Metabolomic exploration of pharmacologically relevant metabolites in *Moringa oleifera* and *Moringa ovalifolia* through the use of UPLC-qTOF- MS and multivariate models

By

Charlene Sampu Makita

A Thesis submitted to the Faculty of Science

University of the Witwatersrand

Johannesburg, 2017

DECLARATION

I declare that this Thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

(Signature of Candidate)
----------2nd----------Day of--------June----------2017

Supervisor:
Prof Luke Chimuka (WITS, School of Chemistry)

Co-supervisor:
Dr. Edwin Ntakadzeni Madala (The University of Johannesburg, Biochemistry Department)

Abstract

Plant metabolomics is considered a holistic qualitative and quantitative analysis of primary and secondary metabolites under specific conditions. Metabolomics provides functional information that is important in the biological studies of plants. Moreover, chromatographic techniques in combination with mass spectrometry are presently commonly utilized analytical technologies that are used in metabolomics studies.

Moringa oleifera (M. oleifera) is an acclaimed medicinal herb and is considered one of the most superior plants based on its nutritional and medicinal attributes. Signature compounds such as chlorogenic acids and flavonoids found in *Moringa* plants are considered beneficial as they are responsible for the health properties found in these plants. Out of the 13 species belonging to the *Moringaceae* family, *M. oleifera* is the most widely studied.

Furthermore, plants are known to produce compounds, such as flavonoids, as a response to biotic stresses (pest resistance) and abiotic stresses (harsh environmental conditions such as drought or cold). Other species belonging to the *Moringaceae* family such as *Moringa ovalifolia* (*M. ovalifolia*), located in the dry desert and semi-desert areas of Namibia, has not been reported much in literature. Moreover, there is very little work done on other classes of compounds that are possibly present in *M. ovalifolia*. Therefore, the significant correlation between desert plants and their possible increased flavonoid content is of interest in this study and is worth investigating. Moreover, in an effort to compare *Moringa* species in that regard, a holistic approach to metabolite fingerprinting using UPLC-qTOF-MS was thus employed for the characterization of possible metabolite markers and taxonomical differences in *M. ovalifolia* and *M. oleifera* plant species (**Paper I)**.

From the results, 17 flavonoid compounds were identified. Interestingly, M. *oleifera* and *M. ovalifolia* had a similar aglycone profile, however, they had a different sugar moiety. Rutinoside sugar moieties were only found to be present in *M. ovalifolia* indicating that it is only capable of attaching the rutinoside sugar moiety to its flavonoid skeleton. *M. oleifera,* however, is capable of attaching many sugar moieties to its flavonoid skeleton. The outcomes of this research additionally

demonstrated for the first time the significance of sugar attachment for taxonomical classification of related species and that the contrasts between the grouped species are likely because of genetic variations as opposed to environmental influences.

This study further provided potential chemotaxonomic biomarkers and a classification model was generated to classify and differentiate between closely related *Moringa* species utilizing multivariate analysis (**Paper III**).

Moreover, chlorogenic acids which are esters formed between derivatives such as cinnamic acid and quinic acid molecules, were identified in *M. ovalifolia* plants species for the first time (**Paper II**). Chlorogenic acids are difficult to identify and differentiate since they are structurally complex. This study demonstrated the effectiveness of an UPLC-ISCID-MS/MS based platform to aid in the profiling of isomers and derivatives of chlorogenic acids present in *M. ovalifolia*. This approach further showed *M. ovalifolia* as a valuable source of molecules with therapeutic potential.

Based on the classification model generated in **Paper II**, the biomarkers identified were utilized in order to differentiate between 12 *M. oleifera* cultivars (**Paper III)**. Three sought after flavonoid compounds, namely, quercetin rutinoside (rutin), kaempferol rutinoside and isorhamnetin rutinoside were identified as potential chemotaxonomic markers amongst the 12 *Moringa* cultivars. Metabolite distribution patterns of the 12 cultivars were analyzed using a metabolomics approach with the aid of UHPLC-qTOF-MS in combination with multivariate data models such as principal component analysis (PCA), hierarchical clustering analysis (HCA) and box-whiskers plot. According to the results, three main cultivars, namely: TOT4977, CHM and TOT5330 were identified as potential cultivars for pharmacological and nutritional purposes according to the presence and abundance of the three studied rutinoside bearing flavonoid molecules.

A dedication to my parents Dr. Makita and Dr. Mwacha and to my siblings Jessica Makita and Cindy Makita.

Acknowledgements

- To everyone that has played an important role in accomplishing this work:

  My first acknowledgement goes to my academic supervisor Prof Luke Chimuka. Thank you for your continued support and guidance academically for the past four years. I feel privileged to have worked under your supervision.

  To my co-supervisor Dr. Edwin Ntaka Madala. Thank you for taking me under your wing for the past two years. Not only have you taught me the art of metabolomics but you helped me build a strong work ethic that will indeed help me greatly in my future pursuits. Thank you very much.

  I would like to extend my gratitude to Prof Ewa Cukrowska, Prof Hlanganani Tutu from the Environmental Analytical Chemistry Research Group, Prof Ashwell Ndhlala (from the Agricultural Research Council) and Prof Paul Steenkamp (from CSIR) for analysis of my data and dedicated assistance throughout this journey.

- To my friends:

  I would like to acknowledge my fellow colleagues from the Environmental Analytical Chemistry Research Group who have contributed immensely to my professional and personal time at Wits University: X. Maputsoe, O. Molefe, M. Kgatitsoe, Dr. N. Tavengwa and S. Ncube.

  I am indebted to the following people from the Department of Biochemistry and Plant Research Group: R. Ramabulana and E. Ncube for helping and giving me the opportunity to grow in knowledge.

- Above all I would like to thank **THE MOST HIGH (Yahuah Elohim) AND YAHOSHUA.**

With you I was able to do *exceedingly abundantly above all that I could ask or think (Ephesians 3:20). Halleluyah!*

List of Publications emanating from the study

1. 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. S. Afr. J. Bot. 105: 116-122.

2. Makita, C., Chimuka, L., Steenkamp, P., Cukrowska, E., Madala, E., Kandawa-Schutz, M., Ndhlala, A.R. (2016). UPLC-qTOF-MS profiling of pharmacologically important chlorogenic acids and associated glycosides in *Moringa ovalifolia* leaf extracts. Afr. J. Bot. 108: 193-199.

3. Makita, C., Chimuka, L., Steenkamp, P., Cukrowska, E., Madala, E., Kandawa-Schutz, M., Ndhlala, A.R. Rutinoside-bearing flavonoids are chemo-taxonomical markers for efficient identification of pharmacologically potent *Moringa oleifera* Lam. Cultivars. Front. Pharmacol. (Submitted).

Contributions of the author

Paper I. Principal author involved in planning, performed the methanol extraction, characterization, evaluation of the results and writing of the article. Co-authors revised the draft manuscript and made suggestions for improvement.

Paper II. Principal author involved in planning, performed the methanol extraction, characterization, evaluation of the results and writing of the article. Co-authors revised the draft manuscript and made suggestions for improvement.

Paper III. Principal author involved in planning, performed the methanol extraction of the twelve cultivars, characterization and evaluation of the results and writing of the article. Co-authors revised the draft manuscript and made suggestions for improvement.

All papers are reproduced with kind permission from the respective copyright holders.

List of conference seminars and seminar presentations

1. Makita C., Chimuka L., Cukrowska E., Madala N.E. Comparitive Analysis of the Flavonoid content in *Moringa* species with the aid of UPLC-qTOF MS fingerprinting. 42nd National Convention South African Chemical Institute, 29th - 4th December 2015. Durban, South Africa. *Oral Presentation*.

2. Makita C., Chimuka L., Cukrowska E., Madala N.E. Comparitive Analysis of the Flavonoid content in *Moringa* species with the aid of UPLC-qTOF MS fingerprinting 6th SEANAC Conference, 19th - 22nd June 2016. Kasane, Botswana. *Poster Presentation*.

TABLE OF CONTENTS

List of Figures

List of Tables

List of abbreviations

BPI	Base peak intensity
CE-MS	Capillary electrophoresis-mass spectrometry
CGAs	Chlorogenic acids
Da	Dalton
ESI	Electrospray ionization
ESI-MS	Electrospray ionization mass spectrometry
FT-IR	Fourier-transform infrared
GC-MS	Gas chromatography- mass spectrometry
ISCID	In-source collision induced dissociation
HCA	Hierarchichal clustering analysis
HPLC	High performance liquid chromatography
LC-MS	Liquid chromatography- mass spectrometry
LSE	Liquid-solid extraction
MAE	Microwave- assisted extraction
METLIN	METabolite LInk
MS/MS	Tandem mass-spectrometry
MVDA	Multivariate data analysis
m/z	Mass-to-charge ratio
NMR	Nuclear magnetic resonance
LC-qTOF-MS	Liquid chromatography- quadrupole time-of-flight mass - spectrometry
MDASA	*Moringa* development association of South Africa
OPLS-DA	Orthogonal projection to latent structures discriminant analysis
PCA	Principal component analysis
PC1	First principal component
PC2	Second principal component
PLE	Pressurised liquid extraction
RT	Retention time
SFE	Supercritical fluid extraction
SPE	Solid phase extraction

UAE	Ultrasound-assisted extraction
UGTs	Glycosyltransferase
UHPLC-MS	Ultra-high pressure liquid chromatography
UPLC-qTOF-MS	Ultra-performance liquid chromatography quadrupole time-of-flight Mass- spectrometry
UV	Ultraviolet

Chapter 1: Introduction

This chapter gives a background to the study of this thesis as well as the outline on how the work is presented. The references arising from this section are listed under the Literature Review.

1.1 Background of the study

The application of medicinal plants as alternative medicine has various aspects directed at improving human health care. The large amount of analytical targeted analyses of indigenous plants have contributed to a better knowledge of plant chemistry (Flamini, 2013; Roullier-gall *et al.*, 2014). In recent time, a return is noticed to the traditional folk medicine and use of extracts of natural origin as medicines or health-promoting agents resulted in increased interest in substances with a specific biological activity (Sasidharan *et al*., 2011). Plants produce a variety of primary and secondary metabolites that play important roles in ecological interactions in the environment. It is well established that plants are producing these compounds in order to fight competing species, ensure appropriate space, energy, water and mineral resources for living and as a response to biotic and abiotic stress (Ramakrishna and Ravishankar, 2011).

Natural medicinal and nutritional plants, such as *Moringa oleifera*, is gaining popularity in many African countries. This plant is known for its dietary esteem and its extraordinary therapeutic properties. Different parts of this plant have been used for different purposes, for instance: the leaves, stem bark, roots, fruit and flowers may be used as anti-microbial agents. The pods and seeds may perform as cardiac stimulants (Ewansiha *et al*., 2014). Earlier studies have confirmed *M. oleifera* leaves to contain strong antioxidant and scavenging activities (Chumark *et al*., 2008).

Of particular interest are antioxidants derived from plants since these compounds play an important role in preventing diseases caused by oxidative stress, cancer, heart disease and aging (Kumar and Pandey, 2013). Phenolic compounds, like flavonoids, are commonly considered as being responsible for antioxidant properties of sage plants and for their high and diverse biological activity (Kumar and Pandey). Thus, it is important and reasonable to study these components in new varieties of different plants. Such important plants growing in sub-tropical and tropical regions, including South Africa and Namibia, are flowering trees from the *Moringaceae* family. In addition, metabolite fingerprints could be used as possible metabolite markers in plant taxonomy.

The study of metabolomics is an expanding technique which involves a novel approach of identifying metabolite profiles in a biological system such as cell, tissue or organism and is very relevant for Natural Product research. Due to its effective and efficient characteristics, this method is ideal for the study of herbal medicine. Utilising a metabolomics approach has shown potential for the study of pharmacological and nutritional related plants such as *M. oleifera*. The composition of bioactive secondary plant metabolites present in *M. oleifera* is determined by factors such as environmental factors and genetic factors, therefore, metabolomics allows for the profiling of the constituents resulting from these factors (Roullier-gall *et al*., 2014). Analytical analyses require the aid of powerful and high resolution systems such as FTIR-MS (Fourier Transform Infrared - Mass Spectrometry), GC-MS (Gas Chromatography - Mass Spectrometry), LC-MS (Liquid Chromatography - Mass Spectrometry) and NMR (Nuclear Magnetic Resonance) spectroscopy in combination with multivariate statistical analysis. The intention of using this process is not only confined to the identification of metabolites, but can extend in determining the biological differences between organisms by comparing the metabolite distribution patterns. Metabolite patterns may provide further knowledge of the physical state, biochemical processes and medicinal attributes of the organism, which we attempted to achieve in the study of *M. oleifera* plant extracts (Dunn and Ellis, 2005; Fukusaki and Kobayashi, 2005; Idborg *et al*., 2005).

Furthermore, the use of chemometrics has become an important discipline in analytical chemistry which includes improving and optimizing experimental methods, calibration of analytical instrumentation and utilising advanced procedures for analysis of chemical data (Duarte, 2006). With these approaches and the addition of multivariate data, one can further highlight the similarities and differences among the data (Ali *et al.,* 2012).

LC-MS in particular is the preferred method for identification of non-volatile compounds such as amino acids, phenolic compounds and fatty acids (Tolin *et al*., 2012; Jaitz *et al*., 2010; Della Corte *et al*., 2013). Many researchers have published LC-MS metabolite profiling studies utilising UPLC (ultra-performance liquid chromatography) with qTOF-MS (quadrupole time-of-flight mass spectrometry). A semi-targeted analysis approach with the aid of LC-MS has been connected to

numerous *M. oleifera* studies with the attempt to annotate flavonoid and chlorogenic acid molecules (Madala *et al*., 2014; Makita *et al*., 2016; Ncube *et al*., 2014). In contrast, however, very few studies have used this technique for semi-targeted metabolomics analysis of chlorogenic aids found in *Moringa ovalifolia* (*M. ovalifolia*), a species equally belonging to the *Moringaceae* family and is native to Namibia. Similarly, very few studies address this technique for the discrimination of *Moringa* cultivars from different geographical locations based on their metabolic content using UPLC-qTOF-MS together with multivariate statistical analysis. We show in this study that this work is at the cutting edge of plant metabolomics, as we have characterized different *Moringa* cultivars from various geographic roots, but now cultivated at the same place, through the identification of common metabolites. The background literature pertaining to this study is provided in depth in the Literature Review section (**Chapter 2**).

Chapter 2: Literature Review

Chapter 2 focuses on the literature review. The topics addressed outline practical and scientifically based approaches for conducting plant metabolomics studies. Extraction and analytical techniques utilized for identification and assurance of metabolites are discussed in detail.

2.1 Introduction to metabolomics

The study of metabolomics is a scientific examination of chemical processes due to the presence of metabolites present in a biological sample. It is an "omics" approach that has emerged in recent years as a tool specifically for the analyses and understanding of metabolites and their metabolic processes (Rochfort, 2005; Tweeddale *et al*., 2006; Weckwerth, 2003). Metabolomics examinations takes a preview of the physiological status of the metabolites present in the cell or tissue at a particular point in time (Tugizimana *et al.*, 2013; Ncube *et al*., 2014; Ramabulana *et al*., 2015). This approach can be used to describe the key links between genotype and phenotype by describing genetic and environmental modifications due to biological systems (Fiehn 2002; Suberu *et al*., 2016). This differs from traditional hypothesis driven "omics" approaches such as genomics, proteonomics and transcriptomics as metabolomics is a data driven scientific discipline (Ivanisevic *et al*. 2016; Kell and Oliver, 2004). Moreover, metabolomics compared to traditional "omics" approaches reflect changes in the phenotype of a particular organism or tissue (Gomez-casati *et al*., 2013). The advantages of this approach is that it's relatively inexpensive, rapid and an automated technique (Gomez-casati *et al*., 2013).

Metabolomic analysis focuses on understanding complex molecular problems in a biological system and adaptation to the environment causes stresses on biological systems (Moco *et al*., 2007; Rochfort *et al*., 2010). This phenomenon thus causes alterations to the genome as well as to the primary and secondary metabolites (Moco *et al*., 2006). Metabolites are organic compounds present in a biological system that can further characterize the biochemical phenotype of a cell or tissue under specific conditions (Dunn and Ellis, 2005; Fiehn, 2002; Fukusaki and Kobayashi, 2005; Idborg *et al*., 2005). In plants, identifying primary and secondary metabolites is beneficial for the enhancement of crops as well as for Natural Product research, ecology and plant biochemistry (Moco *et al*., 2006).

Researchers can go as far as using metabolomics to analyse the response of cells due to the environmental changes by quantifying the change in concentration of the metabolites (Sumner *et al.*, 2003). Metabolites such as amino acids, organic acids, fatty acids, vitamins and inorganic compounds are particularly challenging to analyse due to their complex nature (Rochfort *et al.*, 2010). According to Rochfort *et al*., (2010), the plant system is estimated to comprise of over 200, 000 metabolites. In one system alone, 50 flavonoids were identified *in M. oleifera* Lam leaves (Rodríguez-Pérez *et al*., 2015). In this manner, different logical techniques and terms have been utilized as a part of metabolomics studies primarily to accomplish an intensive investigation of the plant metabolome, and is listed in Table 1 below.

Table 1: Terms utilised for metabolomic analysis (Allwood *et al*., 2011; Tugizimana *et al.*, 2013; Dunn and Ellis, 2005).

Terminology	Definition
Metabolome	The complete set of metabolites present in an organism or system
Metabolomics	Identification and quantification of metabolites in a biological system
Metabolite	Compounds involved in cellular metabolomics reactions
Metabolite profiling	Involves the identification and quantification of metabolites with the aid of chromatography and advanced detection methods
Metabolite fingerprinting	Involves metabolite profiles obtained from crude samples without identification
Metabolite "targeted approach"	Qualitative and quantitative analysis of selected metabolites
Metabolite "untargeted approach"	Comprehensive analysis of all possible metabolites in a sample

2.1.2 Targeted, Semi-targeted and Untargeted approaches

Studies that are hypothesis driven are intended to acquire holistic metabolic profiling data on a wide variety of metabolites with the intent to identify novel and unobserved changes in the metabolome which can provide information of mechanisms, biological functions or applied as a biomarker (Lammerhofer and Weckwerth, 2013). In analytical chemistry, there are three analytical strategies applied in metabolomics, namely: untargeted, semi-targeted and targeted. Untargeted and semi-targeted approaches, or known as metabolic profiling or metabolite profiling, address a specific hypothesis and the data obtained is more semi-quantitative. Semi-targeted analysis in particular, addresses the metabolites to be assayed prior to the experimental stages and further analytical methods are developed to detect these metabolites based on accuracy, precision, sensitivity and selectivity (Lammerhofer and Weckwerth, 2013). Semi-targeted metabolomics have utilised platforms such as Liquid Chromatography Mass- Spectometry (LC-MS) to screen specific metabolites pertaining to a specific pathway or class (Ncube *et al*., 2014; Khoza *et al*., 2014; Ramabulana *et al*., 2016). Targeted approaches address biochemical pathways and mechanisms and also confirm biomarkers in a system (Scalbert *et al*., 2009). The data obtained in a targeted approach is more quantitative and is validated with the aid of internal standards and reference compounds (Scalbert *et al*., 2009). (Huhman and Sumner, 2002; Jander *et al*., 2004). The major differences between untargeted, semi-targeted and targeted studies is sample preparation, analytical signals identified and quantification of metabolites (Scalbert *et al*., 2009; De Vos *et al*., 2008).

Furthermore, this study reports the differences between *Moringa* cultivation species at a metabolic level. In addition the study also shows the utilisation of a semi-targeted approach for the analysis of flavonoid and chlorogenic metabolites in *Moringa* plant extracts (see **Chapter 4, Papers I-III**). Comparison of the approaches based on the analytical signals is summarised in Figure 1 below.

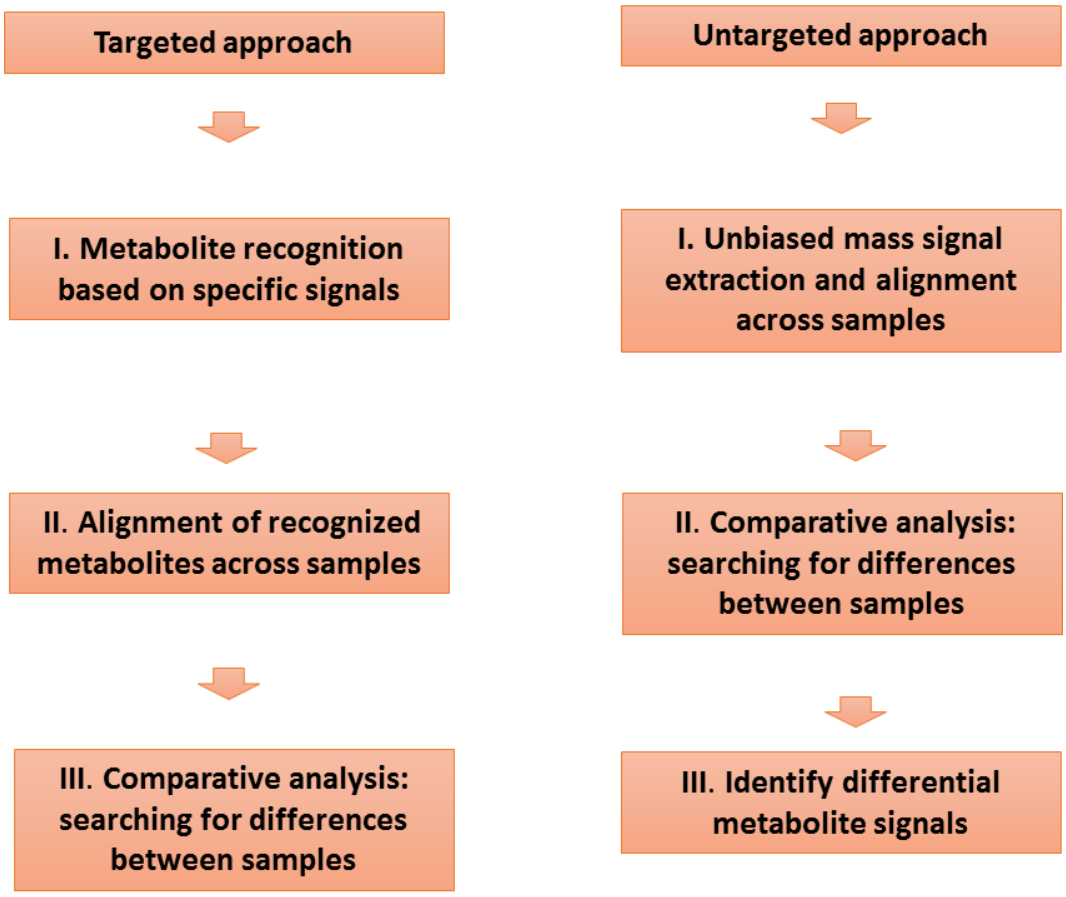


Figure 1: A flow diagram comparing targeted and untargeted approaches utilised in plant metabolomics studies (De Vos *et al*., 2008).

2.1.3 Field applications of metabolomics applications

Metabolomic studies can be used not only to illustrate cellular functions in biological systems caused by bioactive compounds, but it can also be used in quality assessment programs that entail quality of crops and nutrition. Moreover, metabolomics applications have been used in drug development and medical diagnosis. In analytical chemistry, it has been used in the development of methods in order to understand the fate of trace organics. Lastly, in biological sciences, the use of metabolomics has lead to the advancements of genotyping, biomarker discovery and gene function elucidation (Bino *et al*., 2005; Hall, 2006). Furthermore, the focus of this study aims at using metabolomics approaches to explore chemical diversity in metabolomics strategies amongst natural products utilising chemotaxonomy approaches.

2.2 Metabolomics workflow

The plant metabolome is the final downstream product of the genome and consists of highly complex, low molecular weight primary and secondary metabolites (Suberu *et al*., 2016; Roberts *et al*., 2012; Hagel and Facchini, 2008). Due to the multifaceted nature of the metabolome, procedures for metabolomics analysis is necessary and should be considered for the investigation of the whole plant metabolome (Tugizimana *et al*., 2013). Furthermore, with the use of high throughput analytical technologies including various extraction methods and comprehensive profiling of biological compounds, this process is achievable (Hagel and Facchini, 2008).

Metabolomics is considered an investigative tool and consists of three experimental stages: (1) sample preparation (metabolite extractions), (2) data acquisition and lastly (3) data mining with the aid of chemometric methods as shown in Figure 2 below (Tugizimana *et al*., 2012).

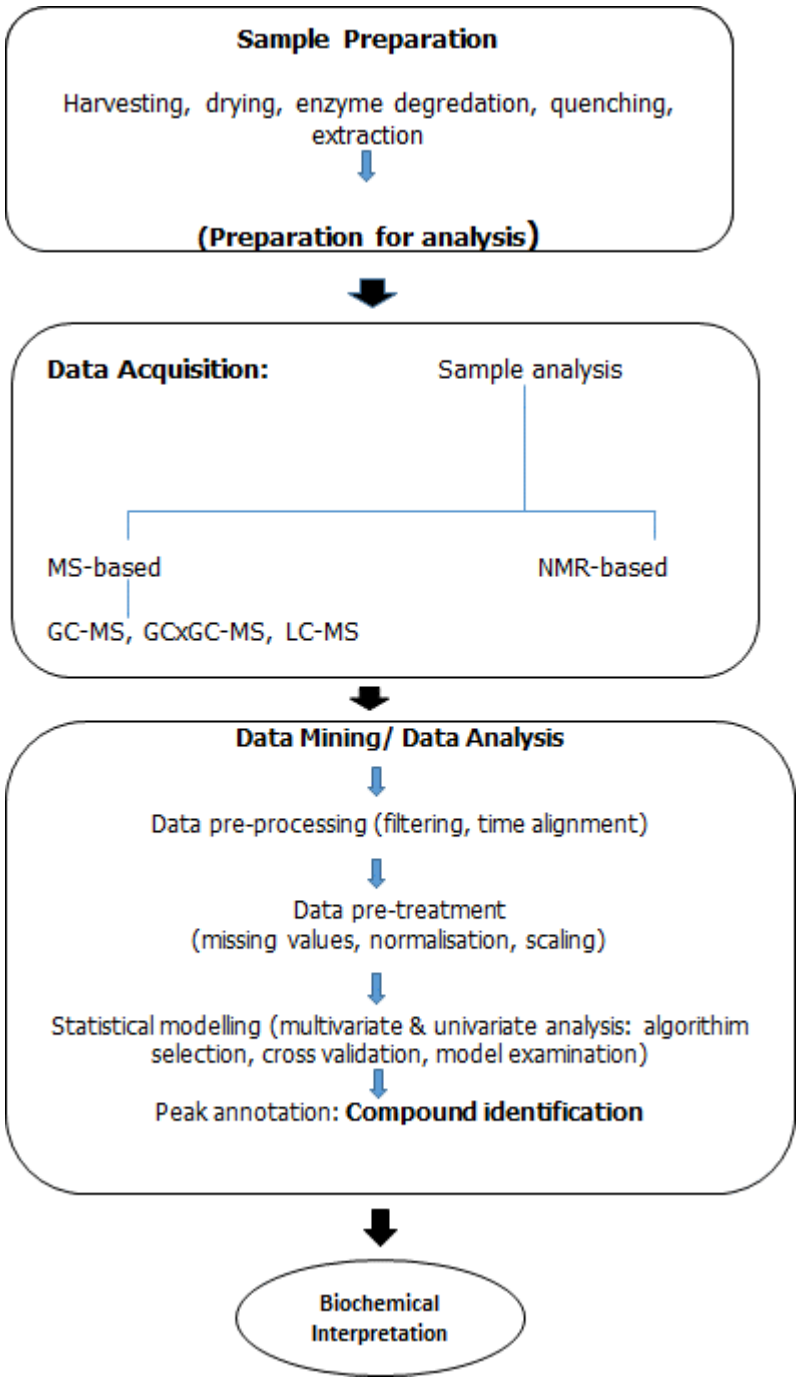


Figure 2: Schematic flowchart of the metabolomics workflow (Tugizimana *et al.*, 2013).

2.2.1 Sample preparation

Sample preparation is a crucial step as practical considerations need to be administered from the start for correct isolation of compounds (Kim and Verpoorte 2010). Sample preparation involves four experimental steps namely: harvesting, drying, extracting plant material and extract preparation for analysis (Kim and Verpoorte, 2010).

Harvesting is considered a critical step and plant samples should be administered with care as metabolic changes can lead to sample variability (Tugizimana *et al.*, 2013). Sample degradation due to enzyme degradation and oxidation causes variability which may lead to inaccurate results (Kim and Verpoorte, 2010). To avoid enzyme degradation of harvested material, techniques such as quenching halts the physiological processes (Kim and Verpoorte, 2010). Quenching methods include treating the samples with acid, utilisation of an enzyme inhibitor and high concentration of organic solvents (Fiehn, 2002; Kim and Verpoorte, 2010).

Following quenching, extraction is the next step. Plant metabolite structures are unique and differ significantly. They form a multitude of compounds that vary based on size, solubility, volatility, polarity, quantity and stability (Dunn and Ellis, 2005; Moco *et al*., 2007; Sumner *et al*., 2003). There are different extraction methods which include liquid-liquid extraction (LLE), supercritical fluid extraction (SFE), solid phase extraction (SPE), pressurised liquid extraction (PLE), microwave assisted extraction (MAE) and ultrasound-assisted extraction (UAE) (Garcia-salas *et al.*, 2010).

The various methods routinely employed to extract metabolites are vast, however, the most common method used for metabolite extraction is liquid-solid extraction (LSE). Chemical properties of organic and aqueous solvents should be considered as different metabolites require specific extraction methods (Kim and Verpoorte, 2010). For instance, a combination of methanol/water is best suited for the extraction of semi-polar compounds such as phenolic acids, flavonoids, alkaloids and glycosylated sterols, whilst chloroform is best suited for the extraction of polar

carotenoids (Moco *et al*., 2007). Therefore, it was specifically applied in this study as described in subsequent chapters (see **Chapter 4 and the Appendix section**).

2.2.2 Data acquisition

A number of analytical platforms are used in metabolomics studies providing advantages such as sensitivity, selectivity and reproducibility (Tugizimana *et al*., 2013). Such analytical platforms include: Gas chromatography-mass spectrometry (GC-MS) (Olivier and Loots, 2012); capillary electrophoresis-mass spectrometry (CE-MS) (García-pérez *et al*., 2008; Ramautar *et al.*, 2006), liquid chromatography-mass spectrometry (LC-MS) (Lu *et al*., 2008), nuclear magnetic resonance (NMR) spectroscopy and fourier-transform infrared (FT-IR). Of these, analytical methods GC-MS, LC-MS and NMR are the most popular used in plant metabolomics as they are capable of determining a variety of metabolites (Allwood and Goodacre, 2010; Tugizimana *et al*., 2013; Fiehn, 2002; Sardans *et al*., 2011). For the scope of this thesis, LC-MS was used as an effective analytical platform due to its sensitivity, catering to semi-polar to polar compounds and its frequent use in the identification of secondary plant metabolites (Tugizimana *et al.*, 2013). In fact, LC-MS has been proven to be the preferred analytical platform for metabolic profiling and metabolomics studies (Metz *et al*., 2007). The full usability of this instrument is discussed below:

2.2.2.1 Ultra-performance liquid chromatography coupled to MS (UHPLC-MS)

The separation of metabolites by LC is with the aid of conventional chromatography on a column. More specifically, separation of complex metabolites can be achieved with the use reverse phase C_{18} column (Allwood and Goodacre, 2010; Dunn and Ellis, 2005). Reports have shown numerous applications of this column for the analysis of multiple secondary metabolites from several biological sources (Seger *et al*., 2013; Ncube *et al*., 2014; Madala *et al*., 2014; Makita *et al*., 2016).

LC interfaced with MS (LC-MS) system combines the separating power of LC in plant crude extracts with MS which aids in confirming the identity of compounds of interest (Patel *et al*., 2010). Mass spectrometry has proven to be a preferred method for molecular analysis as it provides accurate information on the molecular weight and fragmentation patterns of analyte molecules of interest. It further provides the detection of accurate masses and assigning chemical formulas of the detected ions (Allwood and Goodacre, 2010; Spaggiari *et al.* 2013). Several multi-detection methods for the analysis of flavonoids, chlorogenic acids and glucosinolates have been employed over time. Recently, large prospects with the use of LC-MS have been reported by scientists such as Khoza *et al*., (2014), who reported the use of UPLC-qTOF-MS to analyze pharmacologically important metabolites such as chlorogenic acids and flavonoids in *M. oleifera* (Khoza *et al*. 2014).

This hyphenated analytical platform packed with smaller porous particles illustrates high-quality separations and detection capabilities used to characterize compounds in complex samples. The advantages of this technique include:

- small injection volume
- reduced time of analysis
- fast screening of complex samples where analytes of interests are unknown
- high separation power (Bouhifd *et al*., 2013).

Electrospray ionization (ESI) has gained momentum due to its versatility and its simple, vigorous interface that is capable of analyzing various polar molecules at a range of 100-200,000 Dalton (Bristol Biochemistry Research Center, 2013; Anon, 2007). The ionization system is considered as a powerful strategy for fingerprinting applications for complex compounds found in plant extracts (Sawaya *et al*., 2004). It is also a method that involves transferring ions into a gas phase. Prior to the development of ESI-MS, there were several ionization methods used by researchers such as electron ionization and chemical ionization, however, these techniques failed to overcome the susceptibility of analyte fragmentation (Grayson, 2011).

The ESI-mass spectometer consists of three main components such as ion source, mass analyzer and detector (Figure 3). Samples are introduced to the mass

spectrometry via LC. Once in source, samples are subject to ionization and intact molecular and fragmentation ions are produced (Banerjee and Muzumdar, 2012). After they acquire some energy, they are then transferred to the mass analyzer. The mass analyzer then sorts them according to their mass to charge ratio (*m/z* value). Various mass analyzers may be utilized, however, in this study we used quadruple time of flight mass spectrometry (qTOF-MS) (Banerjee and Muzumdar, 2012) (see **Chapter 4, Papers I-III)**.

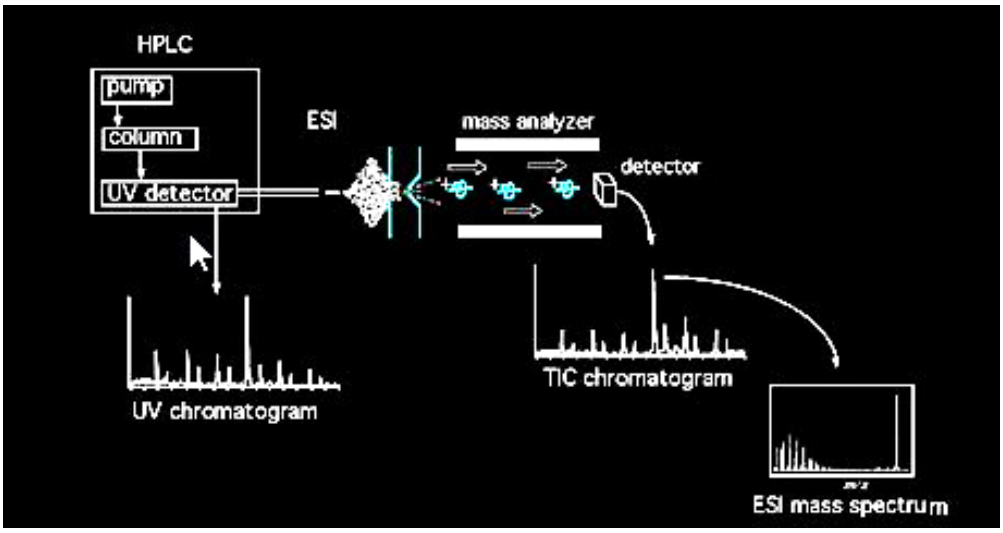


Figure 3: The component of the ESI-mass spectrometer (http://www.cif.iastate.edu/mass-spec/ms-tutorial).

ESI can be applied in both negative and positive polarities. There are three vital processes that involve transferring the sample from the HPLC to gaseous ions that will be taken to the mass spectrometry. These include:

- The production of charged ions at the tip of the Taylor cone
- Decreasing charged droplets
- Constructing highly charged gaseous ions (Baily, 1988; Cloupeau and Prunet-Foch, 1990).

In ESI, samples are dissolved in a polar solvent where it is passed to an electrospray needle which contains high voltage between the needle and nozzle. The high voltage causes the sample fluid to form into what is known as the Taylor cone, which contains positive and negative ions at the tip (Anon, 2007; Banerjee and

Muzumdar, 2012). The electric field causes the Taylor cone to produce a spray of charged droplets. During evaporation, the droplets shrink and are carried by dry gas such as N_2 to the front of the ionization source (Anon, 2007; Bristol Biochemistry Research Center, 2013). In the mass analyzer, small droplets are created and pass through a small aperture where they are transferred into a high vacuum. Once the sample arrives at the spray chamber it is in the form of a mist or droplets. At maximum temperatures, drying gas such as N_2 evaporates any remaining solvent from the droplets. Charged ions are moved to the capillary due to voltage gradients between the spray needle and the entrance of the capillary as well as the difference in atmospheric pressure (Anon, 2007; Bristol Biogeochemistry Research Center, 2013). During solvent evaporation, "Coulombic explosion" occurs, which refers to the droplet being ripped into pieces due to the surface tension notwithstanding the charge. At this point, smaller charged droplets are formed (Anon, 2007; Bristol Biogeochemistry Research Center, 2013). The formation of charged analytes molecules enables high molecular weight components to be analyzed (Anon, 2007; Bristol Biogeochemistry Research Center, 2013). A summary of the technique is shown in Figure 4 below.

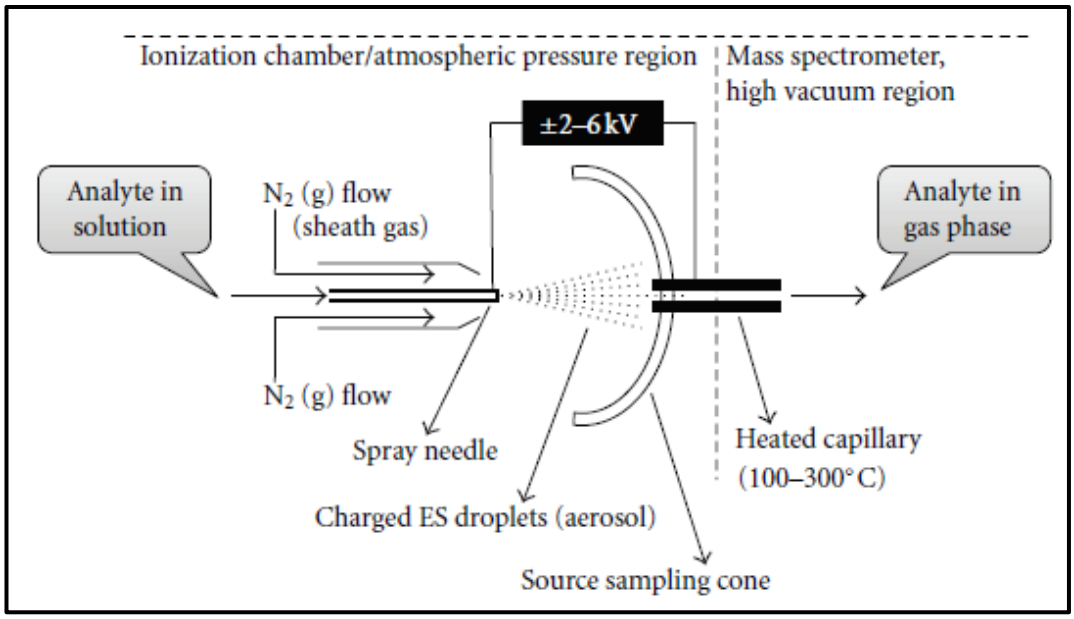


Figure 4: A schematic diagram of the ion-source (Banerjee and Muzumdar, 2012).

Furthermore, ESI has an added advantage of enabling fragmentation when required by the use of instrumentation such as ISCID (In Source Collision Induced Dissociation). The advantages of this instrumentation at several energies provides

controlled and reproducible fragmentation that can further be searched against complex libraries (Fanali *et al*., 2013). An UHPLC-qTOF-MS-ISCID based semi-targeted metabolomics approach was effectively employed to investigate pharmacologically relevant metabolites in *M. ovalifolia* plants (Chapter 4, **Paper II).**

2.2.3 Data analysis/mining

Superior instrumentation, like LC-MS, have the ability to produce large data (Cook and Rutan, 2014; Dunn, 2008). Procedures such as pre-treatment and pre-processing have been utilised to reduce variations in data (Cook and Rutan, 2014; Tugizimana *et al.*, 2013). Pre-treatment is considered an important step as it involves minimising and extracting features from the raw data (Kalogeropoulou, 2011). Different pre-processing steps are utilised in order to produce clean data in the form of normalised peak areas that reflect the metabolite concentration (Van den Berg *et al*., 2006). The steps that are involved in pre-treatment include: noise filtering, peak detection, de-isotoping, peak alignment, peak identification and data normalization (Berg *et al*., 2013; Castillo *et al*., 2011). Subsequently, data undergoes pre-treatment methods such as centering, scaling and transformation (Berg *et al*., 2006). Various software packages are available specifically for metabolomic data processing such as Markerlynx^{TM} (Waters Corporation, Milford, MA, USA), MetAlign (Plant Research International, The Netherlands), XCMS online (www.xcmsonline.scripps.edu), Metaboanalyst (www.metaboanalyst.ca). The cleaned data matrix is then analyzed statistically. The quality of the results is completely reliant on multivariate data analysis and how the analyst processes the data as covered in the subsequent sections below:

2.2.3.1 PCA (Principal component analysis)

Principal component analysis is one of the common statistical tools used specifically for unsupervised approaches (Happyana *et al.*, 2012). This technique

simplifies complex data and represents multivariate data in a low dimensional space (Happyana *et al*., 2012). Moreover, PCA has the proficiencies for finding relationships and variances in the data, therefore, constructing a model of how a chemical system performs, as seen in Figure 5 (Happyana *et al*., 2012). Mathematically in PCA modelling, principal components (PC) represent variances in a data set that are algebraically expressed in the form of variables, where the first principal component (PC1) contains the largest possible variance (PC1), the second largest variance is presented in the second principal component (PC2) and the same applies to subsequent PC's (Tugizimana *et al*., 2013; Liu *et al*., 2010). During PCA analysis, orthogonal linear transformations occur between correlated variables (Madala *et al*., 2012). The score plot provides information and visualization of the relationship between sample groups such as trends, groupings and outliers (Tugizimana *et al*., 2013). A more detailed description and visual comparison of *M. oleifera* cultivars as seen on the PCA score plot can be found in **Chapter 4, Paper III**.

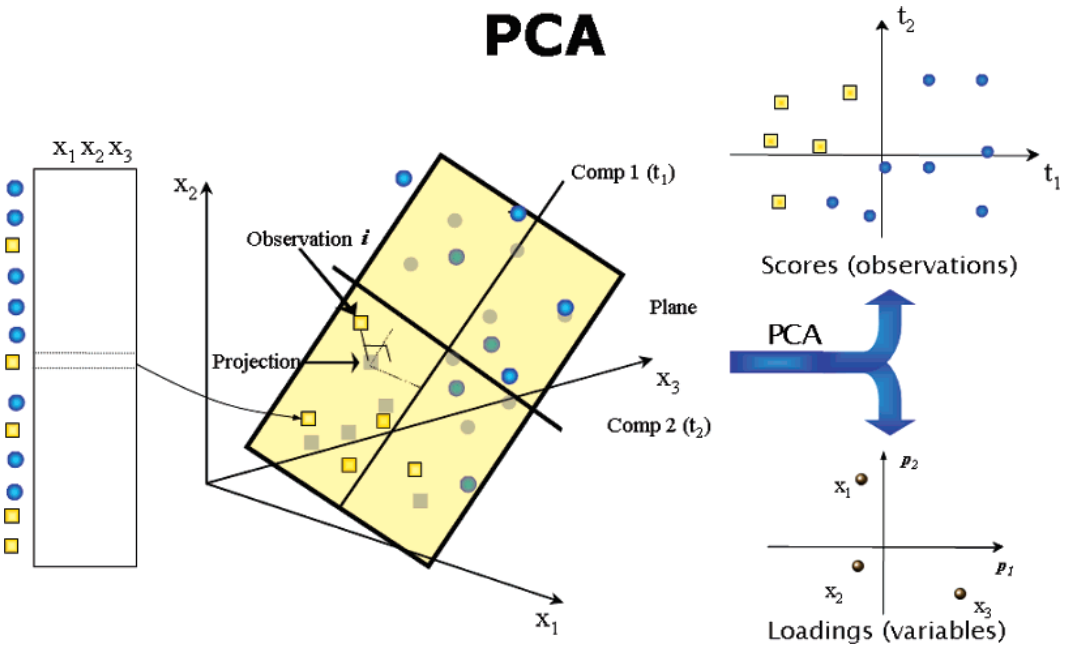


Figure 5: The PCA score plot demonstrates observations in the model plane and the loadings plot demonstrates the influence of the variables present in the model plane (Trygg and Holmes, 2007).

2.2.3.2 Hierarchical clustering analysis (HCA)

HCA is known as a pattern recognition technique that organizes sample data into tree structures, with the main structures further dividing into sub-clusters primarily based on information in the numeric data (Tan *et al*., 2005). It is known as an unsupervised pattern recognition technique because clusters are formed from objects, or more specifically, metabolites without any prior knowledge of what those metabolites are. There are two types of HCA, namely, agglomerative and divisive. Agglomerative is a method in which the object begins as a cluster and pairs of clusters are further formed as the hierarchy moves up until one major cluster remains at the end (Costa, 2014). The second method is a divisive method, which is an opposite approach that involves objects beginning in a single cluster and dividing into smaller clusters (Costa, 2014). HCA is represented in the form of a dendogram. HCA requires a distance metric for linking two clusters in order to determine which clusters should be split or merged (Costa, 2014). The most common distant measure is a Euclidian distance which is the geometric distance between two points.

Previous studies have shown that the combination of PCA and HCA can be utilized as a visual aid to differentiate significant differences between plant cultivars. For instance, in a study by Opara *et al*., (2010), apart from the 82 unique biomarkers identified, HCA and PCA revealed significant differences between the clusters of cultivars (Opara *et al*., 2010). In a similarly study by Tikunov *et al*., (2015), PCA revealed metabolite differences as well as within-type variation which was demonstrated by the clear separation between 94 tomato cultivars (Tikunov *et al.* 2005). HCA visually demonstrated a distinct cluster separation of the cherry tomatoes from the beef and round type tomatoes (Tikunov *et al*., 2005). Likewise, a similar approach was utilized to determine significant differences between 12 *Moringa* cultivars based on their unique metabolite content which is illustrated in **Chapter 4, Paper III**.

2.2.3.3 XCMS online modelling

XCMS is a free open source software and is considered one of the most efficient tools for pre-processing MS-data. It is a refined statistical analysis tool as it offers

many options for data handling and visualization of mass spectrometry based, untargeted metabolomics data (Kalogeropoulou, 2011; Gowda *et al*., 2014). The pre-processing steps processed through the Metlin database include peak detection and matching, Rt alignment, matched filtration and filling in any missing peaks (Gowda *et al*., 2014). The XCMS platform provides the ability to compare two groups and multi-groups comparisons with the aid of visual and statistical interactive tools such as the interactive plot in Figure 6 below. The interactive cloud plot expresses metabolite features (bubbles) based on their significance level (p-value), fold change, *m/z* ratio, Rt and intensity (Gowda *et al*., 2014). Further description of this technique can be found in the **Appendix** section.

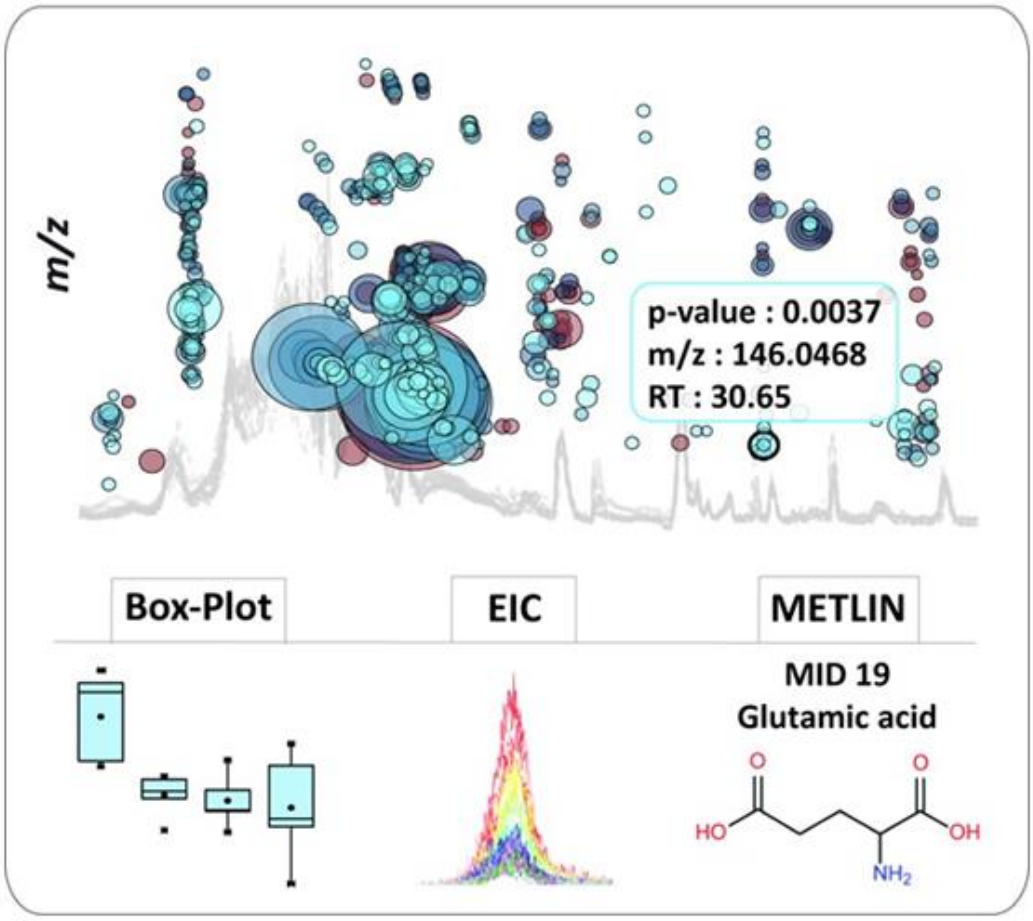


Figure 6: Metabolic data from the interactive cloud plot taken from XCMS online programming (Gowda *et al*., 2014).

2.2.4 Metabolite identification

Metabolite identification is ever progressing, with topics centred on addressing what methods best define a valid metabolite identification (Creek e*t al.*, 2014). Four

levels of identification have been established by the Chemical Analysis Working Group of the Metabolomic Standards of Initiative (MSI; http://www.metabolomics-msi.org/). Identified metabolites are compared to authentic chemical standards (Level 1). At levels 2 and 3, the metabolites identified can be compared to literature as opposed to using authentic standards as confirmation. Lastly, level 4 relates to unknown identified compounds. Metabolite identification in an untargeted approach with the aid of LC/MS involves comparing metabolites in many samples in relative abundancies without prior screening based on their unique mass-to-charge ratio (m/z) and retention time (m/z-Rt pair) (Vinaixa *et al*., 2016). Studies have applied this technology to determine closely related plant taxa, different varieties of individual taxa and plants at various development stages (Cao *et al*., 2011) (As mentioned in **Section 2.2.2.1**). In this current study, metabolomics was utilised to investigate different *M. oleifera* cultivars (as seen in **Chapter 4, Paper III**) and comparison of *M. oleifera* and *M. ovalifolia* plant species (as seen in **Chapter 4, Paper I)**.

2.3 Application of metabolomics and chemotaxonomy

2.3.1 Natural products

Natural products are low molecular weight compounds with a unique structural diversity. They have been studied in many disciplines such as medicine, pharmacology, biology and chemistry and have contributed to many areas of research and development such as nutrition, agriculture, cosmetics, biotechnology, food chemistry and environmental chemistry. The use of metabolomics in plant research has gained momentum due to the benefits that it has on human health, food, pharmaceuticals, plant breeding and nutrition assessment. Furthermore, the contingency and diversity of metabolites present in plants compared to other organisms contribute to the importance of metabolomics in plant research (Akiyama *et al*., 2008).

Fruits and vegetables comprise of a significant number of bioactive molecules that impart health maximizing benefits. These bioactive compounds are a major source of dietary antioxidants that have several properties such as increasing plasma antioxidant capacity, diminishing mortality related diseases such as atherosclerosis,

cardio and cerebrovascular diseases as well as lowering high blood pressure (De Lacerda De Oliveira *et al*., 2014).

Bioactive compounds available in these plants, fruits and vegetables come in a wide variety and are structurally unique and diverse. The chemical compounds include vitamins, phytochemical compounds, isopernoids (terpenoids), protein/amino acids, carbohydrates and its derivatives, lipidic compounds and minerals *(*Ajila *et al*., 2011; Liu, 2004., De La Rosa *et al*., 2009).

Plants such as *M. oleifera,* is said to have medicinal and nutritional benefits (Pakade *et al*., 2012). Several researchers have also reported the leaves of *M. oleifera* to reduce severe medical conditions such as infectious diseases, cardiovascular, gastrointestinal, haematological and hepatorenal disorders (Rodríguez-Pérez *et al*., 2015). The therapeutic potential of *Moringa* leaves stems from phenolic compounds such as flavonoids (Rodríguez-Pérez *et al*., 2015). Bioactive components present in plant foods (fruits and vegetables) are very diverse and include a wide range of chemical compounds with varying structures, such as vitamins (Figure 7). The protective effects of consuming fruits and vegetables is primarily due to secondary phenolic metabolites (Crozier *et al*., 2009).

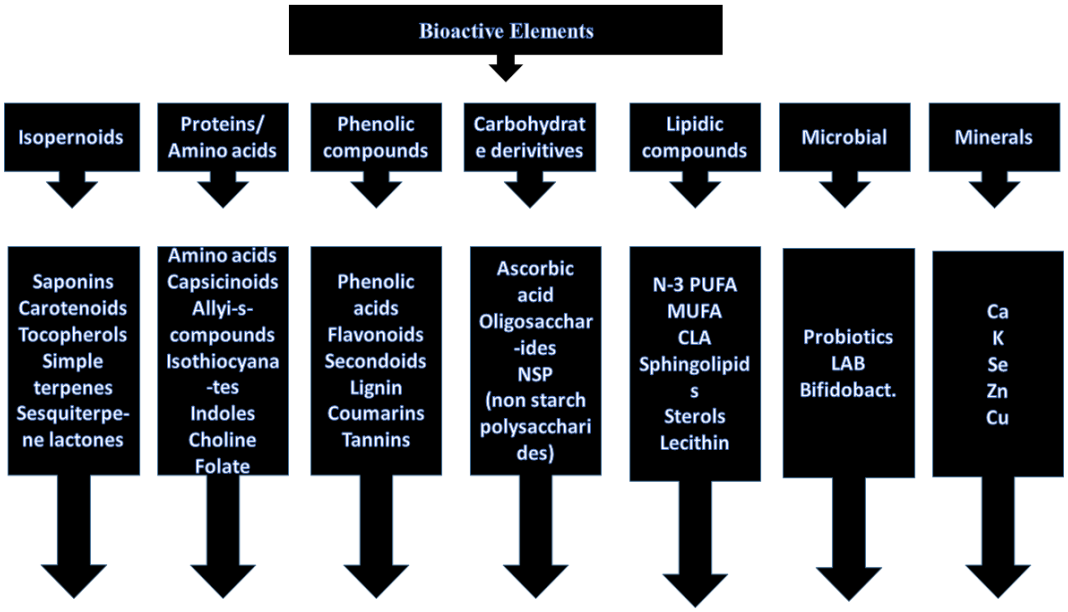


Figure 7: Bioactive phytochemicals in plant foods (Liu, 2013).

2.3.2 Chemotaxonomy

Plant classification is achieved by comparing the morphological, physiological and chemical characteristics in plant species (Lee *et al*., 2015). Chemotaxonomy is an advantageous tool for classification of plant species as this method is primarily based on identifying the differences in chemical compounds of various plant species (Lee *et al*., 2015). This approach also highlights the presence, absence or amount of secondary metabolites (Seaman, 1982).

This tool has been utilized for the classification of plant species based on their phylogenetic genus (Kim *et al*., 2012). Metabolomics (especially LC-MS) in combination with chemometrics has been extensively used and reported for chemotaxonomical classification of medicinal plants (Kim *et al*., 2012). Furthermore, analytical tools such as ^{1}H-NMR, LC-MS and GC-MS have used metabolomics approaches for chemotaxonomic purposes in plants belonging to the same genus and have been reported (Farag *et al*., 2012; Gao *et al*., 2012; Georgiev *et al.,* 2012; Kim *et al*., 2010; Xiang *et al.,* 2011). The addition of metabolomics provides a snapshot of the metabolome as a whole, and therefore provides a more holistic approach (Tugizimana *et al*., 2013; Wink *et al*., 2010).
Metabolomic techniques in combination with chemometric tools such as PCA, HCA and OPLS-DA delivers an easier, faster and consistent method for chemotaxonomy classification (Schulz and Baranska, 2007).

Moreover, at present their neither exists studies for metabolomics techniques for chemotaxonomic approaches for the classification of different cultivars, nor does there exist studies applying LC-MS for plant taxonomic purposes, of which this Thesis addresses in **Chapter 4**.

2.4 Plant secondary metabolites

2.4.1 Flavonoids

Secondary plant metabolites, such as flavonoids, have many biological functions. They are naturally occurring antioxidants and have been found to be responsible for a variety of pharmacological activities (Chen *et al.,* 2013). They are likewise referred to cure degenerative ailments, for example, cardiovascular infections,

malignancies and different diseases (Cook and Samman, 1996; Rice-Evans *et al*., 1995). Moreover, previous reports have shown flavonoids capable of inducing human protective enzyme systems. Flavonoids are classified under six subgroups namely: chalcones, flavones, flavonols, flavandiols, anthocyanins and proanthocyanidins and are synthesized via phenylpropanoid pathway (Ferreyra *et al*., 2012). The basic skeleton structure consists of two aromatic rings (A and B) connected through a pyrone ring (C) as demonstrated in Figure 8. Significant flavonoid compounds such as kaempferol, quercetin and isorhamnetin are present in plant cells as 3-O glycosides such as glucose, glucoroniside, galactoside and rutinoside (Ono *et al.*, 2010). According to Rodriguez-Perez *et al*., (2015), the most predominant subclass of flavonoids found in *M. oleifera* is derived from kaempferol and quercetin derivatives (Rodríguez-Pérez *et al*., 2015). These sugar moeities generally bind to the hydroxyl group on 3', 4' or 7' of the aglycone. Structurally, C-glycosides differ from O- glycosides whereby the glycosyl moeities attach on the C-6' or C-8 position of the aglycone A ring as depicted in Figure 8.

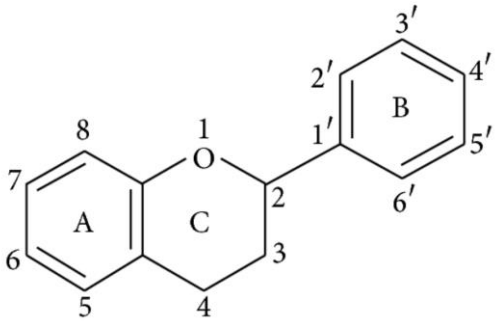


Figure 8: Skeleton of Diphenylpropane (De la Rosa *et al*., 2009).

The growing interest of C-glycosyl flavonoids has grown immensely during the past decade due to the widespread biological applications of these flavonoids. Glycosylation, which involves the sugar conjugation from a sugar donor to an acceptor, often occurs after the completion of aglycone biosynthesis (Kim *et al*., 2013). The sugar attachment at certain positions of the aglycone provides diversification of these metabolites (Ono *et al*., 2010). Glycosylation is mediated by glycosyltransferase (UGTs) by the attachment of a selection of sugars to flavonoid aglycones. These sugars include glucose, galactose, arabinose and more (Kim *et al*., 2013). The presence of these enzymes is of particular importance

because they modulate the bioactive activity of compounds in plants (Méndez and Salas, 2001).

With the increasing interest in flavonoids as bioactive compounds concerning both nutritional and health promotions, researchers have characterized flavonoids with stimulating biological effects (Robak and Gryglewski, 1996). Due to the complexity of the metabolome, certain analytic platforms are needed to cover variety of flavonoids. These techniques are discussed in detail in **Section 2.2.2.** Further information of the flavonoid metabolites that were identified and characterized in the *Moringa* species is found in **Chapter 4, Paper I.**

2.4.2 Chlorogenic acids (CGA)

Chlorogenic acids (CGA) are esters of *trans*-cinnamic esters, that occasionally occur in nature, such as caffeic acid (3, 4-dihydroxycinnamic acid), ferulic acid (3-mehoxy, 4-hydroxtcinnamic acid) and p-coumaric acids (4-hydroxycinamic acid) (Clifford, 2000; Clarke and Macarae, 1985). They are said to contain antioxidant properties that prevent oxidative degeneration in food, cells and organs (Belay and Gholap, 2009). Previously, diets rich in cholorogenic acids have been reported to confer resistance against diseases coupled with oxidative stress such as cancer, cardiovascular, aging and neurodegenerative diseases (Manach, 2004; Fujioka and Shibamoto, 2008). As previously mentioned, CGA's are abundant in nature in the form of *trans* isomers, however, when exposed to UV radiation they're converted into *cis*-isomers. The most common naturally and commercially available CGA is 5-O- caffeoylquinic acid (5-CGA). Chlorogenic acids are products of either the phenylpropanoid pathway or shikimate pathway (Figure 9) that's induced based on environmental stresses such as wounding and infection caused by pathogens, UV radiation and high visible light vessels (Hermann, 1995; Haard and Chism, 1996). From an analytical perspective, CGAs are challenging to identify due to their structural similarities and thus putative characterization of these compounds are achieved with the aid of ion trap-MS based platforms (Ncube *et al*., 2014). Further information of the chlorogenic metabolites identified and characterized in *M. ovalifolia* species is found in **Chapter 4, Paper II.**

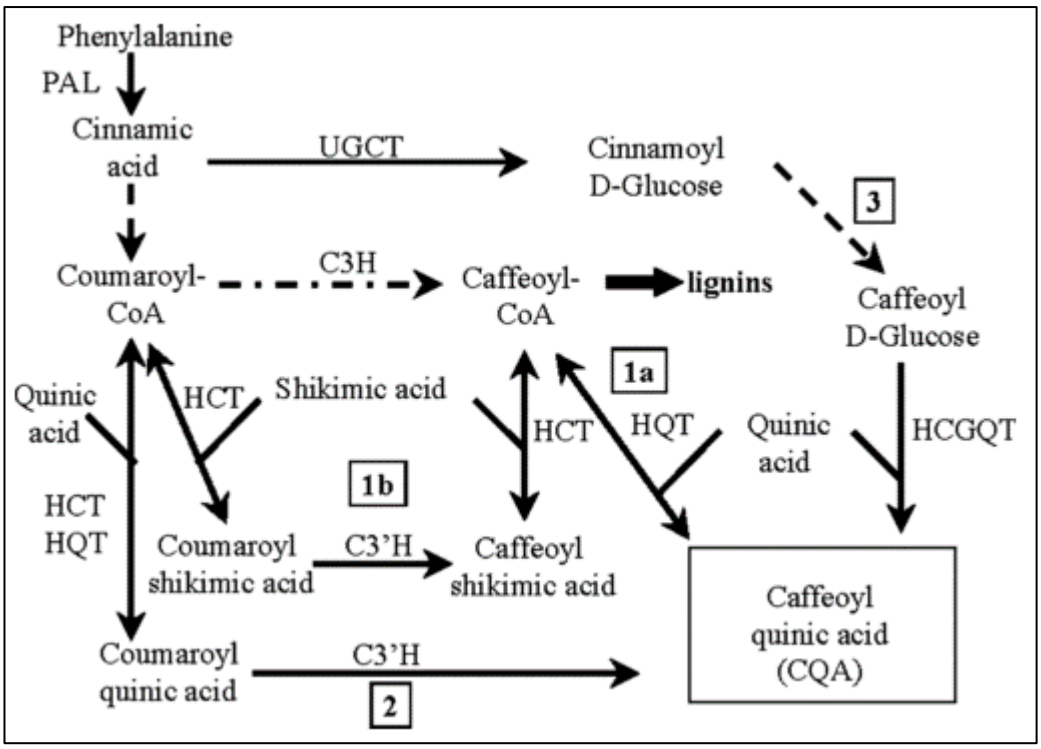


Figure 9: Proposed pathways for chlorogenic biosynthesis in plants (Mahesh *et al*., 2007).

2.5 *Moringa* family

The *Moringaceae* family is the monogeneric flowering plant family, with *Moringa* being the single plant variety. This is characterized by 13 species of dicotyledonous tropical and sub-tropical flowering trees (Padayachee and Baijnath, 2012). The 13 species of *Moringa* include *Moringa arborea, Moringa borzeana Mattei, Moringa concanensis, Moringa drouhardii, Moringa hildebrandtii, Moringa longituba, Moringa oleifera Lam, Moringa ovalifolia, Moringa peregrine, Moringa pygmaea, Moringa ruspoliana and Moringa stenipetala* (Tropicos.org, 2012).

Moringa is known to be called by different names in many regions such as "drumstick," "horseradish tree", "clarifier tree" and "mother's best friend (Natural News, 2011). The *Moringa* tree is known as a multipurpose tree because every part of the tree can be used for food, medicinal and industrial purposes (Moyo *et al.* 2011).

Figure 10: *Moringa oleifera* pictures taken from Hammanskraal, Gauteng, South Africa.

Figure 11: *Moringa ovalifolia,* picture taken from Tsumeb, Namibia.

2.6 Nutritional and medicinal value

The leaves, specifically, can be used as a herb in salads and can be preserved for long periods of time without refrigeration while still maintaining its nutritional

value (Rodríguez-Pérez *et al*. 2015). In South Africa, *Moringa* Development Association of South Africa (MDASA) was formed in 2013 that continuously endorse the consumption of *Moringa* leaves in the form of a nutritional supplement to be used by the local folk in their daily diets (www.mdasa.org). The leaves are reportedly said to contain four times more calcium than milk, three times more potassium than bananas and seven times the vitamin C than oranges (Razis *et al.*, 2014; Teixeira *et al*., 2014).

The plant is also reported to contain vitamins such as Beta carotene (vitamin A), Thiamine (vitamin B1), Riboflavin (vitamin B2), Niacin (vitamin B3), Pyrodixine (vitamin B6), Biotin (vitamin B7) and Tocopherol (vitamin E) (Pakade, 2012). Moreover, *Moringa* also contains useful minerals such as calcium, copper, iron, potassium, magnesium and zinc. According to Farooq *et al.* (2012), the leaves are abundant in iron so much so that they are prescribed for pregnant expecting mothers.

Apart from the vitamins and minerals *Moringa* possesses, it is known to contain over 46 antioxidants, therefore, making it a pre-dominant source of cancer preventing agents (Anwar *et al*., 2007). Antioxidants are known to alleviate the body of free radicals (Fridovich, 1978). According to previous reports, flavonoids such as quercetin and kaempferol prevail in *Moringa* (Rodriguez-Perez *et al*., 2015). Other major antioxidants found in *Moringa* include: β- sitosterol, caffeolyquinic acid and zeatin (Moringa4life, 2011). A summary of its uses are shown in Figure 12 below. Moreover, *Moringa* was utilised as a model plant in this current study (Figure 10 and 11).

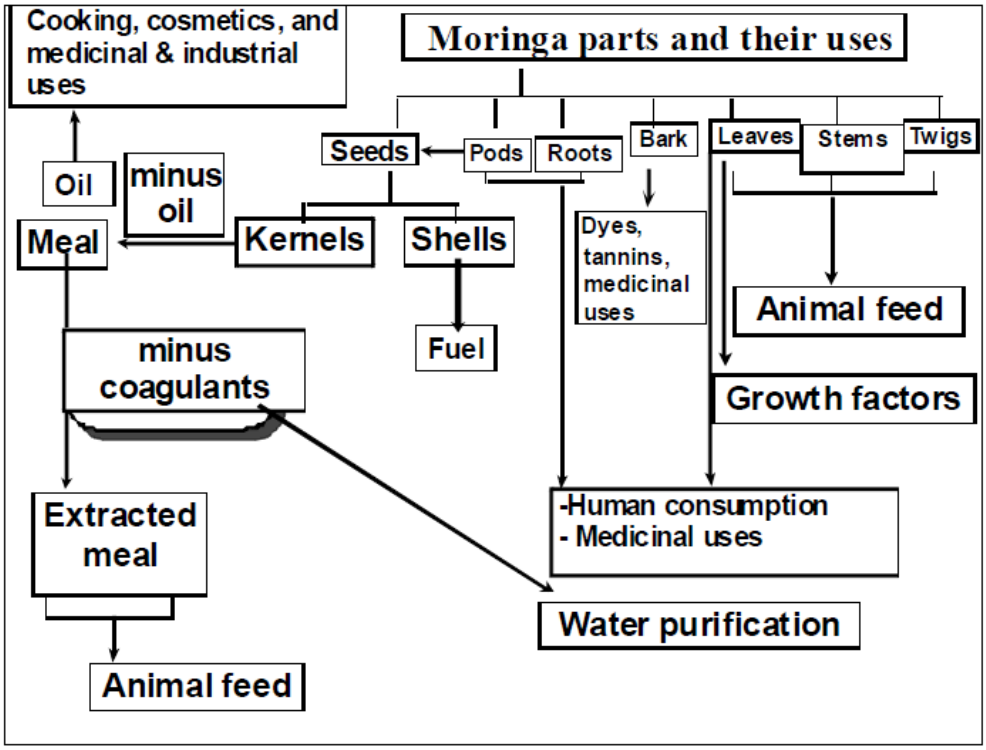


Figure 12: Summary of the different uses of *Moringa* (Foidl *et al*., 2011).

Furthermore, the purpose of this study was to apply a UHPLC-MS method to evaluate and compare the metabolite profiles in *M. oleifera* and *M. ovalifolia* (Figure 10 and 11) as well as to determine the differences between twelve *M. oleifera* cultivars at the metabolome level (Figure 13).

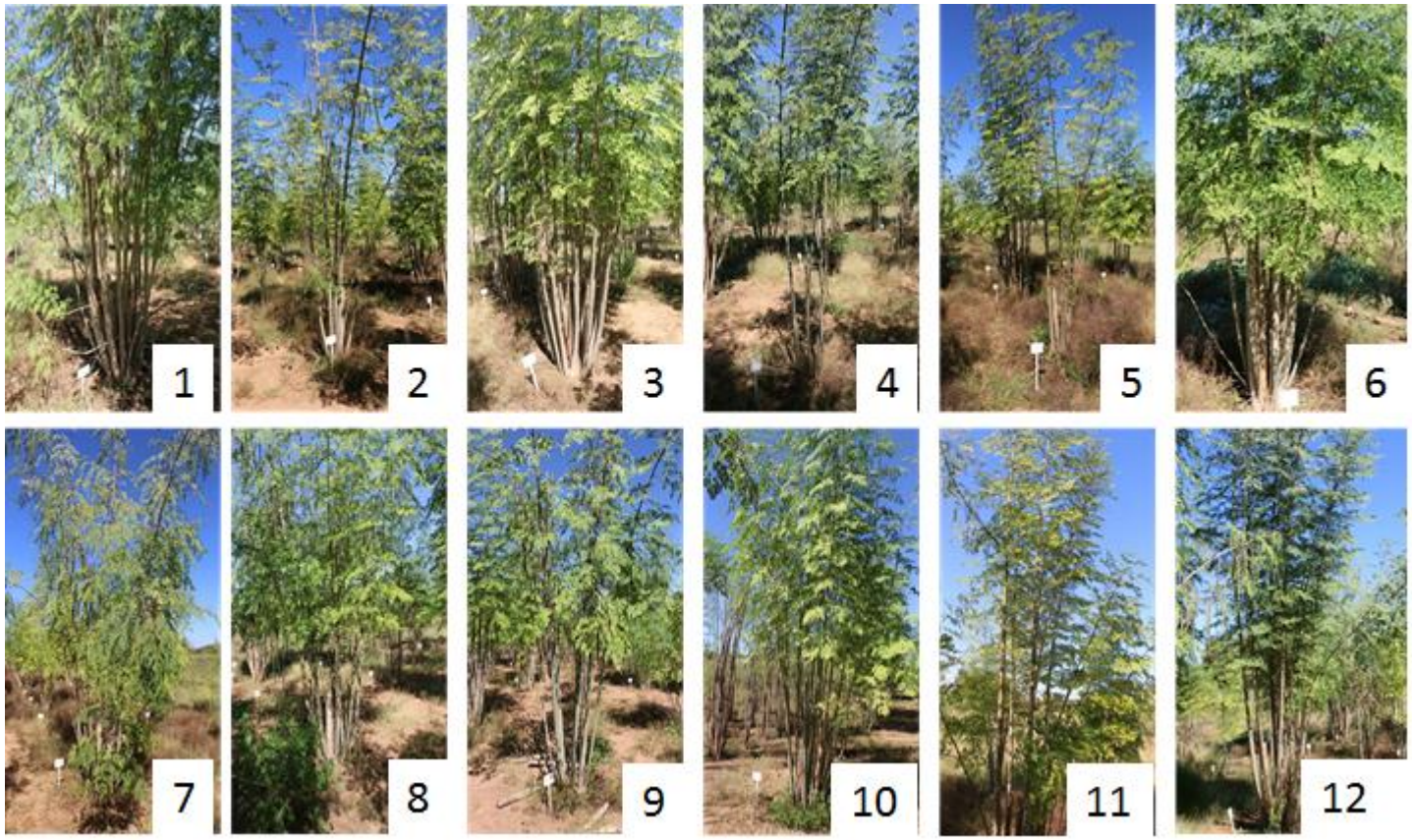


Figure 13: Twelve *Moringa oleifera* cultivars from different locations but harvested at the Agricultural Research Council (ARC) at 8 months and analyzed for chemotaxonomy purposes. 1- CHM; 2-LIMPOPO; 3-SH; 4-TOT4100; 5-TOT4880; 6-TOT4893; 7-TOT5330; 8-4977; 9-TOT5028; 10-TOT5077; 11-TOT5169; 12-TOT7266.

2.7 References

Ajila, C. M., Brar, S. K., Verma, M., Tyagi, R. D., Godbout, S. Valéro, J. R. (2011). Extraction and analysis of polyphenols: Recent trends. Crit. Rev. Biotechnol. 31: pp 227-249.

Akiyama, K., Chikayama, E., Yuasa, H., Shimada, Y., Tohge, T., Shinozaki, K., Hirai, M.Y., Sakurai, T., Kikuchi, J., Saito, K. (2008). Prime: a Web site that assembles tools for metabolomics and transcriptomics. In. Silico. Biol. 8: pp 339–345.

Allwood, J. W., Goodacre, R. (2010). An introduction to Liquid Chromatography – Mass Spectrometry instrumentation applied in plant metabolomic analyses. Phytochem. Anal. 21, pp 33–47.

Allwood, J. W., Vos, R. C. H. De, Moing, A., Deborde, C., Erban, A., Kopka, J., Goodacre, Royston & Hall, R. D. (2011). Plant metabolomics and its potential for systems biology research: Background concepts, technology, and methodology. In Daniel Jameson, Malkhey Verma, and Hans V. Westerhoff, editors: Methods in Enzymology, Vol. 500, Burlington: Academic Press, 2011, pp. 299-336.

Anon. (2007). LC/MS applications guide. Dionex Corporation.

Anwar, F., Latif, S., Ashraf, M., Gilani, AH. (2007). *Moringa oleifera* Lam: a food plant with multiple medicinal uses. Phytother. Res. 21: pp 17–25.

Baily, A.G. (1988). Electrostatic spraying of liquids. John Wiley & Sons, New York, NY, USA.

Banerjee, S., Mazumdar, S. (2012). Electrospray ionization mass spectrometry: A technique to access the information beyond the molecular weight of the analyte. Int. J. Anal. Chem. 2012: pp 1-40.

Belay, A., Gholap, A.V. (2009). Characterization and determination of chlorogenic acids (CGA) in coffee beans by UV-Vis spectroscopy. Afr. J. Pure Appl. Chem. 3: pp 234-240.

Berg, M., Vanaerschot, M., Jankevics, A., Cuypers, B., Breitling, R., Dujardin, J., (2013). LC-MS metabolomics from study design to data analysis using a versatile pathogen as a test case. Comput. Struct. Biotechnol. 4: pp 1-8.

Bino, R.J, De Vos, R.C.H., Lieberman, M., Hall, R.D, Bovy, A., Jonker, H.H., Tikunov, Y., Lommen, A., Moco, S., Levin, I. (2005). The light-hyper responsive high pigment-2dgnmutation of tomato: alterations in the fruit metabolome. New Phytol. 166: pp 427-438.

Bouhifd, M., Hartung, T., Hogberg, H.T., Kleensang, A., Zhao, L. (2013). Review: toxicometabolomics. J. Appl. Toxicol. 33: pp 1365–83.

Bristol Biogeochemistry Research Centre, University of Bristol (2005) High performance liquid chromatography mass spectrometry (HPLC/MS). <http://www.bris.ac.uk/nerclsmsf/techniques/hplcms.html. Accessed 09 Jan. 2016. 168>.

Cao, G., Zhang, Y., Feng, J. A., Cai, H., Zhang, C. R., Ding, M. J. (2011). A rapid and sensitive assay for determining the main components in processed fructus corni by UPLC–Q-TOF-MS. Chromatographia. 73: pp 135–141.

Castillo, S., Gopalacharyulu, P., Yetukuri, L., Orešič, M. (2011). Algorithms and tools for the preprocessing of LC–MS metabolomics data. Chemom. Intell. Lab. Syst. 108: pp 23–32.

Chen, S., Xiang, Y., Deng, J., Liu, Y., Li., S. (2013). Simultaneous analysis of anthocyanin and non-anthocyanin flavonoid in various tissues of different Lotus. PLOS ONE 8.

Chumark, P., Khunawat, P., Sanvarinda, Y., Phornchirasilp, S., Morales, N.P., Phivthong-ngam, L., Ratanachamnong, P., Srisawat, S., Pongrapeeporn, K.U.S. (2008). The in vitro and ex vivo antioxidant properties, hypolipidaemic and antiatherosclerotic activities of water extract of *Moringa oleifera* Lam. leaves. J. Electrostat. 116: pp 439–446.

Clarke, R.J., Macarae, R. (1985). Coffee: Chemistry (Vol. 1). Essex, England. Elsevier Science, pp. 1-7.

Clifford, M.N. (2000). Chlorogenic acids and other cinnamates – nature, occurrence, dietary burden, absorption and metabolism. J. Sci. Food Agric. 80: pp 1033-1043.

Cloupeau, M., Prunet-Foch, B. (1990). Electrostatic spraying of liquids: main functioning. J. Electrostat. 25: pp 165–184.

Cook, N.C., Samman, S. (1996). Review: flavonoids-chemistry, metabolism, cardioprotective effects and dietary sources. J. Nutr. Biochem. 7: pp 66–76.

Cook, D.W., Rutan, S.C. (2014). Chemometrics for the analysis of chromatographic data in metabolomics investigations. J. Chemom. 28: pp 681-687.

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

Crozier, A., Jaganath, I. B. and Clifford, M. N. (2009). Dietary phenolics: chemistry, bioavailability and effects on health. Nat. Prod. Rep. 26: pp 1001-1043.

Costa, C.B. (2014). Development of an integrated computational platform for metabolomics data analysis and knowledge extraction. Master dissertation. Universidado do Minho. Braga, Portugal.

Della Corte, Mattivi, F., Vrhovsek, U. (2013). Development of a targeted LC-MS/MS method for the determination of different classes of grape lipids. Ln: P. Jeandet Eight symposium in vino analytic Scientia: analytical chemistry for wine, brandy and spirits. Reims 47.

De La Rosa, L. A., Alvarez-Parrilla, E. and Gonzalez-Aguilar, G. A. (2009). Fruit and vegetable phytochemicals: Chemistry, nutritional value and stability. Ln: E. Alveres-Parilla. Wiley, Iowa, USA.

De Lacerda De Oliveira, L., Veras De Carvalho, M., Melo, L. (2014). Health promoting and sensory properties of phenolic compounds in food. Rev. Ceres. Vicosa. 61: pp 764–779.

De Vos, R.C.H. Tikunov, Y., Bovy, A.G., Hall, R.D. (2008). Flavour metabolomics: holistic versus targeted approaches in flavour research. Expression of Multidisciplinary Science. Proceedings of the 12th Weurman Symposium. Blank, I., Wust, M., Yeretzian, C. Eds; Zurcher Hoschschule fur Angewandte and Institute fur Chemie and Biologishce Chemie. pp 573-580.

Duarte, A.D.C., Capelo, S. (2006). Application of chemometrics in separation science. J. Liq. Chromatogr. Rel. Technol. 29: pp 1143–1176.

Dunn, W. B., Ellis, D. I. (2005). Metabolomics: Current analytical platforms and methodologies. Trends Anal. Chem. 24: pp 285–294.

Dunn, W.B. (2008). Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. Physiol. Biol. 5: pp 1–24.

Ewansiha, J. U., Anna, H. L., Ajunwa, O.M., Sally, S. M. (2014). The harvesting time and temperature relationship with antimicrobial activity of *Moringa oleifera* Lam (Drum stick). J. Med. Plant Res. 2: pp 33-37.

Fanali, S., Haddad, P. R., Poole, C., Schoenmakers, P., & Lloyd, D. (2013). Liquid Chromatography: Applications. Elsevier Inc. pp 1-667.

Farag, M.A., Mohsen, M., Heinke, R., Ludger A.Wessjohann, L.A. (2014). Metabolomic fingerprints of 21 date palm fruit varieties from Egypt using UPLC/PDA/ESI–qTOF-MS and GC–MS analyzed by chemometrics. Food Res. Int. 64: pp 218–226.

Farooq, F., Rai, M., Tiwari, A., Khan, A.A., Farooq, S. (2012). Medicinal properties of *Moringa* oleifera: An overview of promising healer. J. Med. Plants Res. 6: pp 4368– 4374.

Ferreyra, M.L.F., Rius, S.P., Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front. Plant Sci. 2012: pp 3–222.

Fiehn, O. (2002). Metabolomics–the link between genotypes and phenotypes. Plant. Mol. Biol. 48: pp 155-171.

Flamini, R. (2013). Recent applications of Mass Spectrometry in the study of grape and wine polyphenols. Int. Sch. Res. Notices. 2013: pp 1–45.

Foidl, N., Makkar, H.P.S., Becker, K. (2001). The potential use of *Moringa oleifera* for agriculture and industrial uses. In: The Miracle tree/the multiple attributes of *Moringa oleifera*. Fuglie, L.J. (Ed.). CTA, USA.

Fridovich, I. (1978). Biol. Oxygen Radicals, Sci., 201: pp 875-880. In: H. Sies, Editor, Oxidative Stress, Academic Press, London (1985).

Fukusaki, E., Kobayashi, A. (2005). J. Biosci. Bioengineer. 100: pp 347-354.

Fujioka, K., Shibamoto, T. (2008). Chlorogenic acid and caffeine contents in various commercial brewed coffees. Food Chem. 106: pp 217-221.

Gao,W., Yang, H., Qi, L.W., Liu, E.H., Ren, M. T., Yan,Y. T., Chen, J., Li, P. (2012). Unbiased metabolite profiling by liquid chromatography-quadrupole time-of-flight mass spectrometry and multivariate data analysis for herbal authentication: Classification of seven *Lonicera* species flower buds. J. Chromatogr. A. 1245: pp. 109-116, 2012.

Garcia-Perez, I.; Vallejo, M.; Garcia, A.; Legido-Quigley, C.; Barbas, C. (2008). Metabolic fingerprinting with capillary electrophoresis. J. Chromatogr. A. 1204: pp 130-139.

Garcia-Salas, P., Morales-Soto, A., Segura-Carretero, A., Fernández-Gutiérrez, A. (2010). Phenolic compound extraction systems for fruit and vegetable samples. Molecules 15: pp 8813-8826.

Georgiev, M. I., Ali, K., Alipieva, K., Verpoorte, R., Choi, Y. H. (2011). Metabolic differentiations and classification of Verbascum species by NMR-based metabolomics. Phytochem. 72: pp 2045-2051.

Gomez-Casati, D.F., Zanor, M.I., Busi, M.V. (2013). Metabolomics in Plants and Humans: Applications in the prevention and diagnosis of diseases. Biomed. Res. Int. 2013.

Gowda, H., Ivanisevic, J., Johnson, C.H., Kurczy, M.E., H. Paul Benton, H.P., Rinehart, D., Nguyen, T., Ray, J., Kuehl, J., Arevalo, B., Westenskow, P.D., Wang, J., Arkin, A.P., Deutschbauer, A.M., Patti, G.J., Siuzdak, G. (2014). Interactive XCMS Online: Simplifying advanced metabolomic data processing and subsequent statistical analyses. Anal. Chem. 86: pp 6931−6939.

Grayson, M. (2011). John Bennett Fenn: a curious road to the prize. J. Am. Soc. Mass Spectrom. 22: pp 1301–1308.

Haard, N. Chism, G. (1996) Characteristics of edible tissues. In: Food Chemistry (ed.Fennema, O.). pp. 943–1011. Marcel Dekker, Inc., New York, USA.

Hall, R.D. (2006). Plant metabolomics : from holistic hope, to hype , to hot topic. New Phytol. doi : 10.1111/j.1469-8137.2005.01632.x

Hagel, J.M., Facchini, P.J. (2008). Plant metabolomics: analytical platforms and integration with functional genomics. Phytochemistry. Rev. 7: pp 477-497.

Happyana, N., Muntendam, R., Kayser, O. (2012). Metabolomics as a bioanalytical tool for characterization of medicinal plants and their phytomedical preparations. Pharm. Biotechnol. Drug Discov. Clin. Appl. 2012: pp 527–552.

Hermann, K.M. (1995). This shikimate pathway: early steps in the biosynthesis of aromatic compounds. Plant Physiol. 107: pp 7-12.

Huhman, D.V., Sumner, L.W. (2002). Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. Phytochemistry, 59: pp 347–360.

Idborg, H., Zamani, L., Edlund, P.O., Schuppe-Koistinen, I., Jacobsson, S.P. (2005). Metabolic fingerprinting of rat urine by LC/MS Part 1. Analysis by hydrophilic interaction liquid chromatography–electrospray ionization mass spectrometry J. Chromatogr. B. 828: pp 9-13.

Ivanisevic, J., Stauch, K.L., Petrascheck, M., Benton, H.P., Epstein, A.A., Gorantla, S., Tran, M., Hoang, L., Kurczy, M.E., Michael, D., Gendelman, H.E., Fox, H.S., Siuzdak, G. (2016). Metabolic drift in the aging brain. Aging 8: pp 1–21.

Jander, G., Norris, S.R., Joshi, V., Fraga, M., Rugg, A., Yu, S., Li, L., Last, R.L. (2004). Application of a high-throughput HPLC-MS/MS assay to *Arabidopsis* mutant screening; evidence that threonine aldolase plays a role in seed nutritional quality. Plant J. 39: pp 465–475.

Jaitz, L., Siegl, K., Eder, R., Rak, G., Abranko, L., Koellensperger, G. (2010). LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. Food Chem. 122: pp 366–372.

Kalogeropoulou, A. (2011). Pre-processing and analysis of high-dimensional plant metabolomics data. MSc Thesis. East Anglia, Norwich, England.

Kell, D.B., Oliver, S.G. (2004). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays. 26: pp 99–105.

Kim, H.K., Verpoorte, R. (2010). Sample preparation for plant metabolomics. Phytochem. Anal. 21: pp 4–13.

Kim, Y.M., Lee, J., Park, S.H., Lee, C., Lee, J.W., Lee, D.,Kim, N.,Lee, D.,Kim, H.Y., Lee, C.H. (2012). LC-MS-based chemotaxonomic classification of wild-type Lespedeza sp. and its correlation with genotype. Plant Cell Rep. 31: pp 2085–2097.

Kim, H.S., Kim, B., Sung, S., Kim, M., Mok, H., Chong, Y., Ahn, J.H. (2013). Engineering flavonoid glycosyltransferases for enhanced catalytic efficiency and extended sugar-donor selectivity. Planta. 238: pp 638.

Khoza, B.S., Chimuka, L., Mukwevho, E., Steenkamp, P.A., Madala N.E. (2014). The effect of temperature on Pressurised Hot Water Extraction of pharmacologically important metabolites as analysed by UPLC-qTOF-MS and PCA. J. Evid. Based Complementary Altern. Med. 2014: pp 9.

Kumar, S., Pandey, A. (2013). Chemistry and biological activities of flavonoids: An Overview. Hindawi. 2013.

Lämmerhofer, M., Weckwerth, W. (2013). Metabolomics in Practice: Successful strategies to generate and analyze metabolic data. John Wiley & Sons. Hoboken, NJ, USA.

Lee, S., Oh, D., Lee, S., Ga Kim, R., Lee, J.S., Son, Y.K., Bae, C., Yeo, J., Choong Hwan Lee, C.H. (2015). Chemotaxonomic metabolite profiling of 62 indigenous plant species and its correlation with bioactivities. Molecules. 20: pp 19719-19734.

Liu, R. H. (2004). Potential synergy of phytochemicals in cancer prevention: mechanism of action. J. Nutr. 134: pp 3479S-3485S.

Liu, C.Y., Cao, M., Frederich, M., Choi, Y.H., Verpoorte, R., Van der Kooy, F. (2010). Metabolomic investigation of the ethnopharmacological use of *Artemisia afra* with NMR spectroscopy and multivariate data analysis. J. Ethnopharmac. 128: pp 230-235.

Liu, R.H. (2013). Health-promoting components of fruits and vegetables in the diet. Adv. Nutr. 4: pp 384S-392S.

Lu, X., Zhao, X., Bai, C., Zhao, C., Lu, G., Xu, G. (2008). LC–MS-based metabonomics analysis. J. Chrom. B. 866: pp 64–76.

Madala, N.E., Steenkamp, P.A., Piater, L.A., Dubery, I.A. (2012). Collision energy alteration during mass spectrometric acquisition is essential to ensure unbiased metabolomic analysis. Anal. Bioanal. Chem. 404: pp 367–72.

Madala, N.E., Tugizimana, F., Steenkamp,P.A. (2014). Development and optimization of an UPLC-QTOF-MS/MS method based on an in-source collision induced dissociation approach for comprehensive discrimination of chlorogenic acids isomers from *Momordica* plant species. J. Anal. Methods Chem. 7.

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. S. Afr. J. Bot. 105: pp 116-122.

Manach, C., Scalbert, A., Morand, C., Rémésy, C., Jime, L. Jime´nez, L. Polyphenols: food sources and bioavailability (2004). Am. J. Clin. Nutr. 79: pp 727–47.

Méndez, C., Salas, J.A. (2001). Altering the glycosylation pattern of bioactive compounds. Trends Biotechnol. 19.

Metaboanalyst 3.0- a comprehensive tool suite for metabolomic data analysis. <http://www.metaboanalyst.ca/. Accessed 18 February 2015.>

Metz, T.O., Zhang, O., Page, J.S., Shen, Y., Callister, S.J., Jacobs, J.M., Smith, R.D. (2007). The future of liquid chromatography-mass spectrometry (LC-MS) in metabolic profiling and metabolomic studies for biomarker discovery. Biomark Med. 1: pp 159–185.

Moco, S., Bino, R.J., Vorst, O., Verhoeven, H.A., Groot, J. De, Beek, T.A. Van, Vervoort, J., Vos, C.H.R. (2006). A liquid chromatography mass spectrometry-based metabolome database for tomato. Breakthrough Technologies. 141: pp 1205–1218.

Moco, S., Vervoort, J., Bino, R.J., De Vos, R.C.H., Bino, R. (2007). Metabolomics technologies and metabolite identification. Trends Anal. Chem. 26: pp 855–866.

Moringa4Life < www.Moringa4life.co.za, accessed 12 August 2011.>

Moyo, B., Masika, P.J., Hugo, A., Muchenje, V. (2011). Nutritional characterization of *Moringa* (*Moringa oleifera* Lam.) leaves. African J. Biotech. 10: pp 12925-12933.

Natural News <http://www.naturalnews.com/022272.html#ixzz1MKFQIKwe, accessed 2 January 2015>.

Ncube, E. N., Mhlongo, M. I., Piater, L.A., Steenkamp, P.A., Dubery, I.A., Madala, E.N. (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. Chem. Cent. J. 8: pp 66.

Olivier, I., Loots, D.T. (2012). A metabolomics approach to characterise and identify various *Mycobacterium* species. J. Microbiol. Methods. 88: pp 419–426.

Ono, E., Homma, Y., Horikawa, M., Kunikane-Doi, S., Imai, H., Takahashi, S., Kawai, Y., Ishiguro, M., Fukui, Y., Nakayama, T. (2010). Functional differentiation of the glycosyltransferases that contribute to the chemical diversity of bioactive flavonol glycosides in grapevines (*Vitis vinifera*). Plant Cell. 22: pp 2856–2871.

Opara, U.L., Jacobson, D., Al-saady, N.A. (2010). Analysis of genetic diversity in banana cultivars (*Musa cvs* .) from the South of Oman using AFLP markers and classification by phylogenetic, hierarchical clustering and principal component analyses . Biomed and Biotechnol. 11: pp 332–341.

Padayachee, B., Baujnath. (2012). An overview of the medicinal importance of *Moringaceae*. J. Med. Plant Res. 6: pp 5831-5839.

Pakade, V., Cukrowska, E., Chimuka, L. (2012). Metal and flavonol contents of *Moringa oleifera* grown in South Africa. S. Afr. J. Sci. 109: pp 1–7.

Patel, K.N., Patel, J.K., Patel, M.P., Rajput, G.C., Patel, H.A. (2010). Introduction hyphenated techniques and their applications in pharmacy. Pharm. Methods 1: pp 2–13.

Ramabulana, T., Mavunda, R.D., Steenkamp, P.A., Piater, L.A., Dubery, I.A., Madala, N.E. (2015). Plant physiology and biochemistry secondary metabolite perturbations in *Phaseolus vulgaris* leaves due to gamma radiation. Plant Physiol. Biochem. 97: pp 287–295.

Ramabulana, T., Mavunda, R.D., Steenkamp, P.A., Piater, L.A., Dubery I.A., Madala, N.E. (2016). Perturbation of pharmacologically relevant polyphenolic compounds in *Moringa oleifera* against photo-oxidative damages imposed by gamma radiation. J. Photochem. Photobiol. B. 156: pp 79-86.

Ramakrishna, A., Ravishankar, G.A. (2011). Influence of abiotic stress signals on secondary metabolites in plants. Plant Signal Behav. 6: pp 1720-1731.

Ramautar, R., Demirci, A., de Jong, G.J. (2006). Capillary electrophoresis in Metabolomics. Trends Anal. Chem. 25: pp 455–466.

Razis, A.F.A., Ibrahim, M.D., Kyntayya, S.B. (2014). Health benefits of *Moringa oleifera*. Asian Pacific J. Cancer Prevention. 15: pp 8571-8576.

Rice-Evans, C.A., Miller, N.J., Bolwell, P.G., Broamley, P.M., and J. B. Pridham, J.B. (1995). The relative antioxidant activities of plant-derived polyphenolic flavonoids. Free Radic. Res. 22: pp 375–383.

Robak, J., Gryglewski, R.J. (1996). Bioactivity of flavonoids. Pol. J. Pharmacol. 6: pp 555–564.

Roberts, L.D., Souza, A.L., Gerszten, R.E., Clish, C.B. (2012). Targeted Metabolomics: Curr. Protoc. Mol. Biol. 98: pp 1-30.

Rochfort, S. (2005.) Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J. Nat. Prod. 68: pp 1813-1820.

Rochfort, S.J., Trenerry, C.V. (2010). Natural Products Reseach and Metabolomics. DPI-Werribee Centre, Werribee, VIC, Australia [a] 2010 Elsevier Ltd.

Rodriguez-Perez, C., Quirantes-Pine, R., Fernandez-Gutierrez, A., Segura-Carretero, A. (2015). Optimization of extraction method to obtain a phenolic compounds-rich extract *Moringa oleifera* Lam leaves. Ind. Crops Prod. 66: pp 246-254.

Roullier-Gall., C., Witting, M., Gougeon, R.D., Schmitt-Kopplin, P. (2014). High precision mass measurements for wine metabolomics. Front. Chem. 2: pp 1.

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

Sasidharan, S., Chen, Y., Saravanan, D.,Sundram, K.M., Latha, L.Y. (2011). Extraction, isolation and characterization of bioactive compounds from plants extracts. Afr. J. Tradit. Complement Altern. Med. 8: pp 1–10.

Sawaya, A.C.H.F., Tomazela, D.M., Cunha, I.B.S., Bankova, V.S., Marcucci, M.C.,Angela R. Custodio, A.R., Eberlin, M.N. (2004). Electrospray ionization mass spectrometry fingerprinting of propolis. Analyst. 129: pp 739-744.

Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B.S., van Ommen, B., Pujos-Guillot, E., Verheij, E., Wishart, D., Wopereis, S. (2009). Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. Metabolomics. 5: pp 435–458.

Schulz, H., Baranska, M. (2007). Identification and quantification of valuable plant substances by IR and Raman spectroscopy. Vib. Spectrosc. 43: pp 13–25.

Seaman, F. C. (1982). Sesquiterpene lactones as taxonomic characters in the Asteraceae. Bot. Rev. 48: pp 121-594.

Seger, C., Sturm, S., & Stuppner, H. (2013). Mass spectrometry and NMR spectroscopy: modern high-end detectors for high resolution separation techniques--state of the art in natural product HPLC-MS, HPLC- NMR, and CE-MS hyphenations. Nat. Prod. Rep. 30: pp 970–87.

Spaggiari, D., Fekete, S., Eugster, P. J., Veuthey, J., Geiser, L., Rudaz, S., Guillarme, D. (2013). Contribution of various types of liquid chromatography-mass spectrometry instruments to band broadening in fast analysis. J. Chromatogr. A. 1310: pp 45–55.

Suberu, J., Gromski, P.S., Nordon, A., Lapkin, A. (2016). Multivariate data analysis and metabolic profiling of artemisinin and related compounds in high yielding varieties of *Artemisia* annua field-grown in Madagascar. J. Pharm. Biomed. Anal. 117: pp 522–531..

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

Tan, P. Steinback, M., Kumar, V. (2005). An Introduction to Data Mining, Addison- Wesley. pp. 487-495.

Teixera, E.M.B., Carvalho, M.R.B., Neves, V.A., Silva, M.A., Pereira-Arantes, L. (2014). Chemical characteristics and fractionation of proteins from *Moringa oleifera* Lam. leaves. Food Chem. 147: pp 51–54.

Tikunov, Y., Lommen, A., Vos, R.C.H. De, Verhoeven, H.A., Bino, R.J., Hall, R.D., Bovy, A. (2005). A novel approach for nontargeted data analysis for metabolomics. Large-Scale profiling of tomato fruit volatiles. Plant Physiol. 139: pp 1125–1137.

Tolin, S., Pasini, G., Simonato, B., Mainente, F., and Arrigoni, G. (2012). Analysis of commercial wines by LC-MS/MS reveals the presence of residual milk and egg white allergens. Food Control. 28: pp 321–326.

Tropicos.org.(2012).MissouriBotanicalGarden.
<http://www.tropicos.org/Name/4002223, accessed 18 January 2016.>

Trygg, J., Holmes, E., Lundstedt, T. Chemometrics in Metabonomics. (2007). J. Proteome. Res. 6: pp 469–79.

Tweeddale, H., Notley-McRobb, L., Ferenci, T. (2006). Effect of slow growth on metabolism of Escherichia coli, as revealed by global metabolite pool ("metabolome") analysis. J. Bacteriol. 180: pp 5109-5116.

Tugizimana, F., Piater, L., Dubery, I. (2013). Plant metabolomics: A new frontier in phytochemical analysis. S. Afric. J. Sci. 109.

Van den Berg, R. A., Hoefsloot, H.C.J., Westerhuis, J. A., Smilde, A.K., Van der Werf, M.J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 7: pp 142.

Vinaixa, M., Schymanski, E., Neumann, S., Navarro, M., Salek, R., Yanes, O. (2016). Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects. Trends Anal. Chem. 78: pp 23-35.

XCMS- The original and most widely used metabolomics and lipidomic platform <https://xcmsonline.scripps.edu/landing_page.php?pgcontent=mainPage, Accessed March 2015>.

Xiang, Z., Wang, X. Q., Cai, X. J., Zeng, S. (2011). Metabolomics study on quality control and discrimination of three *Curcuma* Species based on gas chromatograph-mass spectrometry. Phytoch. Analysis. 22: pp 411-418.

Weckwerth, W. (2003). Metabolomics in systems biology. Annu. Rev. Plant Biol. 54: pp 669-689.

Wink, M., Botschen, F., Gosmann, C., Schäfer, H., Waterman, P.G. (2010). Chemotaxonomy seen from a phylogenetic perspective and evolution of secondary metabolism. Annu. Rev. Plant Biol. 40: 364-43.

Chapter 3: Objectives of the Study

This section gives the general and specific objectives, justification and novelty of the study.

3.1 Main objectives

To apply a UHPLC-MS based metabolomics approach to evaluate and compare the metabolic profiles in *M. oleifera* and *M. ovalifolia* as well as to determine the differences between *M. oleifera* cultivars at the metabolome level.

3.2 Specific objectives

- Report the use of in-source collision induced dissociation (ISCID) to develop an LC–qTOF-MS method for the profiling of the chlorogenic acid (CGA) in the leaves of *M. ovalifolia*, a species endemic to Namibia.

- Using a semi-targeted metabolite approach for the comparison of the flavonoid metabolite differences in *M. oleifera* and *M. ovalifolia* with the aid of UPLC-qTOF-MS.

- Use multivariate data modelling to evaluate data from LC-qTOF-MS for the metabolic profiling of 12 *M. oleifera* cultivars using LC-qTOF-MS for chemotaxonomic classification.

3.3 Research question

Are both *Moringa* species similar in their metabolite profile and can the metabolomics approach be used to evaluate the medicinal and nutritional value? Furthermore, can metabolomics be used to identify the superior *Moringa* cultivar utilising LC-MS?

3.4 Hypothesis

- $H_{1.}$ A UHPLC-MS based metabolomics approach can be used to compare species of *M. oleifera* and *M. ovalifolia*. The same approach can be used to compare cultivars of *M. oleifera* cultivars sampled from different locations, but cultivated in the same area.

3.5 Novelty

There have been numerous reports on *M. oleifera* with regards to its nutritional and medicinal benefits, however, there is no information on *M. ovalifolia* and its medicinal properties. Furthermore, this will be the first time the metabolite composition of the two species are compared and hence, we are reporting the significance of taxonomical grouping of related species based on their sugar attachment. Moreover, a semi-targeted approach will be undertaken to identify pharmacological relevant metabolites found in *M. ovalifolia*. Furthermore, comparison of the metabolic content present in 12 *Moringa* cultivars with the aid of multivariate statistical tools provides an indication of the biochemical patterns of variation amongst the cultivars, and to advance our understanding of taxonomic and evolutionary relationships with the species with hopes of determining a superior *Moringa* cultivar for nutritional and medicinal purposes.

Chapter 4: Publications

This chapter lists three publications submitted for examination. Each manuscript in this chapter is formatted into the style requirement as per the requirements of the particular journal published. Furthermore, the reference styles of each journal publication differs. Papers I and II have been submitted and published. Paper III is a completed manuscript intended to be submitted to Frontiers in Pharmacology.

Paper I

This paper titled "Comparative analysis of the flavonoid content in *Moringa oleifera* and *Moringa ovalifolia* with the aid of UHPLC-qTOF-MS fingerprinting," was submitted to *South African Journal of Botany* and the presented format is for that journal. It explores the difference in the flavonoid content in the two *Moringa* species based on their sugar attachment to the aglycone.

South African Journal of Botany 105 (2016) 116–122

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

Comparative analyses of flavonoid content in *Moringa oleifera* and *Moringa ovalifolia* with the aid of UHPLC-qTOF-MS fingerprinting

Charlene Makita [a], Luke Chimuka [a], Paul Steenkamp [b,c], Ewa Cukrowska [a], Edwin Madala [b,*]

[a] Molecular Science Institute, School of Chemistry, University of the Witwatersrand, P/Bag 3, Braamfontein, 2050 Johannesburg, South Africa
[b] Department of Biochemistry, University of Johannesburg, P. O. Box 524, Auckland Park, 2006, Johannesburg, South Africa
[c] CSIR Biosciences, Natural Products and Agroprocessing Group, Building 20, Pretoria 0001, South Africa

ARTICLE INFO

Article history:
Received 15 July 2015
Received in revised form 21 December 2015
Accepted 23 December 2015
Available online xxxx

Edited by AR Ndhlala

Keywords:
Flavonoids
Glycosylation *Moringaceae*
Moringa oleifera
Moringa ovalifolia
UHPLC-QTOF-MS

ABSTRACT

Moringa species are multipurpose plants with nutritional, medicinal and industrial benefits. The flavonoids content of *Moringa oleifera* and *Moringa ovalifolia* was studied using an Ultra high-performance liquid chromatography coupled with electrospray ionization quadrupole time-of-flight mass spectrometry (UHPLC-ESI-qTOF-MS). The results revealed that the two species contain at least 17 flavonoids compounds between them. However, *M. oleifera* was found to contain the most flavonoids than *M. ovalifolia* which contained only three of the total flavonoids. Furthermore, all flavonoids in *M. ovalifolia* were shown to be glycosylated with only rutinoside. Based on the current findings, the two species seem to have a different composition of flavonoids, therefore suggesting an underlying variation at the genetic level for flavonoid biosynthesis. The difference in the flavonoids composition of the two species as seen from the results is mainly due to glycosylation capabilities, with *M. oleifera* being more superior in this case. Prior to this study, there has been no comprehensive investigation into the flavonoid content (or any phytochemical studies) of *M. ovalifolia* and its comparison to other *Moringa* species. Furthermore, vicenin-2 a molecule that has recently been linked to various medicinal properties has been identified in *M. oleifera*. Overall, *M. oleifera* (as compared to *M. ovalifolia*) is expected to exhibit wider pharmacological activities owing to its glycosylation complexity.

1. Introduction

The study of the flavonoid composition in natural plants has stimulated considerable interest because of the health benefits (Cook and Samman, 1996). New natural flavonoids have been isolated and characterized with many biological activities (Robak and Gryglewski, 1996). Scientific evidence has confirmed that the regular intake of dietary flavonoids (specifically flavanols, proanthocyanidins and anthocyanins) from natural plants reduces the effects of oxidative damage such as cardiovascular diseases, diabetes and other diseases associated with aging (Nijveldt et al., 2001). Amongst such, *Moringa oleifera*, a plant native to India has been proven to be a multipurpose plant containing medicinal and nutritional benefits (Pakade et al., 2012). Several researchers have also reported treatment of severe medical conditions such as malaria, bronchitis, fever and even symptoms associated with HIV/AIDS (Rodriguez-Perez et al., 2015). Such benefits are correlated to the large contingency of phenolic acids and flavonoid molecules of these plants (Rodriguez-Perez et al., 2015).

Flavonoids are dispersed secondary plant metabolites with various metabolic functions. These compounds differ from one another and exist as either free forms (aglycones), unsaturated or commonly linked to a sugar moiety (glycosides). The sugars can be attached as monosaccharides, disaccharides or oligosaccharides through *C*- or *O*-glycosylation (Ferreyra et al., 2012). Flavonoids such as kaempferol, quercetin and isorhamnetin are the most common and exist in abundance in plant tissue as glycosides (Ono et al., 2010).

The current development in plant metabolomics using mass spectrometry (MS) has made it possible to characterize a number of natural compounds including flavonoids (Geng et al., 2009; Iswaldi et al., 2011; Rodriguez-Perez et al., 2015). This technique has given researchers the ability to compare samples based on their similarities and differences in a semi-automated and untargeted manner (Bedair and Sumner, 2008; Schripsema, 2010; Sumner and Hall, 2013). It also provides an accurate analysis of a wide-range of metabolites of different polarities compared to standard LC methods (Madala et al., 2014). In this current study, ultra high-performance liquid chromatography (UHPLC) coupled with high resolution quadrupole time-of-flight mass spectrometer (q-TOF-MS) was utilized. This technique is advantageous compared to other LC methods because of its high sensitivity and selectivity. To the best of our knowledge, this is the first report on the flavonoid composition of *M. ovalifolia*. Therefore, the objective of this research was to evaluate the differences in flavonoids composition of *M. ovalifolia* and *M. oleifera* using ultra high-performance liquid chromatography

* Corresponding author. Tel.: +27 11 559 4573; fax: +27 11 559 2605.
E-mail address: emadala@uj.ac.za (E. Madala).

http://dx.doi.org/10.1016/j.sajb.2015.12.007
0254-6299/

coupled with high resolution quadrupole time-of-flight mass spectrometer (UPLC-ESI-qTOF-MS).

2. Materials and methods

2.1. Plant collection

The leaves of *M. oleifera* were sampled from the *Moringa* farm in Lebowakgomo, Limpopo, South Africa. Additionally, *M. oleifera* samples were also collected from three different locations in Namibia (Windhoek, Rundu and Katima Mulilo). The *M. ovalifolia* samples were sampled from the following locations in Namibia: Rocky hills located in the *Moringa* Safari Farm in the Otjozondjuba Region (site 1), Sprokieswoud "enchanted forest" located in Etosha National Park (site 2), Dolomite Hills located near Halalo (site 3) and Tsumeb located along B1 road side (site 4). All *M. ovalifolia* leaves were pooled together.

For experimentation, all the leaves were plucked by hand and laid on shelves, 30 cm apart, in a dry cool place with good ventilation and in shade. Prior to extraction, leaves were ground to a fine powder with a pestle and mortar.

2.2. Metabolite extraction

The ground leaf powder (2 g) was extracted with 20 mL of 80% aqueous methanol (MeOH) by sonicating for 30 min using an ultrasonic cleaning bath (SB-120DT, Loyal Key Group, Hong Kong). Following, the samples were centrifuged at 5000 rpm for 10 min at room temperatures (25 °C) to remove the tissue debris from the homogenate. The supernatant liquid was decanted into a round bottom flask and the solvent evaporated to approximately 1.5 mL using a rotary evaporator under reduced pressure at 60 °C. The extract was then transferred to a 2 mL Eppendorf tube and further dried using a speed vacuum concentrator at 45 °C under negative vacuum (Eppendorf, Merck, South Africa). The dried extracts were reconstituted in 1 mL 50% MeOH and finally filtered through 0.22 μm nylon filters. The resulting extracts were then stored in a freezer at −20 °C to avoid degradation until they were analyzed on the UHPLC-qTOF-MS.

2.3. UHPLC conditions

For the evaluation of metabolites in *M. oleifera* and *M. ovalifolia*, UHPLC coupled to a SYNAPT G1 HDMS mass spectrometer was used to generate accurate mass data (Waters, MA, USA). The chromatographic separation of the *Moringa* extracts (1 μL) was accomplished using a 30 min long gradient chromatographic method on a Waters BEH C_{18} reverse phase column (150 mm × 2.1 mm, 1.7 μm). For chromatographic elution, a binary solvent mixture was used consisting of 0.1% formic acid in deionized water (Eluent A) and 0.1% formic acid in acetonitrile (Eluent B). The chromatographic gradient was as follows: the initial conditions were 98% eluent A followed by multiple gradients to 5% A at 26 min. The conditions were kept constant for 1 min and then changed to the initial conditions. The flow rate was set at 0.4 mL/min and maintained for 1 min after which the column was finally re-equilibrated for 3 min prior to the next run. Chromatographic separation was monitored initially using a photodiode-array (PDA) detector coupled in tandem to an electrospray ionization mass spectrometer (ESI-MS) detector. The PDA detector was set to scan 200–500 nm and collecting 20 spectra per second.

2.4. MS-conditions

For MS detection, an electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-qTOF-MS) detector was used. Leucine enkephalin (556.227/554.2615 Da) was used as reference calibrant to obtain typical mass accuracy of at least 5 MDa at a constant flow rate of 0.1 mL/min. The mass spectrometer was operated in negative and positive ionization mode with a capillary voltage of 2.5 kV, the sampling cone at 30 V, and the extraction cone at 4 V, multichannel plate detector potential at 1600 V, source temperature at 120 °C, desolvation temperature at 450 °C, cone gas flow at 50 L/h, and desolvation gas flow at 550 L/h. The scan time was 0.2 s and an inter-scan delay of 0.02 s covering the mass range of 100 to 1000 Da. The fragmentation patterns of the compounds were obtained in a MS file with three functions. The settings were as follows: Function 1 unfragmented using tune page settings; Function 2 a Trap Collision Energy Ramp (TCER) of 10–30 eV; Function 3 a TCER of 15–50 eV.

3. Results and discussion

Metabolites were extracted from dried leaves and subsequent targeted flavonoid profiling of extracted material was carried out for both species. From the crude UHPLC–MS chromatograms (Fig. 1), a total of 17 chromatographic peaks showing typical flavonoid fragmentation patterns were detected (Table 1). From the structures of the identified flavonoid (Fig. 2), it can be seen that almost all harbor kaempferol, quercetin, apigenin or isorhamnetin aglycone core moieties (Table 1). The results further reveal the presence of a wide spectrum of sugar moieties attached to these flavonoid core structures, showing interesting glycosylation capabilities of the two species.

3.1. Kaempferol-O-glycosides

In *M. oleifera* leaves from South Africa, six kaempferol-*O*-glycoside molecules (**2**, **8**, **9**, **13**, **14**, and **15**) were tentatively identified. In *M. oleifera* leaves from Namibia, three out of the six kaempferol-*O*-glycoside molecules were identified (**2**, **9** and **15**). Molecule **2** showed a precursor ion at *m/z* 651.1537 $[M-H]^-$ was identified as kaempferol acetyl dihexose. The product ions shown in the MS spectrum were indicative of the elimination of hexosyl and acetyl-hexosyl-hexose moieties, respectively (Table 1). Similarly, the same molecule with similar fragmentation patterns was identified in *M. oleifera* sampled from Namibia. To the best of our knowledge, this is the first time kaempferol acetyl dihexose has been detected in *M. oleifera* species, which is noteworthy because of its beneficial effects on human kind. Molecule **9** was identified as kaempferol hexose with a precursor ion at *m/z* 447.0917 $[M-H]^-$ in *M. oleifera* samples from both locations. The product ion shown in the MS spectrum was due to a loss of a hexosyl moiety. The fragmentation pattern from Molecule **9**, shown in Table 1, was also reported from a herb extract of a Cress plant (*Lepidium sativum*) (Justesen, 2000). Molecule **13** was identified as kaempferol hydroxy-methylglutarate-hexose with a precursor ion at *m/z* 591.1358 $[M-H]^-$ The MS spectrum showed product ions indicating the elimination of a hydroxy-methylglutarate and hydroxy-methylglutarate hexosyl moieties, respectively. The fragmentation pattern of Molecule **13**, shown in Table 1, is similar to previously published data on *M. oleifera* by Rodriguez-Perez et al. (2015). Compound **14** was identified as kaempferol acetyl hexose with a precursor ion at *m/z* 489.0955 $[M-H]^-$ based on the precursor and fragmented ions shown in Table 1, indicating a loss of an acetyl moiety and cleavage of a hexosyl moiety. The fragmentation pattern of Molecule **14** was similar to that identified in strawberry fruit (Kajdžanoska et al., 2010). Molecule **15** was tentatively identified, in the *M. oleifera* samples from both locations, as kaempferol malonyl hexose with a precursor ion at *m/z* 533.0917 $[M-H]^-$ due to the decarboxylation and elimination of the malonyl-hexosyl moiety, respectively (Table 1). The fragmentation profile of Molecule **15** was tentatively identified and found to be similar to one found in strawberry fruits (Aaby et al., 2007).

In *M. ovalifolia*, only one kaempferol-*O*-glycoside molecule was observed. Molecule **8** was identified as kaempferol rutinoside with a precursor ion at *m/z* 593.1460 $[M-H]^-$ with a fragmentation ion indicating the elimination of a rutinoside moiety (Table 1). The fragmentation pattern of Molecule **8** was reported by Vagiri et al. (2015) in the

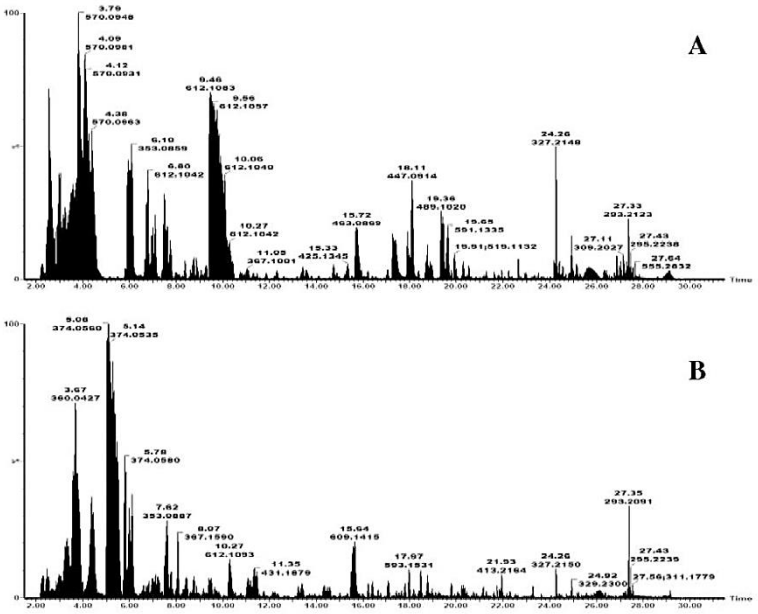


Fig. 1. UHPLC-qTOF-MS chromatogram of *Moringa oleifera* (A) and *Moringa ovalifolia* (B) leaf extracts showing flavonoid peaks detected in negative ionization mode.

identification of kaempferol rutinoside in black currant leaves (*Ribes nigrum*, L). The aglycone radical formation phenomenon was not observed in the case of kaempferol fragmentation. The identification of kaempferol is important since this molecule is gaining application in the management of cancer-related diseases (Chen and Chen, 2013).

3.2. Quercetin-O-glycosides

In *M. oleifera* sampled from South Africa, six quercetin-*O*-glycoside molecules (**3**, **4**, **5**, **6**, **7**, and **12**) were identified, similarly, four out the six quercetin-*O*-glycoside molecules were tentatively identified in *M. oleifera* sampled from Namibia. Molecule **3** was identified as quercetin acetyl dihexose with a precursor ion at *m/z* 667.1467 $[M-H]^-$ in both *M. oleifera* samples. The fragmented ions obtained were due to the elimination of an acetyl, acetyl hexose and acetyl-hexosyl-hexose moieties, respectively (Table 1). Molecule **3** was reported by Barros et al. (2012) in the identification of quercetin acetyl dihexose in the flower (*C. monogyna*). Molecule **5** was identified as quercetin hexose with a precursor ion at *m/z* 463.0844 $[M-H]^-$. The fragmented product ion was due to the elimination of a hexosyl moiety (Table 1)

Table 1
Tentative characterization the bioactive compounds of *Moringa oleifera* and *Moringa ovalifolia*.

Compound no.	Rt (min)	Mass (*m/z*)	*M. oleifera* (South Africa)	*M. oleifera* (Namibia)	*M. ovalifolia* (Namibia)	Aglycone	Sugar attachment	Molecular formula	Diagnostic fragments
1	8.69	593.1465	√	√		Apigenin	6,8 *C*-dihexose	$C_{27}H_{30}O_{15}$	473.1038; 353.0592
2	9.98	651.1537	√	√		Kaempferol	Acetyl dihexose	$C_{29}H_{32}O_{18}$	489.0937; 284.0288
3	12.57	667.1467	√			Quercetin	Acetyl dihexose	$C_{29}H_{32}O_{18}$	624.1457; 462, 0851; 300.0931
4	14.07	609.1457			√	Quercetin	Rutinoside	$C_{27}H_{30}O_{16}$	300.0199
5	14.21	463.0844	√	√		Quercetin	Hexose	$C_{12}H_{20}O_{12}$	301.0831
6	16.41	607.1299	√	√		Quercetin	Hydroxy-methylglutaroyl hexose	$C_{27}H_{28}O_{16}$	463.0796; 301.0268
7	16.46	505.0932	√	√		Quercetin	Acetyl hexose	$C_{23}H_{21}O_{13}$	301.0281
8	16.48	593.1460			√	Kaempferol	Rutinoside	$C_{26}H_{28}O_{16}$	285.0309
9	16.62	447.0917	√	√		Kaempferol	Hexose	$C_{21}H_{20}O_{11}$	285.0931
10	16.98	623.1655			√	Isorhamnetin	Rutinoside	$C_{27}H_{30}O_{16}$	315.0421
11	17.17	477.1031	√			Isorhamnetin	Hexose	$C_{22}H_{22}O_{11}$	315.1032
12	17.35	549.0897	√	√		Quercetin	Malonyl hexose	$C_{23}H_{22}O_{11}$	505.0961; 300.0905
13	18.10	591.1358	√			Kaempferol	Hydroxy-methylglutaroyl hexose	$C_{27}H_{28}O_{15}$	447.0931; 285.0330
14	18.11	489.0955	√	√		Kaempferol	Acetyl hexose	$C_{23}H_{22}O_{11}$	285.1036
15	18.48	533.0917	√	√		Kaempferol	Malonyl hexose	$C_{24}H_{22}O_{14}$	489.0407; 285.1036
16	18.48	621.1465	√			Isorhamnetin	Hydroxy-methylglutaroyl hexose	$C_{28}H_{28}O_{11}$	459.1411; 315.1410
17	18.41	519.1108	√	√		Isorhamnetin	Acetyl hexose	$C_{24}H_{24}O_{13}$	315.0502

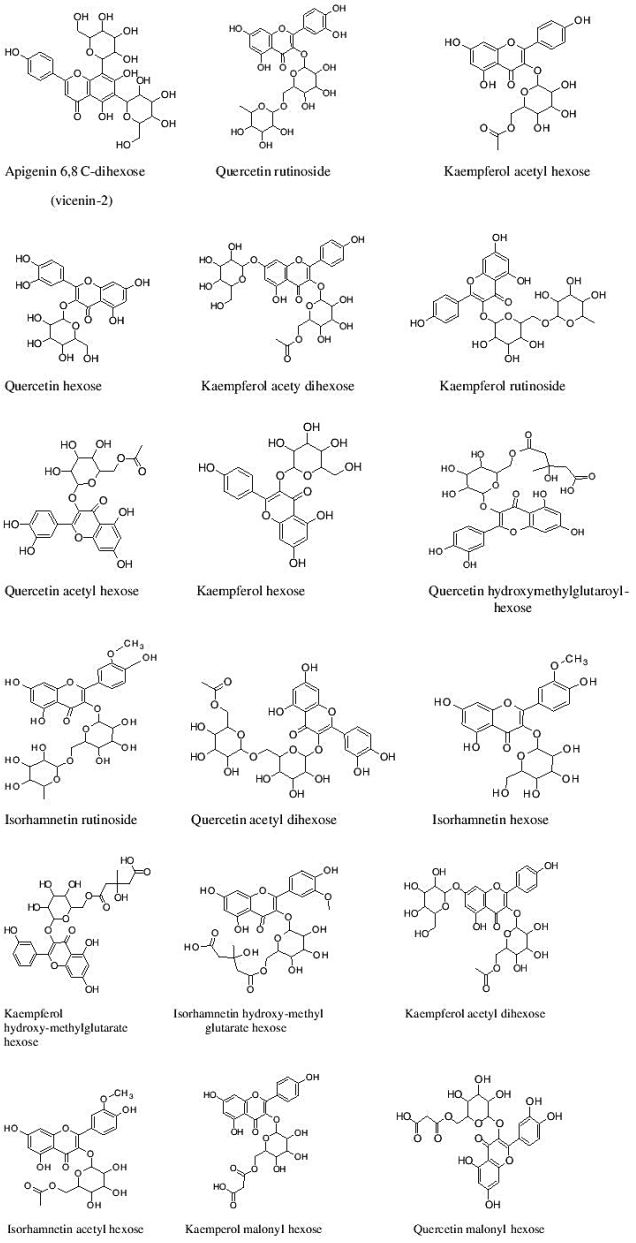


Fig. 2. Structures of flavonoid aglycones identified in the leaves of *Moringa oleifera* and *Moringa ovalifolia*.

(Justesen, 2000). The fragmentation pattern of Molecule **5** was reported by Barros et al. (2012) in the identification of quercetin hexose in *Crataegus monogyna* flowers. Molecule **6** was identified as quercetin hydroxy-methylglutaroyl-hexose with a precursor ion at *m/z* 607.1299 $[M-H]^-$. The product ions obtained were due to the loss of hydroxy-methyl-glutaroyl and hydroxy-methyl-glutaroyl hexosyl moieties (Table 1). The fragmentation pattern of Molecule **6** is consistent with already published data in *M. oleifera* leaves species from Madagascar (Rodriguez-Perez et al., 2015). Molecule **7** was identified as quercetin acetyl hexose with a precursor ion at *m/z* 505.0932 $[M-H]^-$ in both *M. oleifera* samples. The product ion obtained indicated the elimination of an acetyl-hexosyl moiety (Table 1). Molecule **7** was reported by Tuseveski et al. (2013) in the identification of quercetin acetyl hexose in the seeds of *Hypericum perforatum* (*H. perforatum*). Molecule **12** was identified as quercetin malonyl hexose with a precursor ion at *m/z* 549.0897 $[M-H]^-$ in both *M. oleifera* samples. The product ions obtained from the MS spectrum were due to decarboxylation (44 Da) and loss of malonyl-hexosyl moiety (Table 1). The fragmentation pattern of Molecule **12** is consistent with already published data in *M. oleifera* leaves species from Madagascar (Rodriguez-Perez et al., 2015).

In *M. ovalifolia*, one quercetin-*O*-glycoside molecule was identified. Molecule **4** was identified as quercetin rutinoside (rutin) with a precursor ion at *m/z* 609.1457 $[M-H]^-$. The product ion shown in Table 1 was primarily due to the elimination of a rutinoside molecule. The fragmentation pattern of Molecule **4** is consistent with previously published data on wild rose hip *Rosa canina* by Hvattum (2002).

It was observed in our study that Molecule **5** and Molecule **12** gave, in addition to the deprotonated aglycone ion (300.0219 and 300.0217), a mass signal at 301.0291 and 301.0316 (Results not shown). Generally, research has shown that quercetin is the main dietary flavonoid that is linked to many properties that suppress certain aliments associated with chronic diseases such as quercetin rutinoside (rutin). Due to the antioxidant activity of rutin, this molecule displays certain biological potentials such as protecting liver cells and suppressing hemoglobin oxidation. Furthermore, the associated anti-inflammatory properties are useful as a treatment for certain chronic diseases (Grinberg et al., 1994; Janbaz et al., 2002; Rotelli et al., 2003; Obied et al., 2005; Rodriguez-Perez et al., 2015).

3.3. Isorhamnetin-O-glycoside

Isorhamnetin is the methylated version of quercetin and has been reported in *M. oleifera* plants (Rodriguez-perez et al., 2015). In this regard, methylation of isorhamnetin is an important process of quercetin metabolism (Hanhineva et al., 2008; Rodriguez-perez et al., 2015).

In *M. oleifera* from South Africa, four isorhamnetin-*O*-glycoside molecules were identified (**10**, **11**, **16** and **17**). Molecule **11** was identified as isorhamnetin hexose with a precursor ion at *m/z* 477.1031 $[M-H]^-$. The product ion obtained from the MS spectrum showed a loss of a hexosyl moiety (Table 1). Galati et al. (2003) also reported Molecule **11** for the identification of isorhamnetin hexose in *Opuntia ficus indica* (L.) *Mill*. Molecule **16** was identified as isorhamnetin hydroxy-methylglutaroyl hexose with a precursor ion at *m/z* 621.1465 $[M-H]^-$. The product ions obtained corresponded to the loss of a hexosyl, hydroxy-methylglutaroyl hexose moieties (Table 1). Molecule **17** at Rt 18.41 min was identified as isorhamnetin acetyl hexose with a precursor ion at *m/z* 519.1108 $[M-H]^-$. The Isorhamnetin aglycone obtained was primarily due to the loss of the of acetyl-hexosyl moiety (Table 1). Molecule **17** was reported also by Gutzeit et al. (2007) for the identification of isorhamnetin acetyl hexose in Sea Buckthorn juice concentrate (*Hippophaë rhamnoides* L. ssp. *rhamnoides*).

Only one isorhamnetin-*O*-glycoside molecule (**10**) was identified in *M. ovalifolia*. This molecule was identified as isorhamnetin rutinoside with a precursor ion at *m/z* 623.1655 $[M-H]^-$. Its MS spectrum showed a product ion representing an isorhamnetin aglycone due to the elimination of a rutinoside moiety (diagnostic fragment shown in Table 1). Molecule **10** was similarly reported, as with Molecule **17**, by Gutzeit et al. (2007). In Molecule **10**, the radical ion formation was detected also for derivatives of isorhamnetin visible at a fragment of *m/z* 314.0419.

It is known that conjugated metabolites possess different biological activities than the parent compounds and in fact, according to Lemańska et al. (2004), the decrease in antioxidant activity of quercetin was due to methylation of hydroxyl groups (Lemańska et al., 2004). Though naturally methylated, epidemiological studies have also shown isorhamnetin and its glycosides to possess anti-inflammatory properties (Antunes-Ricardo et al., 2015).

3.4. Apigenin-C-glycoside

During chromatographic separation, flavonoids are known to elute in clusters since such compounds have very similar polarities. Moreover, during MS analyses, flavonoid hexose molecules undergo a unique fragmentation pattern which is characterized mainly by their tendency to readily lose the sugar moieties (Geng et al., 2009). However, in the current study, one flavonoid molecule which did not fulfill all the above was identified in *M. oleifera* as Molecule **1** (Table 1, Fig. 2). This molecule was identified as *C*-hexose flavonoid (due to its inability to lose sugar moieties) and based on the accurate mass and MS fragmentation pattern, was identified as apigenin 6, 8 *C*-dihexose (vicenin-2) with a precursor ion at *m/z* 593.1465 $[M-H]^-$. The MS spectrum showed product ions at 473.1038 $[M-H-120]^-$ and *m/z* 353.0592 $[M-H-240]^-$, resulting from internal sugar fragmentation (diagnostic fragment illustrated in Table 1) (Cuyckens and Claeys, 2004; Gobbo-Neto et al., 2008; Guimarães et al., 2013; Choi et al., 2014). Recently, a molecule with a similar precursor and product ions was identified in *M. oleifera* as Multiflorin B (Rodriguez-Perez et al., 2015). Although these two molecules resulted in the same fragmentation patterns, the absence of a kaempferol aglycone ion at *m/z* 284/285 in the identity made by Rodriguez-Perez et al. (2015) is an indication that this molecule is not Multiflorin B. As such, we believe this molecule was wrongly annotated and is most likely vicenin-2. Generally, apigenin has been conveyed to possess pharmaceutical properties such as anti-inflammatory, antioxidant and even anti-Alzheimer's activity (Choi et al., 2014). Similar precursor ions and fragmentation patterns were reported in *Passiflora edulis* (*Passifloraceae*) (Othify et al., 2015). Recently, vicenin-2 isolated from *Artemisia capillaris*, has been reported to possess potent anti-diabetic properties (Nurul-Islam et al., 2014). Interestingly, reports have confirmed vicenin-2 to exhibit antioxidant properties that protect human lymphocytes against the effect of radiation at minimal and non-lethal doses (Vrinda and Devi, 2001). The study thus shows that large scale profiling of plant metabolites provides a broader understanding of plant biochemistry and aids in the identification of novel metabolites (Sumner et al., 2015). Furthermore, a similar fragmentation pattern was identified.

3.5. Comparison with total flavonoids

In a previous study by Matshediso et al. (2015), the influence of extraction temperature in pressurized hot water extraction (PHWE) on the total amount of quercetin and kaempferol extracted was conducted in *M. oleifera* and *M. ovalifolia* plant samples and their results show that *M. oleifera* had higher total amount of these compounds than *M. ovalifolia*. Our study supports these findings, but goes further to show the actual complete structures of the flavonoid molecules with the respective sugar attachments. More importantly, it is very evident that the glycosylation machinery of *M. oleifera* is more complex than that of *M. ovalifolia*. Therefore, the results of the current study highlight the importance of glycosylation patterns in taxonomic studies.

From these results, it is very evident that *M. oleifera* is capable of carrying various sugar attachments which may affect the function of the

consortium of flavonoid molecules. For instance, molecules harboring hydroxy-methylglutarate (and on its own) are known to suppress the synthesis of cholesterol (Siperstein and Fagan, 1966). Thus, a flavonoid containing a certain sugar with further attachment such as hydroxymethylglutarate is capable of exhibiting dual function (being an antioxidant and antagonist for cholesterol synthesis or accumulation). Generally, flavonoids are glycosylated for storage purposes in different organelles of plants. However, the same sugars also help for bioavailability of these compounds for human consumption (thus can be absorbed easily during digestions). Our results further show, for the first time, the importance of sugar attachment for taxonomical grouping of related species. Our results clearly demonstrate that the two species in question have similar qualitative flavonoid (aglycone) profile but differ significantly in the sugar attachment. Based on our findings, *M. ovalifolia* is capable of only attaching rutinoside on the flavonoid skeleton. Surprisingly, none of the flavonoids detected in *M. oleifera* was found to contain rutinoside, however, traces of rutin were previously shown to accumulate in the leaves but absent in stems and petioles of *M. oleifera* (Amaglo et al., 2010).

3.6. Possible causes of the differences in the flavonoid content

Environmental and genetic factors have been attributed to be some of the causes of plant metabolome differences (Alonso-Amelot, 2008; Ncube et al., 2011; Bernal et al., 2013; Khanum et al., 2013; Brunetti et al., 2013). In our study, however, *M. oleifera* plants were harvested from both South Africa and Namibia while *M. ovalifolia* was collected only from Namibia. Therefore, the observed differences between the species are most likely due to genetic differences rather than environmental variations. Recently, *M. oleifera* cultivars growing under similar environmental conditions have been shown to contain varying concentrations of pharmacologically relevant metabolites (Ndhlala et al., 2014). Moreover, the distribution and presence of acetyl-(4-α-L-rhamnopyranosyloxy)-benzyl glucosinolates isomers was shown to be significantly different within plants of the same ecotype, with some not even exhibiting traces of these isomers (Förster et al., 2015). While environmental conditions cannot be ruled out completely, in this case the differences in flavonoid content are too huge and thus genetic variation is speculated to be playing a major role. Moreover, the leaves were harvested in areas with natural soil conditions such that any possible environmental conditions could be more climatic ones. Importantly, environmental impact under normal conditions can only play a major role when comparing the amount of the compounds and not the presence (quality), unless under extreme environments. In another ongoing work in our research group we have found that the flavonoid content in *M. stenopelata* (collected from Namibia) is much closer to that of *M. oleifera*, suggesting again the inferiority of *M. ovalifolia*,with regards to its glycosylation capabilities, when compared to the other two species (Results not shown). A study by Kanayama et al. (2013) investigated the seasonal changes in abiotic stress tolerance and concentration of tocopherol, sugar and ascorbic acid in sea buckthorn leaves and stems. Results of this study revealed that ascorbic acid accumulated in winter more than during other seasons. Another study by Ncube et al. (2011) investigated a comparative study of the antimicrobial and phytochemical properties between outdoor grown and micropropagated *Tulbaghia violacea* harvested plants, and found that the former showed more robust antibacterial activity. Overall, the amounts of phenolic acids and neolignan of entire leaves were found to increase with altitude, while the total amounts of flavonoids in the leaf cuticles decreased in altitude and seasonal changes of phenolic compounds in *Buxus sempervirens* leaves and cuticles by Bernal et al. (2013).

4. Conclusion

Our results show that the two *Moringa* species in question have similar qualitative flavonoid (aglycone) profiles but differ significantly in the sugar attachment (glycosylation). Furthermore, the sugar attachment (type of sugar in particular) is seen to play a pivotal role in taxonomical differentiation of the two species, owing to the genetically distinctive glycosylation capabilities. As such, this data suggests a possible evolutionary strategy in which plants (*M. oleifera* in this case) may/can increase the quantity of metabolites (flavonoids in this case) through glycosylation. Furthermore, the current work reaffirms what has previously been reported, in that other metabolite structural diversification strategies, like isomerism, can also be a way in which plants prepare themselves against a wide spectrum of stresses by increasing metabolite reserves. As such, these diversifications, in this case glycosylation, can render flavonoids from *M. oleifera* to exhibit a wider pharmacological spectrum as compared to those from *M. ovalifolia*.

Acknowledgements

South African National Research Foundation (NRF), University of Johannesburg and the University of the Witwatersrand are thanked for financial support. Dr. Lizelle Piater is thanked for critical reading of the manuscript.

References

Aaby, K., Ekeberg, D., Skrede, G., 2007. Characterization of phenolic compounds in strawberry fruits by different HPLC detectors and contribution of individual compounds to total antioxidant aapacity. Journal of Agricultural and Food Chemistry 55, 4395–4406.
Alonso-Amelot, M.E., 2008. In: Arl-ur-Rahaman (Ed.)High Altitude Plants, Chemistry of Acclimation and Adaptation in Studies in Natural Products Chemistry Vol. 34. Elsevier Publishers, UK, pp. 883–903.
Amaglo, N.K., Bennet, R.N., LoCurto, R.B., Rosa, E.A.S., LoTurco, V., Giuffrid, A., LoCurto, A., Crea, F., Timpo, G.M., 2010. Profiling selected phytochemicals and nutrients in different tissues of the multipurpose tree *Moringa oleifera* L. grown in Ghana. Food Chemistry 122, 1047–1054.
Antunes-Ricardo, M., Gutiérrez-Uribe, J.A., Martínez-Vitela, C., Sergio O. Serna-Saldívar, S.O., 2015. Topical anti-inflammatory effects of isorhamnetin glycosides isolated from *Opuntia ficus-indica*. BioMed Research International 2015.
Barros, L., Dueñas, M., Carvalho, A.M., Ferreira, I.C., Santos-Buelga, C., 2012. Characterization of phenolic compounds in flowers of wild medicinal plants from Northeastern Portugal. Food and Chemical Toxicology 50, 1576–1582.
Bedair, M., Sumner, L.W., 2008. Current and emerging mass-spectrometry technologies for metabolomics. Trends in Analytical Chemistry 27, 238–250.
Bernal, M., Llorens, L., Julkunen-Tiito, R., Badosa, J., Verdaguer, D., 2013. Altitudal and seasonal changes of phenolic compounds in *Buxus sempervirens* leaves cuticles. Plant Physiology and Biochemistry 70, 471–482.
Brunetti, C., George, R.M., Tattin, M., Field, K., Davey, M.P., 2013. Metabolomics in plant environmental physiology. Journal of Experimental Botany 64, 4011–4020.
Chen, A.Y., Chen, Y.C., 2013. A review of the dietary flavonoid, kaempferol on human health and cancer chemoprevention. Food Chemistry 15, 2099–2107.
Choi, J.S., Nurul Islam, M., Yousof Ali, M., Kim, E.J., Kim, Y.M., Jung, H.A., 2014. Effects of C-glycosylation on anti-diabetic, anti-Alzheimer's disease and anti-inflammatory potential of apigenin. Food and Chemical Toxicology 64, 27–33.
Cook, N.C., Samman, S., 1996. Review: flavonoids chemistry, metabolism, cardioprotective effects, and dietary sources. The Journal of Nutritional Biochemistry 7, 66–76.
Cuyckens, F., Claeys, M., 2004. Mass spectrometry in the structural analysis of flavonoids. Journal of Mass Spectrometry 39, 1–5.
Ferreyra, M.L.F., Rius, S.P., Casati, P., 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. Frontiers in Plant Science 2012, 3–222.
Förster, N., Ulrichs, C., Schreiner, M., Müller, C.T., Mewis, I., 2015. Development of a reliable extraction and quantification method for glucosinolates in *Moringa oleifera*. Food Chemistry 166, 456–464.
Galati, E.M., Mondello, M.R., Giuffrida, D., Dugo, G., Miceli, N., Pergolizzi, S., Taviano, M.F., 2003. Chemical characterization and biological effects of Sicilian *Opuntia ficus indica* (*L.*) Mill. fruit juice: antioxidant and antiulcerogenic activity. Journal of Agricultural and Food Chemistry 51, 4903–4908.
Geng, P., Sun, J., Zhang, R., He, J., Abliz, Z., 2009. An investigation of the fragmentation differences of isomeric flavonol-O-glycosides under different collision-induced dissociation based mass spectrometry. Rapid Communications in Mass Spectrometry 23, 1519–1524.
Gobbo-Neto, L., Gates, P.J., Lopes, N.P., 2008. Negative ion "chip-based" nanospray tandem mass spectrometry for the analysis of flavonoids in glandular trichomes of *Lychnophoraericoides* Mart. (Asteraceae). Rapid Communications in Mass Spectrometry 22, 3802–3808.
Grinberg, L.N., Rachmilewitz, E.A., Newmark, H., 1994. Protective effects of rutin against hemoglobin oxidation. Biochemical Pharmacology 48, 643–649.
Guimarães, R., Barros, L., Dueñas, M., Calhelha, R.C., Carvalho, A.M., Santos-Buelga, C., Queiroz, M.J., Ferreira, I.C., 2013. Nutrients, phytochemicals and bioactivity of wild Roman chamomile: a comparison between the herb and its preparations. Food Chemistry 136, 718–725.

Gutzeit, D., Wray, V., Winterhalter, P., Jerz, G., 2007. Preparative isolation and purification of flavonoids and protocatechuic acid from sea buckthorn juice concentrate (*Hippophae˜ rhamnoides* L. ssp. *rhamnoides*) by high-speed counter-current chromatography. Chromatographia 65, 1–7.
Hanhineva, K., Rogachev, I., Kokko, H., Mintz-Oron, S., Venger, I., Karenlampi, S., Aharoni, A., 2008. Non-targeted analysis of spatial metabolite composition in strawberry (Fragaria × ananassa) flowers. Phytochemistry 69, 2463–2481.
Hvattum, E., 2002. Determination of phenolic compounds in rose hip (*Rosa canina*) using liquid chromatography coupled to electrospray ionasation tandem mass spectrometry and diodide-array detection. Rapid Communications in Mass Spectrometry 166, 655–662.
Iswaldi, I., Arraez-Roman, D., Rodriguez-Medina, I., Beltran-Debon, R., Joven, J., 2011. Identification of phenolic compounds in aqueous abd ethanolic rooibos extracts (*Aspalathuslinearis*) by HPLC-ESI-MS (TOF/IT). Analytical and Bioanalytical Chemistry 400, 3643–3654.
Janbaz, K.H., Saeed, S.A., Gilani, A.H., 2002. Protective effect of rutin on paracetamol- and CCl_4-induced hepatotoxicity in rodents. Fitoterapia 73, 557–563.
Justesen, U., 2000. Negative atmospheric pressure chemical ionization low-energy collision activation mass spectrometry for the characterization of flavonoids of extracts of fresh herbs. Journal of Chromatography A 902, 369–397.
Kajdžanoska, M., Gjamovski, V., Stefova, M., 2010. HPLC-DAD-ESI-MS^n identification of phenolic compounds in cultivated strawberries from Macedonia. Macedonian Journal of Chemistry and Chemical Engineering 2, 181–194.
Kanayama, Y., Sato, K., Ikeda, H., Tamura, T., Nishiyama, M., Kanahama, K., 2013. Seasonal changes in abiotic stress tolerance and concentration of tocopherol, sugar and ascorbic in sea buckthorn leaves and stem. Scientia Horticulturae 164, 232–237.
Khanum, R., Muntaz, A.S., Kumar, S., 2013. Predicting the impacts of climate change on medicinal asclepiads of Pakistan using maxent modelling. Acta Oecologica 49, 23–31.
Lemańska, K., der Woude, van, Szymusiak, H., Boersma, M.G., Gliszczyńska-Swigło, A., Rietjens, I.M., Tyrakowska, B., 2004. The effect of catechol O-methylation on radical scavenging characteristics of quercetin and luteolin—a mechanistic insight. Free Radical Research 38, 639–647.
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 Sonitrosoacetophenone-elicited *Nicotiana tabacum* and *Sorghum bicolor* cells. Springer Plus 3, 254.
Matshediso, P.G., Cukrowska, E., Chimuka, L., 2015. Development of pressurised hot water extraction (PHWE) for essential compounds from *Moringa oleifera* leaf extracts. Food Chemistry 172, 423–427.
Ncube, B., Ngunge, V.N.P., Van Staden, J.F., 2011. A comparative study of the antimicrobial and phytochemical properties between outdoor grown and micropropagated *Tulbaghia violacea* Harv. plants. Journal of Ethnopharmacology 134, 775–780.
Ndhlala, A., Mulaudzi, R., Ncube, B., Abdelgadir, H.A., du Plooy, C.P., Van Staden, J., 2014. Antioxidant, antimicrobial and phytochemical variations in thirteen *Moringa oleifera* Lam. cultivars. Molecules 19, 10480–10494.
Nijveldt, R.J., van Nood, E., van Hoorn, E., Boelens, P.G., van Norren, K., van Leewun, P.A., 2001. Flavonoids: a review of probable mechanisms of action and potential applications. The American Journal of Clinical Nutrition 74, 418–425.
Nurul-Islam, Md, Ishita, I.J., Jung, H.A., Choi, J.S., 2014. Vicenin 2 isolated from *Artemisia capillaris* exhibited potent anti-glycation properties. Food and Chemical Toxicology 69, 55–62.
Obied, H.K., Allen, M.S., Bedgood, N.R., Prenzler, P.D., Robards, K., Stockman, R., 2005. Bioactivity and analysis of biofenols recovered from olive mill waste. Journal of Agricultural and Food Chemistry 53, 823–837.
Ono, E., Homma, Y., Horikawa, M., Kunikane-Doi, S., Imani, H., Takahashi, S., Kawaii, Y., Ishiguro, Y., Nakayama, T., 2010. Functional differentiation of glycosyltransferases that contribute to chemical diversity of bioactive flavanol glycosides in grapevines (*Vitis vinifera*). Plant Cell 22, 2856–2871.
Othify, A., George, C., Elsayed, A., Farag, A.M., 2015. Mechanistic evidence of *Passiflora edulis* (Passifloraceae) anxiolytic activity in relation to its metabolite fingerprint as revealed via LC-MS and chemometrics. Food & Function.
Pakade, V., Cukrowska, E., Chimuka, L., 2012. Metal and flavonol contents of *Moringa oleifera* grown in South Africa. South African Journal of Science 1–7.
Robak, J., Gryglewski, R.J., 1996. Bioactivity of flavonoids. Polish Journal of Pharmacology 6, 555–564.
Rodriguez-Perez, C., Quirantes-Pine, R., Fernandez-Gutierrez, A., Segura-Carretero, A., 2015. Optimization of extraction method to obtain a phenolic compounds-rich extract *Moringa oleifera* Lam leaves. Industrial Crops and Products 66, 246–254.
Rotelli, A.E., Guardia, T., Juarez, A.O., De la Rocha, N.E., Pelzer, L.E., 2003. Comparative study of flavonoids in experimental models of inflammation. Pharmacological Research 48, 601–606.
Schripsema, J., 2010. Application of NMR in plant metabolomics: techniques, problems and prospects. Phytochemical Analysis 21, 14–21.
Siperstein, M.D., Fagan, V.M., 1966. Feedback control of mevalonate synthesis by dietary cholesterol. The Journal of Biological Chemistry 241, 602–609.
Sumner, L.W., Hall, R.D., 2013. Metabolomics across the globe. Metabolomics 9258–9264.
Sumner, L.W., Lei, Z., Nikolau, B.J., Saito, K., 2015. Modern plant metabolomics: advanced natural product gene discoveries improved technologies and future prospects. Natural Product Reports 32, 105–352.
Tuseveski, O., Stanoeva, J.P., Stefova, M., Simic, S.G., 2013. Phenolic profile of dark-grown and photoperiod-exposed *Hypericum perforatum* L. hairy root cultures. The Scientific World Journal 2013.
Vagiri, M., Conner, S., Stewart, D., Andersson, S.C., Verrall, S., Johansson, E., Rumpunen, K., 2015. Phenolic compounds in blackcurrant (*Ribes nigrum* L.) leaves relative to leaf position and harvest date. Food Chemistry 172, 135–142.
Vrinda, B., Devi, P.U., 2001. Radiation protection of human lymphocyte chromosomes in vitro by orientin and vicenin. Mutation Research 498, 39–46.

Paper II

This paper titled "UPLC-qTOF-MS profiling of pharmacologically important chlorogenic acids and associated glycosides in *Moringa ovalifolia* leaf extracts," was submitted *to South African Journal of Botany* and the presented format is for that journal. It investigates major chlorogenic isomers with the aid of UPLC-ISCID-MS/MS in *Moringa ovalifolia* leaf samples

South African Journal of Botany 108 (2017) 193–199

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

UPLC-qTOF-MS profiling of pharmacologically important chlorogenic acids and associated glycosides in *Moringa ovalifolia* leaf extracts

C. Makita [a], L. Chimuka [a], E. Cukrowska [a], P.A. Steenkamp [b,c], M. Kandawa-Schutz [d], A.R. Ndhlala [e], N.E. Madala [b,*]

[a] Molecular Sciences Institute, School of Chemistry, University of the Witwatersrand, P/B 3, 2050, Johannesburg, South Africa
[b] Department of Biochemistry, University of Johannesburg, P.O. Box 524, Auckland Park, 2006 Johannesburg, South Africa
[c] CSIR Biosciences, Natural Products and Agroprocessing Group, Pretoria 0001, South Africa
[d] School of Chemistry, University of Namibia, P/B 13301 Windhoek, Namibia
[e] Agricultural Research Council, Vegetable and Ornamental Plants (VOP), Private Bag X293, Pretoria 0001, South Africa

ARTICLE INFO

Article history:
Received 13 April 2016
Received in revised form 30 August 2016
Accepted 9 October 2016
Available online 29 October 2016

Edited by K Doležal

Keywords:
Moringa ovalifolia
UPLC-qTOF-MS
Chlorogenic acids
Caffeoylquinic acids
Dicaffeoylquinic acids
p-Coumaroylquinic acids
Feruloylquinic acids
Caffeoylquinic acid glycosides

ABSTRACT

Moringa ovalifolia Dinter & A. Berger (Moringaceae) is a succulent-stemmed plant, endemic to the desert and semi-desert areas of central Namibia and southwestern Angola. Just like other species in the Moringaceae family, *M. ovalifolia* is believed to be rich in health-promoting phytochemicals. However, there are very limited scientific reports on the phytochemical composition and associated biological activities of this plant. Chlorogenic acids (CGAs), major phenolic compounds of *Moringa* species, have been shown to be effective natural remedies for the management of chronic ailments such as diabetes and cardiovascular diseases. Using a UPLC-ISCID-MS/MS method optimized to mimic the MS^n fragmentation of an ion trap-based MS but generating accurate mass data, various isomers of chlorogenic acids and their associated derivatives in the leaves of *M. ovalifolia* were profiled. *M. ovalifolia* was shown to contain *cis* and *trans* isomers of 3-acyl, 4-acyl and 5-acyl *p*-coumaroylquinic (*p*CoQA), caffeoylquinic (CQA) and feruloylquinic acids (FQA) (**1–18**), a single isomer of 3,5-*di*CQA (**19**), 3-CQA-glycoside (**20**) and two regional isomers of the (3′ and 4′) glycosides of 4-CQA (**21, 22**). To the best of our knowledge, this is the first report on the presence of these compounds in *M. ovalifolia*. The results of the current study confirmed the richness of an underutilized *M. ovalifolia* as a source of pharmacological relevant metabolites with potential medicinal applications.

1. Introduction

Chlorogenic acids (CGAs) are natural antioxidants and form part of the most abundant polyphenols in the human diet. They are usually produced by plants as a part of the defence mechanism response against environmental stresses triggered by microbial pathogens, mechanical wounding and direct exposure to UV or visible light (Clifford, 1999, 2000; Madala et al., 2014). CGA are esters of one or more cinnamic acids and quinic acid. The most common cinnamic acids encountered are caffeic, *p*-coumaric and ferulic acid, which give rise to *p*-coumaroylquinic (*p*CoQA), caffeoylquinic (CQA) and feruloylquinic acids (FQA) (Clifford, 1999, 2000). The *trans*-isomers are predominant but *cis*-isomers are known particularly in tissue subject to intense UV irradiation (Clifford et al., 2008). Apart from being strong antioxidants, CGAs and their derivatives are of particular interest as several epidemiological studies have associated a diet rich therein with certain health-promoting properties such as anti-inflammatory, anti-diabetes, anti-viral and anti-cancer activities (Eliel and Ramirez, 1997).

Moringa ovalifolia is a succulent-stemmed plant belonging to the family Moringaceae. It is endemic to the dry, desert and semi-desert areas of central Namibia and southwestern Angola where the species is utilized for several medical and malnutritional conditions. Just like the other species in the family, *M. ovalifolia* is rich in protein, calcium, iron and vitamin C. In many parts of the world, several *Moringa* species are consumed as a nutritional supplement in the form of capsules or leaf powder (Pakade et al., 2013; Khoza et al., 2016). The family Moringaceae is known for its exceptional nutritional values and healing purposes (Pakade et al., 2013). Information on pharmacological activities, phytochemical composition and nutritional values of *M. ovalifolia* is scarce. Most researchers concentrate on the most famous species within the family, *Moringa oleifera* (Makita et al., 2016). Analytical studies have identified a range of phytochemicals, including glucosinolates, sterols, proanthocyanidins, flavonoids and cinnamates in various tissues and organs (Bennett et al., 2003; Ncube et al., 2014; Ramabulana et al., 2015; Khoza et al., 2016; Makita et al., 2016; Mhlongo et al., 2016). It is assumed but not proven that these

* Corresponding author.
E-mail address: emadala@uj.ac.za (N.E. Madala).

http://dx.doi.org/10.1016/j.sajb.2016.10.016
0254-6299/

phytochemicals are also responsible for the perceived benefits posed by members of the *Moringa* species (Siddhuraju and Becker, 2003; Ndong et al., 2007; Verma et al., 2009; Kashiwada et al., 2012; Vongsak et al., 2013a, 2013b, 2013c).

This study was aimed at profiling pharmacologically important chlorogenic acids and associated glycosides (cinnamates) in *M. ovalifolia* with the aid of a recently developed in-source collision-induced dissociation approach of ultra-performance liquid chromatography mass spectrometry technique.

2. Materials and methods

2.1. Chemicals

All chemicals utilized in this study were of analytical grade quality and were obtained from various international suppliers. Briefly, the organic solvents utilized were UPLC/MS grade quality methanol (Romil, MicroSep, South Africa) and acetonitrile (Romil, MicroSep, South Africa). Water was purified by a Milli-Q Gradient A10 system (Millipore, Billerica, MA, USA). Leucine encephalin and formic acid were all purchased from Sigma Aldrich, Germany. The alcoholic cider (Savanna) and coffee were purchased from local liquior store and food outlet, respectively. The caffeoylquinic standards were purchased from PhytoLab, Germany.

2.2. Plant collection

M. ovalifolia leaf samples were collected from four locations in Namibia as follows: Rocky Hills in Moringa Safari Farm about 60 km from Okahandja to Karibib on a D2156 route in Otjozondjupa Region in central Namibia (Site 1), Sprokieswoud (Enchanted Forest), west of Okaukuejo in Etosha National Park (Site 2), Dolomite Hills near Halali in Etosha National Park (Site 3) and 3 km south of Tsumeb along B1 road side (Site 4). Respective herbarium specimens with the following voucher numbers: MNT 01, MNT 02, MNT 03 and MNT 04 representing plants from the four different sites were prepared and stored at the University of Namibia. A preliminary study to determine any major chemical differences between these was carried out and no significant differences were obtained.

2.3. Metabolites extraction

The *Moringa* leaves were air-dried and crushed using a pestle and mortar at room temperature to a fine-coarse powder with a relatively homogeneous particle size to make a composite sample. Metabolites were extracted from a homogeneous sample consisting of equal plant materials from the four different sites. Mixed ground leaves (2 g) were mixed with 80% aqueous methanol (MeOH) (20 mL) by means of sonicating for 30 min using an ultrasonic water bath (SB-120DT, Loyal Key Group (Hong Kong) Co. Ltd). To remove the tissue debris, the homogenates were centrifuged at 5000 × *g* for 10 min. The supernatant was concentrated under low pressure using a rotary evaporator at 55 °C. The resulting concentrate was transferred to a 2 mL Eppendorf tube and dried to completeness using the speed vacuum concentrator centrifuge at 55 °C. The resulting pellet was re-constituted to 1.0 mL by redissolving in 50% MeOH and filtered through 0.22 μm nylon filters. The extracts were kept at − 20 °C until analysed on the UPLC-qTOF-MS.

2.4. UPLC conditions

A Waters Acquity UPLC coupled to a SYNAPT G1 HDMS mass spectrometer was used to generate accurate mass data. The chromatographic separation of the extracts was accomplished on a Waters BEH C_8 column (150 mm × 2.1 mm, 1.7 μm) and with column temperature controlled at 60 °C. A binary solvent mixture was used consisting of water (eluent A) containing 10 mM formic acid (natural pH of 2.3) and acetonitrile (eluent B). The initial conditions were 98% A at a flow rate of 0.4 mL/min and were maintained for 1 min, followed by multiple gradients to 5% A at 26 min. The conditions were kept constant for 1 min and then changed to the initial conditions and re-equilibrated for 3 min prior to the next injection. The total run time was 30 min and the injection volume was 1 μL.

2.5. MS conditions

The SYNAPT G1 mass spectrometer was used in V-optics and operated in electrospray mode to enable detection of phenolic compounds. Leucine enkephalin (50 pg/mL) was used as the reference calibrant to obtain typical mass accuracies between 1 and 5 mDa. The mass spectrometer was operated in negative ion mode with a capillary voltage of 2.5 kV, the sampling cone at 30 V and the extraction cone at 4 V. The scan time was 0.1 s, covering the 100 to 1000 Dalton mass range. The source temperature was 120 °C and the desolvation temperature was set at 450 °C. Nitrogen was used as the nebulisation gas at a flow rate of 700 L/h. The raw UPLC-MS data was extracted and analysed using Mass Lynx XS software (Waters Corporation, Manchester, UK).

The CGA were profiled by collecting MS/MS (typical MS^2) data of the masses of interest. CQA were monitored at *m/z* 353, pCoQA at *m/z* 337, FQA at *m/z* 367, and dicaffeoylquinic acids (*di*CQA) and CQA-glycosides at *m/z* 515. Accordingly, the optimization process made use of authentic standards: 3CQA, 4CQA, 5CQA, 3,5*di*CQA, 3,4*di*CQA and 4,5*di*CQA (Phytolab, Germany). Moreover, extracts of alcoholic cider and coffee, commodities in which the CGA have been thoroughly characterized by LC-ion trap MS (Clifford, 2003; Clifford et al., 2003, 2007; Jaiswal and Kuhnert, 2010; Kuhnert et al., 2012; Hussain et al., 2014) were also used to serve as excellent surrogate standards for the *trans* 3-acyl, 4-acyl and 5-acyl *p*CoQA, CQA, FQA and *di*CQA which they contain. This helped to validate the developed LC-ion trap MS method.

To achieve the fragmentation pattern of the various compounds reported in Clifford et al. (2003), the trap collision energy was increased (3–60 eV) to afford a stable fragmentation pattern and was optimized for each compound class. To further enable distinction between the CQA and *di*CQA, the cone voltage was raised (10–100 V) to enable the formation of the following stable ions: Q1[quinic acid-H]− at *m/z*

Table 1
Summary of the *m/z* ions (quasi-molecular ion peak, C_1, Q_1, C_2, Q_2 and others) used for the identification of CGAs in *M. ovalifolia* by UPLC-qTOF-MS/MS.

Molecule no.	Molecule identity	Rt (min)	Precursor ion (*m/z*)	Product ions (*m/z*)
1	*trans* 3-pCoQA	5.67	337.09	163.04;119.04
2	*trans* 4-pCoQA	7.96	337.09	173.04; 119.04
3	*trans* 5-pCoQA	7.62	337.09	191.05; 119.04
4	*trans* 3-CQA	4.61	353.09	191.05; 179.05; 135.04
5	*trans* 4-CQA	6.10	353.09	191.05; 135.04
6	*trans* 5-CQA	5.78	353.09	191.05; 135.04
7	*trans* 3-FQA	6.63	367.09	193.04
8	*trans* 4-FQA	9.55	367.09	173.04
9	*trans* 5-FQA	9.28	367.09	191.05
10	*cis* 3-pCoQA	5.54	337.09	163.04; 199.04
11	*cis* 4-pCoQA	7.26	337.09	173.04; 119.04
12	*cis* 5-pCoQA	10.58	337.09	191.05; 119.04
13	*cis* 3-CQA	4.58	353.09	191.05; 179.05; 135.04
14	*cis* 4-CQA	5.64	353.09	191.05; 179.05; 135.04
15	*cis* 5-CQA	8.32	353.09	191.05; 135.04
16	*cis* 3-FQA	6.63	367.09	193.04
17	*cis* 4-FQA	8.78	367.09	173.04
18	*cis* 5-FQA	12.49	367.09	191.05
19	3,5 *di-CQA*	6.30	515.13	353.08; 191.05; 135.04
20	3-CQA-glycoside	4.78	515.14	353.08; 179.02; 341.07; 135.04
21	4-CQA-glycoside isomer 1	5.18	515.12	353.08; 341.08; 179.02; 173.04; 135.04
22	4-CQA-glycoside isomer 2	5.47	515.12	353.08; 341.08; 323.06; 179.02; 173.04; 135.04

191,C1[caffeic acid-H]− at *m/z* 179,Q2[quinic acid-H-H2O]− at *m/z* 173, and C2 [caffeic acid-H-CO2]− at *m/z* 135 by in-source collision induced dissociation (ISCID). This was followed by MS/MS fragmentation of the product ions (Table 1), to simulate a typical MS^3 experiment (secondary and tertiary fragment).

2.6. UV irradiation

To distinguish between the *trans* and *cis* isomers, approximately 1 mL of extracts and alcoholic cider were UV-irradiated for 30 min under a shortwave UV lamp (245 nm). It has been reported by Clifford et al. (2008) that *trans*-CGA undergoes isomerization reaction leading to the formation of *cis* isomers when exposed to UV. This can be monitored by comparing the extracts before and after UV exposure. The peaks that increase in intensity after UV irradiation are regarded as *cis* isomers Clifford et al. (2008). When available, some authentic standards were also analysed under similar analytical conditions as the plant extracts. The alcoholic cider also gave access to authentic pCoQA for which standards are not easily available.

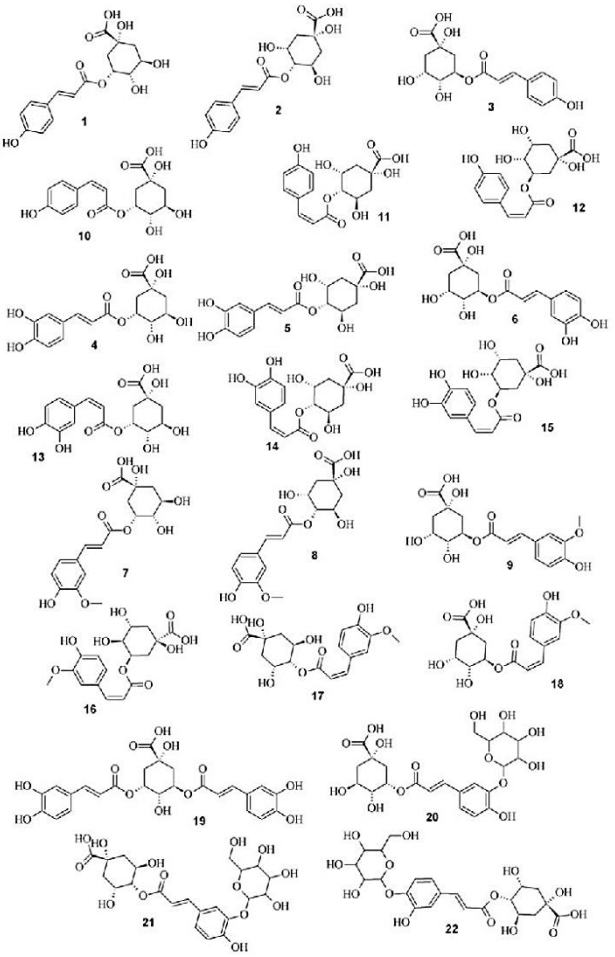


Fig. 1. Structures of different forms of mono-acylated chlorogenic acids (**1–18**), *di*-acylated chlorogenic acid (**19**) and caffeoylquinic acid glycosides (**20–22**) (IUPAC numbering) detected in *M. ovalifolia* leaves.

3. Results and discussion

3.1. Development of the ISCID method

CGA molecules, although naturally found in plants, occur as a complex mixture of regional and geometric isomers (Fig. 1). Identification of these compounds are analytically challenging due to the nature of their structural forms making certain authentic standards unavailable. Researchers, however, have developed a more reliable and validated LC trap-based (IT) MS method for profiling and annotation of CGA molecules (Clifford et al., 2008; Hussain et al., 2014). In the absence of an IT-MS, it has been necessary to develop an alternative MS method and here it is shown that the analysis of CGA and assignments to regional isomer level is possible with the aid of in-source collision-induced dissociation (ISCID) approach of UPLC-qTOF-MS (Madala et al., 2014). In the current study, initial profiling of *M. ovalifolia* extracts using qTOF-MS indicated the presence of, on average, 2 isomers of *p*CoQA, CQA, FQA and *di*CQA as shown in Figs. 2–5.

The UPLC and ISCID-MS parameters were adjusted systematically until a perfect chromatographic separation (Figs. 2 & 4) and fragmentation patterns (Figs. 3 and 5) obtained corresponded to those seen at MS^2 and MS^3 in the ion trap MS. Thus, allowing the unknown CGA in *M. ovalifolia* extracts to be identified not only by their chromatographic retention and UV spectrum but also by the established hierarchical keys that utilize the molecular ions (*m/z* 337, *m/z* 353, *m/z* 367 and *m/z* 515) and diagnostic MS^2 fragmentation (*m/z* 163, *m/z* 173, *m/z* 179, *m/z* 191, *m/z* 193) (Figs. 3 and 5; Table 1) (Clifford et al., 2003; Hussain et al., 2014).

3.2. Identification of CGAs in samples

Here, *M. ovalifolia* exhibited a very rich content of CGA molecules (Fig. 1) which includes the 3-acyl, 4-acyl and 5-acyl regional isomers of *p*CoQA, CQA and FQA (Fig. 5; Table 1). It is clear from UPLC chromatograms (Fig. 2) that for each regional isomer, there are two peaks (with the isobaric *pseudo*-molecular mass and fragmentation patterns)

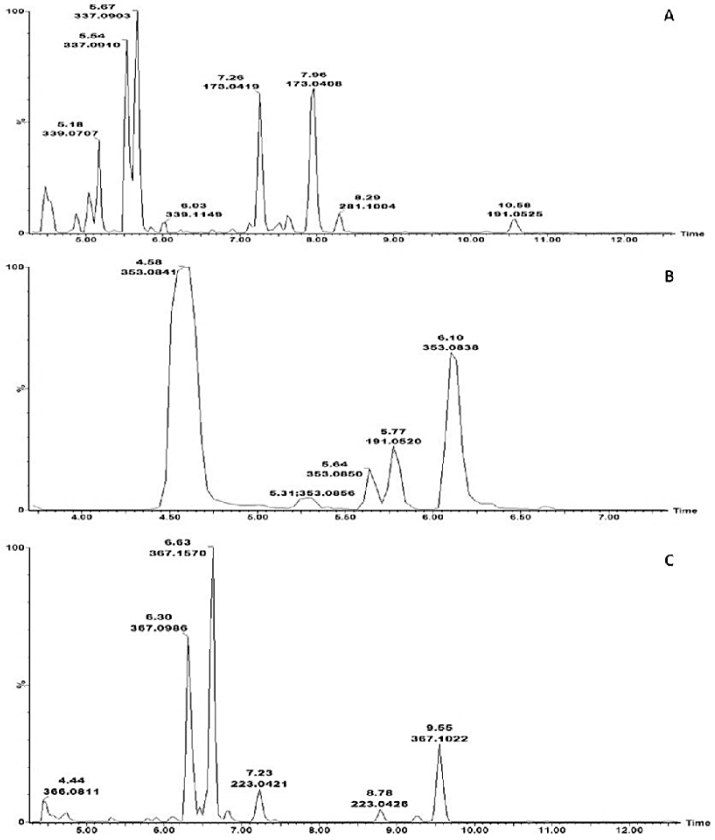


Fig. 2. Representative single ion chromatogram (SIC) of UPLC-MS/MS showing different fragmentation pattern of the regional isomers of pCoQA (A), CQA (B) and FQA (C).

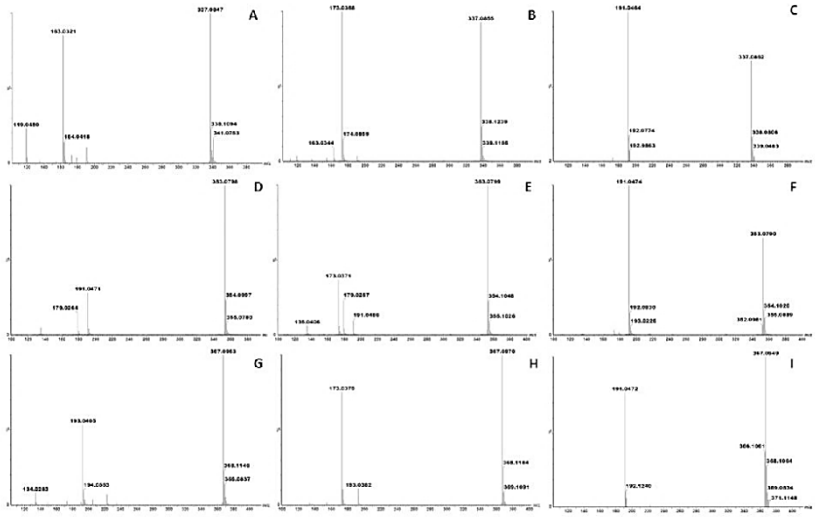


Fig. 3. Representative UPLC-MS/MS spectra showing fragmentation pattern of the 3-pCoQA (A), 4-pCoQA (B), 5-pCoQA (C), 3-CQA (D), 4-CQA (E), 5-CQA (F), 3-FQA (G), 4-FQA (H) and 5-FQA (I).

suggesting that *cis* and *trans* isomers are present in each case (Clifford et al., 2008; Makola et al., 2016). For the sake of simplicity, only the MS spectra of *trans* isomers were shown (Fig. 2).

For the characterization of mono-acyl CGA, the different regional isomers yielded the following fragmentation patterns: *trans* pCoQA and *cis* pCoQA exhibited a precursor ion at *m/z* 337 $[M–H]^-$ and product ions at *m/z* 163 (coumaric acid); *m/z* 119 (decarboxylated coumaric acid); *m/z* 173 (dehydrated quinic acid) and *m/z* 191 (quinic acid) ions (Fig. 3; Table 1). The CQA isomers exhibited a precursor ion at *m/z* 353 $[M–H]^-$ with product ions at *m/z* 191, *m/z* 173, *m/z* 179 (caffeic acid); *m/z* 135 (decarboxylated caffeic acid) ions. FQA *cis* and *trans* isomers exhibited a precursor ion at *m/z* 367 $[M–H]^-$ with product ions at *m/z* 193 (ferulic acid); *m/z* 173, *m/z* 134 (decarboxylated ferulic acid) and *m/z* 191 (Fig. 3; Table 1).

Interestingly, four peaks were detected in the extracts of *M. ovalifolia* with a precursor ion at *m/z* 515 $[M–H]^-$ (Fig. 4), an ion characteristic of *di*CQA and CQA-glycosides. One of these peaks was identified as 3,5-*di*CQA **(19)** by comparison to authentic standard (Fig. 5; Table 1). The other three isobaric peaks yielded a fragment ion at *m/z* 341, an ion not produced by di-CQA but characteristic of CQA-glycosides, and corresponding to the [caffeoyl-glycoside] fragment. One of the peaks exhibited an intense fragment at *m/z* 179 and *m/z* 341 and it was identified as 3-CQA-glycoside (**20**). The remaining two ions yielded fragment ions at *m/z* 353 and *m/z* 173 confirming that these are glycosides of 4-CQA (Figs. 4 and 5).

Interestingly, apart from the *m/z* 515 peak, it has been shown recently that MS fragmentation of these CQA-glycosides gives rise to ions at *m/z* 341. In addition, an equally intense peak at *m/z* 323 ([caffeoyl

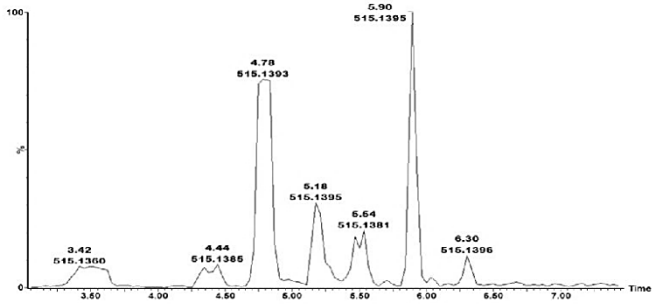


Fig. 4. Representative single ion monitoring chromatogram (SIC) of UPLC-MS/MS showing different separation or elution of *di*CQA and CQA-glycosides both isobaric precursor ions at *m/z* 515.

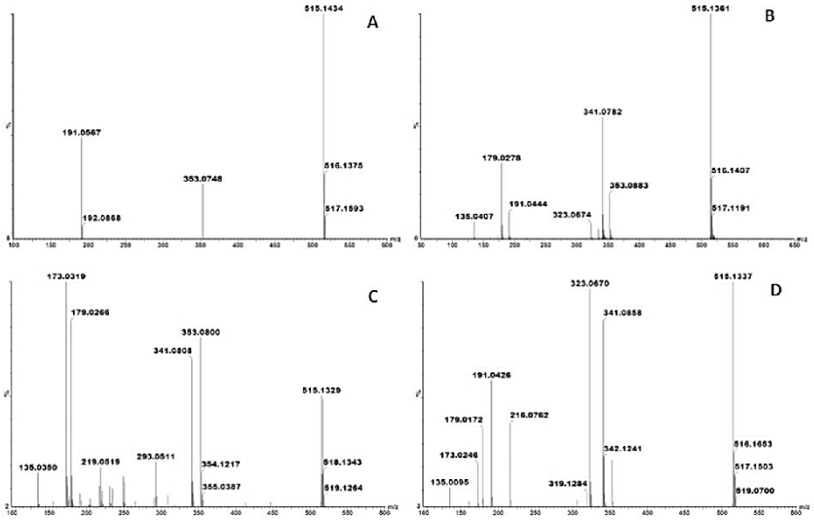


Fig. 5. Representative UPLC-MS/MS spectra showing the fragmentation patterns of 3,5-*di*CQA (A), 3-CQA-glycoside (B), 4-CQA-glycoside isomers 1 (C) and 4-CQA-glycoside isomer 2 (D).

glucoside–H–H_2O]−) is a characteristic of glucosyl attachment at C-3 (Jaiswal et al., 2014). As such, in the current study, these molecules were putatively identified as 4-O-(3′-O-caffeoyl glucosyl) quinic acid (**21**) and 4-O-(4′-O-caffeoyl glucosyl) quinic acid (**22**) since it also produced similar fragmentation to 4-CQA (Fig. 5). Similar CQA-glycosides have been reported in *Chrysanthemum* spp. (Clifford et al., 2007). In contrast to previously published data (IUPAC, 1976), it has been observed that there is a small but significant difference in the fragmentation of *cis*-3-pCoQA compared with its *trans* counterpart. The MS^2 fragment ions at *m*/*z* 163 and *m*/*z* 191 are of almost identical intensity for the *trans* isomer, but in the equivalent spectrum for the *cis* isomer *m*/*z* 163 has over fivefold the intensity of *m*/*z* 191. The reason for this distinctive behaviour is uncertain. However, there were no significant differences in the behaviour of the other eight pairs of geometric isomers investigated.

Although the *di*CQA, *p*CoQA and FQA, and the *cis* isomers were easily detected by the mass spectrometer, they were minor components that were not sufficiently well resolved in the UV traces to enable quantification. The presence of the geometrical isomers was confirmed by subjecting the surrogate standards to UV irradiation thus generating *cis* isomers from the *trans* isomers that predominate in these extracts. Furthermore, *di*-CQA in *M. ovalifolia* as 3,5-*di*CQA was identified. This could be the first report of *di*-CQA in *Moringa* spp. The associated *mono*- and *di*-*cis* isomers were not detected and it is possible they were co-eluted if ever present. Clifford et al. (2008) detected only two of the three possible *cis*-isomers of 3,5-*di*CQA and these had capacity factors of only 1.02 and 1.04 relative to the *di*-*trans* isomer. The presence of these *cis* geometrical isomers poses an undisputed analytical challenge during MS identification (Makola et al., 2016) and, as such, the current study is expected to provide guidelines for future identification of these compounds in other plants.

In keeping with previously published data, the study found out that *cis*-5-acyl CGA elute from reversed phase column packings appreciably later than their *trans* counterparts whereas the *cis*-3-acyl CGA and *cis*-4-acyl CGA elute a little earlier than their *trans* counterparts (Clifford, 2003; Clifford et al., 2008). Compared with the *trans* isomers, Clifford et al. (2008) reported the relative capacity factors as 1.27, 1.44 and 1.13 for *cis*-5-*p*CoQA, *cis*-5-CQA and *cis*-5-FQA, respectively. This was attributed to the ability of the more hydrophobic isomer to form an internal hydrogen bond that was not formed in the more hydrophilic isomer as indicated by modelling at the MM-2 level (Clifford et al., 2008).

Plants producing CGA generally contain a relatively large number of related compounds differing in the identity of the cinnamate(s) along with positional and geometric isomers for which authentic pure standards are not available. Accordingly, their identification presents a challenge, albeit one that has become easier with the development of LC-ion trap MS methods and associated hierarchical schemes for data interpretation (Clifford et al., 2003, 2005; Hussain et al., 2014). However, it seems likely that the full complement of CGA in *Moringa* species has not yet been determined. In this study, the use of in-source collision-induced dissociation (ISCID) to develop a UPLC-qTOF-MS method for the profiling of the CGA in the leaves of *M. ovalifolia* is reported. The study is very important as it contributes to the understanding of various *Moringa* species in as far as chemical composition is concerned especially that it is consumed in many countries as a herbal supplement. Moreover, we recently reported almost all the identified isomers to be antioxidative components of *M. oleifera* (Ramabulana et al., 2015), which can further be exploited for the management of various human diseases associated with oxidative stress.

4. Conclusion

The results of the current study revealed that *M. ovalifolia* contains *cis* and *trans* 3-acyl, 4-acyl and 5-acyl *p*CoQA, CQA and FQA, 3-CQA-glycosides, 4-CQA-glycosides, and at least one isomer of 3,5-*di*CQA. This offers a step towards the validation of the use of the plant species in managing medical and nutritional conditions as the compounds reported here have been previously reported to exert a number of biological activities including being antioxidants. The study has also further

demonstrated the capabilities of UPLC-MS as a powerful technique to characterize both geometrical and regional isomers from complex plant extracts. Although substantial data on the human absorption and metabolism of the *trans* CGA have been reviewed, there have been no explicit studies of the corresponding *cis* isomers or the CQA-glycosides. These derivatives are not commonly featured in the diet but would be a consequence of the use of *M. ovalifolia* and deserve further study. Finally, the presence of both positional and geometrical isomers of CGAs in this plant could be an evolutionary strategy by *Moringa* plants for self-protection against various biotic and abiotic stressors. As such, the large contingency of these compounds can also be exploited for covering a wider pharmacological space.

Acknowledgements

The authors would like to thank the following for financial support: National Research Foundation–Namibian Bilateral Collaborations Grant, Department of Science and Technology, and University of Johannesburg. The funding for the UPLC-TOF-MS instrument was obtained from the South African NEP/RISP programme. Prof. Mike Clifford from the University of Surrey is thanked for his help with chlorogenic acids identification.

References

Bennett, R.N., Mellon, F.A., Foidl, N., Pratt, J.H., Dupont, M.S., Perkins, L., Kroon, P.A., 2003. Profiling glucosinolates and phenolics in vegetative and reproductive tissues of the multi-purpose trees *Moringa oleifera* L. (horseradish tree) and *Moringa stenopetala* L. Journal of Agricultural and Food Chemistry 51, 3546–3553.

Clifford, M.N., 1999. Chlorogenic acids and other cinnamates—nature, occurrence and dietary burden. Journal of the Science of Food and Agriculture 79, 362–372.

Clifford, M.N., 2000. Chlorogenic acids and other cinnamates—nature, occurrence, dietary burden, absorption and metabolism. Journal of the Science of Food and Agriculture 80, 1033–1042.

Clifford, M.N., 2003. The Analysis and Characterization of Chlorogenic Acids and Other Cinnamates. In: Santos-Buelga, C., Williamson, G. (Eds.), Methods in Polyphenol Analysis. Royal Society of Chemistry, Cambridge, pp. 314–337.

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

Clifford, M.N., Knight, S., Kuhnert, N., 2005. Discriminating between the six isomers of dicaffeoylquinic acid by LC-MS^n. Journal of Agricultural and Food Chemistry 53, 3821–3832.

Clifford, M.N., Wu, W., Kirkpatrick, J., Kuhnert, N., 2007. Profiling the chlorogenic acids and other caffeic acid derivatives of herbal *Chrysanthemum* by LC-MS^n. Journal of Agricultural and Food Chemistry 55, 929–936.

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

Eliel, E.L., Ramirez, M.B., 1997. (-)-Quinic acid: configurational (stereochemical) descriptors. Tetrahedron-Asymmetry 8, 3551–3554.

Hussain, S., Malik, F., Mahmood, S., 2014. Review: an exposition of medicinal preponderance of *Moringa oleifera*. Pakistan Journal of Pharmaceutical Sciences 27, 397–403.

IUPAC, 1976. Commission on the Nomenclature of Organic Chemistry (CNOC) and IUPACIUB Commision on Biochemical Nomenclature (CBN): nomenclature of cyclitols. Biochemical Journal 153, 23–31.

Jaiswal, R., Kuhnert, N., 2010. Hierarchical scheme for liquid chromatography/multi-stage spectrometric identification of 3,4,5-triacyl chlorogenic acids in green robusta coffee beans. Rapid Communications in Mass Spectrometry 24, 2283–2294.

Jaiswal, R., Muller, H., Muller, A., Karar, M.G.E., Kuhnert, N., 2014. Identification and characterisation of chlorogenic acids, chlorogenic acid glycosides and flavonoids from *Lonicera henryi* L. (Caprifoliaceae) leaves by LC-MS ^{n}. Phytochemistry 108, 252–263.

Kashiwada, Y., Ahmed, F.A., Kurimoto, S., Kim, S.Y., Shibata, H., Fujioka, T., Takaishi, Y., 2012. New alpha-glucosides of caffeoyl quinic acid from the leaves of *Moringa oleifera* Lam. Journal of Natural Medicine 66, 217–221.

Khoza, B.S., Gbashi, S., Steenkamp, P.A., Njobeh, P.B., Madala., N.E., 2016. Identification of hydroxylcinnamoyl tartaric acid esters in *Bidens pilosa* by UPLC-tandem mass spectrometry. South African Journal of Botany 103, 95–100.

Kuhnert, N., Karaköse, H., Jaiswal, R., 2012. In: Nollet, L.M.L., Toldra, F. (Eds.), Analysis of Chlorogenic Acids and Other Hydroxycinnamates in Food, Plants and Pharmacokinetic Studies. CRC Press, United States.

Madala, N.E., Tugizimana, F., Steenkamp, P.A., 2014. Development and optimization of an UPLC-QTOF-MS/MS method based on an in-source collision induced dissociation approach for comprehensive discrimination of chlorogenic acids isomers from *Momordica* plant species. Journal of Analytical Methods in Chemistry 2014.

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.

Makola, M.M., Steenkamp, P.A., Dubery, I.A., Kabanda, M.M., Madala, N.E., 2016. 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.

Mhlongo, M.I., Piater, A.P., Madala, N.E., Steenkamp, P.A., Dubery, I.A., 2016. Phenylpropanoid defences in *Nicotiana tabacum* cells: overlapping metabolomes indicate common aspects to priming responses induced by lipopolysaccharides, chitosan and flagellin-22. PLoS One 11.

Ncube, E.N., Mhlongo, M.I., Piater, L.A., Steenkamp, P.A., Dubery, I.A., Madala, E.N., 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 Century Journal 8, 1.

Ndong, M., Uehara, M., Katsumata, S., Suzuki, K., 2007. Effects of oral administration of *Moringa oleifera* Lam. on glucose tolerance in Goto–Kakizaki and Wistar rats. Journal of Clinical Biochemistry and Nutrition 40, 229–233.

Pakade, V., Cukrowska, E., Chimuka, L., 2013. Comparison of antioxidant activity of *Moringa oleifera* and selected vegetables in South Africa. South African Journal of Science 109, 17–21.

Ramabulana, T., Mavunda, R.D., Steenkamp, P.A., Piater, L.A., Dubery, I.A., Madala, N.E., 2015. Secondary metabolite perturbations in *Phaseolus vulgaris* leaves due to gamma radiation. Plant Physiology and Biochemistry 97, 287–295.

Siddhuraju, P., Becker, K., 2003. Antioxidant properties of various solvent extracts of total phenolic constituents from three different agroclimatic origins of drumstick tree (*Moringa oleifera* Lam.) leaves. Journal of Agricultural and Food Chemistry 51, 2144–2155.

Verma, A.R., Vijayakumar, M., Mathela, C.S., Rao, C.V., 2009. *In vitro* and *in vivo* antioxidant properties of different fractions of *Moringa oleifera* leaves. Food and Chemical Toxicology 49, 2196–2201.

Vongsak, B., Gritsanapan, W., Wongkrajang, Y., Jantan, I., 2013a. 2013b. In vitro inhibitory effects of *Moringa oleifera* leaf extract and its major components on chemiluminescence and chemotactic activity of phagocytes. Natural Product Communications 8, 1559–1561.

Vongsak, B., Sithisarn, P., Gritsanapan, W., 2013b. Simultaneous determination of cryptochlorogenic acid, isoquercetin, and astragalin contents in *Moringa oleifera* leaf extracts by TLC-densitometric method. Evidence Based Complementary Alternative Medicine 2013, 917609.

Vongsak, B., Sithisarn, P., Gritsanapan, W., 2013c. Simultaneous HPLC quantitative analysis of active compounds in leaves of *Moringa oleifera* Lam. Journal of Chromatographic Science 2013, 917609.

Paper III

This paper titled “Rutinoside-bearing flavonoids are chemo-taxonomical markers for efficient identification of pharmacologically potent *Moringa oleifera* Lam. Cultivars,” was submitted to Frontiers in Pharmacology. The presented format is for that journal. It investigates the distribution of three sought after rutinoside bearing flavonoids in twelve *Moringa* cultivars for pharmacological and nutritional purpose.

Rutinoside-bearing flavonoids are chemo-taxonomical markers for efficient identification of pharmacologically potent *Moringa oleifera* Lam. Cultivars

Charlene Makita[1], Ewa Cukrowska[1], Ashwell Rungano Ndhlala[2], Hafiz Abdelgadir[2], Luke Chimuka[1], Paul Steenkamp[3,4], Ntakadzeni Edwin Madala[3*].

[1]Molecular Science Institute, School of Chemistry, University of Witwatersrand, P/Bag 3, Braamfontein 2050, Johannesburg, South Africa.

[2]Agricultural Research Council, Vegetable and Ornamental Plants (VOP), Private Bag X293, Pretoria 0001, South Africa

[3]Department of Biochemistry, University of Johannesburg, P. O. Box 524, Auckland Park, 2006, Johannesburg, South Africa.

[4]CSIR Biosciences, Natural Products and Agro-processing Group, Building 20, Pretoria, 0001, South Africa.

*Correspondence: emadala@uj.ac.za; Tel.: +27-115-594-573

Abstract: Recent evidence has shown that consumption of plants rich in phenolic compounds has health-promoting benefits. Amongst these plants, include *Moringa oleifera* Lam. *(Moringacea*e) which has gained significant attention due to its high contingency of secondary metabolites. It has been also shown that the quantity of nutritional and health promoting secondary metabolites in *M. oleifera* is cultivar specific. As such, in the current study, the distribution of three highly sought after flavonoid (quercetin rutinoside (rutin), kaempferol rutinoside and isorhamnetin rutinoside) was investigated using untargeted metabolomic approach with the aid of UPLC-qTOF-MS. The results of this pilot study, suggest only three out of the twelve cultivars investigated were found to contain these flavonoid molecules in significant amounts. The current results, therefore, suggest rutinoside-bearing

flavonoids as a potential marker for taxonomical grouping of different cultivars of *M. oleifera* which can definitely help with selection of pharmacologically potent cultivars.

Keywords: *Moringa oleifera*, UPLC-qTOF-MS, chemotaxonomy, rutinoside, pharmacology.

1.0 Introduction

Moringa oleifera is known to be an essential commodity in countries such as India, Pakistan, Philippines, Hawaii and many African countries (Anwar *et al*., 2007). For many people, this plant has been considered a food supplement due to its high content of nutritional chemicals (Rajangam *et al.,* 2001). Apart from the nutritional content it contains, it has numerous biological activities and operates as a cardiac stimulant, aphrodisiac, diuretic and possesses cholagogic properties (Pakade *et al*., 2012; Devaraj *et al*., 2016). The chemical constituents of this plant have been shown to include other medicinal related compounds such as saponins, flavonoids, cardiac glycosides and alkaloids (Stohs and Hartman, 2015).

During the last decade, there has been an increase in the cultivation of *Moringa* trees due to its high biomass under intense farming conditions (Makkar and Becker, 1996). Plant breeding does not only improve the agronomic and medicinal traits of plants, but techniques such as marker-assisted systems can be used as a useful tool to aid in hybrid selection (Canter *et al.,* 2005). It has been shown that *M. oleifera* cultivars differ from each other and show diverse differences in growth performance, leaf-mass production and secondary plant metabolite content (Förster *et al*., 2015). The use of plant metabolomics has made it possible to analyse the metabolite composition between plant species and cultivars, therefore, giving us a broad perspective of the phenotypic and genotypic differences amongst plant varieties as a result of gene expression (Farag *et al*., 2014). Metabolomic approaches in combination with chemometric data tools such as multivariate data analysis (MVDA) models allows key biological information to be interpreted from complex data sets (Kwon *et al*., 2014; Madala *et al*., 2014).

Apart from the genetic makeup, cultivation practises were also shown to have the most significant influence in growth performance, leaf-mass production and secondary plant metabolite content (Förster *et al*., 2015). Only recently has the differences in the composition of secondary plant metabolites amongst cultivars been the basis for cultivation of cultivars, which this present study addresses (Ferreira *et al.,* 2008). Our previous studies on *Moringa* revealed that not all *Moringa* species share similar metabolites (Makita *et al*., 2016). The results further showed for the first time the importance of sugar attachment for taxonomical grouping of related species and that the differences between the species are likely due to genetic differences rather than environmental variations (Makita *et al*., 2016).

Therefore, in this current study, an investigation into the distribution patterns of highly sought-after, rutinoside-bearing flavonoids in twelve cultivars of *M. oleifera* was performed using UPLC-qTOF-MS. A comparative study using complementary multivariate statistical analysis platforms such as principal component analysis (PCA), hierarchal clustering analysis (HCA) and box-and-whiskers plots were utilised. This revealed underlying differences of the metabolic profile, relationship and health promoting attributes amongst twelve diverse *M. oleifera* cultivars. Results from this study thus provided new knowledge about the bioactive compounds found in these cultivars and probed what could be seen as possible superior and optimal cultivars among agronomists and farmers.

2. Materials and Methods

2.1. Plant collection

Seeds of the twelve *M. oleifera* Lam. cultivars were obtained from different geographical locations around the world [through the World Vegetable Centre (AVRDC)] and grown at the Agricultural Research Council (ARC) (Roodeplaat, South Africa). Different cultivar details and where they were originally obtained is shown below.

a) Cultivars TOT4893, TOT4951, TOT4977, TOT5028, TOT5077, TOT5169, TOT5330 and TOT7266 originated from Thailand
b) TOT4100 originated from Taiwan
c) TOT4800 originated from USA

d) SH originated from Silver Hill, South Africa
e) CHM originated from Limpopo province, South Africa

Briefly, ground leaf powder (2 g) was extracted with 20 mL of 80% aqueous methanol (MeOH) with sonication for 30 min. The mixture was centrifuged for 15 minutes at 25ºC at a high speed of 5000xg with a swinging –bucket to remove any excess tissue debris. Thereafter, the liquid supernatant was transferred to a round bottom flask and reduced to approximately 1 mL with the aid of a Büchi rotary evaporator at 65ºC. The resulted extract was then transferred to a 2 mL Eppendorf tube and further dried to completeness under vacuum at 45ºC. The dried extracts were reconstituted in 1 mL of 50% MeOH and finally filtered through 0.22 µm nylon filters. The resulting extracts were then stored in a freezer at -20 ºC so as to avoid degradation until they were analysed on the UPLC- qTOF-MS.

2.2. Metabolite separation and analysis

Metabolite separation and detection was achieved using an Acquity UPLC system hyphenated to SYNAPT G1 HDMS mass spectrometer (Waters, MA, USA), of which, three repeats of the extracts from each cultivar was analysed. The chromotagraphic separation of the *M. oleifera* extracts was accomplished on a Waters BEH C_{18} reverse phase column (150 mm x 2.1 mm, 1.7 µm) (Waters, MA, USA) and the column temperature controlled at 60°C. The mobile phase consisted of 0.1% formic acid in deionized water (Eluent A) and 0.1% formic acid in acetonitrile (Eluent B). The gradient was programmed as follows: the initial conditions were 98% eluent A followed by multiple gradients to 5% at 26 minutes. The conditions were kept constant for 1 minute and then changed to the initial conditions. The injected volume of each extract was 1 µL at a constant flow rate of 0.4 mL/min which was maintained for 1 minute. The chromatographic method was a total run time of 30 minutes. Chromatographic separation was monitored using a photodiode-array detector (PDA), which was set to scan 200-500 nm and collecting 20 spectra per second, was coupled in tandem to an electrospray ionization mass spectrometer (ESI–MS) detector.

MS detection parameters were applied with the following parameters: an electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-qTOF-

MS) detector was used and Leucine enkephalin (556.227/554.2615 Da) was used as reference calibrant at a constant flow rate of flow rate of 0.1 mL/min to obtain typical mass accuracies of at least 3 ppm. The mass spectrometer was operated in both negative and positive ionization modes with a capillary voltage of 2.5 kV, the sampling cone at 30 V, and the extraction cone at 4 V, multichannel plate detector potential at 1600 V, source temperature at 120 °C, desolvation temperature at 450°C, cone gas flow at 50 L/h, and desolvation gas flow at 550 L/h. The scan time was 0.2 seconds and an interscan delay of 0.02 s covering the 100 to 1000 Dalton mass range. For identification purposes, different fragmentation experiments were achieved by alteration of the trap collision energy levels to mimic MS^E experiments. MassLynx ^{TM} software (Waters Corporation, MA, USA) were used to visualise the raw data from data matrix generated from UHPLC-qTOF-MS for accurate statistical modelling.

2.3 Multi-variate data analysis (MVDA)

To create a data matrix for multivariate data modelling, the acquired raw data was imported to SIMCA version 13.0 (Soft independent modelling of class analogy) software (Umetrics Corporation, Umea, Sweden). Initially, the UPLC-MS data was exported and analysed with the aid of the MarkerLynx XS™ software (Waters, Manchester, UK). The following parameters were used for maximum data output: The following parameters were chosen: retention time (Rt) of 1–27 min, mass range of 100–1000 Da, mass tolerance of 0.02 Da and Rt window of 0.2 min. The analyses conditions/parameters were kept constant for both negative and positive data. The dataset obtained from MarkerLynx XS™ was transported to the SIMCA 13.0 software to generate principal component analysis (PCA) and hierarchal clustering analysis (HCA). Pareto scale was utilised for both models. PCA and HCA are utilised as visual tools to determine the differences between samples based on their metabolic profiles. The validity of the generated models were determined by metabolic diagnostic tools such as: cumulative model variation in the matrix X, R2X (cum) which is the goodness-of-fit-parameter the proportion of the variance of the response variable that is explained by the model, R2Y (cum) and predictive

ability parameter (R2Y (cum)), also known as the total variation fraction of matrix X predicted by an extracted components.

2.4. Statistical analysis and metabolite identification

The UHPLC-ESI-MS data was analysed using MarkerLynx ^{TM} software to obtain peak alignment, peak finding, peak integration and retention time (RT) with the following parameters: Rt range of 1-27 min, mass range of 100-1000 Da, mass tolerance of 0.05 D, Rt window of 0.2 min. This highly dimensional data was transferred to Microsoft Excel. The peak area under the curve with respect to the masses (*m/z*) from the twelve *M. oleifera* cultivars was used to create Box-and-Whiskers plots with the aid of SPSS version 22 software (IBM, United States).

3.0 Results and Discussion

3.1 Multivariate data analyses and visual differences of cultivars

There is a strong body of evidence supporting multivariate statistical analysis for identification of primary and secondary metabolites as biochemical markers for use in breeding programs for different types of plants (Arbona *et al*., 2009). The favoured analytic platform for the analysis of polar to non-polar metabolites extracted from methanol is UHPLC-qTOF-MS (Ncube *et al*., 2014). Visual inspection of the stacked base peak intensity (BPI) chromatograms of the 12 *Moringa* cultivars (Fig. 1) show a high degree of instrument reproducibility. The visual inspection of the BPI chromatograms did not show any clear differences between the cultivars, however, through single ion monitoring we noticed clear differences in peak intensities and presence or absence of some of the peaks amongst cultivars shown in Fig. 2.

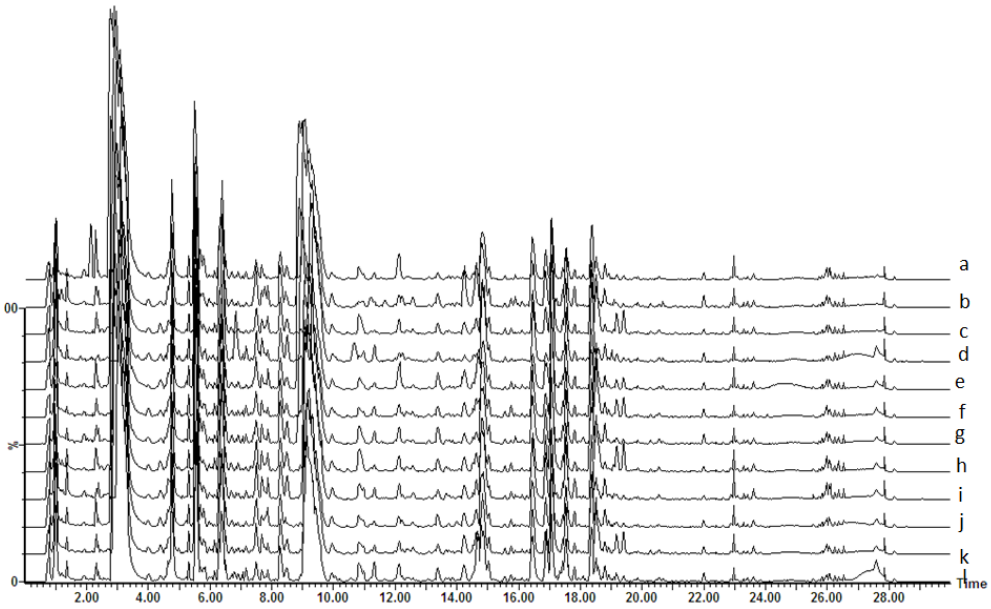


Fig.1: UHPLC-MS-BPI chromatograms (ESI$^-$) of twelve *M. oleifera* cultivars: 4880; b- 5028; c-LIM; d- SH; e- 4977; f-5077; g-5169; h- TOT7266; i- 4100; j-4893; k- TOT5330; l-CHM.

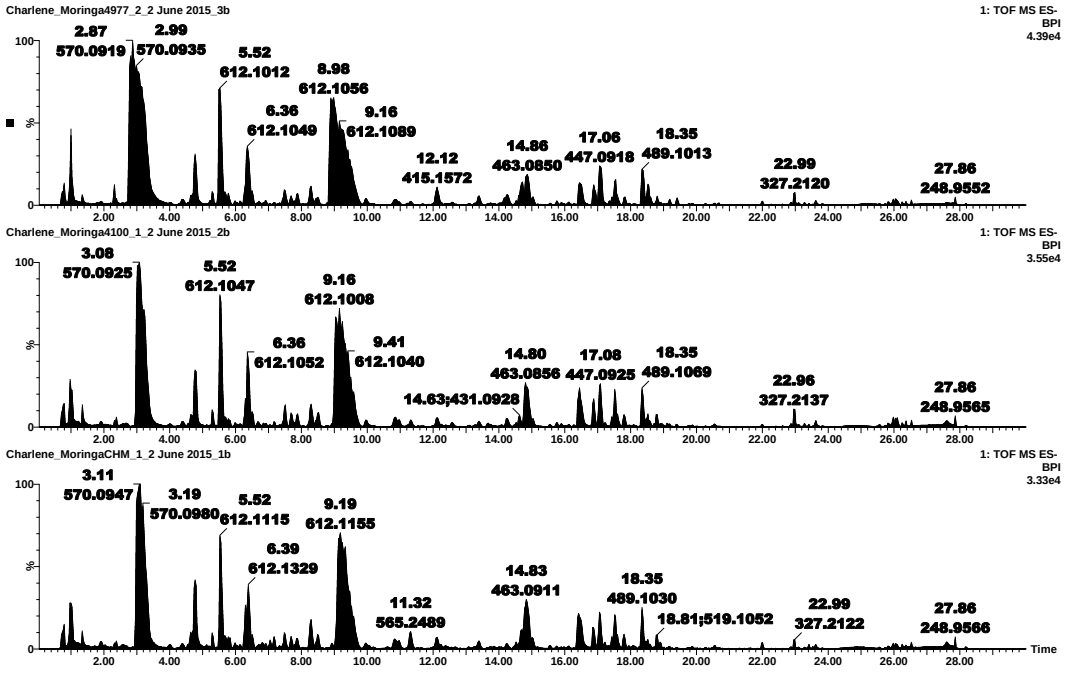


Fig. 2: UHPLC-qTOF-MS chromatographic analyses of cultivars 4977, 4100 and CHM.

ESI negative chromatograms show clear differences in peak intensities between these three representative cultivars. Furthermore, the use of metabolite fingerprinting approach with the aid of UPLC-qTOF-MS revealed potential bio-markers (Fig. 1). Amongst these potential bio-markers, three were of importance since they exhibited an interesting distribution/trend across different cultivars. Upon further investigation through accurate mass, fragmentation pattern and comparison to authentic standards, these three peaks were identified as kaempferol rutinoside, rutin and isorhamnetin rutinoside (Fig. 4, 5 and 6). Rutin eluted at retention time (RT) 14.07 min with precursor ion $[M - H]^-$ at *m/z* 609.1441 and produced fragmentation product ions at *m/z* 300.0217 representing quercetin aglycone (Fig. 5a and 6a) (Hvattum *et al*., 2002; Martucci *et al*., 2014). Another peak at retention time 16.48 min and precursor ion at *m/z* 593.1526 was identified as kaempferol rutinoside. This molecule produced a product ion at *m/z* 285.0364 representing the kaempferol aglycone (Fig. 5b and 6b) (Vagiri *et al*., 2015). The third peak was identified as isorhamnetin rutinoside eluting at retention time 16.98 min and precursor ion $[M-H]^-$ at *m/z* 623.1693 (Gutzeit *et al*., 2007). Similar to the other former two metabolites, the identity of this molecule was further confirmed by the presence of an isorahmanetin aglycone ion at *m/z* 315.0472 (Fig. 5c and 6c). Ironically, even though the ESI negative data was sufficient to carry out positive annotation of these metabolites, ESI positive data was found to be more efficient to deduce the sugar sequence and for strengthening of our metabolite identities (Fig. 5 and 6).

In our recent work, kaempferol rutinoside, rutin and isorhamnetin rutinoside compounds were only found in *M. ovalifolia* (obtained from Namibia) but absent from two *M. oleifera* ecotypes (obtained from South Africa and Namibia) as demonstrated in Paper I. However, the presence of rutin has been previously reported in *M. oleifera* (Amaglo *et al*., 2010), suggesting that it could be a cultivar-specific component. As such, this highlights the importance of determining which *Moringa* cultivars contain these pharmacologically relevant metabolites (Habtemariam and Varghese, 2015; Wang *et al*., 2016; Farag *et al*., 2014; Fujimura *et al*., 2011; Makita *et al*., 2016; Zhang *et al*., 2014; Förster *et al*., 2015; Girennavar *et al*., 2008).

The visual inspection of the BPI chromatograms shows clear differences in peak intensities and presence or absence of some of the peaks (Fig. 1). Such differences point out to the differences in cultivars on a metabolic level. Therefore, to provide comparative interpretations and essential unbiased confirmation of the metabolic differences amongst cultivars, multivariate statistical analysis was utilised to analyse the LC-MS data sets.

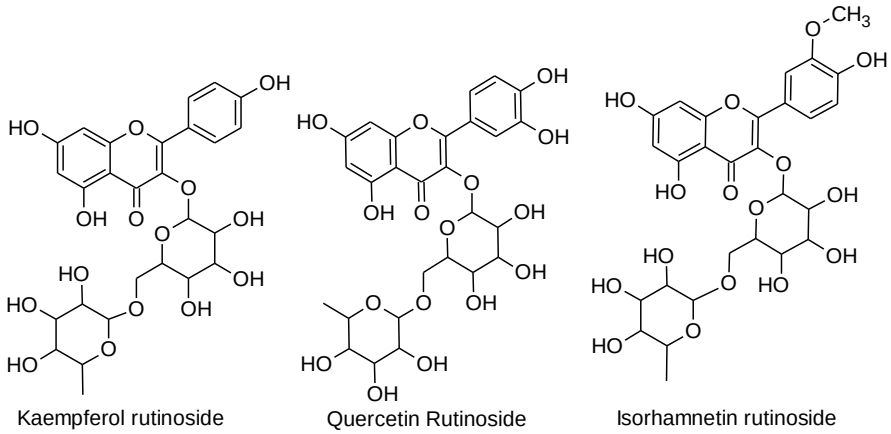


Fig. 3: Structures of kaempferol rutinoside, quercetin rutonoside (rutin) and isorhamnetin rutinoside identifies in the leaves of *M. oleifera* cultivars.

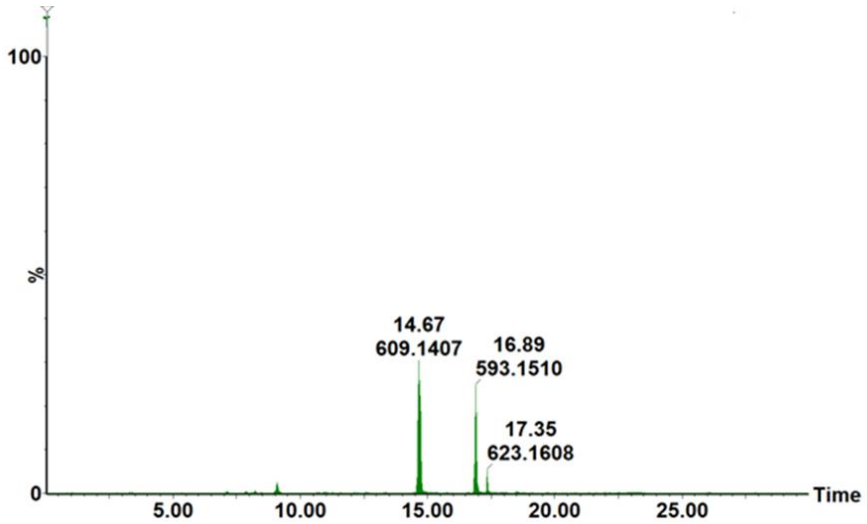


Fig 4: Single ion monitoring chromatogram showing the presence of quercetin rutinoside, kaempferol rutinoside and isorhamnetin rutinoside

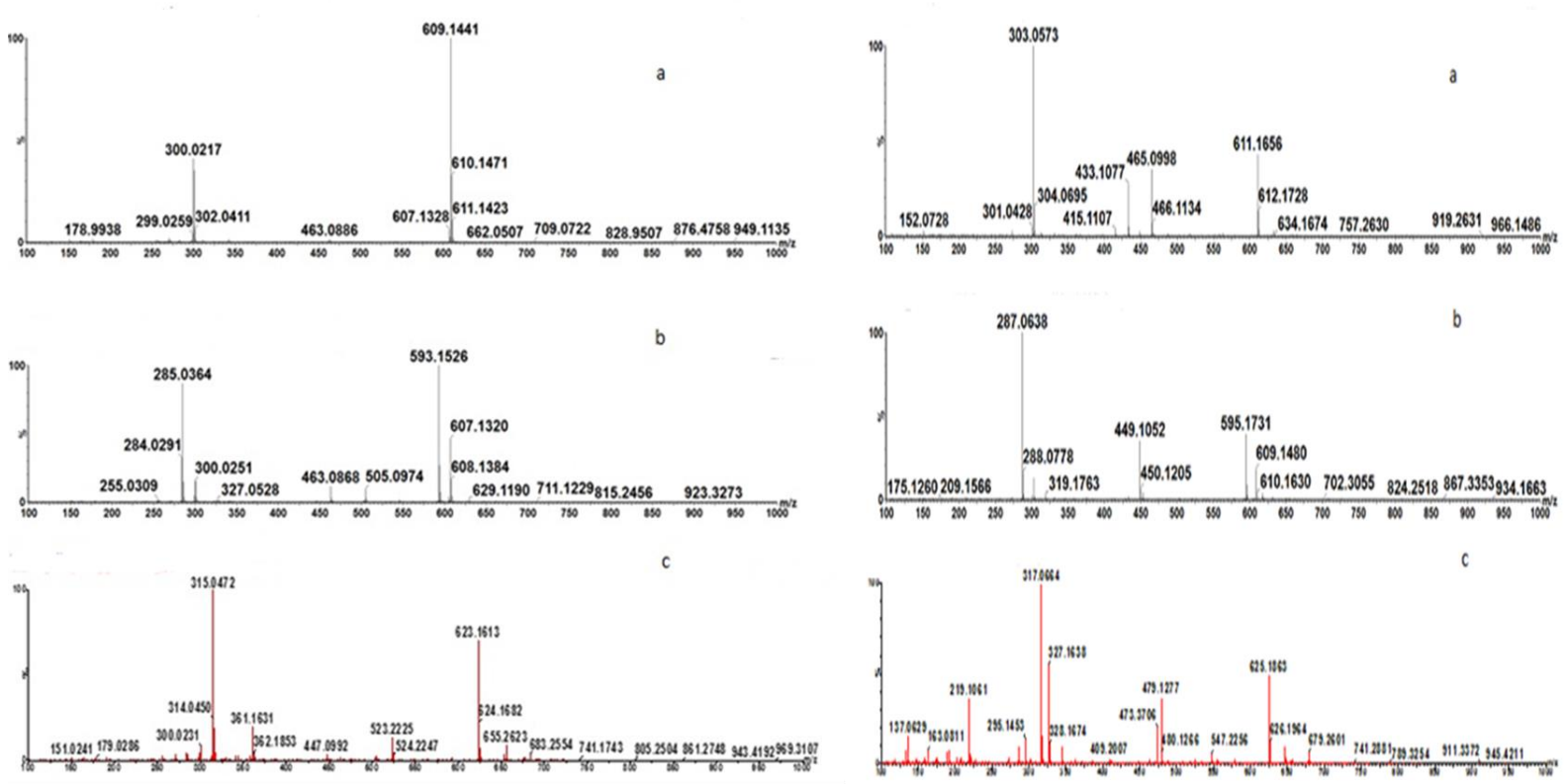


Fig. 5: UPLC-qTOF-MS chromatogram of *M. oleifera* leaf extracts showing (a) quercetin rutinoside (b) kaempferol rutinoside and (c) isorhamnetin rutinoside flavonoid peaks detected in negative ionization mode.

Fig. 6: UPLC-qTOF-MS chromatogram of *M. oleifera* leaf extracts showing (a) quercetin rutinoside (b) kaempferol rutinoside and (c) isorhamnetin rutinoside flavonoid peaks detected in positive ionization mode.

3.2 Multivariate PCA and HCA of UHPLC–qTOF-MS data

To further investigate the differentiation amongst the twelve cultivars, the data was analysed with multivariate data analyses such as PCA and HCA. By definition, PCA is known as an unsupervised model which is used to display clustering and differentiation between or within the sample groupings (Trygg and Holmes, 2007). Furthermore, a chemotaxonomic approach was utilised for classification of the cultivars. In this study, the grouping of the cultivars is of primary importance. The score plot of the twelve cultivars showed PC1versus PC2 explaining 13.2% and 11% of the variation respectively (Fig. 7). The PCA revealed congestion of cultivars as well as differential clustering of samples into two distinct groups, but group 2 still showed subdivision.

Within group 1, tight clusters of cultivars were formed, suggesting that the cultivars could be related. Group 2 showed three sub-divisions although not tightly clustered. From this deduction, we can also conclude that the cultivars grouped in group 2 contain a chemical composition that is different from each other and from the cultivars in group 1 (cultivar TOT4893). With the aid of HCA, we can get a clear picture of how groups relate to one another and how individual plots amongst a group relate to each other such as those observed in the *Vernonia* Schreb plant species where using multivariate analysis approaches such as PCA and HCA segregated ten plant species based on their metabolic profiles (Martucci *et al*., 2014).

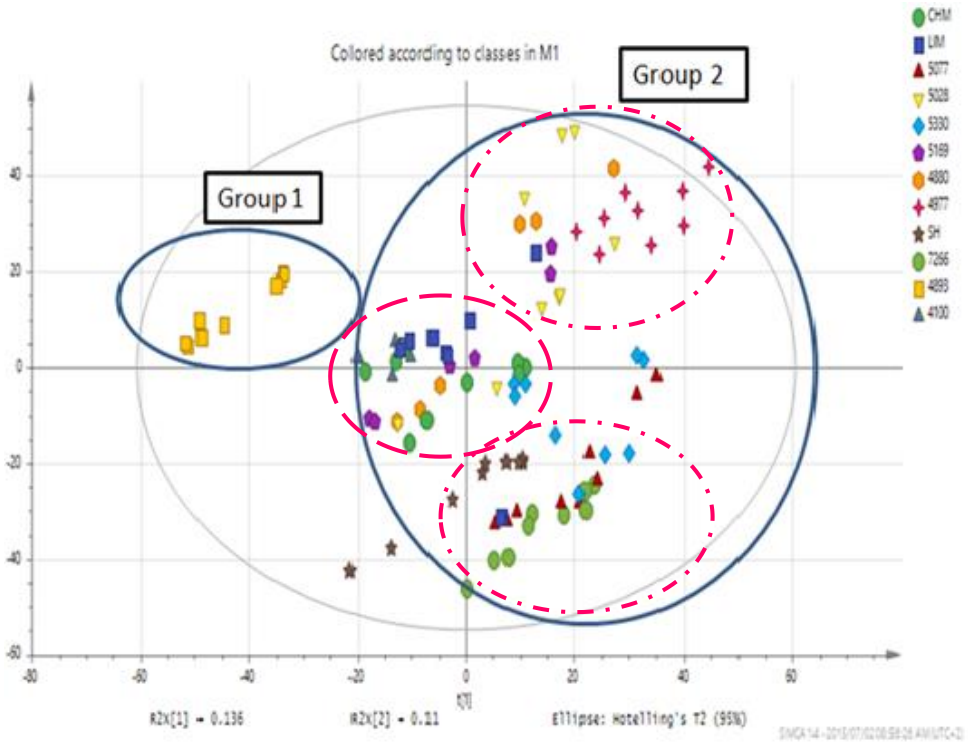


Fig. 7: PCA Score plot analyses of the twelve *M. oleifera* cultivars**.** PCA score plot illustrating the different clustering of samples based on their metabolic differences.

3.3 Hierarchical cluster analysis

Hierarchical cluster analysis (HCA) was conducted to reveal the relationship of the cultivars investigated, further providing a more detailed image of each metabolites contribution to the differentiation between cultivars. Confirmation of our findings were associated with the correlated variables found in PCA (Suberu *et al*., 2016). The output of the HCA computation is commonly displayed as a tree-like diagram called a dendogram, which is displayed in Fig. 8. The dendogram was divided into two distinct clades. The most distinct species, "4893" clustered completely separate from the other species. The rest of the cultivars that fell in clade 2 had subsequent subclades (Fig. 7). The distinct separation of the two clades mimicked the similar pattern of grouping of groups 1and 2 visualised in the PCA score plot (Fig. 7). The subsequent clades illustrated in clade 2 and the large grouping of cultivars from group 2 shown in the PCA (Fig. 8) could indicate that although the metabolites are similar in the cultivars, they display different behaviour in each of the cultivars (as

seen pictorially in the box and whiskers plot section 3.5. This further highlights that chemometric analysis in combination with metabolite composition could be an effective tool to explore the phylogenetic relationships within species (Dong *et al.*, 2015).

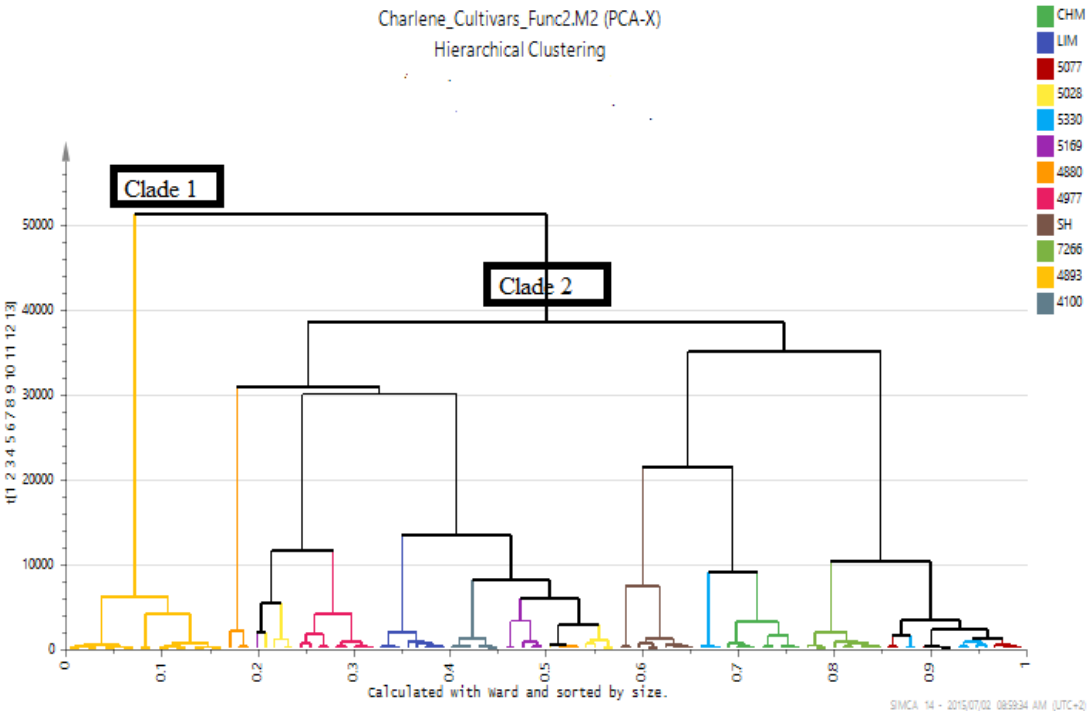


Fig. 8: Hierarchichal clustering analysis (HCA) of twelve *M. oleifera* cultivars**:** HCA depicting the metabolic profiles and relationship between twelve *M. oleifera* cultivars.

The replicates of each cultivar, in the second clade, were clustered into one group. The clustering of the cultivars in the second clade could be due to the genotype of the cultivars based on their geographical locations, however, it's important to note that even though the cultivars were sampled from different locations geographically, they were harvested at the same location under similar conditions. This indicates that the variation between them is primarily associated with genetic differences between the cultivars. This was contrary to eight *M. oleifera* cultivars sampled from different states in India. According to the dendogram, the cultivars were grouped into four sub-clusters but without any distinct geographical pattern

(Saini *et al*., 2012). Authors attributed this phenomenon due to the close proximity of the states where they were sampled which caused the spread of planting materials and high rates of gene flow through cross-pollination (Saini *et al*., 2012). The biological variation amongst the cultivars harvested from different locations is less than cultivar-to-cultivar variance (Dong *et al*., 2015). Further analysis is needed to probe the metabolomic differences between the cultivars. In order to unearth these differences, the cultivars that showed a distinct grouping, from the second clade, were studied further with the aid of box and whiskers plots.

3.4 Distribution of the compounds in cultivars using Box-and-Whiskers plots

As previously stated, the distribution of these three metabolites across the studied cultivars was found to vary significantly (Fig. 5, 6 and 10). Furthermore, the presence of these three molecules was shown to be remarkably higher in cultivars TOT4977, CHM and TOT5330 as opposed to the other cultivars were significant variation occurs. These results suggest that the production of these highly-sought compounds is cultivar specific. Recently, other pharmacologically relevant metabolites were also found to be ecotype specific in *Citrus grandis* L. Osbeck (Zhang *et al*., 2014) and in *M. oleifera* in particular (Förster *et al*., 2015), suggesting a possible difference in the genetic make-up across the different cultivars currently cultivated around the world. Further assessment of these results show that rutin appeared as the most abundant flavonoid compound in cultivars TOT4977, CHM and TOT5330 (Fig. 10), when compared to the other two rutinoside-bearing flavonoids. Isolation of flavonoids, rutin in particular, has also been shown to present a challenging exercise owing to the fact that it is easily transformed and it also sensitive to certain extraction methods (Dawidowicz *et al*., 2016). This could be one reason why these metabolites are rarely reported in *M. oleifera* plants (Rodriguez-Perez *et al.*, 2016; Khoza *et al*., 2014). Rutin accumulation was, however, reported to be significantly different between matured leaves and other organs in *Dimorphandra mollis* (Lucci *et al*., 2009), again suggesting its restricted distribution even within a plant. Similarily, kaempferol rutinoside showed a similar distribution amongst cultivars TOT4977, CHM and TOT5330, whilst, the lowest amount was present in cultivars TOT4100, TOT4893 and TOT5169 exhibited a low peak intensity amongst these cultivars (Fig. 10). The

presence of isorhamnetin rutinoside, in these results, was found to be lowest in cultivars TOT4977, CHM and TOT5330. Interestingly, although cultivars TOT4977, CHM and TOT5330 fell in the same clade on the HCA plot (Fig. 8), cultivars TOT4977 and TOT5330 were originally sampled from the world vegetable center (AVRDC) and CHM was sampled from Limpopo. Therefore, it would appear from the HCA that they have some common points of origin in evolutionary history and given the even spread amongst these cultivars demonstrated in the PCA (Fig. 7), the results may highlight that genetic modification is an important contributor to metabolite variability.

Furthermore, as previously stated, these cultivars, which also fell in cluster 2 in HCA, accumulated the highest amount of quercetin rutinoside as the dominant flavonoid. Amongst these cultivars, the concentration of quercetin rutinoside was the highest in cultivar TOT 4977. Thus this cultivar could be a resource for extraction of quercetin rutinoside due to the much higher amount found compared to the other cultivars. In the flavonoid biosynthetic pathway of seed plants, naringin is hydroxylated at the 3- position by flavonone 3-hydroxlase (F3H) to yield dihydrokaempferol then converted to dihydroquercetin by flavonoid 3'-hydroxylase (F3'H) which is finally converted to quercetin (Fig. 9). We speculate that the flavonoid biosynthetic pathway in cultivars TOT4977, CHM and TOT5330 lead directly to quercetin which would explain the noticeably higher levels of quercetin in these cultivars compared to kaempferol and isorhamnetin (Wang *et al*. 2016).

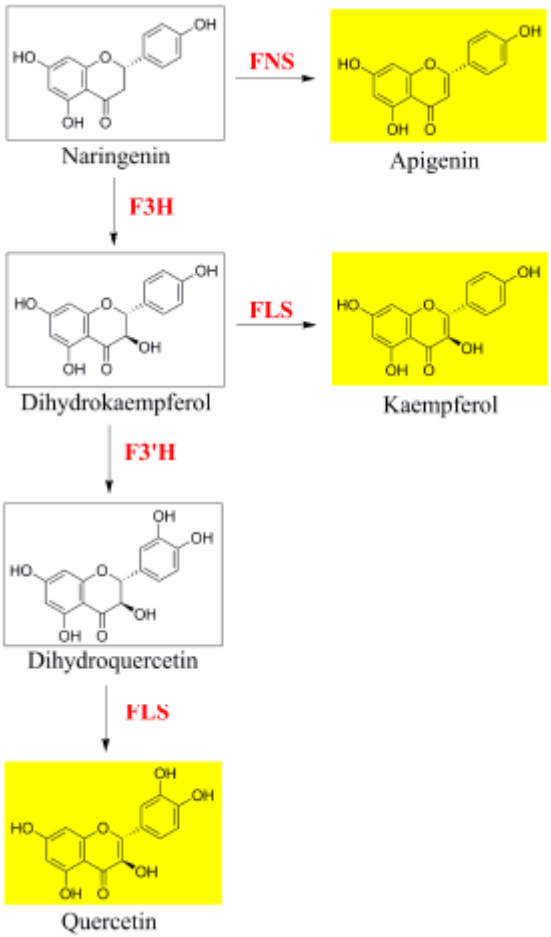
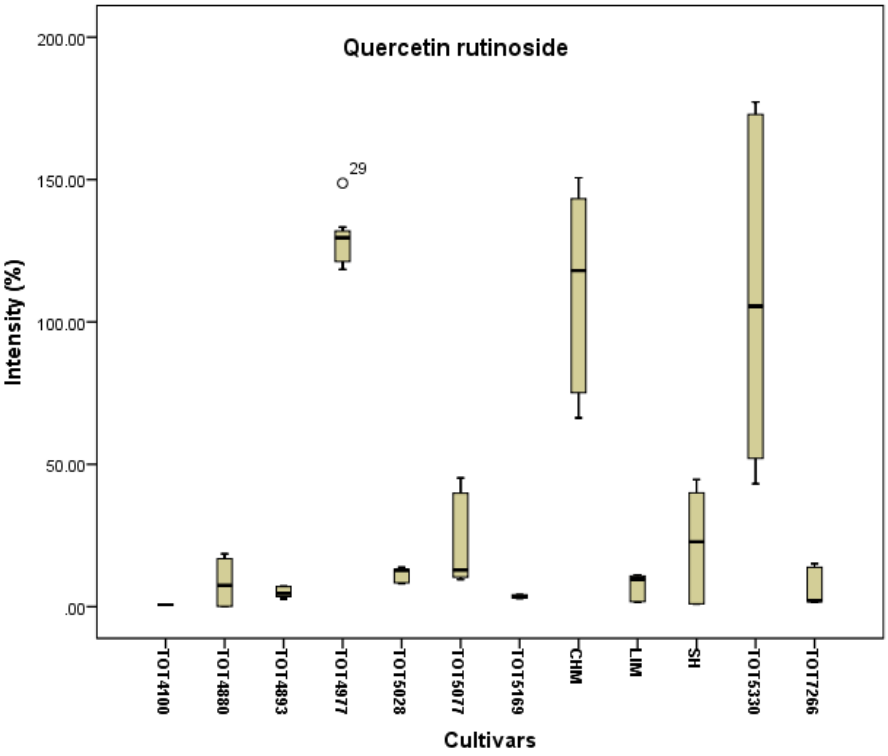


Fig. 9: Schematic diagram of the flavonoid biosynthetic pathway.

The variability of metabolites is thought to be due to environmental conditions (De Graaf *et al.*, 2015), however, in the current study all these cultivars were cultivated at the same place and time to exclude environmental factors (Ndhlala *et al*., 2014). Therefore, we can conclude that the observed chemical differences, under similar environmental conditions, are due to the genetic makeup of the cultivars (Förster *et al*., 2015). Moreover, cultivars containing low amounts of these compounds could be due to their previous environmental conditions where their original parental lines were grown, which might have influenced their genetic makeup. For instance, the glucosinolate content was found to increase under moderate drought stress in *Brassica carinata* A. Braun and *Brassica rapa* L., ssp. Rapa (Gutbrodt *et al.*, 2011). Similarily, anthocyanins and quercetin glycosides in apple fruits increased after exposure to the sun (Treutter *et al*., 2010), a phenomenon which can ultimately become a permanent phenotype.


Quercetin rutinoside
Intensity (%)
200.00
150.00
100.00
50.00
.00
29
TOT4100
TOT4880
TOT4893
TOT4977
TOT5028
TOT5077
TOT5169
CHM
LIM
SH
TOT5330
TOT7266
Cultivars

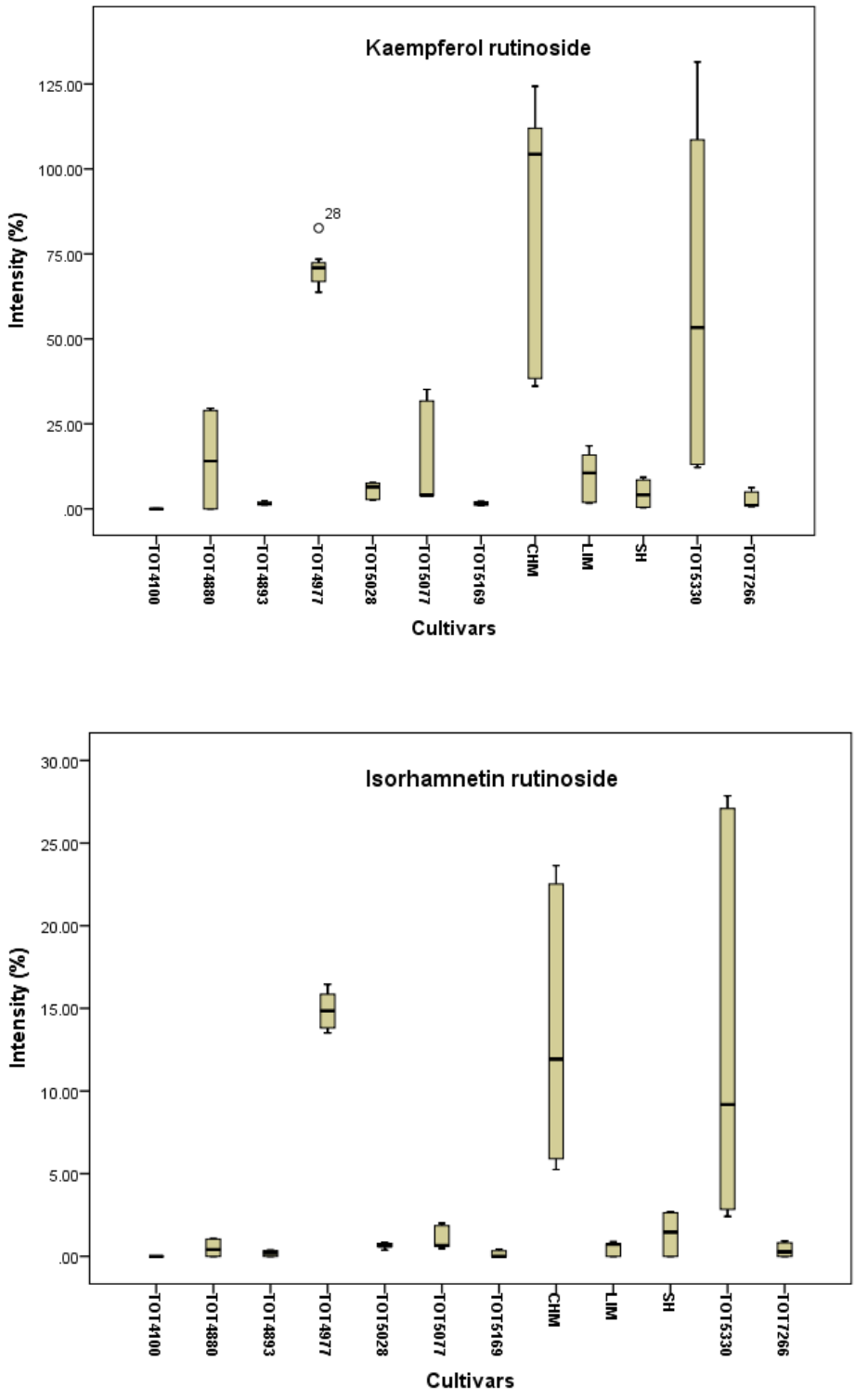


Fig. 10. Box-whiskers plots showing the distribution of kaempferol rutinoside, quercetin rutinoside and isorhamnetin rutinoside across 12 different *M. oleifera* cultivars.

3.5 Significance of kaempferol rutinoside, rutin and isorhamnetin rutinoside

Studies have shown the positive correlation of plant derived foods containing flavonoids to have a positive effect on human health (Calderón-Montaño *et al*., 2011). In fact, countless epidemiological studies have revealed the nutritional and medicinal properties associated with kaempferol rutinoside, rutin and isorhamnetin rutinoside (Calderón-Montaño *et al*., 2011; Nijveldt *et al*., 2001; Ayyanar *et al*., 2012). Pre-clinical trials have also shown that kaempferol and its glycosides, such as rutinoside, possess pharmacological properties associated with antioxidant, anti-inflammatory, anti-diabetic, anti-allergic to name a few (Calderón-Montaño *et al*., 2011). Isorhamnetin rutinoside has been shown to possess cytotoxic effects that have been proven to reduce human leukaemia cells (Boubaker*et al*., 2011). Lastly rutin, one of the most sought after compounds, is not only known to treat chronic ailments such as diabetes, hypertension and hypercholesterolemia but is also known to aid in the cure of neurodegenerative diseases caused by oxidative stress (Sharma *et al*., 2013; Park *et al*., 2014). The rutinoside bearing molecules in the tested cultivars confirm that the leaves of *M. oleifera* from cultivars TOT4977, CHM and TOT5330 to be a valuable source of the nutritional and medicinal applications stated earlier. Furthermore, this information could aid in improving competitiveness of the local sector as opposed to sourcing foreign products for nutritional purposes (Mendieta- Araica *et al*., 2011). In a study by Förster *et al.,* (2015) ecotypes TOT4880 and TOT7267 were recommended for biomass, nutritional and medicinal purposes, however, kaempferol, rutin and isorhamnetin rutinoside were not identified in these cultivars in that present study (Förster *et al*., 2015).

4.0 Conclusions

In conclusion, we attempted to correlate kaempferol rutinoside, rutin and isorhamnetin ruinoside as chemotaxonomic markers for distinguishing between 12 cultivars of *M. oleifera* from different locations. Principal component analysis (PCA) and hierachichal clustering analysis (HCA) based on the metabolites in the leaves of the different *M. oleifera* cultivars gave us a visual representation on how they relate to each other. With the aid of box-whiskers plots it gave us an indication that the variety of the rutinoside bearing molecules in these cultivars is due to the

unique genetic makeup of each cultivated *M. oleifera* cultivars. This study highlights the importance of determining plant secondary metabolites for cultivation purposes and that optimum cultivation of *M. oleifera* for nutritional and medicinal purposes should be highly dependent on the cultivar. Therefore, *Moringa oleifera* cultivars TOT4977, CHM and TOT5330 could be used as potential sources for neutraceuticals and medicinal purposes. Additional quantification of kaempferol rutinoside, rutin and isorhamnetin rutinoside should be implemented to investigate the concentration of these molecules. It is also recommended that the superior cultivars should be tested *in vitro* and possibly in field trials in *M. oleifera* cultivation areas.

Acknowledgments: National Research Foundation (NRF), Pretoria, University of Johannesburg and the University of Witwatersrand are thanked for financial support.

Author Contributions: CM, LC and EC conceived of the study, CM conducted the experiments and PS and EM the MS analyses. EM, LC, EC, AN, HA and PS supervised the project. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

5.0 References

Amaglo, N.K., Bennet, R.N., LoCurto, R.B., Rosa, E.A.S., LoTurco, V., Giuffrid, A., LoCurto, A., Crea, F., Timpo, G.M. (2010). Profiling selected phytochemicals and nutrients in different tissues of the multipurpose tree *Moringa oleifera* L. grown in Ghana. *Food Chemistry*. 122: 1047-1054.

Anwar, F., Latiff, S., Ashraf, M., Gilani, A.H. (2007). *Moringa* oleifera: a food plant with multiple medicinal uses. *Phytotherapy Research*. 21: 17-25.

Arbona, V., Iglesias, D.J.,Talo, M., Aurelio Gomez-Cadenas, A. (2009). Plant phenotype demarcation using nontargeted LC-MS and GC-MS metabolite profiling. *Journal of Agricultural and Food Chemistry*. 57: 7338-7347.

Ayyanar, M., Subash-Babu, P. (2012). *Syzygium cumini* (L.) Skeels: A review of its phytochemical constituents and traditional uses. *Asian Pacific Journal of Tropical Biomedicine*, 2*:* 240–246.

Boubaker, J., Bhouri, W., Sghaier, M.B., Ghedira, K., Dijoux, M.G., Chekir-Ghedira, L. (2011). Ethyl acetate extract and its major constituent, isorhamnetin 3-

O-rutinoside, from Nitraria retusa leaves, promote apoptosis of human myelogenous erythroleukaemia cells. *Cell Proliferation*. 44: 453–461.

Calderón-Montaño, M., Burgos-Morón, E., Pérez-Guerrero, C., López-Lázaro, M. A review on the dietary flavonoid. (2011). *Mini-Reviews in Medicinal Chemistry*. 11: 298-344.

Canter, P.H., Thomas, H., Ernst, E. (2005). Bringing medicinal plants into cultivation: opportunities and challenges for biotechnology. *Trends in Biotechnology*. 23: 180-185.

Dawidowicz, A., Bernacik, K., Typek, R. (2016). Does the plant matrix type have impact on rutin transformation during its extraction? *Food Analytical Methods*. 5: 2042–2051.

De Graaf, R.M., Krosse, S., Swolfs, A.E.M., te Brinke, E., Prill, N., Leimu, R., van Galen, P.M., Wang, Y., Aarts, M.G.M., van Dam, N.M. (2015). Isolation and identification of 4-α-rhamnosyloxy benzyl glucosinolate in Noccaea caerulescens showing intraspecific variation. *Phytochemistry*. 110: 166-171.

Devaraj, S.K., Santhosh Kumar, D.S., Doriya K., Gopalakrishnan, L. (2016). *Moringa oleifera*: A review on nutritive importance and its medicinal application. *Food Science and Human Wellness*. 5: 49-56.

Dong, W., Tan, L., Zhao, J., Hu, R., Lu, M. (2015). Characterization of fatty acid, amino acid and volatile compound compositions and bioactive components of seven coffee (Coffea robusta) cultivars grown in Hainan Province, China. *Molecules*. 20.

Farag, M.A., Mohsen, M., Heinke, R., Ludger A.Wessjohann, L.A. (2014). Metabolomic fingerprints of 21 date palm fruit varieties from Egypt using UPLC/PDA/ESI–qTOF-MS and GC–MS analyzed by chemometrics. *Food Research International*. 64: 218–226.

Ferreira, M.P., Farais, D.F., Oliveira, J.T.D.A., Carvalho, A.D.F.U. (2008) *Moringa oleifera*: bioactive compounds and nutritional potential. *Nutrition Research*. 21: 431-437.

Förster, N., Ulrichs, C., Schreiner, M., Arndt, N., Schmidt, R., Mewis, I. (2015). Ecotype variability in growth and secondary metabolite profile in *Moringa oleifera*: Impact of sulfur and water Availability. *Journal of Agriculture and Food Chemistry*. 63: 2852–2861.

Fujimura, Y., Kurihara, K., Ida, M., Kosaka, R., Miura, D., Hiroyuki Wariishi, H., Maeda-Yamamoto, M., Nesumi, A., Saito T., Kanda, T., Yamada, K., Tachibana,

H. (2011). Metabolomics-driven nutraceutical evaluation of diverse green tea cultivars. *PLoS One*. 6.

Girennavar, B., Jayaprakasha, G.K., Mclin, S.E., Maxim, J., Yoo, K.S., Patil, B.S. (2008). Influence of electron-Bea irradiation on bioactive compounds in grapefruits (*Citrus paradisi Macf*.). *Journal of Agriculture and Food Chemistry*. 56: 10941-10946.

Gutbrodt, B., Mody, K., Dorn, S. (2011). Drought changes plant chemistry and causes contrasting responses in lepidopteran herbivores. *Oikos*. 120: 1732-1740.

Gutzeit, D.,Wray, V.,Winterhalter, P., Jerz, G. (2007). Preparative isolation and purification of flavonoids and protocatechuic acid from sea buckthorn juice concentrate (*Hippophae¨ rhamnoides* L. ssp. *rhamnoides*) by high speed counter-current chromatography. *Chromatographia*. 65: 1–7.

Habtemariam, S., Varghese, G.K. (2015). Extractability of rutin in herbal tea preparations of *Moringa stenopatala* leaves. *Beverages*. 1: 169-182.

Hvattum, E., (2002). Determination of phenolic compounds in rose hip (Rosa canina) using liquid chromatography coupled to electrospray ionasation tandem mass spectrometry and diodide-array detection. *Rapid Communications in Mass Spectrometry*. 166: 655–662.

Khoza, B.S., Chimuka, L., Mukwevho, E., P. A. Steenkamp, P.A., N. E.Madala, N.E. (2014). The effect of temperature on Pressurised Hot Water Extraction of pharmacologically important metabolites as analysed by UPLC-qTOF-MS and PCA. *Evidence-Based Complementary and Alternative Medicine*. 2014.

Kwon, Y., Myung, S.A., Park , J.S., Liu , J.R., In, D.S., Min, B.W., Suk Weon Kim, S.W. (2014). Discrimination of cultivation ages and cultivars of ginseng leaves using Fourier transform infrared spectroscopy combined with multivariate analysis. *Journal of Ginseng Research*. 38: 52-58.

Lucci, N., Mazzafera, P. (2009). Distribution of rutin in Fava D'anta (*Dimorphandra mollis*) seedlings under stress. *Journal of Plant Interactions*. 4: 203-208.

Madala, N.E., Piater, L.A., Steenkamp, P.A., Ian 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. *Springer Plus*. 3: 254.

Makkar, H.P.S., Becker, K. (1996). Nutritional value and anti-nutritional components of whole and ethanol extracted *Moringa oleifera* leaves. *Animal Feed Science and Technology*. 63: 211–228.

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.

Martucci, M.E.P., De Vos, R.C.H., Carollo, C.A., Gobbo-Neto, L. (2014). Metabolomics as a potential chemotaxonomical tool: Application in the genus *Vernonia* Schreb. *Plos One*. 9: 1-5.

Mendieta-Araica, B., Sporndly, E., Reyes-Sanchez, N., Sporndly, R. (2011). Feeding *Moringa oleifera* fresh or ensiled in dairy cows-effects on milk flavour. *Tropical Animal Health and Production*. 43: 1039-1047.

Ndhlala, A.R., Mulaudzi, R., Ncube, B., Abdelgadir, H.A., du Plooy, C.P. Van Staden, J. (2014). Antioxidant, antimicrobial and phytochemical variations in thirteen *Moringa oleifera* Lam. cultivars. *Molecules*. 19: 10480-10494.

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: 66.

Nijveldt, R.J., Van Nood, E., van Hoorn, E., Boelens, P.G.., Van Norren, K., van Leewun, P.A. (2001). Flavonoids: A review of probable mechanisms of action and potential applications. *The American Journal of Clinical Nutrition*. 74: 418-425.

Pakade, V., Cukrowska, E., Chimuka, L. (2012). Metal and flavonol contents of *Moringa oleifera* grown in South Africa. *South African Journal of Science*. 109: 1–7.

Rajangam, J., Azahakia, R.S., Manavalan, A., Thangaraj, T., Vijayakumar, A., Muthukrishan, N. (2001). Status of production and utilization of *Moringa* in Southern India. In: Development potential for *Moringa* products. Workshop proceedings. October 29- November 2, 2001, Dar es Salaam, Tanzania.

Saini, R.K., Saad, K.R., Rayishankar, G.A., Giridhar, P., Shetty, N.P. (2013). Genetic diversity of commercially grown *Moringa oleifera* Lam. cultivars from India by RAPD, ISSR and cytochrome P450-based markers. *Plant Sytematics and Evolution*. 299: 1205-1213.

Sharma, S., Ali, A., Ali, J., Sahni, J.K., Baboota, S. (2013). Rutin: therapeutic potential and recent advances in drug delivery. *Expert Opinion on Investigational Drugs*. 22*:* 1063-79.

Stohs, S.J., M.J. Hartman, M.J. (2015). Review of the safety and efficacy of *Moringa oleifera*. *Phytotherapy Research*. 29: 796–804.

Suberu, J., Gromskib, P.S., Nordonb, A., Lapkina, A. (2016). Multivariate data analysis and metabolic profiling of artemisinin and related compounds in high yielding varieties of *Artemisia* annua field-grown in Madagascar. *Journal of Pharmaceutical and Biomedical Analysis*. 117: 522–531.

Treutter, D. (2010). Managing phenol contents in crop plants by phytochemical farming and breeding visions and constraints. *Molecular Sciences*. 11: 807-857.

Trygg, J., Holmes, E., Lundstedt, T. Chemometrics in Metabonomics. (2007). *Journal of Proteome Research*. *6:* 469–79.

Vagiri, M., Conner, S., Stewart, D., Andersson, S.C., Verrall, S., Johansson, E., Rumpunen, K. (2015). Phenolic compounds in blackcurrant (*Ribes nigrumL*.) leaves relative to leaf po- sition and harvest date. *Food Chemistry*. 172: 135–142.

Wang, Y., Fu, J., Zhang, C., Hongbo Zhao, H. (2016). HPLC-DAD-ESI-MS analysis of flavonoids from leaves of different cultivars of Sweet Osmanthus. *Molecules.* 21: 12224.

Zhang, M., Nan, H., Wang, H., Jiang, X., Li, Z. (2014).Comparison of flavonoid compounds in the flavedo and juice of two Pummelo cultivars (*Citrus grandis* L. Osbeck) from different cultivation regions in China. *Molecules*, *19*: 17314-173.

Chapter 5: Conclusion and Future Prospects

In this chapter, conclusions based on experimental work based on the publications in **Chapter 4** is discussed. The work carried out in this research as well as future prospects is outlined.

5.1 General conclusion

In summary, the use of UPLC –qTOF-MS for semi-targeted metabolic profiling and metabolic identification in *Moringa* was reported in this thesis. The aim of this thesis was to develop a metabolic fingerprinting approach by UPLC-qTOF-MS that can be used for taxonomic classification of *Moringa* plant species.

Retention time information together with accurate mass measurements led us to the identification and characterization of pharmacologically active metabolites such as flavonoids and chlorogenic acids in *M. oleifera* and *M. ovalifolia* plant species. Moreover, flavonoids and chlorogenic acids were reported for the first time in the leaves of *M. ovalifolia*, which is endemic to Namibia. In *M. ovalifolia*, the profiling and annotation of CGA molecules with the aid of in-source-collision–induced dissociation (ISCD) together UPLC-qTOF-MS, successfully profiled geometric and regional isomers in this plant further highlighting the use of *M. ovalifolia* for medicinal and nutritional purposes. The outcomes of the results recommend that metabolite profiling is viewed as a valuable apparatus for recognizing the metabolic conditions of these two species and further understanding the relationship of flavonoid and chlorogenic metabolites as per these two species. The results also highlighted the complexity and importance of correct taxonomic classification illustrated by the correct identification of the sugar attachment to the flavonoid aglycone.

Furthermore, a specific subset of LC-MS flavonoid ions such as quercetin rutinoside (rutin), kaempferol rutinoside and isorhamnetin rutinoside that were found in *M. ovalifolia* was further used as biomarkers for twelve *M. oleifera* plant cultivars in an attempt to create a grouping model to decipher similarities and differences between these cultivars developing at a similar place, however, initially from various areas.

Chemotaxonomic metabolite profiling of 12 *M. oleifera* cultivars using UHPLC-qTOf-MS combined with multivariate statistical analysis added to the data complementarity and facilitated in determining the relationship between the cultivars. The box-whiskers plots with the aid of biomarkers identified specific

Moringa cultivars that could be used for nutritional and therapeutic purposes. Considering the outcome of the statistical analyses, the classification model with the aid of the biomarkers appears to represent a novel approach for metabolic fingerprinting amongst related species. Furthermore, this study has demonstrated the effectiveness of an UHPLC-MS-based platform in comparing the metabolite profiles of the *Moringa* species. As the distribution of these annotated metabolites were found to vary from one cultivar to cultivar, the results verify that not all *M. oleifera* species that are cultivated and harvested are similar and that they vary from one another. In addition, a semi-targeted metabolomics approach based on an UHPLC-qTOF-MS platform and multivariate data models was found to be effective and shows the potential to identify pharmacological and nutritionally relevant and important metabolites.

That being said, the current study provided an in-depth insight into the metabolome of different *M. oleifera* extracts and *M. ovalifolia* extracts. And thus, paves the way for the advancement in analytical manipulation and holistic utilization of all properties of *Moringa*. This study further reflects the need to further elucidate underlyning biochemical differences for general plant research purposes.

5.2 Future prospects

- To improve the robustness of this system, future studies should also focus on studying *Moringa cultivars* from other species such as *Moringa Stenepotala* to determine a superior *Moringa* species and to analyse the concentration of other secondary metabolites such as chlorogenic acids and glucosinolates amongst the *Moringa* species.

- Using plant metabolome studies to differentiate various *Moringa* powder products on the market and to ascertain quality of the various products.

- Using plant metabolomics studies to look at different *M. oleifera* plants growing in the SADC region and in other regions of Africa and compare superiority in terms of geopgraphical location and cultivar or specie types.

- To compare nutritional and herbal compounds in *Moringa* with other African indigenous plants such as Rooibos, Amaranthus (black jack) and others. This is important so as to know how *Moringa* compliments to these in nutritional and herbal compounds.

Appendix

Analysis of genetic diversity and confirmation of potential biomarkers of twelve *Moringa oleifera* Lam. cultivars

Charlene Makita[1], Ewa Cukrowska[1], Ashwell Rungano Ndhlala[2], Hafiz Abdelgadir[2], Luke Chimuka[1], Paul Steenkamp[3,4], Ntakadzeni Edwin Madala[3*].

[1]Molecular Science Institute, School of Chemistry, University of Witwatersrand, P/Bag 3, Braamfontein 2050, Johannesburg, South Africa.

[2]Agricultural Research Council, Vegetable and Ornamental Plants (VOP), Private Bag X293, Pretoria 0001, South Africa

[3]Department of Biochemistry, University of Johannesburg, P. O. Box 524, Auckland Park, 2006, Johannesburg, South Africa.

[4]CSIR Biosciences, Natural Products and Agro-processing Group, Building 20, Pretoria, 0001, South Africa.

Abstract

Moringa oleifera Lam. (*Moringaceae*), known as "drum stick" is a multi-purpose plant that has many nutritional and medicinal properties. Motivated by pharmaceutical and nutritional considerations, this metabolic profiling study aims to identify superior *Moringa* cultivars for that purpose. Moreover, this study was also aimed to identify genetic variations amongst the cultivars. Further classification by hierarchical clustering analysis, principal component analysis, cloud plots and heat maps demonstrated significant differences between the samples based on their biochemical differences. Such metabolomics approaches can be used in the context for identification of superior *Moringa* cultivars.

1. Introduction

M. oleifera Lam., belonging to the *Moringaceae* family is native to India and is commonly cultivated in Asia and in many African countries. In Africa, countries such as Malawi, Ghana and Senegal are said to have the highest cultivation of *M. oleifera* trees **(**Pakade *et al*., 2013**)**. The trees are cultivated mainly for their leaf material because they exhibit high contents of

metabolites such as phenolic acids and flavonoids associated with various pharmacological activities (Förster *et al.,* 2015). In fact, a number of studies to determine the nutritional composition have confirmed the usage of *M. oleifera* as a dietary supplement (Habtemariam *et al*., 2015). Not only is *M. oleifera* known for its potential as a nutritional supplement but is also known for its medicinal properties such as antimicrobial, anti-inflammatory, detoxifying and anti-cancerous properties (Förster *et al.,* 2015).

Currently, the search for rutin amongst neutraceutically important plants has gained momentum primarily due to its medicinal and nutritional properties and has previously been used as a biomarker for the selection of a superior *Moringa* cultivar (Makita *et al*., 2016). Despite the astounding properties that rutin possesses, it has been recognised that the importance of rutin has been over-emphasised relative to the many other (unknown) contributory compounds that could make the *Moringa* cultivar nutritionally and pharmacologically superior.

During the last decade, there has been an increase in the cultivation of *Moringa* trees due to its high biomass under intense farming conditions (Makkar and Becker, 1996). Plant breeding does not only improve the agronomic and medicinal traits of plants, but techniques such as marker-assisted systems can be used as a useful tool to aid in hybrid selection (Canter *et al.,* 2005). The use of plant metabolomics has made it possible to analyse the metabolomic composition between plant species and cultivars, therefore, giving us a broad perspective of the phenotypic and genotypic differences amongst plant varieties as a result of gene expression (Farag *et al.,* 2014). Metabolomic approaches in combination with chemometric data tools such as multivariate data analysis (MVDA) models allows key biological information to be interpreted from complex data sets as shown in Paper **III** (Kwon *et al*., 2014; Madala *et al*., 2014). In this study, a comparative study using various contrasting yet complementary multivariate statistical analysis platforms such as principal component analysis (PCA), hierarchal clustering analysis (HCA), cloud plots and heat maps from the XC-MS online software were utilised to reveal underlying differences of the metabolic profile, relationship and health promoting attributes amongst twelve diverse *M. oleifera* cultivars and four representative *Moringa* cultivars 4893, 5028 and TOT7266. Results from this study will provide new knowledge about the bioactive compounds found in these cultivars and further provide information to agronomists and farmers in the search for a superior and optimal *Moringa* cultivar.

2.0 Material and Methods

2.1 Plant collection

Twelve *M. oleifera* Lam. cultivars were collected from different geographical locations around the world and cultivated at the Agricultural Research Council (ARC) in Roodeplaat, Pretoria (South Africa). Eight cultivars from the World Vegetable Centre (AVRDC) in Thailand which included: TOT4893 ("4893 in text"), TOT4951 ("4951" in text), TOT4977 ("4977" in text), TOT 5028, TOT5077 ("5077" in text), TOT5169 ("5169" in text), TOT5330 ("5330" in text) and TOT7266. One cultivar from Taiwan which included: TOT4100 ("4100" in text). One cultivar from USA TOT4880 ("4880" in text). Three cultivars from South Africa which includes: Silver Hill (SH), CHM and Limpopo ("LIM" in text). The same management practices were applied to the cultivars. They were grown with no fertilizer and watered three times a week and was subjected to the same management practices. The plat cultivars were harvested at 8 months after planting.

2.2 Metabolite extraction

The leaves of the twelve *Moringa* cultivars were air-dried then the dried leaves were crushed using pestle and mortar to a fine powder. The metabolites were extracted from the crushed leaves (2 g) using 80% MEOH (20 mL). The samples were sonicated for 30 min using an ultrasonic bath (SB-120DT, Loyal Key Group (Hong Kong) Co. Ltd). The tissue debris was separated from the homogenate by means of centrifugation at 5000xg for 10 min at room temperature (25ºC). The supernatant liquid was decanted in a clean tube and dried to approximately 1.5 mL using a rotary evaporator apparatus at 55ºC. The 1.5 mL extract was then transferred to a 2 mL Eppendorf tube and subsequently dried to completeness using a speed vacuum concentrator at 45ºC under negative vacuum (Eppendorf, Merck, South Africa). The resulting pellet was reconstituted in 1 mL 50% MeOH and finally filtered through 0.22 µm nylon filters. The resulting extracts were then stored in a freezer at -20ºC to avoid degradation until they were analysed on the UPLC- qTOF-MS.

2.3 UPLC conditions

For the evaluation of metabolites in *M. oleifera* cultivars, UPLC coupled to a SYNAPT G1 HDMS mass spectrometer was used to generate accurate mass data (Waters, MA, USA). The column used for chromatographic separation of the *Moringa* extracts (1 µL) was accomplished using a 30 min long gradient chromatographic method, on a Waters BEH C_{18} reverse phase column (150 mm x 2.1 mm, 1.7 µm). For chromatographic elution, a binary solvent mixture was used consisting of 0.1% formic acid in deionized water (Eluent A) and 0.1% formic acid in acetonitrile (Eluent B). The chromatographic gradient was as follows: the initial conditions were 98% eluent A followed by multiple gradients to 5% A at 26 minutes. The conditions were kept constant for 1 minute and then changed to the initial conditions. The flow rate was set at 0.4 mL/min and maintained for 1 minute and the column was finally re-equilibrated for 3 min prior the next run. Chromatographic separation was monitored initially using a photodiode-array (PDA) detector coupled in tandem to an electrospray ionization mass spectrometer (ESI–MS) detector. The PDA detector was set to scan 200-500 nm and collecting 20 spectra per second.

2.4 MS-conditions

Chromatographic separation (of the 1.5 mL extract) was performed on an ACQUITY UPLC system (Waters Corporations, Milford, MA) using a conditioned autosampler at 4 °C. For MS detection, an electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-qTOF-MS) detector was used. Leucine enkephalin (556.227/554.2615 Da) was used as reference calibrant to obtain typical mass accuracy of at least 5 mDa at a flow rate of 0.1 mL/min. The optimal comditions of analysis are as follows: The mass spectrometer was operated in negative and positive ionization mode with a capillary voltage of 2.5 kV, the sampling cone at 30 V, and the extraction cone at 4 V, multichannel plate detector potential at 1600 V, source temperature at 120°C, desolvation temperature at 450°C, cone gas flow at 50 L/h, and desolvation gas flow at 550 L/h. The scan time was 0.2 seconds and an inter-scan delay of 0.02 s covering the mass range of 100 to 1000 Da. The fragmentation patterns of the compounds were obtained in a MS file with three functions. The settings are as follows: Function 1 un-fragmented using tune page settings; Function 2 a Trap Collision Energy Ramp (TCER) of 10 - 30 V; Function 3 a TCER of 15-50 V.

2.5 Multi-variate data analysis (MVDA)

To create a data matrix for multivariate data modelling, the acquired raw data was imported to SIMCA version 13.0 (Soft independent modelling of class analogy) software (Umetrics Corporation, Umea, Sweden) and XCMS online (https://xcmsonline). Initially, the UPLC-MS data was exported and analysed with the aid of the MarkerLynx XS™ software (Waters, Manchester, UK). The following parameters were used for maximum data output: The following parameters were chosen: retention time (Rt) of 1–27 min, mass range of 100–1000 Da, mass tolerance of 0.02 Da and Rt window of 0.2 min. The analyses conditions/parameters were kept constant for both negative and positive data. The dataset obtained from MarkerLynx XS™ was transported to the SIMCA 13.0 software to generate principal component analysis (PCA) and hierarchal clustering analysis (HCA). Pareto scale was utilised for both models. PCA and HCA are utilised as visual tools to determine the differences between samples based on their metabolic profiles. The validity of the generated models were determined by metabolic diagnostic tools such as: cumulative model variation in the matrix X, R2X (cum) which is the goodness-of-fit-parameter the proportion of the variance of the response variable that is explained by the model, R2Y (cum) and predictive ability parameter (R2Y (cum)), also known as the total variation fraction of matrix X predicted by an extracted components (Mhlongo *et al*., 2016; Ni *et al*., 2008; Sadeghi-bazigani *et al*., 2011).

Moreover, the data was analysed using interactive XCMS Online which is a web- interfaced, automated, metabolomics data processing software intended to identify biomarkers of which the relative intensity varies between sample groups (Mhlongo *et al*., 2016). The Mass Lynx raw data (.raw) was converted with the aid of Data bridge software (Waters, MA, USA) and transferred to the XCMS Online statistical package which is freely available at (https://xcmsonline.scripps.edu). The method parameters were chosen for UHPLC/UHD-QTOF specificities and were as follows: (i) feature detection set as centWave method, minimum peak width = 5, maximum peak width = 20, (ii) Rt correction set as Obiwarp method, Profstep = 1, alignment set as m/z width = 0.015, min- fraction = 0.5, bw = 5, and statistics set as statistical test = Unpaired parametric t-test (Welch t- test), paired t-test and post-hoc analysis with the threshold p-value = 0.01 and fold-change = 1.5. The software further calculates the p-values as well as fold-changes of the metabolites (variables) across different samples of varying biological background. MVDA such as PCA score plots, HCA, Heat maps and Cloud plots were generated to detect major differences between the *Moringa* cultivars (Mhlongo *et al*., 2016).

3.0 Results and Discussion

3.1 Multivariate data analysis and visual chromatograms

The combination of spectroscopy and multivariate data analysis is an efficient tool to obtain quanitative and qualitative information. Chemometric data evaluation stemming from BPI chromatograms as well as PCA and HCA analysis of the 12 *Moringa* cultivars were discussed in detail in Chapter 4, Paper III. To reduce the level of complexity and the number of variables in the data set, three representative cultivars namely 4893, 5028 and TOT7266 were compared and analysed further with the aid of XC-MS Online software. Cultivar 4893 was taken from the first clade in the HCA dendogram, 5028 was taken from the first subgroup under clade 2 and 4893 was taken from the second subgroup under clade 3 (**Refer to Paper III, Fig. 8**).

3.2 Principal component analysis score plot with the aid of XCMS online software

Although principal component analysis **(**PCA) and HCA are effective tools for determining differences between cultivars, it does have its own limitations. Therefore, XCMS online was used to further complement SIMCA 13.0 to aid in further identifying more metabolites between the three cultivars. Another PCA score plot was performed utilising the representative cultivars 4893, 5028 and TOT7266. The relationship between the cultivars were properly differentiated only when they were analysed independent of the rest of the cultivars. The PC1 (35%) vs. PC2 (22%) score plot of the data produced three independent clusters of the cultivars again with 4893 separating from the rest of the cultivars (Fig. 1). The distribution of the three cultivars was similar to that observed among the twelve cultivars in **Paper III, Fig. 7**. The independent clustering gives an indication of the differences in the chemical composition amongst the cultivars, which may be responsible for differences in their bioactivity. Therefore, we conducted further analysis to determine which analytes contributed to the variance in bioactivity with the aid of multi-group cloud, heat maps and box whiskers plots.

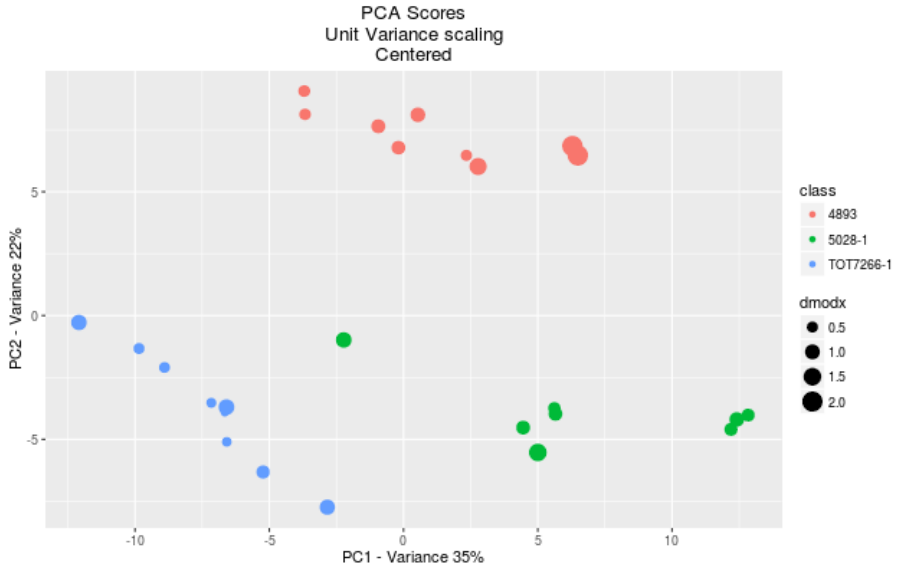


Fig. 1: PCA analyses of cultivars 4893, 5028 and TOT7266: PCA score plot illustrating the different clustering of samples corresponding to their different biochemical profiles.

3.3 Mutli-group cloud plot

A multi-group cloud plot from the XCMS online software was generated to visualise and compare statistical significant metabolites between cultivars. Multi-group cloud plot analysis can be considered as an important tool to distinguish important metabolites with specific functional roles and to characterize metabolic patterns specific to the different *Moringa* cultivars. Multi-group cloud plots present significant metabolite features across different analysed groups or data classes (Gowda *et al*., 2014). Each significant compound is represented as a bubble and the feature assignments are represented by m/z, retention time, p-value and direct fold change (Gowda *et al*., 2014). The size of each bubble is represented by the direct fold change: the bigger the bubble, the bigger the fold change (Patti *et al*., 2013; Godwa *et al*., 2014). Statistical significant metabolites are based on the colour intensity of each bubble and was calculated by the Welch's t-test with unequal variance (Gowda *et al*., 2014; Benton *et al*., 2008). The colour intensity of each bubble is represented by the p-value ($p < 0.01$). Low p-values have a darker intensity while features with higher p-values are brighter (Patti *et al*., 2013; Gowda *et al*., 2014; Benton *et al*., 2008). The y-axis corresponds to the m/z ratio while the x-axis is the retention time (Rt) of the compound. Each feature, with a black border is connected to a METLIN database to provide tentative identification and characterization of untargeted metabolites based on their m/z values (Gowda *et al*., 2014). Cultivars 4893, 5028 and TOT7266 were compared and generated by XCMS online and are presented as multi-group cloud plots (Fig. 2) XCMS Online identified 92 features. According to the cloud plot,

statistical significant metabolites consisted of flavonoids, glucosinolates and chlorogenic acids. Characterization of the untargeted metabolites was confirmed by the METLIN database as well as previously published data (Makita *et al*., 2016). Box-plots were illustrated to further confirm the specificity and significant secondary metabolites found in the cultivars.

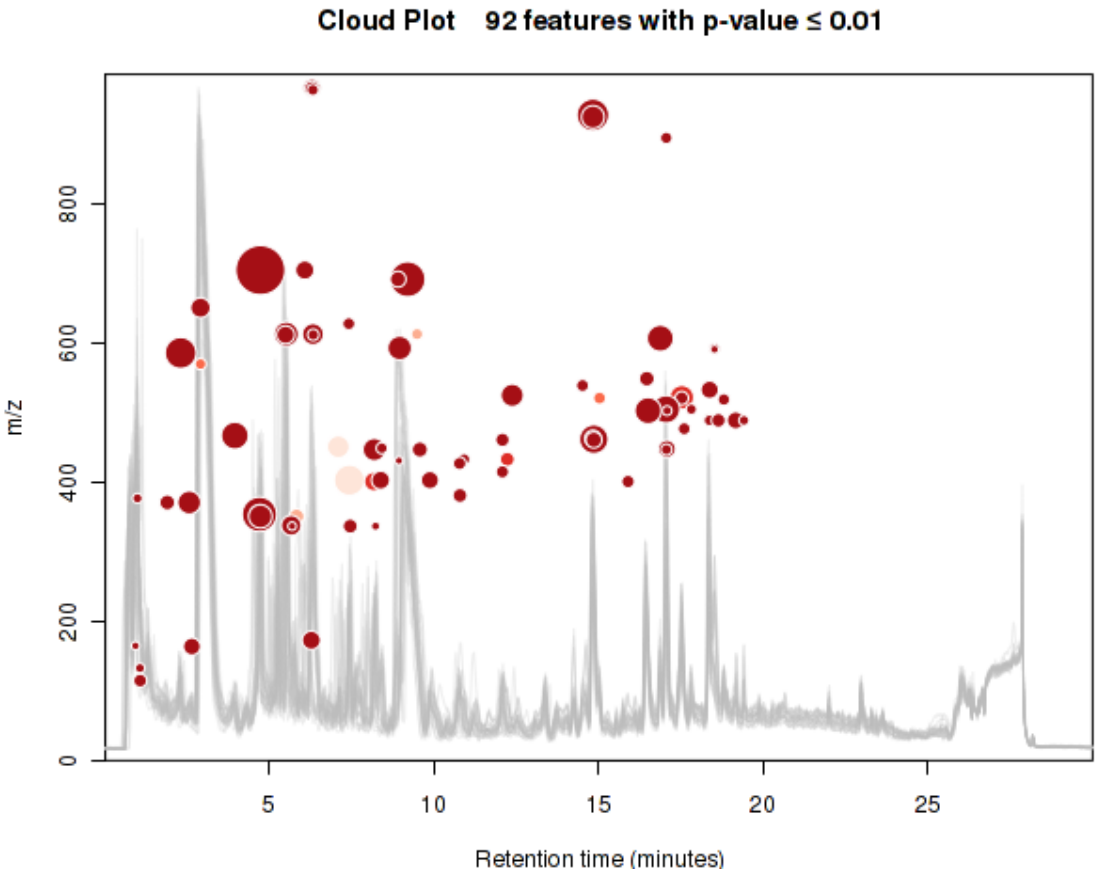


Fig. 2: XCMS interactive cloud plot. Shown are cloud plots of cultivars 4893, 5028 and TOT 7266. Each dot represents a specific feature, specifically, an ion with its m/z and retention time (Rt). The annotated biomarkers displayed in the interactive cloud plot were associated with the highest statistical significance ($p \leq 0.01$).

3.4 Heat-map and box whiskers plots

The correlation between each cultivar can be further visualised in the form of a heat map which also includes a hierarchical clustering dendogram. Just like the cloud plot, the heat map gives an overview of all the samples in the dataset by emphasising the holistic differences in the complex metabolic data (Fujimura *et al*., 2011). The HCA confirmed two distinct clusters defined by the three cultivars (Fig. 3). The HCA was not in absolute agreement with the PCA shown in Fig. 1. Previously, we had observed that cultivar 4893 was different from the other

Moringa cultivars by forming an independent cluster (Chapter 4, Paper III**)**. In Fig. 3, however, we see that although two clusters are present in the HCA, 4893 appears in both clades relating to TOT7266 and 5028. Interestingly, the three cultivars originally were sampled from the world vegetable center (AVRDC) in Thailand, therefore it would appear from the HCA that they have some common points of origin in evolutionary history, and given the even spread amongst these cultivars demonstrated in the PCA, the results highlight that genetic modification is an important contributor to metabolite variability.

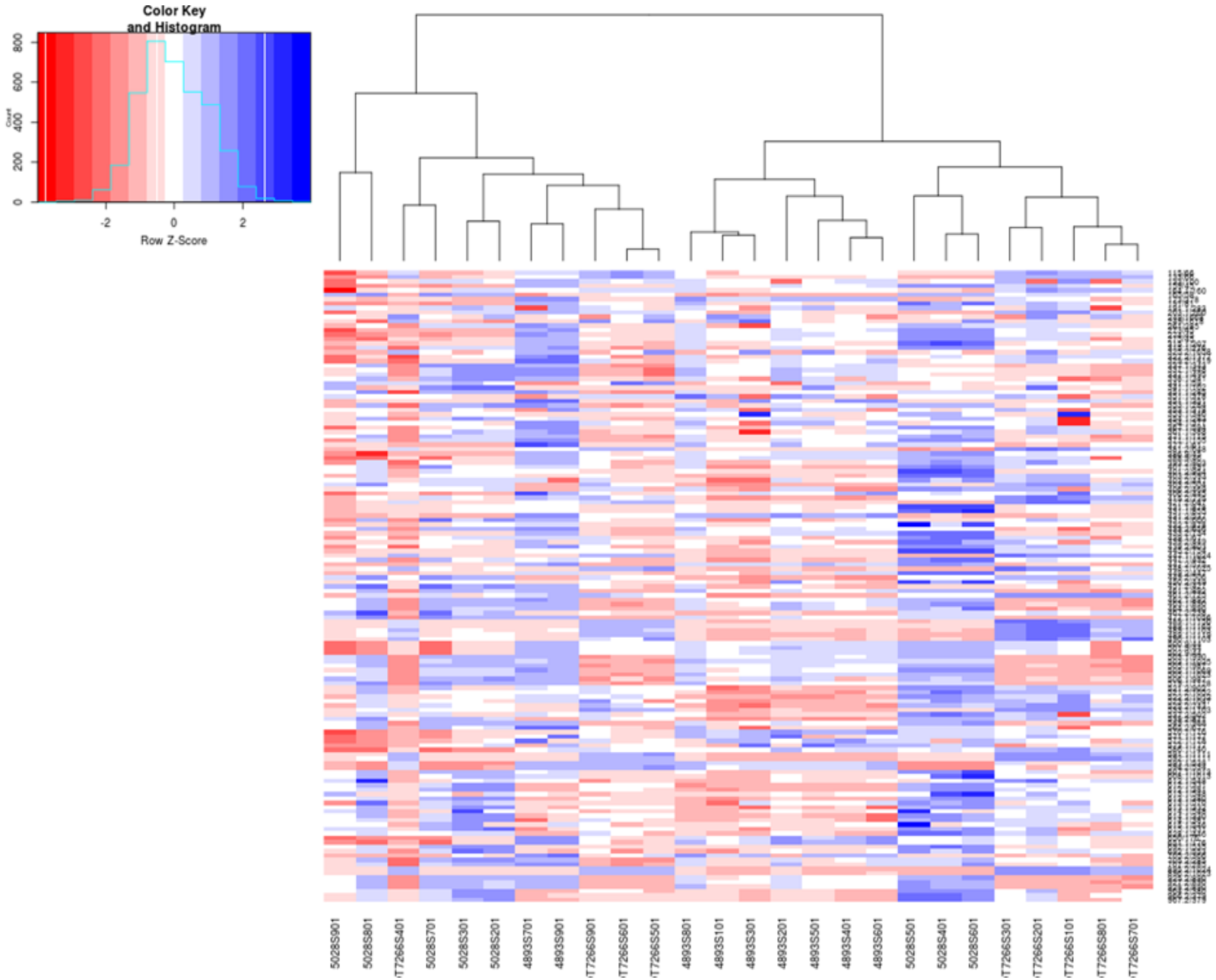


Fig. 3: Interactive heat map with metabolomics data visualisation. Metabolite features indicating significant metabolites ($p \leq 0.01$) across cultivars 5028, 4893 and TOT 7266. The scale value of each feature is represented by a red-blue colour scale. The red indicates high abundance and the blue-low abundance. Each symbolizes a metabolite feature and each row symbolizes a sample (Ivaniseck *et al*., 2014).

3.5 Annotation of metabolites

The annotation of the metabolites were provided by the METLIN metabolic database which produced a list of putative identities based on accurate mass measurements (http://metlin.scripps.edu). Furthermore, confirmation of these annotations was implemented by generating fragmentation patterns from extracted ion chromatograms (EIC) in negative ion (Ncube *et al*., 2014; Makita *et al*., 2016). A total of 15 m/z ions were tentatively annotated as shown in Table 1 below. The annotated molecules were flavonoid molecules and glucosinolates. Interestingly, a recent report by Rodriguez-Perez *et al*., (2015) revealed a similar qualitative profile of these compounds in *M. oleifera* extracts sampled from Madagascar.

Table 1: Tentative characterization the bioactive compounds of *M. oleifera*

Compound No.	Rt (min)	Mass (m/z)	Annotation	Molecular Formula	Database
1	8.69	593.1465	Apigenin 6,8 C-dihexose	$C_{27}H_{30}O_{15}$	Metlin/ Tandem MS
2	9.98	651.1537	Kaempferol acetyl dihexose	$C_{29}H_{32}O_{18}$	Metlin/ Tandem MS
3	14.21	463.0844	Quercetin Hexose	$C_{12}H_{20}0_{12}$	Metlin/ Tandem MS
4	16.41	607.1299	Quercetin hydroxy-methylglutaroyl hexose	$C_{27}H_{28}O_{16}$	Metlin/ Tandem MS
5	16.46	505.0932	Quercetin acetyl-hexose	$C_{23}H_{21}O_{13}$	Metlin/ Tandem MS
6	16.62	447.0917	Kaempferol hexose	$C_{21}H_{20}0_{11}$	Metlin/ Tandem MS
7	17.17	477.1031	Isorhamnetin hexose	$C_{22}H_{22}O_{11}$	Metlin/ Tandem MS
8	17.35	549.0897	Quercetin malonyl hexose	$C_{23}H_{22}0_{11}$	Metlin/ Tandem MS
9	18.10	591.1358	Kaempferol hydroxy-methylglutaroyl hexose	$C_{27}H_{28}O_{15}$	Metlin/ Tandem MS
10	18.11	489.0955	Kaempferol acetyl-hexose	$C_{23}H_{22}O_{11}$	Metlin/ Tandem MS
11	18.48	533.0917	Kaempferol malonyl hexose	$C_{24}H_{22}O_{14}$	Metlin/ Tandem MS
12	18.41	519.1108	Isorhamnetin acetyl-hexose	$C_{24}H_{24}O_{13}$	Metlin/ Tandem MS
13			Glucomoringin		Metlin
14			4-(2′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate		

15	4-(3′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate

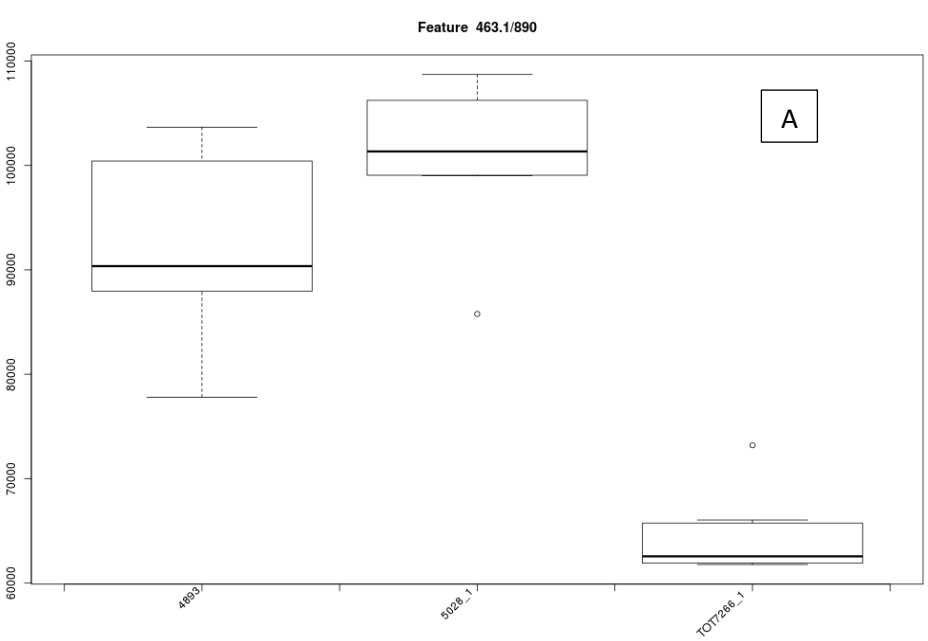
3.5.1 Comparison of annotated metabolites across the three *Moringa* cultivars

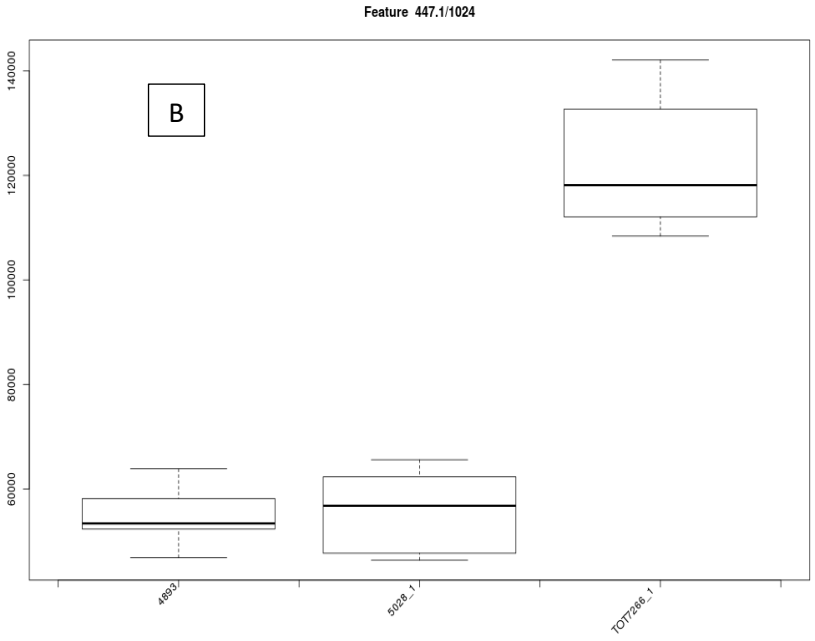
3.5.1.1 Flavonoids

Flavonoid glycoside molecules were tentatively annotated in the three cultivars and characterised (as shown in Table 1). Flavonoid compounds are biosynthesized through the phenylpropanoid pathway (Marques and Farrah, 2009; Sato *et al*., 2011). Flavonoids are neutraceutically and pharmacologically relevant to human health which arises from the antioxidant activities of these polyphenolic acids (Yao *et al*., 2004). Box and whiskers plots were generated for the relative concentration and differentiation of the annotated metabolites across the three cultivars (Fig. 4, 5 and 6).

According to the box-whiskers plot, the distribution of the flavonoid glycosides identified varied amongst the cultivars. Compounds such as quercetin hexose (m/z 463) and kaempferol hexose (m/z 447) had the highest relative peak intensity amongst cultivars TOT7266 and 5028 (Fig. 4). According to Watson *et al*., (2014), these compounds are found to be the most ubiquitous in foods, with quercetin having the ability to neutralise free radicals. Interestingly, when comparing the distribution of quercetin hexose and kaempferol hexose, the cultivars have contrasting differences in relative concentration. For instance, the concentration of quercetin hexose in 5028 was higher compared to the rest of the cultivars, but in kaempferol hexose, the concentration was lower compared to the rest of the cultivars. The same applied when comparing the concentration of quercetin acetyl hexose (m/z 505) and kaempferol acetyl hexose (m/z 489) across the three cultivars (Fig. 5). A similar scenario was reported in strawberry fruits where the content of quercetin-3-glucuronide was significantly higher than flavonols kaempferol 3-glucuronide (Aaby *et al*., 2007). On the other hand, we see that the opposite is true for the compound vicenin-2 (known as apigenin 6, 8-di-C-glucoside) as cultivar 5028 has the lowest concentration of this compound compared to others (Fig. 4). Vicenin-2 is known to treat human skin from UV-radiations (Das *et al*., 2013), which may be useful in the cosmetic industry. It also has the potential to be a pharmacological agent due to its anti-inflammatory, antioxidant and anti-tumor effects. Cultivar 5028, therefore, is not suitable for cosmetic applications, but rather cultivar 4893 which exhibited the highest amount of vicenin-

2. The variation in the flavonoid content amongst cultivars was opposite to what was reported by Ndlhala *et al*., (2014), which showed cultivars TOT5028 and TOT4893 exhibiting the lowest flavonoid content amongst the twelve cultivars, however, according to these results, the variation differs amongst the cultivars.


Feature 463.1/890
A
110000
100000
90000
80000
70000
60000
4893_
5028_1
TOT7266_1


Feature 447.1/1024
B
140000
120000
100000
80000
60000
4893_
5028_1
TOT7266_1

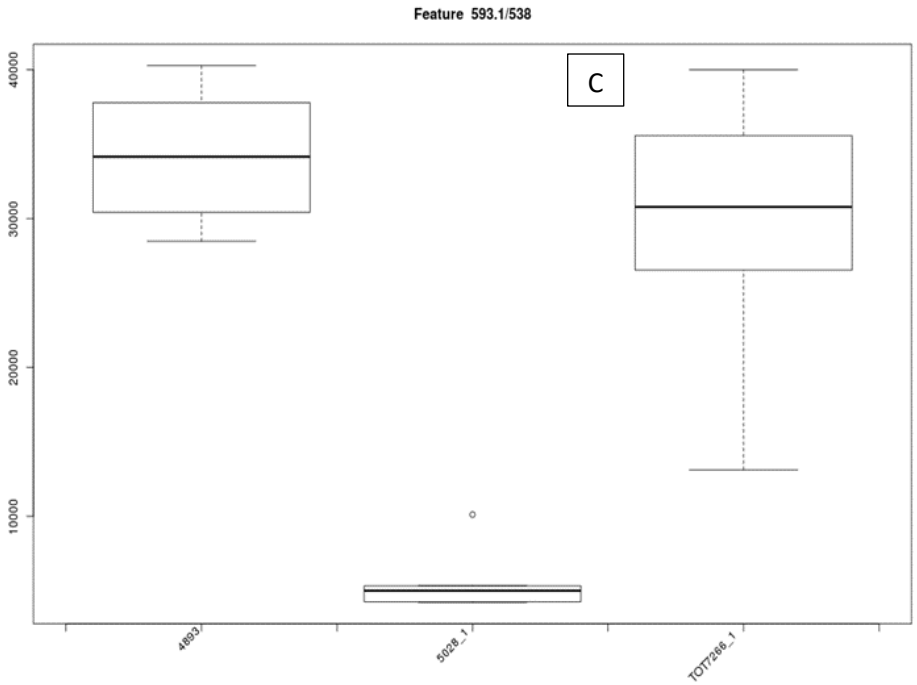


Fig. 4: Box and whiskers plots of the relative peak areas of annotated flavonoid biomarkers: Shown are the relative concentration of (A) Quercetin hexose; (B) Kaempferol hexose; (C) Vicenin-2 (Apigenin).

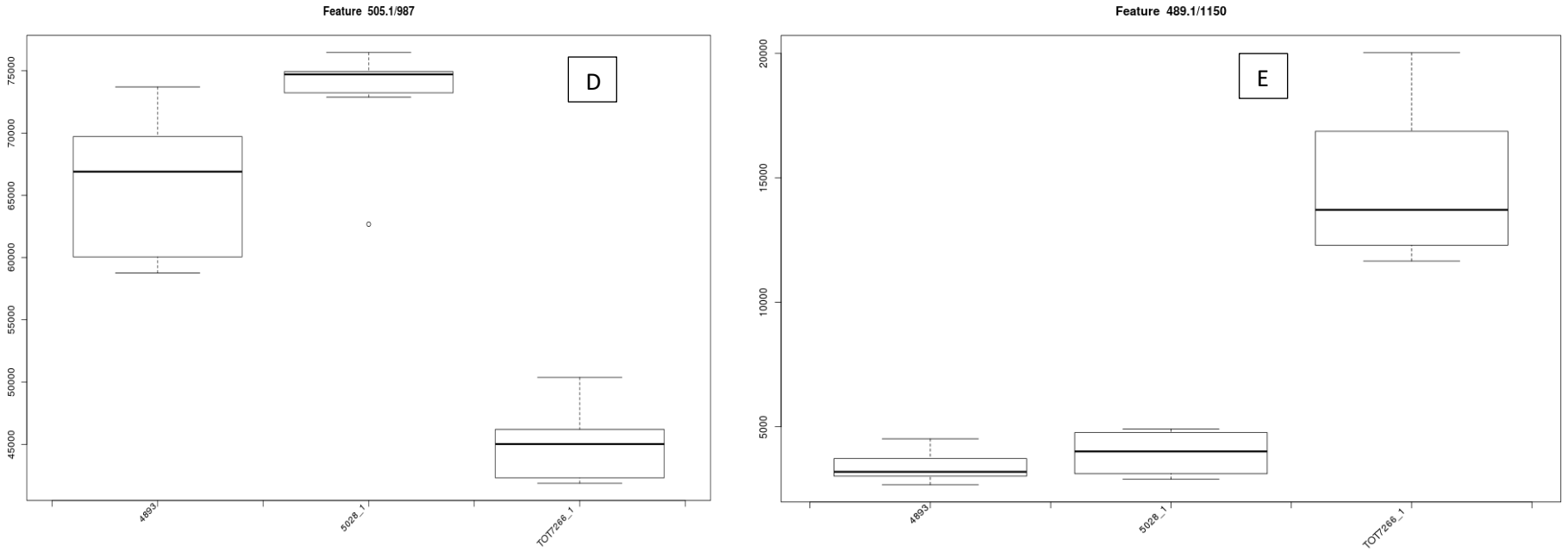


Fig. 5: Box and whiskers plots of the relative peak areas of annotated flavonoid biomarkers: Shown are the relative concentration of (D) Quercetin acetyl hexose; (E) Kaempferol acetyl hexose.

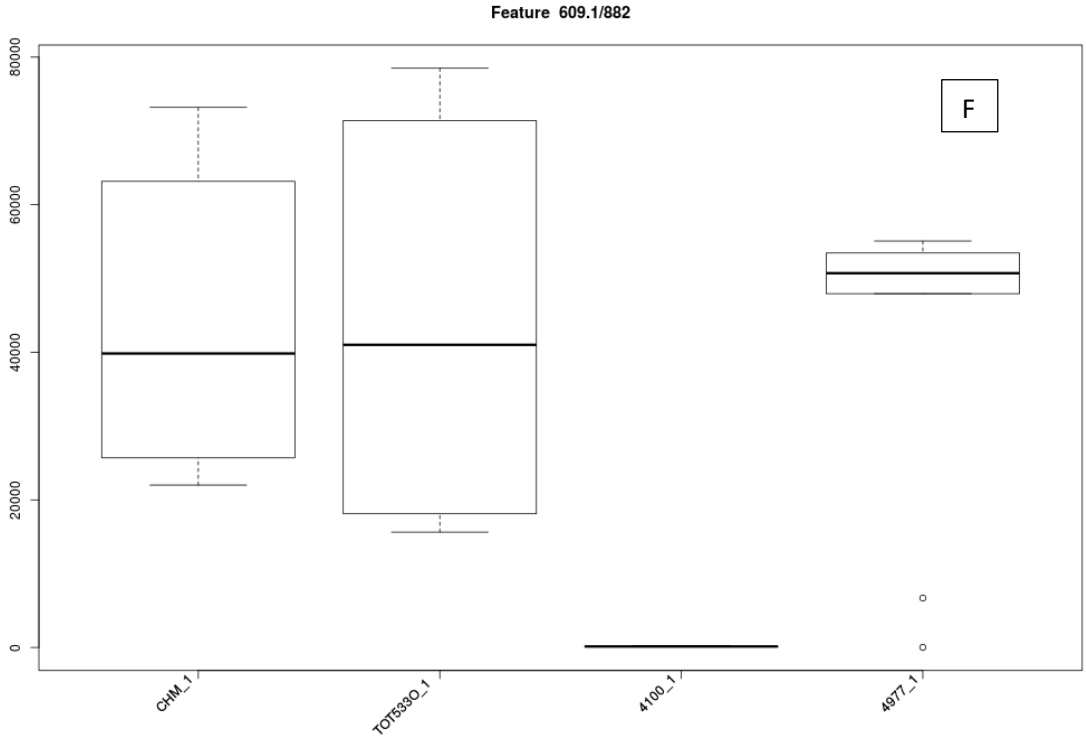


Fig. 6: Box and whiskers plots of the relative peak areas of annotated flavonoid biomarkers: Shown are the relative concentration of (F) Quercetin rutinoside.

According to previous studies, the main compound responsible for the antioxidant activity in *Moringa* and its various other medicinal properties is primarily caused by rutin (quercetin rutinoside). Similarly in elderberries, rutin was found to be the most abundant amongst the flavonols identified (Habtemariam *et al*., 2015 and Mudgea *et al*., 2016). In our previous work, rutin as well as other rutinoside bearing compounds, such as kaempferol rutinoside and isorhamnetin rutinoside, were considered as chemotaxonomical biomarkers for the identification of superior cultivars for pharmacological purposes. With the aid of box-whiskers plots cultivars TOT4977, CHM and TOT5330 were considered to be the superior cultivars amongst the twelve cultivars and 4100 the weakest due to the absence of rutin. The absence of rutin in cultivars 5028, 4893 and TOT7266 indicated that they are also not suitable for pharmacological purposes and don't fall under the category of superior cultivars. Interestingly, a report by Mudgea *et al*., (2016) indicated an elderberry sample (sample 425) contained a very small amount of rutin and a considerable amount of quercetin. Authors attributed the degradation of rutin was responsible for the high content of quercetin, of which we found was

true in cultivars 4100, 5028, 4893 and TOT7266. This further indicates that these varieties could be genetically similar (Mudgea *et al*., 2016).

Cultivars 4893, 5028 and TOT7266 have different phytochemical profiles. On average, their total relative flavonoid distribution is lower in comparison to the rutinoside bearing cultivars mentioned in Chapter 4. Researchers have found that plants grown in warmer climates tend to have a higher flavonol content (Mudgea *et al*., 2016), however, the seedlings of the rutinoside bearing cultivars were grown under the same environmental conditions and so the significant differences in their high flavonoid profile is based on genotype that is influencing the high metabolic profile of the rutinoside bearing cultivars.

3.5.1.2 Glucosinolates

Glucosinolates are sulfur-rich secondary metabolites that are primarily known for their anti-cancer properties (Malabed and Noel, 2013). According to Table 1, three significant glucosinolates such as glucomoringin (m/z 570.176), 4-(2′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate (m/z 612.1043) and 4-(3′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate (m/z, m/z 612.1087) were found to be the highest metabolites identified in cultivars 4893, 5028 and TOT7266 were identified. The relative concentration were confirmed using box and whiskers plots shown in Fig. 7 and Fig. 8.

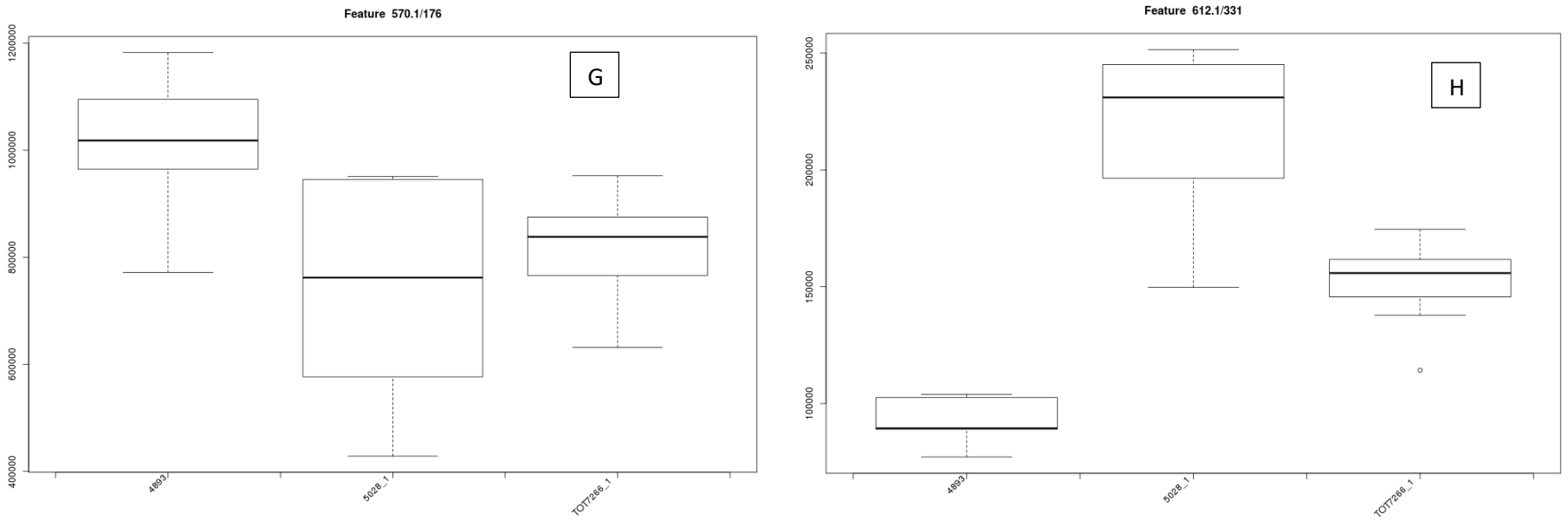


Fig. 7: Box and whiskers plots of the relative peak areas of annotated glucosinolate biomarkers: Shown are the relative concentration of (G) Glucomoringin; (H) 4-(2′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolates.

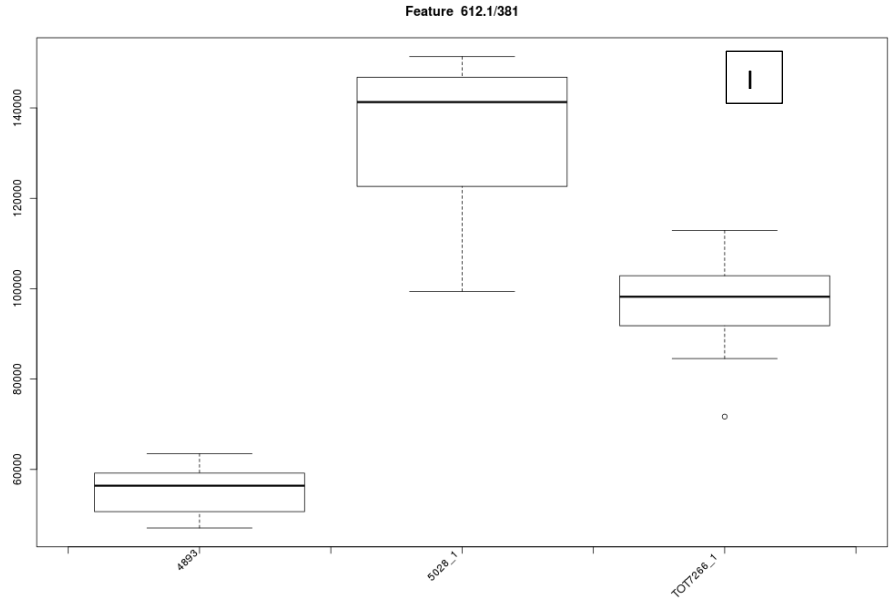


Fig. 8: Box and whiskers plots of the relative peak areas of annotated glucosinolate biomarkers: Shown are the relative concentration of (I) 4-(3′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate.

According to the box and whiskers plots, glucomoringin was found to be the highest glucosinolate and was found in cultivar 4893 which would explain the tight clustering and grouping of this cultivars illustrated in the HCA and PCA score plots in (Fig.1 and Paper III, Fig. 7 and 8). The presence of glucomoringin as well as its acetyl derivatives are known to possess many biological activities such as antimicrobial, antiproliferative and anticarcinogentic properties (Cheenpracha *et al*., 2010; Maldini *et al*., 2014). Cultivar 5028 is capable of producing high yields of 4-(2′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate and 4-(3′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate (Fig. 9). According to Martinez-Ballesta *et al*., (2013), environmental conditions such as abiotic stress, salinity, drought, extreme temperatures, light and nutrient deprivation was considered to increase the glucosinolate profiles of Brassica *sp*. These results further suggests that cultivar 4893 could be considered for the treatment of certain diseases and ailments, however, when it was compared to the rutinoside bearing cultivars, results showed that they contained higher glucosinolate concentration, more so in cultivar 4977. Interestingly, previously we had considered cultivar 4100 as the weakest cultivar, however, in this current study where the focus was not only on flavonoids, the cultivar was found to contain a moderately high concentration of 4-(2′-O-

Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate and 4-(3′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate as shown in the figures below (Fig. 9).

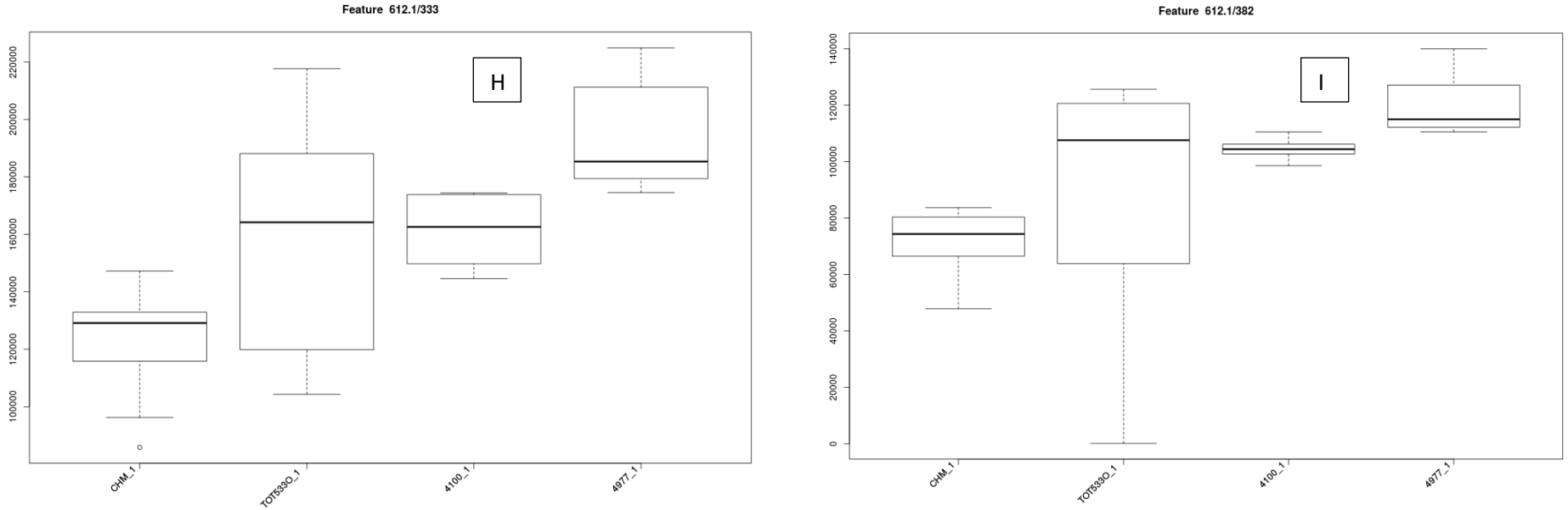


Fig. 9: Box and whiskers plots of the relative peak areas of annotated glucosinolate biomarkers: Shown are the relative concentration of (H) 4-(2′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolates; (I) 4-(3′-O-Acetyl-α-L-rhamnopyranosyloxy) benzyl glucosinolate.

The information in the data implies that the production of highly sought after compounds, for instance, for cosmetic or pharmaceutical applications is cultivar specific. The cultivars used in this study were grown in the same conditions after collecting the seeds from different geographical locations. Therefore, it is expected that the metabolomic differences amongst species observed could not be due to environmental factors such as climate, soil condition and biotic stress, but rather due to the inherit character of each cultivar, that is, their genetic composition (Da *et al*., 2015). Due to the genetic diversity exhibited from these cultivars, this information can allow plant breeders to produce more useful and productive *Moringa* crop varieties by exploiting genes within the plant species itself.

Although many plant metabolomic studies have been implemented highlighting genotype, climatic variables and cultivation methods, there hasn't been many studies demonstrating the relationship between bioactive metabolites and numerous cultivars in a plant species like *Moringa*. Fujimura *et al*., (2011) implemented a similar metabolomics driven study amongst diverse green cultivars. This study illustrates the effectiveness of metabolic profiling by using a chemometric approach with the aid of multivariate statistical analysis to determine the best suitable *Moringa* cultivar for pharmaceutical, nutritional and cosmetic purposes.

4.0 Conclusion

The current study demonstrates chemotaxonomy as a powerful tool for investigating superior cultivars based on genetic diversity. Based on our results, we can see that the cultivars investigated have significantly different metabolomes. From our results, we can specify differential properties between cultivars using statistical platforms which can be utilised for future application for identification of superior cultivars. Importantly, with all the platforms utilised in the study, profiles of cultivars 5028, 4893 and TOT7266 and rutinoside bearing cultivars were found to differ. Thus, biomarkers such as flavonoids and glucosinolates can be used as potential biomarkers for selection of superior cultivars.

5.0 References

Aaby, K., Ekeberg, D., Skrede, G. (2007). Characterization of phenolic compounds in strawberry fruits by different HPLC detectors and contribution of individual compounds to total antioxidant capacity. *Journal of Agricultural Food Chemistry*. 55: 4395-4406.

Benton, H.P., Wong, D.M., Trauger, S.A., Siuzdak, G. (2008). XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. *Analytical Chemistry*. 80: 6382- 6389.

Canter, P.H., Thomas, H, Ernst E. (2005). Bringing medicinal plants into cultivation: opportunities and challenges for biotechnology. *Trends in Biotechnology*. 2005: 23.

Cheenpracha, S., Park, E.J., Yoshida, W. Y., Barit, C., Wall, M., Pezzuto, J. M., Chang, L.C. (2010). Potential anti-inflammatory phenolic glycosides from the medicinal plant *Moringa oleifera* fruits. *Bioorganic Medicinal Chemistry*. 18: 6598–6602.

Da, H., Gu, X.J., Xiao, P.G. (2015). Medicinal plants: Chemistry, biology and omics. Woodhead publishing series in Biomedcine 73 Elsevier. Langford Lane, Kidlington, UK.

Das, S., Das, J., Paul, A., Samadder, A., Khuda-Bukhsh, A.R. (2013). Apigenin, a bioactive flavonoid from Lycopodium clavatum, stimulates nucleotide excision repair genes to protect kkin keratinocytes from ultraviolet B-Induced reactive oxygen species and DNA Damage. *Journal of Acupuncture and Meridian* Studies. 6.

Farag, M.A., Mohsen, M., Heinke, R., Ludger A.Wessjohann, L.A. (2014). Metabolomic fingerprints of 21 date palm fruit varieties from Egypt using UPLC/PDA/ESI–qTOF-MS and GC–MS analyzed by chemometrics. *Food Research International*. 64: 218–226.

Förster, N., Ulrichs, C., Schreiner, M., Arndt, N., Schmidt, R., Mewis, I. (2015). Ecotype variability in growth and secondary metabolite profile in *Moringa oleifera*: Impact of Sulfur and Water Availability. *Journal of Agriculture and Food Chemistry*. 63: 2852–2861.

Fujimura, Y., Kurihara, K., Ida, M., Kosaka, R., Miura, D., Hiroyuki Wariishi, H., Maeda-Yamamoto, M., Nesumi, A., Saito, T., Kanda, T., Yamada, K., Tachibana, H. (2011). Metabolomics-driven nutraceutical evaluation of diverse green tea cultivars. *PLoS One*. 6.

Gowda, H., Ivanisevic, J., Johnson, C.H., Kurczy, M.E., Benton, H.P., Rinehart, D., Nguyen, T., Ray, J., Kuehl, J., Arevalo, B., Westenskow, P.D., Wang, J., Arkin, A.P., Deutschbauer, A.M., Patti, G.J., Siuzdak, G. (2014). Interactive XCMS Online: Simplifying advanced metabolomic data processing and subsequent statistical analyses. *Analytical Chemistry*. 86: 6931-6939.

Habtemariam, S., Varghese, G.K. (2015). Extractability of Rutin in herbal tea preparations of *Moringa stenopatala* leaves. *Beverages*. 1: 169-182.

Ivanisevic, J., Stauch, K.L., Petrascheck, M., Benton, H.P., Epstein, A.A., Gorantla, S., Tran, M., Hoang, L., Kurczy, M.E., Michael, D., Gendelman, H.E., Fox, H.S., Siuzdak, G. (2016). Metabolic drift in the aging brain. *Aging*. 8: 1–21.

Kwon, Y., Myung, S.A., Park , J.S., Liu , J.R., In, D.S., Min, B.W., Suk Weon Kim, S.W. (2014). Discrimination of cultivation ages and cultivars of ginseng leaves using Fourier transform infrared spectroscopy combined with multivariate analysis. *Journal of Ginseng Research*. 38: 52-58.

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: 254.

Makkar, H.P.S., Becker, K. (1996). Nutritional value and anti- nutritional components of whole and ethanol extracted *Moringa oleifera* leaves. *Animal Feed Science and Technology*. 63: 211–228.

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.

Malabed, R.S., Noel, M.G. (2013). Characterization of the glucosinolates and isothiocyanates in Malunggay (*Moringa oleifera* L.) extracts and determination of their myrosinase activity and anticancer properties. Conference presentation presented at the Research Congress. De La Salle University Manila March, 7-9.

Maldini, M., Maksoud, S., Natella, F., Montoro, P., Petretto, G.L., Foddai, M., Nicola, G.R., Chessa, Pintore, G. (2014). *Moringa oleifera*: A study of phenolics andglucosinolates by mass spectrometry. J*ournal of Mass Spectrometry*. 49: 900-910.

Marques, V., Farah, A. (2009). Chlorogenic acids and related compounds in medicinal plants and infusions. *Food Chemistry*. 113: 1370–1376.

Martínez-Ballesta, M., Moreno, D.A., Carvajal, M. (2013). The physiological importance of glucosinolates on plant response to abiotic stress in Brassica. *International Journal of Molecular Science.* 14: 11607-11625.

Mhlongo, M.I., Piater, L.A., Madala, N.E., Paul A. Steenkamp, P.A., Ian A. Dubery, I.A. (2016). Phenylpropanoid defences in Nicotiana tabacum Cells: Overlapping metabolomes indicate common aspects to priming responses induced by lipopolysaccharides, chitosan and flagellin-22. *PLOS ONE*. 11.

Mudgea, E., Applequistb, W.L., Finleya, Listera, J.P., Townesmithb, A.K., Walkerb, K.M., Browna, P.N. (2016). Variation of select flavonols and chlorogenic acid content of elderberry collected throughout the Eastern United States. *Journal of Food Composition and Analysis*. 47: 52–59.

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: 66.

Ndhlala, A.R., Mulaudzi, R., Ncube, B., Abdelgadir, H.A., du Plooy, C.P., Van Staden, J. (2014). Antioxidant, antimicrobial and phytochemical variations in thirteen *Moringa oleifera* Lam. cultivars. *Molecules*. 19: 10480-10494.

Ni, Y., Su, M., Lin, J., Wang, X., Qiu, Y., Zhao, A. (2008). Metabolic profiling reveals disorder of amino acid metabolism in four brain regions from a rat model of chronic unpredictable mild stress. *FEBS Letters*. 2008: 582: 2627–2636.

Pakade, V., Cukrowska, E., Chimuka, L. (2013). Metal and flavonol contents of *Moringa oleifera* grown in South Africa. *South African Journal of Science*. 109: 1-7.

Patti, G.J., Tautenhahn, R., Rinehart, D., Cho, K., Shriver, L.P., Manchester, M., Nikolskiy, I., Johnson, C.H., Mahieu, N.G., Siuzdak, G. (2013). A view from above: cloud plots to visualize global metabolomic data. *Analytical Chemistry*. 85: 798-804.

Rodriguez-Perez, C., Quirantes-Pine, R., Fernandez-Gutierrez, A., Segura-Carretero, A. (2015). Optimization of extraction method to obtain a phenolic compounds-rich extract *Moringa oleifera* Lam leaves. *Industrial Crops and Products.* 66: 246-254.

Sadeghi-bazargani, H., Bangdiwala, S.I., Mohammad, K. (2011). Compared application of the new OPLS-DA statistical model versus partial least squares regression to manage large numbers of variables in an injury case-control study. *Scientific Research and Essays*. 6: 4369–4377.

Sato, Y., Itagaki, S., Kurokawa, T., Ogura, J., Kobayashi, M., Hirano, T., Sugawara, M., & Iseki, K. (2011). In vitro and in vivo antioxidant properties of chlorogenic acid and caffeic acid. *International Journal of Pharmaceutics*. 403: 136–138.

Watson, R.R. (2014). Polyphenols in plants: Isolation, purification and extract preparation. Academic Press, Elsevier. Tucson, AZ. USA.

Yao, L.H., Jiang, Y.M., Shi, J., Tomas-Barber, F.A., Datta, N., Singanusong, R., Chen, S.S. (2004). Flavonoids in food and their health benefits. *Plant Foods for Human Nutrition*. 59: 113–12.