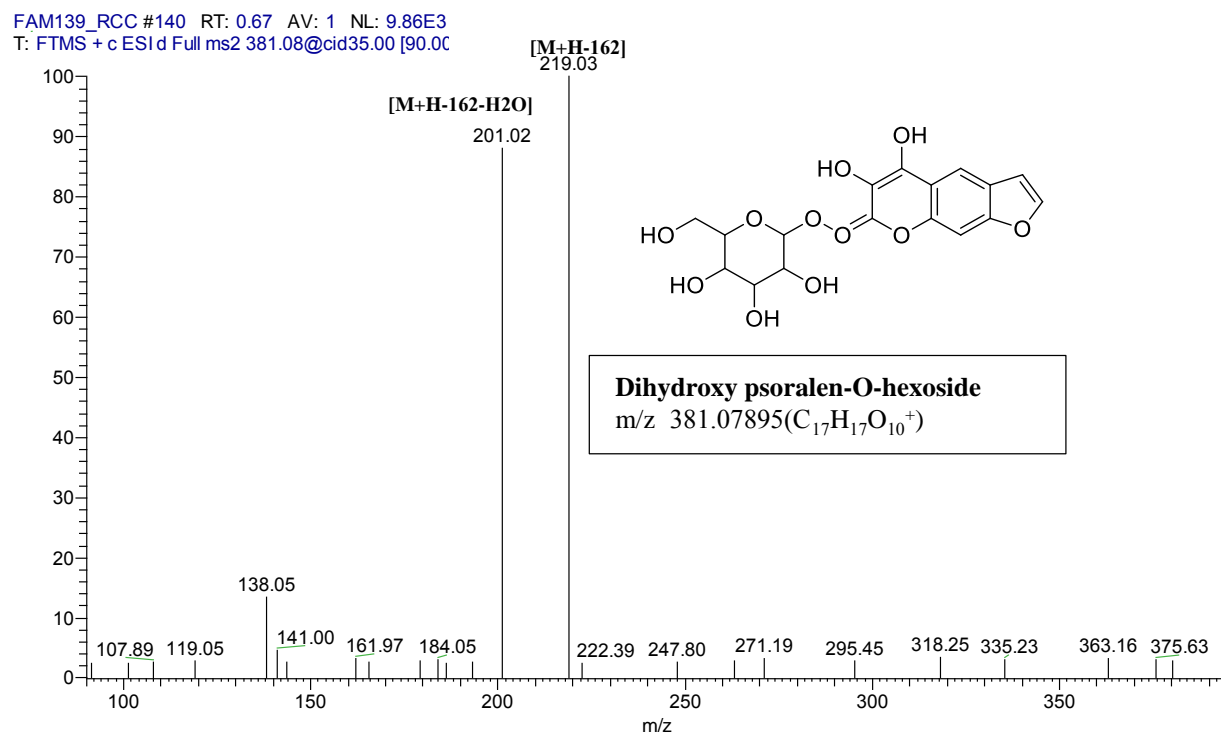
FAM139\_GCA #4106 RT: 20.32 AV: 1 NL: 2.11E4  
T: FTMS + c ESI d Full ms2 527.46@cid35.00 [135.0]



**Figure 64:** MS/MS spectrum peak (P81) in positive ion mode

### 1.3.7 Miscellaneous

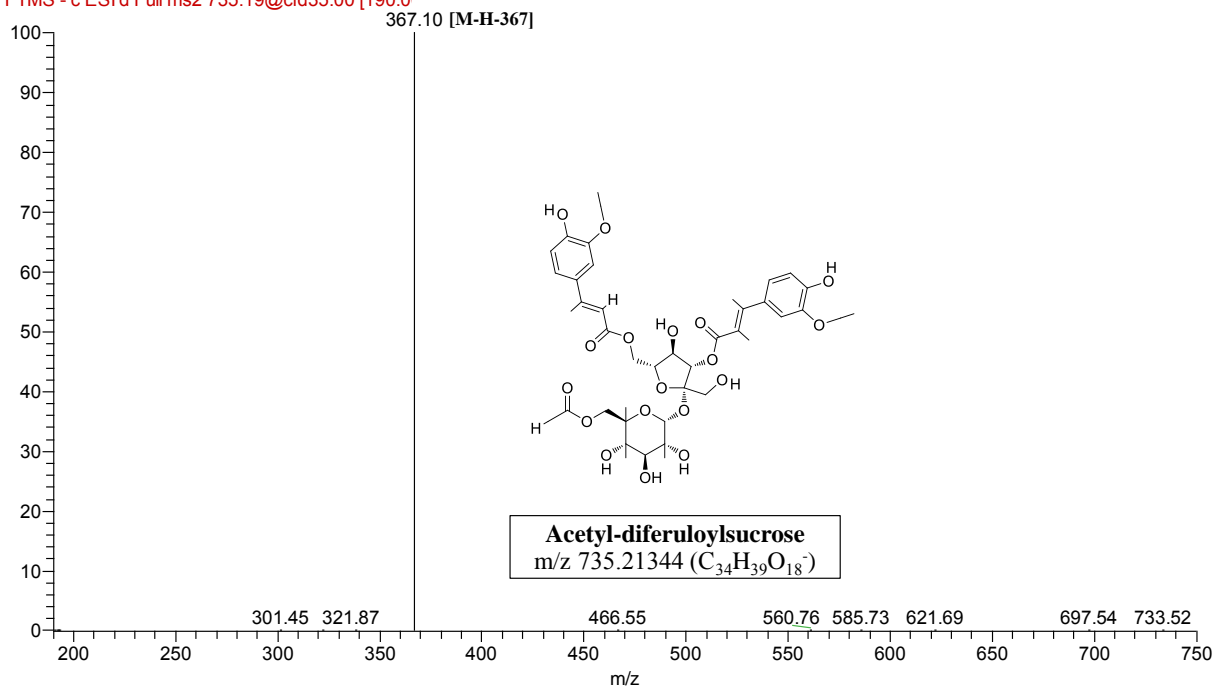
Coumarin (P1) was detected in both arabica and canephora, roasted and unroasted as a-dihydroxy psoralen-*O*-hexoside (M+H) at  $m/z$  381 and product ions at  $m/z$  219 (M+H-162), and  $m/z$  203 (M+H-162-OH) detected for the first time in coffee, **Figure 65** [104]. Phenylpropanoid esters of sucrose are common secondary metabolites in plants but not reported before in coffee. Feruloyl ester of sucrose identified at  $m/z$  735 (P33) in negative mode with product ion at  $m/z$  367 was detected in green arabica and canephora species **Figure 66** [105].



**Figure 65:** MS/MS spectrum peak (P1) in positive ion mode

FAM139\_GCCn #2468 RT: 8.82 AV: 1 NL: 2.63E5

F: FTMS - c ESI d Full ms2 735.19@cid35.00 [190.0



**Figure 66:** MS/MS spectrum peak (P33) in negative ion mode

#### 1.4 Multivariate data analysis and fingerprinting of coffee samples analyzed using UPLC/MS.

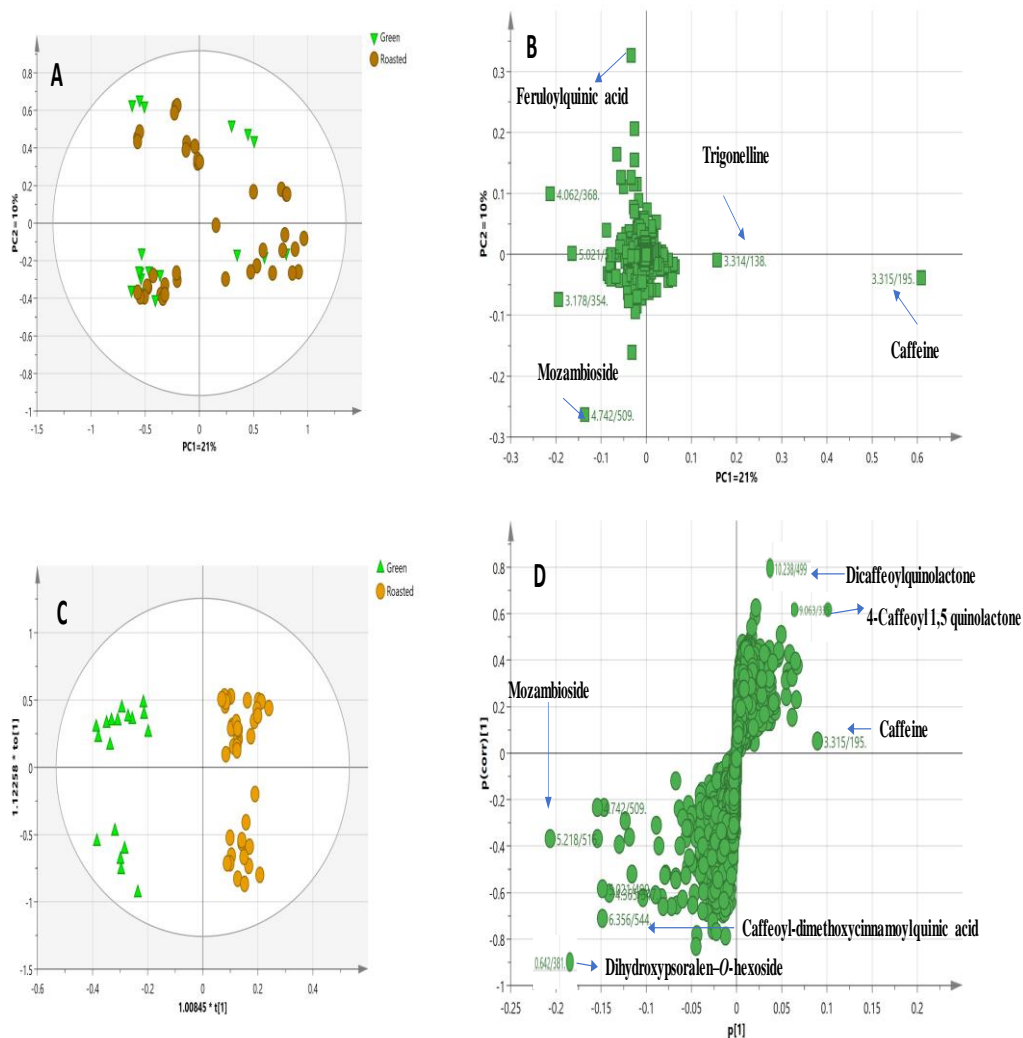
Although differences in metabolites composition could be revealed from the visual inspection of UPLC/MS chromatograms of coffee specimens, we attempted to analyze the dataset holistically extracted from the UPLC/MS using multivariate data analyses. Several models were constructed to classify coffee samples according to their species, roasting indices, and different blends, i.e., cardamom.

PCA was first applied as an unsupervised tool for exploring the relationship among all samples, whereas OPLS-DA was used to investigate samples according to genotype and roasting process and identification of markers for specimens. All models were evaluated based on R<sup>2</sup> and Q<sup>2</sup> values denoting for variance coverage and model prediction power, respectively, to avoid an overfitted model, especially in the case of OPLS.

#### 1.4.1. Roasted versus non-roasted coffee.

PCA was first applied to the 19 coffee samples, including authenticated and commercial preparations, with specimens, denoted different symbols for roasted (RCA, RCC, LRCM, LRS, HRKC, BRK, LRCK, BRA, LRSQ, LRCS, LRCQ, ICA, ICC) versus unroasted ones (GCA, GCC, GCU, GCE, GCS, GCK). Based on the UPLC/MS dataset, the PCA score plot **Figure 67A** showed values for  $R^2=0.49$  and  $Q^2=0.38$ , indicating an acceptable model, though not showing clear segregation of roasted from green samples, with some overlap between roasted and unroasted specimens. The loading plot **Figure 67B** revealed few markers that were responsible for such segregation, including caffeine (P6) and, to less extent, trigonelline (P5) enriched in roasted samples, while CGA isomers such as feruloyl quinic acid (P11) was found abundant in unroasted specimens. These results fall in accordance with previous reports in coffee, which revealed enrichment of unroasted coffee with phenolic compounds that tend to degrade upon roasting [42][29], [97], [98] [57]. Similarly, diterpenes represented by mozambioside (P39) was revealed as a marker in green samples and likely to be degraded upon roasting in black coffee [98].

In an attempt to identify more markers, the OPLS model was established, which showed  $R^2$  &  $Q^2$  of 0.92 and 0.73 respectively, with better samples segregation, of **Figure 67C**. The S-plot showed other markers for HCAs, i.e., caffeoyl-dimethoxy cinnamoyl quinic acid (P24) and coumarin derivative dihydroxypsoralen-O-hexoside (P1) for green specimens. In contrast, caffeoyl-quinolactone (P13) and dicaffeoylquinolactone (P18) were distinguished for the roasted samples and likely to generate during roasting due to dehydration reactions for the quinic acid derivatives and the formation of quinides or lactones of chlorogenic acids (CGL), which are responsible for the characteristic bitter taste of roasted coffee **Figure 67D** [97] 113].



**Figure 67:** Multivariate PCA and OPLS analyses of coffee specimens labelled as (A) and (C) for green and roasted specimens, respectively. (A) Principal component analysis (PCA) score plot, (B) PCA loading plot, (C) Orthogonal projections to latent structures discriminant analysis (OPLS-DA) score plot, and (D) OPLS-DA S-plot show covariance  $p[1]$  against the correlation  $p(\text{cor})[1]$  of the variables of the discriminating component of the OPLS-DA models. Selected variables are highlighted in the S-plot and are discussed in the text.

#### 1.4.2. Instant versus roasted samples

As instant coffee production usually involves additional processing steps rather than just roasting, PCA was run to determine the variability between roasted and instant samples **Figure 68A**.

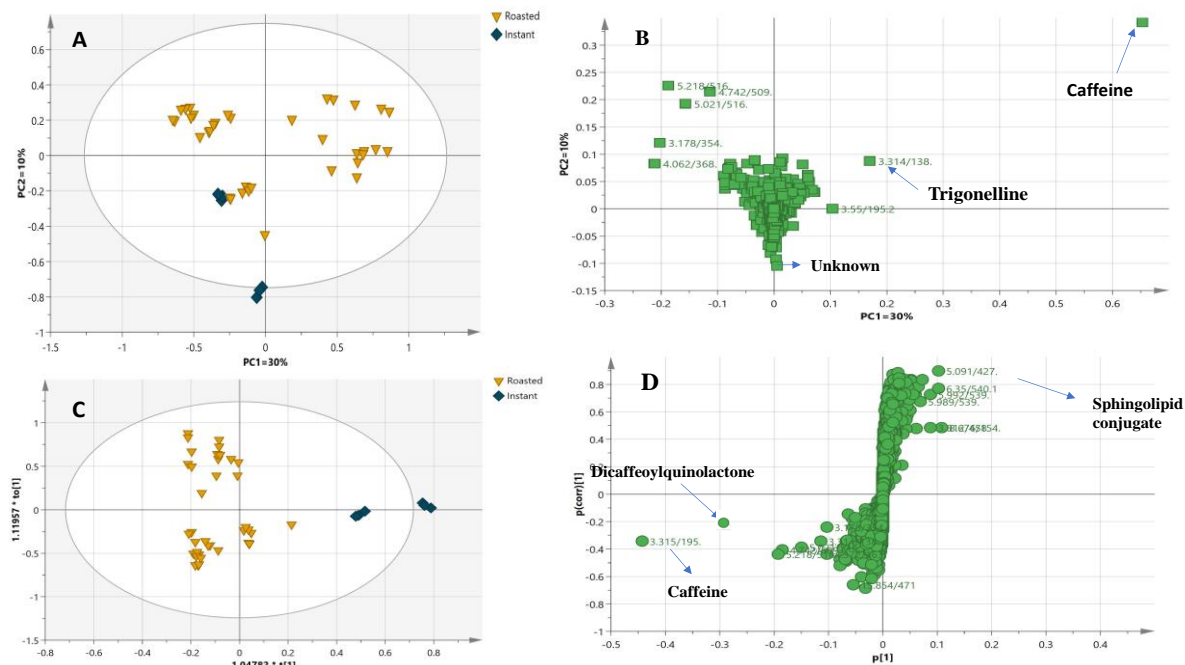
PCA was applied to 12 coffee samples, including an instant sample (ICA), an instant sample with cardamom (ICC), along with 10 roasted specimens denoted different symbols (RCA, RCC, LRCM, LRS, HRKC, BRK, LRCK, BRA, LRSQ, LRCS).

The PCA model showed that instant sample without cardamom was well separated from other samples on the two PC projections representing 40% of the total variance **Figure 68B**.

The low PCs value might be attributed to the reason that the instant sample blended with cardamom, *i.e.*, ICC, was not segregated from the plain instant sample. Also, the roasting degree between coffee samples was variable. Hence, an OPLS model was constructed, which showed better separation parameters (PC1 & PC2= 85%) and R2 and Q2 (0.85 and 0.78), respectively **Figure 68C**.

Moreover, the S- plot revealed that caffeine(P6) and dicaffeoyl quinolactone(P19) were the major markers for roasted samples, while sphingolipids were interestingly predicted to be the main markers for the instant sample **Figure 68D**.

Sphingolipids are highly complex molecules that are involved in cell membrane function, such as growth arrest, proliferation, apoptosis, and signaling molecules. Another factor that might be contributing to the separation of instant samples from roasted samples is acrylamide' enrichment in instant than roasted coffee that was reported by many studies. However, it is hardly detected by LCMS due to its low molecular mass better suited for detection by GC/MS [58].



**Figure 68:** Multivariate PCA and OPLS analyses of coffee specimens labelled as (A) and (C) for Instant and roasted specimens, respectively (A) Principal component analysis (PCA) score plot, (B) PCA loading plot, (C) Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot, and (D) OPLS-DA S-plot, show covariance  $p[1]$  against the correlation  $p(\text{cor})[1]$  of the variables of the discriminating component of the OPLS-DA models. Selected variables are highlighted in the S-plot and are discussed in the text.

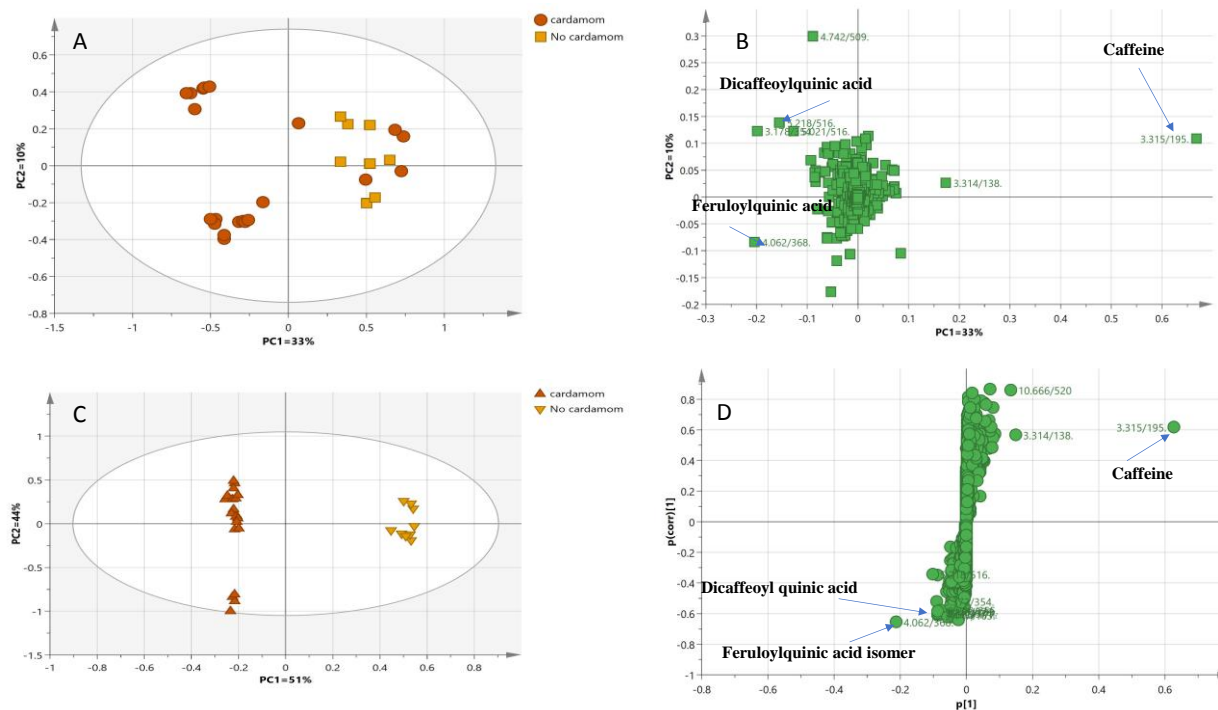
### 1.4.3. Cardamom versus plain samples

PCA was applied to 10 samples, including cardamom blended and plain samples, with specimens denoted different symbols for cardamom blended (LRCM, HRKC, LRCK, LRSQ, LRCS) versus plain ones (RCA, RCC, BRA, BRK, LRS). The score plot model showed clustering of samples blended with cardamom mainly on the left side, while plain samples along with a few cardamoms blended were placed on the right side along PC1 to cover 33% of the total variance of 43%, **Figure 69A**. The loading plot revealed that coffee samples blended with cardamom were richer in phenolic acids cardamom, i.e., dicafeoyl (P19) and feruloyl quinic acid (P11) versus plain samples more abundant in alkaloids such as caffeine (P6) **Figure 69B** [106].

Additionally, the OPLS score plot showed improved discrimination between investigated samples with better PCs responsible for 91% of the total variance ( $R^2=0.99$  &  $Q^2=0.73$ ), **Figure 69C**. Finally, the S-Plot model confirmed PCA loading results regarding the higher abundances



of caffeine in plain coffee while chlorogenic acids, *i.e.*, dicaffeoyl quinic acid and feruloyl quinic acids, were more in coffee blended with cardamom products, **Figure 69D**. The obtained results were aligned with previous analysis of cardamom by HPLC that has been reported presence of phenolic compounds, *i.e.*, tannic, caffeic, gallic, and dicaffeoylquinic acids [106].



**Figure 69:** Multivariate PCA and OPLS analyses of coffee specimens labelled as (cardamom) and (no cardamom) for roasted cardamom and plain roasted specimens, respectively (A) Principal component analysis (PCA) score plot, (B) PCA loading plot, (C) Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) score plot, and (D) OPLS-DA S-plot. Show covariance  $p[1]$  against the correlation  $p(\text{cor})[1]$  of the variables of the discriminating component of the OPLS-DA models. Selected variables are highlighted in the S-plot and are discussed in the text.

## **2 Chapter II: Fingerprinting of coffee seeds *via* UV spectroscopy coupled to chemometrics.**

## 2.1 UV fingerprinting of coffee seeds

UV-vis technique was further applied to identify metabolome components' in coffee using the same extraction method employed for UPLC/MS. UV spectral bands provide information mostly about conjugated bioactive components characteristic profile (phenolic acids, methylxanthines, chlorogenic acids, pentacyclic alcohol, and acids [107]. Absorption bands are usually due to chromophores which are more abundant in aromatic conjugated compounds like flavonoids and phenolic compounds. The phenolic compounds in coffee are mainly chlorogenic acids to absorb at  $\lambda$  max 220 and 325 nm. These bands correspond to the transition from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular weight orbitals (LUMO) with  $\pi$  to  $\pi^*$  type [46]. Additionally, unsaponifiable terpenes, i.e., cafestol and kahweol, are reported to absorb in UV at wavelength  $\lambda$ max 225, 288 nm respectively, while caffeine and trigonelline  $\lambda$  max occurs between (271-275). UV presents cheaper and simpler techniques and consequently can be used as a reference for other herbal plants [108].

Despite the many advantages of the UV platform, few metabolomic studies have been conducted on a limited number of coffee samples in the context of authentication [19, 109]. This may be attributed to the rather complicated UV spectra due to the absorption interferences between coffee constituents. Therefore, chemometrics analysis, including the multivariate data analysis, was applied to exclude interferences between coffee constituents to fingerprint without a prior purification step. In this study, UV spectral fingerprinting 200-700 nm (absorption region after 450 nm was omitted from analysis as it was devoid of any absorption peaks) coupled to multivariate analysis was applied for the differentiation between coffee specimens. Comparison between the classification potential of UV fingerprinting compared to UPLC/MS to determine its potential to be employed for QC determination as an alternative cheaper tool for UPLC/MS.

### 2.1.1 Roasted versus non-roasted

The first UV based model included green and roasted samples of 4 authenticated samples (GCA, GCC, RCA, RCC) and 15 commercial samples (LRCM, LRS, HRKC, BRK, LRCK, BRA, LRSQ, LRCS, LRCQ, ICA, ICC, GCU, GCE, GCS, GCK) **Table 5**. The PCA score plot could classify most of the samples mostly along PC1 (91%), **Figure 70A**. The loading plot revealed intense absorption of green samples in the region between 220-350 nm, suggesting their

richness in phenolic acids, i.e., hydroxycinnamic acid derivatives. In contrast, roasted samples were segregated based on their higher absorption between 375-450 nm, which could be explained by the brown color factor or presence of melanoidins ( $\lambda=420\text{nm}$ ). The chemical structure of melanoidins has not been reported to date, mostly proposed as nitrogenous polymers or high molecular weight compounds that formed via Aldol-condensation reaction. Melanoidins have been reported to have different absorption bands in UV during roasting, early roasting ( $\lambda=280\text{nm}$ ), medium roasting ( $\lambda=330$ ), and final roasting ( $\lambda=420\text{nm}$ ). Hence, the UV can be used for determining the roasting index in coffee for the coffee industry. An OPLS model was constructed for maximal separation of the samples, and green samples were distinctly clustered from roasted and showed good validation parameters ( $R^2$  and  $Q^2 = 1$ ), **Figure 70B**. Inspection of the corresponding loading S-plot revealed that the discriminant wavelengths of green samples from roasted were at (210-230) nm and (300-330) nm attributed mostly to the absorption of phenolic acids. These results provided by the model were synchronous with that obtained from the S-plot results of the UPLC/MS model that revealed enrichment of green coffee with chlorogenic acids, **Figure 70C**.

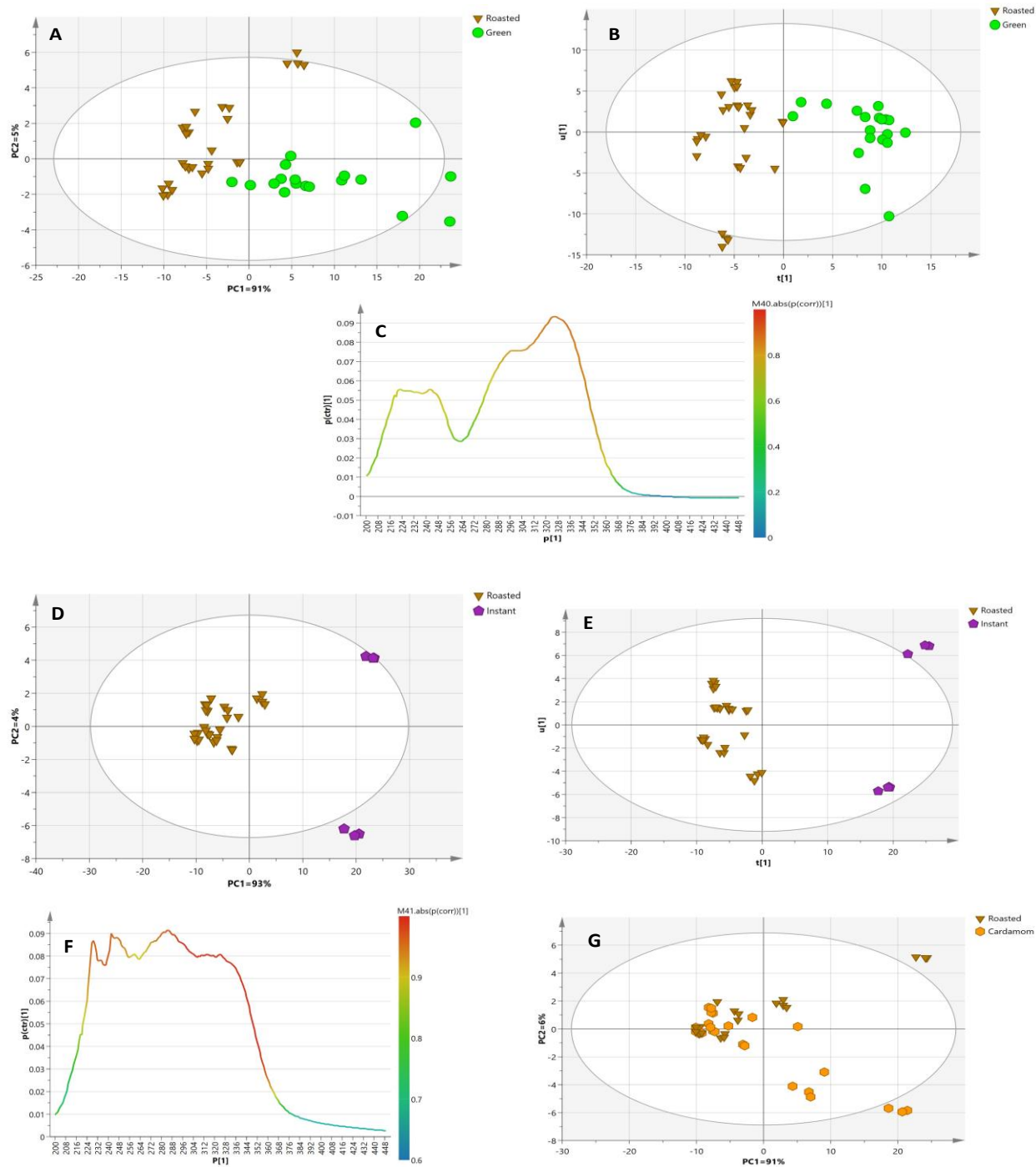
### 2.1.2 Instant versus roasted samples

A second model was performed to include 11 roasted samples only (LRCM, LRS, HRKC, BRK, LRCK, BRA, LRSQ, LRCS, LRCQ, RCA, RCC) along with the two instant samples (ICA, ICC). The instant samples were well separated along PC1 (93%), appearing as outliers in the upper part of the plot, **Figure 70D**. The corresponding loading plot revealed that instant samples absorbed more at 220 nm and 290 nm, likely attributed to fatty acids and sphingolipids. Interestingly, a UV max at 275 nm was detected likewise at 3.5-fold that in roasted arabica sample, which is likely attributed to that of acrylamide, suggesting that instant coffee has more acrylamide than roasted coffee consistently with the literature[58]. These results highlight how UV complement results derived from UPLC/MS by revealing potential coffee markers not detected in the later techniques, including melanoidins and acrylamide. For further confirmation, an OPLS model was built for better differentiation between the two sets of samples with  $R^2$  and  $Q^2$  values computed to be 0.97 and 0.91, respectively, **Figure 70E**. Upon investigation of the S-plot of the OPLS-DA loading plot, characteristic UV regions of instant samples from the roasted ones were identified in the range of 220- 290 nm, likely contributed

from absorption wavelengths of sphingolipid ( $\lambda_{\text{max}}=230\text{nm}$ ) and or acrylamide, **Figure 70F**, [58].

### **2.1.3 Cardamom versus plain samples**

A PCA model was constructed from plain coffee samples (RCA, RCC, BRK, BRA, LRS) and cardamom blended samples (LRCM, HRKC, LRCK, LRSQ, LRCS), with though no clear segregation of cardamom coffee samples from roasted samples with kind of overlap between the two specimens. The loading plot showed absorption of plain samples in the range of 350-450 nm suggesting higher melanoidins content, while the blended cardamom samples, along with some roasted samples, were more rich in phenolic acids absorption range (220-350nm), indicating the cumulative phenolic content for both cardamom and roasted coffee, **Figure 70G [106]**.

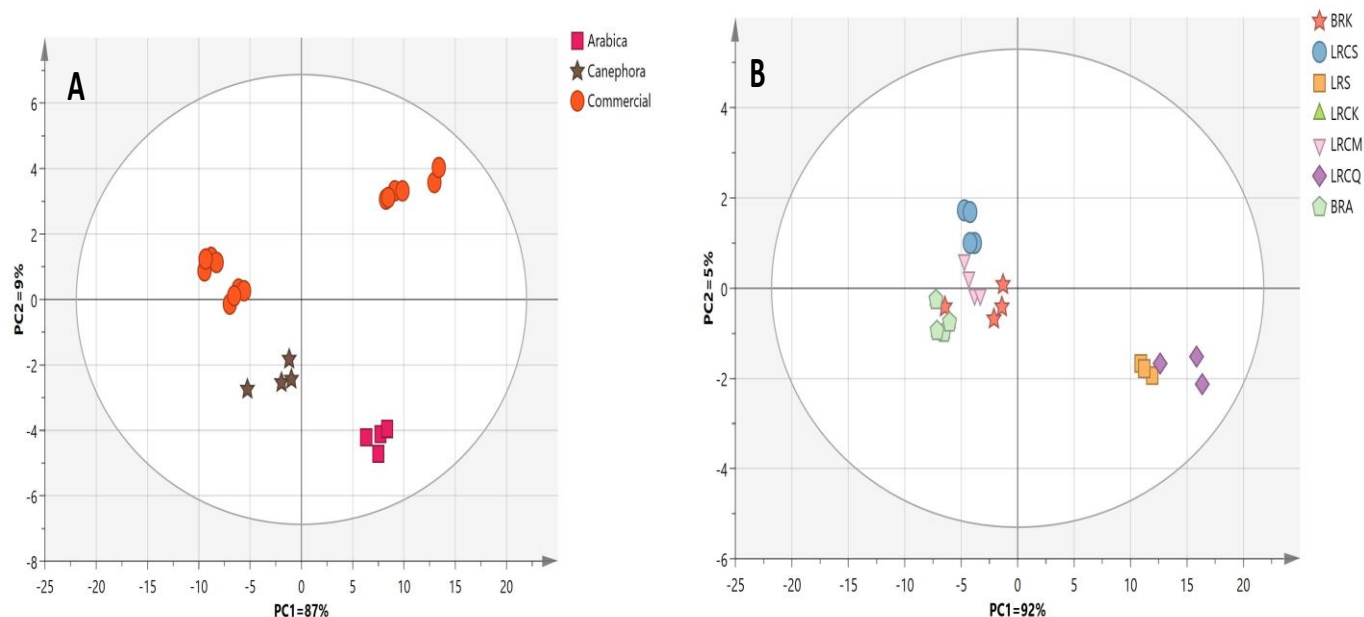


**Figure 70:** PCA score plot (A) of roasted and green, OPLS based on roasting (B), S-line based on roasting (C) PCA of instant and roasted samples (D), OPLS for roasted and instant samples (E), S-plot for instant and roasted samples (F), PCA for plain roasted and instant samples (G). The S-plot shows the correlation( $cor$ ) and covariance  $p[1]$  between variables(wavelengths) and the predictive score of the discriminating component of OPLS-DA. The discriminant wavelengths of the important variables list are highlighted and discussed in the text.

#### 2.1.4 Geographical origin and species

Geographical origin and species were investigated by three separate models employing a UV dataset. In the PCA score plot, the model constructed with the roasted authenticated samples (RCA and RCC) along with the roasted commercial (BRK, LRCS, LRS, LRCK, LRCM, LRCQ, BRA), **Table 5 & Figure71A**. Results revealed that some roasted commercial samples clustered with roasted arabica species on the right side, while others were observed on the left side with roasted canephora along with PC1=87% and with (R2 & Q2=0.99). The loading plot revealed that arabica coffee and commercial samples on the right side absorbed more in the region (220-350) while canephora coffee alongside commercial samples absorbed in a higher wavelength range (350-450nm). To confirm such results, a separate model for roasted commercial samples only was established. PCA score plot showed the lightly roasted Saudi (LRS), and Qatar samples (LRCQ) segregated together on the right side of the plot along PC1.

On the other hand, the rest of the commercial samples from Aswan and Kuwait, respectively, i.e., BRA, HRKC, BRK, etc., were grouped on the left side. The loading plot revealed the max absorption range for LRS and LRCQ samples was between 220-350 nm, suggesting their richness in phenolic acids. Asides, the rest of the samples showed a higher absorption range (370-450nm) that could be explained by their enrichment in melanoidins during the roasting process that might be related to a high roasting degree, **Figure71B**. Accordingly, LRS and LRCQ were closer to the arabica samples than the other commercial samples suggesting they are derived from pure arabica species. However, such a difference may likely be inferred from processing rather than the origin of coffee seeds.



**Figure 71:** PCA score plot derived from modelling roasted arabica coffee (RCA), roasted canephora (RCC), and 7 roasted commercial samples (BRK, LRCS, LRS, LRCK, LRCM, LRCQ, BRA) (A). PCA score plot derived from modelling roasted commercial samples against each other (B).

## 2.2 Comparison between UPLC-MS and UV fingerprinting multivariate data analysis models

The classification potential of both UV and UPLC/MS was compared based on their PCA and OPLS results, with both techniques found generally comparable and to complement metabolites detection in specimens. The PCA and OPLS loading plots obtained from UPLC/MS revealed that caffeine and chlorogenic acid lactones contributed to the discrimination of roasted samples versus the abundance of chlorogenic acids and diterpenes in unroasted samples. Additionally, dihydroxypsoralen-O-hexoside, which was identified for the first time in coffee detected as a marker in unroasted samples, **Figure 67**. In contrast, in UV models, interferences of caffeine bands from chlorogenic acids could be predicted, albeit with roasted samples to show tight clustering at higher absorption ranges (350-450) attributed from melanoidins absorption range ( $\lambda=420$  nm), and not detected using UPLC/MS, **Figure 70**, (2.1.1). Other chemicals inferred from UV is acrylamide, showing increased absorption in roasted specimens than green



ones. Detection of acrylamide is of value considering its health risk potential, and whether UV can be used for its direct assay in coffee extracts or brews should be examined.

On the other hand, unroasted coffee samples showed higher absorbance in the region between 220-350 nm, typically for chlorogenic acid absorption ( $\lambda=220,325$  nm) and diterpenes ( $\lambda$  298nm), **section2.1.1**. Likewise, the addition of cardamom to coffee was investigated by both techniques; PCA score plot obtained from the UPLC/MS model showed tighter clustering of instant samples than roasted samples, while the UV model could not distinguish clearly between roasted and cardamom blended samples and suggestive that blending effect is better revealed using UPLC/MS. However, UPLC/MS derived model was not able to detect any acrylamide nor melanoidins that are important markers of processing impact on coffee and its health safety level (section0.). UV model loading plot showed that instant samples had strong absorption for acrylamide ( $\lambda$  =275nm), while roasted were higher in absorbance range at 370-420 nm suggesting variation in melanoidins' formation during the roasting technique for both types (section2.1.2). Accordingly, the instant coffee considered less safe than the other types of coffee. Lastly, only UV was able to better correlate between the commercial and authentic samples, distinguishing LRS and LRCQ as the most close to Arabica species poses UV as a tool for the detection of admixtures in commercial samples (**section 2.1.4**), **Figure71**, **Table 5**.

**3 Chapter III: Determination of total phenolics in coffee specimens and in relation to its in vitro antioxidant assays**

### 3.1 Determination of total phenolics (TPC) of coffee species

UPLC/MS analysis revealed for coffee seeds enrichment in phenolic acids and to be affected by roasting [110]. TPC assay was employed as a proximate method for QC determination of total phenolics and to aid in the standardization of the different coffee specimens expressed as mgGAE/g. The highest levels were detected at 50-52 mgGAE/g in BRA, LRCK, GCK, versus lowest in ICC, and ICA at 3-7.7 mgGAE/g **Table 7**. Results showed that increasing the roasting degree led to a decrease in total phenolics, with the highest TPC level observed in lightly roasted samples and green samples, with a marked decline in heavily roasted and instant samples. Results were consistent with previous reports and confirmed the superiority of green and light roasted coffee as a rich source of free polyphenols compared to processed (instant) and roasted coffee [111, 112]. Differences between roasted samples may be attributed to the degradation of chlorogenic acids and their contribution to the development of Maillard reaction products, i.e., melanoidins [113]. Among the examined roasted coffee specimens, instant coffee products, i.e., Arabian coffee blended with cardamom (ICC) and instant *C. arabica* (ICA), was recognized for these lowest levels of phenolics, suggesting their degradation during their further processing steps in case of instant coffee, **Figure 72**.

### 3.2 Antioxidant activity

UPLC/MS, UV and total phenolics assay revealed the enrichment of phenolics in coffee seeds, with qualitative differences among specimens, **Table 6**. Consequently, in-vitro antioxidant assays were attempted to determine whether differences in phenolics would affect the antioxidant potential in these coffee specimens, **Figure 72**.

#### 3.2.1 DPPH assay

All coffee extracts displayed a dose-dependent DPPH radical scavenging activity in the concentration range, i.e., 0.01 – 0.5 µg/mL, with results expressed as IC<sub>50</sub> (µg/mL). The DPPH IC<sub>50</sub> values ranged from 27.3µg/mL in lightly roasted *C. arabica* (BRK) to reach 235.9 µg/mL in the heavily roasted *C. arabica* blended with cardamom, (HRKC) compared to Trolox (IC<sub>50</sub>=12.4 µg/mL). The highest antioxidant capacity in lightly roasted coffee is consistent with the results of Bobková *et al* **Table 7** [114].

The higher IC<sub>50</sub> values in heavily roasted samples, i.e., HRKC (235.9 µg/mL) and LRCQ (187 µg/mL), indicate a low antioxidant power and in correlation with their TPC content **Table 7**.

In contrast, lightly roasted samples including BRK, LRS, LRCS showed lower IC<sub>50</sub> values at 27.3 and 43.2 and 48.6 µg/mL, respectively, suggesting their potential antioxidant power and that phenolics are more crucial than melanoidins for determining antioxidant action. Few roasted samples such as RCC and RCA had IC<sub>50</sub> 74.2 and 103.3 µg/mL, respectively, indicating improvement in the antioxidant activity that may be attributed to the production of melanoidins and in correlation with their roasting index, **Table 5**, [23].

Despite the powerful antioxidant effect of cardamom, it did not improve the antioxidant potential in blended coffee, especially HRKC and ICC, in comparison with their plain counterparts, i.e., BRK and ICA, and suggestive that cardamom blend does not exhibit an improved antioxidant action compared to unblended coffee [115].

### 3.2.2 FRAP assay

The FRAP assay measures the ferric reducing power of specimens on ferric-tripyridyltriazine. Compounds with antioxidant activity can convert the ferric-tripyridyltriazine (orange) complex to ferrous tripyridyltriazine (blue colour) at a maximum absorbance of 593 nm [14, 15].

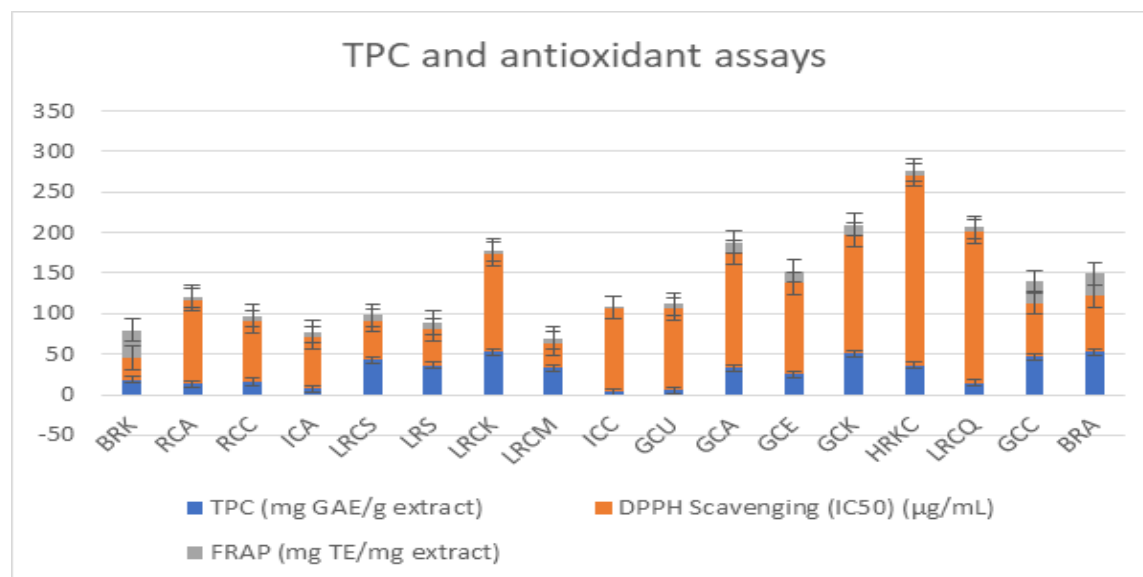
FRAP results were generally in accordance with the DPPH radical scavenging activity to a great extent **Table 7**. Both BRK, BRA, and GCC samples showed the strongest antioxidant effect with FRAP values of 34.1, 28.2, and 26.19 mg TE/mg extract, respectively. In contrast, heavy roasted and instant samples HRKC, RCA, and ICC showed FRAP results at 6.3, 3.9, and 1.5 mg TE/mg extract, respectively. Besides, cardamom addition in the different coffee blend did not result in an increase of FRAP values as in HRKC (6.3 mg TE/mg extract), LRCM (7.9 mg TE/mg extract), and instant coffee products, i.e., ICC at 1.5 mg TE/mg extract., in accordance with DPPH assay.

**Table 7:** Summarized results of TPC, DPPH, FRAP  $\pm$  SD (n=3). The corresponding sample codes are listed in **Table 5**

Coffee samples	TPC (mg GAE/g extract)	DPPH Scavenging (IC <sub>50</sub> ) ( $\mu$ g/mL)	FRAP (mg TE/mg extract)
BRK	18.2 $\pm$ 0.01	27.3 $\pm$ 2.6*	34.1 $\pm$ 0.01*
RCA	13.6 $\pm$ 0.01	103.3 $\pm$ 3.2*	3.9 $\pm$ 0.01*
RCC	15.9 $\pm$ 0.01	74.2 $\pm$ 1.5**	7.1 $\pm$ 0.003*
ICA	7.7 $\pm$ 0.01	62.5 $\pm$ 1.9**	7.9 $\pm$ 0.001*
LRCS	42.8 $\pm$ 0.01	48.6 $\pm$ 1.9*	7.1 $\pm$ 0.002*
LRS	36.5 $\pm$ 0.01	43.4 $\pm$ 1.08*	9.2 $\pm$ 0.001*
LRCK	52.3 $\pm$ 0.01	120.9 $\pm$ 3.1*	5.5 $\pm$ 0.001*
LRCM	32.6 $\pm$ 0.12	30.4 $\pm$ 4.4*	7.9 $\pm$ 0.002*
ICC	3.2 $\pm$ 0.04	104.1 $\pm$ 8.3**	1.5 $\pm$ 0.004*
GCU	15.3 $\pm$ 0.01	100.5 $\pm$ 6.2*	6.3 $\pm$ 0.002*
GCA	33.1 $\pm$ 0.003	142 $\pm$ 1.9*	10.3 $\pm$ 0.001**
GCE	25.3 $\pm$ 0.01	112.1 $\pm$ 1.3*	15.1 $\pm$ 0.003*
GCK	50.9 $\pm$ 0.007	146.1 $\pm$ 1.1*	10.3 $\pm$ 0.001*
HRKC	36.1 $\pm$ 0.03	234.9 $\pm$ 5.8*	6.3 $\pm$ 0.003*
LRCQ	14.7 $\pm$ 0.02	187 $\pm$ 4.1*	5.5 $\pm$ 0.003*
GCC	46.6 $\pm$ 0.004	66.72 $\pm$ 2.8*	26.1 $\pm$ 0.003 *
BRA	52.5 $\pm$ 0.03	69.12 $\pm$ 2.6*	28.2 $\pm$ 0.004*
Gallic acid (mean $\pm$ SD)	100.09 $\pm$ 0.24		
Trolox ( $\mu$ g/mL)		12.4 $\pm$ 1.48*	
Trolox (mean $\pm$ S)			1.12 $\pm$ 0.33*

\*Significant from GCC sample ( $P$  value < 0.05, two-way ANOVA)

\*\*non-significant

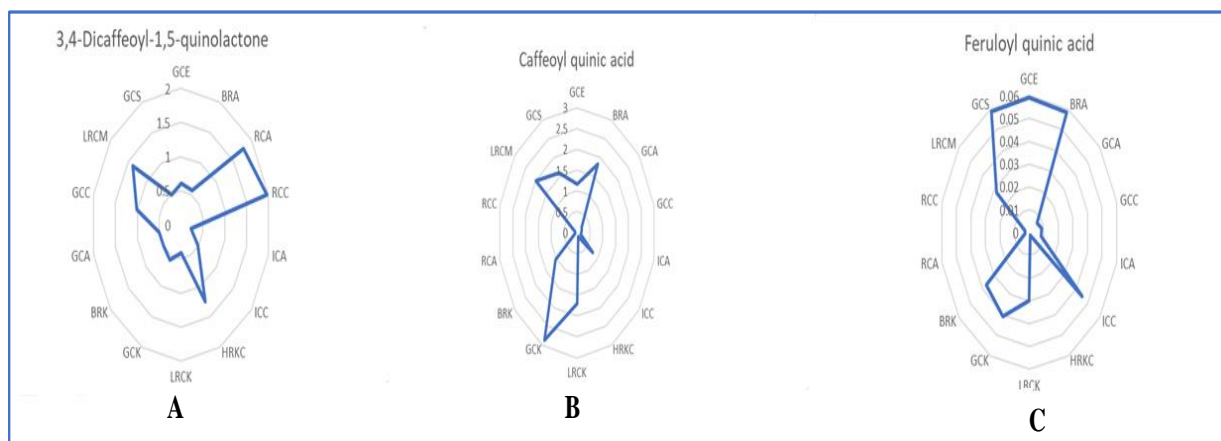


**Figure 72:** Total phenolic content (TPC) of investigated coffee specimen and values are expressed as mg gallic acid equivalent/g extract (mg GAE/g extract), DPPH (2,2-diphenyl-1-picrylhydrazyl) scavenging activity, and values are in IC<sub>50</sub>±SD (µg/mL), and ferric reducing antioxidant power (FRAP) and values are expressed in mg Trolox equivalent per mg extract (mg TE/mg extract). Each bar represents mean±SE (n=3); the corresponding coffee codes are listed in Table 5

### 3.3 Correlation between biological assays and UPLC/MS metabolite profile

In this part, a correlation between bioassay extracts and UPLC/MS data was established to determine the metabolites responsible for these antioxidant activities for different samples. A PLS model was constructed from the data designating the UPLC/MS metabolites as x-variables and the corresponding biological assay parameters (TPC, DPPH, FRAP). The PLS score plot explained 99% of the total variance in Y ( $R^2=0.99$ ) and ( $Q^2=0.92$ ) as a prediction goodness parameter explained 92%, and the loading plot displayed a positive correlation with all the assays. Investigation of the variable importance in the projection (VIP) to recognize the metabolites responsible for the antioxidant effect and pinpoint the relation between the X and Y variables in the pls model. The metabolites which showed a VIP score of more than 1 are significant, and only positive values are represented in the plot. The main potential metabolites that were significant in VIP scores included caffeine, and caffeoylquinic acid at a VIP score of 6.6 & 6.8, respectively. Interestingly, dicaffeoyl quinolactone showed a lower score at 1.9,

suggestive of its lower correlation potential. For further investigation, radar charts were applied to be able to compare the VIP metabolites in our different samples (**Figure 73**).



**Figure 73:** Radar charts created from VIP score of the main metabolites which contributed in antioxidant activity. A: Dicaffeoyl quinolactone showed maximum concentration in roasted samples and light roasted (RCA, RCC, LRCM), respectively. B: Caffeoylquinic acid showed maximum concentration green, and light roasted (BRA, LRCM), respectively. C: Feruloylquinic acid showed maximum concentration in green, and light roasted (GCS, GCE, BRA), respectively. Samples codes are listed in **Table 5**.

## **Chapter IV: Mineral's analysis**



## Mineral analysis

The mineral analysis aimed to identify differences in major elements in coffee specimens and upon roasting. The elemental study of coffee can be considered as an indicator of the authenticity of coffee, despite the small content that does not exceed 5% (m/m). Additionally, minerals can aid to discriminate between coffee geographical origin with different environment and cultivation conditions [116]. Inductively coupled plasma atomic emission spectroscopy (ICP-AES) provides a fully comprehensive, versatile, precise, and rapid analysis of elemental composition. Also, it is characterized by high sensitivity, wider linear dynamic range, and fewer interferences of spectra [117].

To the best of our knowledge, few reports have been conducted using ICP-AES, used previously to investigate the toxic elements in coffee samples from Brazil [92]. Considering the high consumption of coffee, analysis of these elements is of great importance as they can accumulate in the body, major (macro elements) required in large quantities and oppositely the minor (microelements). For heavy metals (Cu, Mn, Zn, Pb), coffee plants absorb and transport them to different parts of plants until reaching the seeds to be the source of contamination and cause serious adverse effects.

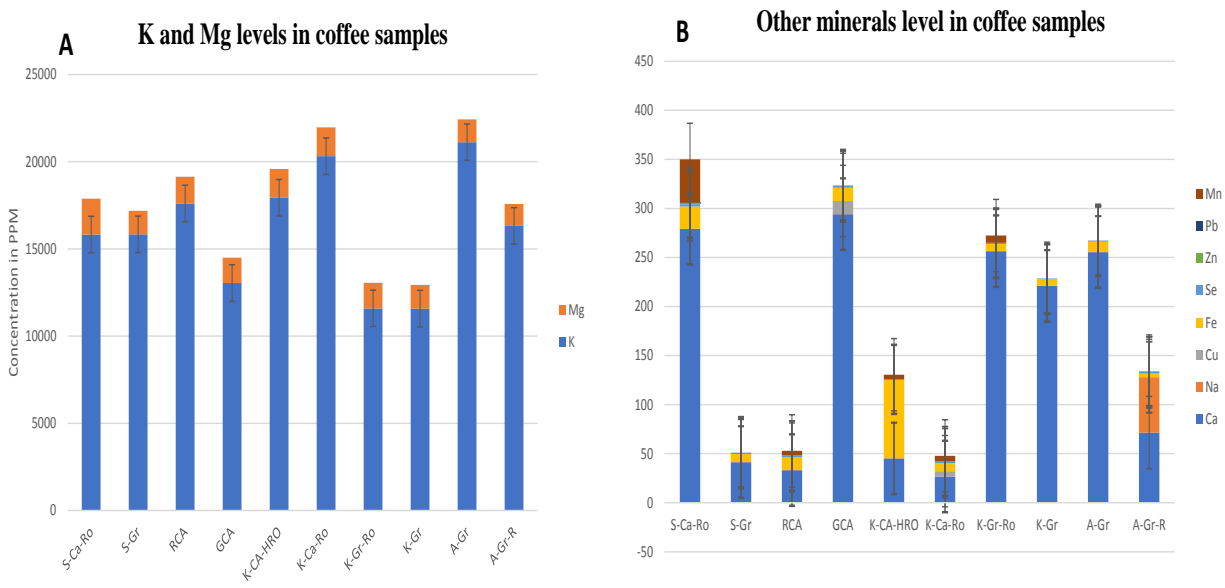
The results for each coffee sample included the determination of macro elements, *i.e.*, Mg, K and Na and microelements, *i.e.*, Zn, Cu, Fe, Ca, Se, Pb, Mn, and expressed in the mean of the three replicates  $\pm$  standard deviation (SD), **Table 8**. K was found to be the most abundant element in all the coffee samples. Next to K, Mg displayed high levels in all samples, especially in cardamom blends such as LRCM, HRKC and LRCK.

Toxic metals such as Pb were not detected in all samples. Moreover, Mn existed in most of the roasted samples and was found absent in green samples. Other minerals were found at comparable levels among all samples (**Table 8, Figure 74**). Authenticated *C. arabica* samples showed that roasting affected Ca levels, found to decrease by 7 folds after roasting in RCA. the same pattern was observed in Egyptian green and roasted coffee samples, *i.e.*, GCE and BRA. Among examined minerals, K showed the highest abundance in all coffee samples, with GCE as the richest also to decrease with roasting as in BRA. However, such a pattern was not observed with potassium levels in roasted sample RCA showing higher values than unroasted GCA. The

obtained results were acceptable for human consumption based on their toxic and nutritional level and agreed with the literature [92].

**Table 8:** Results of mineral analyses in 10 coffee samples. All values in the table are expressed as Mean  $\pm$  SD expressed in ppm.

Samples	Ca	K	Mg	Na	Cu	Fe	Se	Zn	P b	Mn
LRCM	279.2 $\pm$ 2.9	15820 $\pm$ 63 1.5	2060 $\pm$ 0	0.15 $\pm$ 0	0	22.7 $\pm$ 1.9	2.1 $\pm$ 0 .02	1.4 $\pm$ 0.1	0	44.2 $\pm$ 1.3
GCS	41.3 $\pm$ 0 .45	15836.6 $\pm$ 702.3	1353.3 $\pm$ 0	0.16 $\pm$ 0	0	8.1 $\pm$ 0. 28	1.3 $\pm$ 0 .14	0	0	0
RCA	33.1 $\pm$ 1 .9	17610 $\pm$ 20 78.05	1526.6 $\pm$ 5.77	0.17 $\pm$ 0 .0055	0	12.8 $\pm$ 1.40	2.1 $\pm$ 0 .2	0	0	4.5 $\pm$ 0. 07
GCA	294.1 $\pm$ 1.2	13050 $\pm$ 92 4.1	1450 $\pm$ 0	0.14 $\pm$ 0	13.28 $\pm$ 2.52	13.5 $\pm$ 0.1	2.3 $\pm$ 0 .2	0	0	0
HRKC	45.1 $\pm$ 1 5.6	17930 $\pm$ 35 0	1653.3 $\pm$ 5.77	0.14 $\pm$ 0	0	79.7 $\pm$ 3.6	0.9 $\pm$ 0 .13	0	0	4.4 $\pm$ 0. 51
LRCK	26.6 $\pm$ 0 .1	20320 $\pm$ 12 00	1653.3 $\pm$ 15.2	0.16 $\pm$ 0	5.332 3 $\pm$ 0.7 4	8.1 $\pm$ 0. 85	2.2 $\pm$ 0 .1	0	0	5.2 $\pm$ 0. 00
BRK	256.3 $\pm$ 0.9	11596.6 $\pm$ 1413.30	1466.6 $\pm$ 11.5	0.14 $\pm$ 0 .005	0	7.1 $\pm$ 1. 5	1.2 $\pm$ 0 .26	0	0	7.3 $\pm$ 0. 70
GCK	221 $\pm$ 0. 09	11580 $\pm$ 10 59.2	1360 $\pm$ 1 0	0.12 $\pm$ 0	0	6.6 $\pm$ 0. 6	0.8 $\pm$ 0 .09	0	0	0
GCE	255.5 $\pm$ 0.8	21126.6 $\pm$ 974.3	1300 $\pm$ 0	0.19 $\pm$ 0	0	10.4 $\pm$ 0.2	1.2 $\pm$ 0 .02	0	0	0
BRA	71.513 3 $\pm$ 1.08	16326.6 $\pm$ 445	1260 $\pm$ 0	0.17 $\pm$ 0	0	3.5 $\pm$ 0. 4	2.5 $\pm$ 0 .1	0	0	0



**Figure 74:** Mineral's level expressed as ppm in coffee samples, A: K and Mg levels in coffee samples B: Other minerals level in coffee samples (Mn, Pb, Zn, Se, Fe, Cu, Na, and Ca)

## Conclusion

This study represents a multiplex approach that compares two different platforms, i.e., UPLC/MS and UV fingerprinting techniques, as an analytical tool for the identification of secondary metabolites in different coffee species. Specimens represented different variables such as roasting process, geographical origin, and or additives. Both UPLC/MS and UV spectroscopy coupled to multivariate data analysis revealed for differences among authenticated and commercial samples from the Middle East. Such comparative metabolomics approach presents the first detailed profile for green and roasted coffee metabolome. Molecular Networking (GNPS) aided in the identification of metabolites *via* UPLC/MS data analysis and provided annotation of 91 metabolites belonging to different classes, including new chlorogenic acids, diterpenes, fatty acids, and sphingolipids. On the other side, UV fingerprinting provided preliminary data on the absorption ranges of different main chemicals, posing its usage as an alternative tool for UPLC/MS as cheaper and simpler to operate. Interestingly, both techniques were generally comparable towards metabolites detection in specimens with more advantages for UV in the identification of acrylamide and melanoidins. *In-vitro* antioxidant assays provided a measure of how antioxidant activity of different roasted and green coffee samples correlates with difference in metabolites composition. Mineral macro elements and toxic elements were determined using (ICP-AES) revealing an increase in Mg level in cardamom blended samples and enrichment of K in all the samples.

The developed comparative metabolomics approach comprising UV and UPLC/MS reveal differences in secondary metabolite profile in different coffee specimens and identify several markers to be considered in future quality control purposes as well to be considered using other spectroscopic techniques.

## **Recommendation**

The main obstacle in studying metabolomics remains in investigating the extremely complex and enormous diversity of natural plants' metabolites besides their annotations. Besides, limited databases and lack of spectral share in public is another obstacle to be tackled by providing a bioinformatics tool that connects the spectral dataset with the biological activity and can shed light on the underlying active agents. This approach can be applied for the investigation of other coffee products, i.e., coffee silver skin, as it had been neglected as a coffee byproduct of potential as a natural fertilizer, fuel substitute, and a rich source of nutrients. Although coffee has been the main target for phytochemists, some species have not been investigated, such as wild coffee species that need further studies in the future. Moreover, few studies were done on coffee leaves which are considered to be used as an alternative to tea leaves as a healthy beverage owing to their richness in carotenoids, and mangiferin which possess great antioxidant properties and many pharmacological and therapeutic effects.

## References

1. Mahmud, M.M.C., R.A. Shellie, and R. Keast, *Unravelling the relationship between aroma compounds and consumer acceptance: Coffee as an example*. Comprehensive Reviews in Food Science and Food Safety, 2020. 19(5): p. 2380-2420.
2. Samper, L., D. Giovannucci, and L. Vieira, *The powerful role of intangibles in the coffee value chain*, ed. W.I.P.O. (WIPO). Vol. 39. 2017, Geneva, Switzerland.
3. Shokouh, P., et al., *Efficacy of arabica versus robusta coffee in improving weight, insulin resistance, and liver steatosis in a rat model of type-2 diabetes*. Nutrients, 2019. 11(9): p. 2074.
4. De Rosso, M., et al., *UHPLC-ESI-QqTOF-MS/MS characterization of minor chlorogenic acids in roasted Coffea arabica from different geographical origin*. Journal of Mass Spectrometry, 2018. 53(9): p. 763-771.
5. Acidri, R., et al., *Phytochemical profile and antioxidant capacity of coffee plant organs compared to green and roasted coffee beans*. Antioxidants, 2020. 9(2): p. 93.
6. Al Doghaither, H.A. and E. Al-Malki, *The addition of herbal additives influences the antioxidant activity of traditional arabic coffee*. World Applied Sciences Journal, 2017. 35(3): p. 393-398.
7. !!! INVALID CITATION !!!
8. Cavin, C., et al., *Cafestol and kahweol, two coffee specific diterpenes with anticarcinogenic activity*. Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association, 2002. 40(8): p. 1155-1163.
9. Tamanna, N. and N. Mahmood, *Food processing and Maillard reaction products: Effect on human health and nutrition*. International Journal of Food Science, 2015. 2015: p. 526762.
10. Zhang, H., et al., *Melanoidins from coffee, cocoa, and bread are able to scavenge  $\alpha$ -dicarbonyl compounds under simulated physiological conditions* Journal of Agricultural and Food Chemistry, 2019. 67(39): p. 10921-10929.
11. Choi, S., S. Jung, and K.S. Ko, *Effects of coffee extracts with different roasting degrees on antioxidant and anti-Inflammatory systems in mice*. Nutrients, 2018. 10(3): p. 363.
12. Moreira, A., et al., *Coffee melanoidins: Structures, mechanisms of formation and potential health impacts*. Food & Function, 2012. 3: p. 903-15.
13. Marek, G., et al., *Detection and differentiation of volatile compound profiles in roasted coffee arabica beans from different countries using an electronic nose and GC-MS*. Sensors, 2020. 20(7): p. 2124.
14. Liang, N. and D.D. Kitts, *Antioxidant property of coffee components: Assessment of methods that define mechanisms of action*. Molecules, 2014. 19(11): p. 19180-19208.
15. Gorjanović, S., et al., *Antioxidant efficiency of polyphenols from coffee and coffee substitutes-electrochemical versus spectrophotometric approach*. Journal of Food Science and Technology, 2017. 54(8): p. 2324-2331.
16. Sonam, K.S. and S. Guleria, *Synergistic antioxidant activity of natural products*. Annals of Pharmacology and Pharmaceutics, 2017. 2(8): p. 1086.
17. Yashin, A., et al., *Antioxidant and antiradical activity of coffee*. Antioxidants 2013. 2(4): p. 230-245.

18. Otify, A.M., et al., *Metabolites profiling of date palm (Phoenix dactylifera L.) commercial by-products (pits and pollen) in relation to its antioxidant effect: a multiplex approach of MS and NMR metabolomics*. *Metabolomics*, 2019. 15(9): p. 119.
19. Núñez, N., et al., *Authentication of the origin, variety and roasting degree of coffee samples by non-targeted HPLC-UV fingerprinting and chemometrics. Application to the detection and quantitation of adulterated coffee samples*. *Foods*, 2020. 9(3): p. 378.
20. Toci, A.T., et al., *Coffee adulteration: More than two decades of research*. *Critical Reviews in Analytical Chemistry*, 2016. 46(2): p. 83-92.
21. Nunez, N., et al., *High-performance liquid chromatography with fluorescence detection fingerprints as chemical descriptors to authenticate the origin, variety and roasting degree of coffee by multivariate chemometric methods*. *Journal of the Science Food and Agriculture*, 2020.
22. Terrile, A.E., et al., *Chemometric analysis of UV characteristic profile and infrared fingerprint variations of Coffea arabica green beans under different space management treatments*. *Journal of the Brazilian Chemical Society*, 2016. 27(7): p. 1254-1263.
23. Bobková, A., et al., *The effect of roasting on the total polyphenols and antioxidant activity of coffee*. *Journal of Environmental Science and Health, Part B*, 2020. 55(5): p. 495-500.
24. Wang, M., et al., *Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking*. *Nat Biotechnol*, 2016. 34(8): p. 828-837.
25. Clifford, M.N., et al., *Chlorogenic acids and the acyl-quinic acids: discovery, biosynthesis, bioavailability and bioactivity*. *Natural Product Reports*, 2017. 34(12): p. 1391-1421.
26. Ss, E.-H., et al. *Metabolomic Profiling of Five Agave Leaf Taxa via UHPLC/PDA/ESI-MS Inrelation to Their Anti-Inflammatory, Immunomodulatory and Ulceroprotective Activities*. *Steroids* 2020 04/13/2020.
27. Masike, K., et al., *Highlighting mass spectrometric fragmentation differences and similarities between hydroxycinnamoyl-quinic acids and hydroxycinnamoyl-isocitric acids*. *Chemistry Central Journal*, 2017. 11(1): p. 29.
28. Asamenew, G., et al., *Comprehensive characterization of hydroxycinnamoyl derivatives in green and roasted coffee beans: A new group of methyl hydroxycinnamoyl quinate*. *Food Chemistry: X*, 2019. 2: p. 100033.
29. Jaiswal, R., et al., *Profile and characterization of the chlorogenic acids in green Robusta coffee beans by LC-MS(n): identification of seven new classes of compounds*. *Journal of Agricultural and Food Chemistry*, 2010. 58(15): p. 8722-8737.
30. Stalmach, A., et al., *On-line HPLC analysis of the antioxidant activity of phenolic compounds in brewed, paper-filtered coffee*. *Brazilian Journal of Plant Physiology*, 2006. 18(1): p. 253-262.
31. Panusa, A., et al., *UHPLC-PDA-ESI-TOF/MS metabolic profiling and antioxidant capacity of arabica and robusta coffee silverskin: Antioxidants vs phytotoxins*. *Food Research International (Ottawa, Ont.)*, 2017. 99(Pt 1): p. 155-165.
32. Berti, F., et al., *Hydroxycinnamoyl Amino Acids Conjugates: A Chiral Pool to Distinguish Commercially Exploited Coffea spp*. *Molecules*, 2020. 25(7): p. 1704.

33. Nuhu, A., *Bioactive Micronutrients in Coffee: Recent Analytical Approaches for Characterization and Quantification*. ISRN Nutrition, 2014. 2014: p. 1-13.
34. Patay, É.B., T. Bencsik, and N. Papp, *Phytochemical overview and medicinal importance of Coffea species from the past until now*. Asian Pacific Journal of Tropical Medicine, 2016. 9(12): p. 1127-1135.
35. Yashin, A., et al., *Chromatographic Methods for Coffee Analysis: A Review*. Journal of Food Research, 2017. 6(4): p. p60.
36. Islam, M.T., et al., *An Insight into the Therapeutic Potential of Major Coffee Components*. Current Drug Metabolism, 2018. 19(6): p. 544-556.
37. Higdon, J.V. and B. Frei, *Coffee and health: a review of recent human research*. Critical Reviews in Food Science and Nutrition, 2006. 46(2): p. 101-123.
38. Wanyika, H., et al., *Determination of caffeine content of tea and instant coffee brands found in the Kenyan market*. African Journal of Food Science, 2010. 4: p. 353-358.
39. Kurzrock, T. and K. Speer, *Identification of kahweol fatty acid esters in Arabica coffee by means of LC/MS*. Journal of Separation Science, 2001. 24(10- 11): p. 843-848.
40. Liu, M., et al., *Diterpenoids from Isodon species: an update*. Natural Product Reports, 2017. 34(9): p. 1090-1140.
41. Speer, K. and I. Kölling-Speer, *The lipid fraction of coffee bean*. Brazilian Journal of Plant Physiology, 2006. 18.
42. Kurzrock, T. and K. Speer, *Identification of kahweol fatty acid esters in Arabica coffee by means of LC/MS*. Journal of Separation Science - J SEP SCI, 2001. 24: p. 843-848.
43. Scherer, M., et al., *A rapid and quantitative LC-MS/MS method to profile sphingolipids*. Journal of Lipid Research, 2010. 51(7): p. 2001-2011.
44. Wei, F., et al., *<sup>13</sup>C NMR-Based Metabolomics for the Classification of Green Coffee Beans According to Variety and Origin*. Journal of Agricultural and Food Chemistry, 2012. 60(40): p. 10118-10125.
45. Clifford, M.N., W. Zheng, and N. Kuhnert, *Profiling the chlorogenic acids of aster by HPLC–MSn*. Phytochemical Analysis, 2006. 17(6): p. 384-393.
46. Gebrewold, F., *Chemistry of Caffeine in Coffee and Its Determination Using UV/Vis Spectrophotometer: A Review Article*. Chemistry and Materials Research, 2018. 10(1): p. 23-27.
47. Arya, M. and L.J.M. Rao, *An impression of coffee carbohydrates*. Critical Reviews in Food Science and Nutrition, 2007. 47(1): p. 51-67.
48. Murata, M., H. Okada, and S. Homma, *Hydroxycinnamic Acid Derivatives and p - Coumaroyl-(L)-tryptophan, A Novel Hydroxycinnamic Acid Derivative, from Coffee Beans*. Bioscience, Biotechnology, and Biochemistry, 1995. 59(10): p. 1887-1890.
49. Baggenstoss, J., *Coffee roasting and quenching technology - formation and stability of aroma compounds*, 2008, ETH Zurich.
50. Lang, R. and T. Hofmann, *A versatile method for the quantitative determination of  $\beta$  N-alkanoyl-5-hydroxytryptamides in roasted coffee*. European Food Research and Technology, 2005. 220(6): p. 638-643.
51. Craig, A.P., A.S. Franca, and L.S. Oliveira, *Discrimination between defective and non-defective roasted coffees by diffuse reflectance infrared Fourier transform spectroscopy*. LWT, 2012. 47(2): p. 505-511.



52. Ribeiro, M., et al., *Authenticity of Roasted Coffee using <sup>1</sup>H NMR Spectroscopy*. Journal of Food Composition and Analysis, 2017. 57: p. 24-30.
53. Worley, B. and R. Powers, *PCA as a practical indicator of OPLS-DA model reliability*. Current Metabolomics, 2016. 4(2): p. 97-103.
54. Kurniawan, M.F., et al., *Metabolomic approach for understanding phenolic compounds and melanoidin roles on antioxidant activity of Indonesia robusta and arabica coffee extracts*. Food Science and Biotechnology, 2017. 26(6): p. 1475-1480.
55. Pérez-Míguez, R., et al., *A non-targeted metabolomic approach based on reversed-phase liquid chromatography–mass spectrometry to evaluate coffee roasting process*. Analytical and Bioanalytical Chemistry, 2018. 410(30): p. 7859-7870.
56. Jaiswal, R., et al., *Does roasted coffee contain chlorogenic acid lactones or/and cinnamoylshikimate esters?* Food Research International, 2014. 61: p. 214-227.
57. Asamenew, G., et al., *Comprehensive characterization of hydroxycinnamoyl derivatives in green and roasted coffee beans: A new group of methyl hydroxycinnamoyl quinate*. Food Chemistry X, 2019. 2: p. 100033.
58. Andrzejewski, D., et al., *Analysis of Coffee for the Presence of Acrylamide by LC-MS/MS*. Journal of Agricultural and Food Chemistry, 2004. 52(7): p. 1996-2002.
59. Clifford, M.N., *Chlorogenic acids and other cinnamates – nature, occurrence and dietary burden*. Journal of the Science of Food and Agriculture, 1999. 79(3): p. 362-372.
60. Liang, N. and D.D. Kitts, *Antioxidant property of coffee components: assessment of methods that define mechanisms of action*. Molecules (Basel, Switzerland), 2014. 19(11): p. 19180-19208.
61. Pokorná, J., et al., *Comparison of different methods of antioxidant activity evaluation of green and roast C. Arabica and C. Robusta coffee beans*. Acta Alimentaria, 2015. 44(3): p. 454-460.
62. Gómez-Ruiz, J.A., D.S. Leake, and J.M. Ames, *In vitro antioxidant activity of coffee compounds and their metabolites*. Journal of Agricultural and Food Chemistry, 2007. 55(17): p. 6962-6969.
63. Mf, K., et al. *Metabolomic approach for understanding phenolic compounds and melanoidin roles on antioxidant activity of Indonesia robusta and arabica coffee extracts*. Food science and biotechnology 2017 12/13/2017.
64. Richelle, M., I. Tavazzi, and E. Offord, *Comparison of the antioxidant activity of commonly consumed polyphenolic beverages (coffee, cocoa, and tea) prepared per cup serving*. Journal of Agricultural and Food Chemistry, 2001. 49(7): p. 3438-3442.
65. Ullah, F., et al., *Caffeine prevents d-galactose-induced cognitive deficits, oxidative stress, neuroinflammation and neurodegeneration in the adult rat brain*. Neurochemistry International, 2015. 90: p. 114-124.
66. Hwang, J.-H., et al., *Caffeine prevents LPS-induced inflammatory responses in RAW264.7 cells and zebrafish*. Chemico-Biological Interactions, 2016. 248: p. 1-7.
67. Yang, C.-C. and I.-M. Jou, *Caffeine treatment aggravates secondary degeneration after spinal cord injury*. Brain Research, 2016. 1634: p. 75-82.
68. Shin, H., et al., *Anti-inflammatory effect of chlorogenic acid on the IL-8 production in Caco-2 cells and the dextran sulphate sodium-induced colitis symptoms in C57BL/6 mice*. Food Chemistry, 2015. 168: p. 167–175.

69. Zhao, J., et al., *Ferulic acid enhances nitric oxide production through up-regulation of argininosuccinate synthase in inflammatory human endothelial cells*. Life Sciences, 2015. 145.
70. Doss, H.M., et al., *Targeting inflammatory mediators with ferulic acid, a dietary polyphenol, for the suppression of monosodium urate crystal-induced inflammation in rats*. Life Sciences, 2016. 148: p. 201-210.
71. Guzman, J.D., *Natural Cinnamic Acids, Synthetic Derivatives and Hybrids with Antimicrobial Activity*. Molecules, 2014. 19(12): p. 19292-19349.
72. Zeni, A.L.B., et al., *Ferulic acid exerts antidepressant-like effect in the tail suspension test in mice: Evidence for the involvement of the serotonergic system*. European Journal of Pharmacology, 2012. 679(1): p. 68-74.
73. Sh, J., et al. *Coffee consumption and serum lipids: a meta-analysis of randomized controlled clinical trials*. American journal of epidemiology 2001 02/15/2001.
74. Urgert, R. and M.B. Katan, *The cholesterol-raising factor from coffee beans*. Annual Review of Nutrition, 1997. 17: p. 305-324.
75. Roos, B.D., et al., *Consumption of French-press coffee raises cholesteryl ester transfer protein activity levels before LDL cholesterol in normolipidaemic subjects*. Journal of Internal Medicine, 2000. 248(3): p. 211-216.
76. Tuomilehto, J., *Coffee Consumption and Risk of Type 2 Diabetes Mellitus Among Middle-aged Finnish Men and Women*. JAMA, 2004. 291(10): p. 1213.
77. van Dam, R.M. and F.B. Hu, *Coffee consumption and risk of type 2 diabetes: a systematic review*. JAMA, 2005. 294(1): p. 97-104.
78. Saremi, A., M. Tulloch-Reid, and W.C. Knowler, *Coffee consumption and the incidence of type 2 diabetes*. Diabetes Care, 2003. 26(7): p. 2211-2212.
79. Fang, C.-Y., et al., *Caffeine is responsible for the bloodglucose-lowering effects of green tea and Puer tea extracts in BALB/c mice*. Chinese Journal of Natural Medicines, 2015. 13(8): p. 595-601.
80. Sacramento, J.F., et al., *Disclosing caffeine action on insulin sensitivity: effects on rat skeletal muscle*. European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences, 2015. 70: p. 107-116.
81. Naowaboot, J., et al., *Ferulic acid improves lipid and glucose homeostasis in high-fat diet-induced obese mice*. Clinical and Experimental Pharmacology & Physiology, 2016. 43(2): p. 242-250.
82. Giovannucci, E., *Meta-analysis of coffee consumption and risk of colorectal cancer*. American Journal of Epidemiology, 1998. 147(11): p. 1043-1052.
83. Michels, K.B., et al., *Coffee, tea, and caffeine consumption and incidence of colon and rectal cancer*. Journal of the National Cancer Institute, 2005. 97(4): p. 282-292.
84. Huber, W.W., et al., *Chemoprotection against the formation of colon DNA adducts from the food-borne carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) in the rat*. Mutation Research, 1997. 376(1-2): p. 115-122.
85. Huber, W.W., et al., *Enhancement of the chemoprotective enzymes glucuronosyl transferase and glutathione transferase in specific organs of the rat by the coffee components kahweol and cafestol*. Archives of Toxicology, 2002. 76(4): p. 209-217.
86. Jang, H., et al., *Chlorogenic acid and coffee prevent hypoxia-induced retinal degeneration*. Journal of Agricultural and Food Chemistry, 2014. 62(1): p. 182-191.

87. Y, L., et al. *Chlorogenic acid prevents isoproterenol-induced hypertrophy in neonatal rat myocytes*. Toxicology letters 2014 05/02/2014.
88. R, J., et al. *Chlorogenic acid improves ex vivo vessel function and protects endothelial cells against HOCl-induced oxidative damage, via increased production of nitric oxide and induction of Hmox-1*. The Journal of nutritional biochemistry 2016 2016 Jan.
89. Hong, Q., et al., *Antithrombotic activities of ferulic acid via intracellular cyclic nucleotide signaling*. European Journal of Pharmacology, 2016. 777: p. 1-8.
90. Terpstra, A.H.M., et al., *The hypercholesterolemic effect of cafestol in coffee oil in gerbils and rats*. Journal of Nutritional Biochemistry, 2000.
91. Fernández-Poyatos, M.D.P., et al., *Phenolic characterization, antioxidant activity, and enzyme inhibitory properties of Berberis thunbergii DC. leaves: A valuable source of phenolic acids*. Molecules, 2019. 24(22).
92. Santos, É.J.d. and E. de Oliveira, *Determination of mineral nutrients and toxic elements in Brazilian soluble coffee by ICP-AES*. Journal of Food Composition and Analysis, 2001. 14(5): p. 523-531.
93. Heo, J., et al., *Analysis of Caffeine, Chlorogenic Acid, Trigonelline, and Volatile Compounds in Cold Brew Coffee Using High-Performance Liquid Chromatography and Solid-Phase Microextraction-Gas Chromatography-Mass Spectrometry*. Foods, 2020. 9(12).
94. Bianco, G., et al., *Identification and fragmentation pathways of caffeine metabolites in urine samples via liquid chromatography with positive electrospray ionization coupled to a hybrid quadrupole linear ion trap (LTQ) and Fourier transform ion cyclotron resonance mass spectrometry and tandem mass spectrometry*. Rapid Commun Mass Spectrom, 2009. 23(7): p. 1065-74.
95. Fang, N., S. Yu, and R.L. Prior, *LC/MS/MS characterization of phenolic constituents in dried plums*. Journal of Agricultural and Food Chemistry, 2002. 50(12): p. 3579-3585.
96. Clifford, M.N., et al., *Characterization by LC-MS<sup>n</sup> of four new classes of chlorogenic acids in green coffee beans: Dimethoxycinnamoylquinic acids, diferuloylquinic acids, caffeoyl-dimethoxycinnamoylquinic acids, and feruloyl-dimethoxycinnamoylquinic acids*. Journal of Agricultural and Food Chemistry, 2006. 54(6): p. 1957-1969.
97. Farah, A., et al., *Effect of roasting on the formation of chlorogenic acid lactones in coffee*. Journal of Agricultural and Food Chemistry, 2005. 53(5): p. 1505-1513.
98. Panusa, A., et al., *UHPLC-PDA-ESI-TOF/MS metabolic profiling and antioxidant capacity of arabica and robusta coffee silverskin: Antioxidants vs phytotoxins*. Food Research International, 2017. 99(Pt 1): p. 155-165.
99. Lang, R., et al., *Mozambioside is an arabica-specific bitter-tasting furokaurane glucoside in coffee beans*. Journal of Agricultural and Food Chemistry, 2015. 63(48): p. 10492-10499.
100. Liu, M., et al., *Diterpenoids from Isodon species: An update*. Natural Product Reports, 2017. 34(9): p. 1090-1140.
101. El-Hawary, S.S., et al., *Metabolomic profiling of five Agave leaf taxa via UHPLC/PDA/ESI-MS in relation to their anti-inflammatory, immunomodulatory and ulceroprotective activities*. Steroids, 2020. 160: p. 108648.
102. Speer, K. and I. Kölling-Speer, *The lipid fraction of the coffee bean*. Brazilian Journal of Plant Physiology, 2006. 18(1): p. 201-216.

103. Lang, R. and T. Hofmann, *A versatile method for the quantitative determination of  $\beta$  N-alkanoyl-5-hydroxytryptamides in roasted coffee*. European Food Research and Technology, 2005. 220(5-6): p. 638-643.
104. Farag, M.A., et al., *Metabolite profiling of three Opuntia ficus-indica fruit cultivars using UPLC-QTOF-MS in relation to their antioxidant potential*. Food Bioscience, 2020. 36: p. 100673.
105. Sun, X., et al., *New Sucrose Phenylpropanoid Esters from Polygonum perfoliatum*. Journal of Natural Products, 2000. 63(8): p. 1094-1097.
106. *Chemoprotective Effect of Elettaria Cardamomum against Chemically induced Hepatocellular Carcinoma in Rats by Inhibiting NF- $\kappa$ B, Oxidative Stress, and Activity of Ornithine Decarboxylase*
107. *Characterization and determination of chlorogenicacids (CGA) in coffee beans by UV-Vis spectroscopy* - بحث Google.
108. Terrile, A.E., et al., *Chemometric Analysis of UV Characteristic Profile and Infrared Fingerprint Variations of Coffea arabica Green Beans under Different Space Management Treatments*. Journal of the Brazilian Chemical Society, 2016. 27(7): p. 1254-1263.
109. Cheah, W.L. and M. Fang, *HPLC-based chemometric analysis for coffee adulteration*. Foods, 2020. 9(7): p. 880.
110. Górecki, M. and E. Hallmann, *The antioxidant content of coffee and its in vitro activity as an effect of its production method and roasting and brewing time*. Antioxidants 2020. 9(4): p. 308.
111. Odžaković, B., et al., *Effect of roasting degree on the antioxidant activity of different Arabica coffee quality classes*. Acta Scientiarum Polonorum Technologia Alimentaria, 2016. 15(4): p. 409-417.
112. Król, K., et al., *The content of polyphenols in coffee beans as roasting, origin and storage effect*. European Food Research and Technology, 2020. 246(1): p. 33-39.
113. Corso, M.P., J.A. Vignoli, and M.d.T. Benassi, *Development of an instant coffee enriched with chlorogenic acids*. Journal of Food Science and Technology, 2016. 53(3): p. 1380-1388.
114. Elisha, I.L., et al., *The anti-arthritic, anti-inflammatory, antioxidant activity and relationships with total phenolics and total flavonoids of nine South African plants used traditionally to treat arthritis*. BMC Complementary and Alternative Medicine, 2016. 16(1): p. 307.
115. Abu-Taweel, G.M., *Cardamom (Elettaria cardamomum) perinatal exposure effects on the development, behavior and biochemical parameters in mice offspring*. Saudi Journal of Biological Sciences, 2018. 25(1): p. 186-193.
116. Pohl, P., et al., *Determination of the elemental composition of coffee using instrumental methods*. Food Analytical Methods, 2013. 6(2): p. 598-613.
117. Samsonowicz, M., et al., *Antioxidant properties of coffee substitutes rich in polyphenols and minerals*. Food Chemistry, 2019. 278: p. 101-109.