Investigation of biochemical strategies leading to metabolome complexity of two closely related Coccinia species through LC-QTOF-MS-based metabolite fingerprinting

By

Nengovhela Ndamulelo

(11641037)

Dissertation

Submitted in fulfilment of the requirements

for the degree of

***Magister Scientiae* (MSc)**

In

Biochemistry

In the Faculty of Science

at the

University of Venda, South Africa

Supervisor: Dr NE Madala

Co-supervisors: Dr MI Mhlongo & Mr LM Mathomu

April 2021

Table of Contents

Contents

Declaration

I, **Nengovhela Ndamulelo**, declare that this thesis submitted to the University of Venda for the Master of Science degree in Biochemistry under the School of Mathematical and Natural Sciences has not been submitted to any other University. I can further declare that the work presented in this dissertation in original, with the exception of referenced material which has been properly acknowledged and cited in text.

Signature  Date 27/04/2021

Dedication

I dedicate this work to my late father, Mr Ramaano Richard Nengovhela. Your words of encouragement and prayers for me to keep moving forward with education are coming into fruition. May his soul rest in perfect peace

Acknowledgement

I would like to give thanks to my supervisor, Dr Ntakadzeni Edwin Madala, for giving me an opportunity to learn under him and for grooming me, his knowledge and support throughout my masters years. Most of all thank you for not giving up on me and believing that I could do more than what I thought I was capable of. Thank you for always being honest and straightforward in every aspect with both life and academics.

I would also like to give thanks to my co-supervisors for their mentorship and guidance throughout this dissertation.

I also thank Dr Khathutshelo Magwede for helping with the authentication of the plants at the University of Venda herbarium.

The University of Venda Research and Publication Committee (RPC) fund is also thanked for financial support.

To my family and friends, thank you for all the support and prayers, these helped keep me going throughout this project.

Lastly, I would like to thank God Almighty for seeing me through the whole project

Preface

This document has been compiled in a form of published and manuscripts under preparation.

The outline is as follows:

Chapter 1: General Introduction

Chapter 2: Literature Review

Articles published

Nengovhela, N., Steenkamp, P.A. and Madala, N.E., 2020. LC-MS based metabolite fingerprinting of *Coccinia* plants reveals glyco-isomerization as a structural diversification strategy in flavonoid chemical space. Natl. Acad. Sci. Lett. (2020). https://doi.org/10.1007/s40009-020-00990-4.

Presented in **Chapter 3**.

Nengovhela, N., Mbedzi, D.T., Ndhlala, A.R., Mathomu, L. M., Mhlongo, M.I., Madala, N.E. 2021. LC-MS based metabolite profiling reveals hydroxylcinnamoyl conjugation as a discriminatory chemical factor between two closely related *Coccinia* species. Submitted to South African Journal of Botany.

Presented in **Chapter 4**

Manuscript under review

Ndamulelo Nengovhela & Ntakadzeni Edwin Madala. 2021. May the real rutin please stand up? Highlighting the analytical and biological consequences of structural isomerism of plant metabolites.

Presented in **Chapter 5**

Chapter 6: General conclusion

Executive Summary

Medicinal plants play an important role in both health and the livelihood of people. These plants have been used for medicinal purposes since time memorial, and they are deemed safer than their synthetic counterparts. One such plant is *Coccinia grandis (*Ivy Gourd*),* which has been used in traditional medicine as a household remedy for various diseases. *Coccinia grandis* have been reported to possess antidiabetic properties. This is owing to the presence of secondary metabolites which are naturally produced by plants for adaptation to their immediate environment. Secondary metabolites are a diverse group of chemical compounds which include flavonoids, glycosides and hydroxycinnamic acids. These biologically active compounds chemically and structurally complex and their characterization pose undisputed analytical challenges.

Mass spectrometry techniques have however, over the years been developed as feasible technique to assist in the characterisation and structural elucidation of plant metabolites. Liquid Chromatography Mass Spectrometry (LC-MS) is the most widely and preferred analytical platform due to its high sensitivity, resolution and detection specificity. Moreover, MS is able to produce accurate mass, unfragmented and fragmented data, which is important for compound characterization and structural elucidation. As such, in the current study LC-MS based metabolomics approach was used to screen for phytochemicals from commercially cultivated *C. grandis* and its widely growing relative, *Coccinia rehmannii* which has no reported medical use in literature. To establish the chemo-taxonomic relationship between these two species, multivariate statistical modelling revealed differential metabolites distribution pattern between them.

The results revealed that these two closely related plant species possess a wide variety of important secondary metabolites such as flavonoids and hydroxycinnamic acids. Some interesting subtle differences were also noted, for instance the flavonoids composition of these plants were to have different glycosylation patterns (**Chapter 3**). Here, it was noted that flavonoids attached to *di-* or *tri-* saccharides are prone to isomerization, a phenomenon termed glyco-isomerization. Apart from positional migration of sugar moieties, acylation of sugars by other biologically active molecules such as derivatives of cinnamic acids (i.e. caffeic and coumaric acid) was also noted

as a contributor towards glyco-isomerization. Here, *C. rehmannii* was found to attach these cinnamic acid derivatives to its flavonoids, a phenomenon which was absent in *C. grandis.* It was then concluded that *C. rehmannii* produce more flavonoid molecules than *C. grandis.*

Apart from flavonoids composition, hydroxycinnamoyl conjugation was revealed to be a discriminatory chemical factor between these two closely related *Coccinia* species. Through UHPLC-qTOF-MS/MS metabolite fingerprinting, *C. grandis* was found to produce different chlorogenic acids, thus cinnamic acids conjugated to a quinic acid (**Chapter 4**). Ironically, only few chlorogenic acids were found in *C. rehmannii*. The findings revealed that the two closely related species utilise Hydroxycinnamic acid (HCA) conjugations as another strategy to diversify their metabolite composition, a phenomenon which has not been suggested in chemo taxonomical studies.

Furthermore, the current study revealed that these two plants use both glyco-isomerization and conjugation as an evolutionary strategy to maximise their metabolome complexity. This was further highlighted in (**Chapter 5)** where LC-MS analyses of *C. grandis* revealed multiple peaks sharing same precursor ion (with molecular formula $C_{27}H_{30}O_{16}$) and associated fragmentation patterns. Further analyses revealed all six peaks to be Rutin, a most common and highly pharmacologically acclaimed flavonoid molecule. As demonstrated herein, these results also had analytical implications, such that subsequent developments of other specific methods such as multiple reaction monitoring (MRM) could not distinguish these compounds.

In conclusion, the two *Coccinia* species contain a wide range of phytochemicals, dominated by flavonoids and hydroxycinnamic acids derivatives. This study further demonstrated LC-MS as an effective tool in comparing metabolite profiles between *C. rehmannii* and *C. grandis* certified by the varying amount and composition of flavonoids and hydroxycinnamic acids derivatives between the two species. Two biochemical phenomenon known as glyco-isomerization and conjugation were shown to be well co-ordinated strategies used by these plants to diversify their metabolite composition. LC-MS in combination with multivariate data models was effective in the investigation of metabolites distribution patterns between the two species.

List of Abbreviations and units

BA	Benzoic acid
BPI	Base peak intensity
C4H	Cinnamate-4-hydroxyase
4CL	4-coumaroyl-CoA ligase
CA	Cinnamic acid
CGA	Chlorogenic acid
CID	Collision induced dissociation
CG	*Coccinia grandis*
CR	*Coccinia rehmannii*
CQA	Caffeoylquinic acid
DDA	Data dependent acquisition
ESI	Electrospray ionization
EI	Electron ionization
EICs	Extracted ion chromatograms
FQA	Ferruloylquninc acid
g	gram
GNPS	Global Natural Product Social Molecular
LC-MS	Liquid chromatography mass spectrometer
MN	Molecular network
MRM	Multiple reaction monitoring
MSI	Metabolomics standard initiative
PCA	Principal component analysis
PC	Principal component
PPP	Phenylpropanoid pathway
QA	Quinic acid
Rt	Retention time
UPLC	Ultra performance liquid chromatography
UHPLC-qTOF-MS/MS	Ultra high performance liquid chromatography quadrupole time of flight

Units

%	percent
µl	microlitre
ºC	degree Celsius
mM	millimolar
g/L	gram per litre
mL	mililitre
g	gram
µg/mL	microgram per millilitre
mg/mL	milligram per milliltre
µm	micrometre
mL/min	millilitre per minute
min	minutes
xg	times gravity
mm	millimetre
kV	kilovolt
v	voltage
L/h	litre per hour
pg/mL	picograms per millilitre
bw	band width

List of Figures and Tables

Figure 2.14: A schematic representation of the metabolomics workflow for plant metabolomics studies, showing three main steps which are sample preparation, sample analysis/data acquisition and data analysis modified from (Tugizimana *et al*., 2013).

Figure 3.1: Base peak intensity chromatograms showing distribution of metabolites within two *Coccinia* species (A) *Coccinia grandis* (B) *Coccinia rehmannii.*

Figure 3.2: Principal component analysis (PCA) scores scatterplot and corresponding loadings plot. **(A)** PCA scores scatterplot obtained from XCMS online showing 72% variation between PC 1 and PC 2, indicating the general clustering within the datasets of *Cocccinia grandis and Coccinia rehmannii* leaf samples. **(B)** Loadings plot that indicates the ions responsible for the clustering observed in the scores scatterplot.

Figure 3.3: Cloud plot and total ion chromatogram created by the XCMS online, comparing the metabolite composition between two *Coccinia* species. (**A**) Cloud plot showing comparison of metabolites with p-value ≤ 0.01 and fold change ≥ 2, between of *Coccinia grandis* (CG) and *Coccinia rehmannii* (CR) leaf extracts. (**B**) The total ion chromatogram showing metabolite composition between the two plants, with *Coccinia grandis* metabolites shown in green colour and *Coccinia rehmannii* metabolites in purple colour.

Figure 3.4: Box and whisker plot showing the distribution and significance level of some of the identified metabolites in Table 1 dominated by metabolites made up of hydroxycinnamate derivatives and flavonoids extracted from leaf extracts of two *Coccinia species*.

Figure 3.5: Representative UPLC-qTOF MS/MS chromatograms showing distribution patterns of compounds with pseudo-molecular ion peaks at *m/z* 593, *m/z* 609, *m/z* 739 isolated from (**A**) *Coccinia rehmannii* and (**B**) *Coccinia grandis*.

Figure 3.6: Representative UHPLC-qTOF-MS/MS spectra showing fragmentation patterns of flavonoids found in *Coccinia grandis* and *Coccinia rehmannii*. **(A)** Kaempferol-3-rhamnoside-7glucosyl-(1-2)-rhamnoside **(B)** Quercetin 3-O-rutinoside **(C)** Kaempferol-3-O-rutinoside **(D)** Kaempferol-3-*p*-(6”-caffeoylglucoside **(E)** Kaempferol-3-*p*-Coumaroylglucoside.

Chapter 1: General Introduction

1.1. Background

Medicinal plants continue to receive increased attention from scientific and pharmaceutical communities as potential source of biologically active drugs. They stand out amongst the most promising sources due to their natural origin and biocompatibility as compared to synthetic compounds, mainly because of their lower toxicity and various biological effects (Bassam, 2012). Most of these medicinal plants have been used traditionally for treating various diseases but have not been scientifically researched for their medical efficiency. However, evidence has shown that some plants possess antioxidant and anti-inflammatory properties which can be helpful in management of diseases such as diabetes (Ilkhanizadeh *et al*., 2016). Plants from Cucurbitaceae family such as *Momordica charantia* have been used as anti-diabetic agents (Sharma *et al*., 2010; Aggarwal & Aggarwal, 2011). Other plants from this family such as *Coccinia grandis* have also been used, but there is little information about this plant (Muniappan *et al.*, 2009).

Chemical compounds bio-synthesized by plants are classified into primary and secondary metabolites based on their function (Thirumurugan *et al*., 2018). Primary metabolites are compounds widely distributed in nature, and are directly involved in organism's metabolic pathways necessary for growth, development and reproduction (Kudjordjie *et a*l., 2020). Primary metabolites include lipids, lactic acid and carbohydrates which do not show any pharmacological activities, but play an important role in processes such as photosynthesis and respiration (Geetha & Geetha, 2014; Isah, 2019). Secondary metabolites are derived from primary metabolites but are not directly involved in the metabolic pathways (for growth and development). However, they are essential in the ecological function and they exhibit pharmacological activities with one example of these secondary metabolites being flavonoids (Martin, 2016; Moore *et al*., 2014).

Metabolomics has over the years become a valuable tool in understanding plants secondary metabolites. Metabolomic analytical techniques such as Liquid Chromatography Mass Spectrometry (LC-MS) have been used previously to elucidate

plants secondary metabolites (Madala *et al*., 2014). LC-MS is one of the most widely used tools because it is highly sensitive and is capable of covering a wide range of metabolites in a biological system (Sardans *et al*., 2011; Tugizimana *et al*., 2013). LC-MS is combined with multivariate statistical such as the principal component analysis (PCA) to better understand and interpret data. PCA allows for visualization of similarities and dissimilarities within and between samples by reducing the dimensionality of the dataset but preserving variance in the data (Zhou *et al*., 2012). In this study, Liquid Chromatography Mass Spectrometry was used for characterisation of metabolites between two closely related *Coccinia* plants. Furthermore, metabolomics approaches were followed to understand and interpret the generated complex data.

1.2. Aim

- The aim of this study was to apply LC-MS based metabolic approach to establish the correlation between two closely related *Coccinia* species and to investigate the biochemical strategies used by these plants to increase their already complex metabolome.

1.3. Objectives

- To study the discriminatory chemical factor between two closely related *Coccinia* species through LC-MS based metabolite fingerprinting and statistical data analysis.
- To investigate the biochemical strategies that lead to metabolome complexity of the two *Coccinia* species through LC-MS based metabolite fingerprinting.
- To showcase the chemical modifications of two *Coccinia* plants metabolites through metabolite fingerprinting with the aid of LC-MS based analysis.
- To study the discriminatory chemical factor between two closely related *Coccinia* species using LC-MS and statistical data analysis.

References

Aggarwal, S and Aggarwal, N. (2011). A review of recent investigations on medicinal herbs possessing anti-diabetic properties. Journal of Nutritional Disorders & Therapy, 1, 2.

Sharma, N., Sharma, M., & Bindal, M. C. (2010). Potential antidiabetic herbal drugs: A comparative review of marketed products. Research Journal of Pharmacognosy and Phytochemistry, 2; 115-121.

Thirumurugan, D., Cholarajan, A., Raja, S. S., & Vijayakumar, R. (2018). An introductory chapter: secondary metabolites. Second metab—sources Appl, 1-21.

Kudjordjie, E. N., Sapkota, R., & Nicolaisen, M. (2020). Arabidopsis sculpt distinct root-associated microbiomes through the synthesis of secondary metabolites and defense signaling molecules.

Isah, T. (2019). Stress and defense responses in plant secondary metabolites production. Biological research, 52.

Geetha, T. S., & Geetha, N. (2014). Phytochemical screening, quantitative analysis of primary and secondary metabolites of Cymbopogan citratus (DC) Stapf. leaves from Kodaikanal hills, Tamilnadu. International Journal of pharmtech research, 6; 521-529.

Muniappan, R., Reddy, G. VP & Raman, A. (2009). *Coccinia grandis*. In R.Muniappan (Ed.). Biological control of tropical weeds using arthropods, 10; 175-182

Martin, S. (2016). Identification and mechanistic study of novel drug metabolites by LC-MS. Sheffield Hallam University (United Kingdom).

Moore, B. D., Andrew, R. L., Külheim, C., & Foley, W. J. (2014). Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phytologist, 201; 733-750.

Madala, N. E., Piater, L. A., Steenkamp, P. A., & Dubery, I. A. (2014). Multivariate statistical models of metabolomic data reveals different metabolite distribution patterns in isonitrosoacetophenone-elicited Nicotiana tabacum and Sorghum bicolor cells. SpringerPlus, 3; 1-10.

Sardans, J., Penuelas, J., & Rivas-Ubach, A. (2011). Ecological metabolomics: overview of current developments and future challenges. Chemoecology, 21; 191-225.

Tugizimana F, Piater L, Dubery I (2013). Plant metabolomics: A new frontier in phytochemical analysis. South African Journal of Science, 109; 1-11.

Zhou, B., Xiao, J. F., Tuli, L., & Ressom, H. W. (2012). LC-MS-based metabolomics. Molecular BioSystems, 8; 470-481.

Chapter 2: Literature review

2.1. Phenolic compounds

Phenolic compounds are a large group of plant secondary metabolites that show a diversity of structures, from simple structures, e.g. phenolic acids, through polyphenols such as flavonoids and chlorogenic acids, to polymeric compounds such as tannins (Kala *et al*., 2016). These compounds are important to the quality of plant foods because they are responsible for the color of fruits, juices and wines, and they are substrates for enzymatic browning as well as other properties such as flavor and fragrances (Cheynier, 2012). Phenolic compounds consist of two major groups namely, hydroxybenzoic acids (HBA) and hydroxycinnamic acids (HCA) comprised of an aromatic ring (C6), with at least one or more hydroxyl groups (Pinto and Pollio, 2011; Vogt, 2010). The most common phenolic metabolites include flavonoids, chlorogenic acids, lignins and tannins which are synthesized from primary metabolites (particularly phenylalanine and tyrosine) through the phenylpropanoid pathway (Marchiosi *et al*., 2020). The latter of which is derived from the shikimate pathway (Vogt, 2010; Irchhaiya *et al*., 2015).

Phenolic compounds are classified on the basis of their structure, as either simple or complex phenolics (**Figure 2.1**) (Preethi *et al*., 2016). Simple phenolics include benzoic acid (BA) derivatives and cinnamic acid (CA) derivatives, which are the simplest phenolics in nature with 6- & 9- carbon skeletons (Giada, 2013). A carboxylic group is attached to the benzene ring of these compounds, with cinnamic acids having unsaturated propionic acid side chain attached to the benzene ring (Yang *et al*., 2001; Ochuko *et al*., 2018). Complex phenolics are compounds with higher molecular weight which include flavonoids and tannins as examples found in fruits and vegetables (Cooper, 2007; Pandey & Rizvi, 2009). Flavonoids possess two phenolic rings which attach an oxygenated heterocyclic pyrin ring, which enables further classification flavonoids into; flavones, flavonols and anthocyanins (Kumar and Pandey, 2013; Cheynier, 2005).

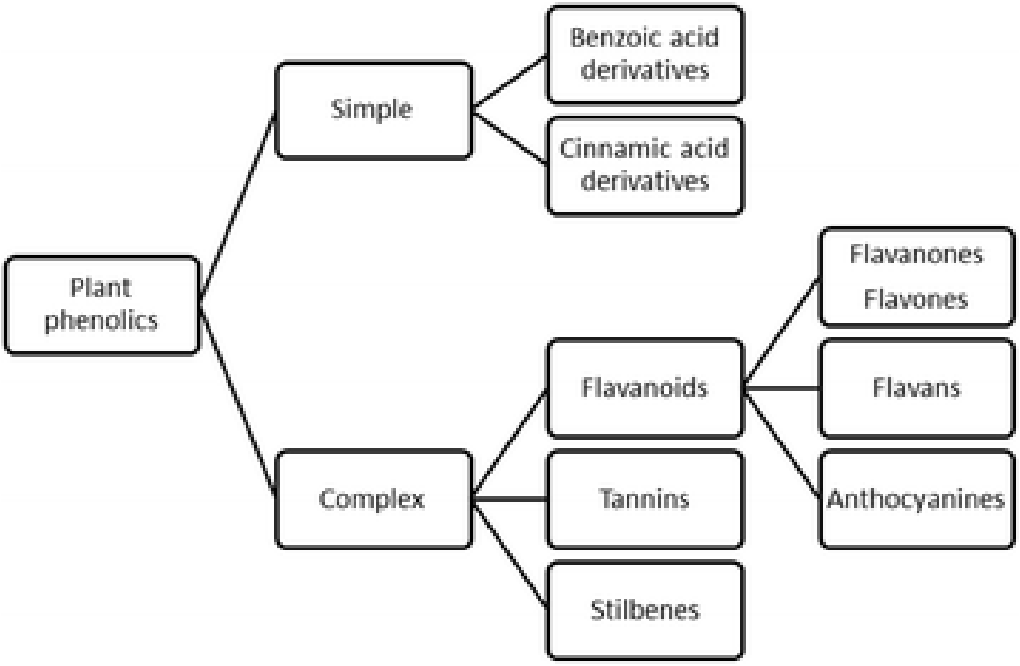


Figure 2.1: Classification of the plant phenols, based on their structural configuration (Preethi *et al*., 2016).

In plants, phenolic compounds are believed to play a key role as defensive compounds against environmental stresses, such as high light, low temperatures, pathogen infection, herbivores feeding and nutrient deficiency, which lead to an excessive production of free radicals and other oxidative species in plants (Lattanzio, 2013; Pieterse *et al*., 2009). Eating such plant as foods is therefore beneficial to humans because they contain a high antioxidant compounds that will reduce the occurrence of chronic diseases such as diabetes, cancer and cardiovascular diseases. Fruits and vegetables with low-amylase and high-glucosidase inhibitory activities can be utilised as food items that could facilitate the control of the early stages of hyperglycaemia, which is associated with type 2 diabetes (Iwai *et al*., 2006; Derong *et al*., 2016).

Pharmacologically, phenolic compounds are required in the prevention of chronic diseases such as cancers and cardiovascular diseases because of their anti-carcinogenic, anti-mutagenic and anti-glycemic properties. They are targeted for the design and development of health-promoting foods and food ingredients (Derong *et al*., 2016; Xu *et al*., 2016). Research has provided evidence where phenolic compounds have been utilized successfully for the benefit of human health. For example, caffeic acid (3, 4-dihydroxycinnamic acid) was reported to possess strong inhibitory effect on colony formation of skin cancer cells (Yang *et al*., 2014).

Caffeic acid also reportedly inhibited tumour incidence within a solar UV-induced skin carcinogenesis mouse model (Dzialo *et al*., 2016).

2.1.1. Biosynthesis of phenolic acids

Phenolic acid biosynthesis involves various pathways such as pentose phosphate, shikimate and phenylpropanoid pathways (PPP) (Shetty, 2004; Sarkar & Shetty, 2014). The biosynthesis mainly occurs via the PPP pathway from aromatic acids 1-phenylalanine and /or 1-tyrosine (Vogt, 2010; Vaishali and Neeraj, 2018). Phenylalanine is catalysed by the action of enzyme phenylalanine ammonia lyase (PAL) to form p-OH-cinnamic acid (**Figure 2.2**) (Carvalho *et al*., 2015) and ammonia. The following step involves reactions that are carried out by enzymes hydroxylases and methyltransferases, which result in formation of hydroxycinnamic acids (HCAs). Cinnamate-4-hydroxyase (C4H) regulates hydroxylation of cinnamic acid for the formation of p-coumaric acid also known as 4-coumarate (Olsen *et al*., 2008; Fraser and Chapple, 2011). The 4-coumaroyl-CoA ligase (4CL) activity results in formation of *p*-coumaroyl-CoA/ 4coumaroyl-CoA, which then acts as a precursor for production of secondary metabolites (Biala and Jasinski, 2018). Thus, phenylpropanoids serve as precursors to production of a wide range of metabolites in plants (Nanda *et al*., 2017; Kumar *et al*., 2019).

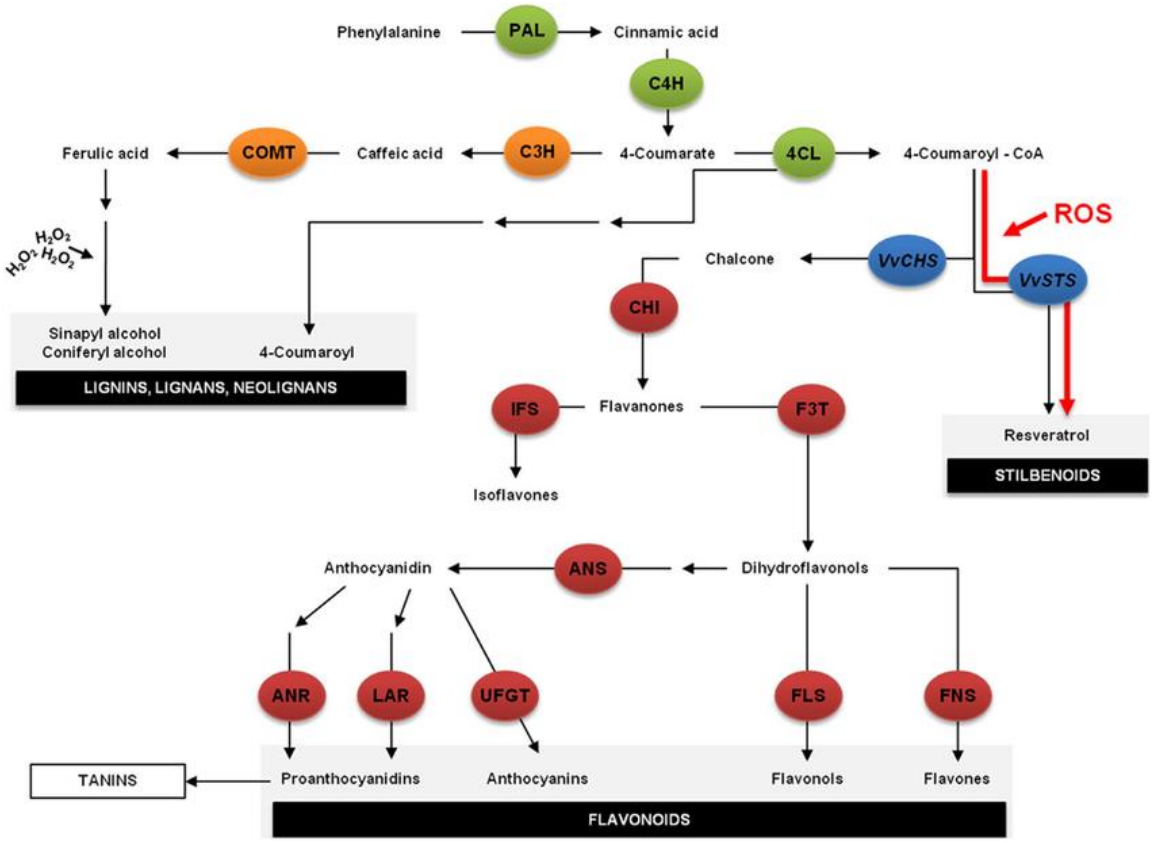


Figure 2.2: Schematic representation showing biosynthesis of secondary metabolites through phenylpropanoid pathway (Carvalho *et al*., 2015).

2.1.2. Flavonoids

Flavonoids are a group of phytochemicals that belong to a class of secondary metabolites and are present in almost all fruits and vegetables. Flavonoids share the basic chemical structure of three-ring moiety (A-, C-and B- rings) with a 15- carbon skeleton (C6-C3-C6) figure 3 (Alkhalidy et al., 2018). They are synthesized from the phenylpropanoid pathway and classified into either flavonols, flavones or isoflavones (Ferreyra *et al*., 2012; Kumar and Pandey, 2013). Flavonoids occur widely as glycosides (sugar attached) (Barbara, 2015), however they can also appear in aglycone form (no sugar attached) (Kumar and Goel, 2019) or in methylated derivatives (Nigel & Renee, 2014). The glycosylation of these flavonoids can be though O- or C-glycosides and they have been shown to attach (through acylation) to other biologically active molecules such as cinnamic acids (Muth *et al*., 2008; Jiang *et al*., Veitich & Grayer, 2011).

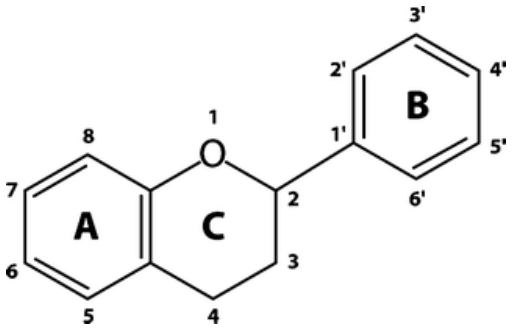


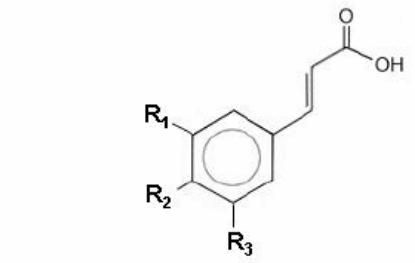
Figure 2.3: General structure of flavonoid indicating three ring moiety (Alkhalidy *et al*., 2018).

In plants systems flavonoids are responsible for the colour of fruits and vegetables and combating oxidative stress, with some of the best-known flavonoids being quercetin and kaempferol (Chandra *et al*., 2016; Panche *et al*., 2016). Flavonoids have been reported to be rich in antioxidant activity which can be beneficial to the immune system (Watson, 2019). It is said that a diet rich in flavonoid-containing foods exhibit beneficial health effect associated with cancer, neurodegenerative and cardiovascular disease prevention (Kozlowska & Dorota, 2014; Ahn-Jarvis, 2019). Onions, tea, strawberries, grapes, citrus fruit and many spices are just a few natural foods rich in flavonoids (Premkumar, 2014). Due to these health benefits of flavonoids, a need arises to evaluate their structure and function relationship. The biological activity and metabolism of flavonoids depend upon the structural requirements including structural configuration, the number of hydroxyl groups, and how functional groups are substituted around their core structure (Kumar and Pandey, 2016; Agati *et al*., 2013).

2.1.3. Hydroxycinnamic acids (HCAs)

HCAs are a major subgroup of phenolic or phenylpropanoids having a C6-C3 skeleton widely distributed in the plant kingdom. These compounds are hydroxyl derivatives of cinnamic acid with examples including caffeic acid, coumaric acid and ferulic acid (**Figure 4**) (Bento-Silva *et al*., 2020; Khoddami *et al*., 2013). They are mostly abundant in tea leaves, fruits and vegetables of which they have been characterized as structural and functional constituents of the cell walls and also play a role as dietary bioactive ingredients (Teixeira *et al*., 2013; Shahidi and Chandrasekara, 2010). The high concentration and wide distribution of HCAs provides them with a key role in biosynthesis of other phenolic compounds. HCAs such as caffeic acid are present in human diet (e.g. coffee) as either free or conjugated forms including amides

(conjugated with amines or amino acids), esters of hydroxyl acids (such as tartaric acid and quinic acid) and also conjugated to sugar derivatives as well as glycosides. The conjugation results in secondary metabolites such as chlorogenic acids and chicoric acids (Gonthier *et al*., 2006; Teixeira, 2013; Sova and Saso, 2020).



R_2 = OH: *p*-Coumaric acid
R_2 = R_3 = OH: Caffeic acid
R_3 = OCH_3, R_2 = OH : Ferulic acid

Figure 2.4: General structure of hydroxycinnamic acids (Bento-Silva *et al*., 2020).

HCAs such as chlorogenic acids (CGAs) have been recently reported as potential therapeutic agents in oxidative stress related diseases such as cancer and inflammatory injuries. Research shows that the antioxidant potency of hydroxycinnamic acids depends on their structural features, which informs their ability to scavenge free radicals (Zhang and Tsao, 2016; El-Seedi *et al*., 2018). Due to their overwhelming purported pharmacological activities, analyses of CGAs have been on the rise, with plants such as coffee deemed to contain the largest contingency of these compounds (Tshabalala *et al*., 2019). Elsewhere, it has been reported that a diet rich in HCAs exhibit beneficial health effects such as protection against cardiovascular diseases (Pragasam *et al*., 2013; Alam *et al*., 2016). In recent years, there has been a rise in studies on the biological activity and structural diversity of plant phenolics (Dai & Mumper, 2010; Wannenmacher *et al*., 2018). However, there are still plant species such as *Coccinia*, of which the phenolic compounds content and composition are still insufficiently explored.

2.2. Glycosylation and its biological purposes

Plants produce a wide variety of secondary metabolites which play key roles in the coordination and defence strategies of plants. The diversity in the structure and function of these metabolites arise through different biochemical processes such as conjugation (Rai *et al*., 2016) and methylation (Wang *et al*., 2019). These processes enable plants to produce secondary metabolites with diverse physical and chemical properties (Wink, 2010). Conjugation of metabolites is one of the processes that causes a change in chemical properties of metabolites, resulting in novel bioactivity and stability (Vaistij *et al*., 2009; Bowles *et al*., 2005). To showcase an example of the diversity of metabolites brought about by chemical conjugation herein, (**Figure 2.5**) shows chicoric acid vs dicaffeoylquinic acid. Chicoric acid shows two caffeic acids conjugated to a tartaric acid and dicaffeoylquinic acid shows two caffeic acids conjugated to a quinic acid, thereby indicating that chemical conjugation plays a vital role in metabolite diversity (Vinholes *et al*., 2015). Glycosylation is one of the most prominent conjugation process as it modifies plant secondary metabolites through the addition of carbohydrate moiety to an acceptor aglycone, thereby, altering hydrophobicity and bioactivity of natural products (Bowles *et al*., 2005; Dias *et al*., 2012).

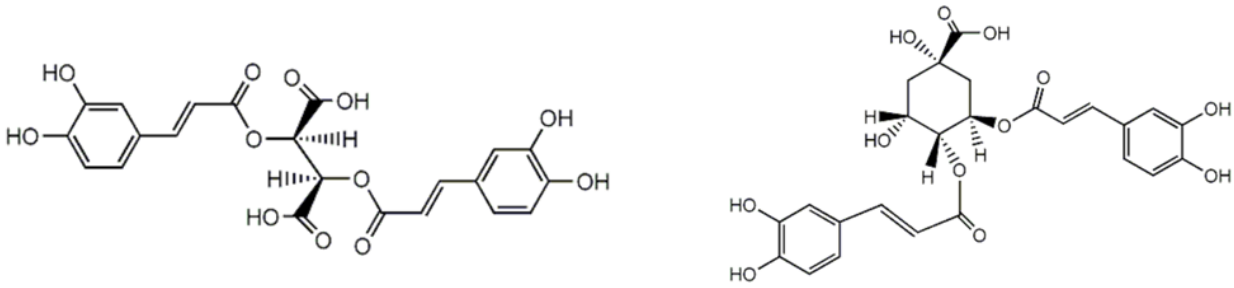


Figure 2.5: Chemical structures showing conjugation of caffeic acid to two different molecules (Vinholes *et al*., 2015).

For instance, flavonoid O-glycosides sugar moiety are substituted to the hydroxyl group of the aglycones, whereas, C-glycosides have their sugar moiety attached to the aglycone through a carbon-carbon bond (Negri & Tabach, 2013). There is little

information known about the enzymes involved in the glycosylation processes, however, more is known about their regulation and coordination with other processes. Many core aglycone moieties undergo regio- & sterio- selective glycosylation which involves different sugar molecules, thus resulting in glycosides harbouring diverse chemical structures and properties (Wink, 2010; Vaistij *et al*., 2009).

Glycosylation is also a major regulator of phenylpropanoid availability in plants. The phenylpropanoid pathway is responsible for the biosynthesis of a huge amount of secondary metabolites such as flavonoids and tannins, derived from phenylalanine and tyrosine (Vogt, 2010; Le Roy *et al*., 2016). The action of enzymes such as UDP-glycosyltransferases (UGTs) further increase the diversity of these molecules resulting in the production of glycosylated hydroxycinnamates and related aldehydes and alcohols (Wilson & Tian, 2019; Tiwari *et al*., 2016). Glycosylation can change solubility and stability phenylpropanoid as well as influence the compartmentalization and biological activity. Hence, it is a powerful mechanism that allows plants to regulate phenylpropanoid localisation, availability and biological activity (Cheynier *et al*., 2012; Le Roy *et al*., 2016). Most phytochemical data have shown that glycosylation of metabolites such as flavonoids is a biological phenomenon which is also prone to further chemical modification isomerization, thereby complicating an already complex metabolome (Ogo *et al*.,2016).

2.2.1. Isomerization

Chemical modifications such as positional and geometrical isomerization in plant metabolites, also contribute to the high complexity of plant metabolome (Masike *et al*., 2017; Masike & Madala, 2018). Isomerization is a chemical process of transforming a chemical compound into its isomeric forms, resulting in compounds with the same atomic composition but different structural configuration (Dhar *et al*., 2018). Plants produce a wide range of diverse chemical compounds, most of which are manufactured through specialized metabolism (Masike *et al.,* 2018; Pyne *et al*., 2019). Glycosides are some of the compounds that have been reported to possess many isomers with same molecular weight but differ in the hydroxylation pattern and substitution of their rings, such isomers are known as positional isomers (**Figure 2.6**) (Masike & Madala, 2018). These compounds are a result of modifications such as

glycosylation, acetylation and addition of biological compounds such as cinnamic acids (Muth *et al*., 2008; Kachlicki *et al*., 2008).

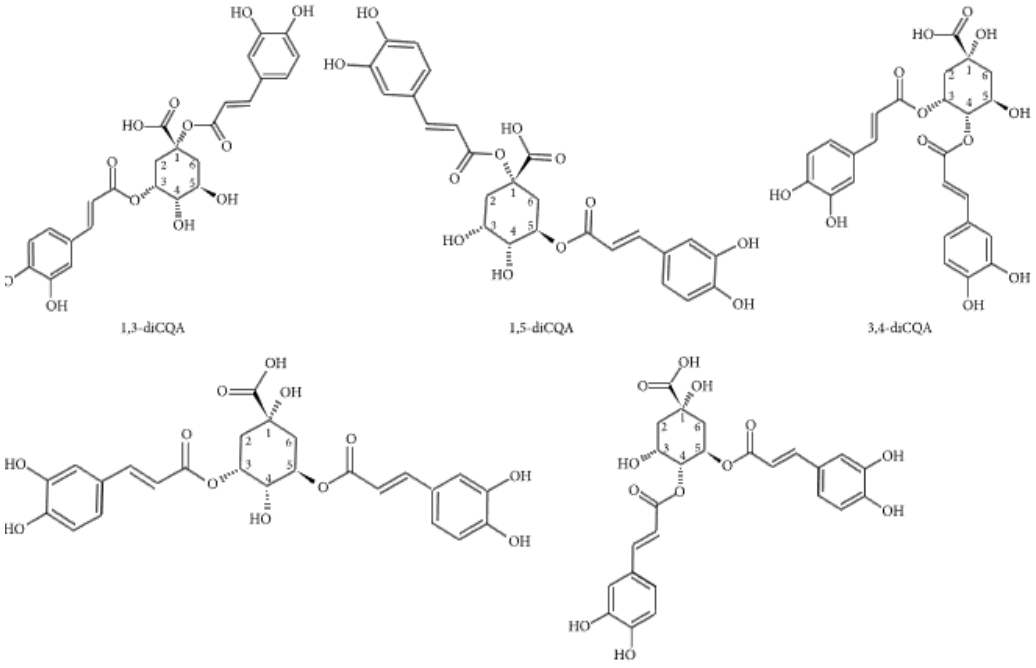


Figure 2.6: Chemical structures of the trans-form positional isomers of dicaffeoylquinic acids (diCQAs) isolated from *Vernonia fastigiata* extracts (Masike *et al*., 2018).

Naturally, almost all HCA derivatives are synthesized with a *trans* configuration, but *cis* isomers have also been noted as a result of photo-isomerization, thereby resulting in structurally diverse metabolomes which are characterized by positional isomers, geometrical isomers (**Figure 2.7**) (Ncube *et al*., 2014) and differential conjugation to various molecules (Clifford *et al*., 2003; Clifford *et al*., 2008; Masike *et al*., 2017; Nobela *et al*., 2019). Isomerization has been shown to be an evolutionary strategy of plants to diversify their metabolite compositions and the presence of such modifications has poised an undisputed analytical challenge, with most of these isomers being identified as mere structural artefacts of one another (Nengovhela *et al*., 2020).

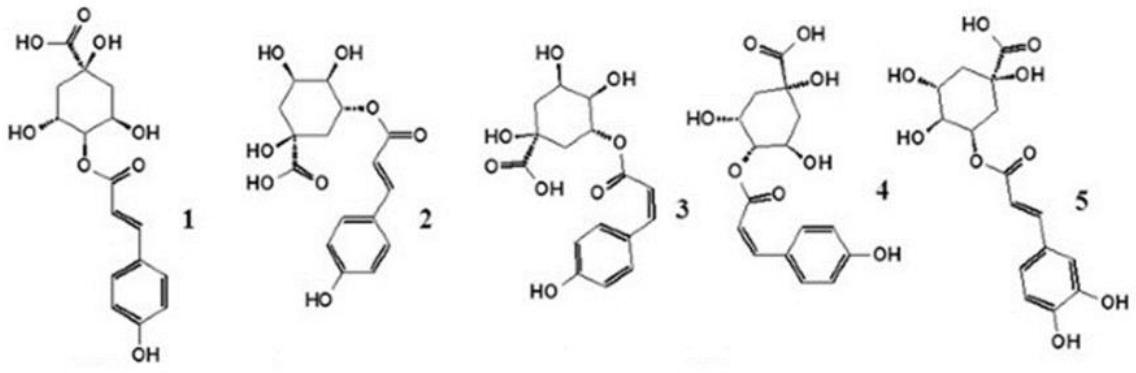


Figure 2.7: Structures of geometrical isomers of p-Coumaroylquinic acid detected in tobacco leaves (Ncube *et al*., 2014).

2.2.2. Glyco-isomerization and bioavailability

Plant metabolites such as flavonoids undergo chemical modification such as glycosylation, which are also prone to further modifications such as isomerization and acetylation. It has been reported that flavonoids also form isomers through attachment of sugar moieties (*di*- or *tri*- saccharides), a phenomenon known as glyco-isomerization (**Figure 2.8**). This phenomenon has been revealed as a means by plants to diversify its metabolites (Ogo *et al*., 2016; Nengovhela *et al*., 2020). It has also been shown that the subtle differences in the position of glycosidic bonds results in isomeric compounds (Avato and Argentieri, 2019). Elsewhere, the attachment (through acylation) of other biologically active molecules such as cinnamic acids was reported to be responsible for production of isomeric compounds (Muth *et al*., 2008; Kachliki *et al*., 2008). The above phenomenons were viewed through the use of LC-MS approaches, for example, by studying isomeric molecules appearing at different retention times.

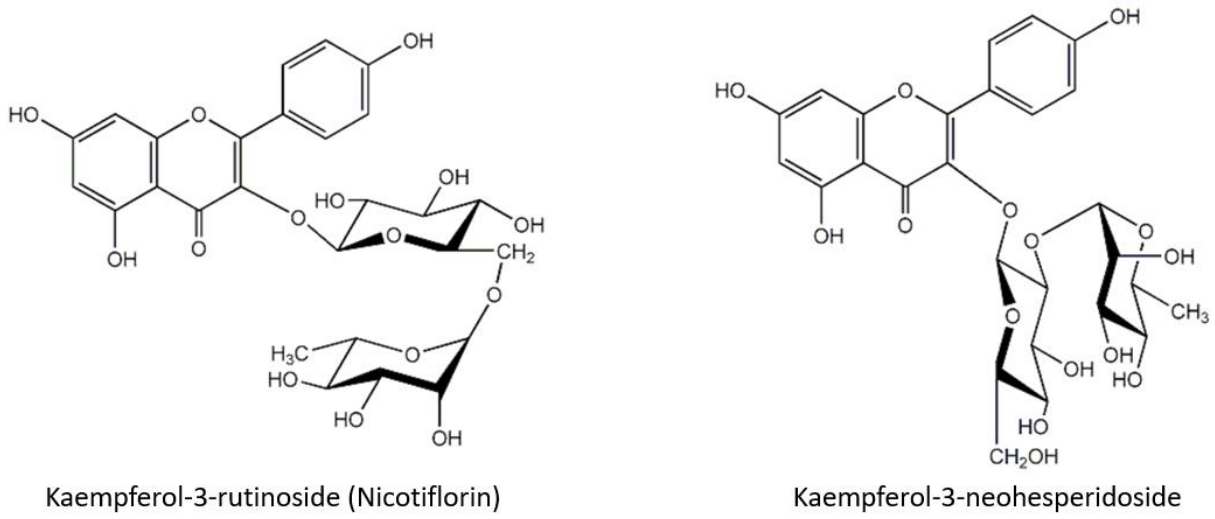


Figure 2.8: Representative of isomeric molecules of flavonoid kaempferol with different sugar moieties attached (Cuyckens *et al*., 2000).

Flavonoids predominantly exist as glycosides. The importance of sugar is for storage and compartmentalization. Flavonoid metabolism involves structural modifications such as hydroxylation and methylation of flavonoid glycosides before they perform their intended functions such as antioxidant activities (Mullen *et al*., 2006; Chen *et al*., 2014). Before flavonoids are consumed by humans, there are several modifications that take place such as “sugar hydrolysis” which is responsible for the removal of sugar moiety, to result in an aglycone which then confers the activities. These modifications perform their function through specialized enzymes such as hydrolases and glucosidases which in turn affects the bioavailability of flavonoids. It is important to note that the sugar attachment is critical for diversification in order to avoid *in vivo* toxicity, and as such many plants are known to produce flavonoids attached to diverse sugar molecules, one example being Moringa, which produces a number of quercetin differing in the number of sugars attached (Makita *et al*., 2016).

Bioavailability is defined as the measure of the rate at which an active substance is absorbed and how it functions when it reaches its site of action or biological destination. It is well known that sugar molecules enhance the bioavailability of active molecules, therefore removal of sugar moiety is consequential to the bioavailability of that molecule. Flavonoids usually undergo sulfurylation, glucuronidation and methylation in the small intestine after ingestion (Mullen *et al*., 2006) and their metabolites tend to show reduced bioactivity when compared with parent compounds

(Rupasinghe *et al*., 2010), although there have been results that reported otherwise as well (Thilakarantha and Rupasinghe, 2013).

The first step is conjugation in the small intestine, followed by the liver where they are further metabolized into glucuronides and sulfate derivatives (**Figure 2.9**) (Ruphasinge *et al*., 2013) which facilitate their excretion via urine and bile (Scalbert et al., 2002). The bioavailability of flavonoids is also a determinant factor of their bioactivity *in vivo*, therefore, increasing the bioavailability is important in enhancing health effects *in vivo* (Thilakarathna *et al*., 2013). Metabolism and the rate of flavonoid absorption in humans has been reported to be dependent primarily on the type and position of a sugar moiety, therefore, it is also imperative to study the biological actions of flavonoids whilst they are in their aglycone form, to see how they carry out their mode of action *in vivo* (Greafe *et al*., 2001).

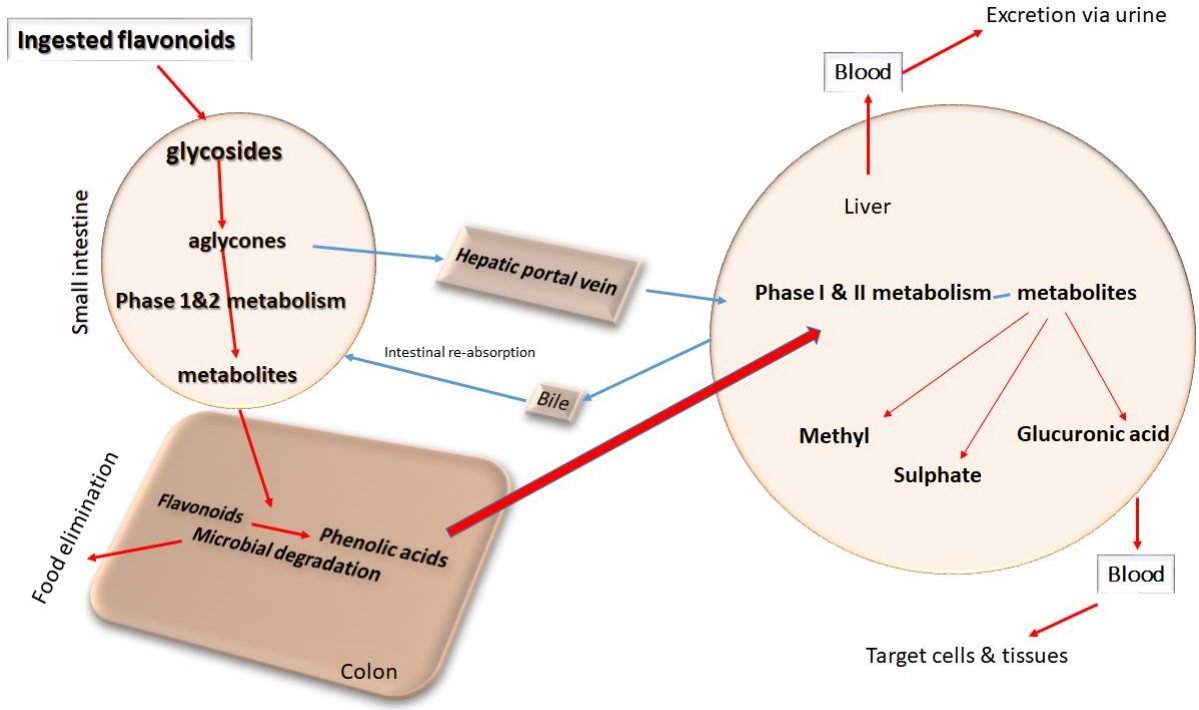


Figure 2.9: Representative of a simplified human flavonoid metabolism. Ingested flavonoids are metabolised in the intestine, transported to the liver where they undergo further metabolism, and the resultant metabolites are then transported to the site of action and then secreted or undergo re-circulation, modified from (Ruphasinge *et al*., 2013).

2.3. *Coccinia* genus

The genus *Coccinia* falls under the family Cucurbitaceae and forms part of the major group of angiosperms. It comprises 25 known species and are commonly referred to as scarlet gourds. Some of the most famous species include; *C*. *grandis*, *C*. *rehmannii, C*. *adoensis*, *C*. *barteri*, *C*. *grandiflora* (Mathews and Endress, 2004). *Coccinia* species are widely distributed in the Sub-Saharan Africa (from semi-arid Savannahs to Rain forests), with one species namely *Coccinia grandis* also found in tropical Asia (Pekamwar *et al*., 2013). The *Coccinia* species are perennial climbers or creeps which is enabled by its rather simple or bifid tendrils. They have long, elastic tendrils with a coil-like spring character that allows the plant to wrap around its host to the entire length (Holstein, 2015; Butler, 2019).

Coccinia species are characterized by simple to deeply lobed, stalked and paired leaves, having a tendril on each node (Holstein & Renner, 2011). Individual plants in *Coccinia* genus either produce flowers with only male organs or with only female organs. Female flowers contain three stamens whereas male flowers lack any sign of female organs (Schmidt & Zizka, 2014). Their fruit is a red-flesh berry with a red skin when ripe, with seeds covered by a flat to lentil shaped juicy hull. Most plants in the *Coccinia* genus possess roots with woody base that produce hypocotyl tubers. These root tubers in one of the species (*C*. *ardoensis*) are used for adaptation to fire, since the species is most dominant in woodlands. Some species such as *C*. *grandis* and *C*. *barteri* produce adventitious roots if their stems reach the ground/soil (Schaefer & Renner, 2011; Zhang *et al*., 2006).

Several plants from the *Coccinia* genus are utilised by tribal communities as a source of food and cultural applications. Some of the plants fruits (e.g. *C.grandis*) are edible when ripe but can also be cooked when unripe, and the young shoots and leaves are eaten as spinach (Maroyi, 2013; Starr *et al*., 2003). These parts of the plant have been reported to be rich in carotenoid content and thus they are highly valued. High carotenoid content is of great importance to counter Vitamin-A deficiency amongst women and young children (Addis *et al*., 2009; Wasantwisut & Vitiyapanach, 2003; WHO, 2009). One of the most utilized species is *C. grandis*, which is also known for its pharmaceutical attributes and has been tipped to be of great value in health applications (Pekamwar *et al*., 2013; Tamislelvan *et al*., 2016).

2.3.1. *Coccinia grandis* (Ivy gourd)

Coccinia grandis or ivy gourd, also known as scarlet gourd, is a tropical vine in the Genus *Coccinia* in the family of Cucurbitaceae. It is a fast growing, aggressive climbing vine that can form dense mats on lands, readily covering small trees and fences (Pekamwar *et al*., 2013). This plant widely grown in all tropical regions due to favourable climate conditions found in those areas. Ivy gourd plant is broad-leafed in nature and it grows as long as 30m. They are climbers with long tendrils, tuberous roots and their leaves are broad and arranged in fluctuated fashion (Csurhes and Edwards, 1998; Renner *et al*., 2007). Their flowers have shapes like stars which then transform into fruits (**Figure 2.10**). It climbs over other plants, supporting itself by means of its tendrils.

Figure 2.10: Physical appearance of *Coccinia grandis*.

C. *grandis* has been reported to contain large amount of beta-carotene and rich in complex carbohydrates, and an array of vitamins B. Several phytochemical constituents of great importance have been reported from different parts of the *C. grandis* plants. Phenolic compounds are generally noted for their antimicrobial activities (Uddin, 2014; Kondhare and Lade, 2017). The fruits of *C. grandis* are known to contain active constituents like lupeol and glycoside cucurbitacin B (Deokate and Khadabadi, 2011; Chanda, 2020). The leaves are bittersweet and astringent and possess major phyto-constituents such as cardenolides, saponins, flavonoids and polyphenols (Sikta *et al*., 2018). The Ivy gourd changes colour from green to pink as

an indication of ripening and loses its weight due to rapid wilting during storage. The whole plant of *C.grandis* has been reported to possess diverse pharmacological activities such as anti-inflammatory, antidiabetic, antioxidant, hepatoprotective and mutagenic activities (Deshpande, 2011; Kulkami, 2012).

2.3.2. *Coccinia rehmannii*

A non-woody, climbing tree from the Cucurbitaceae family which produces stems from a large rootstock found in South Africa and Namibia. It scrambles over the ground and over smother vegetation, covers fences and walls by its simple tendrils (Retief and Meyer, 2017). The leaves of *Coccinia rehmannii* are simple, alternate and paired (**Figure 2.11**) with the younger leaves shallowly lobbed, and the upper leaves whitish. The plant has roots which are woody at the base, grow up to 0.5 m and produces carrot-like tubers. It produces whitish cream-colored to yellow flowers, and its fruits are many seeded globose berries. The seeds are ovate, greyish, a flat surface and often has a darker colour (Holstein, 2015).

Figure 2.11: Physical appearance of *Coccinia rehmannii*.

The plant grows in either full sun or shade and is mostly found in woodlands and semi-desert areas. *C. rehmannii* is sometimes harvested as food of which the leaves are eaten locally as spinach and its fruits are eaten when ripe. Its tubers contain starch and are boiled or roasted then peeled and pounded and mixed with leafy vegetables

for eating (Leffers, 2003). The roots of *C. rehmannii* are also food items, as a source of water and starch. Its use as medicinal plants have been poorly reported and thus is categorized as an underutilized plants, even though some local traditional healers are reported to prepare medicinal concoctions from its bulb/roots for treating different ailments (Van der Walt, 2009; Holstein, 2015; Sigidi *et al*., 2017).

2.3.3. Pharmacological activities of *Coccinia grandis*

2.3.3.1. Anti-malarial activity

Extracts of *Coccinia grandis* have been reported to display excellent anti-plasmodial activity against the *Plasmodium falciparum* (Sundaram *et al*., 2012). Its aqueous leaf extract decreases the serum glutamic pyruvic transaminase (SGPT), serum glutamic oxaloacetic transaminase (SGOT), alkaline phosphatase (ALP), blood urea nitrogen concentration (Umamaheswari & Charttejee, 2007). The antimalarial activity of *C. grandis* extract is controlled by its strong affinity for water and also its ability to dissolve in other polar substances (Pekamwar et al., 2013. The extract of *C. grandis* have also been reported to decrease the parasite count of *Plasmodium berghei* in mice (Bhattacharya *et al*., 2011).

2.3.3.2. Anti-diabetic activity

The regulation of intracellular glucose homeostasis requires pancreatic β-cell function at its optimum. Several studies have shown that losing functional β-cell mass through apoptosis and impaired proliferation resulting to hyperglycaemia, is considered as a characteristic of diabetes mellitus (Inaishi & Saisho, 2020). One critical therapeutic challenge known in patients with the disease is the regulation of functional β- cell mass. Thus, regeneration of islet cell has gained much interest and has been recently considered as a suitable strategy for restoring the loss of β-cell mass in diabetes mellitus (Ashcroft and Rorsman, 2012; Hosseini *et al*., 2015). As evidenced elsewhere, an extract of *C*. *grandis* showed β-cell regeneration through an increase in the number of insulin secreting β-cells in diabetic rats with induced streptozotocin (Attanayake *et al*., 2015). For conformation, islets with increased diameter were found in *C*. *grandis* treated diabetic rats, indicating that treatment with *C*. *grandis* extract increased the

functional mass of the islet and the entire regenerative capacity of the pancreas in diabetic rats (Vetere *et al*., 2014; Attanayake *et al*., 2015).

2.4. Introduction to plant metabolomics

Metabolomics is defined as a large-scale study of metabolite composition within living organisms, under given physiological status (Sumner *et al*., 2003; Verpoorte *et al*., 2013). It is the downstream of the genomic, transcriptomic and proteomic levels (**Figure 2.12**) (Martis *et al*., 2011) and it captures the chemistry of an organism which changes all the time depending on the surrounding environmental conditions (Ramabulana *et al*., 2020). These environmental factors result in chemical shifts characterised by structural modification such as glycosylation and chemical conjugation. A metabolome is the collection of all small metabolites that are produced during metabolism in a biological system. These small metabolites are involved in cellular activity and physiological status of cells (Burgess *et al*., 2014). For a while, little amounts of metabolites have been used for diagnosis of complex metabolic diseases and monogenic disorders such as inborn errors of metabolism (Clish, 2015; Fernando *et al*., 2016). However, current metabolomic technologies expand beyond standard clinical chemistry techniques and can perform precise inspection of thousands of metabolites in a single analytical analysis (Scalbert *et al*., 2009; Pinu *et al*., 2019). Thus, it is an emerging technology that holds promise in the practice of precision (Clish, 2015; Bowne *et al*., 2018). Metabolomics approaches allow for precision medicine in various levels, including; the discovery of new therapeutic targets, biomarker discovery and characterization of metabolic irregularities that cause diseases (Bhalla *et al*., 2005).

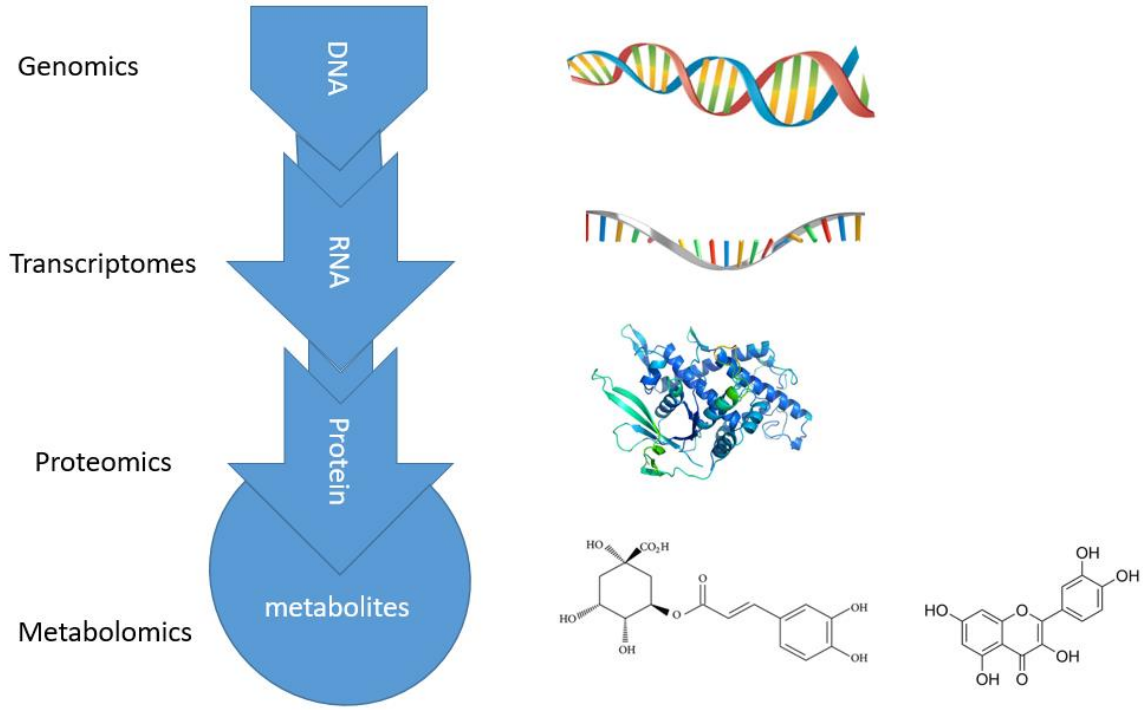


Figure 2.12: Schematic representation of the flow of biological information from genomics to metabolomics, modified from (Martis *et al*., 2011).

Compared to other "omics" approaches, metabolomics uses highly advanced analytical tools that are highly sensitive and selective. Such analytical tools include Infrared (IR) spectroscopy, Nuclear Magnetic Resonance (NMR) Spectrometry, Liquid Chromatography Mass Spectrometry (LC-MS), and Gas Chromatography-Mass Spectrometry (GC-MS) (Kim et al, 2011; Allwood and Goodacre, 2010; Tugizimana *et al*., 2013). The most commonly used technique is LC-MS, which is highly sensitive and can identify a wide range of secondary metabolites (Moco *et al*., 2007; Madala *et al*., 2013). LC-MS combines the chromatographic power of HPLC and the accuracy of an MS detector (**Figure 2.13**) (Das, 2015) to identify several metabolites in a single run. Preference is mostly given to LC-MS due to its many advantages, which include short analysis time and its high separating power.

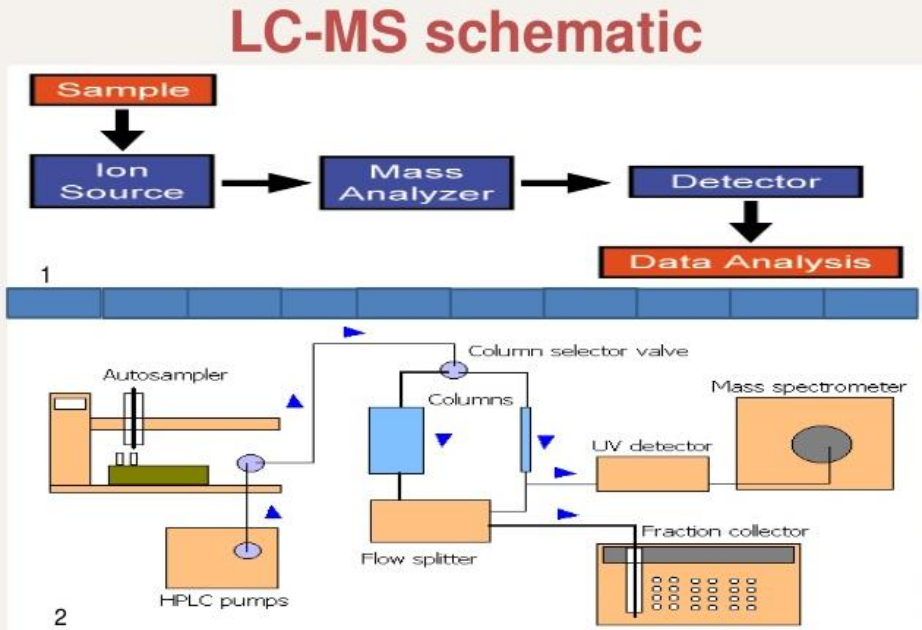


Figure 2.13: A simple schematic representation of an LC-MS system adapted from (Das & Pal, 2015).

LC-MS enables the identification, quantification and analysis of masses of wide variety of non- volatile or semi-volatile organic compounds in a mixture (Lee *et al*., 2013; Ramabulana *et al*., 2015). Accuracy of LC-MS and the ability to acquire specific information is aided by the use of ionizing technologies such as electron ionization (EI) and electrospray ionization (ESI) which are used as interface between LC and MS (Ramabulana *et al*., 2015). LC separates components of mixtures by passing through a chromatographic column and the separated sample species are sprayed onto atmospheric pressure ion source, where they get converted into ions in the gas phase (Korfmacher, 2005). To sort the ions according to their mass/charge ratio, a mass analyser is used and the detector counts the emerging ions and amplifies the signal that is generated from each ion. A mass spectrum (plot of the ion signal as a function of the mass-to-charge ratio) is created as a result and is used to determine the elemental nature of a sample, the masses of a particle and to clarify the chemical structures of molecules (Korfmacher, 2005; Sparkman, 2003).

2.4.1. Workflow of metabolomics

Metabolomics has been used in research for interpretation of metabolic pathways, biomarker discovery and the discovery of metabolites that are involved in plant response toward environmental conditions (abiotic/biotic) (Lewis *et al*., 2008; Madala *et al*., 2013; Ramabulana *et al*., 2015). Metabolites are small molecules that belong to various classes of compounds such as amino acids and lipids. They possess complex physico-chemical properties which mainly differ in size, polarity and solubility (Kortbeek *et al*., 2019). Thus, analysis of these compounds becomes challenging during metabolomics studies, as a result a wide range of metabolomics strategies including metabolite profiling, metabolite fingerprinting and metabolite target analysis have been developed (Allwood *et al*., 2008; Hall, 2006; Tugizimana *et al*., 2013).

Metabolomics study workflow comprises three main experimental steps (**Figure 2.14**) namely sample preparation, sample analysis/data acquisition and data analysis (Rodrigues et al., 2019). Each step is compulsory, and they all require comprehensive optimization to ensure the quality of data generated and meaningful biological interpretation (Tugizimana *et al*., 2013; Mushtaq *et al*., 2014).

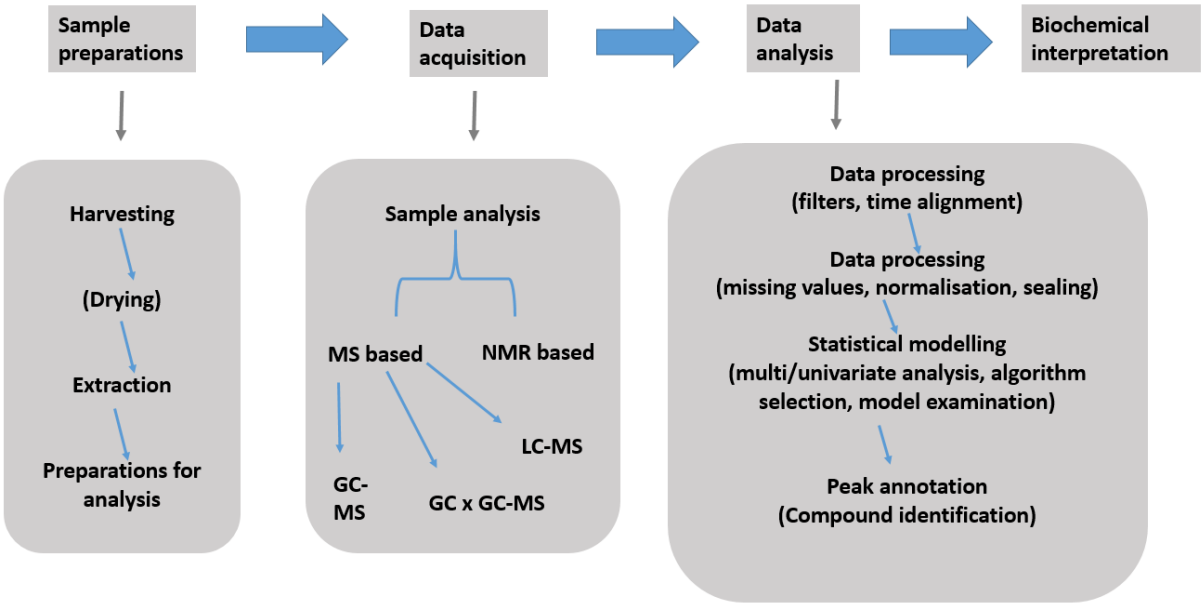


Figure 2.14: A schematic representation of the metabolomics workflow for plant metabolomics studies, showing three main steps which are sample preparation, sample analysis/data acquisition and data analysis modified from (Tugizimana *et al*., 2013).

2.4.1.1. Sample preparation

Preparation of samples is the first and most crucial step because it determines the outcome of the study. This step is highly dependent on the type of sample under study, the overall experimental design of the study and expectation thereof. Specimens that can be used during sample preparation include leaves, roots, stems and fruits of plants because the metabolite content varies throughout the whole plant (Moco *et al*., 2007; Kim & Verpoorte, 2010). There are several steps that are crucial during sample preparations including; harvesting of sample, drying, extraction of metabolites and preparing the samples for analysis. For metabolite extraction, different methods can be employed such as solid phase extraction (SPE) (He *et al*., 2018) and pressurized hot water extraction (PHWE) (Gbashi *et al*., 2017).

There is no single-extraction that can recover all classes of metabolites at once, thus, extraction methods that complement each other can be used in combination to extract wide range of metabolites (Tugizimana *et al*., 2013). There are factors that need to be taken into consideration during preparation of samples for analysis, such as removal of solvent during extraction by evaporation or freeze-drying and/or filtering the samples before introducing them to the analytical apparatus such as UPLC in order to avoid clogging (Mushtaq *et al*., 2014; Kim & Verpoorte, 2010).

2.4.1.2. Data acquisition

Following extraction, metabolite analysis is the second step of the metabolomics workflow where various analytical apparatus are used. Each analytical technique has its advantages and disadvantages and therefore, the choice of analytical instrument depends on the aim of the study and the properties of class compounds. MS-based techniques are one of the most commonly used analytical techniques, because of their high sensitivity and robustness (Allwood *et al*., 2008; Bianchi *et al*., 2018). In order to offer more fast, sensitive and selective metabolic analysis these techniques are coupled with chromatographic analysis such as HPLC (Feihn, 2002; Kumar *et al*., 2017; Sardans *et al*., 2011). In this study, we opted to UHPLC-qTOF-MS/MS technique for its high sensitivity and throughput to characterize and structurally elucidate metabolites of interest.

2.1.4.3. Data analysis

Datasets generated by LC-MS based metabolomics are very complex and, as such, in order to clean and minimize variations in the generated data, the pre-processing and pre-treatment steps of the data are of necessity. The pre-processing step involves noise filtering, detection of peaks and alignment, peak identification and generation of data table from the MS raw data (Zhou *et al*., 2012; Tugizimana *et al*., 2016). Data pre-treatment involves procedures such as centering, scaling and transformation of the data, in order to generate data areas that reflect the metabolite concentrations. Several software for data-pre-processing step include; MetAlign, XCMS, MarkerLynx, and MetaboAnalyst (Van den Berg *et al*., 2006; Zhou *et al*., 2012).

Molecular networking (MN) is another method that allows for organization and visualization of tandem mass spectrometry (MS) data (Pilon *et al*., 2019). MN is increasingly becoming an essential bioinformatics tool for annotating non-targeted mass spectrometry (MS) data. It is able to capture all metabolites detected by an analytical instrument under optimum conditions, thereby allowing chemical relationships between molecules to be established based on their similarity (Nothias *et al*., 2019; Quinn *et al*., 2017).

2.1.4.4. Compound identification

During metabolite identification, it is important to consider aspects such as accurate mass-to-charge ratio (m/z), chromatographic separation (rt) and the fragmentation patterns, as they can be used to deduce molecular formulae of metabolites (Moco *et al*., 2007; Watson, 2013). The molecular formula of an ion can be generated through accurate mass of a compound and used to search through online metabolomics database for metabolite identity. Fragmentation patterns of metabolites are obtained through approaches such as tandem MS (MS/MS) and in order to acquire information about their structure, the data is compared with already existing mass spectral libraries (Sud *et al*., 2007; Pluskal *et al*., 2010).

There are four levels introduced by the Metabolomics Standard Initiative (MSI) for metabolite identification (Summer *et al*., 2007). Level 1 involves identified metabolites

through Level 2 and 3 (putatively annotated and putatively characterized compounds classes) and lastly, level 4 which involves the unknown metabolites that can still be characterized (Salek *et al*., 2013; Creek *et al*., 2014; Viant *et al*., 2017). Examples of Metabolomics libraries that are used in the identification of metabolites include; Knapsack (www.knapsackfamily.com), PubChem (www.pubchem.ncbi.nlm.nih.gov), Chemspider (www.chemspider.com).

References

Ahn-Jarvis, J. H., Parihar, A., & Doseff, A, I. (2019). Dietary flavonoids for immunoregulation and cancer: Food design for targeting disease. Antioxidants, 8; 202.

Addis, G., Baskaran, R., Raju, M., Ushadevi, A., Asfaw, Z., Woldu, Z., & Baskaran, V. (2009). Effect of blanching and drying process on carotenoids composition of underutilized Ethiopian (*Coccinia grandis L. voigt*) and indian (Trigonella foenum-graecum L.) green leafy vegetables. Journal of Food Processing and Preservation, 33; 744-762.

Agati, G., Azzarello, E., Pollastri, S & Tattini, M. (2012). Flavonoids as antioxidants in plants: Location and functional significance. Plant science: an international journal of experimental plant biology. 196; 67-76.

Agati, G., Brunetti, C., Di Ferdinando, M., Ferrini, F., Pollastri, S., & Tattini, M. (2013). Functional roles of flavonoids in photoprotection: new evidence, lessons from the past. Plant Physiology and Biochemistry, 72; 35-45.

Alam, M.A., Subhan, N., Hossain, H., Hossain, M., Reza, H.M., Rahman, M.M. and Ullah, M.O. (2016). Hydroxycinnamic acid derivatives: a potential class of natural compounds for the management of lipid metabolism and obesity. Nutrition & metabolism, 13; 27.

Aggarwal, N and Shishu (2011) A Review of recent investigations on medicinal herbs ossessing anti-diabetic properties. Journal of Nutritional Disorders & Therapy, 1; 2.

Alkhalidy, H., Wang, Y & Liu, Dongmin. (2018). Dietary flavonoids in the prevention of T2D: An Overview. Nutrients, 10; 438.

Anantharaju, P.G., Gowda, P.C., Vimalambike, M.G. et al. (2016). An overview on the role of dietary phenolics for the treatment of cancers. Nutritional Journal, 15; 99.

Allwood, J. W & Goodacre, R. (2010). An introduction to liquid chromatography-mass spectrometry instrumentation applied in plant metabolomic analysis. Phytochemical Analysis, 21; 33-47.

Ashcroft, F. M & Rorsman, P. (2012). Diabetes mellitus and the beta cell: the last ten years, 148; 1160-1171.

Attanayake, A. P., Jayatilaka, K. A. P. W., Pathirana, C., Mudduwa, L. K. B (2015). Antihyperglycemic activity of *Coccinia grandis* (L.) Voigt (Cucurbitaceae) in streptozotocin induced diabetic rats. Indian Journal of Traditional Knowledge, 14; 376-381.

Avato, P., & Argentieri, M. P. (2019). Quality assessment of commercial spagyric tinctures of Harpagophytum procumbens and their antioxidant properties. Molecules, 24; 2251.

Bárbara, M. S., Machado, C. S., Sodré, G. D. S., Dias, L. G., Estevinho, L. M., & De Carvalho, C. A. L. (2015). Microbiological assessment, nutritional characterization and phenolic compounds of bee pollen from Mellipona mandacaia Smith, 1983. Molecules, 20; 12525-12544.

Barnes, S., Benton, H. P., Casazza, K., Cooper, S. J., Cui, X., Du, X & Tiwari, H. K. (2016). Training in metabolomics research. I. Designing the experiment, collecting and extracting samples and generating metabolomics data. Journal of Mass Spectrometry, 51; 461-475.

Bento-Silva, A., Koistinen, V. M., Mena, P., Bronze, M. R., Hanhineva, K., Sahlstrøm, S & Aura, A. M. (2020). Factors affecting intake, metabolism and health benefits of phenolic acids: do we understand individual variability?. European journal of nutrition, 59; 1275-1293.

Bhalla, R., Narasimhan, K., & Swarup, S. (2005). Metabolomics and its role in understanding cellular responses in plants. Plant cell reports, 24; 562-571.

Bhattacharya, B., Samanta, M., Pal, P., Chakraborty, S., Samanta, A. (2010). In Vitro evaluation of antifungal and antibacterial activities of the plant *Coccinia grandis* (L.) Voigt. (FamilyCucurbitaceae). Journal of Phytology, 2; 52-57.

Biala, W & Jasinski, M. (2018). The phenylpropanoid case - It is transport that matters. Frontiers in Plant Science, 9; 1610.

Bianchi, F., Riboni, N., Termopoli, V., Mendez, L., Medina, I., Ilag, L., & Careri, M. (2018). MS-based analytical techniques: advances in spray-based methods and EI-LC-MS applications. Journal of analytical methods in chemistry, 1-24.

Bowles, D., Isayenkova, J., Lim, E. K., Poppenberger, B. (2005) Glycosyltransferases: managers of small molecules. Current Opinion in Plant Biology, 8; 254-63.

Bowne, J., Bacic, A., Tester, M., & Roessner, U. (2018). Abiotic stress and metabolomics. Annual Plant Reviews online, 61-85.

Burgess, K., Rankin, N., & Weidt, S. (2014). Metabolomics. Handbook of Pharmacogenomics and Stratified Medicine, 181-205.

Butler, Rhett A. (2019). "Total number of plant species by country". Mongabay.

Chandrasekara, A., & Shahidi, F. (2010). Content of insoluble bound phenolics in millets and their contribution to antioxidant capacity. Journal of agricultural and food chemistry, 58; 6706-6714.

Carvalho, L.C., Vidigal, P & Amâncio, S. (2015). Oxidative stress homeostasis in grapevine (*Vitis vinifera* L.). Frontiers in Environmental Science, 20.

Chanda, J., Mukherjee, P. K., Kar, A., Maitra, P. K., Singha, S., Halder, P. K., ... & Vishnuvardh, R. (2020). LC–QTOF–MS-based metabolite profiling and evaluation of α-glucosidase inhibitory kinetics of *Coccinia grandis* fruit. Biomedical Chromatography, 34; e4950.

Csurhes, S., & Edwards, R. (1998). National weeds program, potential environmental weeds in Australia, candidate species for preventative control. National Parks and Wildlife Biodiversity Group, Environment Australia, Canberra: 208.

Chandrasekara, A., & Shahidi, F. (2010). Content of insoluble bound phenolics in millets and their contribution to antioxidant capacity. Journal of agricultural and food chemistry, 58; 6706-6714.

Chen, Z., Zheng, S., Li, L., & Jiang, H. (2014). Metabolism of flavonoids in human: a comprehensive review. Current Drug Metabolism, 15; 48-61.

Cheynier V. Polyphenols in foods are more complex than often thought. American Journal of Clinical Nutrition, 81; 223-229.

Cheynier V (2012). Phenolic compounds: from plants to foods, 11; 153–177

Geetha, T. S., & Geetha, N. (2014). Phytochemical screening, quantitative analysis of primary and secondary metabolites of Cymbopogan citratus (DC) Stapf. leaves from Kodaikanal hills, Tamilnadu. International Journal of pharmtech research, 6; 521-529.

Clifford, M. N., Johnston, K. L., Knight, S., & Kuhnert, N. (2003). Hierarchical scheme for LC-MS n identification of chlorogenic acids. Journal of agricultural and food chemistry, 51; 2900-2911.

Clifford, M. N., Kirkpatrick, J. O., Kuhnert, N., Roozendaal, H., & Salgado, P. R. (2008). LC–MSn analysis of the cis isomers of chlorogenic acids. Food Chemistry, 106; 379-385.

Clish, C. B. (2015). Metabolomics: an emerging but powerful tool for precision medicine. Molecular Case Studies, 1; a000588.

Cooper, K. A., Campos-Giménez, E., Alvarez, D. J., Nagy, K., Donovan, J. L. and Williamson, G. (2007). Rapid reversed phase ultra-performance liquid chromatography analysis of the major cocoa polyphenols and inter-relationships of

their concentrations in chocolate. Journal of Agricultural and Food Chemistry, 55; 2841-2847.

Creek, D. J., Dunn, W. B., Fiehn, O., Griffin, J. L., Hall, R. D., Lei, Z., ... & Wolfender, J. L. (2014). Metabolite identification: are you sure? And how do your peers gauge your confidence?. Metabolomics, 10; 350-353.

Cuyckens, F., Ma, Y. L., Pocsfalvi, G., & Claeysi, M. (2000). Tandem mass spectral strategies for the structural characterization of flavonoid glycosides. Analusis, 28; 888-895.

Dai, J., & Mumper, R. J. (2010). Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. Molecules, 15; 7313-7352.

Das, R., & Pal, T. K. (2015). Method development & validation of LCMS/MS for atorvastatin and olmesartan in human plasma to trace drug Interaction of formulation. Current Pharmaceutical Analysis, 11; 43-52.

Deokate, U.A & Khadabadi, S. S. (2011). Pharmacology and phytochemistry of *Coccinia indica*. Review. Journal of Pharmacognosy and Phytotherapy 3; 155-159.

Deshpande, SV (2011). In vitro antioxidant study of petroleum ether chloroform and ethyl acetate fractions of *Coccinia grandis* stems extract. International journal of chemical sciences, 9; 80-86.

Dhar, A., Vekariya, R. L., & Bhadja, P. (2018). n-Alkane isomerization by catalysis—a method of industrial importance: An overview. Cogent Chemistry, 4; 1514686.

Działo, M., Mierziak, J., Korzun, U., Preisner, M., Szopa, J., & Kulma, A. (2016). The potential of plant phenolics in prevention and therapy of skin disorders. International journal of molecular sciences, 17; 160.

El-Seedi, H.R., Taher, E.A., Sheikh, B.Y., Anjum, S., Saeed, A., AlAjmi, M.F., Moustafa, M.S., Al-Mousawi, S.M., Farag, M.A., Hegazy, M.E.F. and Khalifa, S.A. (2-18). Hydroxycinnamic acids: Natural sources, biosynthesis, possible biological activities, and roles in Islamic medicine. In Studies in Natural Products Chemistry, 55; 269-292.

F, Ferreyra., M, Lorena., S, Rius & P, Casati. (2012). "Flavonoids: biosynthesis, biological functions, and biotechnological applications." Frontiers in plant science 3; 222.

Fernando, I. P. S., Ryu, B., Ahn, G., Yeo, I. K., & Jeon, Y. J. (2020). Therapeutic potential of algal natural products against metabolic syndrome: A review of recent developments. Trends in Food Science & Technology, 97; 286-299.

Fiehn, O. (2002). Metabolomics-the link between genotypes and phenotypes. Functional genomics, 155-171.

Fraser, C. M., & Chapple, C. (2011). The phenylpropanoid pathway in Arabidopsis. The Arabidopsis Book/American Society of Plant Biologists, 9; e0152.

Gbashi, S., Njobeh, P., Steenkamp, P., & Madala, N. (2017). Pressurized hot water extraction and chemometric fingerprinting of flavonoids from Bidens pilosa by UPLC-tandem mass spectrometry. CyTA-Journal of Food, 15; 171-180.

Giada, M. D. L. R. (2013). Food phenolic compounds: main classes, sources and their antioxidant power. Oxidative stress and chronic degenerative diseases—A role for antioxidants. InTech, 87-112.

M. P. Gonthier, C. Remesy, A. Scalbert, V. Cheynier, J.-M. Souquet, K. Poutanen, A.-M. Aura. (2006). Microbial metabolism of caffeic acid and its esters chlorogenic and caftaric acids by human faecal microbiota in vitro, Biomedicine & Pharmacotherapy, 60; 536-540.

Graefe EU, Wittig J, Mueller S, Riethling AK, Uehleke B, Drewelow B, Pforte H, Jacobasch G, Derendorf H, Veit M. (2001). Pharmacokinetics and bioavailability of quercetin glycosides in humans. Journal of Clinical Pharmacology, 41; 492–499.

Hall, R. D. (2006). Plant metabolomics: from holistic hope, to hype, to hot topic. New phytologist, 169; 453-468.

Kala, H. K., Mehta, R., Tandey, R., Sen, K. K., & Mandal, V. (2016). Ten years of research on phenolics (2005–2015): A status report. Pacific Science Review A: Natural Science and Engineering, 18; 1-4.

Harine SJ (2012). Ivy gourd- medicinal and nutritional values. International Journal of Current Research, 9; 47604-47607.

He, K., Echigo, S., Asada, Y., & Itoh, S. (2018). Determination of caffeine and its metabolites in wastewater treatment plants using solid-phase extraction and liquid chromatography–tandem mass spectrometry. Analytical Sciences, 34; 349-354.

Holstein, N and Renner, S. S (2011). A dated phylogeny and collection records reveal repeated biome shifts in the African genus *Coccinia* (Cucurbitaceae). BMC Evolutionary Biology 11; 28.

Holstein N. Monograph of *Coccinia* (Cucurbitaceae). (2015). PhytoKeys, 3; 1-166.

Hong, Y. S., Martinez, A., Liger-Belair, G., Jeandet, P., Nuzillard, J.-M., & Cilindre, C. (2012). Metabolomics reveals simultaneous influences of plant defence system and fungal growth in Botrytis cinerea infected Vitis vinifera cv. Chardonnay berries. Journal of Experimental Botany, 63; 5773–5785.

Ilkhanizadeh, B., Shirpoor, A., Nemati, S., & Rasmi, Y. (2016). Protective effects of ginger (Zingiber officinale) extract against diabetes-induced heart abnormality in rats. Diabetes & metabolism journal, 40; 46-53.

Inaishi, J & Saisho, Y. (2020). Beta cell mass in obesity and its relation to pancreas fat: A mini review. Nutrients, 12; 3846.

Irchhaiya, D. R. (2015). Anumalik yadav, Nitika Gupta, Swadesh Kumar, Nikhil Gupta, Santosh Kumar, Vinay Yadav, Anuj Prakash2 and Himanshu Gurjar. Metabolites in plants and its classification. World Journal of Pharmacy and Pharmaceutical Sciences, 4; 287-305.

Isah, T. (2019). Stress and defense responses in plant secondary metabolites production. Biological research, 52.

Iwai, K., Kim, M. Y., Onodera, A., & Matsue, H. (2006). α-Glucosidase inhibitory and antihyperglycemic effects of polyphenols in the fruit of *Viburnum dilatatum Thunb*. Journal of agricultural and food chemistry, 54; 4588-4592.

Jiang, B., Luo, M.J. & Zhang, Z.W. (2015). Influence of altitudes on phenolic compounds and antioxidant activities of Cabernet Sauvignon and Cabernet Franc wines in Loess Plateau region. Journal of Chinese Institute of Food Science and Technology, 2; 223-230.

P, Kachlicki., J, Einhorn., M, Stobiecki. (2008). Evaluation of glycosylation and malonylation patterns in flavonoid glycosides during LC/MS/MS metabolite profiling. Published in Journal of mass spectrometry, 43; 572-586.

Khan, M. S. I., & Kim, Y. J. (2020). Dielectric barrier discharge (DBD) plasma induced flavonoid degradation kinetics and mechanism in water. LWT, 118; 108777.

Khatun, S., F. Pervin, M.R. Karim, M. Ashraduzzaman and A. Rosma (2012). Phytochemical screening and antimicrobial activity of *Coccinia cordifolia* L. Plant. Pakistan Journal of Pharmaceutical Sciences, 25; 757-761.

Kim, H. K., & Verpoorte, R. (2010). Sample preparation for plant metabolomics. Phytochemical Analysis: An International Journal of Plant Chemical and Biochemical Techniques, 21; 4-13.

Kim, H.K., Choi, Y.H., Verpoorte, R. (2011). NMR-based plant metabolomics: where do we stand, where do we go? Trends Biotechnology, 29; 267-275.

Kondhare, D., & Lade, H. (2017). Phytochemical profile, aldose reductase inhibitory, and antioxidant activities of Indian traditional medicinal *Coccinia grandis* (L.) fruit extract. 3 Biotech, 7; 1-10.

Korfmacher, W.A. (2005). Foundation review: Principles and applications of LC-MS in new drug discovery. Drug Discovery Today, 10; 1357-1367.

Kortbeek, R. W., van der Gragt, M., & Bleeker, P. M. (2019). Endogenous plant metabolites against insects. European Journal of Plant Pathology, 154; 67-90.

Kozlowska, A., & Szostak-Wegierek, D. (2014). Flavonoids-food sources and health benefits. Roczniki Państwowego Zakładu Higieny, 65.

Kudjordjie, E. N., Sapkota, R., & Nicolaisen, M. (2020). Arabidopsis sculpt distinct root-associated microbiomes through the synthesis of secondary metabolites and defense signaling molecules.

Kumar, S., & Pandey, A. K. (2013). Chemistry and biological activities of flavonoids: an overview. The scientific world journal.

Kulkarni SG, Vijayanand P (2012). Effect of Pretreatments on quality characteristics of dehydrated Ivy Gourd (*Coccinia indica* L.). Food and Bioprocess Technology, 5; 593-600.

Kumar, S & Pandey A, K. (2013). Chemistry and biological activities of flavonoids: an overview. Scientific World Journal, 162750.

Kumar, S., Singh, A., & Kumar, B. (2017). Identification and characterization of phenolics and terpenoids from ethanolic extracts of Phyllanthus species by HPLC-ESI-QTOF-MS/MS. Journal of pharmaceutical analysis, 7; 214-222.

Kumar, N., & Goel, N. (2019). Phenolic acids: Natural versatile molecules with promising therapeutic applications. Biotechnology Reports, 24; e00370.

Lamichhane, S., Sen, P., Dickens, A. M., Hyötyläinen, T., & Orešič, M. (2018). An overview of metabolomics data analysis: current tools and future perspectives. Comprehensive analytical chemistry, 82; 387-413.

Lattanzio, V. (2013). Phenolic compounds: Introduction. In book: Natural products, 1543-1580.

Lee, T., Ho, Y.S., Yeo, H.C., Lin, J.P.Y., Lee, D.Y. (2013). Precursor mass prediction by clustering ionization products in LC-MS based metabolomics. Metabolomics, 9; 1301-1310.

Leffers, A., 2003: Gemsbok Bean and Kalahari Truffle - Traditional Plant use by Jul'hoansi in NorthEastern Namibia. Windhoek: Gamsberg MacMillan; Namibia.
.
Le Roy, J., Huss, B., Creach, A., Hawkins, S., & Neutelings, G. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. Frontiers in plant science, 7, 735.

Lewis, G. D., Asnani, A., & Gerszten, R. E. (2008). Application of metabolomics to cardiovascular biomarker and pathway discovery. Journal of the American College of Cardiology, 52; 117-123.

Lin, D., Xiao, M., Zhao, J., Li, Z., Xing, B., Li, X & Chen, S. (2016). An overview of plant phenolic compounds and their importance in human nutrition and management of type 2 diabetes. Molecules, 21; 1374.

Madala N.E, Tugizimana, F, Steenkamp P.A, Piater L.A. and Dubery I.A (2013). The short and long of it: shorter chromatographic analysis suffice for sample classification during UHPLC-MS-based metabolic fingerprinting. Chromatographia, 76; 279-285.

Makita C., Chimuka L., Steenkamp P., Cukrowska E., Madala E. (2016). Comparative analyses of flavonoid content in Moringa oleifera and Moringa ovalifolia with the aid of UHPLC-qTOF-MS fingerprinting South African Journal of Botany, 105; 116-122.

Masike, K., & Madala, N. (2018). Synchronized survey scan approach allows for efficient discrimination of isomeric and isobaric compounds during LC-MS/MS analyses. Journal of analytical methods in chemistry, 2018.

Matthews M. L., Endress P. K. (2004) Comparative floral structure and systematics in Cucurbitales (Corynocarpaceae, Coriariaceae, Tetramelaceae, Datiscaceae, Begoniaceae, Cucurbitaceae, Anisophylleaceae). Botanical Journal of the Linnean Society, 145; 129–185.

Maroyi, A. (2013). Traditional use of medicinal plants in south-central Zimbabwe: review and perspectives. Journal of ethnobiology and ethnomedicine, 9; 1-18.

Martis, Elvis A; Ahire, Deepak C; Singh, Ruchi O. (2011). Metabolomics in drug discovery: A Review, International Journal of Pharmacy and Pharmaceutical Science, 1; 67-74.

Martín, J. F., García-Estrada, C., & Zeilinger, S. (Eds.). (2014). Biosynthesis and molecular genetics of fungal secondary metabolites. Springer.

Masike, K.; Khoza, B.S.; Steenkamp, P.A.; Smit, E.; Dubery, I.A.; Madala, N.E (2017). A metabolomics-guided exploration of the phytochemical constituents of *Vernonia fastigiata* with the aid of pressurized hot water extraction and Liquid Chromatography-Mass Spectrometry. Molecules, 22; 1200.

Mazumdar, S., Akter, R., & Talukder, D. (2015). Antidiabetic and antidiarrhoeal effects on ethanolic extract of Psidium guajava (L.) Bat. leaves in Wister rats. Asian Pacific Journal of Tropical Biomedicine, 5; 10-14.

Moco S, Vervoort J, Bino RJ, De Vos RCH, Bino R (2007). Metabolomics technologies and metabolite identification. Trends Analytical Chemistry, 26; 855-866.

Moore, B. D., Andrew, R. L., Külheim, C., & Foley, W. J. (2014). Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phytologist, 201; 733-750.

Mullen, W., Edwards, C. A., & Crozier, A. (2006). Absorption, excretion and metabolite profiling of methyl-, glucuronyl-, glucosyl-and sulpho-conjugates of quercetin in human plasma and urine after ingestion of onions. British journal of nutrition, 96; 107-116.

Mushtaq, M. Y., Choi, Y. H., Verpoorte, R., & Wilson, E. G. (2014). Extraction for metabolomics: access to the metabolome. Phytochemical Analysis, 25; 291-306.

Muth D, Marsden-Edwards E, Kachlicki P, Stobiecki M (2008). Differentiation of isomeric malonylated flavonoid glyconjugates in plant extracts with UPLC-ESI/MS/MS. Phytochemical Analysis, 5; 444-452.

Ncube, E. N., Mhlongo, M. I., Piater, L. A., Steenkamp, P. A., Dubery, I. A., & Madala, N. E. (2014). Analyses of chlorogenic acids and related cinnamic acid derivatives from Nicotiana tabacum tissues with the aid of UPLC-QTOF-MS/MS based on the in-source collision-induced dissociation method. Chemistry Central Journal, 8; 1-10.

Nengovhela, N., Steenkamp, P. A., & Madala, N. E. (2020). LC-MS Based Metabolite Fingerprinting of Coccinia Plants Reveals Glycoisomerization as a Structual Diversification Strategy in Flavanoid Chemical Space. National Academy Science Letters, 1-5.

Negri, G & Tabach, R (2013). Saponins, tannins and flavonols found in hydroethanolic extract from Periandra dulcis roots. Rev Bras Farmacogn 23; 851-860.

Nobela, O. N., Bansal, R. C., & Justo, J. J. (2019). A review of power quality compatibility of wind energy conversion systems with the South African utility grid. Renewable Energy Focus, 31; 63-72.

Ochuko Erukainure, Sanni Olakunle, Shahidul Islam. (2018). Clerodendrum volubile: Phenolics and Applications to Healt. In book: polyphenols: mechanisms of action in human health and disease, 53-68.

Y Ogo, T Mori, R Nakabayashi, K Saito and F Takaiwa (2016). Transgenic rice seed expressing flavonoid biosynthetic genes accumulate glycosylated and/or acylated flavonoids in protein bodies. Journal of Experimental Botany, 67; 95–106.

Panche, Archana & Diwan, Arvind & Chandra, S. (2016). Flavonoids: An overview. Journal of Nutritional Science, 5; 47.

Pandey, K. B., & Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. Oxidative medicine and cellular longevity, 2; 270-278.

Pekamwar S. S, Kalyankar T.M, and Kokate S.S. (2013). Pharmacological activities of *Coccinia Grandis*: Review Journal of Applied Pharmaceutical Science, 3; 114-119.

Pieterse, C. M., Leon-Reyes, A., Van der Ent, S., & Van Wees, S. C. (2009). Networking by small-molecule hormones in plant immunity. Nature chemical biology, 5; 308-316.

G Pinto and A Pollio (2011). “Plant polyphenols and their anti-cariogenic properties,” Molecules, 16; 1486-1507.

Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J., & Wishart, D. (2019). Systems biology and multi-omics integration: viewpoints from the metabolomics research community. Metabolites, 9; 76.

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC bioinformatics, 11; 1-11.

Pragasam, Samuel & Murunikkara, Vachana & Sabina, Evan & Rasool, Mahaboob khan. (2012). Ameliorative effect of p-coumaric acid, a common dietary phenol, on adjuvant-induced arthritis in rats. Rheumatology international, 33; 325-334.

Premkumar, Louis. (2014). Fascinating Facts about Phytonutrients in Spices and Healthy Food: Scientifically proven facts. Xlibris Corporation.

Pyne, M. E., Narcross, L., & Martin, V. J. (2019). Engineering plant secondary metabolism in microbial systems. Plant physiology, 179; 844-861.

Rai, A., Umashankar, S., Rai, M., Kiat, L. B., Bing, J. A. S., & Swarup, S. (2016). Coordinate regulation of metabolite glycosylation and stress hormone biosynthesis by TT8 in Arabidopsis. Plant physiology, 171; 2499-2515.

Ramabulana T, Mavunda RD, Steenkamp PA, Piater LA, Dubery IA., Madala NE (2015). Secondary metabolite perturbations in Phaseolus vulgaris leaves due to gamma radiation. Plant Physiology Biochemistry, 97; 287-295.

Ramabulana, A. T., Steenkamp, P., Madala, N., & Dubery, I. A. (2020). Profiling of chlorogenic acids from bidens pilosa and differentiation of closely related positional isomers with the aid of UHPLC-QTOF-MS/MS-based in-source collision-induced dissociation. Metabolites, 10; 178.

Rasool Hassan, B. A. (2012). Medicinal plants (importance and uses). Pharmaceut Anal Acta, 3; 2153-2435.

Renner, SS, Schaefer, H & Kocyan, A (2007), 'Phylogenetics of Cucumis (Cucurbitaceae): Cucumber (C. sativus) belongs in an Asian/Australian clade far from melon (C. melo)', BMC Evolutionary Biology, 7; 58–69.

Rehab A. Hussein and Amira A. El-Anssary, (2018). Plants Secondary Metabolites: The Key Drivers of the Pharmacological Actions of Medicinal Plants. Herbal Medicine, 1; 13.

Retief, E. & Meyer, N.L. (2017). Plants of the Free State inventory and identification guide Strelitzia. National Botanical Institute, Pretoria, 38; 1-1236.

Rupasinghe, H. V., Ronalds, C. M., Rathgeber, B., & Robinson, R. A. (2010). Absorption and tissue distribution of dietary quercetin and quercetin glycosides of apple skin in broiler chickens. Journal of the Science of Food and Agriculture, 90; 1172-1178.

Salek, R. M., Steinbeck, C., Viant, M. R., Goodacre, R., & Dunn, W. B. (2013). The role of reporting standards for metabolite annotation and identification in metabolomic studies. GigaScience, 2, 2047-217X.

Sardans, J., Penuelas, J., & Rivas-Ubach, A. (2011). Ecological metabolomics: overview of current developments and future challenges. Chemoecology, 21; 191-225.

Sarkar, D & Shetty, K. (2014). Metabolic stimulation of plant phenolics for food preservation and health. Annual review of food science and technology, 5; 395-413.

Scalbert, A., Morand, C., Manach, C., & Rémésy, C. (2002). Absorption and metabolism of polyphenols in the gut and impact on health. Biomedicine & Pharmacotherapy, 56; 276-282.

Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., & Wopereis, S. (2009). Mass-spectrometry-based metabolomics: limitations

and recommendations for future progress with particular focus on nutrition research. Metabolomics, 5; 435-458.

Schmidt, M & Zizka, G. (2014). Plant species associated with different levels of species richness and of vegetation cover as indicators of desertification in Burkina Faso (West Africa). Flora et Vegetatio Sudano-Sambesica, 17; 3-8.

Sharma, N., Sharma, M., & Bindal, M. C. (2010). Potential antidiabetic herbal drugs: A comparative review of marketed products. Research Journal of Pharmacognosy and Phytochemistry, 2; 115-121.

Schaefer, H., & Renner, S. S. (2011). Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). Taxon, 60; 122-138.

Seyler, L., Kujawinski, E. B., Azua-Bustos, A., Lee, M. D., Marlow, J., Perl, S. M., & Cleaves II, H. J. (2020). Metabolomics as an emerging tool in the search for astrobiologically relevant biomarkers. Astrobiology, 20; 1251-1261.

Shetty K, Wahlqvist ML (2004). A model for the role of the proline-linked pentose-phosphate pathway in phenolic phytochemical bio-synthesis and mechanism of action for human health and environmental applications. Asia Pacific Clinical Nutrition, 13; 1-24.

Sigidi, M. T., Traoré, A. N., Boukandou, M. M., Tshisikhawe, M. P., Ntuli, S. S., & Potgieter, N. (2017). Anti-HIV, pro-inflammatory and cytotoxicity properties of selected Venda plants.

Sova, M., and Saso, L. (2020). Natural sources, pharmacokinetics, biological activities and health benefits of hydroxycinnamic acids and their metabolites. Nutrients 12; 1–30.

Sparkman, O. D. (2003). Mass spectrometry basics Christopher G. Herbert and Robert AW Johnstone. Journal of the American Society for Mass Spectrometry, 14; 925.

Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E. A., Glass, C. K & Subramaniam, S. (2007). Lmsd: Lipid maps structure database. *Nucleic acids research*, 35; 527-532.

Starr, F., Starr, K., & Loope, L. (2003). *Coccinia grandis*. Ivy gourd Cucurbitaceae Biological Resources Division Haleakala Field Station, Maui, Hawail, 2; 1-6.

Sundaram R, Samuel JI, Palavesam S (2012). In vitro anti-plasmodial activity of ethanolic extracts of South Indian Medicinal plants against Plasmodium falciparum. Asian Pacific Journal of Tropical Disease, 2; 180-183.

Sumner, L. W., Mendes, P., & Dixon, R. A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry, 62; 817-836.

Thilakarathna, S. H., & Rupasinghe, H. V. (2012). Anti-atherosclerotic effects of fruit bioactive compounds: A review of current scientific evidence. Canadian Journal of Plant Science, 92; 407-419.

Niranjan, S., Ranju, G., Sharma, U. K., Sharma, U. S., Jana, G. K., & Abhishek, S. (2010). Evaluation of anti-inflammatory activity of *Coccinia indica* leaves extracts. Journal of Pharmacy Research, 3; 2172-2173.

Tamilselvan, N., Thirumalai, T., Elumalai, E. K., Balaji, R., & David, E. (2011). Pharmacognosy of Coccinia grandis: A review. Asian Pacific Journal of Tropical Biomedicine, 1; 299– 302.

Teixeira, J., Gaspar, A., Garrido, E. M., Garrido, J., & Borges, F. (2013). Hydroxycinnamic acid antioxidants: an electrochemical overview. BioMed research international, 251754.

Tiwari, Pragya, Rajender Singh Sangwan, and Neelam S. Sangwan. (2016). "Plant secondary metabolism linked glycosyltransferases: an update on expanding knowledge and scopes." Biotechnology Advances 34; 714-739.

Tiwari, P., Sangwan, R. S., & Sangwan, N. S. (2016). Plant secondary metabolism linked glycosyltransferases: an update on expanding knowledge and scopes. Biotechnology Advances, 34; 714-739.

Thilakarathna, S. H & Rupasinghe, H. P. (2013). Flavonoid bioavailability and attempts for bioavailability enhancement. Nutrients, 5; 3367-3387.

Thirumurugan, D., Cholarajan, A., Raja, S. S., & Vijayakumar, R. (2018). An introductory chapter: secondary metabolites. Second metab-sources Appl, 1-21.

Tugizimana F, Piater L, Dubery I (2013). Plant metabolomics: A new frontier in phytochemical analysis. South African Journal of Science, 109; 1-11.

Tugizimana, F., Steenkamp, P. A., Piater, L. A., & Dubery, I. A. (2016). A conversation on data mining strategies in LC-MS untargeted metabolomics: Pre-processing and pre-treatment steps. Metabolites, 6; 40.

Tshabalala, T., Ncube, B., Madala, N. E., Nyakudya, T. T., Moyo, H. P., Sibanda, M., & Ndhlala, A. R. (2019). Scribbling the cat: a case of the “miracle” plant, Moringa oleifera. Plants, 8; 510.

Uddin, N., Hasan, M. R., Hossain, M. M., Hasan, M. M., Roy, A., Islam, T & Rana, M. S. (2014). Antioxidant, brine shrimp lethality and antimicrobial activities of methanol and ethyl-acetate extracts of Citrus macroptera Montr. fruit using in vitro assay models. Journal of Pharmaceutical Research International, 1725-1738.

Umamaheswari, M., & Chatterjee, T. K. (2008). In vitro antioxidant activities of the fractions of *Coccinia grandis* L. leaf extract. African Journal of Traditional, Complementary and Alternative Medicines, 5; 61-73.

Vaistij, F. E., Lim, E. K., Edwards, R., & Bowles, D. J. (2009). Glycosylation of secondary metabolites and xenobiotics. In Plant-derived natural products. Springer, New York, NY, 209-228.

Van der Walt, A. M., Loots, D. T., Ibrahim, M. I. M., & Bezuidenhout, C. C. (2009). Minerals, trace elements and antioxidant phytochemicals in wild African dark-green leafy vegetables (morogo). South African Journal of Science, 105; 444-448.

Viant, M. R., Kurland, I. J., Jones, M. R., & Dunn, W. B. (2017). How close are we to complete annotation of metabolomes?. Current opinion in chemical biology, 36; 64-69.

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics, 7;1-15.

Vaishali M.P & Neeraj M. (2018) Chapter 12 - Anticancer potential of flavonoids: chemistry, biological activities, and future perspectives studies in natural products chemistry, 59; 401-430.

Veitch, N. C., & Grayer, R. J. (2011). Flavonoids and their glycosides, including anthocyanins. *Natural product reports*, *28*; 1626-1695.

Verpoorte, R., Choi, Y. H., Mustafa, N. R., & Kim, H. K. (2008). Metabolomics: back to basics. Phytochemistry Reviews, 7; 525-537.

Vetere A, Choudhary A, Burns SM, Wagner BK (2014). Targeting the pancreatic β-cell to treat diabetes. Nature Reviews Drug Discovery, 14; 278-289.

Vinholes, J., Silva, B. M., & Silva, L. R. (2015). Hydroxycinnamic acids (HCAS): Structure, biological properties and health effects. Advances in Medicine and Biology, 88; 105-130.

Vogt, T. (2010). Phenylpropanoid biosynthesis. Molecular plant, 3; 2-20.

Wang, S., Alseekh, S., Fernie, A. R., & Luo, J. (2019). The structure and function of major plant metabolite modifications. Molecular plant, 12; 899-919.

Wannenmacher, J., Gastl, M., & Becker, T. (2018). Phenolic substances in beer: Structural diversity, reactive potential and relevance for brewing process and beer quality. Comprehensive Reviews in Food Science and Food Safety, 17; 953-988.

Watson, D. G. (2013). A rough guide to metabolite identification using high resolution liquid chromatography mass spectrometry in metabolomic profiling in metazoans. Computational and structural biotechnology journal, 4, e201301005.

Watson, R. R., & Preedy, V. R. (Eds.). (2019). Bioactive food as dietary interventions for diabetes. Academic Press.

WHO. Global prevalence of vitamin A deficiency in populations at risk 1995–2005. WHO global database on vitamin A deficiency. Geneva, World Health Organization.

Wink, M. (2010). Introduction: biochemistry, physiology and ecological functions of secondary metabolites. Annual plant reviews 40: Biochemistry of plant secondary metabolism, 1-19.

Wilson, A. E., & Tian, L. (2019). Phylogenomic analysis of UDP-dependent glycosyltransferases provides insights into the evolutionary landscape of glycosylation in plant metabolism. The Plant Journal, 100; 1273-1288.

Hejun Wu and Saiyan Chen (2016). An overview of plant phenolic compounds and their mportance in human nutrition and management of Type 2 diabetes. molecules, 21; 1374.

Xu, Y., Fan, M., Ran, J., Zhang, T., Sun, H., Dong, M & Zheng, H. (2016). Variation in phenolic compounds and antioxidant activity in apple seeds of seven cultivars. Saudi Journal of Biological Sciences, 23; 379-388.

Yang, G., Fu, Y., Malakhova, M., Kurinov, I., Zhu, F., Yao, K & Dong, Z. (2014). Caffeic acid directly targets ERK1/2 to attenuate solar UV-induced skin carcinogenesis. Cancer Prevention Research, 7; 1056-1066.

Yang, C. S., Landau, J. M., Huang, M. T., & Newmark, H. L. (2001). Inhibition of carcinogenesis by dietary polyphenolic compounds. Annual review of nutrition, 21, 381-406.

Yuliana, N. D., Jahangir, M., Verpoorte, R., & Choi, Y. H. (2013). Metabolomics for the rapid dereplication of bioactive compounds from natural sources. Phytochemistry reviews, 12; 293-304.

Li-Bing Zhang, Mark P. Simmons, Alexander Kocyan, Susanne S. Renner (2006). Phylogeny of the Cucurbitales based on DNA sequences of nine loci from three genomes: Implications for morphological and sexual system evolution. Molecular Phylogenetics and Evolution, 39; 305–322

Zhang, Hua & Tsao, Rong. (2016). Dietary polyphenols, oxidative stress and antioxidant and anti-inflammatory effects. Current Opinion in Food Science. 8; 33-42.

Zhou, B., Xiao, J. F., Tuli, L., & Ressom, H. W. (2012). LC-MS-based metabolomics. Molecular BioSystems, 8; 470-481.

Chapter 3: LC-MS Based Metabolite Fingerprinting of *Coccinia* Plants Reveals Glyco-isomerization as a Structural Diversification Strategy in Flavonoid Chemical Space

Abstract

Flavonoids are ubiquitous plant metabolites with many purported pharmacological properties. Emergence of advanced analytical instrumentations such as liquid chromatography combined with mass spectrometry (Liquid chromatography mass spectrometry) has allowed for unparalleled elucidation of chemical compositions of various plants. Through such analyses, various chemical modifications such as glycosylation and other forms of isomerization have been shown to be an evolutionary strategy of plants to diversify their metabolite compositions (metabolome). Herein, metabolite fingerprinting of two closely related plant species Vis *Coccinia grandis* and *Coccinia rehmannii* revealed that flavonoids attached to either *di*- or *tri*-saccharides are prone to isomerization, a phenomenon termed glyco-isomerization herein. Furthermore, modification known as sugar acylation by known biologically active metabolites such as derivatives of cinnamic acids (i. e., caffeic acid and coumaric acid) has been noted in *C. rehmannii* but not in *C. grandis*. The findings of the current study reveal glycosylation patterns of flavonoids as an evolutionary strategy used by plants to diversify its metabolome, a phenomenon of which the biological consequence is still to be investigated. Several secondary metabolites were identified within the two species, however, the ones which stood out and were of great interest are the ones which displayed glyco-isomerization.

3.1. Introduction

Flavonoids are biologically active compounds that exist as either free aglycone or glycosides attached to various forms of sugar molecules (Xiao *et al*., 2016; Day *et al*., 2000). The glycosylation of these flavonoids can be either through *O*-glycosides or C-glycosides. Interestingly, the attachment of these sugar molecules has been shown to have biological effect (Negri and Tabach, 2013). However, most phytochemical data have shown flavonoids glycosylation to be a complex biological phenomenon, which is also prone to further chemical modification such as isomerization, acetylation, malonylation, and acylation, thereby complicating the already complex metabolome (Ogo *et al*., 2016). Elsewhere, attachment (through acylation) of other biologically active molecules such as cinnamic acids are responsible for the production of isomeric compounds (Muth *et al*., 2008; Kachliki *et al*., 2008). As such, the presence of these modifications has poised an undisputed analytical challenge with most of these isomers being identified as mere structural artefacts. Currently, instrumentation such as LC–MS has been shown to be useful in deciphering these isomerization (Makita *et al*., 2017).

Liquid chromatography mass spectrometry (LC-MS) is combined with online platforms such as XCMS for the annotation and characterization of metabolites. XCMS online is a cloud-based processing platform of chromatography mass spectrometry which is considered the most efficient tool for pre-processing of MS-data. It enables processing and visualization of untargeted metabolomics data, and provides a platform for multi-groups comparison with the aid of visual statistical interactive tools (Wikoff *et al*., 2008; Wikoff *et al*., 2009). The XCMS platform also combines both univariate and multivariate analysis of metabolites through interactive tools such as PCA and cloud plots (Hao *et al*., 2018). The cloud plot data is expressed in the form of bubbles (metabolic features) based on characteristics such as p-value (significance level), fold change, m/z ratio, retention time (Rt) and the signal intensity for each feature (Gowda *et al*., 2014). The characterization and structural identification of metabolite features in metabolomics experiments is facilitated by integration of cloud plot to the METLIN database. From this, users are able to interrogate metabolic matches based on accurate *m/z* measurements with just a “mouse click”. Using a mouse to scroll over the metabolite feature (bubble), a user can visualize properties of the feature such as

box and whisker plots, extracted ion chromatograms (EICs) spectrum and METLIN database matches which appear on the main panel (Tautenhahn *et al*., 2012; Smith *et al*., 2005).

For large, complex and high dimensional datasets, multivariate statistical analysis are often used. Statistical models such as principal component analysis (PCA) which is a mathematical model mostly used for metabolomics studies (Lamichhane *et al*., 2018; Barnes *et al*., 2016). PCA model is an unsupervised (require no user intervention) statistical model which highlights the similarities and differences between samples, caused by variability in the data (Cook & Rutan, 2014). PCA is used to reduce the dimensionality of a wide range of datasets by converting the large data into smaller variables that contains most of the information in the larger dataset (Jaadi, 2019). Smaller datasets are easy to visualize and explore, which makes it easier and faster to analyse. The idea of PCA is therefore to reduce the number of variables (dimensionality) in a dataset, while retaining as relevant information as much as possible (Jolliffe & Cadima, 2016).

PCA transforms a set of correlated variables into new sets of uncorrelated variables called principal components (PCs). The principal components are obtained in the order of decreasing importance, with each PC subsequently explaining less of the total variation in the dataset (Jaadi, 2019). Thus, the first principal component (PC1) describes the maximum variation in the dataset, and the second principal component (PC2) describes the variation that is not explained by the first principal component (Jansen *et al*., 2010). The dimension of the data is decreased through this and the PCA converts the complex dataset into simple graphical representations called score plots, where samples are grouped according to similarities or differences (Xia *et al*., 2015; Barnes *et al*., 2016). The underlying cause of different groupings on a PCA model can be further explored using other models such as loadings plot. The two models (scores and loadings plot) complement each other because the position of a sample group in a given direction in a score plot is influenced by variables lying in the same direction in the loadings plot. Thus, metabolites that are either up- or down-regulated in the loadings plot can be selected and further exploited to elucidate their significance and identity (Madala *et al*., 2014; Masike *et al*., 2017). The aim of this

study was to study the distribution pattern of metabolites across two *Coccinia* species and to evaluate their metabolome complexity through chemical diversity.

3.2. Materials and method

3.2.1. Metabolite extraction

To showcase the significance of chemical modifications, herein. LC-MS based analyses of 80% methanolic extract of two *Coccinia* plants Vis *C. grandis* and *C. rehmannii* revealed they contain structurally similar flavonoids with very similar aglycones core moieties (quercetin and kaempferol). Herein, we also show that the two plants have different number of flavonoids, this owing to the glycosylation and acylation patterns of these flavonoids. Both plants were collected around villages in Venda outside of Thohoyandou, with assistance from Dr Khathutshelo Magwede (Department of Zoology, University of Venda). Extraction of metabolites was accomplished with a method proposed by (Makita *et al*., 2016). Briefly, 2 g of ground sample was used for extraction by mixing with 20 mL of 80% methanol. The debris was removed by centrifuging at 5000xg for 20 min, followed by drying using a rotary evaporator and reconstitution using 1mL of 50% methanol.

3.2.2. LC-MS analysis

LC-MS analysis was conducted using a method previously shown elsewhere (Masike *et al*., 2017). Here, a volume of 2 µL was injected on a UPLC instrument joined to a SYNAPT G1 HDMS. The chromatographic separation was done utilising a Waters HSS T3 C18 column (150 mm x 2.1 mm, 1.8 µm) and the column temperature controlled at 60 °C. Herein, a binary solvent mixture was used consisting of water (Eluent A) and acetonitrile (Eluent B), both containing 10 mM formic acid. A gradient method starting from 2 % (B) to 98 % B (B) was used at constant flow rate of 0.4 mL/min0. The analyte elations were monitored using a mass spectrometry detector. Here the MS conditions were as follows: ESI negative modes, capillary voltage of 2.5 kV, the sampling cone at 30 V and the extraction cone at 5.0 V, source temperature was 120 °C and the desolvation temperature was set at 450 °C. Nitrogen gas was used as the nebulisation gas at a flow rate of 550 L/h and cone gas was set at 50 L/h. Leucine enkephalin (50 pg/mL) was used as reference calibrant to obtain typical mass

accuracies. MassLynx 4.1 (SCN 872) software was used to analyse and visualize the data acquired from the LC-MS.

3.2.3. XCMS analysis

Further analyses was accomplished with the aid of interactive XCMS Online (https://xcmsonline.scripps.edu). The web interface allows users to either upload data sets using a java applet or select pre-uploaded data sets on XCMS Online. Following the upload of raw data files, users can select and customize parameters depending on the instrument platform in which the data were acquired. The parameters are defined for statistical analysis (paired/unpaired) based on the type of experiment and data. The uploaded data files are then processed for peak detection, retention-time correction, chromatogram alignment, metabolite feature annotation and statistical evaluation (Gowda *et al*., 2014). Metabolite features were defined as ions with unique m/z and retention time values and putative identification was accomplished through METLIN standard database matching. Parameter settings for XCMS processing of our demonstration data acquired by HILIC were as follows: centWave for feature detection (Δ *m/z* = 10 ppm, minimum peak width = 5 s, and maximum peak width = 20 s); obiwarp settings for retention-time correction (profStep = 1); and parameters for chromatogram alignment, including *mz*wid = 0.015, minfrac = 0.5, and bw = 5. The results output included extracted ion chromatographs (EICs), boxplots, cloud plots and PCA, which were exported directly from XCMS Online.

3.3. Results and discussion

3.3.1. LC-MS analysis

In this study, metabolite profiling in leaf extracts of two *Coccinia* species was studied using a high-throughput analytical method known as UPLC-MS. The difference in intensities of metabolites between the two species was clearly indicated by the base peak intensity (BPI) chromatograms in figure 1. It was expected that the distribution of secondary metabolites will vary due to their geographical origins, as the two species were collected from various places. A total of 23 (Table 3.1) metabolites were detected from the chromatogram of methanolic extracts of *C. grandis & C. rehmannii* (Figure

3.1). Some of detected metabolites were identified to be hydroxycinnamic acid and flavonoid isomers, of which some were found to be attached to hydroxycinnamic acid. The latter is further discussed in section 3.3.3.

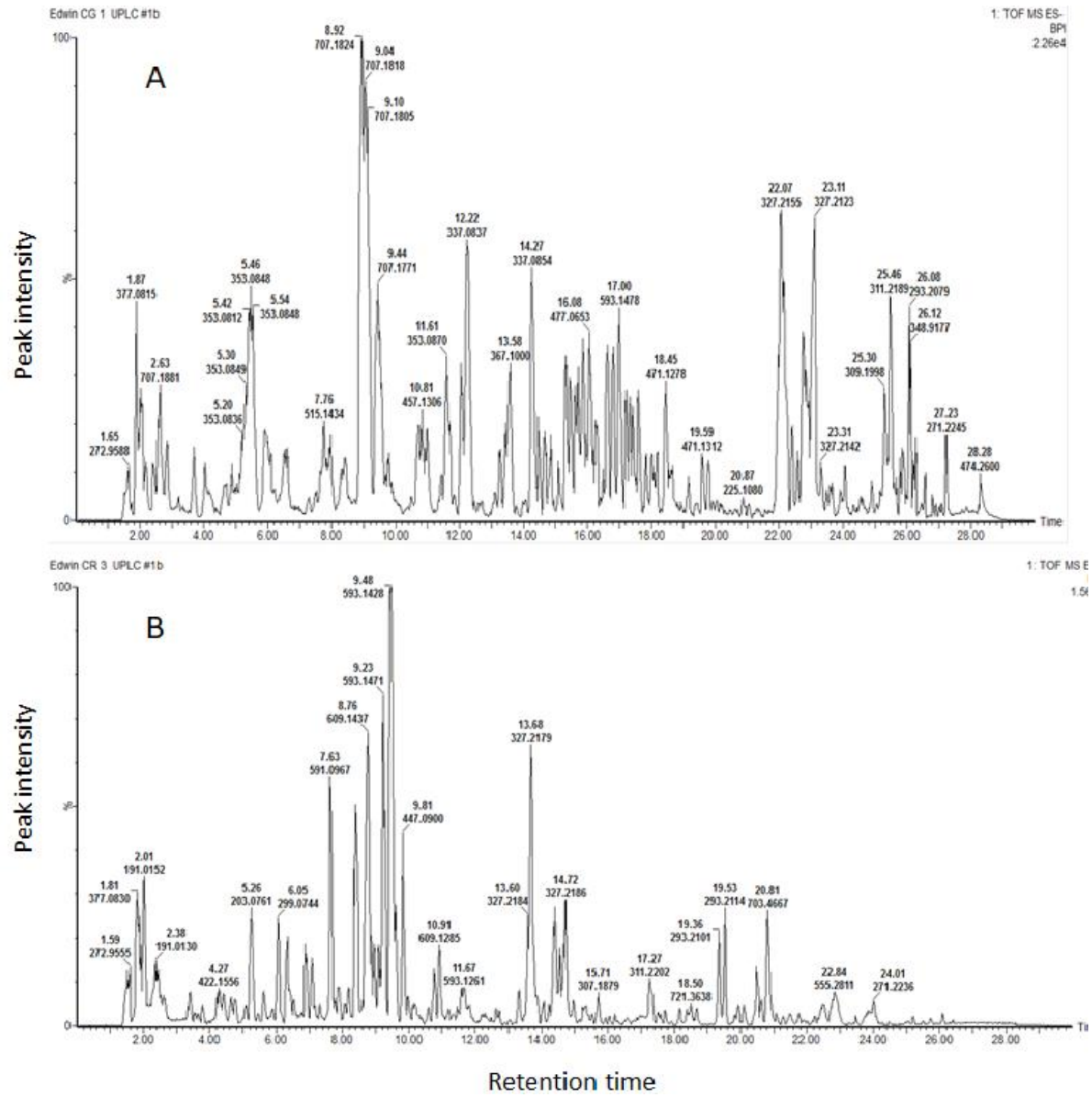


Figure 3.1: Base peak intensity chromatograms showing distribution of metabolites within two *Coccinia* species (A) *Coccinia grandis* (B) *Coccinia rehmannii.*

3.3.2. Multivariate statistical analysis of phytochemicals from leaf extracts of *Coccinia* plants

Liquid chromatography mass spectrometry enables visual description of the differences in metabolic composition between samples of different plants. However, the phytochemical composition and metabolite distribution across different parts of the

plant can be facilitated by univariate and multivariate data modelling, particularly PCA and interactive cloud plots (Ncube, 2016). In this study a comparative study to facilitate determination of phytochemical composition and metabolite distribution pattern between the leaf extracts of the two *Coccinia* species (*C. grandis* & *C. rehmannii*), was performed using PCA. The generated multidimensional data from PCA allowed for assessment and detection of natural groupings and natural separation from the score plots (Figure 3.2A), showing metabolite distribution variation across the leaves of the two plants. The PCA plots revealed differential clustering of samples into two distinct groups, group 1 (*C. grandis*-CG) represented by red colour and group 2 (*C.rehmannii*-CR) represented by blue colour. From the score plots (Figure 3.2A), PC1 accounts for (61%) variation along the x-axis and reveals differences between sample groups. PC2 accounts for (11%) variation along the y-axis and shows differences within the samples.

In total, the model explains 72% of the variation between the two plants showing the robustness of this model. Samples from the *C. grandis* are revealed to be clustered together which indicates that these samples share common metabolite profiles. *C. rehmannii* samples can be seen further away from samples of *C. grandis* and separated from each other which suggests that samples in group 2 (*C. rehmannii*) contain chemical composition that are different form each other and from those in group 1 (*C. grandis*). The grouping of CG is because they were collected from a single area but CR were collected from various places. Thus, it can be concluded from these observations that PCA revealed specific phytochemical distribution within the two *Coccinia* plants, suggesting that metabolic profiles across the two plants differs. To further investigate mass ions associated with distribution patterns depicted by the PCA, another model called loadings plot was evaluated. The loadings plot (Figure 3.2B) reveals the relationship between metabolite features that relate to the sample groupings observed in the score plot. The loadings plot helps to determine metabolite features whose levels are significantly different between different conditions. From these investigations, 23 metabolites (Table 3.1) comprising HCA derivatives and flavonoids were subsequently annotated based on their chromatographic elution order, mass fragmentation and comparison with already published data. The annotation of the metabolites was also carried out with the aid of interactive cloud plot directly linked to the METLIN database.

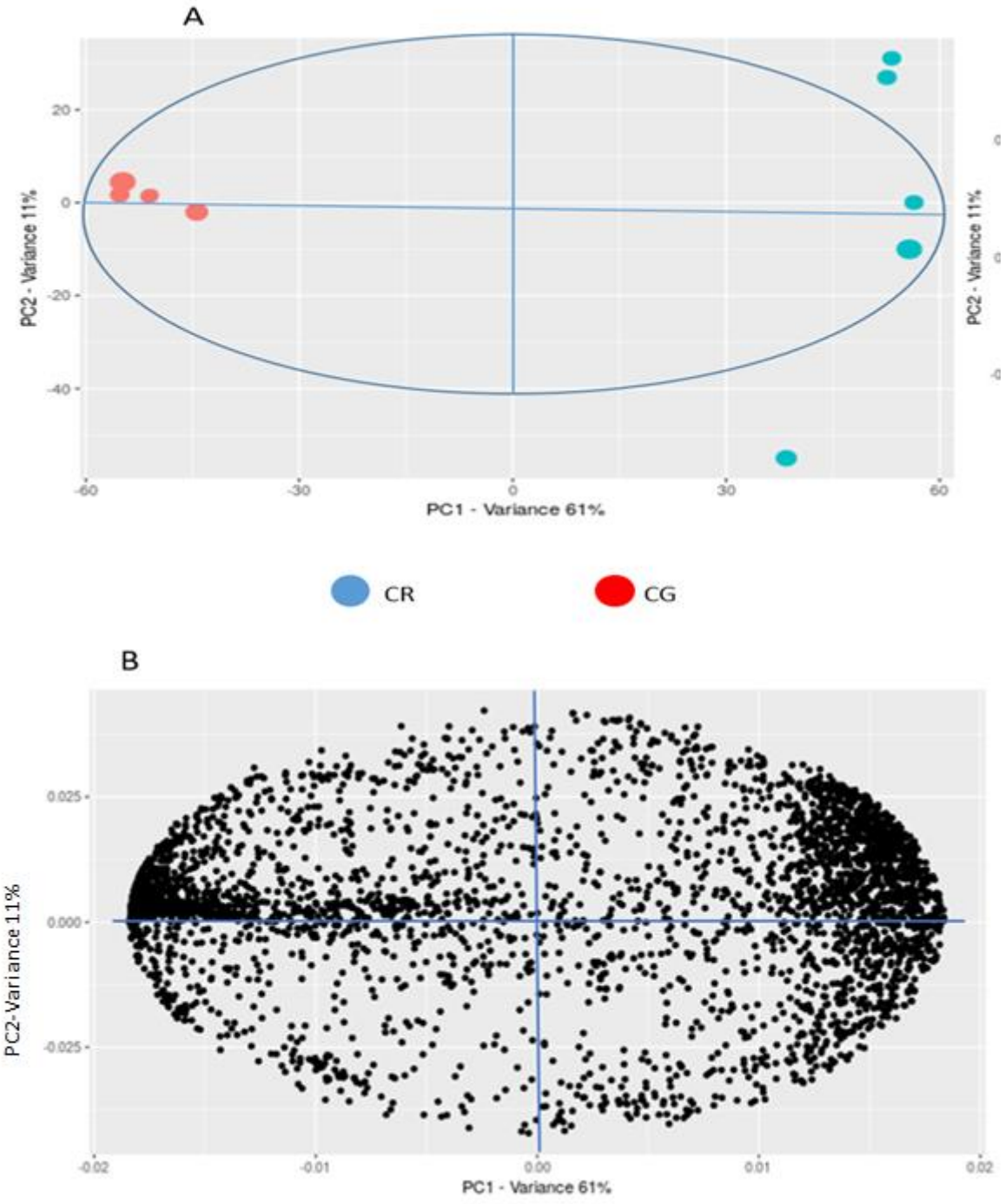


Figure 3.2: Principal component analysis (PCA) scores scatterplot and corresponding loadings plot. **(A)** PCA scores scatterplot obtained from XCMS online showing 72% variation between PC 1 and PC 2, indicating the general clustering within the datasets of *Coccinia grandis and Coccinia rehmannii* leaf samples. **(B)** Loadings plot that indicates the ions responsible for the clustering observed in the scores scatterplot.

In this study, XCMS online-based cloud plot was constructed to visualize metabolite variation between samples extracted from the two *Coccinia* plants (Figure 3.3). Cloud plot displays features whose intensities have been altered between sample groups. The cloud plot displays data characteristics such as *m/z* value, p-value, fold change, retention time (Rt) and signal intensity for each metabolite feature having a m/z ratio and retention time. Researchers have previously used cloud plot for different experiments. For example, it was used to visualize the metabolite variation between samples extracted at different temperature (Masike *et al*., 2017). It was also used elsewhere for an untargeted experiment, comparing lymphoblastic leukaemia cell lines (Gowda *et al*., 2014). Tuatenhahn and colleagues used cloud plot as a visualization scheme to present multi-dimensional data from septic mice (Tautenhahn *et al*., 2012). From the cloud plot (Figure 3.3A), the difference of metabolite composition between the two species is clearly observed and an indication that both plants possess interesting metabolites was detected. Features that are up regulated (on top) represent metabolites form the *C. rehmannii* (CR) leaf extracts and are represented in green colour. Those that are down-regulated (bottom) represent *C. grandis* (CG) metabolites in red colour. Clustering of metabolite features on the cloud plot is evidence that those compounds possess similar m/z values and retention times. Metabolite composition between the two plants was also shown in the total ion chromatogram (Figure 3.2B), taken from XCMS online. In (Figure 3.3B), peaks for *C. grandis* metabolites were shown in green colour and peaks representing *C. rehmannii* metabolites were shown in purple colour. From these observation, it can be concluded that number of significant metabolite features from *C. grandis* extracts are greater (in number) than those from *C. rehmannii* extracts, thereby supporting the observation from the PCA.

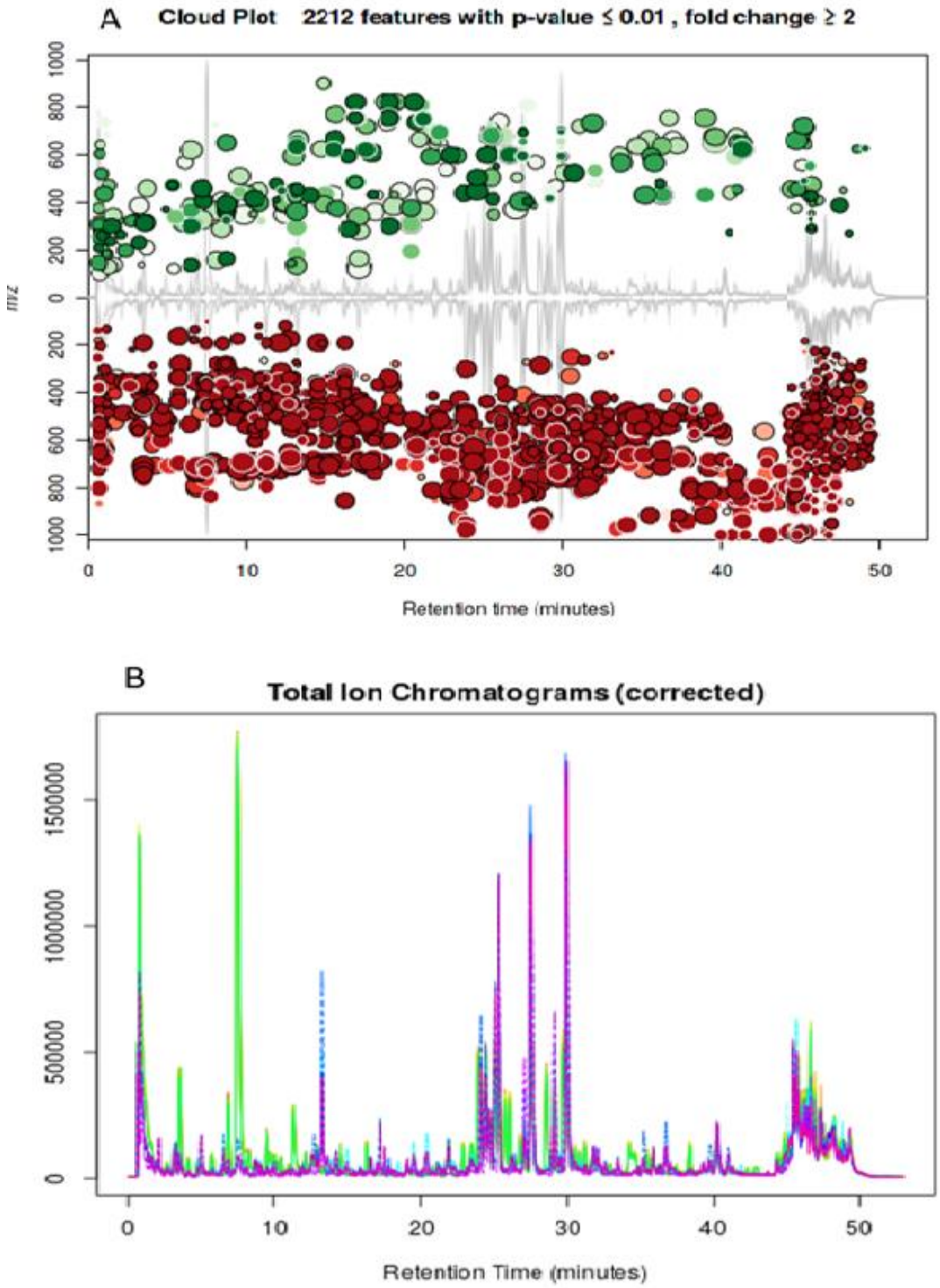


Figure 3.3: Cloud plot and total ion chromatogram created by the XCMS online, comparing the metabolite composition between two *Coccinia* species. (**A**) Cloud plot showing comparison of metabolites with p-value ≤ 0.01 and fold change ≥ 2, between of *C. grandis* (CG) and *C. rehmannii* (CR) leaf extracts. (**B**) The total ion chromatogram showing metabolite composition between the two plants, with *Coccinia grandis* metabolites shown in green colour and *Coccinia rehmannii* metabolites in purple colour.

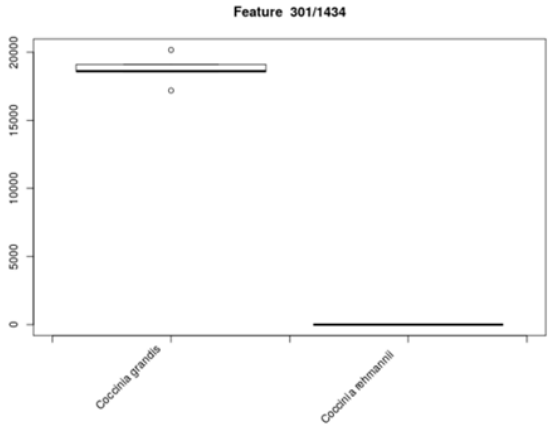
From the cloud plot (Figure 3.3A) features associated with *C. rehmannii* are projected on the top of the figure in green and those associated with *C. grandis* are projected at the bottom of the figure in red. The size of each bubble corresponds to the log fold change of the feature, the larger the bubble the bigger the fold change. The statistical significance of the fold change, as calculated by a Welch t-test with unequal variances, is represented by the intensity of the feature's colour where features with low *p*-values are brighter compared to features with high *p*-values. The y coordinate of each feature corresponds to the mass-to charge ratio of the compound as determined by mass spectrometry.

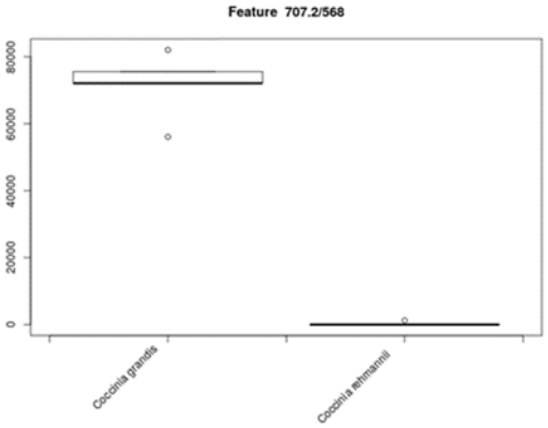
Table 3.1: Table listing *m/z* ratio, retention time and the probable identity of metabolites from the cloud plot of *Coccinia grandis* vs *Coccina rehmannii*, showing how the metabolites are distributed between the two species.

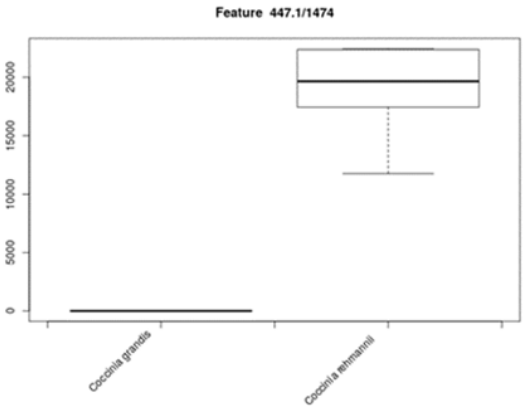
NO	Mass (m/z)	Retention time (rt)	Molecule name	*C. rehmannii*	*C. grandis*
1	755.206	16.91	Kaempferol-3-O-rutinoside-7-O-*β*-D-glucopyranoside	✓	
2	137.024	3.32	Salicylic acid	✓	
3	707.184	10.15	Kaempferide-3-O-rhamnoside-7-(6"-succinylglucoside)		✓
4	447.094	24.57	Quercetin 3-O-a-L-rhamnoside	✓	
5	301.036	23.9	Quercitin	✓	✓
6	353.088	18.25	Caffeoylquinic acid		✓
7	567.116	20.52	Isoorientin 2"-p-hydroxybenzoate	✓	
8	431.099	35.18	Apigenin 8-C-glucoside	✓	
9	467.051	11.28	Kaempferol 3-p-Coumarate		✓
10	583.114	15.12	Quercetin-3-(6"-p-hydroxybenzoylgalactoside)	✓	✓
11	535.087	4.69	Sagecoumarin		✓
12	729.166	16.45	Isoorientin4'-o-glucoside2"-o-p-hydroxybenzoate		✓
13	179.035	13.19	Caffeic acid		✓
14	193.051	14.75	Ferulic acid		✓
15	609.1503	6.98	Quercetin-3-O-rutinoside (**Isomer 1**)	✓	✓
16	609.142	7.30	Quercetin-3-O-rutinoside (**Isomer 2**)	✓	✓
17	609.125	11.66	Kaempferol-3-(6"-caffeoylglucoside) (**Isomer 1**)	✓	
18	609.127	12.01	Kaempferol-3-(6"-caffeoylglucoside) (**Isomer 2**)	✓	
19	593.156	7.79	Kaempferol-3-O- rutinoside (**Isomer 1**)	✓	✓

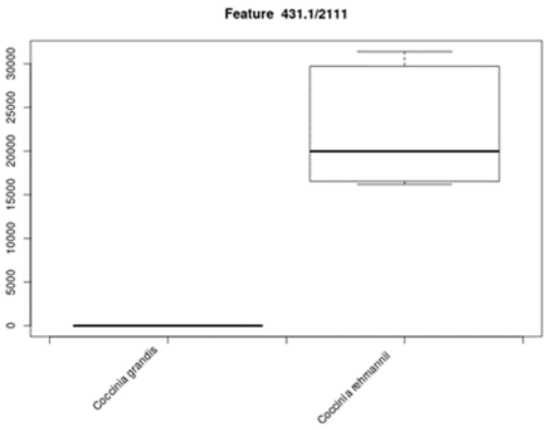
20	593.147	8.50	Kaempferol-3-O- rutinoside (**Isomer 2**)	✓	✓
21	593.146	9.21	Kaempferol-3-O- rutinoside (**Isomer 3**)	✓	✓
22	739.209	6.49	Kaempferol-3-rhamnoside-7glucosyl-(1-2)-rhamnoside (**Isomer 1**)	✓	✓
23	739.215	6.81	Kaempferol-3-rhamnoside-7glucosyl-(1-2)-rhamnoside (**Isomer 2**)	✓	✓

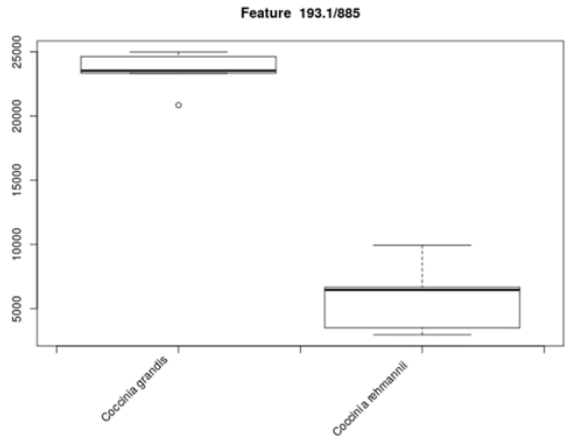
Identification of the 23 (Table 3.1) metabolites was also facilitated by cloud plot linked to the METLIN database. By clicking or scrolling through the bubbles (metabolite feature) properties such as retention time, p-value, fold change and possible names are displayed on the screen. Using XCMS online, box and whisker plots (Figure 3.4) were generated for the identified compounds. Box plots can be used to observe the distribution level of plant metabolites, for instance, within two different *Coccinia* species in this case. The level of distribution of each metabolite within the two species was noted to be varying widely, with some metabolites being greatly significant in *C. grandis* and others in *C. rehmannii*. The significance levels were noted by metabolites having *p*-values of less than 0.01 and a fold change of greater than 2 as indicated on the cloud plot figure 3.3A. Thus, it can be concluded that both *C. grandis* and *C. rehmannii* contain pharmacologically relevant metabolites, which are statistically significant across the datasets.

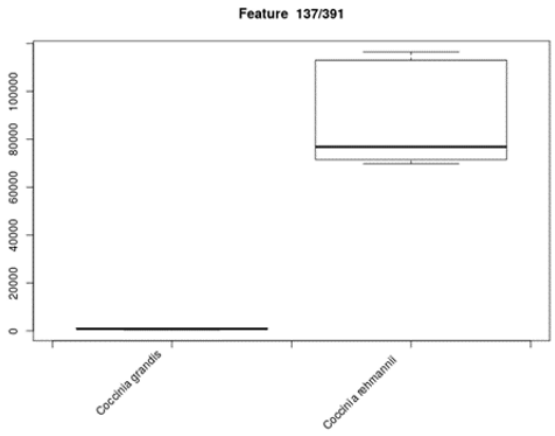


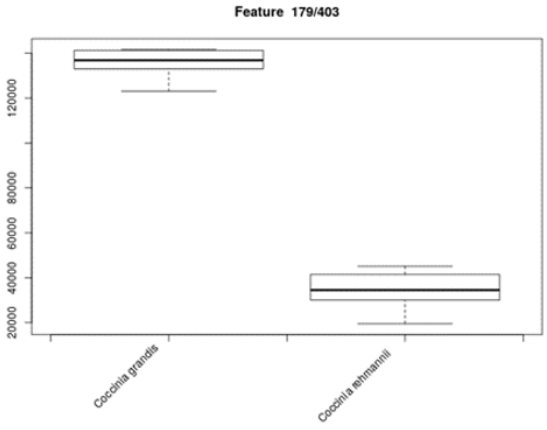


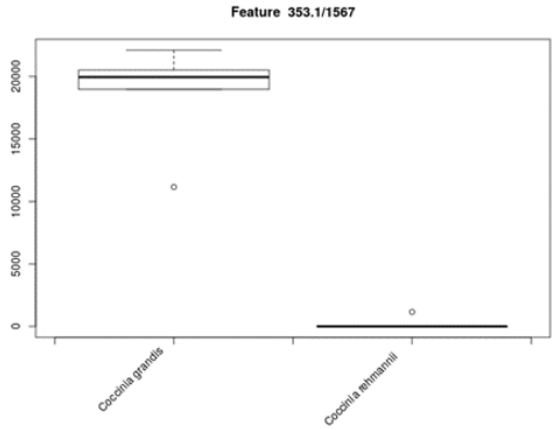


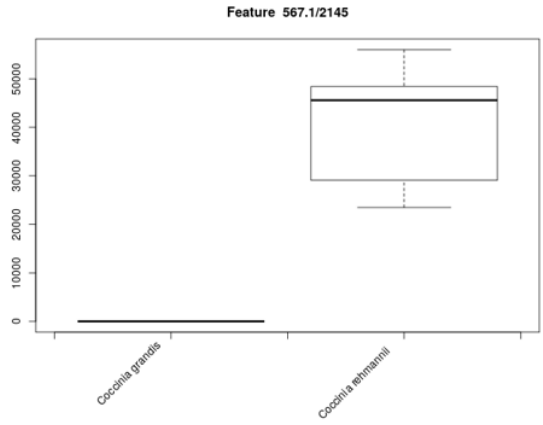


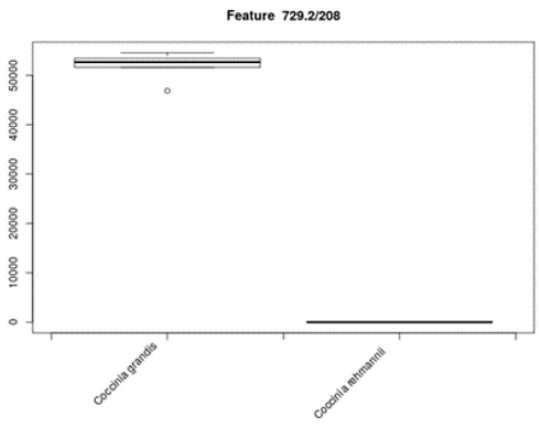


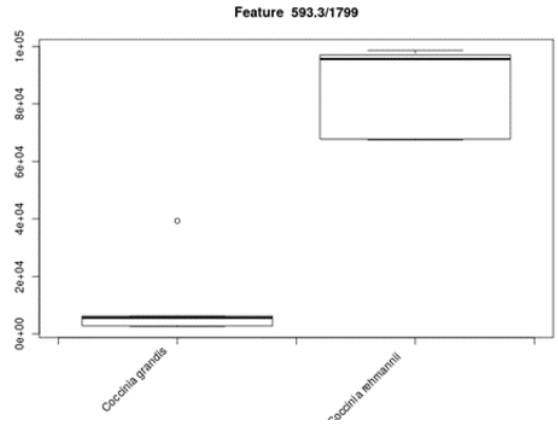


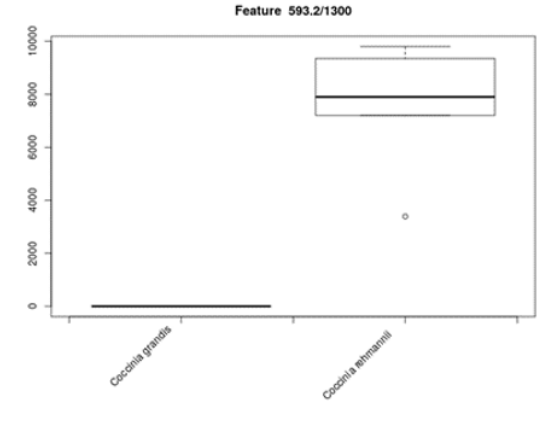


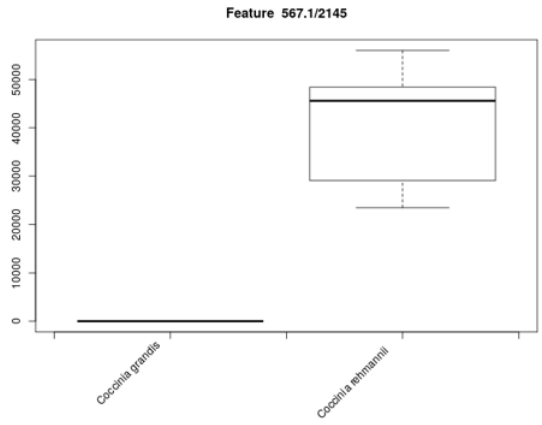


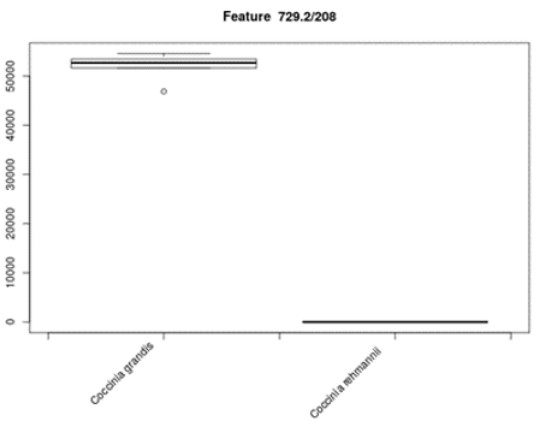


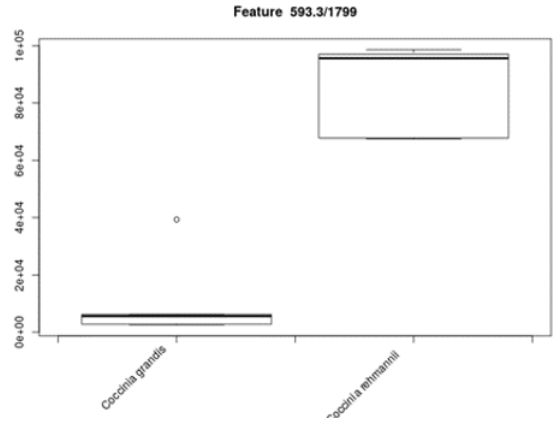


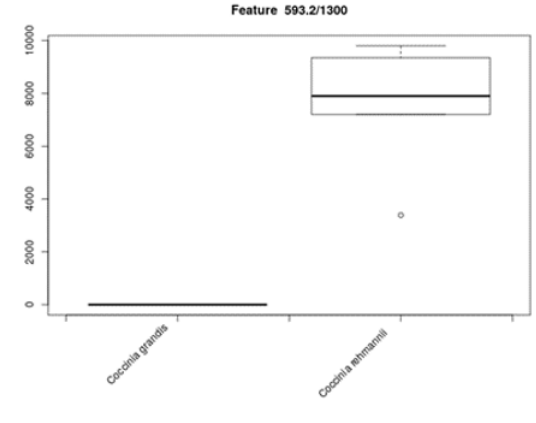


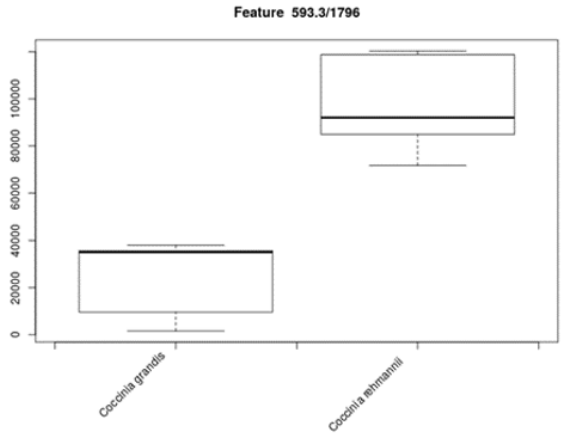


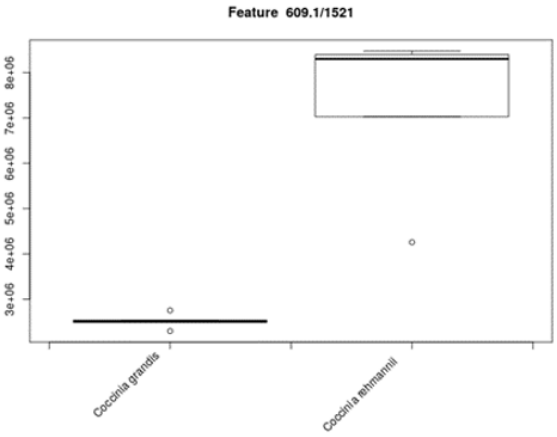


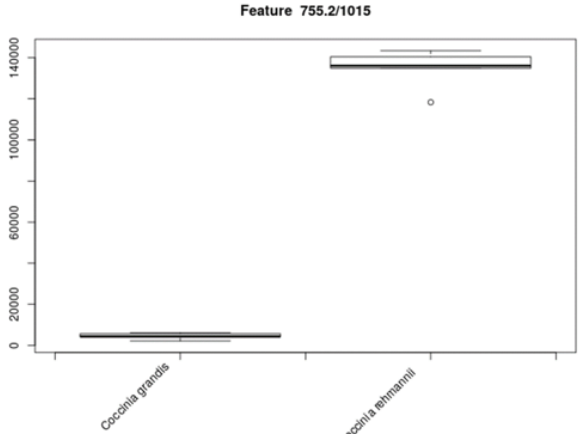


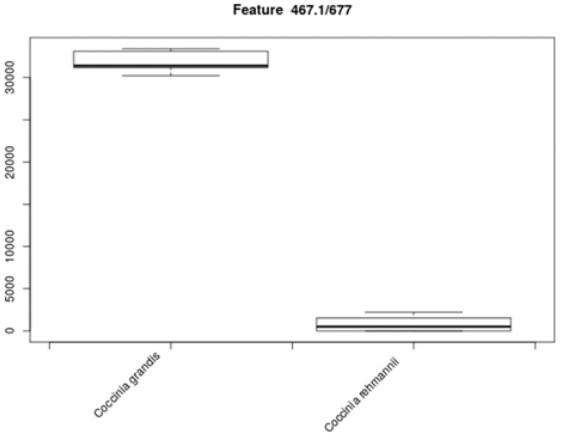


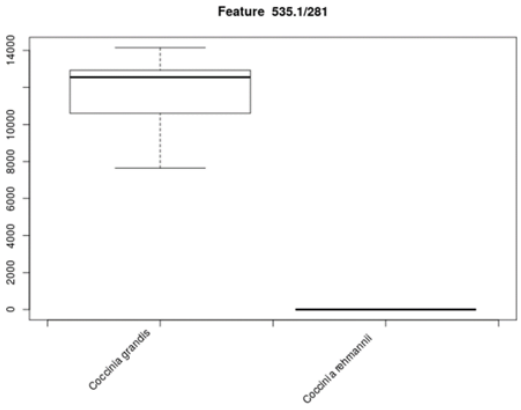


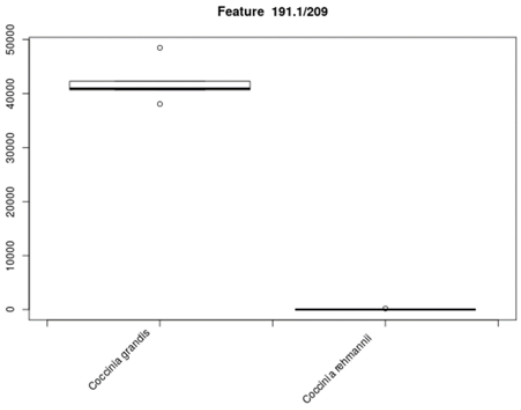


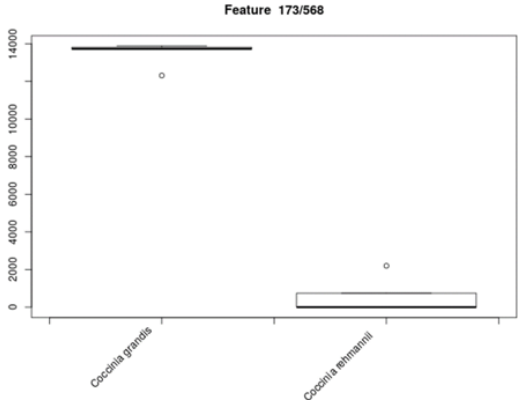












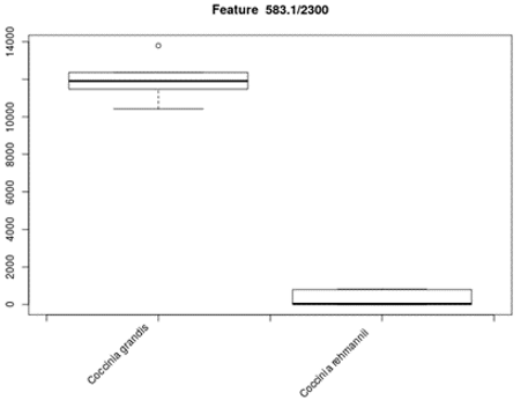


Figure 3.4: Box and whisker plot showing the distribution and significance level of some of the identified metabolites in Table 3.1 dominated by metabolites made up of hydroxycinnamate derivatives and flavonoids extracted from leaf extracts of two *Coccinia species*.

3.3.3. LC-MS/MS^2 analysis

Analyses of the un-targeted LC-MS data of the two understudied plants revealed the presence of peaks representing molecules containing quercetin and kaempferol flavonoids in the structural skeleton. This was further shown using fragmentation data achieved with various collision energy levels. Interestingly, when selected single ion chromatograms of these flavonoids were generated (using the pseudo-molecular ion peaks (m/z, $[M-H]^-$), several peaks appeared. The presence of multiple peaks with similar pseudo-molecular ion peaks and similar fragmentation patterns is an indication that these molecules are isomeric.

To further decipher these observations, a more sensitive and targeted approach through tandem mass spectrometry (MS/MS) was done (Figure 3.5A & B). The results of these studies revealed that indeed the selected molecules are capable of forming isomeric structures. The elemental composition of these compounds was generated, and it was later established that these molecules have been characterised elsewhere and most of them contain a complex glycoside, with most of them showing glycosylation through disaccharides attachments. For instance, some of the peak at *m/z* 593 and 609 were shown to have the following elemental compositions $C_{27}H_{30}O_{15}$ (Kaempferol rutinoside) and $C_{27}H_{30}O_{16}$ (Quercetin rutinoside), respectively. Rutinoside is a disaccharide sugar which has been shown to possess interesting biological activities (Orcic *et al*., 2011). The elemental composition and fragmentation of the compounds was generated with the aid of the UPLC-qTOF MS/MS spetra (Figure 3.6), which assisted in the differentiation of isomers.

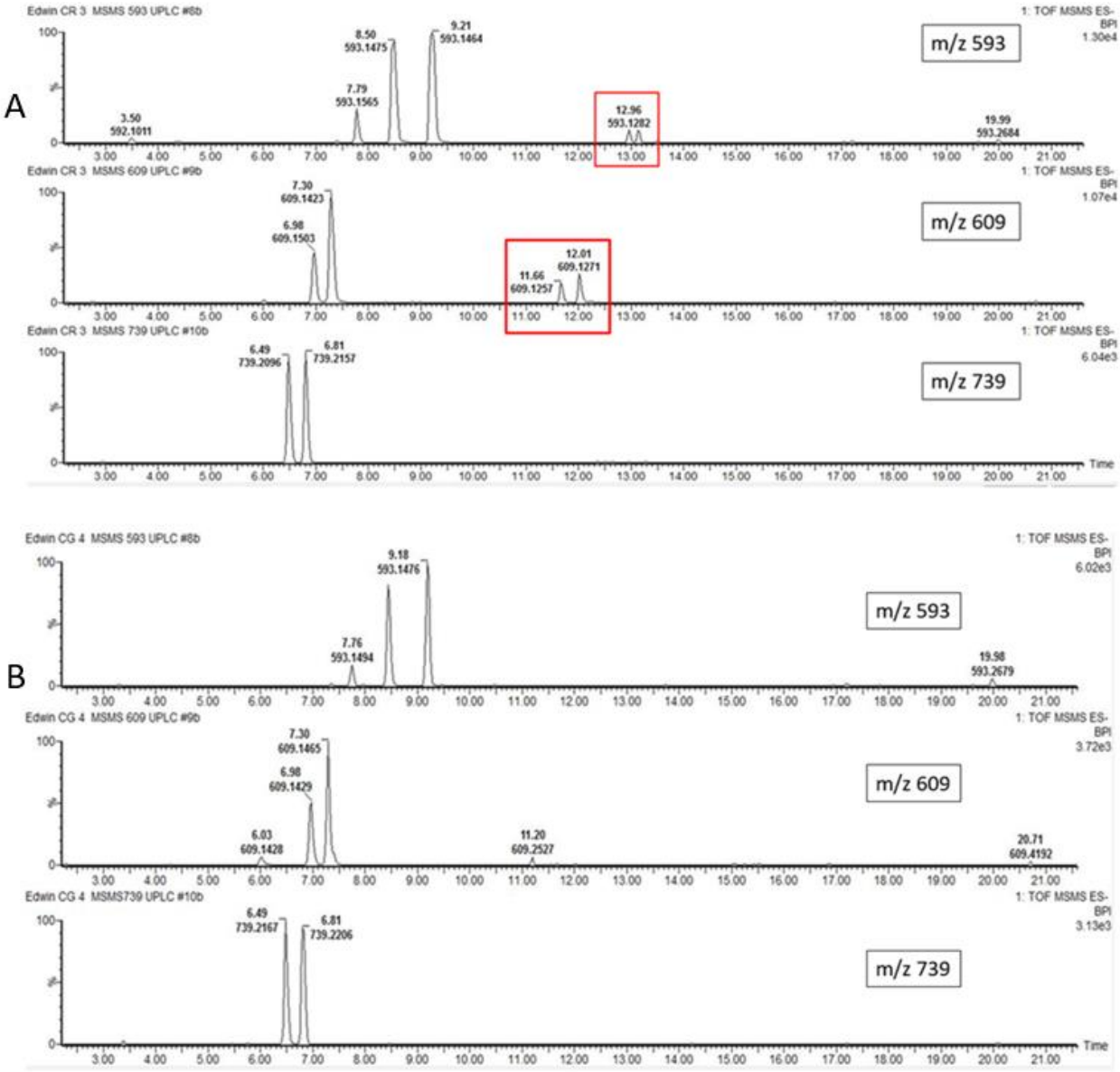


Figure 3.5: Representative UPLC-qTOF MS/MS chromatograms showing distribution patterns of compounds with pseudo-molecular ion peaks at *m/z* 593, *m/z* 609, *m/z* 739 isolated from (**A**) *Coccinia rehmannii* and (**B**) *Coccinia grandis*.

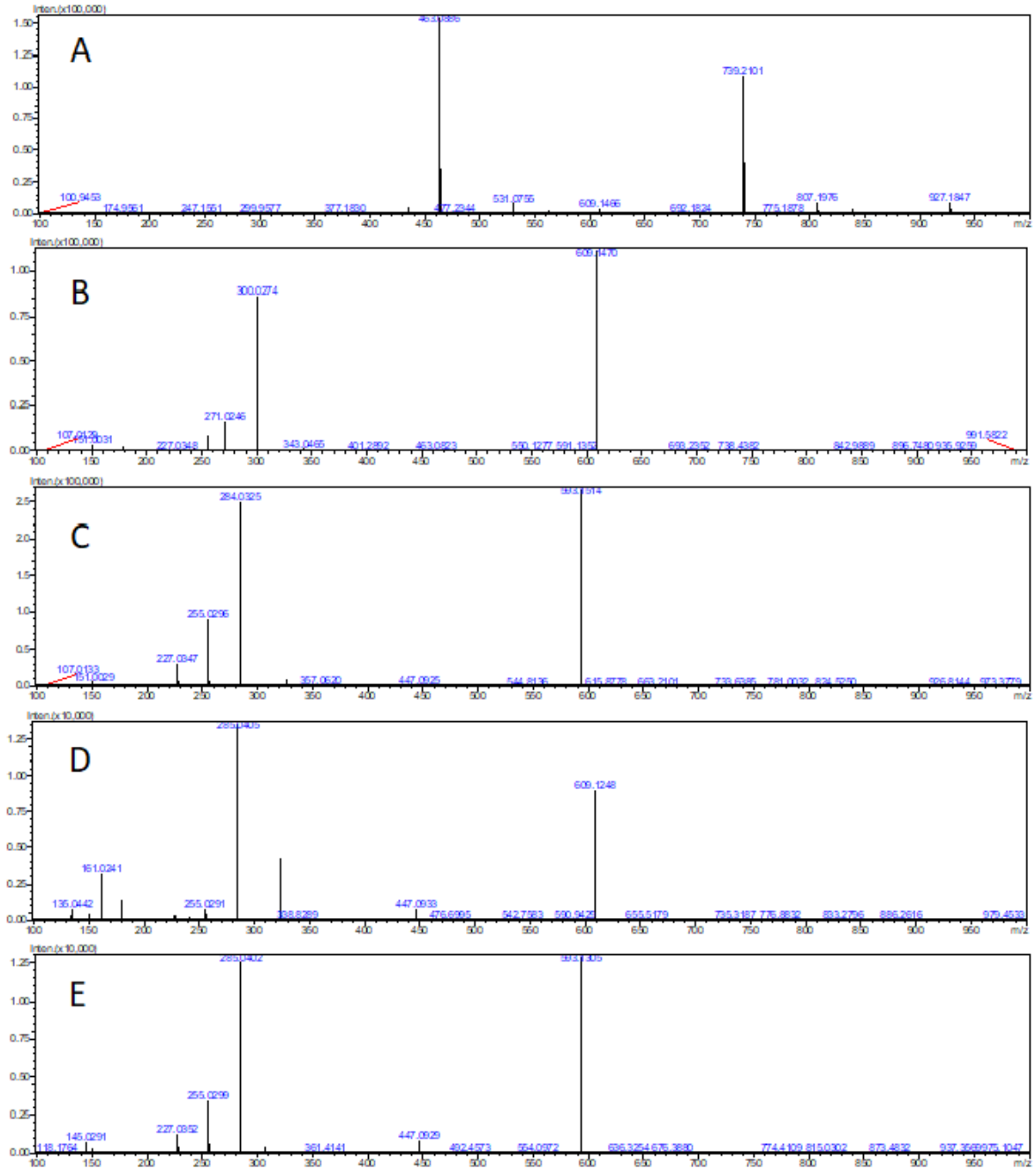


Figure 3.6: Representative UHPLC-qTOF-MS/MS spectra showing fragmentation patterns of flavonoids found in *Coccina grandis* and *Coccinia rehmannii*. **(A)** Kaempferol-3-rhamnoside-7glucosyl-(1-2)-rhamnoside **(B)** Quercetin 3-O-rutinoside **(C)** Kaempferol-3-O-rutinoside **(D)** Kaempferol-3-*p*-(6"-caffeoylglucoside **(E)** Kaempferol-3-*p*-Coumaroylglucoside.

In both plants, peaks representing kaempferol rutinoside at *m/z* 593 appeared at three retention times (7.7 min, 8.5 min and 9.2 min) (Figure 3.5A & B). The same can be seen with quercetin rutinoside at *m/z* 609, which appeared at two retention times (6.9 min and 7.3 min). When *m/z* 739 was targeted during MS/MS analyses, two peaks

representing a Kaempferol molecule harboring a trisaccharide attachment (http://www.knapsackfamily.com/knapsackjsp/top.html), were also noted at two different retention times (6.4 min and 6. 8 min). From the above results, it can be seen that all these three molecules contain more than one sugar molecule attached to the next one through a glycosidic bond (Figure 3.7A, B & C). Elsewhere, it has been shown that subtle differences in the position of a glycosidic bond result in isomeric molecules appearing at different retention times when analyzed on LC-MS (Avato & Argentieri, 2019).

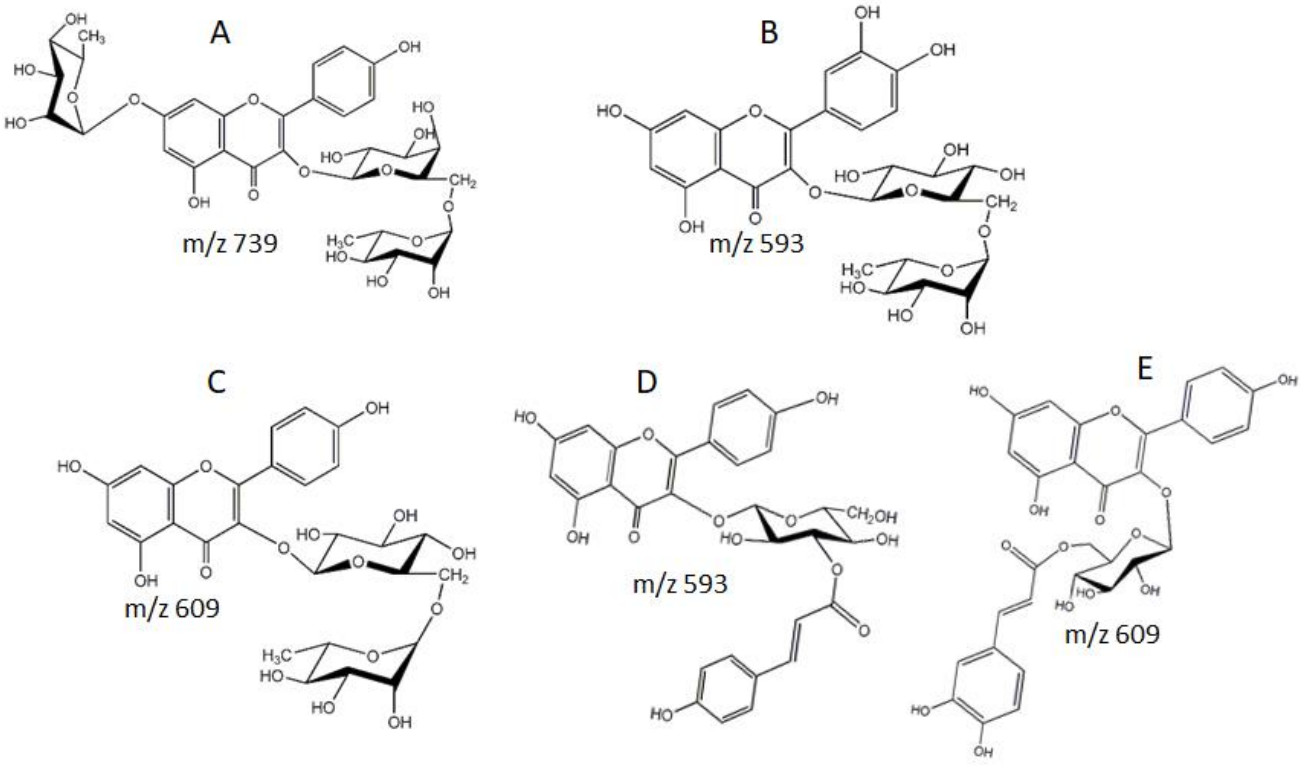


Figure 3.7: Structures of different compounds detected from the methanol extracts of *Coccinia* plants.

From the same results, additional peaks for either molecules appearing at *m/z* 593 and 609 were also seen on the chromatograms of *C. rehmannii* extracts as indicated by red boxes (Figure 3.5). Upon closer investigation, these molecules were found to contain hydroxycinnamic acid (HCA) moieties (Figure 3.7D & E) instead of the second sugar in the case of disaccharides discussed above. Attachment of HCA derivatives on the flavonoids have been shown elsewhere (Kim *et al*., 2019). Acylation of flavonoids by HCA molecules is not known, however, recently it has been shown that sugar glycosylation pattern has biological significance (Reily *et al*., 2019) and, as such,

modification of glycosylation patterns by these HCA molecules warrants further investigation.

The two *Coccinia* species are capable of producing flavonoid molecules that vary greatly due to modifications they undergo. In this study, flavonoids have been shown to be prone to isomerization through attaching to either *di-* or *tri-* saccharides. Furthermore, isomerization through sugar acylation by derivatives of cinnamic acids was also noted, herein (Figure 3.7). The distribution and significance level of these interesting metabolites was shown to vary between the two species (Figure 3.8). The box and whisker plots reveal flavonoid significance level between the two species, with most of the flavonoid being highly significant in the *Coccinia rehmannii* specie. In addition, flavonoids which exhibited the attachment to derivatives of cinnamic acids were only noted in *C. rehmannii* and not in *C. grandis*, further revealing the differences between the two closely related species.

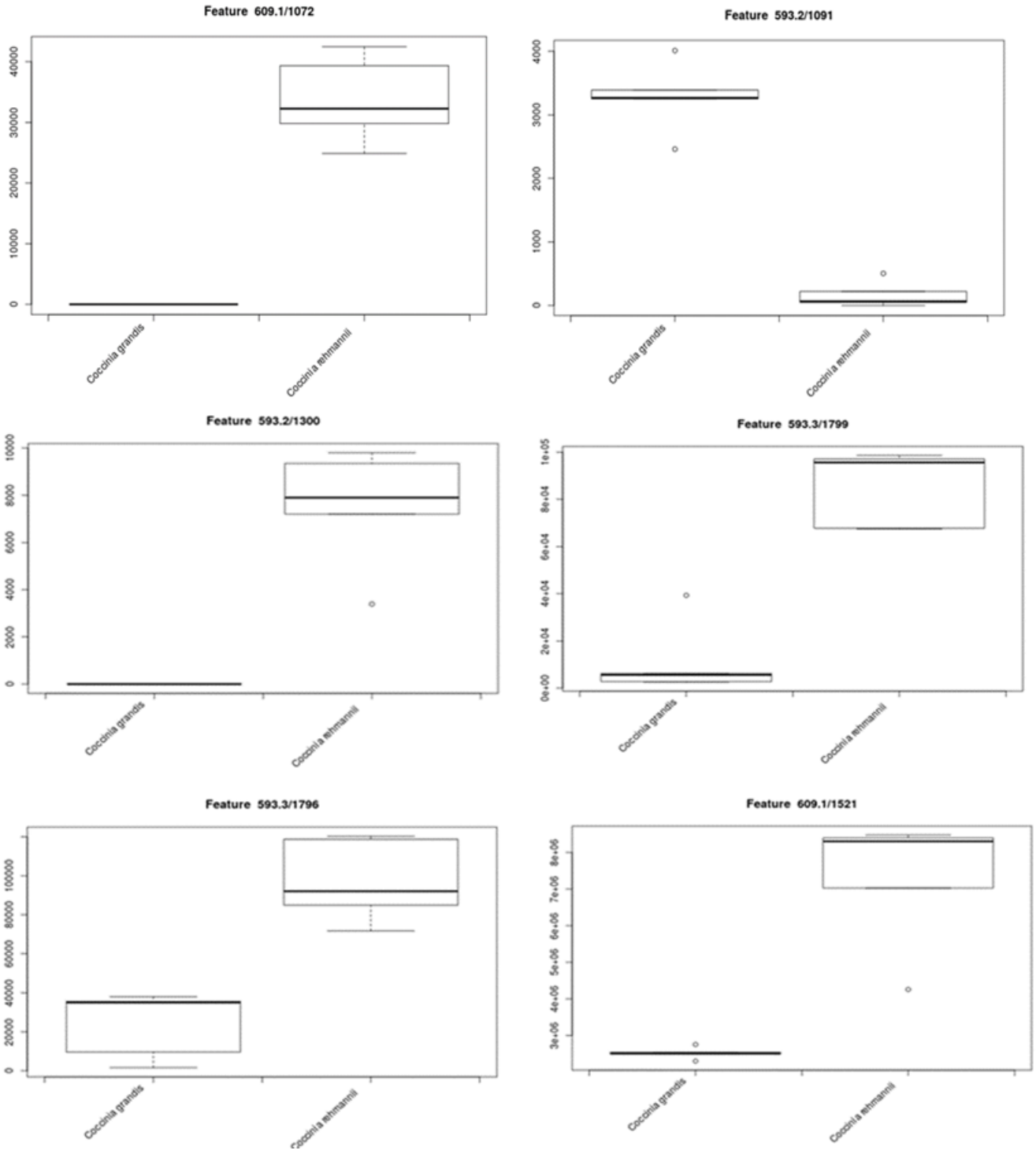


Figure 3.8: Box and whisker plots showing the distribution and significance level of flavonoids that display glyco-isomerization between the two *Coccinia* species.

3.4. Conclusions

In conclusion, the findings of the current study highlight glyco-isomerization as an evolutionary strategy utilised by plants to maximise its metabolite pool. The presence of these diverse glycosylation patterns could have overwhelming biological consequences, especially nutraceutical properties when consumed as food or medicine. For instance, it is well known that sugar molecules enhance bioavailability of active molecules (flavonoids in this case). From the results of the current study, it can also be seen that the composition of the flavonoids molecules differs from the two investigated plants, with *C. rehmannii* being the superior species. Here, *C. rehmannii* was shown to possess some flavonoids with terminal sugars acylated with hydroxycinnamic acid (HCA) molecules (Caffeic & Coumaric acids), a phenomenon which was absent in *C. grandis*. The presence of HCA derivatives could also allow formation of geometrical isomers (*trans* – *cis*) which could result is diverse or complex metabolome which could also affect the pharmacological properties thereof.

References

Barnes, P. W., Flint, S. D., Tobler, M. A., & Ryel, R. J. (2016). Diurnal adjustment in ultraviolet sunscreen protection is widespread among higher plants. Oecologia, 181; 55-63.

Cook, D. W., & Rutan, S. C. (2014). Chemometrics for the analysis of chromatographic data in metabolomics investigations. Journal of Chemometrics, 28; 681-687.

Day AJ, Canada FJ, Diaz JC (2000). Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. FEBS Letters. 468; 166–170.

Gowda, H., Ivanisevic, J., Johnson, C. H., Kurczy, M. E., Benton, H. P., Rinehart, D & Siuzdak, G. (2014). Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. Analytical chemistry, 86; 6931-6939.

Hao, L., Wang, J., Page, D, *et al*. (2018). Comparative evaluation of MS-based metabolomics software and its application to Preclinical Alzheimer's disease. Scientific Report, 8; 9291.

Jaadi,Z. A. (2019). Step by step explanation of Principal Component Analysis, September 4.

Jansen, J. J., Smit, S., Hoefsloot, H. C., & Smilde, A. K. (2010). The photographer and the greenhouse: how to analyse plant metabolomics data. Phytochemical Analysis: An International Journal of Plant Chemical and Biochemical Techniques, 21; 48-60.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, 374; 2065.

Jianbo Xiao (2016). Dietary flavonoid aglycones and their glycosides: Which show better biological significance? Journal Critical Reviews in Food Science and Nutrition, 57; 1874-1905.

Kachlicki, P., Einhorn, J., Stobiecki, M. (2008). Evaluation of glycosylation and malonylation patterns in flavonoid glycosides during LC/MS/MS metabolite profiling. Published in Journal of mass spectrometry, 43; 572-586.

Kim, J. Y., Swanson, R. T., Alvarez, M. I., Johnson, T. S., Cho, K. H., Clark, D.G., & Thomas, A. (2019). Down regulation of p-coumarate 3-hydroxylase in petunia uniquely alters the profile of emitted foral volatiles, 9; 8852.

Lamichhane, S., Sen, P., Dickens, A. M., Hyötyläinen, T., & Orešič, M. (2018). An overview of metabolomics data analysis: current tools and future perspectives. Comprehensive analytical chemistry, 82; 387-413.

Lee, T., Ho, Y. S., Yeo, H. C., Lin, J. P. Y., Lee, D. Y. (2013). Precursor mass prediction by clustering ionization products in LC-MS based metabolomics. Metabolomics, 9; 1301-1310.

Madala, N. E., Piater, L. A., Steenkamp, P. A., & Dubery, I. A. (2014). Multivariate statistical models of metabolomic data reveals different metabolite distribution patterns in isonitrosoacetophenone-elicited Nicotiana tabacum and Sorghum bicolor cells. SpringerPlus, 3; 1-10.

Makita, C., Chimuka, L., Steenkamp, P., Cukrowska, E., Madala, E. (2016). Comparative analyses of flavonoid content in *Moringa oleifera* and *Moringa ovalifolia* with the aid of UHPLC-qTOF-MS fingerprinting South African Journal of Botany, 105; 116-122.

Masike, K., Khoza, B. S., Steenkamp, P. A., Smit, E., Dubery, I. A., Madala, N. E (2017). A metabolomics-guided exploration of the phytochemical constituents of *Vernonia fastigiata* with the aid of pressurized hot water extraction and Liquid Chromatography-Mass Spectrometry. Molecules, 22; 1200.

Muth, D., Marsden-Edwards, E., Kachlicki, P., Stobiecki, M. (2008). Differentiation of isomeric malonylated flavonoid glyconjugates in plant extracts with UPLC-ESI/MS/MS. Phytochemical Analysis, 5; 444-452.

Muthana, S., Campbell, C & Gildersleeve, J. (2011). Modifications of Glycans: Biological Significance and Therapeutic Opportunities. ACS chemical biology, 7; 31-43.

Ncube, E. N., Steenkamp, P. A., Madala, N. E., & Dubery, I. A. (2016). Stimulatory effects of acibenzolar-S-methyl on chlorogenic acids biosynthesis in Centella asiatica cells. Frontiers in plant science, 7; 1469.

Negri, G & Tabach, R. (2013). Saponins, tannins and flavonols found in hydroethanolic extract from Periandra dulcis roots. Revista Brasileira de Farmacognosia, 23; 851-860

Xiao J, Capannoglu E, Jassbi AR, and Miron A (2016). Advance on the Flavonoid C-glycosides and Health Benefits. Critical Reviews in Food Science and Nutrition, 56; 29–45.

Y Ogo, T Mori, R Nakabayashi, K Saito and F Takaiwa (2016). Transgenic rice seed expressing flavonoid biosynthetic genes accumulate glycosylated and/or acylated flavonoids in protein bodies. Journal of Experimental Botany, 67; 95–106.

Orčić, D. Z., Mimica-Dukić, N. M., Francišković, M. M (2011). Antioxidant activity relationship of phenolic compounds in Hypericum perforatum. Chemistry Central Journal, 5; 34.

Avato, P and Argentieri, M. A. (2019). Quality assessment of commercial Spagyric Tinctures of Harpagophytum procumbens and their antioxidant properties. Molecules 24: 2251.

Reily, C., Stewart, T. J., Renfrow, M. B., & Novak, J. (2019). Glycosylation in health and disease. Nature reviews. Nephrology, 15; 346–366.

Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R & Siuzdak, G. (2005). METLIN: a metabolite mass spectral database. Therapeutic drug monitoring, 27; 747-751.

Tautenhahn, R., Patti, G. J., Rinehart, D., & Siuzdak, G. (2012). XCMS Online: a web-based platform to process untargeted metabolomic data. Analytical chemistry, 84; 5035-5039.

Wikoff, W. R., Pendyala, G., Siuzdak, G., & Fox, H.S. (2008). Metabolomic analysis of the cerebrospinal fluid reveals changes in phospholipase expression in the CNS of SIV-infected macaques. The journal of clinical investigation, 118; 2661-2669.

Wikoff, W. R., Anfora, A. T., Liu, J., Schultz, P. G., Lesley, S. A., Peters, E. C., & Siuzdak, G. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. Proceedings of the national academy of sciences, 106; 3698-3703.

Xia, J., Sinelnikov, I. V., Han, B., & Wishart, D. S. (2015). MetaboAnalyst 3.0-making metabolomics more meaningful. Nucleic acids research, 43; 251-257.

Chapter 4: LC-MS based metabolite profiling reveals hydroxylcinnamoyl conjugation as a discriminatory chemical factor between two closely related *Coccinia* species

Abstract

The fruits of *Coccinia grandis* (L.) Voigt have been used to treat several aliments in traditional medicine whereas *Coccinia rehmannii* Cogn. fruits are consumed as food due to their high starch content. *C. rehmannii* plant are also known to contain vast amount indispensable bioactive compounds that are used in various cosmetics, foods, and medicines. Among these compounds are hydroxycinnamic acids (HCAs) which are usually conjugated/esterified to quinic acid (QA), tartaric acid, citric acid, shikimic acid and sugars and are collectively referred to as chlorogenic acids (CGAs). Over the years, metabolite profiling has been used in plant taxonomy and species classification. Herein, UPLC-qTOF-MS/MS semi-targeted approached was use to profile HCA conjugates from two closely related plant species i.e. *C. grandis* and *C. rehmannii* leaves in order to distinguish them from each other. Examination of their relative metabolite contents showed that the two closely related *Coccinia* species utilise contrasting means of HCA conjugations as means of metabolite diversification strategy. *C. grandis* was found to conjugate its HCA derivatives to a QA, whilst *C. rehmannii* conjugates its HCA derivatives to mainly flavonoid glycosides and to QA, even though present in low concentration in this plant. These results shows that plants from the same genus can utilise different chemical diversification means for storage of their metabolites and that processes such as conjugation and isomerization can be exploited as chemotaxonomic markers. Future studies can also be designed to look at the pharmacological potency of these plants as a function of their contrasting chemical differences.

4.1. Introduction

Coccinia species from Curcubaticeae family have been shown to have several pharmacological activities (Rahmatullah *et al*., 2012). In Africa, two *Coccinia* species have been reported, namely *Coccinia rehmannii* Cogn. known as Wild cucumber in English and *Coccinia grandis* (L.) Voigt kwon as Ivy gourd in English, with the former referred to as an African species and the latter as an introduced species from India (Holstein, 2015). In traditional medicine, fruits of *C. grandis* have been used to treat leprosy, fever, asthma, bronchitis, and jaundice due to its antioxidant and anti-inflammatory properties (Pekamwar *et al*., 2013; Taur & Patil, 2011). Unripe fruits of *C. rehmannii* are taken mainly by children as food sources due to its nutritional composition (Imbumi, 2004). The tubers of *C. rehmannii* are rich in starch and are eaten raw, boiled or roasted as food. The roasted tuber is then peeled and pounded, and mixed with leafy vegetables as a nutritious preparation (Leffers, 2003). Previously, it has been reported that a diet rich in hydroxycinnamic acids (HCAs) exhibit beneficial health effects such as protection against cardiovascular diseases (Alam *et al*., 2016). These compounds also present potential therapeutic effects which can be useful in treatment of cancer, diabetes and lung diseases (El-Seedi *et al*., 2018). Furthermore, these molecules have been reported to confer anti-oxidative activities, owing to some of their structural features such as the catechol group which allows them to scavenge radical ions through hydrogen atom transfer mechanism (Makola *et al*. 2016a)

Hydroxycinnamic acids such as coumaric acid, ferulic acid, sinapic acid and caffeic acid are phenolic compounds that are widely distributed in plants and they are found as free compounds or conjugated to other molecules such as quinic acid (QA), tartaric acid, citric acid and sugars, to name a few. When conjugated/esterified to other molecules HCA are referred to as chlorogenic acids (CGAs) (Roleira *et al*., 2018; Masike *et al*., 2017a; Nobela *et al*., 2018). Naturally, almost all HCA derivatives are synthesized with a *trans* configuration but *cis* isomers have also been noted as a result of photo-isomerization (Clifford *et al*., 2008). HCA conjugation/esterification can occur at different positions on the other molecule. For examples, HCA can acylate quinic acid at position 1, 3, 4 and 5, thereby resulting in structurally diverse metabolomes characterized by positional isomers and geometrical isomers (*trans/cis*). Moreover, HCA can be conjugated to different molecules resulting in isobars (i.e., caffeic acid

conjugation to quinic acid and isocitric acid, both respective group of compounds have a molecular mass of 354 Da) (Masike *et al*., 2017b). Recently, Nengovhela *et al*., (2020) reported HCAs flavonoid conjugates in *C. grandis* and *C. rehmannii*. Interestingly, the nature and chemistry of HCA derivatives has been shown to be plant specific and, as such, CGAs have been used as chemotaxonomic markers (Masike *et al*., 2017a). As such, the analyses of CGAs pose an undisputed analytical challenge, since the isomers of CGAs exhibit similar fragmentation patterns which makes them impossible to distinguish under un-optimized analytical conditions (Nobela *et al*., 2018). However, the use of both elution orders and the fragmentation attainable through multi-stage MS fragmentation has been used successfully (Clifford *et al*., 2003). As such, fragmentation pattern of closely related isomers can be regarded as a diagnostic tool to differentiate both positional and geometrical isomers and the acquisition and reporting of these results are essential for future references (Clifford and Madala, 2017).

Tandem (MS/MS) spectra for the unknown ions is crucial during the identification of the unknown compounds through direct comparison with spectral databases or through data-mining tools such as molecular networking (Wang *et al*., 2016). However, both these approaches require the acquisition of MS/MS data and a good MS/MS data acquisition strategy has to produce many ions with high quality from the sample. One such strategy in metabolomics is the data dependant acquisition (DDA) approach which target particular ions observed in MS^1 scans for downstream fragmentation (MS^2). DDA eliminates the need to re-run the sample twice to acquire firstly the MS^1 and secondly the MS/MS data. Thus it eliminates the need to run the sample in MS mode to identify the target precursor ions then re-run the sample in MS/MS mode to acquire the MS/MS data of the same ions (Martin, 2016).

When analysing complex biological samples top-n-based DDA method inhibit high MS/MS coverage metabolites by selecting only a small fraction of ions to be fragmented. Therefore, Wang et al., (2017) proposed a novel DDA method which would improve MS/MS acquisition performance in LC-MS untargeted metabolomics, with the use of target directed DDA (t-DDa). The results showed that the t-DDA exhibit better MS/MS coverage and efficiency when compared to the conventional DDA and it is capable of highlighting the features of interest regardless of their peak intensities (Wang *et al*., 2017).

Molecular Network (MN) is a computational strategy that has become a key in visualizing and annotating the complex non-targeted mass spectrometry data. MN is a tool that clusters molecules into their molecular families based on potential similarities within their MS/MS spectra (common MS^2 fragmentation) (Wang *et al*., 2016; Pilon *et al*., 2019). The similarity between molecules can be noted by the peaks they have, that is, the more peaks the MS/MS spectra of the molecules share, the more similar the molecules are (Wolfender *et al*., 2019). A molecular network is built/drawn using fragmented molecules, which serve as "nodes". These nodes connect to other nodes that they share spectral similarity with and after a collection of all similar nodes, a molecular network can be drawn (Wolfender *et al*., 2019).

Molecular networking has been used previously in metabolomics. For instance, Garg et al (2015) applied MS/MS based MN to analyse complex mixtures. That is, they showed that molecular networking can be used for rapid identification metabolites, for managing and comparing MS data from complex mixtures of organisms. This was achieved by showing data analysis from hard and covals and a human lung associated with cystic firbrosis (Garg *et al*., 2015). In plant studies, MN has also been used to view similarities in fragmentation patterns of a saponin glycoside and a putatively found unknown metabolite corresponding to Scillanoside L-1 (Raheem *et al*., 2019). The molecular network of MS/MS data was also established to find the nearest correlated structure which was comparable to the unknown compound (Raheem *et al*., 2019).

Mass spectrometry techniques are capable of high throughput characterization of natural products, however, a challenge arises when analysing MS based data of complex biological mixture. Said et al (2020) aimed to overcome such limitations and simplify analysis of complex MS data by conducting a cocoa tea based human trial through implementation of MS based MN (Said *et al*., 2020). The study highlighted the use of MS based MN in illustrating the tracking of different structural motifs of ingested Cocoa phenolics in a human based study.

Moreover, due to the limitation of traditional MS technique, other methods such as ion mobility mass spectrometry have been used to differentiate between closely related CGAs (Zheng *et al*., 2017). However, ion mobility MS are very expensive and rare in scientific communities and, as such, other methods such as the use of metal

adduct formation during in source collision induced dissociation based MS method has been shown to unequivocally distinguish between geometrical isomers of CGAs (Makola *et al*., 2016b). In the current study a simple high resolution MS in combination with collision induced dissociation was used and it was also found to be effective in previous studies (Ncube *et al*., 2014; Nobel *et al*., 2018)

Previous research has shown *Coccinia* plants to produce HCAs however, very little scientific information on the complete composition of HCAs within *Coccinia* species has been documented (Chanda *et al*., 2020). In most cases, CGAs exist as conjugates of quinic acid but recent report by Nengovhela et al. (2020) showed that *Coccinia* plants attaches their HCAs on flavonoids compound and this further contributes towards metabolite complexity through biological phenomenon such as glyco-isomerization (i.e. differential acylation of sugar moieties). Therefore, the current study investigated the structural differences of HCA-containing metabolites in these two species with emphasis on isomers of CGAs through the use of ultra-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (UPLC-qTOF-MS/MS) based on collision induced dissociation fragmentation patterns.

4.2. Materials and Methods

4.2.1. *Plant material*

Fresh leaves *Coccinia grandis* (L.) Voigt and *Coccinia rehmannii* Cogn. were collected from the villages around Thohoyandou area, Limpopo Province (22°99’6’’S; 30°40’08’’E for *C. grandis* and 22°61’7’’S; 30°55’’6’E for *C. rehmannii*). The plants were authenticated with the help of botanist, Dr Khathutshelo Magwede. Herbarium specimens of these two plant species were logged at the University of Venda’s herbarium collection. Prior metabolite extractions, the leaves of the collected plants were air dried at low humidity, and controlled temperature (26°C) in complete darkness for a period of 72h to complete dryness.

4.2.2. *Metabolites extraction*

Metabolite extraction was accomplished by a method proposed by Makita et al., (2016). The dried plant leaves were grinded to fine powder and 2 g was mixed with 20 mL (1:10 m/v) of 80% aqueous methanol. To remove the debris, the homogenate was centrifuged at 5000 g for 20 min and the supernatant was transferred to new clean tubes. The extracts were further diluted 1:1 (v:v) to final volume of 20 mL using methanol, followed by transfer of 10 mL of the diluted extracts into cylindrical quartz glass vials (2 × 10 cm). Thereafter, the samples were filtered into a 2 mL vial fitted with a 0.2 mL conical bottom glass insert using a syringe fitted with a 0.2 µmm microfilter.

4.2.3. *Liquid Chromatography Mass Spectrometry analysis*

Sample analysis was conducted on an LC-QTOF-MS, model LC-MS 9030 instrument with a Shim Pack Velox C18 column (100 mm × 2.1 mm with particle size of 2.7 µm) (Shimadzu, Kyoto, Japan) placed in a column oven set at 40 °C. Herein, a binary solvent mixture was used consisting of 0.1% formic acid in water (Eluent A) and 0.1% formic acid in acetonitrile (Eluent B) at a constant flow rate of 0.4 mL/min.

Using conditions used previously (Madala and Kabanda, 2021), a mass spectrometer detector was used for monitoring analyte elations, under the following conditions: ESI negative modes; interface voltage of 3.5 kV; nitrogen gas was used as nebulizer at flow rate 3 L/min, heating gas flow at 10 L/min; heat block temperature at 400 °C, CDL temperature at 250 °C; detector voltage of 1.70 kV and the TOF temperature at 42 °C.

4.2.4. *MS/MS experiments*

For tandem MS (MSMS) experiments, a mass calibration solution of sodium iodide (NaI) was used to obtain typical mass accuracies with a mass error below 1 ppm. A range of *m/z* 100-1000 was used for high resolution for MS and MS/MS experiments. Argon gas was used as a collision gas for MSMS experiments along with MS^E mode using collision energy ramp of 12 eV to 25 eV for generation of fragments.

4.2.5. *Classical Molecular Networking of Coccinia species*

A molecular network was created using the method created by Wang et al. 2016 using the online workflow (https://ccms-ucsd.github.io/GNPSDocumentation/) on the GNPS website (http://gnps.ucsd.edu). Chromatographic data (mzXmL format) representing metabolites extracted from *Coccinia* plants were filtered by removing all MS/MS fragment ions within +/- 17 Da of the precursor m/z. Only the top 6 fragment ions in the +/- 50Da window throughout the spectrum were chosen to window filter MS/MS spectra. The precursor ion mass tolerance and MS/MS fragment ion tolerance were set at 0.05 Da. Molecular network was then created where edges were filtered in order to have a cosine score which is above 0.7 and more than 6 matched peaks. Edges between two nodes were kept in the network only if each of the nodes happen to appear in each other's respective top 10 most similar nodes. The maximum size of a molecular family was set to 100, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data.

4.3. *Results and discussion*

In this study, UHPLC-qTOF-MS/MS was used to profile the hydroxycinnamoyl acids using fragmentation methods presented elsewhere (Clifford *et al*., 2003; Ramabulana *et al*., 2016; Ramabulana *et al*., 2020). Using collision induced dissociation (CID) dependant collision energy remping, fragmentation patterns consistent to typical ion trap MS based MS^2 profiles were achieved and were also found to be sufficient to discriminate between isomers. Additionally, chromatographic elution order of the different regional isomers of CGAs were also used to aid in identification of these molecules (Clifford *et al*., 2008). As a consequence, a total of 21 CGAs with precursor ions at *m/z* 337, *m/z* 353, *m/z* 367 and *m/z* 515 were detected in the chromatogram of the methanolic extracts of *C. grandis* and *C. rehmannii* (Figure 4.1).

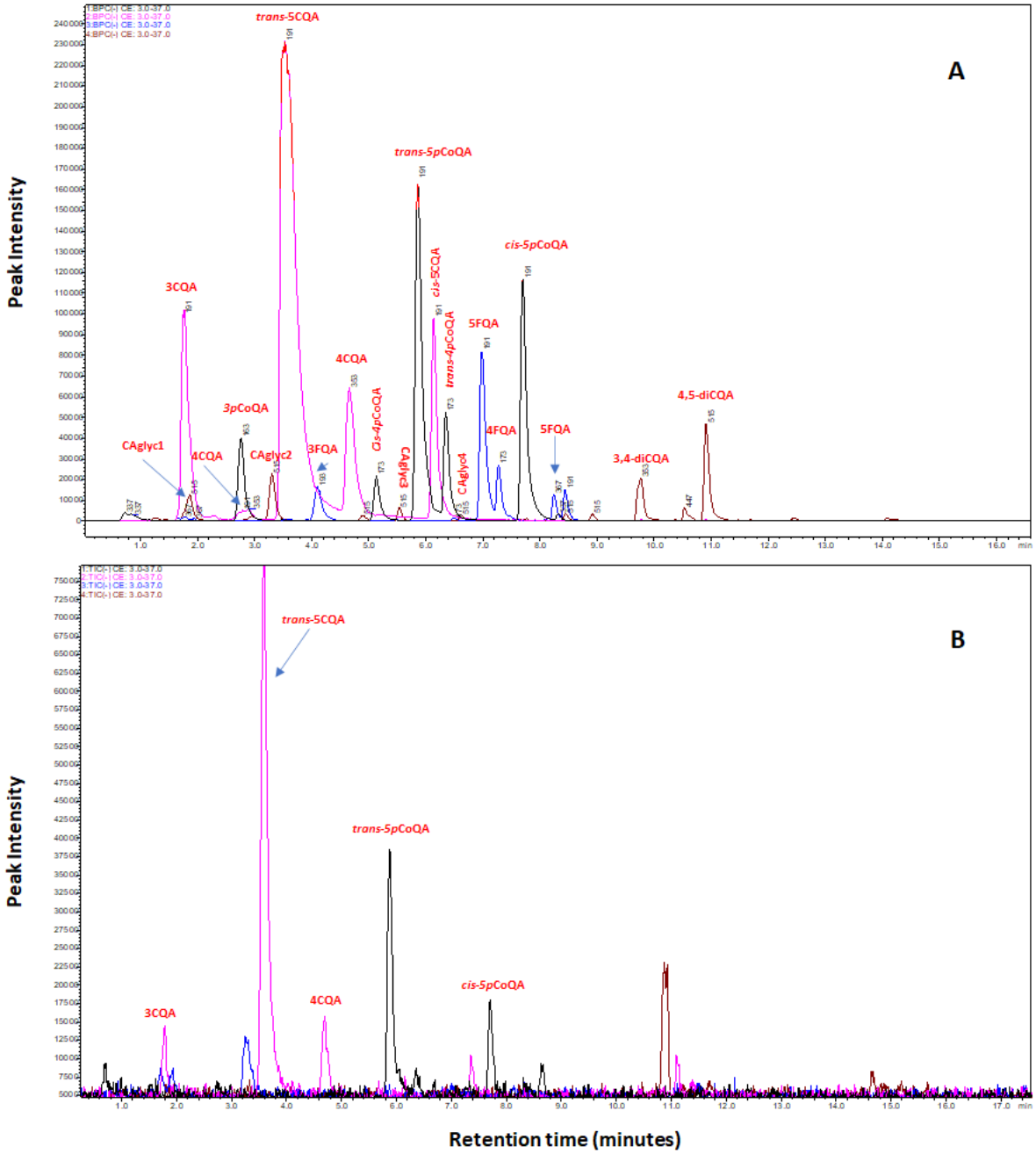


Figure 4.1: Representative UHPLC-qTOF-MS/MS chromatogram showing distribution patterns of chlorogenic acids found in *Coccinia grandis* (**A**) and in *Coccinia rehmannii* (**B**).

As part of the profile, other molecules which showed signs of being conjugate to HCAs were also identified. These were found to be flavonoid glycosides, of which the structures were further acylated by HCAs as reported recently (Nengovhela *et al*., 2020) (Figure 4.2). Similarly, MS^2 like spectra achieved through CID based approaches were also found to be sufficient for characterization of these molecules as well.

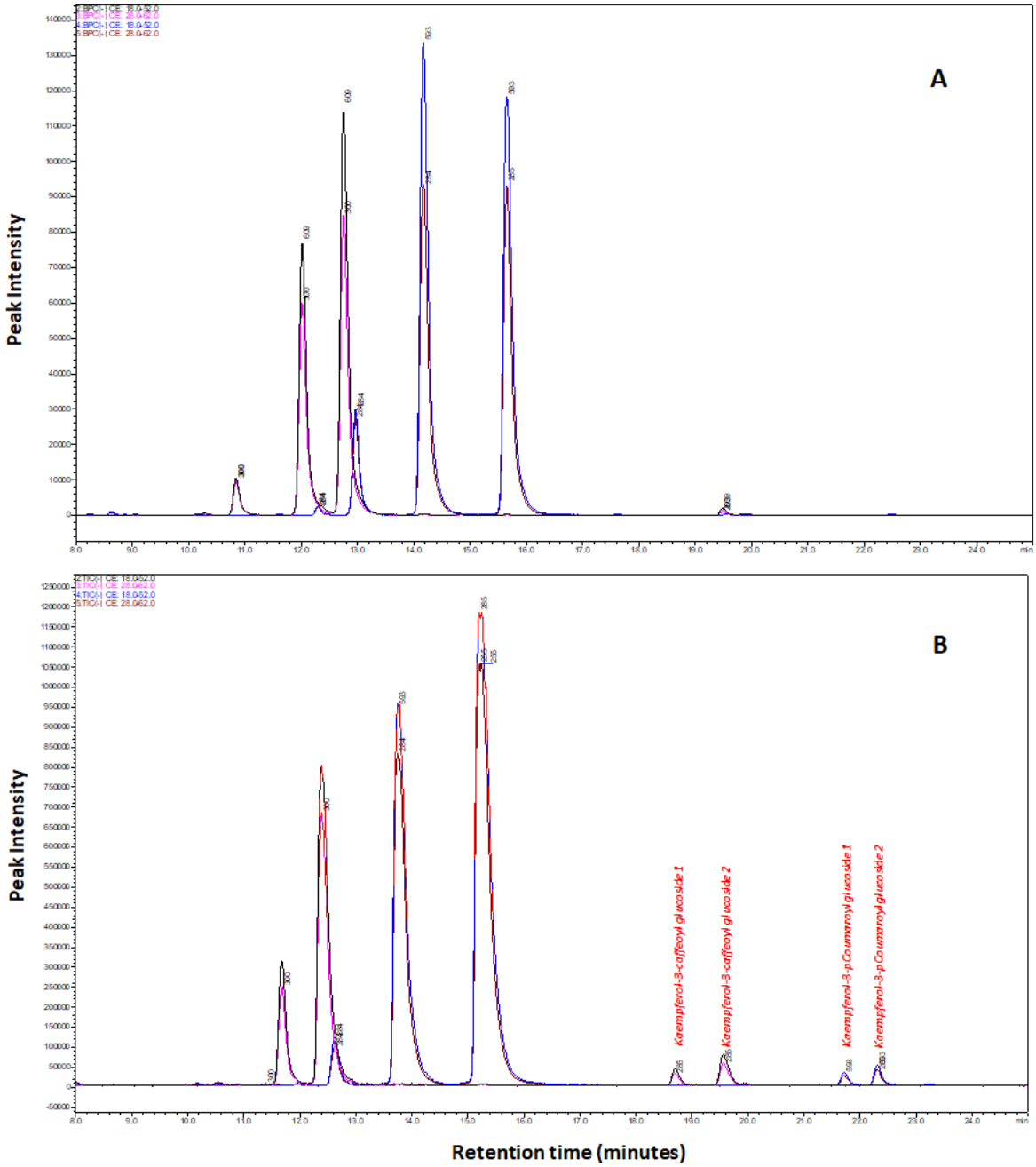


Figure 4.2: Representative UHPLC-qTOF-MS/MS chromatogram showing distribution patterns of flavonoids found in *Coccinia grandis* **(A),** flavonoids in *Coccinia rehmannii* which some flavonoids glycosides conjugate with hydroxycinnamic **(B)**.

4.3.1. *Characterization of p-coumaroylquinic acid*

Four peaks with precursor ions at *m/z* 337 (**1-5**) were detected in the chromatogram of *C. grandis* leaf extracts, whilst only one was peak was detected in *C. rehmannii*. Based on accurate mass and accompanying fragmentation patterns (MS^2) patterns, these molecules were identified as *para*-Coumaroylquinic acids

(*p*CoQA). Due to the presence of the product ions at *m/z* 191, additional peaks at *m/z* 163 and *m/z* 119, the first molecule (**1**) was identified as 3-pCoQA. The presence of product ion at *m/z* 173 was consistent with a derivative acylated on the 4-OH of the quinic acid and, as such, two peaks (**2 &4**) were identified as 4-pCoQA. Two peaks (**3 & 5**) which showed an exclusive product ion peak at m/z 191 were identified as 5-*p*-CoQA. A typical CID spectrum showing fragmentation pattern of *p*CoQA is shown in Figure 4.3A.

4.3.2. *Characterization of mono-acyl caffeoylquinic acids*

Five peaks were detected with precursor ion $[M-H]^-$ at *m/z* 353 (**6-10**). Similarly a peak with product ions at *m/z* 191 and *m/z* 179 at 50% of the base peak was identified as 3-CQA (**6**) (Figure 4.3B). Two peaks (**7 & 9**) were identified as isomers of 4-CQA due to the presence of a peak at m/z 173 (dehydrated product, which is only possible if acylation is on the 4-OH of the quinic acid). Two peaks with a single product ion at *m/z* 191 were identified as isomers 5-CQA (**8 & 10**).

4.3.3. *Characterization of mono-acyl feruloylquinic acids*

A similar approach as above was used to identify isomers of feruloylquinic acids (**11-15**), having a precursor ion $[M-H]^-$ at *m/z* 367. Based on fragmentation patterns, 3-FQA (**11**) produced a base peak at *m/z* 193 and an additional fragment ion at *m/z* 134 (Figure 4.3C). 4-FQA was identified by the presence of a base peak at *m/z* 173 and through this scrutiny two isomers (**13 & 14**) were identified. Two isomers with a base peak at *m/z* 191 were identified as geometrical isomers of 5-FQA (**12 & 15**).

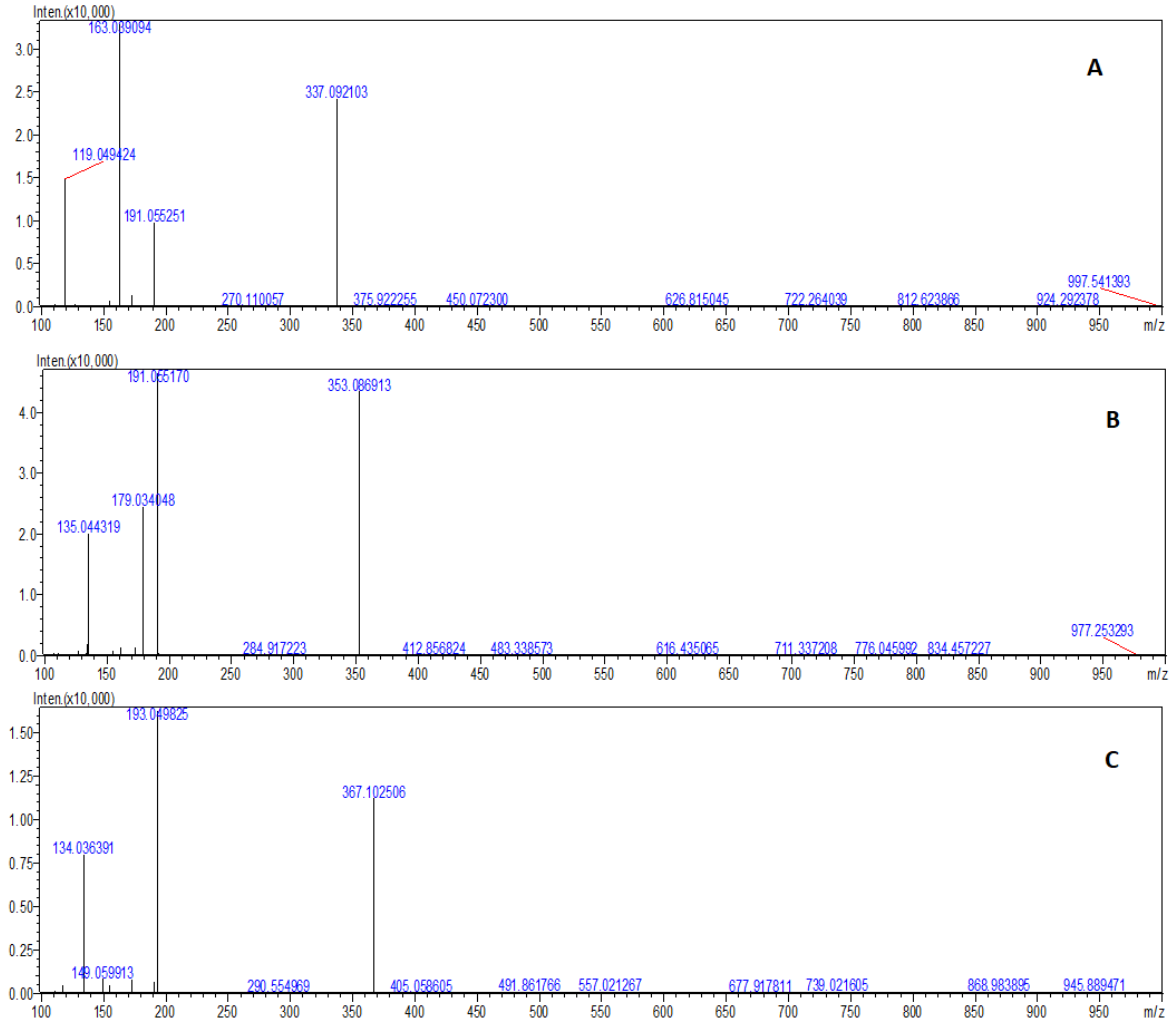


Figure 4.3: Representative ESI negative spectrum showing the fragmentation pattern of p-coumaroylquinic acid (**A**), Caffeoylquinic acid (**B**) and Feruloylquinic acid (**C**).

4.3.4. *Characterization of di-caffeoylquinic acids (diCQA)*

Di-caffeoylquinic acids are known for having a molar mass (Mr) of 516. In this study, two molecules of diCQA were identified (Table 1) having a precursor ion $[M-H]^-$ at 515 (**20-21**). Using the same approach as the one used for characterizing mono CGAs, the two molecules were identified as diCQAs with acylation on position 4, which suggest they could be geometrical isomers of either 3,4-diCQA or 4,5-diCQA (Figure 4.4A & B). Both compounds showed MS^2 base peaks at *m/z* 353, indicating loss of one caffeoyl residue, and the presence of secondary ions at *m/z* 173 and at *m/z* 191 in MS^2 was enough to account for 4-acyl.

4.3.5. *Characterization of caffeoylglucosides*

Four compounds with a precursor ion $[M-H]^-$ at *m/z* 515 were identified as caffeoylquinic acid glucoside isomers (**16-19**). These molecules were identified by production of MS^2 base peaks at *m/z* 353 $[CQA-H]^-$ and secondary ions at *m/z* 323 [caffeoyl glucoside $–H_2O-H^+]^-$, *m/z* 191 [quinic acid-H]$^-$, *m/z* 179 [caffeic acid-H]$^-$ and *m/z* 161 [caffeic acid-H_2O-H]$^-$ respectively (Figure 4.4C). The presence of four peaks with similar fragmentation patterns is an indication that they could either be positional or geometrical isomers thereof.

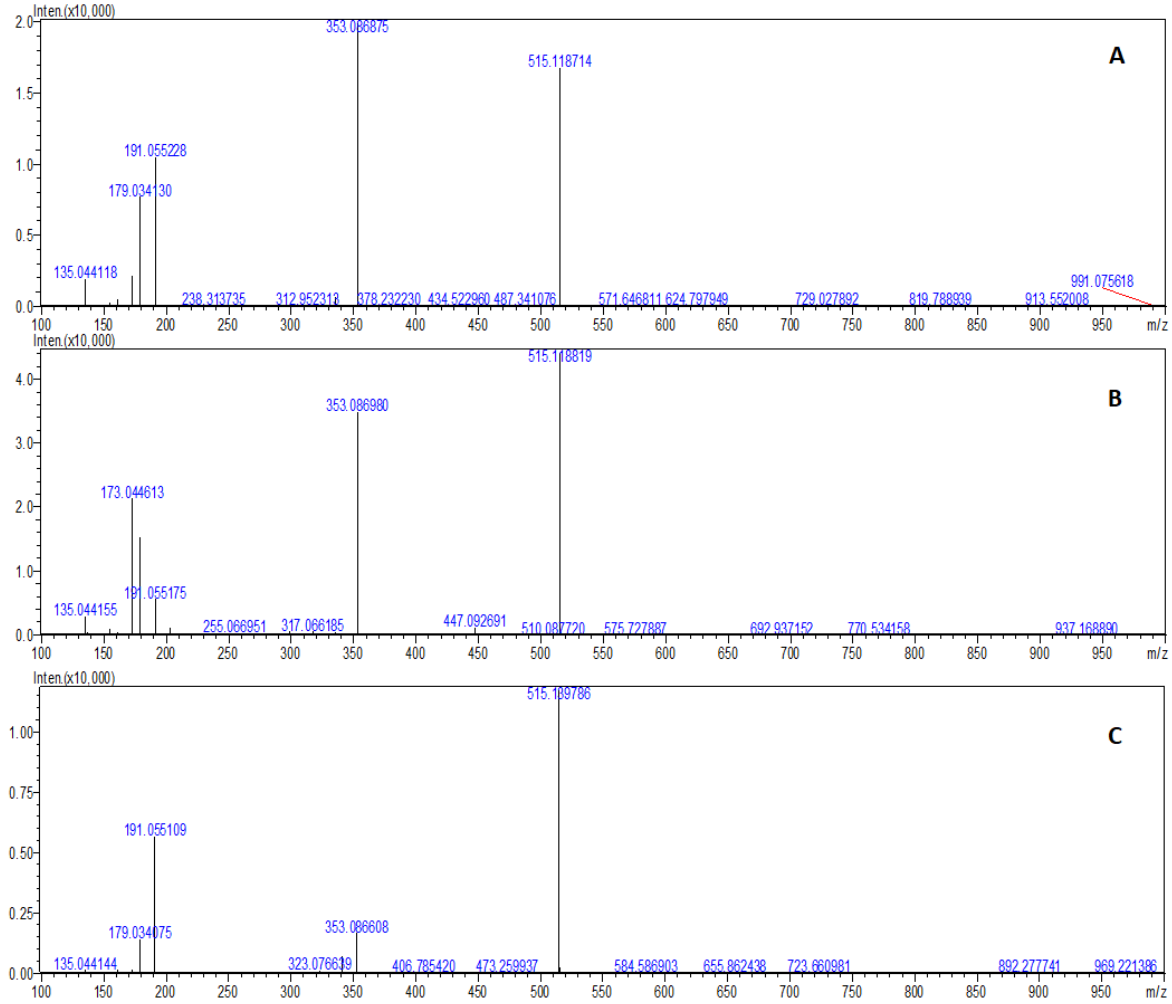


Figure 4.4: Representative ESI negative spectrum showing fragmentation pattern of 3,4 di-caffeoylquinic acid (**A**), 4,5 di-caffeoylquinic acid (**B**) and 3-caffeoylquinic acid glycoside (**C**).

4.3.6. *Characterization of flavonoids conjugated to HCA molecules*

Four flavonoid molecules conjugated to HCAs were identified in chromatogram of *C. rehmannii* leaf extracts. Two of each appearing with precursor ion at *m/z* 593 and the other two at *m/z* 609 respectively (Figure 4.2 & 4.5). Compounds **22** and **23** are isomers of a kaempferol glycoside which was further conjugated to a coumaroyl unit, whereas compound **24** and **25** are isomers of a kaempferol conjugated to a caffeoyl unit (Table 4.1). Interestingly none of these molecules were identified on *C. grandis* leaf extracts.

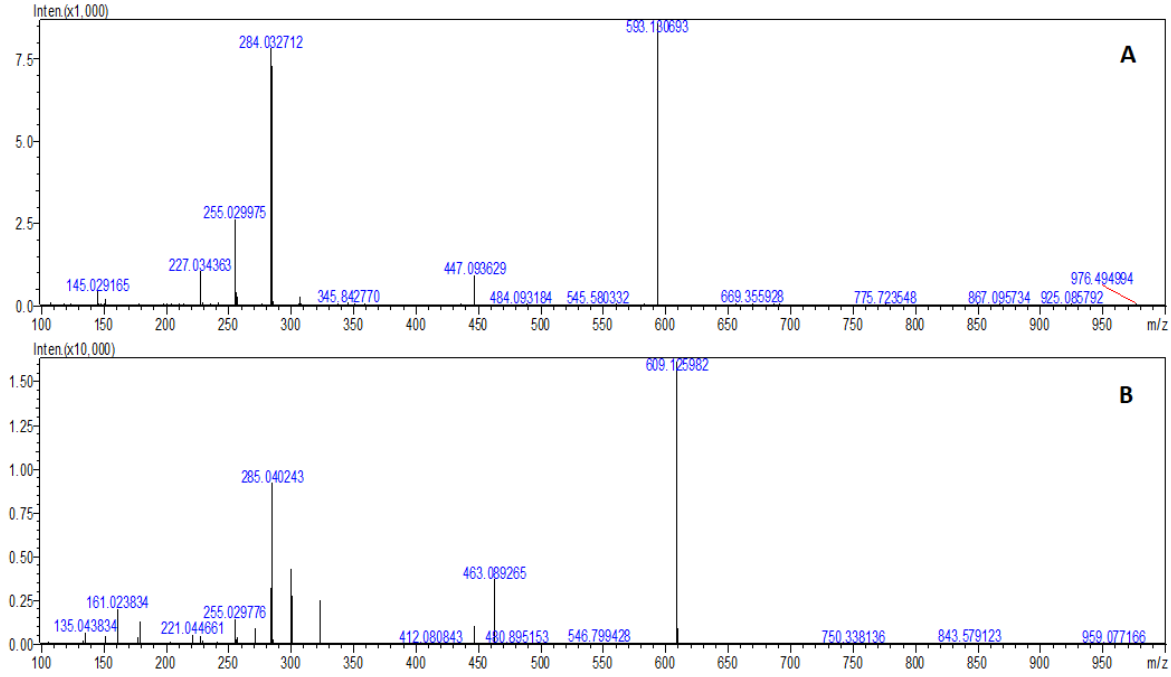


Figure 4.5: Representative ESI negative spectrum showing fragmentation patterns of kaempferol glycoside attached to a coumaric acid (**A**) and kaempferol glycoside attached to caffeic acid (**B**).

Table 4.1: Chlorogenic acids and flavonoids conjugated to HCA identified in both *Coccinia rehmannii* and *Coccinia grandis.*

NO	Mass (*m/z*)	Retention time (min)	Fragment ions (min)	Molecular formula	Metabolite identity	Abbreviation	*C. rehmannii*	*C. grandis*
1	337.0920	2.76	191.0553, 163.0391, 119.0494	$C_{16}H_{18}O_8$	3-*p*-Coumaroylquinic acid	3-*p*CoQA		●
2	337.0922	5.12	173.0446, 191.0553, 119.0494	$C_{16}H_{18}O_8$	4-*p*-Coumaroylquinic acid	4-*p*CoQA		●
3	337.0921	5.86	191.0552	$C_{16}H_{18}O_8$	5-*p*-Coumaroylquinic acid	5-*p*CoQA	●	●
4	337.0920	6.35	173.0445, 119.0494	$C_{16}H_{18}O_8$	4-*p*-Coumaroylquinic acid	4-*p*CoQA		●
5	337.0921	7.69	191.0552	$C_{16}H_{18}O_8$	5-*p*-Coumaroylquinic acid	5-*p*CoQA		●
6	353.0871	1.76	191.0553, 179.0340, 135.0442	$C_{16}H_{18}O_9$	3-Caffeoylquinic acid	3-CQA		●
7	353.0870	2.90	173.0446, 135.0443, 179.0340	$C_{16}H_{18}O_9$	4-Caffeoylquinic acid	4-CQA		●
8	353.0870	3.53	191.0553	$C_{16}H_{18}O_9$	5-Caffeoylquinic acid	5-CQA	●	●
9	353.0871	4.65	173.0440, 191.0552, 135.0442, 179.0341	$C_{16}H_{18}O_9$	4-Caffeoylquinic acid	4-CQA		●
10	353.0868	6.13	191.0552	$C_{16}H_{18}O_9$	5-Caffeoylquinic acid	5-CQA		●
11	367.1026	4.10	193.0498, 134.0363	$C_{17}H_{20}O_9$	3-Feruloylquinic acid	3-FQA		●
12	367.1026	6.98	191.0553, 134.0364	$C_{17}H_{20}O_9$	5-Feruloylquinic acid	5-FQA		●
13	367.1026	7.22	173.0446, 193.0497, 134.0363	$C_{17}H_{20}O_9$	4-Feruloylquinic acid	4-FQA		●
14	367.1023	7.47	173.0448, 191.0546, 134.0357	$C_{17}H_{20}O_9$	4-Feruloylquinic acid	4-FQA		●
15	367.1028	8.43	191.0553, 173.0444	$C_{17}H_{20}O_9$	5-Feruloylquinic acid	5-FQA		●

16	515.1398	1.86	353.0866, 191.0551, 179.0341, 135.0441	$C_{15}H_{18}O_9$	Caffeoylglucoside	CAglyc-1		●
17	515.1397	3.30	353.08667, 323.0764, 191.0553, 161.0235	$C_{15}H_{18}O_9$	Caffeoylglucoside	CAglyc-2		●
18	515.1398	5.53	353.0876, 191.0554, 179.0335, 135.0455	$C_{15}H_{18}O_9$	Caffeoylglucoside	CAGlyc-3		●
19	515.1397	6.62	323.0768, 191.0551, 161, 0238	$C_{15}H_{18}O_9$	Caffeoylglucoside	CAGlyc-4		●
20	515.1187	9.75	353.0869, 191.0552, 179.0341, 135.0441	$C_{25}H_{24}O_{12}$	3,4-di-Caffeoylquinic acid	3,4-di-CQA		●
21	515.1188	10.91	353.0870, 191.0552, 179.0340, 173.0446, 135.0442	$C_{25}H_{24}O_{12}$	4,5-di-Caffeoylquinic acid	,5-di-CQA		●
22	593.1282	12.96	284.0322, 119.0504	$C_{30}H_{26}O_{13}$	Kaempferol-3-*p*-Coumaroylglucoside	Isomer 1	●	
23	593.1288	12.98	284.0322, 119.0503	$C_{30}H_{26}O_{13}$	Kaempferol-3-*p*-Coumaroylglucoside	Isomer 2	●	
24	609.1257	11.66	463.0883, 285.0405,135.0444	$C_{30}H_{26}O_{14}$	Kaempferol-3-*p*-(6”-caffeoylglucoside)	Isomer 1	●	
25	609.1271	12.01	463.0887, 285.0403, 135.0443	$C_{30}H_{26}O_{14}$	Kaempferol-3-*p*-(6”-caffeoylglucoside)	Isomer 2	●	

4.3.7. Molecular networking

Molecular networking arranges the similarities of fragmentation spectra as cosine score and further groups the similar molecules in a cluster (network). Compounds within a molecular network are grouped according to their common MS/MS fragmentation patterns, which enables the distribution of structurally related molecules based on their product ion scans and abundances. In this study, the chemical diversity of HCA derivatives was revealed by a data dependant acquisition based tandem MS based MN of *C. grandis* and *C. rehmannii* (Figure 6). All molecular networks were calculated using data acquired at the same ionization mode and collision energy. The obtained molecular networks revealed the clustering of hydroxycinnamic acid-containing molecules. Herein, it can be seen that almost all the CGA molecules formed a very strong cluster.

From literature, it is known that CGAs share similar structural moieties such as the quinic acids from which the different HCAs acylates. Thus, upon MS/MS fragmentation, it is expected that all CGAs will produce similar product ions, hence clustering together on the molecular network (Figure 4.6A). Interestingly, it can be seen from the network achieved using extracts of C. rehmannii that the HCA derivatives conjugated to the flavonoids moiety also formed a unique cluster and this further explains the recently coined phenomenon of glyco-isomerization (Nengovhela et al., 2020). Briefly, the acylation of flavonoids glycosides by HCAs derivatives further increases/diversify the number of flavonoids that are found in *C. rehmannii*. This observation was absent in *C. grandis* and this further supports the MS/MS data which showed this species to only conjugates its metabolites to quinic acid, forming chlorogenic acids (Figure 4.6B).

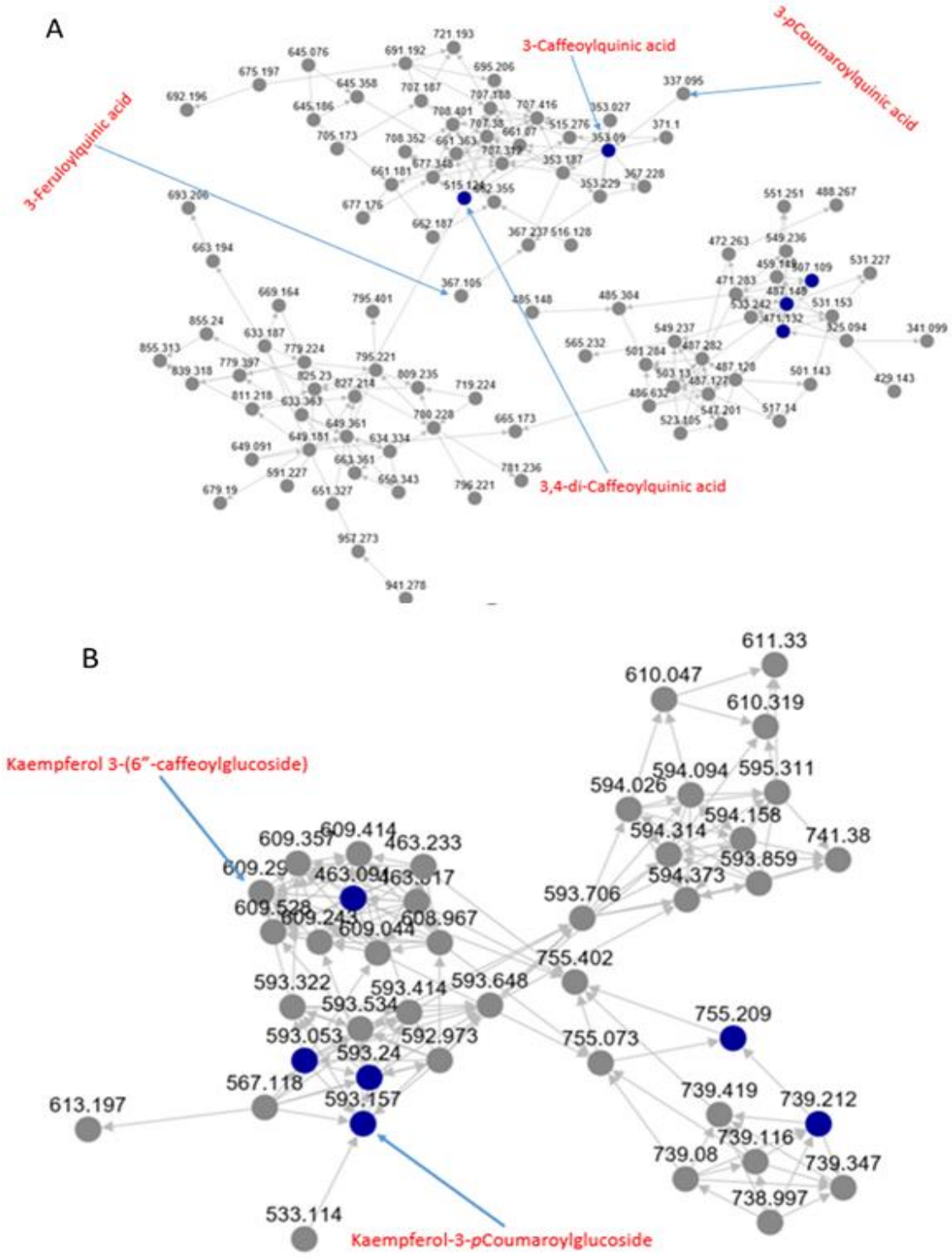


Figure 4.6: Molecular network applied to metabolites extracted from the two *Coccinia* species (**A**) *Coccinia grandis* and (**B**) *Coccinia rehmannii*.

4.4. Chemo-taxonomical and medicinal relevance of the results

Some plants are very difficult to distinguish through visual inspection and, as such, scientific methods such as chemo-taxonomical approaches are favoured in order to establish the fundamental differences thereof. However, in some plants the differences are not so significant or obvious and sophisticated technological means are needed for conclusive delineation. As seen from the results presented herein, the two *Coccinia* species produce cinnamic acid derivatives. However there are fundamental differences owing to how these cinnamic acids are conjugated. Here, *C. grandis* was shown to conjugate its cinnamic acids to quinic acid thereby forming the typical chlorogenic acids and, on the contrary, *C. rehmannii* conjugates its cinnamic acid pool onto the flavonoids glycosides. In our recent report, we have shown that cinnamic acid forms part of an interesting biologically controlled process deployed by plants during diversification of its metabolite pool (Nengovhela *et al*., 2020).

Acylation of flavonoids by cinnamic acids derivatives created a chemo-diverse pool of isomeric metabolites of differing in polarity, a phenomenon which has been recently referred to as glyco-isomerization (Nengovhela *et al*., 2020). Similarly, due to multiple OH group on the quinic acid, acylation of cinnamic acid result in formation of multiple regional/positional isomers of chlorogenic acids (Ncube *et al*., 2014). As seen from the chromatograms presented herein, the different isomers occupy different chromatographic spaces which further points to differences in polarities (Figure 4.1 & 4.2). As reported previously, polarity of metabolites is critically important for bio-availability thereof, thus how well a metabolite is absorbed and utilised by a human body once consumed (Erlund, 2004). Cinnamic acid also present another pharmacological dimension due to their susceptibility of photo-isomerization (*trans*/*cis* conversion). It has been well documented that when cinnamic acids are exposed to sunlight (or UV light), they undergo structural modification from their natural *trans* configuration to *cis* configuration (Clifford *et al*., 2008), a phenomenon which also lead a dramatic polarity switching (Masike *et al*., 2017a). Therefore, the above is an indication that this difference in cinnamic acid conjugation is a good taxonomical marker and could have adverse pharmacological implication as well.

4.5. Conclusions

The results of the current study have further showcased LC-MS as a feasible technique of metabolite identification. Herein, it has been successfully shown through LC-MS based metabolite fingerprinting that the two closely related *Coccinia* plants utilise contrasting means of cinnamic acid conjugations as means of metabolite diversification strategy. Here, *C. grandis* produced large number of chlorogenic acid derivatives as means of storing their cinnamic acids and *C. rehmannii* conjugated their cinnamic acids onto flavonoids molecules, even though few chlorogenic acids were also detected at trace level. Moreover, most of the metabolites exited as isomers of one another which could easily be distinguished based on their fragmentation patterns and elution profiles which were consistent with previously published data. Molecular networking (MN) was used herein, to substantiate the MS/MS results which revealed the two conjugation strategies used by these two *Coccinia* plants. The molecular networks show *lc-ms* to produce CGA molecules that forms a very tight cluster due to fragmentation similarities. On the other hand the two flavonoids harbouring HCAs were found to cluster with other flavonoids molecules produced by *C. rehmannii.* From these results it can be said that MN holds great potential for highlighting pharmacologically relevant metabolites through direct comparison with known (or already characterised) metabolites found within a plant species. Furthermore, these differences in conjugation and formation of isomeric molecules of different polarities could be further investigated for pharmacological purposes such as bio-availability during human consumption. These differences (in conjugations of HCAs) can also be used as chemo-taxonomical marker to distinguish between these closely related species. From the results, *C. grandis* produces QA-based CGAs and *C. rehmannii* produces mainly HCAs conjugated to flavonoid glycosides.

References

Alam, M.A., Subhan, N., Hossain, H., Hossain, M., Reza, H.M., Rahman, M.M. and Ullah, M.O. (2016). Hydroxycinnamic acid derivatives: A potential class of natural compounds for the management of lipid metabolism and obesity. Nutrition and Metabolism, 13; 27.

Chanda, J., Mukherjee, P.K., Kar, A., Maitra, P.K., Singha, S., Halder, P.K., Gajbhiye, R. and Vishnuvardh, R. (2020). LC–QTOF–MS-based metabolite profiling and evaluation of α-glucosidase inhibitory kinetics of *Coccinia grandis* fruit. Biomedical Chromatography, 34; e4950.

Clifford, M.N., Johnston, K.L., Knight, S. and Kuhnert, N. (2003). Hierarchical scheme for LC-MS^n identification of chlorogenic acids. Journal of Agricultural and Food Chemistry, 51; 2900-2911.

Clifford, M.N., Kirkpatrick, J., Kuhnert, N., Roozendaal, H. and Salgado, P.R. (2008). LC-MS^n analysis of the *cis* isomers of chlorogenic acids. Food Chemistry, 106; 379-385.

Clifford, M.N. and Madala, N.E. (2017). Surrogate standards: A cost-effective strategy for identification of phytochemicals. Journal of Agricultural and Food Chemistry, 65; 3589-3590.

El-Seedi, H.R., Taher, E.A., Sheikh, B.Y., Anjum, S., Saeed, A., AlAjmi, M.F., Moustafa, M.S., Al-Mousawi, S.M., Farag, M.A., Hegazy, M.E.F. and Khalifa, S.A. (2018). Hydroxycinnamic acids: Natural sources, biosynthesis, possible biological activities, and roles in Islamic medicine. Studies in Natural Products Chemistry, 55; 269-292.

Erlund, I. (2004). Review of the flavonoids quercetin, hesperetin, and naringenin. Dietary sources, bioactivities, bioavailability, and epidemiology. Nutrition Research, 24; 851-874.

Garg, N., Kapono, C.A., Lim, Y.W., Koyama, N., Vermeij, M.J., Conrad, D., Rohwer, F. and Dorrestein, P.C. (2015). Mass spectral similarity for untargeted metabolomics

data analysis of complex mixtures. International Journal of Mass Spectrometry, *377*; 719-727.

Holstein N. Monograph of *Coccinia* (Cucurbitaceae). (2015). PhytoKeys. 3; 1-166.

Imbumi MD. (2004). *Coccinia grandis* (L.) Voigt In: Grubben, G.J.H. & Denton, O.A. (Editors). PROTA 2: Vegetables/Légumes. [CD-Rom]. PROTA, Wageningen, Netherlands.

Ilkhanizadeh, B., Shirpoor, A., Nemati, S. and Rasmi, Y. (2016). Protective effects of ginger (*Zingiber officinale*) extract against diabetes-induced heart abnormality in rats. Diabetes and Metabolism Journal, 40; 46-53.

Leffers, A. (2003). Gemsbok Bean and Kalahari Truffle: Traditional plant use by Julhoansi in North-Eastern Namibia. Gamsberg Macmillan, Windhoek, Namibia.

Madala, N.E. and Kabanda, M.M. (2021). LC-MS based validation and DFT investigation on the antioxidant properties of clovamide:• OH and• OOH scavenging and Cu (II) chelation mechanisms. Journal of Molecular Structure, 130349.

Makita C., Chimuka L., Steenkamp P., Cukrowska E., Madala E. (2016). Comparative analyses of flavonoid content in *Moringa oleifera* and M*oringa ovalifolia* with the aid of UHPLC-qTOF-MS fingerprinting. South African Journal of botany, 105; 116-122.

Makola, M.M., Madala, N.E., Dubery, I.A., Steenkamp, P.A. and Kabanda, M.M., (2016a). Influence of the geometric isomers on the radical scavenging properties of 3, 5-dicaffeoylquinic acid: A DFT study in vacuo and in solution. Journal of Theoretical and Computational Chemistry, 15; 1650052.

Makola, M.M., Steenkamp, P.A., Dubery, I.A., Kabanda, M.M. and Madala, N.E., (2016b). Preferential alkali metal adduct formation by cis geometrical isomers of dicaffeoylquinic acids allows for efficient discrimination from their trans isomers during ultra-high-performance liquid chromatography/quadrupole time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry, 30; 1011-1018.

Martin, S., (2016). Identification and mechanistic study of novel drug metabolites by LC-MS. Sheffield Hallam University (United Kingdom).

Masike, K., Tugizimana, F., Ndlovu, N., Smit, E., du Preez, L., Dubery, I. and Madala, E. (2017a). Deciphering the influence of column chemistry and mass spectrometry

settings for the analyses of geometrical isomers of L-chicoric acid. Journal of Chromatography B, 1052; 73-81.

Masike, K., Mhlongo, M.I., Mudau, S.P., Nobela, O., Ncube, E.N., Tugizimana, F., George, M.J., Madala, N.E. (2017b). Highlighting mass spectrometric fragmentation differences and similarities between hydroxycinnamoyl-quinic acids and hydroxycinnamoyl-isocitric acids. Chemistry Central Journal, 11; 1-7.

Ncube, E.N., Mhlongo, M.I., Piater, L.A., Steenkamp, P.A., Dubery, I.A., Madala, N.E., 2014. Analyses of chlorogenic acids and related cinnamic acid derivatives from *Nicotiana tabacum* tissues with the aid of UPLC-QTOF-MS / MS based on the in-source collision-induced dissociation method. Chemistry Central Journal, 8; 1-10.

Nengovhela, N., Steenkamp, P.A. and Madala, N.E. (2020). LC-MS based metabolite fingerprinting of Coccinia plants reveals glycoisomerization as a structual diversification strategy in flavanoid chemical space. National Academy of science letters, 1-5.

Nobela, O., Renslow, R.S., Thomas, D.G., Colby, S.M., Sitha, S., Njobeh, P.B., du Preez, L., Tugizimana, F. and Madala, N.E. (2018). Efficient discrimination of natural stereoisomers of chicoric acid, an HIV-1 integrase inhibitor. The Journal of Photochemistry and Photobiology B: Biology, 189; 258-266.

Pekamwar, S.S. & Kalyankar, T.M. & Kokate, S.S. (2013). Pharmacological activities of *Coccinia grandis*: Review. Journal of Applied Pharmaceutical Sciences, 3; 114-119.

Pilon, A. C., Gu, H., Raftery, D., Bolzani, V. D. S., Lopes, N. P., Castro-Gamboa, I., & Carnevale Neto, F. (2019). Mass spectral similarity networking and gas-phase fragmentation reactions in the structural analysis of flavonoid glycoconjugates. Analytical chemistry, 91; 10413-10423.

Raheem, D.J., Tawfike, A.F., Abdelmohsen, U.R., Edrada-Ebel, R. and Fitzsimmons-Thoss, V. (2019). Application of metabolomics and molecular networking in investigating the chemical profile and antitrypanosomal activity of British bluebells (Hyacinthoides non-scripta). Scientific reports, 9; 1-13.

Rahmatullah, M., Azam, N.K., Khatun, Z., Seraj, S., Islam, F., Rahman, A., Jahan, S. and Aziz, S. (2012). Medicinal plants used for treatment of diabetes by the Marakh sect of the Garo tribe living in Mymensingh district, Bangladesh. AJTCAM. 9; 380-385.

Ramabulana, A.T., Steenkamp, P., Madala, N. and Dubery, I.A. (2020). Profiling of chlorogenic acids from *Bidens pilosa* and differentiation of closely related positional isomers with the aid of UHPLC-QTOF-MS/MS-based in-source collision-induced dissociation. Metabolites, 10, 178.

Ramabulana, T., Mavunda, R.D., Steenkamp, P.A., Piater, L.A., Dubery, I.A., Ndhlala, A.R. and Madala, N.E. (2017). Gamma radiation treatment activates glucomoringin synthesis in *Moringa oleifera.* Revista Brasileira de Farmacognosia, 27; 569-575.

Roleira, F.M., Varela, C.L., Costa, S.C. and Tavares-da-Silva, E.J. (2018). Phenolic derivatives from medicinal herbs and plant extracts: anticancer effects and synthetic approaches to modulate biological activity. Studies in Natural Product Chemistry, 57; 115-156.

Said, I.H., Truex, J.D., Heidorn, C., Retta, M.B., Petrov, D.D., Haka, S. and Kuhnert, N. (2020). LC-MS/MS based molecular networking approach for the identification of cocoa phenolic metabolites in human urine. Food Research International, 132; 109119.

Taur, D.J., Patil, R.Y., 2011. Mast cell stabilizing, antianaphylactic and antihistaminic activity of *Coccinia grandis* fruits in Asthma. Chinese Journal of Natural Medicine, 9; 359-362.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y & Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nature biotechnology, 34; 828-837.

Wang, Y., Feng, R., Wang, R., Yang, F., Li, P., & Wan, J. B. (2017). Enhanced MS/MS coverage for metabolite identification in LC-MS-based untargeted metabolomics by target-directed data dependent acquisition with time-staggered precursor ion list. Analytica chimica acta, 992; 67-75.

Wolfender, J. L., Litaudon, M., Touboul, D., & Queiroz, E. F. (2019). Innovative omics-based approaches for prioritisation and targeted isolation of natural products–new strategies for drug discovery. Natural product reports, 36; 855-868.

Zheng, X., Renslow, R.S., Makola, M.M., Webb, I.K., Deng, L., Thomas, D.G., Govind, N., Ibrahim, Y.M., Kabanda, M.M., Dubery, I.A. and Heyman, H.M., 2017. Structural elucidation of cis/trans dicaffeoylquinic acid photoisomerization using ion mobility spectrometry-mass spectrometry. Journal of Physical Chemistry Letters, 8; 1381-1388.

Chapter 5: May the real rutin please stand up? Highlighting the analytical and biological consequences of structural isomerism of plant metabolites

Abstract

The emergence of big data science such as the field of metabolomics has led to discovery and identification of large number of metabolites, some of which have pharmacological relevance. These metabolites are characterised by diverse physical and chemical properties owing to the structural diversity thereof. Some of the structural differences are very subtle and requires sophisticated analytical instrumentation such as liquid chromatography mass spectrometry (LC-MS). This has revealed interesting chemistry associated with isomers of plant metabolites, for instance, the overwhelming polarity change caused by minimal structural rearrangement. Herein, we show through a simple flavonoid molecule, rutin, that isomerism is an evolutionary strategy used by plants to diversify their metabolomes. However, due to structural similarities associated with these isomers, their identification requires special skills and advanced analytical instrumentation. Our observation highlight herein that isomerism could serve a feasible premise from which smart medicine of varying polarities can be developed for application in different *in vivo* environment of varying polarities. Finally, it important that scientist start to look at isomers closely rather ignoring them as structural artefacts because their existence could have favourable pharmacological consequences.

Significance

Due to emergence of scientific disciplines such as plant-based metabolomics, more and more metabolites are being discovered and described. Correct identification of metabolites requires a thorough understanding of the associated chemistry and of the limitations of the advanced analytical instrumentation and statistical methods employed.

Multiple forms of isomerism present a particular challenge which as young scientists working on the chemical characterization of under-utilized green leafy vegetables as an alternative source of nutraceuticals we have encountered many times, and sadly have found numerous examples in the literature where obvious errors have escaped the scrutiny of reviewers and editors, and have become established dogma.

Biological activity, or lack thereof, is a function of the three-dimensional structure of a molecule, determining for example its ability to bind to an enzyme, receptor or transporter, and the number of possible isomers can be surprisingly high as recently discussed, for example there are at least 22 C_6–C_3 phenyl-propanoids with an accurate mass of 182.0579 Da (Kay *et al*., 2020). Potentially amongst all these isomers, one is active and the rest are pharmacologically redundant.

As with many research groups we favor Liquid Chromatography Mass Spectrometry (LC–MS) because of its sensitivity and robustness, and ability to quantify as well as characterize many components in a single run, but even ultra-accurate MS cannot discriminate between isomers, although ion trap MS can through optimized fragmentation discriminate acyl-quinic acid (chlorogenic acid) regio-isomers,(Clifford *et al*., 2003) and the regio- and stereo-isomers of caffeoyl-glucose (Jaiswal *et al*., 2014). The associated geometric isomers investigated had identical fragmentation but can be distinguished after UV-irradiation (Clifford *et al*., 2008; Zheng *et al*., 2017).

It has long been known that sugars epimerise, for example the conversion of glucose to its C4 epimer galactose. The controlled chemical epimerisation is currently attracting attention because the ‘rare’ sugars that result may have valuable physico-chemical and biological properties (Wang *et al*., 2020). We have recently shown that plants have similarly controlled biological mechanisms for sugar epimerisation, using

a very sophisticated chemical strategy known as glyco-isomerization facilitating diversification of its metabolome (Nengovhela *et al*., 2020).

Glyco-isomerization encompasses epimerization, glycal migration and regional/positional glycosylation of aglycones, and we have shown that these processes are accompanied by dramatic changes in glycoside polarity, clearly illustrated during chromatographic analyses where molecules with similar mass are seen to elute in different chromatographic spaces, suggesting differences in hydrophobicity. Sometimes these differences are so great as to suggest that these are unrelated substances that by chance are isobaric.

However, with the aid of advanced MS procedures such as collision induced dissociation, ion mobility and adduct formation, it is now possible to recognise these compounds as isomers, discriminate between them and at least partially define their structural differences, such information of crucial chemophenetic importance,(Zidorn, 2019) and potentially also with reference to bioactivity. Previously such subtle differences in structure have gone largely unnoticed.

Flavonoids are abundant in plants and their content has been directly associated with the pharmacological potency of a host plant. Possibly the most studied flavonoid molecule is the widely distributed rutin (quercetin-3-*O*-rutinoside) which has been associated with various pharmacological properties. This molecule has an aglycone moiety (quercetin) bearing the disaccharide sugar rutinose (α-L-rhamnopyranosyl-(1→6)-β-D-glucopyranose). Quercetin-3-*O*-robinobioside is also known, with galactose replacing glucose. Quercetin-3-neohesperidoside (Al-Madhagy *et al*., 2019) is 1α-2 rhamnose-glucose compared with 1α-6, suggesting that positional isomerism of the glycoside part could also be a contributing factor towards the observed isomerism.

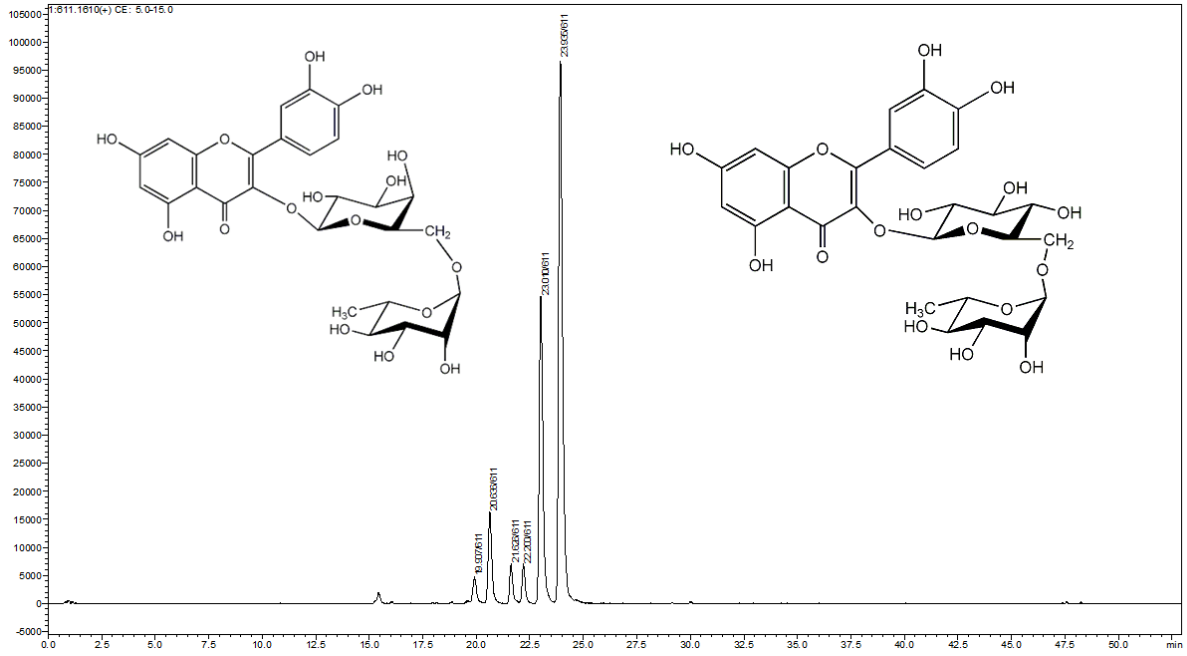


Figure 5.1: A representative liquid chromatography-quadrupole time of flight mass spectrometry (LC-QTOF-MS), base peak ion chromatogram showing MS/MS-based detection of isomeric molecules with precursor ion at m/z 611.161 which is typically associated with Rutin. Inserted are the structures of quercetin-3-robinobioside (left) and quercetin-3-rutinoside (right), which are presumed to be amongst the six peaks detected herein.

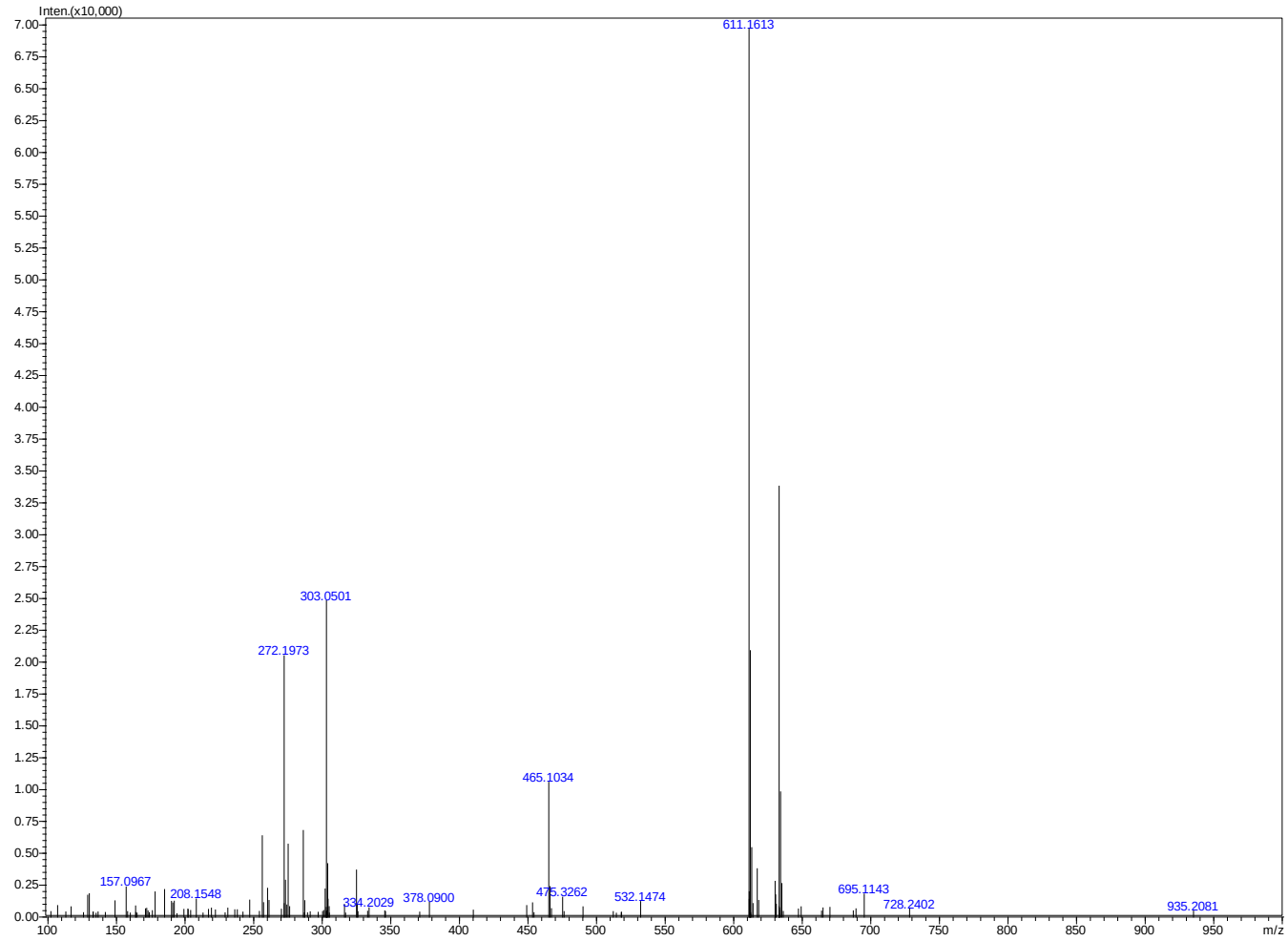


Figure 5.2: Representative MS spectrum showing fragmentation of Rutin Molecule generated using positive electro-spray ionization mode.

Using two plants that are closely related, namely *Coccinia rhemannii* and *Coccinia grandis* we have shown that they differ in their glyco-isomerization, that is the way they glycosylate their flavonoids. Accordingly, we designed a quantitative LC–MS method based on multiple reaction monitoring (MRM) developed with a commercially available rutin standard. This MRM-based method permits simultaneous detection and characterisation of the multiple fragmentation routes of a molecule, thereby allowing closely-related isomers to be distinguished from distinctly different isobaric substances, supported by preliminary reverse-phase separation by polarity. For some unknown reasons, this method also increases the sensitivity of the MS instrument, thereby allowing amplification of the signals associated with molecules which otherwise would be regarded as noise.

In the case of Rutin herein, the MRM-based analyses of the *C. grandis* extracts detected a maximum of six peaks sharing accurate mass indicating a molecular formulae of $C_{27}H_{30}O_{16}$, and similar fragmentation patterns, whereas *C. rhemannii* yielded only three, with rutin dominant in both, appearing at retention time 23.9 min (Figure 5.1). The UPLC-qTOF MS/MS spetra (Figure 5.2), showing fragmentation of the molecule assisted in the differentiation between the six isomers. Closer inspection of the fragmentation (Table 5.1) illustrates that these fragmentations differ with regard to the relatively intensity of the fragments at *m/*z 465 due to the loss of a terminal methylpentose (for example rhamnose, –146 Da) and *m/*z 303, due to the loss of the disaccharide (for example rutinoside or robinobioside, –308 Da) implying that there are some structural differences. Rutin and peak 1 have a ratio of *ca* 1.55, whereas for peaks 2, 4 and 5 this ratio is *ca* 1.8, and for peak 3 *ca* 3.1. Interestingly, while *C. grandis* has few flavonoids that *C. rhemannii*, it has more of the rutin isomers.

Table 5.1: The calculated peak ration between the dominant product ion peaks generated from the six isomeric molecules detected in the extracts of *Coccinia grandis* as info-graphically displayed in fig. 1.

	Peak 1	2	3	4	5	6 = rutin
m/z 303	1.32	6.25	0.99	0.83	3.7	3.5
m/z 465	0.86	3.4	0.32	0.45	2.1	2.2
Ratio	1.53	1.84	3.09	1.84	1.76	1.59

Therefore, in the absence of authentic standards, other methods such as visual comparison of fragmentation patterns or the use of surrogate standards (Clifford & Madala, 2017), each and every one of the peaks in Figure 5.1 could have been assigned as Rutin.

This misidentification could have unbearable consequences for both pharmacological and chemophenetic applications. Therefore, it is imperative that scientists should start to take note of the existence of these isomers and their unparallel biological potential. Elsewhere, the removal of sugar prior flavonoid analysis is habitual, however, we have shown that some glycosylation patterns are unique for some plants, and can then be used as taxonomic markers.

For instance *Moringa oleifera* and *Moringa ovalifolia* have been shown to produce three aglycone flavonoids *vis* quercetin, kaempferol and isorhamnetin but *M.ovalifolia* was only capable of attaching rutinoside sugar whilst *M.oleifera* was capable of attaching various types of sugars which were also shown to undergo further conjugation to other biologically relevant molecules such as hydroxymethylglutaroyl, malonyl and acetyl, thereby resulting in multiple flavonoids of differing polarities (Makita *et al*., 2016). However, these interesting differences could be abolished if sugars were removed prior LC-MS analyses, thus the two plants could be chemo-taxonomically indistinguishable. We hope that this letter will manage to show through a simple, but now complicated, molecule such as Rutin that plant metabolomes are complex entities and thankfully to the advancement of technology, more of these subtle differences can be brought to the fore leading to precise scientific conclusions. Lastly, this and other works further highlight glyco-isomerization as an evolutionary strategy of plants to maximize its metabolite complexity thereby increasing its defensive arsenal by deploying a “better safe than sorry” phenomenon (Ramabulana *et al*., 2016). As explained herein, formation of isomers always resulted in formation of metabolites with different polarities, an observation which can be exploited for pharmacological purposes where activities can be evaluated using models of different polarities such as blood or adipose tissue environments. Lastly, taxonomy of a plant is crucial and has implications in other scientific disciplines such as pharmacology and, as such, it is important to reiterate that glycosylation is a genetically coded process and by ignoring/neglecting its outcomes can have adverse biological consequences.

References

Al-Madhagy, S. A., Mostafa, N. M., Youssef, F. S., Awad, G. E. A., Eldahshan, O. A., & Singab, A. N. B. (2019). Metabolic profiling of a polyphenolic-rich fraction of Coccinia grandis leaves using LC-ESI-MS/MS and in vivo validation of its antimicrobial and wound healing activities. Food & Function, 10: 6267-6275.

Clifford, M.N., Johnston, K.L., Knight, S. and Kuhnert, N. (2003). Hierarchical scheme for LC-MS n identification of chlorogenic acids. Journal of agricultural and food chemistry, 51: 2900-2911.

Clifford, M. N., Kirkpatrick, J., Kuhnert, N., Roozendaal, H., & Salgado, P. R. LC-MSn analysis of the cis isomers of chlorogenic acids. Food Chemistry. 2008; 106: 379-385.

Clifford, M.N. and Madala, N.E (2017). Surrogate standards: A cost-effective strategy for identification of phytochemicals. Journal of agricultural and food chemistry, 65: 3589-3590.

Jaiswal, R., Matei, M. F., Glembockyte, V., Patras, M. A., & Kuhnert, N. (2014). Hierarchical key for the LC-MSn identification of all ten regio- and stereoisomers of caffeoyl glucose. Journal of Agricultural and Food Chemistry, 2014; 62: 9252-9265.

Kay, C. D., Clifford, M. N., Mena, P., McDougall, J. G., Andres-Lacueva, C., Cassidy, A. Wishart, D. S. (2020). Special Article Recommendations for standardizing nomenclature for dietary (poly) phenol catabolites. American Journal of Clinical Nutrition, 2020; 112: 1051-1068.

Makita, C., Chimuka, L., Steenkamp, P., Cukrowska, E. and Madala, E. (2016). Comparative analyses of flavonoid content in Moringa oleifera and Moringa ovalifolia with the aid of UHPLC-qTOF-MS fingerprinting. South African Journal of Botany, 105: 116-122.

Nengovhela, N., Steenkamp, P. & Madala, N.E. (2020). LC-MS based metabolite fingerprinting of *Coccinia* plants reveals Glycoisomerization as a structual diversification strategy in Flavanoid chemical space. National Academy Science Letters, 1-5.

Ramabulana, T., Mavunda, R.D., Steenkamp, P.A., Piater, L.A., Dubery, I.A. and Madala, N.E. (2016). Perturbation of pharmacologically relevant polyphenolic

compounds in Moringa oleifera against photo-oxidative damages imposed by gamma radiation. Journal of Photochemistry and Photobiology B: Biology, 156: 79-86.

Wang, Y., Carder, H. M., & Wendlandt, A. E. (2020). Synthesis of rare sugar isomers through site-selective epimerization. Nature, 578: 403-408.

Zheng, X. Y., Renslow, R. S., Makola, M. M., Webb, I. K., Deng, L. L., Thomas, D. G., Baker, E. S. (2017). Structural elucidation of cis/trans dicaffeoylquinic acid photoisomerization using ion mobility spectrometry-Mass Spectrometry. Journal of Physical Chemistry Letters, 8: 1381-1388.

Zidorn, C. (2019). Plant chemophenetics: A new term for plant chemosystematics/plant chemotaxonomy in the macro-molecular era. Phytochemistry. Journal of phytochemistry, 163: 147-148.

Chapter 6: General conclusions

In this study, LC-MS-based metabolomics approaches were successfully used to investigate, profile and compare metabolite composition from two closely related *Coccinia* species (*C. grandis* and *C. rehmannii*). These two plants were found to be rich in flavonoid molecules (quercetin and kaempferol) and hydroxyl cinnamic acid (HCA) derivatives (also referred to as cinnamic acids). Here, a new phenomenon referred to as glyco-isomerization was found to contribute towards metabolome complexity of these two plants. Briefly, this glyco-isomerization can be defined as a biochemical strategy used by plants through their glycosylation machinery to alternatively/differentially attach sugar molecules onto the core aglycone molecule (flavonoids in this case). Herein, flavonoids with either *di*- and/or *tri*- saccharides were found to be prone to this glyco-isomerization. Interestingly, some of the formed isomers, probably positional isomers, were characterised by similar sugars which were attached to different position on the aglycone and were found to elute at different regions of the chromatogram. The difference in elution order during reverse phase chromatography is a characteristic of differences in polarity, which can have effects depending on the bio-availability thereof, during human consumption. Apart from positional isomerism and differential glycosylation, other molecules such as different forms of cinnamic acid (coumaric acid and caffeic acid) were also found to attach to these sugar moieties, thereby contributing towards diversification of metabolites in these plants. Ironically, only *C. rehmannii* was found to attach these cimmanic acids on its flavonoids glycoside molecules, a phenomenon which was absent in *C. grandis.*

Through the use of multivariate statistical models such as XCMS online software, it was possible to annotate most of the profiled flavonoid metabolites. Further analysis of the data revealed *C. grandis* to produce chlorogenic acids, which are esters of cinnamic acids to a quinic acid moiety. Interestingly, there were very few chlorogenic acids detected in *C. rehmannii.* Therefore, all the findings of the current study show that the two plants use different strategies to store and diversify their cinnamic acid composition, with *C. rehmannii* conjugating its HCAs onto flavonoids molecules and *C. grandis* producing chlorogenic acids.

The analytical challenges posed by the presence of isomers in plants were also highlighted by monitoring the presence of Rutin. Several chromatographic peaks with similar MS signals to Rutin were identified and couldn't be distinguished. Finally, the current study is a classical example which shows metabolome complexity to be a strategy of plants to diversify their metabolites. The two understudied plants were shown to use glyco-isomerization and conjugation to diversify its metabolites composition. Therefore, these two biochemical processes can be regarded as taxonomical markers that can be used to distinguish between these two closely related plants. Further studies can be designed to investigate the pharmacological importance of these plants, such as evaluation of their anti-diabetic properties, owing to the presence of HCA and flavonoids derivatives.