

HPLC Method Development for Characterisation of the Phenolic Composition
of *Cyclopia subternata* and *C. maculata* Extracts and Chromatographic
Fingerprint Analysis for Quality Control

by

Alexandra Elizabeth Schulze



*Thesis presented in partial fulfilment of the requirements for the degree of
Master of Science in Food Science*

Department of Food Science
Faculty of AgriSciences
Stellenbosch University

Supervisor: Dr D. de Beer

Co-supervisors: Prof A. de Villiers, Prof M. Manley and Prof E. Joubert

December 2013

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2013

Abstract

The phenolic composition of *Cyclopia* species is believed to be partially responsible for the numerous health promoting properties associated with their extracts. Current quality control measures do not accommodate variation in phenolic profiles of *Cyclopia* species. In this study, comprehensive high performance liquid chromatography (HPLC) methods were developed for the improved characterisation of the phenolic composition of aqueous extracts of two *Cyclopia* species (*C. subternata* and *C. maculata*). The methods were developed to be suitable for both routine quantitative analysis on conventional HPLC instrumentation, and the construction of chromatographic fingerprints for further data analysis. The latter entailed similarity analysis and prediction of total antioxidant capacity (TAC).

Using a methodical approach, two separate HPLC methods, using diode array detection (DAD), were developed and validated for the analysis of aqueous extracts prepared from unfermented (green) and fermented plant material of *C. subternata* and *C. maculata*. Separation was achieved using the same method parameters (column, temperature, mobile phases), except for differing mobile phase gradients. Hyphenation of the developed HPLC methods with mass spectrometry (MS) and tandem MS allowed the confirmation of phenolic compounds previously identified in *Cyclopia*, and the tentative identification of several additional compounds in *Cyclopia* species, which are reported here for the first time. These included apigenin-6,8-di-*C*-glucoside, 3-hydroxyphloretin-3',5'-di-*C*-hexoside, eriodictyol-di-*C*-glucoside, iriflophenone-di-*O,C*-hexoside, hydroxymangiferin and hydroxyisomangiferin. Subsequently, a large number of aqueous extracts of randomly selected green *C. subternata* ($n = 64$) and *C. maculata* ($n = 50$) plant material samples were analysed. Large quantitative variations were observed on intra- and inter-species levels. *Cyclopia maculata* extracts contained almost six times more mangiferin than extracts from *C. subternata*.

HPLC-DAD analysis produced duplicate fingerprints for each extract which were consequently used for further analysis. The chromatographic fingerprint of a bioactive extract of each species was included in the respective data sets. Similarity analysis was conducted between the fingerprints from the randomly selected extracts and the corresponding active extract. For each species several extracts were determined to have similar "activity" as that of the active extract ($n = 15$ for *C. subternata* and $n = 45$ for *C. maculata*). Compounds potentially responsible for the activity were tentatively identified with the aid of principal component analysis (PCA) in combination with similarity analysis. PCA was more effective in identifying small differences between fingerprints than similarity analysis based on the correlation coefficients (r) alone.

Furthermore, multivariate data analysis was used to construct partial least squares (PLS) regression models for the prediction of TAC from fingerprint data of each species, and available data from two microplate TAC assays. The construction of the models was successful with reasonable errors ($< 10\%$), and permitted the determination of compounds of interest for future research. These included compounds of known identity that had large positive contributions toward the predictions of TAC, or unknown compounds that had small UV

signals, but relatively large positive contributions to the models.

Uittreksel

Die talle gesondheidsbevorderingseienskappe van ekstrakte van *Cyclopia* spesies word gedeeltelik geassosieer met hul fenoliese samestelling. Huidige kwaliteitskontrolemaatreëls is nie in staat om die variasie wat in die fenoliese profiele van die spesies voorkom, te akkommodeer nie. Omvattende hoë druk vloeistof chromatografiese (HPLC) metodes is vir twee *Cyclopia* spesies, naamlik *C. subternata* en *C. maculata*, in hierdie studie ontwikkel vir beter karakterisering van die fenoliese samestelling van waterekstrakte van dié spesies. Die metodes moes ook geskik wees vir roetine analise van *C. subternata* en *C. maculata* ekstrakte op konvensionele HPLC instrumentasie, en vir die opstel van chromatografiese vingerafdrukke (fenoliese samestellingsprofile) vir verdere data analise, soos gelykvormighedsanalise en die voorspelling van die totale antioksidantkapasiteit (TAC).

Twee HPLC metodes, wat van 'n ultraviolet-diode detektor (DAD) gebruik maak, is ontwikkel deur 'n sistematiese benadering te volg. Die onderskeie metodes is vir die ontleding van waterekstrakte van groen (on-gefermenteerde) en gefermenteerde plantmateriaal van *C. subternata* en *C. maculata* gevalideer. Ongeag die spesie is optimale skeiding met dieselfde kolom, mobiele fase en kolom-temperatuur bereik, maar met verskillende mobiele fase gradiënte. Analise met massaspektrometrie (MS) en tandem MS het die teenwoordigheid van fenoliese verbindings, wat voorheen in *Cyclopia* spesies geïdentifiseer is, bevestig. Verder is ook 'n aantal verbindings vir die eerste keer in *Cyclopia* tentatief geïdentifiseer. Dit sluit apigenien-6,8-di-*C*-glukosied, 3-hidroksiefloretien-3',5'-di-*C*-heksosied, eriodiktiol-di-*C*-glukosied, iriflofenoon-di-*O,C*-heksosied, hidroksiemangiferien en hidroksie-isomangiferien in. Vervolgens is 'n groot aantal ewekansig gekose waterekstrakte van beide groen *C. subternata* ($n = 64$) en *C. maculata* ($n = 50$) plantmateriaal geanaliseer, en groot kwantitatiewe variasie op intra- en inter-spesievlak waargeneem. *Cyclopia maculata* ekstrakte het byvoorbeeld byna ses maal die mangiferieninhoud van *C. subternata* ekstrakte gehad.

HPLC-DAD analise van die ekstrakte het duplikaat vingerafdrukke van elke ekstrak geproduseer, wat vir verdere data analise gebruik is. Die chromatografiese vingerafdruk van 'n bioaktiewe ekstrak van elke spesie was by die onderskeie datastelle ingesluit. Gelykvormighedsanalise is tussen vingerafdrukke van die ewekansig gekose ekstrakte en die ooreenstemmende aktiewe ekstrak uitgevoer. Vir elke spesie is 'n aantal "aktiewe" ekstrakte aangewys ($n = 15$ vir *C. subternata* en $n = 45$ vir *C. maculata*). Die verbindings wat potensieel verantwoordelik kan wees vir die aktiwiteit is met behulp van hoofkomponentontleding (PCA) in kombinasie met gelykvormighedsanalise, tentatief aangewys. PCA was egter meer effektief om klein verskille tussen vingerafdrukke aan te dui, in vergelyking met gelykvormighedsanalise wat slegs op die korrelasie koëffisiënt (r) gebaseer is.

Meerveranderlike data analise is gebruik om "gedeeltelike kleinste kwadrate" (PLS) regressiemodelle, vir die voorspelling van die TAC van beide spesies te bou. Die voorspelling is gebaseer op hul vingerafdruk data en TAC data van twee TAC mikroplaat metodes. Die model-konstruksie was suksesvol met aanvaarbare voorspellingsfoute ($< 10\%$). Verbindings van belang kon ook bepaal word. Dit sluit bekende verbindings in wat groot

positiewe bydraes ten opsigte van die voorspelling van TAC getoon het, asook ongeïdentifiseerde verbindings wat klein UV-seine getoon het, maar relatiewe groot bydraes tot die modelle gehad het.

Acknowledgements

I would like to express my most sincere gratitude to the following people:

To my supervisor, Dr Dalene de Beer, I am truly grateful for your practical support and patience throughout this study. Thank you for giving me the freedom to explore on my own, while also providing guidance when I needed it. Your attention to detail, willingness to help at any moment and technical insight were invaluable to the completion of this study.

Prof Lizette Joubert, a co-supervisor with great expertise and high standards. It has been a privilege to work with such a goal-driven and passionate researcher who enabled me to see the bigger picture at all times. It would have been impossible to complete this thesis without your guidance and persistent help.

Prof André de Villiers, a co-supervisor, thank you for your valuable input, especially with the analytical chemistry section of this project. Your guidance has enabled me to successfully complete this thesis and develop a better understanding of my work.

Prof Marena Manley, a co-supervisor who always had the confidence in my abilities to complete this work. Thank you for your kind support and motivation throughout this project. Your direction in the preparation of this thesis and interpretation of chemometric analysis is truly appreciated.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF (DST Innovation Master's Scholarship - Grant 84563).

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant 70525 to E. Joubert).

Dr Marietjie Stander (Central Analytical Facility, Stellenbosch University) for assistance with LC-MS analysis.

Prof Beata Walczak, Małgorzata Kazura and Chemometrics research group (Department of Chemometrics, University of Silesia, Poland) for guidance with pre-treatment of data, and kind hospitality during our research visit.

Dr Christie Malherbe (Post Harvest and Wine Technology, ARC Infruitec-Nietvoorbij) for the ORAC analysis and for providing technical insight and encouragement. "Die nag kom nader!"

Gideon van der Kolf for providing practical assistance with data handling and MATLAB, and endless encouragement, love and support.

Neliswa Matrose for conducting the DPPH analysis and precision tests.

Marieta van der Rijst (Biometry Unit, ARC Infruitec-Nietvoorbij) for normal distribution analysis.

My fellow students and colleagues, Alet Venter, Brigitte du Preez, Ilona Steenkamp, Theresa Beelders & Stephanie Bosman, thank you for your kind support, advice and motivation.

Family and friends, I am very grateful for your continuous encouragement and support throughout my studies.

Contents

Abstract	iii
Uittreksel	v
Acknowledgements	vii
List of Figures	xii
List of Tables	xvi
1 Introduction	1
2 Literature Review	9
2.1 Introduction	10
2.2 Honeybush tea	10
2.2.1 Honeybush as commercial product	10
2.2.2 Quality	12
2.2.3 Chemical composition of honeybush	13
2.3 High performance liquid chromatography	16
2.3.1 Introduction	16
2.3.2 Method development	17
2.3.3 Method validation	18
2.3.4 Detection methods	18
2.3.4.1 Diode array detection	18
2.3.4.2 Mass spectrometry	21
2.3.5 Identification of phenolic compounds	21
2.3.5.1 Elution order	21
2.3.5.2 UV-Vis detection	22
2.3.5.3 MS and MS/MS detection	22
2.3.5.4 Aglycones	23
2.3.5.5 Flavonoid glycosides	25
2.4 Fingerprints	26
2.4.1 Introduction	26
2.4.2 Chromatographic fingerprints	27
2.5 Data analysis of fingerprints	27

2.5.1	Pre-treatment of chromatographic fingerprints	27
2.5.1.1	Alignment	28
2.5.1.2	Normalisation	31
2.5.1.3	Signal-to-noise ratio improvement	31
2.5.1.4	Baseline elimination	31
2.5.2	Multivariate data analysis	32
2.5.2.1	Unsupervised data analysis	33
2.5.2.2	Supervised data analysis	33
2.5.3	The effect of pre-treatment techniques on multivariate data analysis	34
2.5.4	Application of chemometrics to fingerprint analysis	36
2.6	Conclusions	38
3	Development of HPLC methods for the analysis of <i>Cyclopia subternata</i> and <i>Cyclopia maculata</i> phenolics	45
3.1	Abstract	46
3.2	Introduction	47
3.3	Experimental	47
3.3.1	Chemicals	47
3.3.2	Plant material	48
3.3.3	Preparation of aqueous honeybush tea extracts for analyses	48
3.3.4	HPLC-DAD method development	48
3.3.5	Quantification of phenolic compounds using HPLC-DAD	49
3.3.6	Identification of phenolic compounds using LC-DAD-MS and -MS/MS detection	50
3.3.7	Method validation	50
3.4	Results and discussion	51
3.4.1	HPLC-DAD method development	51
3.4.2	Identification of phenolic compounds in two <i>Cyclopia</i> species	54
3.4.3	Method validations	63
3.4.4	Differences in phenolic composition of green honeybush seedlings	65
3.5	Conclusions	69
4	HPLC chromatographic fingerprint analyses of <i>Cyclopia subternata</i> and <i>Cyclopia maculata</i> extracts: Similarity to biologically active extracts and identification of potentially active compounds	72
4.1	Abstract	73
4.2	Introduction	74
4.3	Experimental	75
4.3.1	Plant material and extract preparation	75
4.3.2	HPLC analysis of <i>C. subternata</i> and <i>C. maculata</i> extracts	76
4.3.3	Pre-treatment of chromatograms	76
4.3.4	Similarity analysis	78
4.3.5	Principal component analysis	78

4.4	Results and discussion	79
4.4.1	Pre-treatment	79
4.4.2	Similarity analysis	79
4.4.3	Principal component analysis	84
4.5	Conclusions	91
5	Predicting the total antioxidant capacity of <i>Cyclopia</i> spp. extracts using HPLC chromatographic fingerprints	94
5.1	Abstract	95
5.2	Introduction	96
5.3	Experimental	97
5.3.1	Plant material and HPLC analysis	97
5.3.2	Pre-treatment of fingerprints	97
5.3.3	Total antioxidant capacity assays	97
5.3.4	Data analysis	98
5.3.4.1	Data exploration and outlier detection	98
5.3.4.2	Calibration and validation subset division	98
5.3.4.3	Partial least squares	98
5.4	Results and discussion	101
5.4.1	Pre-treatment of chromatographic data	101
5.4.2	Data exploration and outlier detection	101
5.4.3	Calibration and validation subset division	103
5.4.4	Partial least squares models	103
5.5	Conclusions	111
6	General Discussion and Conclusions	115
A	Quantitative data of <i>Cyclopia subternata</i> and <i>Cyclopia maculata</i>	126
B	Typical and ‘atypical’ <i>C. maculata</i> extract	129

List of Figures

2.1	(a) Natural and (b) cultivated distribution of honeybush species (Joubert <i>et al.</i> , 2011).	11
2.2	UV-Vis spectral properties of the different phenolic classes present in honeybush (a) benzophenone (iriflophenone-3- <i>C</i> -glucoside), (b) dihydrochalcone (phloretin-3',5'-di- <i>C</i> -glucoside), (c) flavanone (eriocitrin), (d) flavone (scolymoside) and (e) xanthone (mangiferin).	22
2.3	Nomenclature used for fragmentation of flavonoid glycosides in mass spectrometry (Domon & Costello, 1988; Cuyckens & Claeys, 2004; Abad-García <i>et al.</i> , 2009).	23
2.4	Structures and fragmentation pathways for selected phenolic aglycones, indicated for positive ionisation mode (Abad-García <i>et al.</i> , 2009; Yu <i>et al.</i> , 2013; Zhang <i>et al.</i> , 2013). Carbon numbering for all classes are similar to indicated for flavones, with the exception of dihydrochalcones (indicated separately).	24
2.5	The main fragmentation pathways observed for <i>O</i> -glycosidic and <i>O</i> -diglycosidic flavonoids (De Rijke <i>et al.</i> , 2006; Abad-García <i>et al.</i> , 2009).	25
2.6	The main fragmentation pathways observed for <i>C</i> -glycosidic flavonoids (Waridel <i>et al.</i> , 2001; Cuyckens & Claeys, 2004; Abad-García <i>et al.</i> , 2009).	26
2.7	Example of signal alignment (Zhang <i>et al.</i> , 2012).	28
2.8	(a) Graphic illustration of DTW principle; (b) Illustration of determined minimal cumulative distances (Pravdova <i>et al.</i> , 2002).	30
2.9	Example of signal-to-noise improvement (Daszykowski <i>et al.</i> , 2009).	32
2.10	Example of baseline elimination of a chromatographic signal (Daszykowski & Walczak., 2006).	32
2.11	PCA score plots (a) before and (b) after alignment. Different symbols represent different plant species (Gong <i>et al.</i> , 2004).	34
2.12	HPLC-DAD chromatograms of <i>Turnera diffusa</i> extracts (a) before and (b) after alignment (Lucio-Gutiérrez <i>et al.</i> , 2012).	35
2.13	PCA score plots for HPLC-DAD analysis of <i>Turnera diffusa</i> extracts (a) before and (b) after alignment (Lucio-Gutiérrez <i>et al.</i> , 2012).	35
2.14	PLS regression plots for HPLC-DAD analysis of <i>Turnera diffusa</i> extracts (a) before and (b) after alignment (Lucio-Gutiérrez <i>et al.</i> , 2012).	36
3.1	Gradient profiles used for <i>C. subternata</i> and <i>C. maculata</i> analyses [Mobile phases: 2% acetic acid (A) and acetonitrile (B)].	49
3.2	Chromatograms for the separation of freeze-dried aqueous extracts of (a) green and (b) fermented <i>Cyclopia subternata</i> phenolic compounds and (c) a standard calibration mixture [See Table 3.1 and Table 3.2 for identities of numbered peaks; a aspalathin, b luteolin].	52

3.3	Chromatograms for the separation of freeze-dried aqueous extracts of (a) green and (b) fermented <i>Cyclopia maculata</i> phenolic compounds and (c) a standard calibration mixture [See Table 3.3 for identities of numbered peaks].	53
3.4	Structures of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented <i>C. subternata</i> and/or <i>C. maculata</i> [* indicates that the position of hexosyl or hydroxyl moieties are not certain].	58
3.5	LC-MS/MS spectrum of 3-hydroxyphloretin-3',5'-di- <i>C</i> - β -hexosyl (14) obtained in positive ionisation mode.	60
3.6	LC-MS/MS spectrum of unidentified compound 20 obtained in negative ionisation mode.	63
3.7	Phenolic composition (g/100 g dried extract) of freeze-dried aqueous extracts of green <i>C. subternata</i> and <i>C. maculata</i> [Bars indicate the mean values \pm standard deviation. Average values are indicated on top of bars with relative standard deviation in brackets; n.q., compounds not quantified; glc, glucoside].	68
4.1	Schematic representation of data handling and pre-treatment [SA = similarity analysis; PCA = principal component analysis].	77
4.2	(a) <i>C. subternata</i> and (b) <i>C. maculata</i> chromatograms before pre-treatment of data.	80
4.3	(a) <i>C. subternata</i> and (b) <i>C. maculata</i> chromatograms after pre-treatment of data.	81
4.4	Heat maps of correlation coefficients between samples of <i>C. subternata</i> (a) before and (b) after fingerprint data pre-treatment (duplicate injections are shown as separate measurements).	82
4.5	Heat maps of correlation coefficients between samples of <i>C. maculata</i> (a) before and (b) after fingerprint data pre-treatment (duplicate injections are shown as separate measurements).	82
4.6	Heat map of correlation coefficients of <i>C. subternata</i> extracts against the bioactive <i>C. subternata</i> extract (duplicate injections are shown as separate measurements).	83
4.7	Comparison of chromatograms of an extract that correlated (a) well ($r > 0.95$) and an extract that correlated (b) poorly ($r < 0.70$) with the chromatogram of the active <i>C. subternata</i> extract.	83
4.8	Heat map of correlation coefficients of <i>C. maculata</i> extracts against a bioactive <i>C. maculata</i> extract (duplicate injections for each sample are included) (n = 82; set 1).	84
4.9	Heat map of correlation coefficients of additional wild harvested <i>C. maculata</i> extracts together with the previous extracts against the bioactive <i>C. maculata</i> extract (duplicate injections are shown as separate measurements) (sets 1 and 2; n = 102).	84
4.10	PCA score plot of PC1 vs. PC2 for fingerprints of <i>C. subternata</i> samples with r values indicated by colour bar (n = 130; duplicates) [129 and 130 = duplicates of active extract].	86
4.11	Loadings of PC1 for fingerprints of <i>C. subternata</i> samples (n = 130; duplicates) [1 = iriflophenone-di- <i>O,C</i> -hexoside; 2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = scolymoside; 6 = phloretin-3',5'-di- <i>C</i> -glucoside].	86
4.12	Loadings of PC2 for fingerprints of <i>C. subternata</i> samples (n = 130; duplicates) [1 = iriflophenone-di- <i>O,C</i> -hexoside; 2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = scolymoside; 6 = phloretin-3',5'-di- <i>C</i> -glucoside; 7 = eriocitrin; 8 = hesperidin].	87

4.13	PCA biplot of PC1 vs. PC2 scores and loadings of the quantitative data for <i>C. subternata</i> samples with <i>r</i> values indicated by colour bar (n = 65; no duplicates) [65 = active extract; Mg = mangiferin; IMg = isomangiferin; Iri-glc = iriflophenone-3- <i>C</i> -glucoside; Iri-di-glc = iriflophenone-di- <i>O,C</i> -hexoside; Phlor = phloretin-3',5'-di- <i>C</i> -glucoside; Erio = eriocitrin; scol = scolymoside; HPdi-glc = 3-hydroxyphloretin-3',5'-di- <i>C</i> -hexoside; Hd = hesperidin].	87
4.14	PCA score plot of PC1 vs. PC2 for fingerprints of <i>C. maculata</i> samples (n = 102) [81 and 82 = duplicates of active extract].	88
4.15	Loadings of PC1 for fingerprints of <i>C. maculata</i> samples (n = 102) [1 = iriflophenone-3- <i>C</i> -glucoside; 2 = mangiferin; 3 = isomangiferin; 4 = hesperidin].	88
4.16	Loadings of PC2 for fingerprints of <i>C. maculata</i> samples (n = 102) [1 = maclurin-3- <i>C</i> -glucoside; 2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = hesperidin].	89
4.17	PCA biplot containing PC1 vs. PC2 scores and loadings for the quantitative data of <i>C. maculata</i> (n = 51; no duplicates) [41 = active extract; Mg = mangiferin; IMg = isomangiferin; Iri-glc = iriflophenone-3- <i>C</i> -glucoside; Erio = eriocitrin; Hd = hesperidin; Mac = maclurin-3- <i>C</i> -glucoside]. . .	89
4.18	PCA score plots of PC1 vs. PC3 for fingerprints of <i>C. maculata</i> samples (n = 102) [81 and 82 = duplicates of active extract].	90
4.19	Loading plot of PC3 for fingerprints of <i>C. maculata</i> samples (n = 102; duplicates) [1 = maclurin-3- <i>C</i> -glucoside; 2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = hesperidin; 6 = extra peak].	90
5.1	Schematic representation of data handling and pre-treatment [TAC = total antioxidant capacity; PCA = principal component analysis; PLS = partial least squares].	99
5.2	Histograms showing the distribution of TAC values from (a) DPPH and (b) ORAC assays for <i>C. subternata</i> and (c) DPPH assay for <i>C. maculata</i> . Histograms of <i>C. subternata</i> and <i>C. maculata</i> data comprised 10 and 11 bins, respectively [TE = Trolox equivalents].	101
5.3	PCA scores plot of PC2 vs. PC3 for chromatograms of <i>C. subternata</i>	102
5.4	PCA score plots of PC2 vs. PC3 for chromatograms of <i>C. maculata</i> (a) before and (b) after removal of the outlier (sample 24).	102
5.5	RMSE values for models with varying complexity to predict TAC _{DPPH} of <i>C. subternata</i> extracts . .	104
5.6	Measured vs. predicted TAC _{DPPH} of <i>C. subternata</i> extracts	104
5.7	RMSE values for models with varying complexity to predict TAC _{ORAC} of <i>C. subternata</i> extracts . .	105
5.8	Measured vs. predicted TAC _{ORAC} of <i>C. subternata</i> extracts	105
5.9	RMSE values for models with varying complexity to predict TAC _{DPPH} of <i>C. maculata</i> extracts . . .	106
5.10	Measured vs. predicted TAC _{DPPH} of <i>C. maculata</i> extracts	106
5.11	Compounds of importance when predicting TAC _{DPPH} of <i>C. subternata</i> extracts shown on a (a) representative chromatogram, as evaluated using (b) regression coefficients, (c) loading scores and the (d) variables of importance for projection (VIP) scores [1 = iriflophenone-di- <i>O,C</i> -hexoside; 2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = apigenin-6,8-di- <i>C</i> -glucoside; 6 = 3-hydroxyphloretin-3',5'-di- <i>C</i> -hexoside; 7 = eriocitrin; 8 = scolymoside; 9 = phloretin-3',5'-di- <i>C</i> -glucoside; 10 = hesperidin; a = b = unidentified; * = compounds with positive regression and high variables of importance for projection scores; # = compounds of interest with low UV-Vis signals]. .	108

5.12	Compounds of importance when predicting TAC _{ORAC} of <i>C. subternata</i> extracts shown on a (a) representative chromatogram, as evaluated using (b) regression coefficients, (c) loading scores and the (d) variables of importance for projection (VIP) scores [1 = iriflophenone-di- <i>O,C</i> -hexoside; 2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = apigenin-6,8-di- <i>C</i> -glucoside; 6 = 3-hydroxyphloretin-3',5'-di- <i>C</i> -hexoside; 7 = eriocitrin; 8 = scolymoside; 9 = phloretin-3',5'-di- <i>C</i> -glucoside; 10 = hesperidin; a = b = c = unidentified; * = compounds with positive regression and high variables of importance for projection scores; # = compounds of interest with low UV-Vis signals].	109
5.13	Compounds of importance when predicting TAC _{DPPH} of <i>C. maculata</i> extracts shown on a (a) representative chromatogram, as evaluated using (b) regression coefficients, (c) loading scores and the (d) variables of importance for projection (VIP) scores [2 = iriflophenone-3- <i>C</i> -glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = apigenin-6,8-di- <i>C</i> -glucoside; 7 = eriocitrin; 10 = hesperidin; 11 = maclurin-3- <i>C</i> -glucoside; d = e = f = g = h = unidentified; * = compounds with positive regression and high variables of importance for projection scores; # = compounds of interest with low UV-Vis signals].	110
B.1	Chromatographic fingerprints of a typical and the 'atypical' <i>C. maculata</i> extract (sample 24) as used for data analysis; extra peak indicated with black star symbol (*).	130
B.2	Chromatogram of a typical <i>C. maculata</i> extract at 288 nm.	131
B.3	Chromatogram of the 'atypical' <i>C. maculata</i> extract (sample 24) at 288 nm; extra peak indicated with black star symbol (*).	131
B.4	UV-spectra of benzophenones (a) compound * (extra peak, presumably iriflophenone-di- <i>O,C</i> -hexoside), (b) maclurin-3- <i>C</i> -glucoside and (c) iriflophenone-3- <i>C</i> -glucoside from the 'atypical' <i>C. maculata</i> extract (sample 24).	132
B.5	UV-spectra of benzophenones (a) iriflophenone-di- <i>O,C</i> -hexoside, (b) maclurin-3- <i>C</i> -glucoside and (c) iriflophenone-3- <i>C</i> -glucoside from a <i>C. subternata</i> extract.	132

List of Tables

2.1	Chemical compounds identified in <i>Cyclopia</i> spp.	14
2.2	Average content (g/100 g extract) of the major phenolic compounds in unfermented and fermented honeybush obtained from De Beer & Joubert (2010) (values outside and inside brackets indicate those of unfermented and fermented honeybush, respectively)	16
2.3	Previous methods employed for separation of phenolic compounds in <i>Cyclopia</i> spp.	19
2.4	Parameters used to validate an analytical method, as adapted from LoBrutto & Patel (2007), Anonymous (2010) and Snyder <i>et al.</i> (2010)	20
2.5	Examples of the application of HPLC fingerprints in chemometric analysis of different herbal products	37
3.1	UV-Vis, LC-MS and LC-MS/MS characteristics of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented <i>C. subternata</i>	55
3.2	UV-Vis, LC-MS and LC-MS/MS characteristics of unidentified compounds in freeze-dried aqueous extracts of green and fermented <i>C. subternata</i>	56
3.3	UV-Vis, LC-MS and LC-MS ^E characteristics of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented <i>C. maculata</i>	62
3.4	Calibration curves for phenolic standards using the method for <i>C. subternata</i> extracts	63
3.5	Calibration curves for phenolic standards using the method for <i>C. maculata</i> extracts	64
3.6	Compound stability in standard calibration mixtures and reconstituted freeze-dried aqueous extracts of green and fermented <i>C. subternata</i> extracts over a 26 h period (injected in <i>ca.</i> 3 h intervals) . . .	64
3.7	Compound stability in standard calibration mixtures and reconstituted freeze-dried aqueous extracts of green and fermented <i>C. maculata</i> over a 24 h period (injected in <i>ca.</i> 3 h intervals)	65
3.8	Intra- and inter-day precision (%RSD) for individual phenolic compounds as determined using standard calibration mixtures, and reconstituted freeze-dried aqueous extracts of green and fermented <i>C. subternata</i>	66
3.9	Intra- and inter-day precision (%RSD) for individual phenolic compounds as determined using standard calibration mixtures, and reconstituted freeze-dried green and fermented <i>C. maculata</i> extracts .	67
5.1	RMSE values of PLS models for <i>C. subternata</i> (TAC _{DPPH} and TAC _{ORAC}) and <i>C. maculata</i> (TAC _{DPPH})	103
A.1	Quantitative data from extracts of unfermented (green) <i>C. subternata</i> stems and leaves (n = 64) . .	127
A.2	Quantitative data from extracts of unfermented (green) <i>C. subternata</i> stems (n = 10)	127
A.3	Quantitative data from extracts of unfermented (green) <i>C. subternata</i> leaves (n = 10)	127

A.4 Quantitative data from extracts of unfermented (green) <i>C. maculata</i> (pot trial) stems and leaves (n = 40)	128
A.5 Quantitative data from extracts of unfermented (green) <i>C. maculata</i> (wild-harvested) stems and leaves (n = 10)	128

Chapter 1

Introduction

Honeybush (*Cyclopia* spp.) is an endemic South African fynbos plant that grows along the coastal regions of the Western and Eastern Cape. Currently the main economic value of honeybush lies in its use as herbal tea, but it offers great opportunity for value-added products such as extracts for the cosmetic, nutraceutical and food ingredient markets. The market for such honeybush products is still small, but expanding (Joubert *et al.*, 2011). To date, over 20 honeybush species have been identified and they grow fairly localised due to their specific growing requirements. Species currently favoured for commercial use are *C. genistoides* and *C. subternata*, which are cultivated, and *C. intermedia*, a largely wild-harvested species which comprises most of the total annual production. With the demand for honeybush surpassing its production, other species, such as *C. maculata*, *C. longifolia* and *C. sessiliflora* have elicited interest for commercialisation (Joubert *et al.*, 2011).

Changing consumer attitudes towards the role of food in health promotion and greater awareness of the role of antioxidants in health are supporting a growing market for herbal extracts as supplements and food ingredients (Crawford, 2012; Dennis 2013). The global nutraceutical market, which includes functional foods and beverages and dietary supplements, was estimated at US \$142-151 billion in 2011, and is expected to reach nearly US \$205 billion by 2017 according to a projected compound annual growth rate of 6.3% (Anon., 2013). Honeybush has the potential to be considered a functional ingredient and nutraceutical product as it contains a relatively high concentration of antioxidants. Studies have focused on the antioxidant, anti-mutagenic, anti-cancer and phytoestrogen properties of honeybush (Joubert *et al.*, 2008a), and more recently its anti-diabetic (Muller *et al.*, 2011) and anti-obesity (Dudhia *et al.*, 2013) properties. As a result honeybush extracts could also be considered condition-specific sources of antioxidants and therefore of higher value (Becker, 2013). These properties are mostly attributed to the phenolic content of honeybush. The presence of high levels of the xanthone mangiferin, and the flavanone hesperidin, amongst others (De Beer & Joubert, 2010), supports the nutraceutical use of honeybush extracts. Recent reviews summarise the various pharmacological properties of these two compounds (Garg *et al.*, 2001; Wauthoz *et al.*, 2007).

Plant extracts are prone to large compositional variation due to intrinsic and extrinsic factors such as seasonal variation, environment, species, cultivation and harvesting techniques, storage conditions and extraction methods, as recently reviewed for tea (*Camellia sinensis*) (Tounekti *et al.*, 2012). Xanthone biosynthesis in *Hypericum perforatum* cells is stimulated by biotic stress (Franklin *et al.*, 2009). Standardised extracts are therefore in higher demand than non-standardised extracts (Crawford, 2012; Becker, 2013). Similarly, honeybush tea demonstrates compositional variation due to species (Ferreira *et al.*, 1998; Kamara *et al.*, 2003, 2004; Joubert *et al.*, 2008b), harvesting date (Joubert *et al.*, 2003), genotype and ecotype (unpublished results) and processing (Joubert *et al.*, 2008b). Existing quality control of honeybush extracts as food ingredients is limited to minimum specifications for total polyphenol content and total antioxidant capacity (TAC). However, for nutraceutical products, quality control should be more stringent. According to new regulations of the European Food Safety Authority (EFSA), nutraceuticals should be accompanied by scientific evidence to support health claims and to obtain authorisation (Anon., 2007). In order to provide scientific evidence, quality control measures have to be implemented to prove the corresponding biological activity.

High performance liquid chromatography (HPLC) is an efficient tool for the analysis of phenolic compounds in herbal extracts (Dobes *et al.*, 2013). It allows for their separation, tentative identification and quantification. Quantification provides a sufficient indication of the phenolic variation between different species as demonstrated for *Cyclopia* (De Beer & Joubert, 2010), as well as the evaluation of specific compounds for batch-to-batch consistency (Monagas *et al.*, 2005). When a specific biological activity is linked to a fixed concentration of

marker compound(s) which have been proven to represent the activity, quantification of these compounds is required for quality control (Ninfali *et al.*, 2009; Rossi *et al.*, 2012). However, when marker compounds cannot be established, and considering the possibility of synergistic effects, more comprehensive representation of extracts is required (Ong, 2004; Ninfali *et al.*, 2009).

A more comprehensive analysis of plant extracts is possible when HPLC fingerprints are considered. Fingerprints are regarded as more representative of the complex chemical profile of herbal species as opposed to one or two marker compounds (Tistaert *et al.*, 2011). Fingerprint data have been introduced in a variety of applications for quality control purposes. These include similarity analysis, which has been used to distinguish extracts of the same species that have been collected from different sources (*Chaunxiang*) (Gong *et al.*, 2003) and origins (*Yiqing*) (Kong *et al.*, 2009), and determination of the stability of (Quingfu Guanjiesshu) samples post-production (Xie *et al.*, 2007) and batch-to-batch consistency (*Gegen Qinlian*) (Chen *et al.*, 2010). Exploratory analysis has been applied using fingerprints to distinguish between different origins (*Niu Huany Jie Du Pill*) (Liang *et al.*, 2009) as well as metabolite variation during growth stages (*Artemisia annua* L.) (Ma *et al.*, 2008). Different classification methods have been applied to differentiate between adulterated and authentic commercial samples (*Citri reticulatae* and *Citri reticulatae Viride*) (Yi *et al.*, 2007), plant origins (*Ganoderma lucidum*) (Chen *et al.*, 2008), cultivation areas (*Angelica acutiloba*) (Tianniam *et al.*, 2009) and steam treatments (*Panax notoginseng*) (Toh *et al.*, 2010). Furthermore, regression techniques have been applied to predict antioxidant capacity of extracts (green tea) (Daszykowski *et al.*, 2007), or to indicate compounds potentially responsible for an activity (*Mallotus* species) (Nguyen Hoai *et al.*, 2009; Tistaert *et al.*, 2009, 2012).

In order to obtain representative fingerprints, HPLC methods with good peak separation are required for herbal extracts. Characterisation studies have been conducted on the phenolic fraction of honeybush species, *C. subternata*, *C. intermedia* and *C. genistoides*, although these focused on organic extracts and involved no quantification (Ferreira *et al.*, 1998; Kamara *et al.*, 2003, 2004; Kokotkiewicz *et al.*, 2012, 2013). Studies relating to aqueous extracts of green (unfermented) and fermented *C. subternata*, *C. intermedia*, *C. genistoides* and *C. sessiliflora* suggested that the species analysed to date all contained the same three major phenolic compounds, namely mangiferin, isomangiferin and hesperidin (Joubert *et al.*, 2003; De Beer & Joubert, 2010). Subsequently, the HPLC methods developed mostly focussed on the quantification of these major phenolic compounds in *Cyclopia* species (Joubert *et al.*, 2003; De Beer & Joubert, 2010). An HPLC method developed for *C. subternata* failed to provide sufficient resolution of several compounds (De Beer & Joubert, 2010). Apart from data for the three major compounds from organic extracts, no other compositional data are available for *C. maculata* (Joubert *et al.*, 2003). The present study will focus specifically on aqueous extracts of *C. subternata* and *C. maculata*, as aqueous extracts are mainly produced by industry (Joubert *et al.*, 2011) and are the main type of extracts investigated for health promoting properties of *Cyclopia* (Joubert *et al.*, 2008). Biological activities studied for these two species, which include anti-diabetic (Mose Larsen *et al.*, 2008) and anti-obesity (Dudhia *et al.*, 2013), have not been linked to specific compounds and quantities, although their phenolic constituents are considered to be responsible for the activity. In order to investigate the relationship between their phenolic compounds and activity, better characterisation of the phenolic composition of these species is required.

As oxidative stress is the cause of numerous chronic diseases, antioxidant rich diets have subsequently been linked to the protection against oxidative damage *in vivo* (Espín *et al.*, 2007; Puchau *et al.*, 2009; Hermsdorff *et al.*, 2011; Badhadoran *et al.*, 2012). Therefore the measurement of antioxidant capacity in food products

is important to the health and nutraceutical industries (George *et al.*, 2005). The total antioxidant capacity (TAC) is also used as a measure of quality of plant extracts (Bell & Ou, 2007; Ninfali *et al.*, 2009). TAC assays include the 2,2-diphenyl-1-picrylhydrazyl (DPPH) (Brand-Williams *et al.*, 1995), total phenols (TP) (Singleton & Rossi, 1965), oxygen radical absorbance capacity (ORAC) (Cao *et al.*, 1993), ferric reducing antioxidant power (FRAP) (Benzie & Strain, 1996) and 2,2'-azino-bis-3-ethylbenzthiazoline-6-sulphonic acid (ABTS) (Re *et al.*, 1999) assays, and are relatively fast and inexpensive. To be able to identify the phenolic compounds that have the largest influence on the antioxidant activity in a herbal product would contribute greatly to more simplistic and rapid future quality control measures. Fingerprints can be used to identify compounds responsible for an activity (antioxidant and/or biological) by incorporating multivariate data analysis techniques, as demonstrated in earlier studies (Nguyen Hoai *et al.*, 2009; Kvalheim *et al.*, 2011; Tistaert *et al.*, 2011).

HPLC fingerprint analysis of *Cyclopia* extracts also has potential as a screening technique for activity since extracts of unknown activity could be compared to active extracts as a first screening step in research and quality control. Current methods for evaluating extract activity involve expensive and time-consuming cell and/or animal model tests. For quality control, a need exists for rapid and less expensive screening methods to evaluate activity of different extract production batches. Such a technique could also be used in the Plant Breeding and Improvement Programme of Honeybush of the Agricultural Research Council (Stellenbosch, South Africa) to identify "bio-active" plant selections for advancement to commercialisation trials.

The aims of this study are informed by the lack of comprehensive HPLC methods for *C. subternata* and *C. maculata*, as well as the need for a more advanced quality control method than currently employed by industry. Therefore the aims are to develop improved HPLC methods suitable for quantification of the phenolic compounds in aqueous extracts, and evaluation of fingerprint analysis as a screening and predictive tool. The fingerprints that will be generated from the analysis of numerous green *C. subternata* and *C. maculata* extracts will be submitted to extensive pre-treatment before they will be used to examine alternative methods for rapid, inexpensive quality control of these extracts for future use. This will include the use of fingerprints together with multivariate data analysis to determine the similarity between extracts of unknown activity and extracts with proven biological activity, also aiming at identifying compound(s) responsible for this activity. Fingerprint analysis will also be employed to predict the TAC of the extracts using a multivariate regression method. This will be used to determine whether there are specific compound(s) that contribute significantly to the prediction of TAC, which could then be regarded as marker(s) when evaluating the TAC of new selections or in quality control. Ultimately, this study will serve as a starting point for the investigation of alternative methods for future quality control measures of *Cyclopia*.

References

- Anonymous (2007). Corrigendum to Regulation (EC) No 1924/2006 of the European Parliament and of the council of 20 December 2006 on nutrition and health claims made on foods. *Official Journal of the European Union*, Article 13.
- Anonymous (2013). Nutraceuticals product market is expected to reach USD 204.8 billion globally in 2017: transparency market research. PRWeb. [WWW document]. URL <http://www.prweb.com/releases/2013/6/prweb10868043.htm>. 8 August 2013.
- Bahadoran, Z., Golzarand, M., Mirmiran, P., Shive, N. & Azizi, F. (2012). Dietary total antioxidant capacity and the occurrence of metabolic syndrome and its components after 3-year follow-up in adults: Tehran lipid and glucose study. *Nutrition and Metabolism*, **9**(70), doi:10.1186/1743-7075-9-70.
- Becker, M. (2013). Market maturation: Innovation & science drive antioxidants forward. *Nutraceuticals World*, **16**(2), 36-46.
- Bell, D. & Ou, B. (2007). The branding of ORAC: where will it lead? *Nutraceuticals World*, **10**(3), 72-74.
- Benzie, I.F.F. & Strain, J.J. (1996). The ferric reducing ability of plasma (FRAP) as a measure of 'antioxidant power': The FRAP assay. *Analytical Biochemistry*, **239**, 70-76.
- Brand-Williams, W., Cuvelier, M.E. & Berset, C. (1995). Use of free radical method to evaluate antioxidant activity. *Lebensmittel-Wissenschaft und Technologie*, **28**, 25-30.
- Cao, G.H., Alessio, H.M. & Cutler, R.G. (1993). Oxygen-radical absorbance capacity assay for antioxidants. *Free Radical Biology and Medicine*, **14**, 303-311.
- Chen, L., Tang, Y.P., Chen, M.J., Deng, H.S., Yan, X.P. & Wu, D.K. (2010). Chemical correlation between *Gegen Qinlian* dispensing granule and its four raw herbs by LC fingerprint. *Phytomedicine*, **17**, 100-107.
- Chen, Y., Zhu, S.B., Xie, M.Y., Nie, S.P., Liu, W., Li, C., Gong, X.F. & Wang, Y.X. (2008). Quality control and original discrimination of *Ganoderma lucidum* based on high-performance liquid chromatographic fingerprints and combined chemometrics methods. *Analytica Chimica Acta*, **623**, 146-156.
- Crawford, M. (2012). Antioxidant market: strong & sophisticated. *Nutraceuticals World*, **15**(2), 38.
- Daszykowski, M., Vander Heyden, Y. & Walczak, B. (2007). Robust partial least squares model for prediction of green tea antioxidant capacity from chromatograms. *Journal of Chromatography A*, **1176**, 12-18.
- De Beer, D. & Joubert, E. (2010). Development of HPLC method for *Cyclopia subternata* phenolic compound analysis and application to other *Cyclopia* spp. *Journal of Food Composition and Analysis*, **23**, 289-297.
- Dennis, J. (2013). International perspectives on herbs & botanicals. *Nutraceuticals World*, **16**(6), 48-53.
- Dobes, J., Zitka, O., Sochor, J., Ruttikay-Nedecky, B., Babula, P., Beklova, M., Kynicky, J., Hubalek, J., Klejdus, B., Kizek, R. & Adam, V. (2013). Electrochemical tools for determination of phenolic compounds in plants. A review. *International Journal of Electrochemical Science*, **8**, 4520-4542.
- Dudhia, Z., Louw, J., Muller, C., Joubert, E., de Beer, D., Kinnear, C. & Pheiffer, C. (2013). *Cyclopia maculata* and *Cyclopia subternata* (honeybush tea) inhibits adipogenesis in 3T3-L1 pre-adipocytes. *Phytomedicine*, **20**, 401-408.
- Espín, J.C., García-Conesa, M.T. & Tomás-Barberán, F.A. (2007). Nutraceuticals: facts and fiction. *Phytochemistry*, **68**, 2986-3008.
- Ferreira, D., Kamara, B.I., Brandt, E.V. & Joubert, E. (1998). Phenolic compounds from *Cyclopia intermedia* (honeybush tea). *Journal of Agricultural and Food Chemistry*, **46**, 3406-4310.

- Franklin, G., Conceição, L.F.R., Kombrink, E. & Dias, A.C.P. (2009). Xanthone biosynthesis in *Hypericum perforatum* cells provides antioxidant and antimicrobial protection upon biotic stress. *Phytochemistry*, **70**, 65-73.
- Garg, A., Garg, S., Zaneveld, L.J.D. & Singla, A.K. (2001). Chemistry and pharmacology of the citrus bioflavonoid hesperidin. *Phytotherapy Research*, **15**, 655-669.
- George, S., Brat, P., Alter, P. & Amiot, M.J. (2005). Rapid determination of polyphenols and vitamin C in plant-derived products. *Journal of Agricultural and Food Chemistry*, **53**, 1370-1373.
- Gong, F., Liang, Y.Z., Xie, P.S. & Chau, F.T. (2003). Information theory applied to chromatographic fingerprint of herbal medicine for quality control. *Journal of Chromatography A*, **1002**, 25-40.
- Hermisdorff, H.H.M., Puchau, B., Volp, A.C.P., Barbosa, K.B.F., Bressan, J., Zulet, M.A. & Martínez, J.A. (2011). Dietary total antioxidant capacity is inversely related to central adiposity as well as to metabolomic and oxidative stress markers in healthy young adults. *Nutrition and Metabolism*, **8**(59), doi:10.1186/1743-7075-8-59.
- Joubert, E., Gelderblom, W.C.A., Louw, A. & De Beer, D. (2008a). South African herbal teas: *Aspalathus linearis*, *Cyclopia* spp. and *Athrixia phylicoides* - A review. *Journal of Ethnopharmacology*, **119**, 376-412.
- Joubert, E., Joubert, M.E., Bester, C., De Beer, D. & De Lange, J.H. (2011). Honeybush (*Cyclopia* spp.): From local cottage industry to global markets - The catalytic and supporting role of research. *South African Journal of Botany*, **77**, 887-907.
- Joubert, E., Otto, F., Grüner, S. & Weinreich, B. (2003). Reversed-phase HPLC determination of mangiferin, isomangiferin and hesperidin in *Cyclopia* and the effect of harvesting date on the phenolic composition of *C. genistoides*. *European Food and Research Technology*, **216**, 270-273.
- Joubert, E., Richards, E.S., Van der Merwe, J.D., De Beer, D. Manley, M. & Gelderblom, W.C.A. (2008b). Effect of species variation and processing on phenolic composition and in vitro antioxidant activity of aqueous extracts of *Cyclopia* spp. (honeybush tea). *Journal of Agricultural and Food Chemistry*, **56**, 954-963.
- Kamara, B.I., Brand, D.J., Brandt, E.V. & Joubert, E. (2004). Phenolic metabolites from honeybush tea (*Cyclopia subternata*). *Journal of Agricultural and Food Chemistry*, **52**, 5291-5395.
- Kamara, B.I., Brandt, E.V., Ferreira, D. & Joubert, E. (2003). Polyphenols from honeybush tea (*Cyclopia intermedia*). *Journal of Agricultural and Food Chemistry*, **51**, 3874-3879.
- Kokotkiewicz, A., Luczkiewicz, M., Sowinski, P., Glod, D., Gorynski, K. & Bucinski, A. (2012). Isolation and structure elucidation of phenolic compounds from *Cyclopia subternata* Vogel (honeybush) intact plant and *in vitro* cultures. *Food Chemistry*, **133**, 1373-1382.
- Kokotkiewicz, A., Luczkiewicz, M., Pawlowska, J., Luczkiewicz, P., Sowinski, P., Witkowski, J., Bryl, E. & Bucinski, A. (2013). Isolation of xanthone and benzophenone derivatives from *Cyclopia genistoides* (L.) Vent. (honeybush) and their pro-apoptotic activity in synoviocytes from patients with rheumatoid arthritis. *Fitoterapia*, **90**, 199-208.
- Kong, W.J., Zhao, Y.L., Xiao, X.H., Jin, C. & Li, Z.L. (2009). Quantitative and chemical fingerprint analysis for quality control of *Rhizoma Coptidis chinensis* based on UPLC-PAD combined with chemometrics methods. *Phytomedicine*, **16**, 950-959.
- Kvalheim, O.M., Chan, H., Benzie, I.F.F., Szetto, Y., Hing-chung Tzang, A., Kam-wah Mok, D. & Chau, F. (2011). Chromatographic profiling and multivariate analysis for screening and quantifying the contributions

- from individual components to the bioactive signature in natural products. *Chemometric and Intelligent Laboratory Systems*, **107**, 98-105.
- Liang, X.M., Jin, Y., Wang, Y.P., Jin, G.W., Fu, Q. & Xiao, Y.S. (2009). Qualitative and quantitative analysis in quality control of traditional Chinese medicines. *Journal of Chromatography A*, **1216**, 2033-2044.
- Ma, C., Wang, H., Lu, X., Xu, G. & Liu, B. (2008). Metabolic fingerprinting investigation of *Artemisia annua* L. in different stages of development by gas chromatography and gas chromatography- mass spectrometry. *Journal Chromatography A*, **1186**, 412-419.
- Monagas, M., Hernández-Ledesma, B., Garrido, I., Martín-Álvarez, P.J., Gómez-Cordovés, C. & Bartolomé, B. (2005). Quality assessment of commercial dietary antioxidant products from *Vitis vinifera* L. grape seeds. *Nutrition and Cancer*, **53**, 244-254.
- Mose Larsen, P., Fey, S.J., Louw, J. & Joubert, E. (2008). An anti-diabetic extract of honeybush. World Intellectual Property Organization WO. 2008/110552A2.
- Muller, C.J.F., Joubert, E., Gabuza, K., De Beer, D., Fey, S.J. & Louw, J. (2011). Assessment of the antidiabetic potential of an aqueous extract of honeybush (*Cyclopia intermedia*) in streptozotocin and obese resistant Wistar rats. In: *Phytochemicals - bioactivities and impact on health* (edited by I. Rasooli). Pp. 313-332. Rijeka: Intech.
- Nguyen Hoai, N., Dejaegher, B., Tistaert, C., Nguyen Thi Hong, V., Rivière, C., Chataigné, G., Phan Van, K., Chau Van, M., Quentin-Leclercq, J. & Vander Heyden, Y. (2009). Development of HPLC fingerprints for *Mallotus* species extracts and evaluation of peaks responsible for their antioxidant activity. *Journal of Pharmaceutical and Biomedical Analysis*, **50**, 753-763.
- Ninfali, P., Gennari, L. & Biagiotti, E. (2009). Improvement in botanical standardization of commercial freeze-dried herbal extracts by using the combination of antioxidant capacity and constituent marker concentrations. *Journal of AOAC International*, **92**, 797-805.
- Ong, E.S. (2004). Extraction methods and chemical standardization of botanicals and herbal preparations. *Journal of Chromatography B*, **812**, 23-33.
- Puchau, B., Zulet, M.A., de Echavarri, A.G., Hermsdorff, H.H.M., Martinez, J.A. (2009). Dietary total antioxidant capacity: A novel indicator of diet quality in healthy young adults. *Journal of the American College of Nutrition*, **28**, 648-656.
- Re, R., Pellegrini, N., Proteggente, A., Pannala, A., Yang, M. & Rice-Evans, C. (1999). Antioxidant activity applying an improved ABTS radical cation decolorization assay. *Free Radical Biology and Medicine*, **26**, 1231-1237.
- Rossi, D., Guerrini, A., Bruni, R., Brognara, E., Borgatti, M., Gambari, R., Maietti, S. & Sacchetti, G. (2012). *trans*-Resveratrol in nutraceuticals: Issues in retail quality and effectiveness. *Molecules*, **17**, 12393-12405.
- Singleton, V.L. & Rossi, J.A. (1965). Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *American Journal of Enology and Viticulture*, **16**, 144-158.
- Tianniam, S., Bamba, T. & Fukusaki, E. (2009). Non-targeted metabolite fingerprinting of oriental folk medicine *Angelica acutiloba* roots by ultra performance liquid chromatography time-of-flight mass spectrometry. *Journal of Separation Science*, **32**, 2233-2244.
- Tistaert, C., Dejaegher, B., Chataigné, G., Nguyen Hoai, N., Rivière, C., Nguyen Thi Hong, V., Chau Van, M., Quetin-Leclercq, J. & Vander Heyden, Y. (2012). Potential antioxidant compounds in *Mallotus* species fingerprints. Part II: Identification with LC-MS. *Analytica Chimica Acta*, **721**, 35-43.

- Tistaert, C., Dejaegher, B., Nguyen Hoai, N., Chataigné, G., Rivière, C., Nguyen Thi Hong, V., Chau Van, M., Quetin-Leclercq, J. & Vander Heyden, Y. (2009). Potential antioxidant compounds in *Mallotus* species fingerprints. Part I: indication, using linear multivariate calibration techniques. *Analytica Chimica Acta*, **652**, 189-197.
- Tistaert, C., Dejaegher, B. & Vander Heyden, Y. (2011). Chromatographic separation techniques and data handling methods for herbal fingerprints: a review. *Analytica Chimica Acta*, **690**, 148-161.
- Toh, D.F., New, L.S., Koh, H.L. & Chan, E.C.Y. (2010). Ultra-high performance liquid chromatography/time-of-flight mass spectrometry (UHPLC/TOFMS) for time-dependent profiling of raw and steamed *Panax notoginseng*. *Journal of Pharmaceutical and Biomedical Analysis*, **52**, 43-50.
- Tounekti, T., Joubert, E., Hernández, I. & Munné-Bosch, S. (2013). Improving the polyphenol content of tea. *Critical Reviews in Plant Science*, **32**, 192-215.
- Wauthoz, N., Balde, A., Balde, E.S., Van Damme, M. & Duez, P. (2007). Ethnopharmacology of *Mangifera indica* L. bark and pharmacological studies of its main C-glucosylxanthone, mangiferin. *International Journal of Biomedical and Pharmaceutical Sciences*, **1**, 112-119.
- Xie, Y., Jiang, Z.H., Zhou, H., Cai, X., Wong, Y.F., Liu, Z.Q., Bian, Z.X., Xu, H.X. & Liu, L. (2007). Combinative method using HPLC quantitative and qualitative analyses for quality consistency assessment of a herbal medicinal preparation. *Journal of Pharmaceutical and Biomedical Analysis*, **43**, 204-212.
- Yi, L.Z., Yuan, D.L., Liang, Y.Z., Xie, P.S. & Zhao, Y. (2007). Quality control and discrimination of *Pericarpium Citri Reticulatae* and *Pericarpium Citri Reticulatae Viride* based on high-performance liquid chromatographic fingerprints and multivariate statistical analysis. *Analytica Chimica Acta*, **588**, 207-215.

Chapter 2

Literature Review

2.1 Introduction

Honeybush is a traditional South African herbal tea with an extensive history of local use, and one of few indigenous South African plants that has been commercialised (Joubert *et al.*, 2011). Its "re-discovery" in the 1990s sparked interest in the development of commercial agriculture and agro-processing industries. The role of research and the progress made towards achieving these goals have been summarised in a recent review (Joubert *et al.*, 2011).

This literature review will present a brief overview of the historical and modern use of honeybush, its geographical distribution and the current status of the industry, focusing on the *Cyclopia* species of current commercial importance. Product quality as presently defined and a summary of the phenolic composition of the species investigated to date will be discussed as both aspects are integral to achieving the objectives of the current work. The section on chemical composition, with emphasis on phenolic composition, will serve as the basis for a general discussion on the use of mass spectrometry for identification of phenolic compounds. Basic principles in high performance liquid chromatography (HPLC) and mass spectrometry (MS) will be covered, together with method development, in view of the use of these methods for quantification and tentative identification of phenolic compounds in honeybush tea. Moreover, advanced chromatographic fingerprint analysis as used in quality control and screening of bioactive extracts will be discussed in terms of data analysis techniques. The latter will cover pre-treatment and multivariate data analysis techniques commonly used in fingerprint analysis. The review will be concluded with a brief section on the effects of pre-treatment methods on multivariate data analysis.

2.2 Honeybush tea

2.2.1 Honeybush as commercial product

Honeybush, with its honey-like and sweet aromatic yellow flowers, belongs to the Cape fynbos biome and grows in the coastal regions of the Western and Eastern Cape of South Africa (Du Toit *et al.*, 1998). More than 20 *Cyclopia* species have been described to date. The earliest mention of honeybush dates back to 1705 according to Kies (1951), while the first mention of its use as a tea, as indicated by the name of 'honigtee', was made by a Swedish botanist, C. Thunberg, in the 1770s (De Lange, 2002). A scientific paper published by the Royal Society of Chemistry (RSC) on an anatomical and chemical study of *C. genistoides* mentioned that this *Cyclopia* species was used as a tea substitute (Greenish, 1881). In addition to *C. genistoides*, several species have historically been used for the preparation of tea, such as *C. vogelii* (later renamed as *C. subternata*), *C. latifolia* and *C. longifolia* (Marloth, 1913, 1925; Kies 1951; Watt & Breyer-Brandwijk, 1962). According to Marloth (1925), all the tea prepared in the Cape Peninsula region was from *C. genistoides*, and in the Overberg and George areas, from *C. subternata*. It is known that *C. sessiliflora* grows along the Langeberg and Warmwaterberg mountains and *C. longifolia* on the Van Stadens River mountains near Port Elizabeth. Compared to these more localised distributions, several other species of economic importance are more widespread (illustrated in Fig. 2.1). *Cyclopia genistoides* grows in the Malmesbury - Darling, Cape Peninsula and Overberg areas, *C. subternata* along the coastal mountain ranges of Tsitsikama, Outeniqua and Langeberg and *C. maculata* along the riverbanks and streams in the south-western and southern region of the Western Cape Province. *Cyclopia intermedia* is the most widespread of the honeybush species currently of interest, and has been found to date on the Witteberg, Anysberg, Swartberg, Touwsberg, Rooiberg, Kammanassie, Kouga, Baviaanskloof,

Langeberg, Outeniqua, Tsitsikamma and Van Stadens mountains (Schutte, 1997; De Lange & von Mollendorff, 2006; Joubert *et al.*, 2011). Honeybush tea first appeared for a limited period on the South African market in the 1960s as ‘Caspa Cyclopia Tea’, but marketing remained largely through mail order and farm stalls until the 1990s. More recently, as the demand for honeybush tea increased and exceeded the supply, cultivation became important. Initially studies on propagation of honeybush, undertaken by the South African National Biodiversity Institute (SANBI) and the Agricultural Research Council (ARC), focussed on *C. intermedia*, *C. genistoides* and *C. subternata*, but research on the commercialisation of *Cyclopia* species has been expanded to include *C. longifolia* and *C. maculata* (Joubert *et al.*, 2011). In 2011, approximately 200 ha of *C. genistoides* and *C. subternata* were under cultivation. *Cyclopia intermedia* is mostly wild-harvested, and contributes most of the total honeybush production as it is preferred for its flavour by some markets (Joubert *et al.*, 2011).

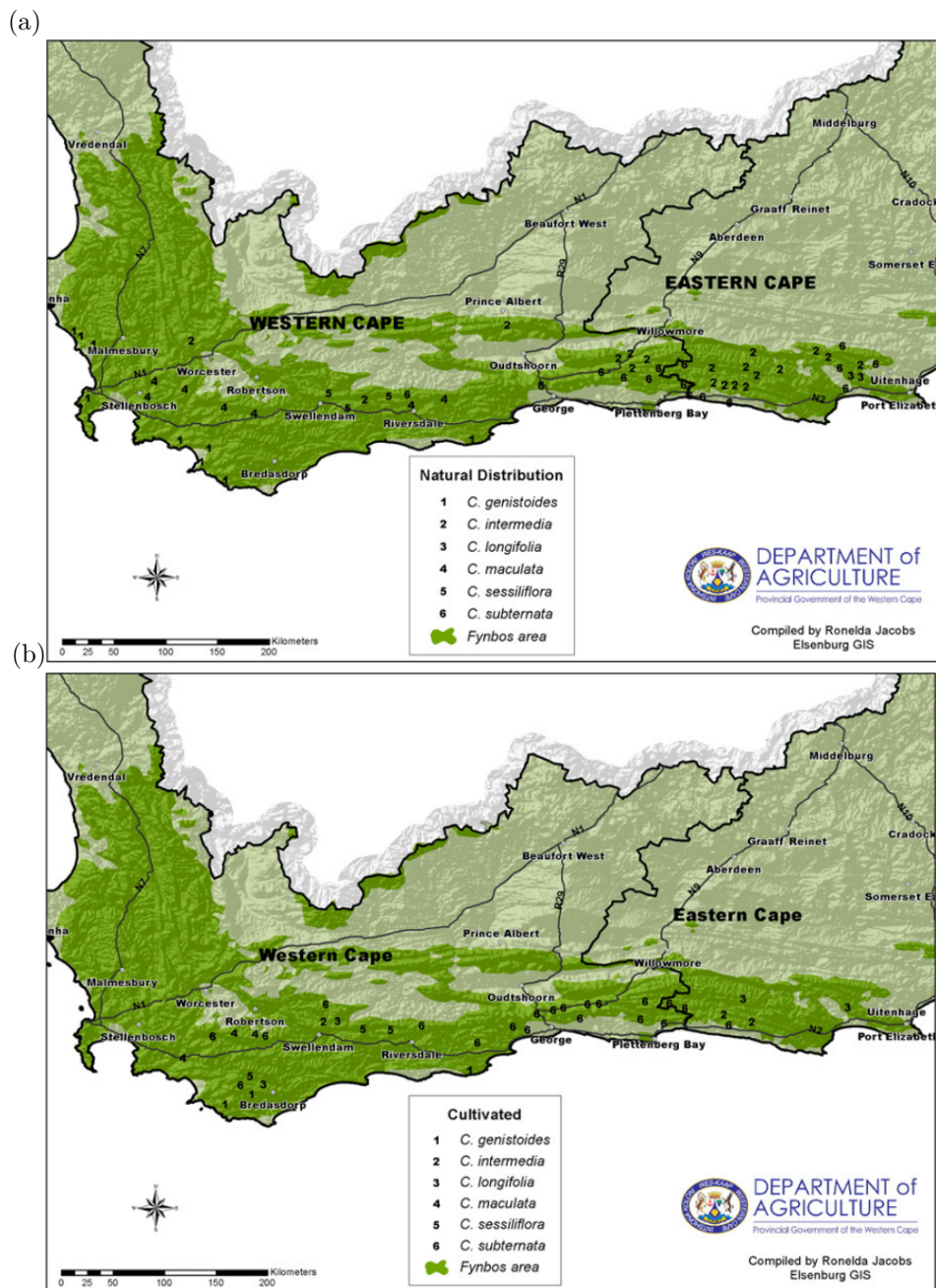


Figure 2.1 – (a) Natural and (b) cultivated distribution of honeybush species (Joubert *et al.*, 2011).

Apart from its use as herbal tea, consumption of honeybush has traditionally also been associated with health promoting properties. These include its use as an expectorant for chronic catarrh and pulmonary tuberculosis (Bowie, 1930) and the alleviation of heart burn and nausea (Marloth, 1925). Research has focussed on its anti-mutagenic, anti-cancer and phytoestrogen properties (Joubert *et al.*, 2008a), which are largely attributed to its phenolic content (discussed in section 2.2.3). Most recent studies indicated anti-diabetic properties of aqueous extracts of *Cyclopia intermedia* (Muller *et al.*, 2011) and *C. subternata* (Mose Larsen *et al.*, 2008) and anti-obesity properties of aqueous extracts of *C. subternata* and *C. maculata* (Dudhia *et al.*, 2013).

Modern uses of honeybush include consumption as a herbal tea or adding honeybush to rooibos tea, followed by consumption as an infusion. Other plant products, such as dried marula, have also been mixed with honeybush. Honeybush has followed trends similar to the use of rooibos in ready-to-drink iced teas, toiletries and cosmetic products (Joubert *et al.*, 2008a). Apart from the traditional fermented honeybush, the unfermented (unoxidised) product, also known as green honeybush, has become available. Its apparent health promoting potential, in particular its higher antioxidant activity compared to the fermented product, provided a market need for antioxidant extracts to be used in food and nutraceutical products (Joubert *et al.*, 2011). A major international food company currently markets honeybush extract, standardised on mangiferin content (a major phenolic compound in honeybush) for use as a food ingredient in drinks, food bars and dairy products (Joubert *et al.* 2011). With a number of studies examining the phenolic composition of *Cyclopia* spp. (Ferreira *et al.*, 1998; Kamara *et al.*, 2003, 2004; De Beer & Joubert, 2010; De Beer *et al.*, 2012; Kokotkiewicz *et al.*, 2012, 2013), compounds with known beneficial health effects were identified, suggesting more value-adding opportunities for honeybush.

2.2.2 Quality

Teas are usually graded or assessed by trained tea tasters to ensure overall quality and to satisfy customers' requirements (Feria-Morales, 2002; Koch, 2011). Quality control of exported honeybush tea currently entails only cut size, moisture content, the presence of foreign matter and insects, microbial safety and the levels of pesticides (Joubert *et al.*, 2010). A sensory wheel has recently been developed as a quality control tool for honeybush tea (Theron, 2012), but to date no official grading system based on sensory properties has been established. Whereas sensory evaluation is adequate when honeybush is used and consumed as a tea, other factors are important when assessing the quality of extracts, especially when their bioactivity is relevant.

Currently, apart from extract standardisation on mangiferin content, quality control of honeybush extracts only entails specifications of minimum levels in terms of total polyphenol content and antioxidant activity. Although this may be adequate when no claims regarding the bioactivity of the extracts are made, stricter regulation by the major markets of South African plant extracts, USA and Europe, will require additional quality control measures. New regulations from the European Food Safety Authority (EFSA) for example prohibits any of the 222 listed health claims for nutraceutical, functional food and drinks, and dietary supplements without substantial scientific evidence and authorisation (Anon., 2007). The first step towards fulfilling these requirements would be to have adequate and relevant quality control procedures in place. Since health promoting properties may be linked to a specific quantity of a compound or a combination of compounds, characterisation of the chemical composition of extracts is essential. Analytical method(s) such as HPLC analysis for quantification of individual compounds of interest (marker compounds) and chromatographic fingerprint analysis are methods of choice in quality control (Liang *et al.*, 2004; Jiang *et al.*, 2010). The following sections will address these

aspects in more detail.

2.2.3 Chemical composition of honeybush

Honeybush tea is caffeine free (Greenish, 1881) and contains low tannin levels (Marloth, 1925; Terblanche, 1982). Its phenolic compounds are considered the most important fraction due to the associated health promoting properties such as antioxidant activity and anti-carcinogenic properties. Only *C. subternata* and *C. intermedia* have been widely studied in relation to their phenolic composition (Ferreira *et al.*, 1998; Kamara *et al.*, 2003, 2004; Kokotkiewicz *et al.*, 2012). Some of the phenolic compounds identified in these two species have been tentatively confirmed in other species, i.e. *C. genistoides* and *C. sessiliflora*, by means of LC-MS (Joubert *et al.*, 2008b; De Beer & Joubert, 2010).

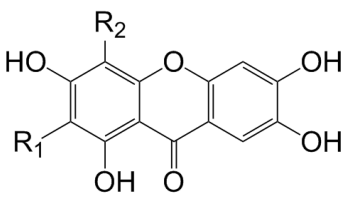
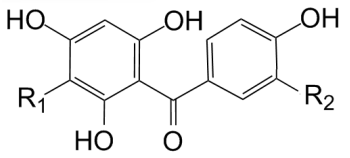
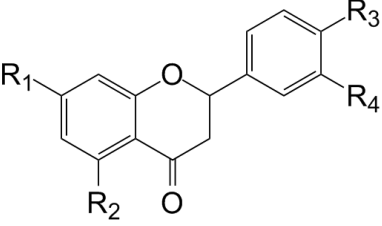
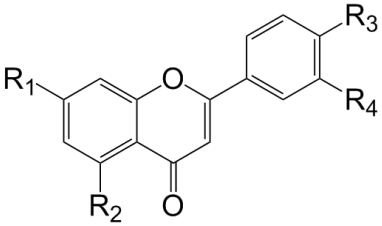
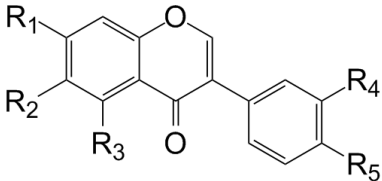
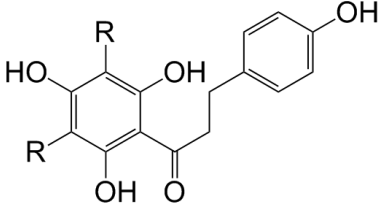
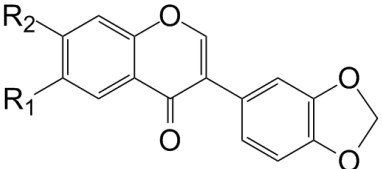
The major phenolic compounds present in all investigated *Cyclopia* species include the xanthones mangiferin and isomangiferin, as well as the flavanone hesperidin. Additional compounds previously identified are the flavanones hesperetin, naringenin, narirutin, eriodictyol, eriocitrin, naringenin-5-*O*- β -D-glucopyranoside, eriodictyol-5-*O*- β -D-glucopyranoside and eriodictyol-7-*O*- β -D-glucopyranoside. Flavones that have been identified include luteolin, 5-deoxyluteolin, scolymoside, isosakuranetin and diosmetin, while the isoflavones include formononetin, formononetin-diglucoside, afrormosin, calycosin, wistin, orobol, pseudobaptigenin and fujikinetin. Several flavonols, i.e. kaempferol glucosides, have been identified, as well as coumestans such as medicagol, flemmichaparin and sophoracoumestan (De Nysschen *et al.*, 1996; Ferreira *et al.*, 1998; Kamara *et al.*, 2003, 2004). In 2012, three previously unidentified compounds (De Beer & Joubert, 2010) were characterised for the first time. These were the benzophenone derivative iriflophenone 3-*C*- β -D-glucoside, the dihydrochalcone phloretin 3'-5'-di-*C*- β -glucoside and the flavone isorhoifolin (Kokotkiewicz *et al.*, 2012). Furthermore, three isoflavone glucosides have also been identified in the callus of *C. subternata*, i.e. ononin, rothindin and calycosin 7-*O*- β -glucoside (Kokotkiewicz *et al.*, 2012) as well as a benzophenone maclurin-3-*C*-glucoside in *C. genistoides* (Kokotkiewicz *et al.*, 2013). Table 2.1 contains a summary of the above mentioned compounds identified in *Cyclopia* spp.

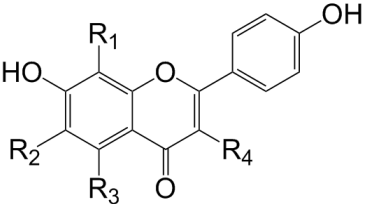
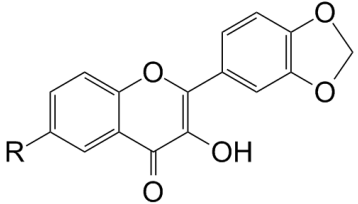
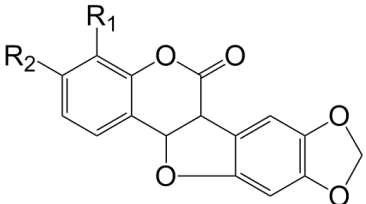
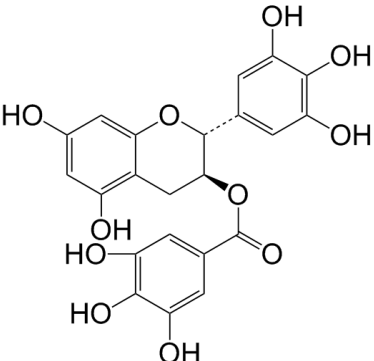
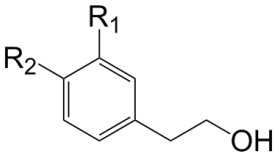
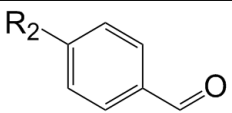
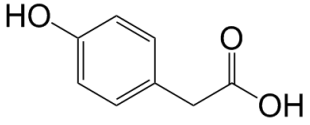
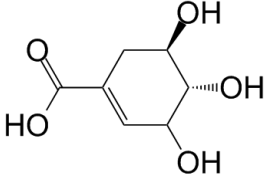
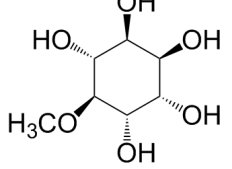
Quantitative data for the major phenolic compounds of honeybush are restricted to a few species (Table 2.2). The lack of an optimised HPLC quantification method for *C. maculata* limited data to only three major compounds present in the unfermented plant material, i.e. mangiferin, isomangiferin and hesperidin (Joubert *et al.*, 2003). From Tables 2.1 and 2.2 qualitative and quantitative differences in the phenolic composition of *Cyclopia* species are obvious. From data obtained from water extracts, it is clear that the mangiferin and isomangiferin contents are highest in *C. genistoides* and lowest in *C. subternata*. *Cyclopia subternata* and *C. sessiliflora* have the highest concentration of eriocitrin and scolymoside, while *C. sessiliflora* has the highest concentration iriflophenone-3-*C*- β -glucoside (De Beer & Joubert, 2010).

It is known that the phenolic composition of plant material is affected by numerous factors such as the environment, processing methods and genetics (Asami *et al.*, 2003; Lattanzio *et al.*, 2006; Komes *et al.*, 2011). Environmental factors such as the length of the day and the intensity of sunlight, which can be related to seasonal effects, as well as geographical location, are great contributors to these changes as demonstrated for *Camellia sinensis* (Yao *et al.*, 2005). It has also been noted that the age of the plant, the presence of flowers or pods, different geographical origin, soil and climate play roles as well (Schutte, 1997; Du Toit & Joubert, 1998).

Processing can also affect the phenolic concentration. De Beer & Joubert (2010) investigated the effect of fermentation on the phenolic composition of several *Cyclopia* species and concluded that the process decreases the content of most phenolic compounds quantified. The data showed that the xanthones mangiferin and

Table 2.1 – Chemical compounds identified in *Cyclopia* spp.

Chemical structure	Compound class, name and substituents
	Xanthone Mangiferin ^a : R ₁ = <i>C</i> -β-D-glucosyl; R ₂ = H Isomangiferin ^a : R ₁ = H ; R ₂ = <i>C</i> -β-D-glucosyl
	Benzophenone Iriflophenone-3- <i>C</i> -β-D-glucoside ^d : R ₁ = <i>C</i> -β-D-glucosyl; R ₂ = H Maclurin-3- <i>C</i> -β-glucoside ^e : R ₁ = <i>C</i> -β-glucosyl; R ₂ = OH
	Flavanone Hesperidin ^{a,c} : R ₁ = <i>O</i> -rutinosyl; R ₂ = R ₄ = OH; R ₃ = OCH ₃ Hesperetin ^a : R ₁ = R ₂ = OH; R ₃ = OCH ₃ ; R ₄ = OH Eriocitrin ^c : R ₁ = <i>O</i> -rutinosyl; R ₂ = R ₃ = R ₄ = OH Eriodictyol ^a : R ₁ = R ₂ = R ₃ = R ₄ = OH Narirutin ^c : R ₁ = <i>O</i> -rutinosyl; R ₂ = R ₃ = OH; R ₄ = H Naringenin ^a : R ₁ = R ₂ = R ₃ = OH; R ₄ = H Prunin ^b : R ₁ = <i>O</i> -glucosyl; R ₂ = R ₃ = OH; R ₄ = H Naringenin-5- <i>O</i> -rutinoside ^b : R ₁ = R ₃ = OH; R ₂ = <i>O</i> -rutinosyl; R ₄ = H Eriodictyol-5- <i>O</i> -glucoside ^b : R ₁ = R ₃ = R ₄ = OH; R ₂ = <i>O</i> -β-D-glucosyl Eriodictyol-7- <i>O</i> -glucoside ^b : R ₁ = <i>O</i> -β-D-glucosyl; R ₂ = R ₃ = R ₄ = OH;
	Flavone Luteolin ^{a,c} : R ₁ = R ₂ = R ₃ = R ₄ = OH Diosmetin ^b : R ₁ = R ₂ = R ₄ = OH; R ₃ = OCH ₃ 5-Deoxyluteolin ^c : R ₁ = R ₃ = R ₄ = OH; R ₂ = H Scolymoside ^c : R ₁ = <i>O</i> -rutinosyl; R ₂ = R ₃ = R ₄ = OH Isorhoifolin ^d : R ₁ = <i>O</i> -rutinosyl; R ₂ = R ₃ = OH; R ₄ = H
	Isoflavone Formononetin ^a : R ₁ = OH; R ₂ = R ₃ = R ₄ = H; R ₅ = OCH ₃ Formononetin diglucoside ^b : R ₁ = <i>O</i> -α-apiofuranosyl-(1''' → 6'')-β-D-glucopyranosyl; R ₂ = R ₃ = R ₄ = H; R ₅ = OCH ₃ Afromosin ^a : R ₁ = OH; R ₂ = R ₅ = OCH ₃ ; R ₃ = R ₄ = H Calycosin ^a : R ₁ = R ₄ = OH; R ₂ = R ₃ = H; R ₅ = OCH ₃ Calycosin-7- <i>O</i> -β-glucoside ^d : R ₁ = <i>O</i> -β-D-glucosyl; R ₂ = R ₃ = H; R ₄ = OH; R ₅ = OCH ₃ Wistin ^b : R ₁ = <i>O</i> -β-D-glucosyl; R ₂ = R ₅ = OCH ₃ ; R ₃ = R ₄ = H Orobol ^c : R ₁ = R ₃ = R ₄ = R ₅ = OH; R ₂ = H Formononetin-7- <i>O</i> -glucoside (Ononin) ^d : R ₁ = <i>O</i> -β-D-glucosyl; R ₂ = R ₃ = R ₄ = H; R ₅ = OCH ₃
	Dihydrochalcone Phloretin-3',5'-di- <i>C</i> -β-D-glucoside ^d : R = <i>C</i> -β-D-glucosyl
	Methylenedioxyisoflavone derivative Pseudobaptigenin ^a : R ₁ = H; R ₂ = OH Fujikinetin ^a : R ₁ = OCH ₃ ; R ₂ = OH Rothindin ^d : R ₁ = H; R ₂ = <i>O</i> -β-D-glucosyl

Chemical structure	Compound class, name and substituents
	Flavonol Kaempferol-5- <i>O</i> -glucoside ^b : R ₁ = R ₂ = H; R ₃ = <i>O</i> - α -D-glucosyl; R ₄ = OH Kaempferol-6- <i>C</i> -glucoside ^b : R ₁ = H; R ₂ = <i>C</i> - β -D-glucosyl; R ₃ = R ₄ = OH Kaempferol-8- <i>C</i> -glucoside ^b : R ₁ = <i>C</i> - β -D-glucosyl; R ₂ = H; R ₃ = R ₄ = OH Kaempferol-di-3- <i>O</i> ,6- <i>C</i> -glucoside ^b : R ₁ = H; R ₂ = <i>C</i> - β -D-glucosyl; R ₃ = OH; R ₄ = <i>O</i> - β -D-glucosyl
	Methylenedioxyflavonol derivative 3',4'-Methylenedioxyflavonol-6- <i>O</i> -apiosyl-glucoside ^b : R = <i>O</i> - α -apiofuranosyl-(1''' \rightarrow 6'')- β -D-glucopyranosyl
	Coumestan Medicagol ^a : R ₁ = H; R ₂ = OH Flemichapparin ^a : R ₁ = H; R ₂ = OCH ₃ Sophoracoumestan B ^a : R ₁ = OCH ₃ ; R ₂ = OH
	Flavan-3-ol (-)-epigallocatechin gallate ^c
	Phenylethanol derivative Tyrosol ^b : R ₁ = H; R ₂ = OH 3-Methoxy-tyrosol ^b : R ₁ = OCH ₃ ; R ₂ = OH Tyrosol-4- <i>O</i> -glucoside ^c : R ₁ = H; R ₂ = <i>O</i> - β -D-glucosyl Phenylethanol-3- <i>O</i> -apiosyl-glucoside ^b : R ₁ = <i>O</i> - α -apiofuranosyl-(1'' \rightarrow 6')- β -D-glucopyranosyl; R ₂ = H
	Benzaldehyde derivative Benzaldehyde-4- <i>O</i> -apiosyl-glucoside ^b : R = <i>O</i> - α -apiofuranosyl-(1'' \rightarrow 2')- β -D-glucopyranosyl
	Phenolic carboxylic acid <i>p</i> -Coumaric acid ^{a,c}
	Organic acid (\pm)-Shikimic acid ^c
	Inositol (+)-Pinitol ^{a,c}

[^aFerreira *et al.*, 1998; Kamara *et al.*, ^b2003, ^c2004; Kokotkiewicz *et al.*, ^d2012, ^e2013]

isomangiferin, were the most affected leading to significant decreases in content (De Beer & Joubert, 2010).

Joubert *et al.* (2010) reported the effect of different pre-drying techniques on the phenolic composition and colour of unfermented *C. subternata* leaves. The different treatments investigated included current commercial practices, i.e. steaming of the plant material directly after comminution (shredding). It was found that this treatment was detrimental to the colour and phenolic content. It was also postulated that scolymoside content increased due to the oxidation of eriocitrin with comminution and steaming.

The most recent evaluation of the effect of fermentation on honeybush composition was conducted by Theron (2012). Several *Cyclopia* species were subjected to different fermentation time and temperature combinations. It was concluded that certain time and temperature combinations are less detrimental to the phenolic composition, although these changes were not constant between species. Fermentation clearly reduced the total polyphenol content, as well as the content of individual polyphenolic compounds. Consequently, the increase in fermentation temperature and time lead to decrease in colour absorbance. Sensory attributes such as taste and mouthfeel were also affected by high temperature fermentation. Optimal fermentation conditions for *C. genistoides* and *C. maculata* were determined to be for 24 h at 80 °C, while *C. subternata* requires 24 h at 90 °C.

A summary of the differences in the content of several of the major phenolic compounds present between unfermented and fermented hot water extracts of different honeybush species is presented in Table 2.2, as adapted from De Beer & Joubert (2010).

Table 2.2 – Average content (g/100 g extract) of the major phenolic compounds in unfermented and fermented honeybush obtained from De Beer & Joubert (2010) (values outside and inside brackets indicate those of unfermented and fermented honeybush, respectively)

Compound	<i>C. intermedia</i>	<i>C. genistoides</i>	<i>C. sessiliflora</i>	<i>C. subternata</i>
Mangiferin	4.35 (0.13)	9.55 (1.67)	4.67 (0.59)	2.73 (0.06)
Isomangiferin	1.40 (0.26)	2.72 (1.12)	1.67 (0.78)	0.86 (0.15)
Eriocitrin	0.13 (0.03)	traces (nd)	0.32 (0.20)	0.32 (0.12)
Hesperidin	0.62 (0.27)	0.71 (0.31)	0.74 (0.38)	0.62 (0.24)
Eriodictyol glucoside	0.07 (nd)	nd (nd)	0.53 (nd)	0.35 (nd)
Scolymoside ^{a,b}	0.04 (nd)	traces (nd)	0.06 (traces)	0.68 (0.20)
Iriflophenone	1.04 (0.05)	0.98 (0.15)	4.22 (0.66)	0.82 (traces)
3- <i>C</i> - β -glucoside ^{a,c}				
Possible hydroxy-cinnamic acid	0.13 (0.11)	traces (nd)	nd (0.05)	0.20 (0.09)

^aKokotkiewicz *et al.* (2012)

^bPrevious designation - compound 11 (De Beer & Joubert, 2010)

^cPrevious designation - compound 8 (De Beer & Joubert, 2010)

nd = not detected; ne = not evaluated

2.3 High performance liquid chromatography

2.3.1 Introduction

High performance liquid chromatography (HPLC) is a well-known technique used for the separation of mixtures of compounds (Niessen, 2006). The technique is mainly used in analytical chemistry and biochemistry with the purpose of identifying known or unknown compounds and quantifying the individual substances present in mixtures (Niessen, 2006). This section will include a brief description of HPLC method development in the context of honeybush tea phenolics, with reference to previous methods and their limitations. A section on

method validation will follow. Furthermore, a summary of HPLC, hyphenated with mass spectrometry (MS) will be included, as well as a section on general MS fragmentation patterns of flavonoids and their interpretation for identification of compounds.

2.3.2 Method development

HPLC involves the mechanism of forcing a liquid sample over a solid adsorbent material packed into a column (stationary phase) under high pressure using a liquid solvent (mobile phase). The column, or stationary phase, typically consists of packing material of small particle size that creates a large surface area for interactions between the analyte and the stationary phase, leading to the separation of a mixture of compounds. Several modes of HPLC exist, differing in terms of the principle interactions involved in separation. For example normal phase HPLC involves the use of a polar stationary phase consisting of silica, and a non-polar mobile phase, for example hexane. This results in non-polar compounds eluting first, and polar compounds being retained longer in the column. Although described as 'normal', it is not the most commonly used form of HPLC. Reversed phase HPLC is much more popular (and the mode of HPLC used in this work) and uses a non-polar stationary phase and polar mobile phases. This results in non-polar compounds having stronger interactions with the stationary phase and therefore polar compounds elute first. The separation of compounds is illustrated as a chromatogram, for example absorbance being plotted as a function of time (Dong, 2006; Snyder *et al.*, 2010; Fanali *et al.*, 2013).

Throughout the process of developing a HPLC method, analytical conditions that provide satisfactory separation of selected sample components are determined (Snyder *et al.*, 2010). The quality of separation increases when the retention time difference between two peaks increase, or when the peaks become narrower. The primary goal during HPLC method development is to obtain good separation (resolution) of every peak of interest from its neighbouring peaks (Snyder *et al.*, 2010).

During method optimisation, only one parameter should be altered at a time in order to determine the relevant effect(s). To enhance separation, any of the following parameters can be altered: the percentage of the strong eluent (or gradient profile in the case of gradient analysis), i.e. the organic modifier in RP-LC; the pH of the mobile phase; the acid or base used to regulate the pH; temperature; the column stationary phase, particle size and/or column dimensions, and so forth. All of these parameters affect the retention times of the compounds, their resolution as well as the overall analysis time. Although many of these parameters and trends can be calculated, some of the finer "tweaking" will almost always amount to a trial and error process (Snyder *et al.*, 2010).

In general, the mobile phase composition, gradient profile and stationary phase directly affect the selectivity, whereas particle size affects the efficiency, resolution and speed (smaller particle sizes are beneficial, but there is a price to pay in terms of pressure). The mobile phase pH affects both the selectivity and efficiency due to its influence on analyte ionisation, whereas temperature has an effect on selectivity, retention (generally retention decreases with an increase in temperature, although there are notable exceptions) and speed. Most current silica-based HPLC columns are limited to temperatures below 60 °C (Snyder *et al.*, 2010; Fanali *et al.*, 2013). Previous literature can assist with choosing the initial conditions for method development.

In terms of HPLC analysis of honeybush phenolics, several columns have been investigated by Joubert *et al.* (2003) during a method development study for selected honeybush phenolic compounds, with the major focus on mangiferin, isomangiferin and hesperidin. These included a Multisphere C18 (3 μm , 125 \times 4 mm) column,

which displayed strong tailing of the phenolic compounds and also failed to separate the isomeric compounds mangiferin and isomangiferin from one another. A Synergi Polar RP column containing an ether-linked phenyl phase with polar end-capping (4 μm , 150 \times 4.6 mm, 80 Å) proved unable to sufficiently separate luteolin and eriodictyol (Joubert *et al.*, 2003). A Synergi Max-RP (C12-bonded silica, 4 μm , 80 Å) was found to provide the best separation (Joubert *et al.*, 2003).

In a further effort to improve the separation of all the major phenolic compounds in *Cyclopia* species including eriocitrin, which was insufficiently resolved from unidentified compounds when present in large quantities (Joubert *et al.*, 2003), different columns were evaluated again to obtain better selectivity (De Beer & Joubert, 2010). Columns used during this investigation were all 150 \times 4.6 mm, and included Synergi Fusion (polar-embedded ligand- and C18 silica, 4 μm , 80 Å), Synergi Max-RP (C12-bonded silica, 4 μm , 80 Å), Synergi Polar-RP (ether-linked phenyl-bonded silica, 4 μm , 80 Å), Synergi Hydro-RP (C18 with polar endcapping, 4 μm , 80 Å), Gemini C18 (hybrid C18-bonded silica and polymer, 5 μm , 110 Å), Luna phenyl-hexyl (phenyl-hexyl bonded silica, 3 μm , 100 Å), Prodigy phenyl-ethyl (phenyl-ethyl bonded silica, 5 μm , 100 Å) and Zorbax Eclipse XDB-C18 (extra dense C18 bonding with double endcapping, 5 μm , 80 Å) (De Beer & Joubert, 2010). In this study, temperature was also increased and the aqueous component of the mobile phase changed from 2% acetic acid to 0.1% formic acid to facilitate transfer of the method to LC-ESI-MS analysis. Separation of eriocitrin and unidentified peaks were optimal on the Zorbax Eclipse XDB-C18 column, whilst good separation for mangiferin, isomangiferin, hesperidin, eriodictyol, luteolin and unidentified compounds of interest was retained.

A concise summary of previous methods reported for the analysis of honeybush phenolics is presented in Table 2.3.

2.3.3 Method validation

In order for an analytical HPLC method to be employed and regarded as reliable, it should be validated. The developed method should meet the criteria of the intended application, and thus method validation delivers assurance in the form of documented confirmation (Snyder *et al.*, 2010). There are several standard parameters that should be evaluated during method validation. These and the main actions to be taken are summarised in Table 2.4 (as adapted from LoBrutto & Patel (2007), Anon. (2010) and Snyder *et al.* (2010)).

2.3.4 Detection methods

Nowadays the most common detectors used in combination with HPLC for phenolic compounds are UV-Vis or diode array detectors (DAD) and/or mass spectrometric (MS, or tandem MS, MS/MS) detectors equipped with atmospheric pressure ionisation chambers such as electrospray ionisation (ESI) sources (Cuyckens & Claeys, 2004; Abad-García *et al.*, 2009, 2012; Snyder *et al.*, 2010; Van der Hooft *et al.*, 2011).

2.3.4.1 Diode array detection

DAD is an essential tool for the detection and provisional identification of phenolic compounds in plants, as these molecules possess characteristic UV-Vis spectra (Abad-García *et al.*, 2009; Liu, 2011). DAD detectors are intended to deliver an output in absorbance which is linearly proportional to the injected sample concentration (according to the Beer-Lambert law), which makes them ideal for quantitative analysis (Snyder *et al.*, 2010).

Table 2.3 – Previous methods employed for separation of phenolic compounds in *Cyclopia* spp.

Objective species	Compounds quantified / identified	Mobile phase	Stationary phase	Gradient	Temperature	Reference
Method 1						
General <i>Cyclopia</i> spp.	Mangiferin, isomangiferin and hesperidin	A = 2% acetic acid and B = acetonitrile	Synergy MAX-RP C12 (4 μm , 80 \AA)	0-6 min (12% B), 7 min (18% B), 14 min (25% B), 19 min (40% B), 24 min (50% B), 29 min (12% B)	25 °C	Joubert <i>et al.</i> , 2003
Comment: Poor resolution of eriocitrin in <i>C. subternata</i> extracts with unidentified compounds, and co-elution of unidentified compounds with mangiferin and isomangiferin.						
Method 2						
<i>Cyclopia subternata</i> (applied to <i>C. genistoides</i> , <i>C. longifolia</i> , <i>C. intermedia</i> and <i>C. sessiliflora</i>)	Mangiferin, isomangiferin, eriocitrin, hesperidin eriodictyol & luteolin (iriflophenone-3- <i>C</i> -glucoside, scolymoside & phloretin-diglucoside were only detected)	A = 0.1% formic acid and B = acetonitrile	Zorbax Eclipse XDB-C18 (5 μm , 80 \AA)	0-6 min (12% B), 7 min (18% B), 14 min (25% B), 19 min (40% B), 24 min (50% B), 29 min (12% B)	30 °C	De Beer & Joubert, 2010
Comment: Co-elution of unidentified compounds with mangiferin and isomangiferin and complicated integration process due to the elution of polymeric compounds underneath eriocitrin, scolymoside and phloretin-3',5'-di- <i>C</i> -glucoside.						
Method 3						
<i>Cyclopia subternata</i> (ethyl acetate fraction)	No quantification. Mangiferin, isomangiferin, hesperidin, narirutin, scolymoside, iriflophenone-3- <i>C</i> -glucoside, phloretin-diglucoside, isohofolin, calycosin, pseudobaptigenin, & formononetin detected	A = 0.1% formic acid and B = 50% acetonitrile + 50% formic acid (0.1%)	Chromolith RP-18C	0 min (12% B), 15 min (25% B), 75-77 min (100% B), 80 min (12% B)	30 °C	Kokotkiewicz <i>et al.</i> , 2012
Comment: Poor resolution of most compounds, except iriflophenone-3- <i>C</i> -glucoside.						

Table 2.4 – Parameters used to validate an analytical method, as adapted from LoBrutto & Patel (2007), Anonymous (2010) and Snyder *et al.* (2010)

Parameter	Definition	Execution
Accuracy	Expresses the degree of closeness between the true value and the value found.	Comparison of results obtained from the proposed analytical procedure with known quantities of the substances in the sample.
Specificity	The ability to assess the exact compound in the presence of components which may be expected to be present in samples.	Suitable identification tests should be able to discriminate compounds from closely related structures which are likely to be present (e.g. LC-MS).
Precision	Expresses the degree of agreement between a series of measurements obtained from repeated sampling of the same homogeneous sample under the proposed conditions.	Usually expressed as the variance, standard deviation or coefficient of variation of a series of repeated measurements. Repeatability: a minimum of 9 determinations covering the specified range (e.g., 3 concentrations with 3 replicates) Intermediate precision: Variations such as days or analysts. Reproducibility: Inter-laboratory trial.
Detection limit	The lowest analyte concentration in a sample which can be detected (but not necessarily quantified as an exact value).	The minimum level at which the analyte can be detected. Compare measured signal of sample with low concentration to that of a blank sample, and determine signal-to-noise ratio (Limit is 2:1).
Quantitation limit	The lowest analyte concentration in a sample which can be quantified with suitable precision and accuracy.	Analysis of samples with known analyte concentrations and determining the minimum level at which the analyte can be quantified with acceptable accuracy and precision. Compare measured signal of sample with low concentration to that of a blank sample, and determine signal-to-noise ratio (Limit is 10:1).
Linearity	The ability (within a given range) of the proposed method to acquire results which are directly proportional (linear) to the concentration of analyte in the sample.	Obtained by injecting compounds or samples at different concentrations or dilutions, and determining the slope of the regression line, correlation coefficient and y-intercept
Range	The interval between the upper and lower analyte concentrations in the sample for which it has been demonstrated that the method has suitable levels of precision, accuracy and linearity.	Normally derived from linearity studies and should cover up to 130% of expected concentrations
Robustness	A measure of the method's capacity to remain unaffected by small, but deliberate variations in method parameters; provides an indication of its reliability during normal usage.	Deliberate variation of: - Mobile phase pH - Mobile phase composition - Different column - Temperature & flow rate

Compared to traditional UV-Vis detectors, DAD offers the advantage of on-line spectral acquisition, which may assist significantly in tentative compound identification. These detectors are potentially capable of relatively high sensitivity for analytes absorbing in the UV-Vis range and have a good linear range. Detectors based on UV-Vis absorbance are reasonably insensitive to slight changes in mobile phase flow rate and temperature (Snyder *et al.*, 2010).

2.3.4.2 Mass spectrometry

Liquid chromatography coupled to mass spectrometry (LC-MS) has become the preferred method for polyphenolic analysis (Abad-García *et al.*, 2009). In recent years, MS detection has become a standard method for obtaining structural information or for the confirmation of unknown compounds in research and development fields. MS detection involves three steps: (i) ionisation of analytes, in LC-MS commonly performed using ESI, (ii) mass separation, most often performed using quadrupole, time-of-flight (TOF) or tandem MS (MS/MS) analysers, and (iii) detection of the separated ions (Snyder *et al.*, 2010). Quadrupole detectors are most common due to their robustness and relatively low cost. TOF-MS instruments, although more expensive, provide high resolution data from which molecular formula information can be obtained. The ionisation polarity of the ionisation interface can be switched between negative and positive ionisation modes. In the case of ESI, negative ionisation mostly provides deprotonated ions, whilst the positive mode typically provides protonated ions or adducts (McMaster, 2005; Niessen, 2006; Snyder *et al.*, 2010). With the aid of collision induced dissociation (CID), more comprehensive structural information can be acquired from MS/MS when used in combination with CID (Cuyckens & Claeys, 2004). In MS/MS, two mass analysers are combined, typically with a CID cell placed between them for fragmentation purposes. Different analysers can be combined, such as for example triple quadrupole and quadrupole-TOF (Q-TOF) instruments. With regard to structural characterisation of phenolic compounds, information that can be obtained using MS and MS/MS includes the nature of the aglycone, as well as the degree of glycosylation (for example mono-, di-, tri- or tetrasaccharides, and so forth) (Cuyckens & Claeys, 2004). Furthermore, the attachment positions of the substituents to the aglycone can also be determined (Cuyckens & Claeys, 2004).

2.3.5 Identification of phenolic compounds

2.3.5.1 Elution order

The elution order of phenolic compounds in RP-HPLC is now well established (Abad-García *et al.*, 2009). Regarding the compounds related to honeybush, flavanones elute before dihydrochalcones, which elute before flavones. Moreover, within the same class, compounds containing more polar substituents such as hydroxyls will elute first, while apolar substituents, such as a methoxy group, results in increased retention (Abad-García *et al.*, 2009). Glycosylated phenolic compounds elute before their corresponding aglycones, while sugar acylation will increase the retention time (Abad-García *et al.*, 2009). For polyphenols, where glycosylation occurs at the same position, the elution order will be di-*O*-glycoside, *O*-galactoside, *O*-glucoside, *O*-rutinoside, *O*-neohesperidoside, *O*-rhamnoside and aglycone. The position of the sugar is regarded as more significant in determining the elution time than the nature of the sugar and therefore the 7-*O*-glycoside will be less retained than the 4'- or 3-*O*-glycosides. As for the type of glycosidic bond, *C*-glycosides elute before *O*-glycosides (Abad-García *et al.*, 2009).

2.3.5.2 UV-Vis detection

Phenolic compounds have characteristic UV-Vis spectral properties that are unique for each of the different classes of phenolics, and may therefore be used for tentative identification of the class of a phenolic compound. In honeybush, the major classes include the xanthenes, benzophenones, flavanones, dihydrochalcones and flavones. Typical UV-Vis spectra for each of these classes are illustrated in Fig. 2.2.

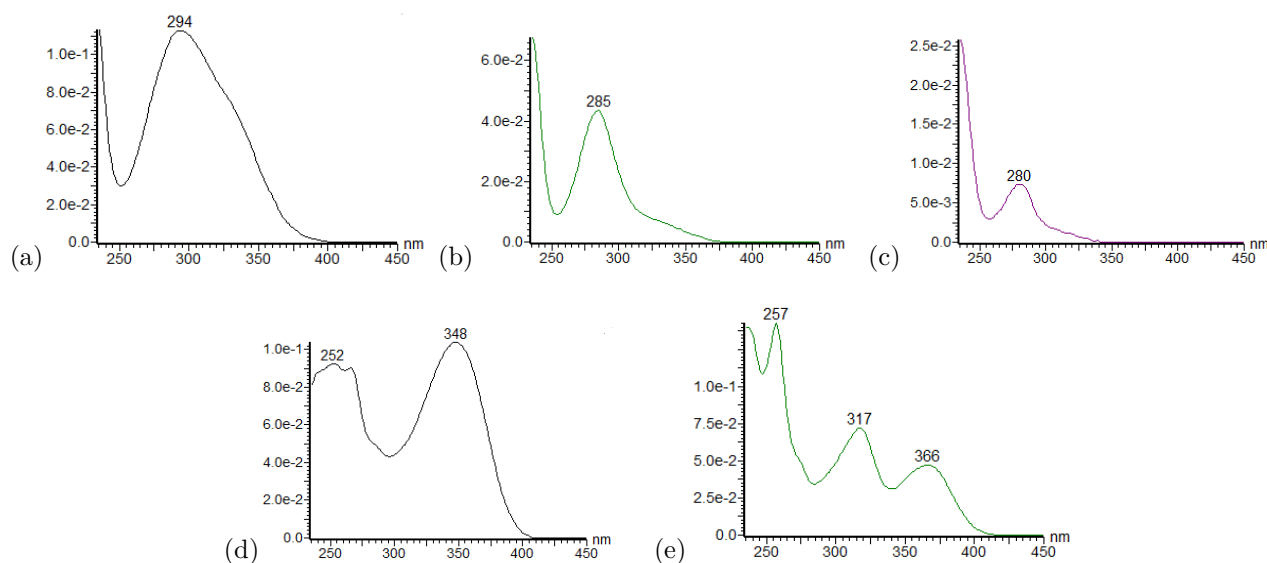


Figure 2.2 – UV-Vis spectral properties of the different phenolic classes present in honeybush (a) benzophenone (iriflophenone-3-*C*-glucoside), (b) dihydrochalcone (phloretin-3',5'-di-*C*-glucoside), (c) flavanone (eriocitrin), (d) flavone (scolymoside) and (e) xanthone (mangiferin).

2.3.5.3 MS and MS/MS detection

Fragmentation of phenolics

Identification of phenolics by MS is based on molecular ion information (including accurate mass information if available) as well as their respective fragmentation patterns. Certain methodological rules, together with relative retention orders and UV spectra, aid in the identification of compounds (Abad-García *et al.*, 2009). These will be summarised in the following section. Nomenclature used in this section is according to Domon & Costello (1988).

Differences between ionisation modes

Positive ionisation mode

Regarding the identification of flavonoid aglycones, the most informative fragmentation involves the cleavage of two *C-C* bonds in the *C*-ring (Fig. 2.3). This results in the formation of the structurally informative $^{i,j}A^+$ and $^{i,j}B^+$ ions. The level of collision energy used during CID, play a significant role in the fragmentation mechanisms observed. Ma *et al.* (1997) and Wolfender *et al.* (2000) reported that different ratios of fragment ions containing the *A* and *B* rings can be used to distinguish between aglycones such as flavonols and flavones. In addition to the formation of $^{i,j}A^+$ and $^{i,j}B^+$ ions, subsequent losses of small molecules and/or radicals from the $[M+H]^+$ ion may generally also follow. Losses of 18 amu (H_2O), 28 amu (CO), 42 amu (C_2H_2O) are frequently detected and are useful for determining the presence of particular functional groups (Cuyckens & Claeys, 2004).

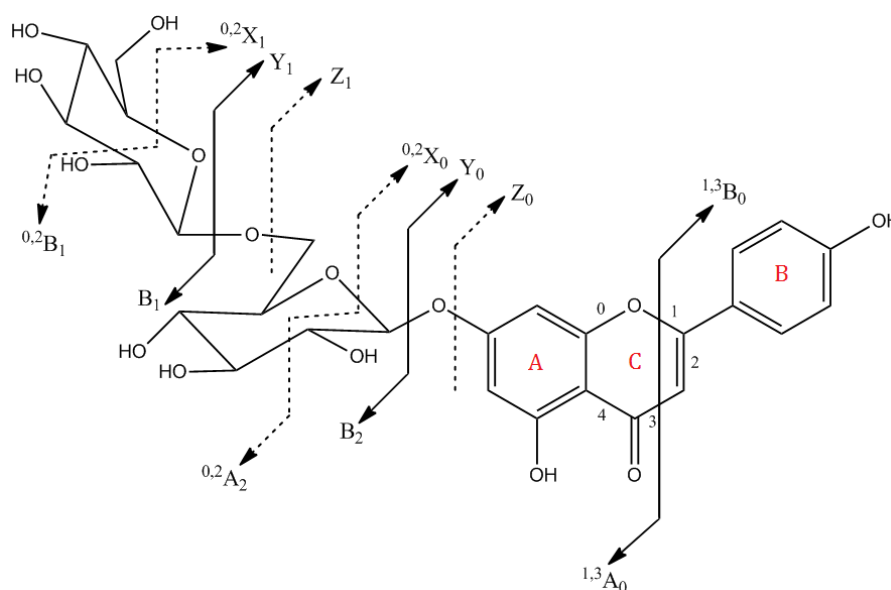


Figure 2.3 – Nomenclature used for fragmentation of flavonoid glycosides in mass spectrometry (Domon & Costello, 1988; Cuyckens & Claeys, 2004; Abad-García *et al.*, 2009).

Negative ionisation mode

During the structural analysis of flavonoids, the use of positive CID spectra is preferred over negative ionisation CID spectra, as the interpretation of the latter is frequently considered to be more difficult (Cuyckens & Claeys, 2004). The negative ionisation mode usually requires higher collision energies compared to positive ion mode to produce satisfactory fragmentation. Occasionally ions observed in the positive ionisation mode, which are helpful in the structural elucidation of a compound, are lacking in negative ionisation mode. However, this mode is considered more sensitive for flavonoid analysis and the fragmentation behaviour is different, delivering additional and complementary data.

Fabre *et al.* (2001) have studied the detailed MS fragmentation mechanisms of numerous aglycones in the negative ionisation mode. The $^{1,3}A^-$ fragment is frequently the major fragment ion in the negative ionisation mode, whereas the $^{0,3}B^-$ ion is usually the major peak for isoflavones, and is therefore regarded as characteristic for this class of phenolic compounds (Cuyckens & Claeys, 2004).

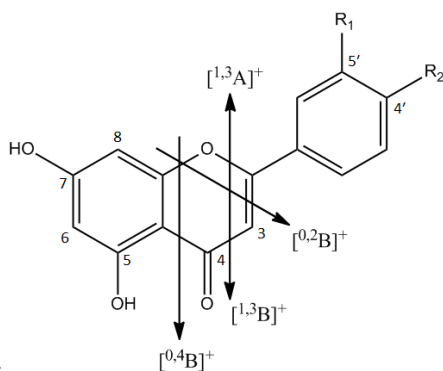
The extent of hydroxylation of the B-ring influences the fragmentation. Flavonols that have two or more hydroxyl groups attached to the B-ring will produce ions corresponding to $[^{1,2}A-H]^-$ and $[^{1,2}B+H]^-$. When the B-ring is unsubstituted, the corresponding collision energy needed to produce fragmentation, is considerably higher, and therefore results in many product ions (Cuyckens & Claeys, 2004). Small neutral losses such as CO (-28 amu), CO₂ (-34 amu), C₂H₂O (-42 amu) can also be observed (Fabre *et al.*, 2001; Cuyckens & Claeys, 2004), as well as some rearrangements such as retro-Diels-Alder (RDA) reactions (Cuyckens & Claeys, 2004).

2.3.5.4 Aglycones

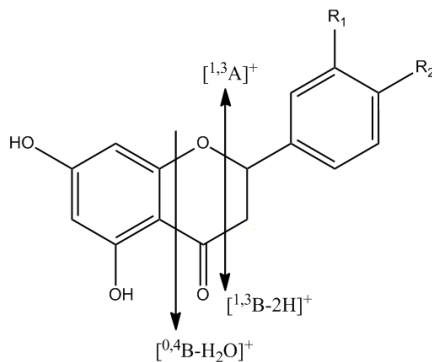
This section briefly describes the fragmentation of aglycones relevant to this study, as illustrated in Fig. 2.4. In the following discussion the focus will be on positive mode ESI detection, since this mode is mostly used to identify compounds.

For hydroxylated flavones such as luteolin, a series of ions can be observed that are characteristic of the C-ring cleavage: $[^{0,4}B_0]^+$, $[^{0,4}B_0-H_2O]^+$, $[^{1,3}A_0]^+$ (base peak), $[^{0,2}B_0]^+$, $[^{1,3}B_0]^+$, $[^{1,3}A_0-C_2H_2O]^+$, $[^{1,3}B_0-CO]^+$, $[^{1,3}B_0-H_2O]^+$ and $[^{1,3}B_0-H_2O-CO]^+$ (Abad-García *et al.*, 2009).

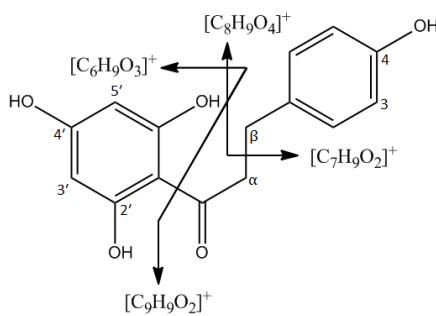
Flavones:



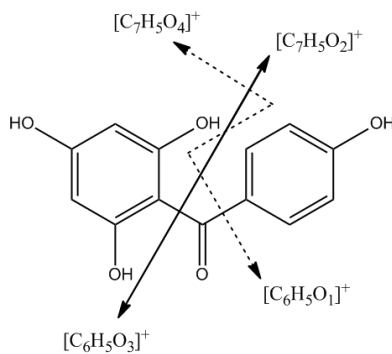
Flavanones:



Dihydrochalcones:



Benzophenones:



Xanthones:

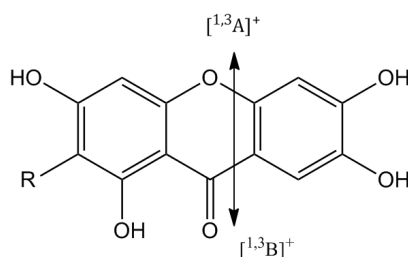


Figure 2.4 – Structures and fragmentation pathways for selected phenolic aglycones, indicated for positive ionisation mode (Abad-García *et al.*, 2009; Yu *et al.*, 2013; Zhang *et al.*, 2013). Carbon numbering for all classes are similar to indicated for flavones, with the exception of dihydrochalcones (indicated separately).

In the case of flavanone aglycones, the ion $[^{1,3}A_0]^+$ at m/z 153 is typically the most intense fragment (the base peak), as in the case for flavonols. Other fragments such as $[^{1,3}A_0 - CO]^+$, $[^{1,3}A_0 - C_2H_2O]^+$, $[^{1,3}B_0 - 2H]^+$ and $[^{0,4}B_0 - H_2O - CO]^+$ can also be detected. According to Abad-García *et al.* (2009), the cleavage of the C-ring in flavanones occurs relatively easily.

Unlike the rest of the flavonoid families, dihydrochalcones are unique in that they do not contain a heterocycle C-ring, but rather a saturated open 3-C chain. This chain fragments easily at low collision energies, resulting in the $[C_9H_9O_2]^+$ and $[C_7H_7O]^+$ product ions. Ions such as $[C_8H_9O_4]^+$ and $[C_6H_7O_3]^+$ are also found to be common (Abad-García *et al.*, 2009).

According to Zhang *et al.* (2013) fragmentation of the benzophenone aglycone entails cleavage of the bonds connecting the aromatic rings as indicated in Fig. 2.4. This results in the product ions $[C_6H_5O_3]^+$ and $[C_7H_5O_2]^+$, as well as $[C_6H_5O_1]^+$ and $[C_7H_5O_4]^+$.

Fragmentation of xanthenes is suggested to occur in the γ -pyrone skeleton. Similar fragmentation as described for flavones is also observed for this class (Yu *et al.*, 2012).

2.3.5.5 Flavonoid glycosides

Flavonoid O-glycosides

For flavonoid O-glycosides, low collision energies are sufficient to produce fragmentation at the glycosidic O-C-linkage. The typical glycosylation sites for the acid-labile glycosidic C-O bonds are located at the C3 and C7 positions (Abad-García *et al.*, 2009).

When cleavage occurs at the glycosidic O-linkages, monosaccharide residues are eliminated with losses of 162, 146 or 132 amu for hexoses, deoxyhexoses and pentoses, respectively (Cuyckens & Claeys, 2004). The two common disaccharides associated with flavonoids are rutinose [rhamnosyl-(α 1 \rightarrow 6)-glucose] and neohesperidose [rhamnosyl-(α 1 \rightarrow 2)-glucose], where the disaccharide is attached to a single hydroxyl group of the aglycone (Abad-García *et al.*, 2009).

In Fig. 2.5 the fragmentation pathway of an O-diglycoside is illustrated, with the Y_1 fragment originating from the loss of a carbohydrate moiety and Y_0 from the loss of two carbohydrate moieties. The corresponding carbohydrate fragments are denoted B_1 and B_0 (De Rijke *et al.*, 2006).

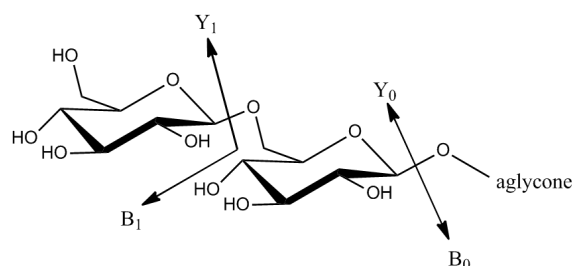


Figure 2.5 – The main fragmentation pathways observed for O-glycosidic and O-diglycosidic flavonoids (De Rijke *et al.*, 2006; Abad-García *et al.*, 2009).

Flavonoid C-glycosides

In the case of a flavonoid C-glycoside, the sugar is directly linked to the flavonoid nucleus through an acid-hydrolysis resistant C-C bond, with C6 and C8 being the most common binding positions (Cuyckens & Claeys, 2004; Abad-García *et al.*, 2009). Flavonoid C-glycosides require higher collision energies to fragment compared

to *O*-glycosides. Under these conditions, fragmentation occurs in the sugar itself, which has the weakest bonds in the molecule (Abad-García *et al.*, 2009). The major fragmentations observed entail cross-ring cleavages of the sugar moiety as illustrated in Fig. 2.6 and the successive loss of water molecules (Cuyckens & Claeys, 2004).

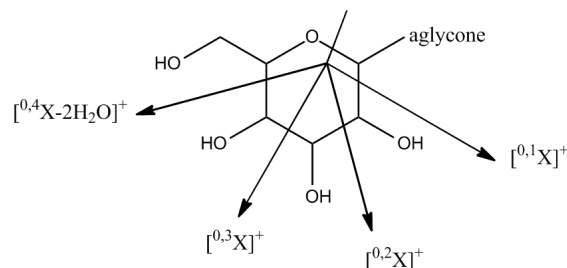


Figure 2.6 – The main fragmentation pathways observed for *C*-glycosidic flavonoids (Waridel *et al.*, 2001; Cuyckens & Claeys, 2004; Abad-García *et al.*, 2009).

As indicated in Fig. 2.6, the $^{0,1}X^+$, $^{0,2}X^+$, $^{0,3}X^+$ and $^{0,4}X^+-2H_2O$ fragments represent $[M+H-150]^+$, $[M+H-120]^+$, $[M+H-90]^+$ and $[M+H-96]^+$ ions for hexoses (Waridel *et al.*, 2001; Cuyckens & Claeys, 2004; Abad-García *et al.*, 2009). The additional loss of water molecules are represented by $[M+H-18]^+$ ions (Abad-García *et al.*, 2009). In the case of flavonoid-di-*C*-glycosides, the simultaneous fragmentation of both sugars is common, leading to a relatively large number of fragments being detected.

Flavonoid *O,C*-diglycosides

O,C-diglycosides contain a sugar linked to a hydroxyl groups of the aglycone, as well as another sugar coupled *via* the *C-C* bond (Abad-García *et al.*, 2009). These types of flavonoids can be identified by the occurrence of both *O*-glycosidic and *C*-glycosidic fragments (Abad-García *et al.*, 2009).

2.4 Fingerprints

2.4.1 Introduction

Quality control of herbal products is of great importance as it is known that the efficacy of these products are dependent on the concentrations of their bioactive compounds (Tistaert *et al.*, 2011). Factors such as climate, cultivation conditions, harvest time, drying, storage and adulteration can cause large variation in the active compound concentrations. Because of these variations, efficacy can be reduced, leading to health risks (Tistaert *et al.*, 2011). With the increasing number of herbal products appearing on the market, efficient quality control of these products is becoming more difficult, requiring the investigation of new quality evaluation systems (Fan *et al.*, 2006). Thus, it is becoming important to evaluate a more comprehensive chemical profile of a sample, instead of only a single compound (Liang *et al.*, 2004). Fingerprinting is one such technique that has gained increasing attention and has been accepted by leading organisations such as the World Health Organisation (WHO), Federal Drug Administration (FDA) and the British Herbal Medicine Association (Anon., 2002; Fan *et al.*, 2006). A fingerprint is a specific profile or pattern which chemically represents a sample, based on the detected compounds (Alaerts *et al.*, 2010, Tistaert *et al.*, 2011). Fingerprints can be obtained by numerous separation methods including capillary electrophoresis (EC), gas chromatography (GC) and HPLC (Alaerts *et al.*, 2010; Gad *et al.*, 2013).

2.4.2 Chromatographic fingerprints

Chromatographic fingerprints are chromatograms which represent the chemical characteristics of a herbal product (Fan *et al.*, 2006). It is generally accepted that samples comprising similar chromatographic fingerprints, have similar chemical and pharmacological properties (Fan *et al.*, 2006; Alaerts *et al.*, 2010). This concept is used to evaluate the identity, authenticity and consistency of herbal products. In the case of HPLC separation, only chromatograms from single wavelengths in the UV-Vis range were previously employed to evaluate consistency between different herbal product batches. This is, however, an inadequate representation of the chemical profile of the product, especially since herbal products are known to be very complex. It has been deemed impossible to develop an appropriate analytical method which can represent the entire chemical characteristic pattern with a single vector chromatogram (Fan *et al.*, 2006). Marker compounds have also been used to determine quality, but might also not always be suitable due to the lack of unique compounds and the possible synergistic effects between the constituents (Liang *et al.*, 2004; Liang *et al.*, 2009; Tistaert *et al.*, 2011). Consequently, chromatograms over multiple wavelengths have to be used to represent the sample adequately (Fan *et al.*, 2006). For this purpose, several chromatograms can be extracted, or the entire UV-Vis spectrum (when using HPLC-DAD) should preferably be used (Tistaert *et al.*, 2011).

Numerous quality control methods have been developed for herbal products using chromatographic fingerprinting as basis, combined with advanced data analysis and multivariate data analysis techniques (Liang *et al.*, 2009). With the aid of chemometrics, the (dis)similarity between different samples can be evaluated (Liang *et al.*, 2009) and compounds associated with bioactivity can potentially be identified (Kvalheim *et al.*, 2011). When employed to screen products for the presence of bioactive compounds, fingerprint analysis provides a more cost effective manner in which to screen samples for activity compared to using animal models (Liang *et al.*, 2009, Kvalheim *et al.*, 2011).

2.5 Data analysis of fingerprints

When a method has been chosen or developed for instrumental analysis (such as HPLC-DAD), it is important to remember that the data will be analysed at a later stage. Depending on the selected analytical technique (for example HPLC-DAD), various factors can influence the results of such a method (such as retention time shifts). Univariate data analysis includes the examination of one changing element at a time, while in multivariate analyses two or more predefined experiments are performed and several influential factors are examined simultaneously (Tistaert *et al.*, 2011). Whether the aim is to visually inspect or statistically compare data, there is a wide variety of tools available to achieve the desired goal.

2.5.1 Pre-treatment of chromatographic fingerprints

The raw chromatographic data obtained from instrumental analyses usually consist of chemical information and random variation. This unwanted variation in data can include shifts in peak retention time, background noise and slope inconsistencies. When these variations are left untreated, it may lead to data analysis providing inaccurate results. Therefore the pre-treatment of chromatographic data is essential (Trygg *et al.*, 2009; Alaerts *et al.*, 2010; Gad *et al.*, 2013).

2.5.1.1 Alignment

In chromatographic analysis, it is common for elution times to shift slightly between different injections due to column ageing and other random instrumental and experimental variation. These unavoidable changes in chromatographic data have to be corrected to enable data analysis of the relevant variations. The proper alignment technique can correct these small changes by aligning the information from all the corresponding peaks of different chromatograms (Fig. 2.7). It has been demonstrated that alignment is one of the most important pre-treatment techniques for chromatographic data (Alaerts *et al.*, 2012).

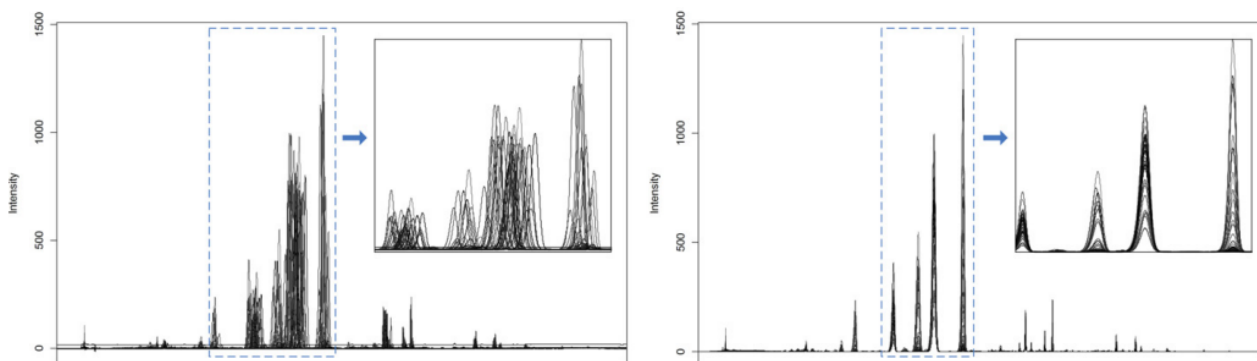


Figure 2.7 – Example of signal alignment (Zhang *et al.*, 2012).

Several alignment techniques have already been developed including correlation optimized warping (COW) (Nielsen *et al.*, 1998), dynamic time warping (DTW) (Sakoe and Chiba, 1971; 1978), parametric time warping (PTW) (Eilers, 2004) and semi-parametric time warping (STW) (Eilers, 2004), with development of automated alignment methods (Daszykowski *et al.*, 2010; Tomasi *et al.*, 2011) in progress. Only a few selected methods will be explained in the next sections as reviewed by Van Nederkassel *et al.* (2006), Skov (2008), Alaerts *et al.* (2010) and Daszykowski *et al.* (2010).

Target selection

Most alignment algorithms use a reference or target chromatogram towards which to align, therefore the process of the selection of the target chromatogram is an extremely important task. Not all signal chromatograms are expected to have the same amount of peaks or comparable shapes, but a good reference signal should represent all the other samples as well as possible in terms of the number of peaks. The target chromatogram should not be biased towards a certain part of the signal and thus only have intermediate shifts in retention time (Jellema, 2009). According to Daszykowski & Walczak (2007) it is best to use correlation coefficients to determine the target chromatogram by evaluating the correlation coefficients of all the chromatograms with each other. The best target signal will be the chromatogram which exhibits the highest correlation coefficient with all of the profiles. It has been shown to provide the most acceptable results when compared to other methods (Daszykowski & Walczak, 2007). Other methods include selecting the first chromatogram that was generated in the sequence of injections, as it was subjected to the least amount of column aging, and is regarded as analysed under the most favourable conditions (Eilers, 2004). This does not guarantee that it will be the most representative within the entire sample set. Another method includes creating a mean chromatogram from the whole sample set, however this could cause deformation of peaks (Daszykowski & Walczak, 2007). Choosing an improper target chromatogram could lead to more unwanted variance and inadequate alignment.

Correlation optimised warping

COW, a method introduced by Nielsen *et al.* (1998), is a piecewise alignment technique that uses dynamic programming to align a sample chromatogram towards a reference or target chromatogram by means of stretching and compressing (warping) segments of the time axis by means of linear interpolation.

Each chromatogram is aligned or warped to the target separately, one segment at a time (piecewise). Dynamic programming aids in determining the correct or optimal pathway for warping by using mathematical methods to solve combinatorial problems (Tomasi *et al.*, 2004). During this procedure all the possible combinations of conversions of the time axis are examined, while only two parameters need to be set (Nielsen *et al.*, 1998). The first input parameter is the number of segments into which the chromatogram should be divided. The second is the slack size, also known as the flexibility of the warping. This is the length (in data points) by which the each segment may increase or decrease due to stretching and compressing from warping (Tomasi *et al.*, 2004; Skov, 2008). After warping each segment, the correlation coefficient between the target chromatogram and the samples is determined and saved, before moving on to the next segment (Nielsen *et al.*, 1998).

The performance of the warping is determined by calculating the sum of all the correlation coefficients. The optimal warping path is established by identifying the path with the highest sum of correlation coefficients (Skov, 2008). It should be noted, however, that the flexibility of the warping increases toward the middle of the vector. Therefore caution should be taken to not over-align the peaks in the middle when it is challenging to align peaks in the beginning or end of the chromatogram (Van Nederkassel *et al.*, 2006). Over-alignment of peaks may lead to distortion of the peak shape, which would have negative implications during data analysis (Daszykowski *et al.*, 2010). As the optimal alignment can be determined by correlation calculations, no knowledge of the data is required beforehand and no peak detection or integration is necessary (Nielsen *et al.*, 1998). COW does not require the signals to be of equal length, although the signals may be shortened to reduce computational efforts (Van Nederkassel *et al.*, 2006).

Dynamic time warping

Sakoe and Chiba (1971, 1978) were the first to present the algorithm for DTW, which was developed further in the following years (Tomasi *et al.*, 2004). In contrast to COW, DTW is a nonlinear warping method, making use of the principle of dynamic programming. DTW warps two signals (reference and sample) by aligning similar events in the reference and sample chromatograms (Pravdova *et al.*, 2002). The target chromatogram is placed on the x-axis while the sample chromatogram is placed on the y-axis of a plot. A grid is positioned on the plot and the algorithm constructs a path, linking these similar events or patterns to each other, while minimising the cumulative distance between them (Pravdova *et al.*, 2002; Tomasi *et al.*, 2004). A graphical representation of DTW is presented in Fig. 2.8.

There are two versions of the DTW path, symmetric and asymmetric. The symmetric algorithm considers both the target and the sample chromatogram as equally significant where the time indexes of the similar events are mapped as a common time index. This allows for the x- and y-axes to be interchangeable, while obtaining the same optimal path (Pravdova *et al.*, 2002). The asymmetric algorithm does not consider the two chromatograms as equal, as one is the reference. This algorithm will not allow for the same optimal path to be obtained when interchanging the chromatograms (Pravdova *et al.*, 2002). Alignment by DTW has an important element which can be viewed as a limitation or constraint; the beginning and the ends of the chromatograms or signals are fixed, resulting in difficulty to align early or late eluting peaks (Tomasi *et al.*, 2004).

2006; Skov, 2008). In most applications, it is recommended that 40 B-splines should be used as default (Van Nederkassel *et al.*, 2006). STW has the same prerequisites as PTW regarding baseline removal, smoothing, and normalisation (Van Nederkassel *et al.*, 2006).

Automatic alignment methods

Manual alignment methods are time-consuming and labour intensive, creating a need for automated alignment methods. Daszykowski *et al.* (2010), developed an automatic alignment method which focuses on the alignment of one-dimensional chromatographic signals. No assumptions are made regarding the shape of the warping function to ensure flexible alignment (Daszykowski *et al.*, 2010). The method is optimised by a few B-spline functions, as in STW. The key to this automated alignment is in determining the coefficients of the spline functions. This automatic alignment method maintains high flexibility of the warping function while the input parameter is selected in an objective manner.

An optimisation strategy has also been introduced by Tomasi *et al.* (2011), with a technique named *icoshift*, which optimises the warping effect with the aid of a so-called ‘discrete-coordinate simplex-like optimisation routine’. This includes establishing the combination of all the relevant segment length and slack sizes (as used for COW). Currently, these two methods have been suggested to be used prior to COW, as a preliminary alignment method to reduce time consumed during manual alignment (Daszykowski *et al.*, 2010; Tomasi *et al.*, 2011).

2.5.1.2 Normalisation

Normalisation is a way of standardising the raw data and is another very important pre-treatment technique (Hendriks *et al.*, 2005). Depending on the type of data analysed, normalisation can either be essential, or not at all. There are several different normalisation methods, and selecting the correct method is also very important. During HPLC analysis, small experimental (for example weighing of samples) and instrumental (for example injection volume) variations are inevitable. These variations should be eliminated before data analysis as they can affect the outcome (Bylesjö *et al.*, 2009). Although there are many normalisation methods, for chromatographic data normalisation is conducted by dividing each signal value for each sample by the sum of all the signal values for that specific sample (Hendriks *et al.*, 2005).

2.5.1.3 Signal-to-noise ratio improvement

When the signal-to-noise ratio (SNR) is relatively small, it can be improved using a technique often referred to as denoising (Fig. 2.9). SNR methods involve reducing or smoothing the random high-frequency components (noise) in the chromatograms (Trygg *et al.*, 2009). Digital filters can also be used in either the time domain (direct) or the frequency domain (indirect) (Daszykowski & Walczak, 2006). It should be noted that smoothing does not increase the accuracy or the precision of the signal, and when applied incorrectly, smoothing can lead to an even noisier signal than before treatment. Therefore care should be taken when applying these kinds of methods (Thekkudan & Rutan, 2009).

2.5.1.4 Baseline elimination

The chromatographic baseline fluctuations should be eliminated, as illustrated in Fig. 2.10 as part of pre-treatment of chromatographic data. The variation in the baseline greatly influences calculations based on

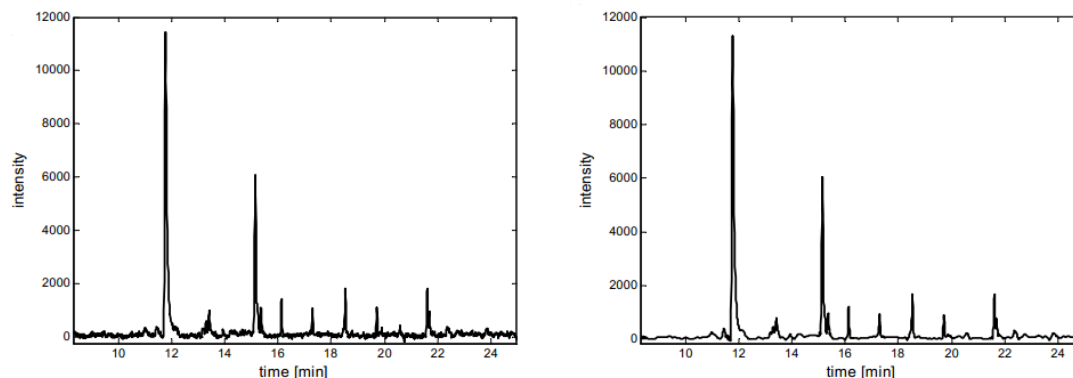


Figure 2.9 – Example of signal-to-noise improvement (Daszykowski *et al.*, 2009).

correlation coefficients and Euclidean distances, such as similarity analysis (Daszykowski & Walczak, 2006). Eilers (2003, 2004) proposed the penalised asymmetric least squares (PALS) method. This method approximates the baseline, after which it is subtracted from the signal. Input parameters (the degree of derivatives and a penalty parameter) are required that involve optimisation by the user (Daszykowski & Walczak, 2006).

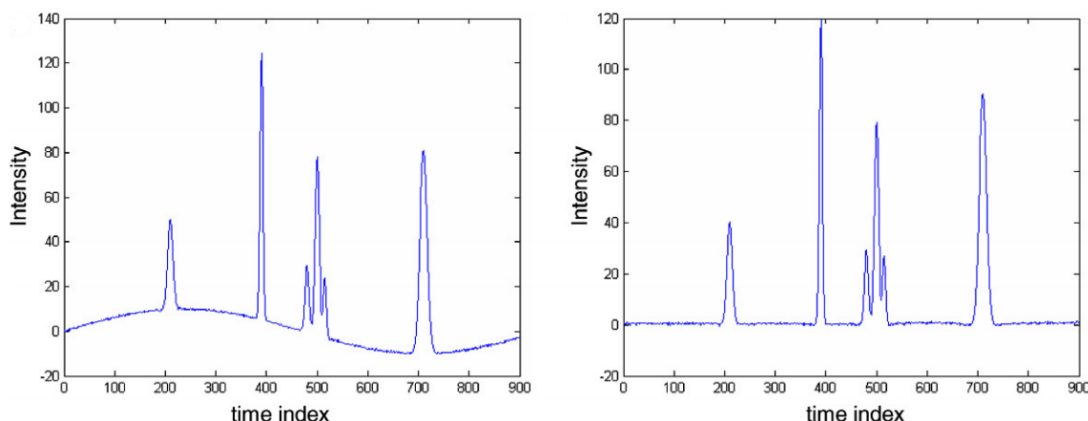


Figure 2.10 – Example of baseline elimination of a chromatographic signal (Daszykowski & Walczak., 2006).

2.5.2 Multivariate data analysis

There are several different types of multivariate data analysis techniques. The desired outcome of the experiment will determine the type of analysis that should be considered. The first type is referred to as unsupervised data analysis. Unsupervised data analysis is used to explore data without having any prior knowledge about the data. This entails the analysis of only matrix \mathbf{X} , for example the fingerprint data matrix. The aim of unsupervised data analysis is to reduce the data complexity and explore the data. Unsupervised data analysis encompasses techniques such as similarity analysis, exploratory analysis and curve resolution (Alaerts *et al.*, 2010). The second is termed supervised data analysis which entails the analysis of, for example, the fingerprint data (matrix \mathbf{X}), as well as another set of data (vector \mathbf{y}), which provides external information on the fingerprints. This type of analysis includes techniques such as pattern recognition or classification, and multivariate calibration (Alaerts *et al.*, 2010).

2.5.2.1 Unsupervised data analysis

Similarity analysis

Similarity analysis is one of the most common multivariate data analysis tools for chromatographic data, as it is based on correlation coefficients (r) and congruence coefficients (c) (Alaerts *et al.*, 2010, 2012; Tistaert *et al.*, 2011). By definition, the correlation coefficient between a pair of chromatograms or fingerprints is ‘equivalent to the scalar product of the mean centered fingerprints’ (Alaerts *et al.*, 2010). The congruence coefficient is the calculated correlation with respect to the origin (not the mean as calculated for r) (Alaerts *et al.*, 2012). Similarity is indicated by values ranging from 0 to 1, with 0 being the least similar and 1 being most similar. Furthermore, when the entire fingerprint is not used, one can base the similarity analysis on individual compounds, including their relative retention time and relative peak area (Alaerts *et al.*, 2010, 2012; Tistaert *et al.*, 2011).

Reference samples can also be included in the similarity analysis, for quality evaluation purposes. The r or c values for samples relative to the reference sample are calculated to determine the (dis)similarity. Both the r and c can be used in hierarchical clustering analysis, which is an exploratory analysis technique. As authentic standards are not always readily available, a mean or median fingerprint or signal is usually composed, and the similarity of individual chromatograms are determined against the compiled chromatogram for quality control (Alaerts *et al.*, 2010, 2012; Tistaert *et al.*, 2011).

Exploratory data analysis

In exploratory analysis insight can be provided into the data by revealing possible structures and clusters in the data, such as for example harvest time, origin and species. These techniques include principal component analysis (PCA) (Wold *et al.*, 1987) and cluster analysis (Vandeginste *et al.*, 1998). PCA is a method that reduces the number of variables by transforming original variables into principal components (PCs). The first PC usually contains the most information (Wold *et al.*, 1987).

Clustering can be divided into hierarchical and non-hierarchical techniques, although both methods group data into different clusters. Hierarchical clustering results in a tree or a dendrogram where the samples are classified, compared to non-hierarchical clustering where the data clusters are merely represented in a table, with no specific order or hierarchy. Therefore hierarchical clustering is used more frequently (Wold *et al.*, 1987).

All exploratory analysis techniques allow for visualisation of the data to reveal groupings, trends, and outlier samples (Alaerts *et al.*, 2010). Typically PCA can provide score and loading plots, which illustrate information from the objects and the variables, respectively. Biplots on the other hand illustrate the information of both the scores and loadings in one plot for easier interpretation.

2.5.2.2 Supervised data analysis

It is widely recommended to perform a form of exploratory analysis before applying any supervised analysis techniques. Visualising the data prior to further investigation is very important before any type of calibration or pattern recognition should be applied. This is to ensure that outliers are not present.

Multivariate classification

Pattern recognition can be used to classify samples into different groups or classes based on their properties. Calibration is used to determine these classes beforehand, and these calibrated classes are then used to predict

the class of unknown samples. Some of the methods commonly used include partial least squares discriminant analysis (PLS-DA) (Vandeginste *et al.*, 1998), classification and regression trees (CART) (Breiman *et al.*, 1984) and linear and quadratic discriminant analysis (LDA and QDA) (Fisher, 1936). Although these methods were built on different principles, testing each of the different methods can be valuable in order to determine which method will provide the best predictive ability for a specific sample set.

Multivariate calibration

In methods of multivariate calibration, a property of the samples, for instance their antioxidant activity, is used to construct a model from their fingerprints. Such models can then be used to predict the antioxidant activity of new samples. Techniques which incorporate multivariate calibration include principal component regression (PCR) (Vandeginste *et al.*, 1998), robust PCR (rPCR) (Hubert & Verboven, 2003), partial least squares (PLS) (Wold *et al.*, 2001), robust PLS (rPLS) (Wakeling & Macfie, 1992) and uninformative variable elimination PLS (UVE-PLS) (Centner *et al.*, 1996). For calibration models it is good practice to cross validate and report the root mean squared error of this cross validation (RMSECV). External validation should also be conducted, with errors reported as the root mean squared error of prediction (RMSEP). The RMSECV and RMSEP are used to determine the reliability of the model and serve as the average errors to expect when using the constructed model (Vandeginste *et al.*, 1998).

2.5.3 The effect of pre-treatment techniques on multivariate data analysis

Several examples will be used to indicate how pre-treatment of chromatographic fingerprint data can influence multivariate data analysis. To indicate the importance of alignment of fingerprints, Gong *et al.* (2004) analysed a chromatographic data set before and after this applying pre-treatment. In this study, the chemical composition of four herbal medicines were determined quantitatively and qualitatively by means of HPLC-DAD. The chromatographic data sets of the herbal medicines were analysed by means of PCA to determine inhomogeneity within the samples. In Fig. 2.11 the score plots of PC1 vs. PC2, before and after alignment, are displayed.

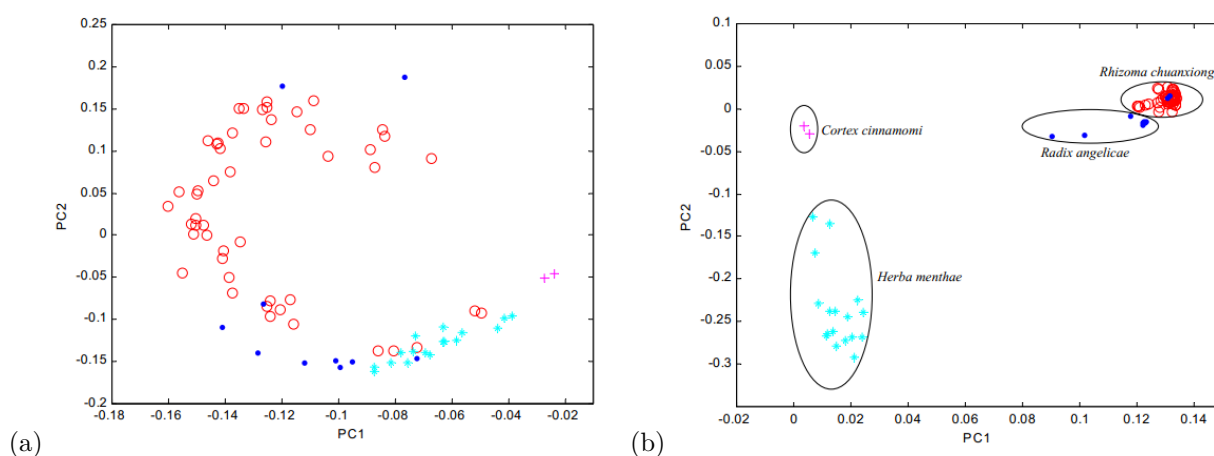


Figure 2.11 – PCA score plots (a) before and (b) after alignment. Different symbols represent different plant species (Gong *et al.*, 2004).

No distinct groupings were observed in the PCA scores plot before alignment (Fig. 2.11) due to the comparison of irrelevant information in the fingerprints. After appropriate alignment of the chromatograms, the PCA scores plot showed clear clustering of the samples representing the four plant species (Gong *et al.*, 2004).

In another study conducted by Lucio-Gutiérrez *et al.* (2012), the antioxidant capacity of *Turnera diffusa* extracts was predicted by means of PLS regression on chromatographic fingerprint data. HPLC-DAD chromatograms from four wavelengths were used to construct a single fingerprint for each sample. Pre-treatment techniques applied to the fingerprints included baseline correction, scaling, and COW for the alignment of peaks. A PLS model was constructed from the data sets before and after pre-treatment, respectively.

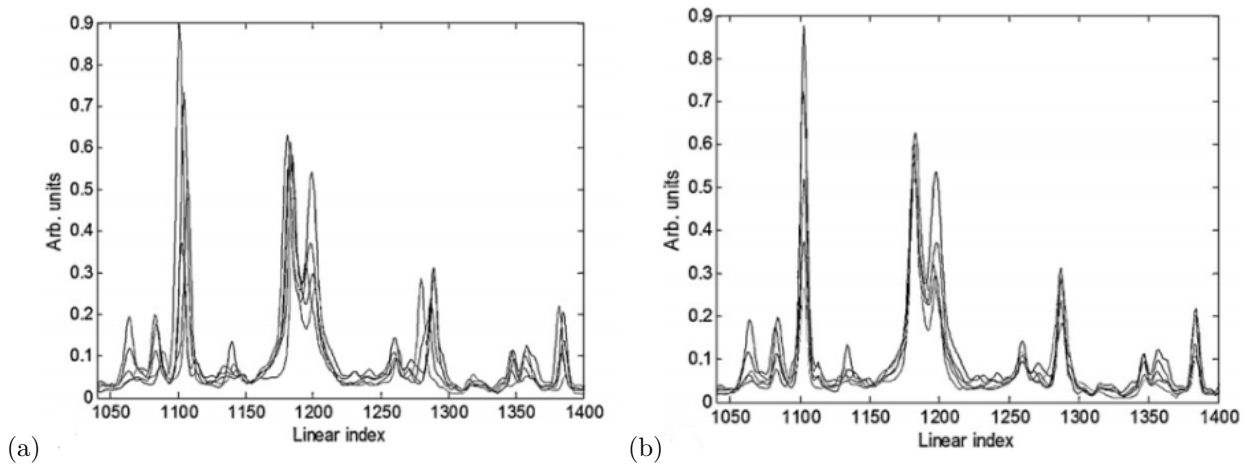


Figure 2.12 – HPLC-DAD chromatograms of *Turnera diffusa* extracts (a) before and (b) after alignment (Lucio-Gutiérrez *et al.*, 2012).

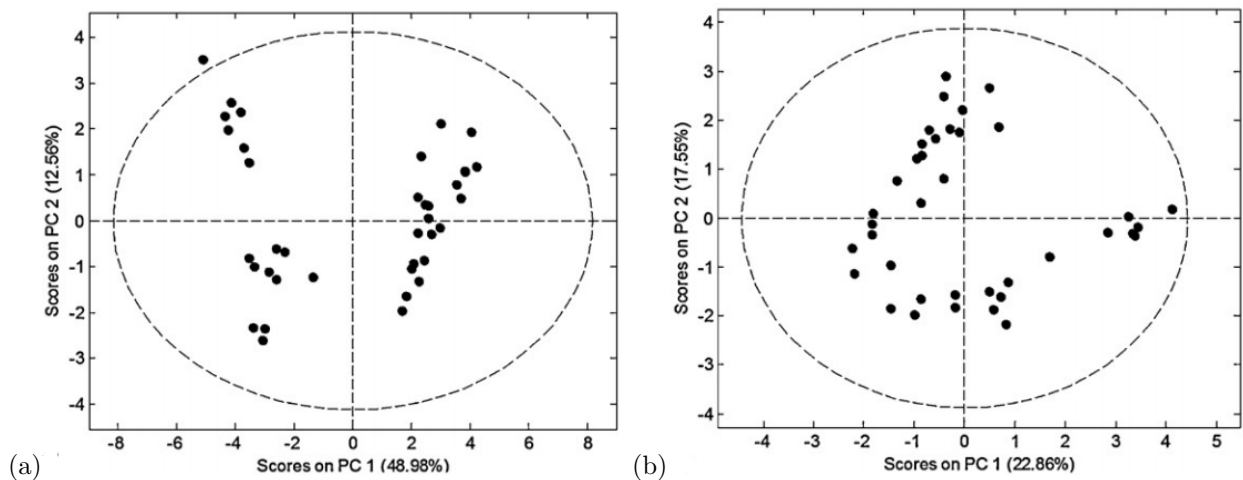


Figure 2.13 – PCA score plots for HPLC-DAD analysis of *Turnera diffusa* extracts (a) before and (b) after alignment (Lucio-Gutiérrez *et al.*, 2012).

As illustrated in Fig. 2.12, the original data are distorted as a result of retention time shifts, while the aligned profiles are no longer distorted. In Fig. 2.13 the score plots of PC1 vs. PC2 are presented before and after alignment. Before alignment, the scores plot indicated irrelevant groupings of samples. After alignment, the scores plot showed a more homogenous scatter of the samples. To determine whether there was any relationship between the antioxidant activity and the scores, they were plotted against each other (Fig. 2.14). Before alignment there was no apparent relationship visible compared to after alignment where a linear relationship was evident (Lucio Gutiérrez *et al.*, 2012).

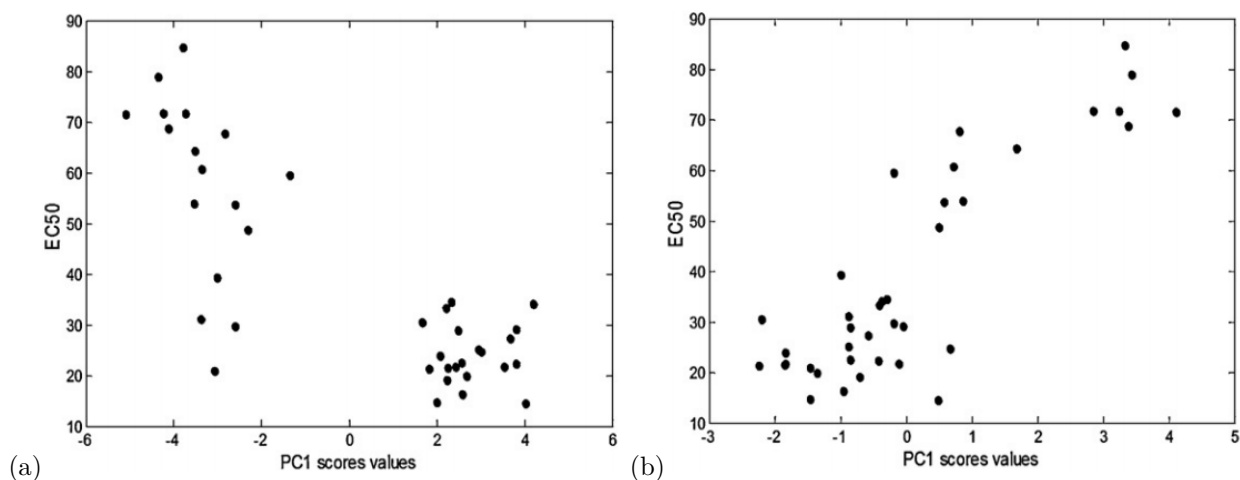


Figure 2.14 – PLS regression plots for HPLC-DAD analysis of *Turnera diffusa* extracts (a) before and (b) after alignment (Lucio-Gutiérrez *et al.*, 2012).

2.5.4 Application of chemometrics to fingerprint analysis

The importance of pre-treating chromatographic data has been stressed. The selection of the correct pre-treatment techniques is dependent on the chemometric method to be applied to the data. Applying incorrect pre-treatment can have detrimental effects on data analysis. Table 2.5 lists examples of the different uses of chemometrics with regard to herbal products, as well as the different pre-treatment techniques that were applied.

Table 2.5 – Examples of the application of HPLC fingerprints in chemometric analysis of different herbal products

Purpose of study	Method of Analysis	Pre-treatment techniques	Chemometric methods	Plant species	References
Geographic origin	UPLC-DAD	Averaged	Similarity analysis Hierarchical cluster analysis Principal component analysis	<i>Coptidis chinensis</i>	Kong <i>et al.</i> , 2009
	HPLC-UV	Alignment Normalisation	Principal component analysis	<i>Salix</i> sp.	Hendriks <i>et al.</i> , 2005
Quality assessment	HPLC-DAD	Averaged Relative peak areas	Similarity analysis	<i>Yiqing</i>	Li <i>et al.</i> , 2010
	HPLC	Alignment Averaged	Partial least squares Partial robust M-regression	<i>Camellia sinensis</i> (green tea)	Daszykowski <i>et al.</i> , 2007
	HPLC	Alignment Averaged	Similarity analysis	<i>Camellia sinensis</i> (green tea)	Alaerts <i>et al.</i> , 2012
	HPLC	Background correction Alignment	Principal component analysis Partial least squares	<i>Turnera diffusa</i>	Lucio-Gutiérrez <i>et al.</i> , 2012
	HPLC-MS	Alignment	Principal component analysis Orthogonal projections to latent structures	<i>Mallotus</i> species	Tistaert <i>et al.</i> , 2012
Taxonomic discrimination	HPLC-DAD	Relative peak areas Single wavelength selection	Similarity analysis Hierarchical cluster analysis	<i>Aconitum kusnezoffii</i> and related species	Zhao <i>et al.</i> , 2009
	HPLC-DAD	Averaged	Similarity analysis Principal component analysis	Chinese <i>Angelica</i>	Lu <i>et al.</i> , 2005

2.6 Conclusions

Quality control of honeybush is of great importance for the food and nutraceutical market. Limited information is currently available on the qualitative and quantitative phenolic composition of *Cyclopia subternata* and *C. maculata* water extracts. HPLC-DAD and -MS remain important tools for the analysis of phenolic composition. HPLC-DAD methods that provide good chemical representation of *C. subternata* and *C. maculata* extracts are required to produce chromatographic fingerprints adequate for multivariate data analysis. Chromatographic fingerprints in combination with chemometric techniques could be used to determine the similarity of new *Cyclopia* extracts to those with proven bioactive properties.

References

- Anonymous (2002). Traditional Medicine Strategy 2002-2005, World Health Organisation, Geneva, Switzerland. http://whqlibdoc.who.int/hq/2002/who_edm_trm_2002.1.pdf (February 2013).
- Anonymous (2007). Corrigendum to Regulation (EC) No 1924/2006 of the European Parliament and of the council of 20 December 2006 on nutrition and health claims made on foods. *Official Journal of the European Union*, Article 13.
- Anonymous (2010). Guidance for Industry: Analytical Procedures and Method Validation, Final Guidance. Food and Drug Administration. Maryland, USA.
- Abad-García, B., Berrueta, L.A., Garmón-Lobato, S., Gallo, B & Vicente, F. (2009). A general analytical strategy for the characterization of phenolic compounds in fruit juices by high-performance liquid chromatography with diode array detection coupled to electrospray ionization and triple quadrupole mass spectrometry. *Journal of Chromatography A*, **1216**, 5398-5415.
- Abad-García, B., Garmón-Lobato, S., Berrueta, L.A., Gallo, B & Vicente, F. (2012). On line characterization of 58 phenolic compounds in *Citrus* fruit juices from Spanish cultivars by high-performance liquid chromatography with photodiode-array detection coupled to electrospray ionization triple quadrupole mass spectrometry. *Talanta*, **99**, 213-224.
- Alaerts, G., Dejaegher, B., Smeyers-Berbeke, J. & Vander Heyden, Y. (2010). Recent developments in chromatographic fingerprints from herbal products: Set-up and data analysis. *Combinatorial Chemistry and High Throughput Screening*, **13**, 900-922.
- Alaerts, G., Van Erps, J., Pieters, S., Dumarey, M., van Nederkassel, A.M., Goodarzi, M., Smeyers-Verbeke, J. & Vander Heyden, Y. (2012). Similarity analysis of chromatographic fingerprints as tools for identification and quality control of green tea. *Journal of Chromatography B*, **910**, 61-70.
- Asami, D.K., Hong, Y., Barrett, D.M. & Mitchell, A.E. (2003). Comparison of total phenolic and ascorbic acid content of freeze-dried and air-dried marionberry, strawberry, and corn grown using conventional, organic, and sustainable agricultural practices. *Journal of Agricultural and Food Chemistry*, **51**, 1237-1241.
- Bowie, J. (1930). Sketches of the botany of South Africa. *South African Quarterly Journal*, **1**, 27-36.
- Bylesjö, M., Cloarec, O. & Rantalainen, M. (2009). Normalization and closure. In: *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, Volume II (edited by S.D. Brown, R. Tauler & B. Walczak). Pp. 109-125. Netherlands: Oxford University Press.
- Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984). *Classification and Regression Trees*. New York, USA: Chapman & Hall.
- Centner, V., Massart, D.L., De Noord, O.E., de Jong, S., Vandeginste, B.G.M. & Sterna, C. (1996). Elimination of uninformative variables for multivariate calibration. *Analytical Chemistry*, **68**, 3851-3858.
- Cuyckens, F. & Claeys, M. (2004). Mass spectrometry in the structural analysis of flavonoids. *Journal of Mass Spectrometry*, **39**, 1-15.
- Daszykowski, M. & Walczak, B. (2006). Use and abuse of chemometrics in chromatography. *Trends in Analytical Chemistry*, **25**, 1081-1096.
- Daszykowski, M. & Walczak, B. (2007). Target selection for alignment of chromatographic signals obtained using monochannel detectors. *Journal of Chromatography A*, **1176**, 1-11.
- Daszykowski, M., Vander Heyden, Y., Boucon, C. & Walczak, B. (2010). Automated alignment of one-dimensional chromatographic fingerprints. *Journal of Chromatography A*, **1217**, 6127-6133.

- De Beer, D. & Joubert, E. (2010). Development of HPLC method for *Cyclopia subternata* phenolic compound analysis and application to other *Cyclopia* spp. *Journal of Food Composition and Analysis*, **23**, 289-297.
- De Beer, D., Schulze, A.E., Joubert, E., De Villiers, A., Malherbe, C.J. & Stander, M.A. (2012). Food ingredients extracts of *Cyclopia subternata* (honeybush): Variation in phenolic composition and antioxidant capacity. *Molecules*, **17**, 14602-14624.
- De Lange (2002). Geskiedenis en ontwikkeling van heuningbostee. *South African Honeybush Tea Association Newsletter*, **6**, 5-12.
- De Lange, H. & von Mollendorff, L. (2006). Huidige stand van plantverbetering in die heuningbostee-bedryf. *South African Honeybush Tea Association Newsletter*, **14**, 8-15.
- De Nysschen, A.M., Van Wyk, B.E., Van Heerden, F.R. & Schutte, A.L. (1996). The major phenolic compounds in the leaves of *Cyclopia species* (Honeybush tea). *Biochemical and Systematic Ecology*, **24**, 243-246.
- De Rijke, E., Out, P., Niessen, W.M.A., Ariese, F., Gooijer, C. & Brinkman, U.A.T. (2006). Review - Analytical separation and detection methods for flavonoids. *Journal of Chromatography A*, **1112**, 31-63.
- Domon, B. & Costello, C.E. (1988). A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal* **5**, 397-409.
- Dong, M.W. (2006). *Modern HPLC for Practicing Scientists*. New Jersey, USA: John Wiley & Sons, Inc.
- Du Toit, J. & Joubert, E. (1998). The effect of pre-treatment on the fermentation of honeybush tea (*Cyclopia maculata*). *Journal of the Science of Food and Agriculture*, **76**, 537-545.
- Du Toit, J., Joubert, E. & Britz, T.J. (1998). Honeybush tea- A rediscovered indigenous South African herbal tea. *Journal of Sustainable Agriculture*, **12**, 67-84.
- Dudhia, Z., Louw, J., Muller, C., Joubert, E., de Beer, D., Kinnear, C. & Pheiffer, C. (2013). *Cyclopia maculata* and *Cyclopia subternata* (honeybush tea) inhibits adipogenesis in 3T3-L1 pre-adipocytes. *Phytomedicine*, **20**, 401-408.
- Eilers, P.H.C. (2003). A perfect smoother. *Analytical Chemistry*, **75**, 3631-3636.
- Eilers, P.H.C. (2004). Parametric time warping. *Analytical Chemistry*, **76**, 404-411.
- Espín, J.C., García-Conesa, M.T. & Tomás-Barberán, F.A. (2007). Nutraceuticals: facts and fiction. *Phytochemistry*, **68**, 2986-3008.
- Fabre, N., Rustan, I., de Hoffmann, E., Quentin-Leclercq, J. (2001). Determination of flavone, flavonol and flavanone aglycones by negative ion LC-ES ion trap mass spectrometry. *Journal of the American Society for Mass Spectrometry*, **12**, 707-715.
- Fan, X., Cheng, Y., Ye, Z., Lin, R. & Qian, Z. (2006). Multiple chromatographic fingerprinting and its application to the quality control of herbal medicines. *Analytica Chimica Acta*, **555**, 217-224.
- Fanali, S., Haddad, P.R., Poole, C., Schoenmakers, P., & Lloyd, D.K. (2013). *Liquid Chromatography: Fundamentals and Instrumentation*. USA: Elsevier.
- Feria-Morales, A. (2002). Examining the case of green coffee to illustrate the limitations of grading systems/expert tasters in sensory for quality control. *Food Quality and Preference*, **13**, 355-367.
- Ferreira, D., Kamara, B.I., Brandt, E.V. & Joubert, E. (1998). Phenolic compounds from *Cyclopia intermedia* (honeybush tea). *Journal of Agricultural and Food Chemistry*, **46**, 3406-4310.
- Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, **7**, 179-188.
- Gad, H.A., El-Ahmady, S.H., Abou-Shoer, M.I. & Al-Azizi, M.M. (2013). Application of chemometrics in authentication of herbal medicines: A review. *Phytochemical Analysis*, **24**, 1-24.

- Gong, F., Liang, Y., Fung, Y. & Chau, F. (2004). Correction of retention time shifts for chromatographic fingerprints of herbal medicine. *Journal of Chromatography A*, **1029**, 173-183.
- Greenish, H.G. (1881). Cape tea. *The Pharmaceutical Journal and Transactions*, **11**, 549-551.
- Hendriks, M.M.W.B., Cruz-Juarez, L., De Bont, D. & Hall, R.D. (2005). Preprocessing and exploratory analysis of chromatographic profiles of plant extracts. *Analytica Chimica Acta*, **545**, 53-64.
- Hubert, M.; Verboven, S. A robust PCR method for high-dimensional regressors. *J. Chemometrics*, 2003, **17**, 438-452
- Jellema, R.H. (2009). Variable shift and alignment. In: *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, Volume II. Pp. 85-108. Netherlands: Oxford University Press.
- Jiang, Y., David, B., Tu, P. & Barbin, Y. (2010). Recent analytical approaches in quality control of traditional Chinese medicines - a review. *Analytica Chimica Acta*, **657**, 9-18.
- Joubert, E., Gelderblom, W.C.A., Louw, A. & De Beer, D. (2008a). South African herbal teas: *Aspalathus linearis*, *Cyclopia* spp. and *Athrixia phylicoides* - A review. *Journal of Ethnopharmacology*, **119**, 376-412.
- Joubert, E., Joubert, M.E., Bester, C., De Beer, D. & De Lange, J.H. (2011). Honeybush (*Cyclopia* spp.): From local cottage industry to global markets - The catalytic and supporting role of research. *South African Journal of Botany*, **77**, 887-907.
- Joubert, E., Manley, M., Maicu, C. & De Beer, D. (2010). Effect of pre-drying treatments and storage on color and phenolic composition of green honeybush (*Cyclopia subternata*) herbal tea. *Journal of Agriculture and Food Chemistry*, **58**, 338-344.
- Joubert, E., Otto, F., Grüner, S. & Weinreich, B. (2003). Reversed-phase HPLC determination of mangiferin, isomangiferin and hesperidin in *Cyclopia* and the effect of harvesting date on the phenolic composition of *C. genistoides*. *European Food and Research Technology*, **216**, 270-273.
- Joubert, E., Richards, E.S., Van der Merwe, J.D., De Beer, D. Manley, M. & Gelderblom. W.C.A. (2008b). Effect of species variation and processing on phenolic composition and in vitro antioxidant activity of aqueous extracts of *Cyclopia* spp. (honeybush tea). *Journal of Agricultural and Food Chemistry*, **56**, 954-963.
- Kamara, B.I., Brand, D.J., Brandt, E.V. & Joubert, E. (2004). Phenolic metabolites from honeybush tea (*Cyclopia subternata*). *Journal of Agricultural and Food Chemistry*, **52**, 5291-5395.
- Kamara, B.I., Brandt, E.V., Ferreira, D. & Joubert, E. (2003). Polyphenols from honeybush tea (*Cyclopia intermedia*). *Journal of Agricultural and Food Chemistry*, **51**, 3874-3879.
- Kies, P. (1951). Revision of the genus *Cyclopia* and notes on some other sources of bush tea. *Bothalia*, **6**, 161-176.
- Koch, I. (2011). Development of a sensory lexicon and sensory wheel for rooibos (*Aspalathus linearis*) and the role of its composition on taste and mouthfeel. MSc Food Science Thesis, Stellenbosch University, Stellenbosch, South Africa.
- Kokotkiewicz, A., Luczkiewicz, M., Pawlowska, J., Luczkiewicz, P., Sowinski, P., Witkowski, J., Bryl, E. & Bucinski, A. (2013). Isolation of xanthone and benzophenone derivatives from *Cyclopia genistoides* (L.) Vent. (honeybush) and their pro-apoptotic activity in synoviocytes from patients with rheumatoid arthritis. *Fitoterapia*, **90**, 199-208.
- Kokotkiewicz, A., Luczkiewicz, M., Sowinski, P., Glod, D., Gorynski, K. & Bucinski, A. (2012). Isolation and structure elucidation of phenolic compounds from *Cyclopia subternata* Vogel (honeybush) intact plant and

in vitro cultures. *Food Chemistry*, **133**, 1373-1382.

- Komes, D., Belščak-Cvitanović, A., Horžić, D., Rusak, S., Likić, S. & Berendika, M. (2011). Phenolic composition and antioxidant properties of some traditionally used medicinal plants affected by the extraction time and hydrolysis. *Phytochemical Analysis*, **22**, 172-180.
- Kong, W., Zhao, Y., Xiao, X., Jin, C. & Li, Z. (2009). Quantitative and chemical fingerprint analysis for quality control of Rhizoma *Coptidis chinensis* based on UPLC-PAD combined with chemometric methods. *Phytomedicine*, **16**, 950-959.
- Kvalheim, O.M., Chan, H., Benzie, I.F.F., Szetto, Y., Hing-chung Tzang, A., Kam-wah Mok, D. & Chau, F. (2011). Chromatographic profiling and multivariate analysis for screening and quantifying the contributions from individual components to the bioactive signature in natural products. *Chemometric and Intelligent Laboratory Systems*, **107**, 98-105.
- Lattanzio, V., Lattanzio, V.M.T. & Cardinali, A. (2006). Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects. In: *Phenolics and Plant Defence* (edited by F. Imperato). India: Research Signpost.
- Li, Y., Wu, T., Zhu, J., Wan, L., Yu, Q., Li, X., Cheng, Z. & Guo, C. (2010). Combinative method using HPLC fingerprint and quantitative analyses for quality consistency evaluation of an herbal medicinal preparation produced by different manufacturers. *Journal of Pharmaceutical and Biomedical Analysis*, **52**, 597-602.
- Liang, X., Xie, P. & Chan, K. (2004). Quality control of herbal medicines. *Journal of Chromatography B*, **812**, 53-70.
- Liang, X., Jin, Y., Wang, Y., Jin, G. Fu, Q. & Xiao, Y. (2009). Qualitative and quantitative analysis in quality control of traditional Chinese medicines. *Journal of Chromatography A*, **1216**, 2033-2044.
- Liu, W.J.H. (2011). *Traditional Herbal Medicine Research Methods: Identification, Analysis, Bioassay, and Pharmaceutical and Clinical Studies*. New Jersey, USA: John Wiley & Sons, Inc.
- LoBrutto, R. & Patel, T. (2007). Method Validation. In: *HPLC for Pharmaceutical Scientists* (edited by Y.V. Kazakevich and R. LoBrutto). Pp. 455-463. New Jersey, USA: John Wiley & Sons, Inc.
- Lu, G., Chan, K., Liang, Y., Leung, K., Chan, C., Jiang, Z. & Zhao, Z. (2005). Development of high-performance liquid chromatographic fingerprints for distinguishing Chinese *Angelica* from related umbelliferae herbs. *Journal of Chromatography A*, **1073**, 383-392.
- Lucio-Gutiérrez, J.R., Garza-Jauréz, A., Coello, J., Masphoch, S., Salazar-Cavazos, M.L., Salazar-Aranda, R. & Waksman de Torres, N. (2012). Multi-wavelength high-performance liquid chromatographic fingerprints and chemometrics to predict the antioxidant activity of *Turnera diffusa* as part of its quality control. *Journal of Chromatography A*, **1235**, 68-76.
- Ma, Y.L., Li, Q.M., Van den Heuvel, H. & Claeys, M. (1997). Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, **11**, 1357-1364.
- Marloth, R. (1913). The chemistry of South African plants and plant products. *Cape Chemical Society*, p.9.
- Marloth, R. (1925). *The Flora of South Africa with Synoptical Tables of the Genera of the Higher Plants*. Pp. 69-72. Cape Town, South Africa: Darter Bros & Co.
- McMaster, M.C. (2005). *LC/MS: A Practical User's Guide*. New Jersey, USA: John Wiley & Sons, Inc.
- Mose Larsen, P., Fey, S.J., Louw, J. & Joubert, E. (2008). An anti-diabetic extract of honeybush. World Intellectual Property Organization WO. 2008/110552A2.

- Muller, C.J.F., Joubert, E., Gabuza, K., De Beer, D., Fey, S.J. & Louw, J. (2011). Assessment of the antidiabetic potential of an aqueous extract of honeybush (*Cyclopia intermedia*) in streptozotocin and obese resistant Wistar rats. In: *Phytochemicals - bioactivities and impact on health* (edited by I. Rasooli). Pp. 313-332. Rijeka: Intech.
- Nielsen, N.P.V., Carstensen, J.M. & Smedsgaard, J. (1998). Aligning single and multiple wavelength chromatographic data analysis using correlation optimised warping. *Journal of Chromatography A*, **805**, 17-35.
- Niessen, W.M.A. (2006). *Liquid Chromatography-Mass Spectrometry*. Florida, USA: CRC Press.
- Pravdova, V., Walczak, B. & Massart, D.L. (2002). A comparison of two algorithms for warping of analytical signals. *Analytica Chimica Acta*, **456**, 77-92.
- Sakoe, H. & Shiba, S. (1971). A dynamic programming approach to continuous speech recognition. Proceedings of the 7th international congress of acoustics. **20**(C13).
- Sakoe, H. & Shiba, S. (1978). Dynamic-programming algorithm optimization for spoken word recognition. *IEEE Trans Acoustics Speech Signal Process ASSP*, **26**, 43-49.
- Schutte, A.L. (1997). Systematics of the genus *Cyclopia* Vent. (Fabaceae, Podalyriaceae). *Edinburgh Journal of Botany*, **54**, 125-170.
- Skov, T. (2008). Mathematical resolution of complex chromatographic measurements. PhD Food Science Thesis, University of Copenhagen, Denmark.
- Snyder, L.R., Kirkland, J.J. & Dolan, J.W. (2010). *Introduction to Modern Liquid Chromatography*, 3rd ed. New Jersey, USA: John Wiley & Sons, Inc.
- Terblanche, S.E. (1982). Report on Honeybush Tea. Department of Biochemistry, University of Port Elizabeth, Port Elizabeth, South Africa.
- Thekkudan, D.F. & Rutan, S.C. (2009). Denoising and signal-to-noise ratio enhancement: Classical filtering. In: *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, Volume II (edited by S.D. Brown, R. Tauler & B. Walczak). Pp. 9-23. Netherlands: Oxford University Press.
- Theron, K.A. (2012). Sensory and phenolic profiling of *Cyclopia* species (honeybush) and optimisation of the fermentation conditions. MSc Food Science Thesis, Stellenbosch University, Stellenbosch, South Africa.
- Tistaert, C., Dejaegher, B., Chataigné, G., Rivière, C., Nguyen Hoai, N., Chau Van, M., Quetin-Leclercq, J. & Vander Heyden, Y. (2012). Potential antioxidant compounds in *Mallotus* species fingerprints. Part II: Fingerprint alignment, data analysis and peak identification. *Analytica Chimica Acta*, **721**, 35-43.
- Tistaert, C., Dejaegher, B. & Vander Heyden, Y. (2011). Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Analytica Chimica Acta*, **690**, 148-161.
- Tomasi, G., van der Berg, F. & Andersson, C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics*, **18**, 231-241.
- Tomasi, G., Savaroni, F. & Engelsen, S.B. (2011). *icoshift*: An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, **1218**, 7832-7840.
- Trygg, J., Gabrielsson, J. & Lundstedt, T. (2009). Background estimation, denoising, and preprocessing. In: *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, Volume II (edited by S.D. Brown, R. Tauler & B. Walczak). Pp. 1-6. Netherlands: Oxford University Press.
- Van der Hooft, J.J.J., Vervoort, J., Bino, R.J., Beekwilder, J. & de Vos, R.C.H. (2011). Polyphenol identification based on systematic and robust high-resolution accurate mass spectrometry fragmentation. *Analytical Chemistry*, **83**, 409-416.

- Van Nederkassel, A.M., Daszykowski, M., Eilers, P.H.C. & Vander Heyden, Y. (2006). A comparison of three algorithms for chromatographic alignment. *Journal of Chromatography A*, **1118**, 199-210.
- Vandeginste, B.G.M., Massart, D.L., Buydens, L.M.C., De Jong, S., Lewi, P.J. & Smeyers-Verbeke, J. (1998). Multivariate Calibration. *Data Handling in science and technology; Handbook of Chemometrics and Quality Metrics B*, **20**, 349-381.
- Wakeling, I.N. & Macfie, H.J.H. (1992). A robust PLS procedure. *Journal of Chemometrics*, **6**, 189-198.
- Waridel, P., Wolfender, J., Ndjoko, K., Hobby, K.R., Major, H.J. & Hostettmann, K. (2001). Evaluation of quadrupole time-of-flight tandem mass spectrometry and ion-trap multiple-stage mass spectrometry for the differentiation of C-glycosidic flavonoid isomers. *Journal of Chromatography A*, **926**, 29-41.
- Watt, J.M. & Breyer-Brandwijk, M.G. (1962). Medicinal and Poisonous Plants of Southern and Eastern Africa, 2nd ed. Edinburgh, UK: E & S Livingstone.
- Wold, S., Esbensen, K & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **2**, 37-52.
- Wold, S., Sjostrom, M. & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**, 109-130.
- Wolfender, J.L., Waridel, P., Ndjoko, K., Hobby, K.R., Major, H.J. & Hostettmann, K. (2000). Evaluation of Q-TOF-MS/MS and multiple stage IT-MDⁿ for the dereplication of flavonoids and related compounds in crude plant extracts. *Analisis*, **28**, 895-906.
- Yao, L., Caffin, N., D'Arcy, B., Jiang, Y., Shi, J., Singanusong, R., Liu, X., Datta, N., Kakuda, Y. & Xu, Y. (2005). Seasonal variations of phenolic compounds in Australia-grown tea (*Camellia sinensis*). *Journal of Agricultural and Food Chemistry*, **53**, 6477-6483.
- Yu, Q., Qi, J., Yu, H., Chen, L., Kou, J., Liu, S & Yu, B. (2013). Qualitative and quantitative analysis of phenolic compounds in the leaves of *Aquilaria sinensis* using liquid chromatography-mass spectrometry. *Phytochemical analysis*, **24**, 349-356.
- Zhang, Y., Han, L., Ge, D., Liu, X., Liu, E., Wu, C., Gao, X. & Wang, T. (2013). Isolation, structural elucidation, MS profiling, and evaluation of triglyceride accumulation inhibitory effects of benzophenone C-glucosides from the leaves of *Mangifera indica* L. *Journal of Agricultural and Food Chemistry*, **61**, 1884-1895.
- Zhang, Z., Liang, Y., Lu, H., Tan, B., Xu, X. & Ferro, M. (2012). Multiscale peak alignment for chromatographic datasets. *Journal of Chromatography A*, **1223**, 93-106.
- Zhao, Y., Zhang, Y., Lin, R & Sun, W. (2009). An expeditious HPLC method to distinguish *Aconitum kusnezoffii* from related species. *Fitoterapia*, **80**, 333-338.

Chapter 3

Development of HPLC methods for the analysis of *Cyclopia subternata* and *Cyclopia maculata* phenolics

Part of this chapter has been included in an article and published:

De Beer, D., Schulze, A.E., Joubert, E., de Villiers, A., Malherbe, C.J. & Stander, M.A. (2012). Food ingredient extracts of *Cyclopia subternata* (honeybush): Variation in phenolic composition and antioxidant capacity. *Molecules*, **17**, 14602-14624.

3.1 Abstract

Cyclopia spp. are used for the production of honeybush tea and aqueous extracts for use in the food industry. With its high concentration of phenolic compounds, honeybush is associated with many health promoting properties. Limitations of current high performance liquid chromatography (HPLC) methods for phenolic compounds present in *C. subternata* and *C. maculata* are responsible for the relatively small number of compounds that have been quantified. Therefore, two HPLC methods were developed for better resolution of poorly separated peaks and validated for quantification of the major phenolic compounds present in *C. subternata* and *C. maculata* aqueous extracts. The method development for *C. subternata* allowed quantification of xanthenes (mangiferin and isomangiferin), flavanones (hesperidin and eriocitrin), a flavone (scolymoside), a benzophenone (iriflophenone-di-*O,C*-hexoside) and dihydrochalcones (phloretin-3',5'-di-*C*- β -glucoside and 3-hydroxyphloretin-3',5'-di-*C*-hexoside) in extracts of both green and fermented plant material. Additional benzophenones (iriflophenone-3-*C*- β -glucoside and maclurin-3-*C*- β -glucoside) could be quantified in green (unfermented) *C. subternata* extracts. In the case of *C. maculata*, the developed method allowed for the quantification of xanthenes (mangiferin and isomangiferin) and flavanones (hesperidin and eriocitrin) present as major compounds in extracts of fermented plant material, and additionally benzophenones (iriflophenone-3-*C*- β -glucoside and maclurin-3-*C*- β -glucoside) in extracts of green (unfermented) plant material. The following compounds were also tentatively identified for the first time in the designated species, using mass spectrometric detection: maclurin-3-*C*- β -glucoside, 3-hydroxyphloretin-3',5'-di-*C*-hexoside, iriflophenone-di-*O,C*-hexoside and vicenin-2 in both species, together with (*R*)- and (*S*)- eriodictyol-di-*C*-hexoside in *C. subternata*, and hydroxylated derivatives of mangiferin and isomangiferin in *C. maculata*. Among them iriflophenone-di-*O,C*-hexoside, apigenin-6,8-di-*C*-glucoside, hydroxyphloretin-3',5'-di-*C*-hexoside, eriodictyol-di-*C*-glucoside, hydroxymangiferin and hydroxymangiferin were reported for the first time in *Cyclopia*. Large intra- and inter-species variation was observed for the quantified compounds in extracts of green *C. subternata* and *C. maculata*. The variation observed was expected due to the plant material being cultivated from seedlings, and has implications for extract standardisation and quality control.

3.2 Introduction

Honeybush (*Cyclopia* spp.) is endemic to the Cape fynbos biome, and grows in the coastal regions of the Western and Eastern Cape of South Africa (Du Toit *et al.*, 1998). Several species are used for the preparation of herbal tea, including *C. subternata*, *C. maculata*, *C. intermedia* and *C. genistoides* (Joubert *et al.*, 2011). Honeybush tea has a high phenolic content and consequently has been the focus of several studies investigating its potential health promoting benefits, i.e. anti-mutagenic, anti-cancer, phytoestrogenic (Joubert *et al.*, 2008b), anti-diabetic (Mose Larsen *et al.*, 2008; Muller *et al.*, 2011) and anti-obesity properties (Dudhia *et al.*, 2013). To fully understand the health promoting potential of *Cyclopia* spp. and identify bioactive compounds for future quality control purposes, greater insight into their phenolic composition is required.

The identities of several phenolic compounds present in *C. subternata* and *C. intermedia* have been established by means of nuclear magnetic resonance (Ferreira *et al.*, 1998; Kamara *et al.*, 2003, 2004; Kokotkiewicz *et al.*, 2012). High performance liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS) analyses indicated that the major compounds are the xanthenes mangiferin and isomangiferin and the flavone, hesperidin, present in all species investigated to date (Joubert *et al.*, 2003; Joubert *et al.*, 2008a; De Beer & Joubert, 2010).

HPLC is a common tool for the analysis of phenolic compounds in plant extracts. The method that is currently employed by the Agricultural Research Council (ARC) Infruitec-Nietvoorbij for the quantification of phenolic constituents of *C. subternata* is limited in terms of resolution of compounds and problematic integration (De Beer & Joubert, 2010). Unidentified compound(s) co-elute with mangiferin and isomangiferin, leading to overestimation, and a complicated integration process is required due to the elution of eriocitrin, scolymoside and phloretin-3',5'-di-*C*- β -glucoside on a polymeric 'hump' presumably containing unresolved polymeric phenolic compounds. Limited research has been conducted on *C. maculata* to date, and currently there is no quantitative method available for the analysis of its extracts.

The main aim of the present study was to develop and validate more comprehensive HPLC methods for the analysis of phenolic compounds in *C. subternata* and *C. maculata* extracts. Given that the plant material is used either in green or fermented form, the methods were validated for extracts of both types. Mass spectrometry was applied to confirm the identity of the known compounds and tentatively identify additional compounds. The new HPLC methods were applied to a large number of hot water extracts of green *C. subternata* and *C. maculata* prepared from individual seedling plants with the seed originating from the wild to ensure variation.

3.3 Experimental

3.3.1 Chemicals

All chemicals were analytical grade and sourced from Sigma-Aldrich (St. Louis, MO, USA) or Merck Chemicals (Darmstadt, Germany); HPLC gradient grade acetonitrile was also purchased from Merck. Authentic reference standards with purity >95% were obtained from Sigma-Aldrich (hesperidin), Extrasynthese (Genay, France: mangiferin, eriocitrin, luteolin) and Chemos (Regenstauf, Germany: isomangiferin). Aspalathin (3-hydroxyphloretin-3'-*C*-glucoside) (>95%) and nothofagin (phloretin-3'-*C*-glucoside) (>95%) were obtained from PROMEC (Medical Research Council of South Africa, Tygerberg, South Africa). Deionised water was prepared using an Elix water purification system (Millipore, Milford, MA, USA) and further purified to HPLC grade using a Milli-Q Academic water purification system (Millipore).

3.3.2 Plant material

Cyclopia subternata seedling plants ($n = 64$) of the same age were harvested from a commercial plantation on Kanetberg Flora in the Barrydale district, South Africa. The harvested shoots (stems and leaves) were dried at 40 °C in a temperature-controlled drying tunnel with forced air circulation to a *ca.* 8-10% moisture content and ground with a Retsch mill (1 mm seive; Retch GmbH, Haan, Germany). These samples represented green *C. subternata*.

Cyclopia maculata seedlings ($n = 40$; 2-year-old) were grown in pots under standardised conditions at ARC Infruitec-Nietvoorbij, Stellenbosch, South Africa. The seeds for the pot trial were collected from a natural population, growing in marsh land near Riversdale, South Africa. The shoots were dried at 30 °C for *ca.* 2.5 days, followed by 6 h at 40 °C and powdered as described for *C. subternata*. These samples represented green *C. maculata*.

For method validation, paired green and fermented plant material was prepared for both *C. subternata* and *C. maculata*. The shoots of single bushes of each species were harvested at Kanetberg Flora and Riversdale, respectively. The shoots of each bush were mechanically shredded into < 3 mm lengths and one half dried immediately in a drying tunnel as described above, representing the green plant material. The other half was "fermented" as follows: the cut plant material was moistened to *ca.* 60-65% moisture content and fermented (oxidised) at 90 °C for 16 h, followed by drying. All the material was sieved to remove mainly coarse pieces of stem and the fraction < 1.4 mm ("tea bag cut") used to prepare hot water extracts (section 3.3.3).

3.3.3 Preparation of aqueous honeybush tea extracts for analyses

An extract of each sample was prepared by adding boiling water (70 mL) to the milled plant material (7 g) in a screw-cap glass bottle, which was placed in a water bath at 93 °C for 30 min. The mixture was stirred by swirling every 5 min. The resulting extract was filtered through Whatman #4 filter paper while warm, followed by cooling to room temperature in a water bath. Extracts of sieved material (10 g) was prepared using the same procedure and 1:10 mass:water ratio. All extracts were frozen and freeze-dried using a VirTis Advantage Plus freeze-drier (SP Scientific, Warminster, PA, USA).

3.3.4 HPLC-DAD method development

Analyses were conducted on an Agilent 1200 series HPLC instrument which consisted of an in-line degasser, quaternary pump, autosampler, column oven and diode array detector (DAD), controlled by Chemstation software (Agilent Technologies Inc., Santa Clara, CA). The solvent gradient described by De Beer & Joubert (2010) was evaluated on four different columns in order to improve separation of co-eluting compounds in *C. subternata* and *C. maculata* extracts. The following columns were evaluated: Zorbax Eclipse XDB-C18 (150 × 4.6 mm; 5 μm; 80 Å) and Zorbax SB-C18 (100 × 4.6 mm; 1.8 μm; 80 Å), both from Agilent Technologies (Waldbronn, Germany), as well as Kinetex C18 (150 × 4.6 mm; 2.6 μm; 100 Å) with core-shell technology and Gemini-NX C18 (150 × 4.6 mm; 3 μm; 110 Å), both from Phenomenex (Santa Clara, CA, USA). Mobile phases evaluated were 1% and 2% acetic acid and 0.1% formic acid as acidic aqueous phases (A), and acetonitrile and methanol as organic modifiers (B). The effect of temperature and gradient lengths were evaluated from 25 to 50 °C, and from 25 to 35 min, respectively.

3.3.5 Quantification of phenolic compounds using HPLC-DAD

Dimethylsulfoxide (DMSO) was used to prepare stock solutions of standards and aliquots were frozen at -20°C until analysis. Extracts were dissolved in purified water (*ca.* 6 mg/mL) and frozen at -20°C until analysis. Ascorbic acid was added to standard mixtures (*ca.* 5 mg/mL final concentration) and defrosted reconstituted extracts (*ca.* 9 mg/mL final concentration) to prevent oxidative degradation, where-after the mixtures were filtered using 0.22 μm pore-size Millex-HV syringe filters (Millipore). The injection volumes were 15 μL for *C. subternata* extracts, 10 μL for *C. maculata* and 10-20 μL for the standards. A Gemini-NX C18 (150 \times 4.6 mm; 3 μm ; 100 \AA) column was used, with 2% acetic acid (A) and acetonitrile (B) as mobile phases. Separation was carried out at 30°C and a flow rate of 1 mL/min. The mobile phase gradient for *C. subternata* was 0-2 min (8% B), 2-27 min (8%-38% B), 27-28 min (38%-50% B), 28-29 min (50% B), 29-30 (50%-8% B), 30-40 min (8% B). The mobile phase gradient for *C. maculata* was 0-2 min (8% B), 2-31 min (8%-38% B), 31-32 min (38%-50% B), 32-33 min (50% B), 33-34 (50%-8% B), 34-44 min (8% B) (Fig. 3.1).

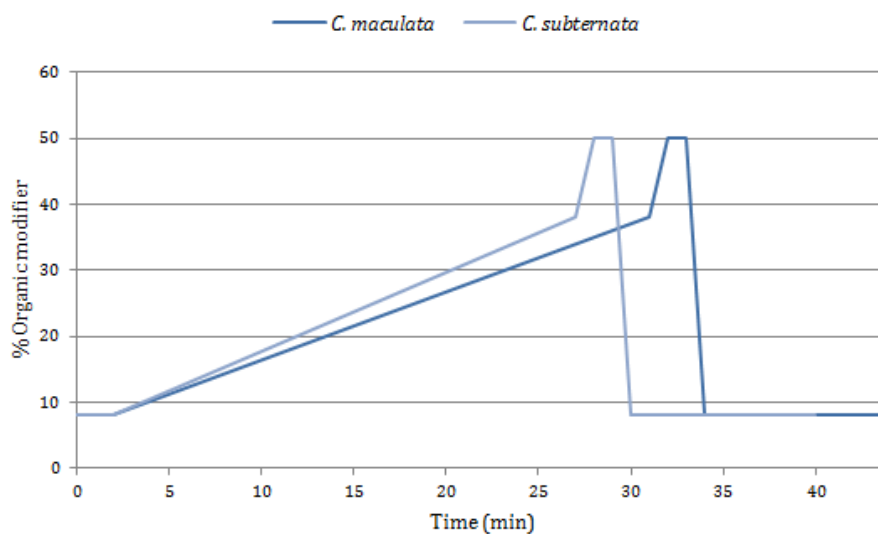


Figure 3.1 – Gradient profiles used for *C. subternata* and *C. maculata* analyses [Mobile phases: 2% acetic acid (A) and acetonitrile (B)].

UV spectra were recorded for all samples from 200 to 550 nm. The xanthenes (mangiferin and isomangiferin) and flavone (scolymoside) were quantified at 320 nm, and the flavanones (eriocitrin and hesperidin), benzophenones (iriflophenone-3-*C*-glucoside, iriflophenone-di-*O,C*-hexoside and maclurin-3-*C*- β -glucoside) and dihydrochalcones (phloretin-3',5'-di-*C*- β -glucoside and 3-hydroxyphloretin-3',5'-di-*C*- β -hexoside) at 288 nm. Separate seven-point calibration curves were set up for all the available authentic reference standards, as well as standards needed to calculate equivalent values in cases where authentic standards were not available. The isomangiferin content of *C. subternata* extracts was quantified using a response factor previously determined for isomangiferin relative to mangiferin, while the authentic standard for isomangiferin was included in the calibration curve used for *C. maculata*. Iriflophenone-3-*C*-glucoside was quantified using a response factor previously determined relative to hesperidin. Iriflophenone-di-*O,C*-hexoside and maclurin-3-*C*- β -glucoside were both quantified as iriflophenone-3-*C*-glucoside equivalents. Scolymoside was quantified as luteolin equivalents. Phloretin-3',5'-di-*C*- β -glucoside and 3-hydroxyphloretin-3',5'-di-*C*- β -hexoside were quantified as phloretin-3'-*C*- β -glucoside (nothofagin) and 3-hydroxyphloretin-3'-*C*- β -glucoside (aspalathin) equivalents, respectively. Nothofagin calibration was performed using a response factor previously determined relative to aspalathin. The standard

calibration mixtures accommodated the expected concentration range (μg injected) of the compounds present in the sample extracts. For *C. subternata*: mangiferin (0.032-2.889 μg), hesperidin (0.009-2.263 μg), luteolin (0.005-1.360 μg), eriocitrin (0.007-1.689 μg) and aspalathin (0.007-1.660 μg). For *C. maculata*: mangiferin (0.020-4.873 μg), isomangiferin (0.005-1.343 μg), eriocitrin (0.007-1.765 μg) and hesperidin (0.009-2.263 μg).

3.3.6 Identification of phenolic compounds using LC-DAD-MS and -MS/MS detection

Liquid chromatography-mass spectrometry (LC-MS) and tandem mass spectrometry (LC-MS/MS) analyses were conducted on a Waters Acquity UPLC equipped with a binary pump, in-line degasser, autosampler, column oven and DAD detector (Waters, Milford, MA, USA). The system was coupled to a Synapt G2 Q-TOF mass spectrometer (Waters) equipped with an electrospray ionisation (ESI) source. The same methods were used as described in section 3.3.5, with the exception that the gradients were delayed by 0.2 min to compensate for differences in dead volumes between instruments. Furthermore, for negative electrospray ionisation (ESI), the percentage of solvent B was also increased by 0.5% at each interval of the gradient, and the column temperature increased to 32 °C. These changes were required to obtain similar separation as on the Agilent HPLC system. An injection volume of 3 μL was used for the standards and samples. UV-Vis spectra were recorded from 235 to 450 nm. The eluent was split 3:2 prior to introduction into the ionisation chamber. MS data were acquired in both positive and negative mode: in the first analysis, MS and MS^E data were acquired (the latter using a collision energy ramp from 25 to 60 V) for both *C. subternata* and *C. maculata*; subsequently, a second analysis was performed on *C. subternata* to acquire MS/MS data using a collision energy of 30 V. The MS parameters were as follows: desolvation temperature, 275 °C; nitrogen flow rate, 650 L/h; source temperature, 120 °C; capillary voltage, 2500 V; cone voltage, 15 V. Data were processed using MassLynx v.4.1 software (Waters). Peak identities were determined by comparison of retention times, UV-Vis spectra and LC-MS spectra with those of authentic standards. LC-MS and -MS/MS spectra were compared to literature to identify or tentatively identify peaks for which no authentic standards were available.

3.3.7 Method validation

One standard calibration mixture, injected at two injection volumes, and the paired green and fermented extracts of *C. subternata* and *C. maculata* were selected for method validation. Specificity of the method was determined by evaluating the UV-Vis and MS spectra of peaks selected for quantification. Linearity of the response was determined by calculating the slope, y-intercept and correlation coefficient (r) of each calibration curve. Stability of the standard calibration mixture and samples were determined by injecting them repeatedly over a 26-h-period for *C. subternata* and over a 24-h-period for *C. maculata*. The %RSD and % change over the period were determined for each compound. Intra-day precision was determined by consecutive repeated injections ($n = 6$) of the standard mixture and samples on the same day. The %RSD was calculated for each of the compounds. The same procedure was repeated over three days, and compared by calculating the %RSD for the averaged values for each day to determine the inter-day precision.

3.4 Results and discussion

3.4.1 HPLC-DAD method development

The HPLC method previously described by De Beer & Joubert (2010) to quantify the major phenolic compounds in *C. subternata*, and applied for analysis of the major phenolic compounds of other *Cyclopi*a species, suffered from a few limitations. These included co-elution of unidentified compound(s) with mangiferin and isomangiferin, and a complicated integration process due to the elution of eriocitrin, scolymoside and phloretin-3',5'-di-*C*- β -glucoside on a 'polymeric hump'. Using the same method for *C. maculata* posed limitations of the same nature. Using the existing method as basis, new species-specific HPLC methods were thus developed and evaluated for the analysis of *C. subternata* and *C. maculata* aqueous extracts. The focus fell on aqueous extracts as this type of extract is mainly produced by industry (Joubert *et al.*, 2011) and is the main type of extract investigated for health promoting properties of *Cyclopi*a (Joubert *et al.*, 2008a).

Four columns were evaluated with 1% acetic acid, 2% acetic acid or 0.1% formic acid as acidic aqueous phase and acetonitrile as organic modifier, using the gradient as described by De Beer & Joubert (2010). The performances of the columns were evaluated with respect to the problematic separation areas in the chromatograms of paired green and fermented *C. subternata* and *C. maculata* extracts ($n = 1$ of each type). The Gemini-NX and Kinetex columns provided the best potential separation of mangiferin and isomangiferin from the unidentified co-eluting compounds, and were therefore selected for further method development. Methanol was subsequently also evaluated as organic modifier after adjusting the gradient to accommodate the differences in solvent strength between methanol and acetonitrile. Methanol did not provide better separation than acetonitrile on either column, and resulted in maximum pressures above the instrumental limit ($P_{\max} = 400$ bar) for the Kinetex column. Thus the Gemini-NX column and acetonitrile were selected for further optimisation.

Of the three acidified aqueous phases evaluated, 2% acetic acid provided the best separation, especially for mangiferin and isomangiferin, from the unidentified compounds. Subsequent evaluation continued with 2% acetic acid and acetonitrile as mobile phases. A scouting gradient changing from 10% to 40% acetonitrile was employed to remove the smaller steps in the previously employed gradient (De Beer & Joubert, 2010). This approach spread out the polymeric hump previously observed, while maintaining good separation of the major phenolic compounds and simplifying integration. In order to improve separation of mangiferin and isomangiferin from unidentified compounds, the initial acetonitrile content was decreased to 8% and the gradient ended at 38%. Furthermore, gradient times from 25 to 35 min, changed in one min intervals, were evaluated. Using gradient times longer than 25 min did not improve separation in the chromatogram of *C. subternata* and a 25 min gradient was thus selected as optimal for the separation of the major phenolic compounds in *C. subternata*. For *C. maculata* a longer gradient time (less steep gradient) improved separation of the major phenolic compounds, with 29 min being the optimal gradient length. Temperature evaluation was performed in the range of 25 to 50 °C and 30 °C provided the lowest degree of co-elution for hesperidin, eriocitrin and iriflophenone-3-*C*- β -glucoside in both species. Similarly, mangiferin and isomangiferin also displayed the best resolution at this temperature. The chromatograms of the aqueous extracts of green and fermented *C. subternata* and *C. maculata*, using the optimised conditions, are illustrated in Figs. 3.2 and 3.3, respectively.

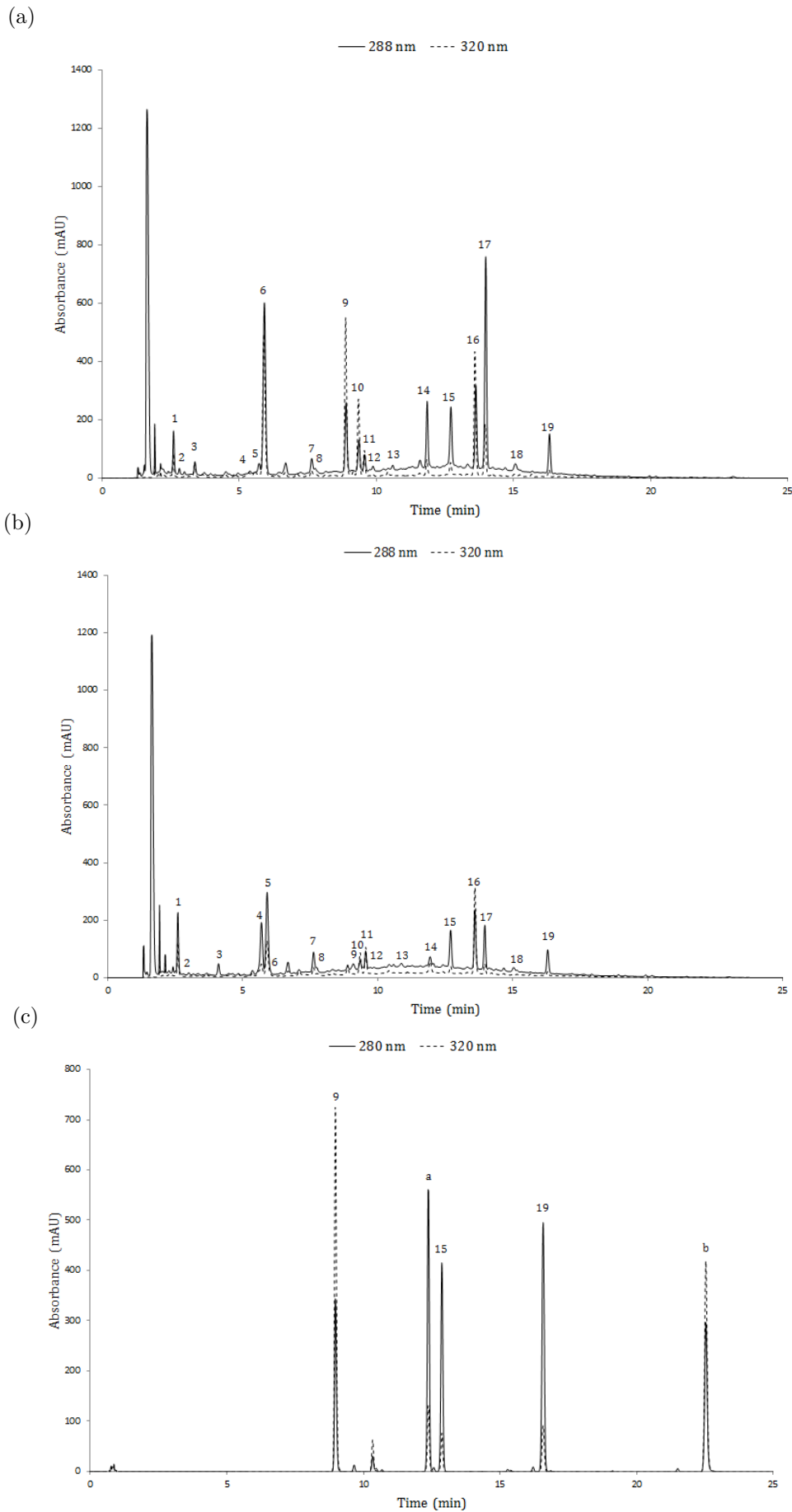


Figure 3.2 – Chromatograms for the separation of freeze-dried aqueous extracts of (a) green and (b) fermented *CyclopiA subternata* phenolic compounds and (c) a standard calibration mixture [See Table 3.1 and Table 3.2 for identities of numbered peaks; a aspalathin, b luteolin].

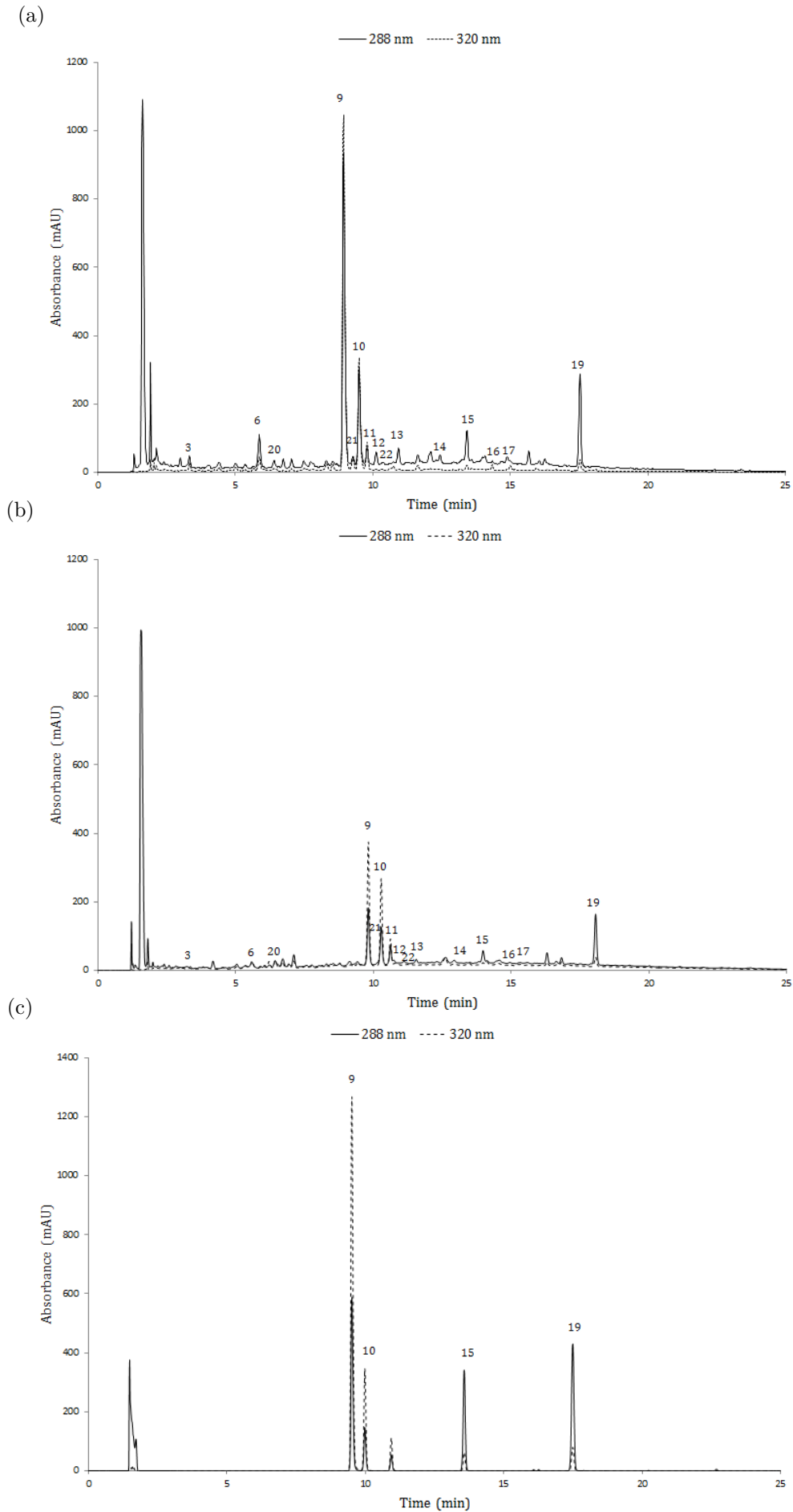


Figure 3.3 – Chromatograms for the separation of freeze-dried aqueous extracts of (a) green and (b) fermented *CyclopiA maculata* phenolic compounds and (c) a standard calibration mixture [See Table 3.3 for identities of numbered peaks].

3.4.2 Identification of phenolic compounds in two *Cyclopia* species

LC-MS and LC-MS/MS analyses of the paired green and fermented *C. subternata* and *C. maculata* extracts were performed in both positive and negative ionisation modes. By comparing retention times, UV-Vis spectral properties, LC-MS spectra and LC-MS/MS fragmentation patterns with those of authentic commercial standards (Fig. 3.2, Tables 3.1 and 3.2), the presence of mangiferin (**9**), eriocitrin (**15**), and hesperidin (**19**) were confirmed in the *C. subternata* extracts. The presence of isomangiferin (**10**) was confirmed by comparison with isolated isomangiferin (De Beer *et al.*, 2009). Similarly, the presence of mangiferin (**9**), isomangiferin (**10**), eriocitrin (**15**), and hesperidin (**19**) in *C. maculata* were confirmed with authentic standards as indicated in Fig. 3.3 and Table 3.3. Additional compounds were tentatively identified by comparison of UV-Vis, LC-MS and LC-MS/MS data with literature. These compounds are discussed in more detail for the different phenolic classes below. Structures for known and tentatively identified compounds are shown in Fig. 3.4.

General fragmentation of *Cyclopia* phenolics

Cyclopia subternata and *C. maculata* flavonoids include both flavonoid *O*- and *C*-glycosides, which can be distinguished with relative ease, based on their divergent mass spectral properties. Fragmentation of *O*-glycosides at low collision energies typically involves cleavage of the *O*-glycosidic bond, with corresponding losses of 162, 146 and 132 amu for hexoses, deoxyhexoses and pentoses, respectively (Abad-García *et al.*, 2009). *C*-glycosides, on the other hand, do not contain labile bonds, therefore higher collision energies are required for fragmentation. Under these conditions most of the fragmentation involves the sugar moiety. Losses of 90, 96, 120 and 150 amu are typical for hexoses, and 60, 90 and 120 amu for pentoses. These are often accompanied by additional loss of water molecules (-18 amu). In the case of flavonoid-di-*C*-glycosides, simultaneous fragmentation of both sugars is common, leading to a relatively large number of fragments typically being detected (for example $[M-H-96-120-H_2O]^-$ for a flavonoid-di-*C*-hexoside).

Cyclopia subternata

In the following discussion of *C. subternata* phenolics, all references will be to Figs. 3.2 and 3.4, and Tables 3.1 and 3.2.

Benzophenones

Compound **6** with m/z 407 ($[M-H]^-$) was detected at a retention time of 6.69 min and has the proposed molecular formula $C_{19}H_{19}O_{10}$ ($[M-H]^-$). The compound presented a major fragment ion at m/z 287 corresponding to $[M-H-120]^-$, which indicates a *C*-glycoside structure. Additionally, other fragments including m/z 317 and 257, corresponding to losses of $[M-H-90]^-$ and $[M-H-150]^-$, respectively, are also characteristic of *C*-glycosides. Based on the UV-Vis spectral data, MS and MS/MS fragmentation patterns, this compound was identified as iriflophenone-3-*C*- β -glucoside, which has previously been isolated from *C. subternata* and described by Kokotkiewicz *et al.* (2012).

Peak **1** with m/z 569 ($[M-H]^-$) was detected at a retention time of 3.10 min and has the proposed molecular formula of $C_{25}H_{29}O_{15}$ ($[M-H]^-$). The compound presented several fragment ions: m/z 479 $[M-H-90]^-$, m/z 449 $[M-H-120]^-$, m/z 317 $[M-H-90-162]^-$, m/z 287 $[M-H-120-162]^-$. This fragmentation pattern possibly indicates a phenolic compound containing 2 hexosyl groups, of which one is a *C*-hexosyl and the other an *O*-hexosyl. This compound also showed similar characteristics to iriflophenone-*C*- β -glucoside (m/z 407, $[M-H]^-$), according to

Table 3.1 – UV-Vis, LC-MS and LC-MS/MS characteristics of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented *C. subternata*

Peak ^a	Mode	t _R (min)	Accurate mass ^b	λ _{max} (nm)	Error (ppm)	Proposed molecular formula	Fragments	Phenolic compound
1	+	3.10	571.1664	290	0.2	C ₂₅ H ₃₁ O ₁₅	373, 355, 337, 325, 313, 289, 271, 259, 231, 219 195* , 177, 165	Iriflophenone-di- <i>O,C</i> -hexoside
	-		569.1488		-2.3	C ₂₅ H ₂₉ O ₁₅	479, 449, 317, 287*	
3	+	4.20	425.1084	234, 315	0.2	C ₁₉ H ₂₁ O ₁₁	261, 243, 231, 219, 195* , 177, 165, 137, 121	Maclurin-3- <i>C-β</i> -glucoside
	-		423.0929		0.5	C ₁₉ H ₁₉ O ₁₁	333, 303, 223, 193* , 165, 151, 109	
4	+	6.32	613.1780	285	1.8	C ₂₇ H ₃₃ O ₁₆	475, 409, 339, 327, 303, 285, 261* , 219	(<i>S</i>)-Eriodictyol-di- <i>C</i> -hexoside
	-		611.1621		1.0	C ₂₇ H ₃₁ O ₁₆	491, 431, 401, 371*	
5	+	6.50	613.1780	285	1.8	C ₂₇ H ₃₃ O ₁₆	475, 409, 339, 327, 303, 285, 261* , 219	(<i>R</i>)-Eriodictyol-di- <i>C</i> -hexoside
	-		611.1621		1.0	C ₂₇ H ₃₁ O ₁₆	491, 431, 401, 371*	
6	+	6.69	409.1136	294	0.2	C ₁₉ H ₂₁ O ₁₀	391, 289, 231, 195* , 177, 165, 121	Iriflophenone-3- <i>C-β</i> -glucoside
	-		407.0981		1.0	C ₁₉ H ₁₉ O ₁₀	317, 287* , 257, 245, 215, 201, 193, 165, 125	
9	+	8.95	423.0920	234, 257,	-1.7	C ₁₉ H ₁₉ O ₁₁	351, 339, 327, 303, 299, 285, 273* , 257	Mangiferin
	-		421.0770	317, 366	-0.2	C ₁₉ H ₁₇ O ₁₁	331, 301* , 271, 259	
10	+	9.30	423.0922	234, 255,	-1.2	C ₁₉ H ₁₉ O ₁₁	405, 357, 341, 327, 303* , 299, 287, 285, 273, 261	Isomangiferin
	-		421.0769	316, 366	-0.5	C ₁₉ H ₁₇ O ₁₁	331, 301* , 273, 271, 259	
11	+	9.40	595.1671	235, 270, 331	1.3	C ₂₇ H ₃₁ O ₁₅	505, 457, 427, 421, 409, 391, 379, 355, 337, 325* , 307, 295	Apigenin-6,8-di- <i>C</i> -glucoside [Vicenin-2]
	-		593.1499		-1.2	C ₂₇ H ₂₉ O ₁₅	503, 473* , 383, 353	
12	+	10.04	451.1229	281	-2.4	C ₂₁ H ₂₃ O ₁₁	289, 163, 153*	Eriodictyol- <i>O</i> -glucoside
	-		449.1069		-3.3	C ₂₁ H ₂₁ O ₁₁	287, 151* , 135	
13	+	10.63	451.1232	281	-1.8	C ₂₁ H ₂₃ O ₁₁	289* , 163, 153	Eriodictyol- <i>O</i> -glucoside
	-		449.1080		-0.9	C ₂₁ H ₂₁ O ₁₁	287, 151* , 135, 107	
14	+	11.54	615.1927	283	0.3	C ₂₇ H ₃₅ O ₁₆	525, 495, 477, 465, 447, 435, 423, 411, 399, 381, 369, 345, 327, 259, 247, 235, 217, 205, 165* , 123	3-Hydroxyphloretin-3',5'-di- <i>C</i> - hexoside
	-		613.1772		0.5	C ₂₇ H ₃₃ O ₁₆	493, 475, 433, 403, 373* , 361, 331, 239, 209	
15	+	12.50	597.1812	283	-1.2	C ₂₇ H ₃₃ O ₁₅	289* , 273, 153	Eriodictyol-7- <i>O</i> -rutinoside
	-		595.1661		-0.3	C ₂₇ H ₃₁ O ₁₅	287* , 151, 135	[Eriocitrin]
16	+	13.29	595.1647	252, 348	-2.7	C ₂₇ H ₃₁ O ₁₅	449, 287*	Luteolin-7- <i>O</i> -rutinoside
	-		593.1509		0.5	C ₂₇ H ₂₉ O ₁₅	285*	[Scolymoside]
17	+	13.53	599.1972	284	-0.7	C ₂₇ H ₃₅ O ₁₅	479, 461, 449, 431, 419, 407, 395* , 383, 365, 353, 329, 301, 107	Phloretin-3',5'-di- <i>C-β</i> -glucoside
	-		597.1831		2.0	C ₂₇ H ₃₃ O ₁₅	477, 459, 417, 387, 357* , 345, 315	
19	+	15.85	611.1970	283	-1.0	C ₂₈ H ₃₅ O ₁₅	449, 303* , 177, 153	Hesperetin-7- <i>O</i> -rutinoside
	-		609.1837		3.0	C ₂₈ H ₃₃ O ₁₅	301*	[Hesperidin]

^aPeak numbers correspond to numbered peaks in Figure 3.2; ^bAccurate mass determined experimentally; bold*: Ion with highest relative intensity

Table 3.2 – UV-Vis, LC-MS and LC-MS/MS characteristics of unidentified compounds in freeze-dried aqueous extracts of green and fermented *C. subternata*

Peak ^a	Mode	t _R (min)	Accurate mass ^b	λ _{max} (nm)	Error (ppm)	Proposed molecular formula	Fragments	Phenolic compound
2	+	3.20	345.1186	234, 272	-0.3	C ₁₅ H ₂₁ O ₉	123* , 165	Unidentified
	-		343.1024		-1.5	C ₁₅ H ₁₉ O ₉	163, 119*	
7	+	7.83	nd	282	nd	nd	nd	Unidentified
	-		457.1353		1.5	C ₂₀ H ₂₅ O ₁₂	163* , 199	
8	+	7.93	597.1812	280	-1.2	C ₂₇ H ₃₃ O ₁₅	405, 393, 363, 339, 327, 321, 285, 273, 261* , 219, 207	Naringenin-di- <i>C</i> -hexoside (?)
	-		595.1665		0.3	C ₂₇ H ₃₁ O ₁₅	475, 415, 385* , 355	
18	+	14.68	581.1837	279	-5.7	C ₂₇ H ₃₃ O ₁₄	273* , 153	Naringenin- <i>O</i> -deoxyhexoside-
	-		579.1719		0.9	C ₂₇ H ₃₁ O ₁₄	433, 417, 271* , 151	<i>O</i> -hexoside (?)

^aPeak numbers correspond to numbered peaks in Figure 3.2; ^bAccurate mass determined experimentally; bold*: Ion with highest relative intensity; nd, not detected

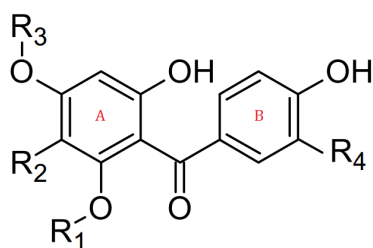
its UV and MS/MS spectra. Due to its mass difference of 162 amu, compound **1** was proposed to be a glycosylated derivative of iriflophenone-*C*- β -glucoside. Its earlier elution time is also in accordance with this hypothesis. Therefore compound **1** was tentatively identified as an iriflophenone-di-*O,C*-hexoside.

Compound **3** with m/z 423 ($[M-H]^-$) was detected at a retention time of 3.38 min. Accurate mass data provided the proposed molecular formula $C_{19}H_{19}O_{11}$ $[M-H]^-$. This information, together with the molecular ion which is 16 amu higher than that of compound **6**, indicates a possible hydroxylated derivative of iriflophenone-*C*-glucoside. This is in accordance with the relative retention times of compounds **3** and **6**, with the former eluting earlier due to its higher polarity (Abad-García *et al.*, 2009). Furthermore, peak **3** presented major fragment ions at m/z 333 and 303 corresponding to $[M-H-90]^-$ and $[M-H-120]^-$, respectively, which indicates a *C*-glycoside structure. These fragment ions were also 16 amu higher than those observed for **6**, whereas the A-ring fragments were identical for both these compounds, confirming that the additional hydroxyl moiety is attached to the B-ring (Fig. 3.4) (Zhang *et al.*, 2013). These data are in accordance with data representing maclurin-3-*C*- β -glucoside, a compound present in the leaves of *Mangifera indica* L. (Tanaka *et al.*, 1984; Zhang *et al.*, 2013) and *C. genistoides* (Kokotkiewicz *et al.*, 2013).

Dihydrochalcones

Compound **17** with m/z 597 ($[M-H]^-$) eluted at 13.53 min and has a proposed molecular formula $C_{27}H_{33}O_{15}$ ($[M-H]^-$). This molecule displayed MS/MS fragments at m/z 477 ($[M-H-120]^-$), 459 ($[M-H-120-H_2O]^-$), 417 ($[M-H-2\times 90]^-$), 387 ($[M-H-120-90]^-$) and 357 ($[M-H-2\times 120]^-$), which correspond with the dihydrochalcone phloretin-3',5'-di-*C*- β -glucoside previously identified in *C. subternata* (Kokotkiewicz *et al.*, 2012). The UV spectrum of this compound is also in agreement with earlier work (Kokotkiewicz *et al.*, 2012).

Compound **14**, which eluted at 11.54 min with a m/z of 613 ($[M-H]^-$) and 615 ($[M-H]^+$), had a UV spectrum and MS/MS fragmentation pattern similar to that of phloretin-3',5'-di-*C*- β -glucoside (**17**, m/z 597 $[M-H]^-$), but displayed a molecular ion 16 amu higher than that of **17**. This indicates a possible hydroxylated derivative of phloretin-3',5'-di-*C*- β -glucoside. Compound **14** has the proposed molecular formula $C_{27}H_{35}O_{16}$ ($[M-H]^+$), which was in good agreement with its accurate mass, 615,1927 ($[M+H]^+$). In fact, these data are in accordance with UV, MS and MS/MS data of a compound previously reported in rooibos tea (*Aspalathus linearis*) (Beelders *et al.*, 2012c). The following fragments were observed for compound **14** in positive ionization mode: m/z 525 ($[M+H-90]^+$), m/z 495 ($[M+H-120]^+$), m/z 477 ($[M+H-120-H_2O]^+$), m/z 465 ($[M+H-150]^+$), m/z 447 ($[M+H-150-H_2O]^+$), m/z 435 ($[M+H-2\times 90]^+$), m/z 423 ($[M+H-2\times 96]^+$), m/z 411 ($[M+H-90-96-H_2O]^+$), m/z 399 ($[M+H-96-120]^+$), m/z 381 ($[M+H-96-120-H_2O]^+$), m/z 369 ($[M+H-96-150]^+$), m/z 345 ($[M+H-120-150]^+$) and m/z 327 ($[M+H-120-150-H_2O]^+$) (Fig. 3.5). This fragmentation pattern is identical to that of the novel C-5'-hexosyl derivative of aspalathin (2',3,4,4',6',-pentahydroxy-3',5'-di-*C*-hexosyldihydrochalcone) reported by Beelders *et al.* (2012b; 2012c). This compound is therefore the 3-hydroxy-derivative of phloretin-3',5'-di-*C*- β -glucoside (**17**). This proposition is entirely consistent with the mass difference of 16 amu observed between **14** and **17**, while the fragments detected at m/z 123 and 165 in the MS/MS spectrum of the former confirm that the additional hydroxyl group is attached to the B-ring. Compound **14** will be referred to as 3-hydroxyphloretin-3',5'-di-*C*-hexoside. It is interesting to note that the mono-glucoside, aspalathin (2',3,4,4',6'-pentahydroxy-3'-*C*- β -D-glucopyranosyldihydrochalcone), considered to be a unique constituent of rooibos tea (*Aspalathus linearis*), was not observed in any of the *C. subternata* or *C. maculata* extracts analysed. Both *A. linearis* and *Cyclopia* spp. belong to the legume family, Fabaceae. To confirm that compound **14** is the same molecule previously



Benzophenones

Iriflophenone-di-*O,C*-hexoside (**1**)*:

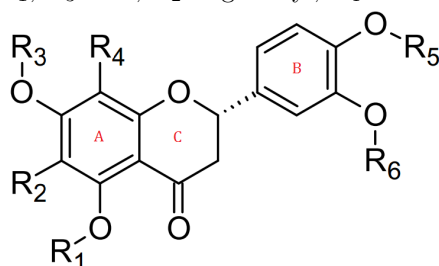
R₁ or R₃ = hexosyl or H; R₂ = hexosyl; R₄ = H

Iriflophenone-3-*C*-β-glucoside (**6**):

R₁, R₃, R₄ = H; R₂ = glucosyl

Maclurin-3-*C*-β-glucoside (**3**):

R₁, R₃ = H; R₂ = glucosyl; R₄ = OH



Flavanones

(*S*)- and (*R*)-Eriodictyol-di-*C*-hexoside (**4,5**)*:

R₁, R₃, R₅, R₆ = H; R₂ or R₄ = hexosyl or H

Eriodictyol-*O*-glucoside (**12,13**)*:

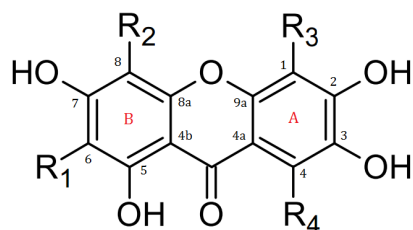
R₁ or R₃ = glucosyl or H; R₂, R₄, R₅, R₆ = H

Eriocitrin (**15**):

R₃ = rutinosyl; R₁, R₂, R₄, R₅, R₆ = H

Hesperidin (**19**):

R₃ = rutinosyl; R₅ = CH₃; R₁, R₂, R₄, R₆ = H



Xanthones

Mangiferin (**9**):

R₁ = glucosyl; R₂, R₃, R₄ = H

Isomangiferin (**10**):

R₂ = glucosyl; R₁, R₃, R₄ = H

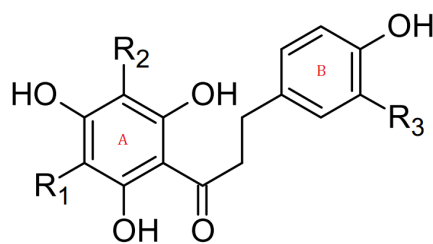
Hydroxymangiferin (**21**):

R₁ = glucosyl; R₂ = H; R₃ or R₄ = OH or H

Hydroxyisomangiferin (**22**):

R₂ = glucosyl; R₁ = H; R₃ or R₄ = OH or H

Figure 3.4 – Structures of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented *C. subternata* and/or *C. maculata* [* indicates that the position of hexosyl or hydroxyl moieties are not certain].



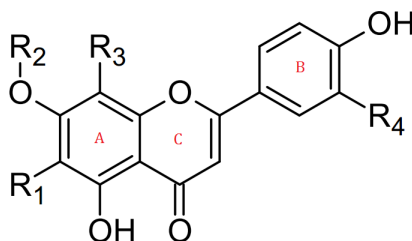
Dihydrochalcones

3-Hydroxyphloretin-3',5'-di-*C*-hexoside (**14**):

$R_1, R_2 = \text{hexosyl}; R_3 = \text{OH}$

Phloretin-3',5'-di-*C*-glucoside (**17**):

$R_1, R_2 = \text{glucosyl}; R_3 = \text{H}$



Flavones

Vicenin-2 (Apigenin-6,8-di-*C*-glucoside) (**11**):

$R_1, R_3 = \text{glucosyl}; R_2, R_4 = \text{H}$

Scolymoside (**16**):

$R_1, R_3 = \text{H}; R_2 = \text{rutosyl}; R_4 = \text{OH}$

Figure 3.4 cont – Structures of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented *C. subternata* and/or *C. maculata* [* indicates that the position of hexosyl or hydroxyl moieties are not certain].

observed in rooibos, a rooibos extract was analysed using the current *C. subternata* HPLC method. Identical retention times and MS data confirmed this was indeed the same compound (results not shown).

Flavone and flavanone *C*-glucosides

Compounds **4** and **5** were characterised by molecular ions at m/z 611 ($[\text{M}-\text{H}]^-$, $t_R = 6.32$ and 6.50 min, respectively) and an identical proposed molecular formula of $\text{C}_{27}\text{H}_{31}\text{O}_{16}$ ($[\text{M}-\text{H}]^-$). Both compounds showed numerous fragments under MS/MS conditions, including m/z 491 ($[\text{M}-\text{H}-120]^-$), 431 ($[\text{M}-\text{H}-2 \times 90]^-$), 401 ($[\text{M}-\text{H}-90-120]^-$) and 371 ($[\text{M}-\text{H}-2 \times 120]^-$). These fragments indicate a di-*C*-hexoside, which together with their UV-spectra, molecular weights and formulae point to eriodictyol-di-*C*-hexoside. The known reversed phase elution order of eriodictyol derivatives (Beelders *et al.*, 2012a; 2012c) allows tentative identification of compounds **4** and **5** as the flavanone derivatives (*S*)-eriodictyol-di-*C*-hexoside and (*R*)-eriodictyol-di-*C*-hexoside, respectively.

Compound **11** with a retention time of 9.40 min and m/z of 593 $[\text{M}-\text{H}]^-$ and 595 $[\text{M}+\text{H}]^+$ has tentatively been identified as apigenin-6,8-di-*C*-glucoside (vicenin-2) with the proposed molecular formula $\text{C}_{27}\text{H}_{29}\text{O}_{15}$ $[\text{M}-\text{H}]^-$. This compound has a UV spectrum corresponding to that of a flavone derivative, with maximum absorbance at 272 and 331 nm. MS/MS fragmentation showed the following fragments: m/z 505 ($[\text{M}-\text{H}-90]^+$), m/z 457 ($[\text{M}-\text{H}-120-\text{H}_2\text{O}]^+$), m/z 427 ($[\text{M}-\text{H}-150-\text{H}_2\text{O}]^+$), m/z 409 ($[\text{M}-\text{H}-90-96]^+$), m/z 379 ($[\text{M}-\text{H}-120-96]^+$),

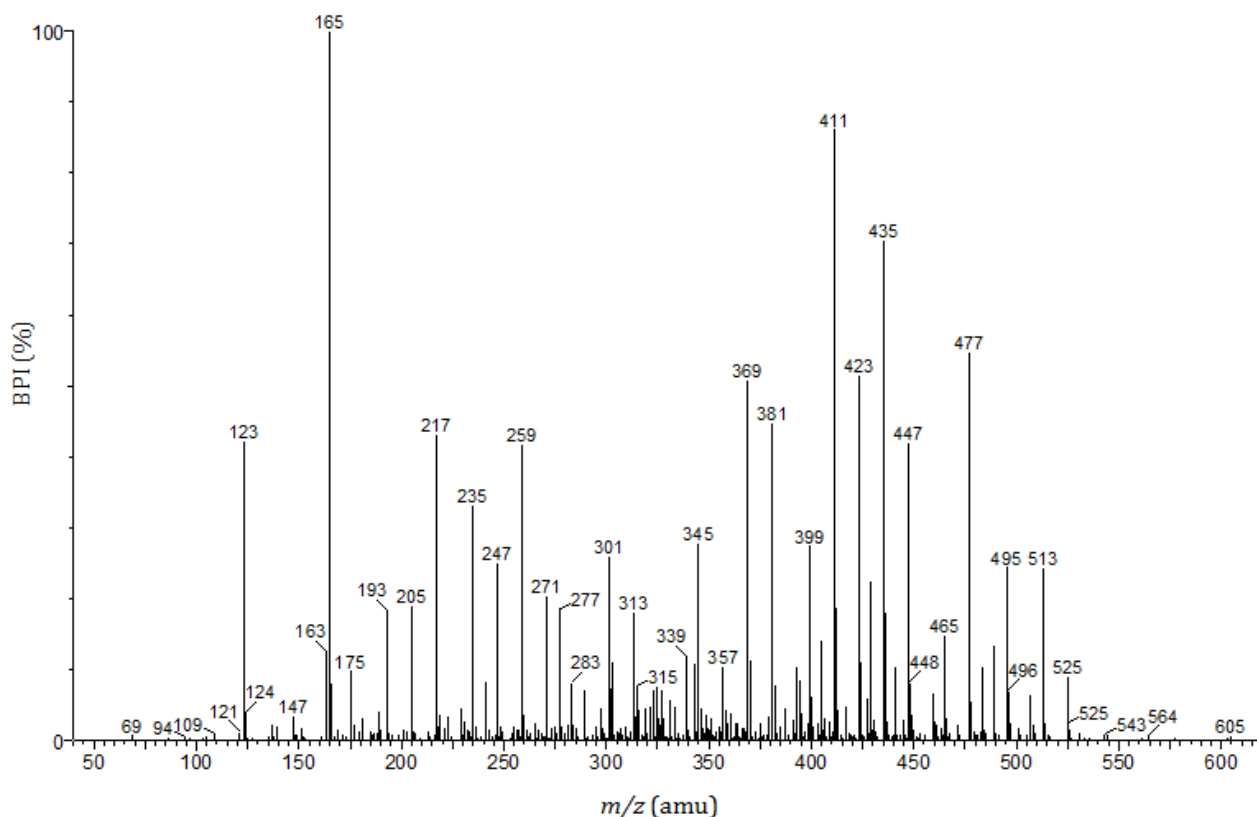


Figure 3.5 – LC-MS/MS spectrum of 3-hydroxyphloretin-3',5'-di-*C*- β -hexosyl (**14**) obtained in positive ionisation mode.

m/z 355 ($[M-H-2 \times 120]^+$), m/z 337 ($[M-H-2 \times 120-H_2O]^+$) and m/z 325 ($[M-H-150-120]^+$). These data are in accordance with earlier reports where apigenin-6,8-di-*C*-glucoside was detected in *A. linearis* (Beelders *et al.*, 2012c; Beltrán-Debón *et al.*, 2011).

Flavone and flavanone *O*-glucosides

Compounds **12** and **13** were detected at retention times of 10.04 and 10.63 min, respectively, and both displayed molecular ions at m/z of 449 ($[M-H]^-$) and UV spectra characteristic of flavanones. MS/MS data showed fragments at m/z 287 ($[M-H-162]^-$) and m/z 151 for both molecules. Based on these data, compounds **12** and **13** were tentatively identified as eriodictyol-*O*-glucoside isomers. These compounds are believed to be isomers of one another, as one displayed a higher intensity of the m/z 287 ion and the other of the m/z 151 ion (Table 3.1). Eriodictyol-5-*O*-glucoside and -7-*O*-glucoside have previously been isolated from the related species, *C. intermedia*, and identified by nuclear magnetic resonance (NMR) spectroscopy (Kamara *et al.*, 2003). Assignment of peaks **12** and **13** as one or the other of these isomers was however not possible based on relative retention or MS/MS spectral information. Alternatively, these compounds may also be the (*R*)- and (*S*)-diastereomers of eriodictyol-5-*O*-glucoside, which, in contrast to the corresponding diastereomers of eriodictyol-7-*O*-glucoside, can be separated by reversed-phase liquid chromatography (RP-LC) (Pan *et al.*, 2008).

Compound **16** exhibited a molecular ion at m/z 593 $[M-H]^-$ and a retention time of 13.29 min. UV and MS/MS spectra of this compound corresponded to those of the flavone derivative, scolymoside (luteolin-7-*O*-rutoside). A major fragment was observed at m/z 285 $[M-H-308]^-$, which is consistent with the loss of the rutoside moiety. The presence of scolymoside has previously been established in *C. subternata* (Kokotkiewicz

et al., 2012; De Beer & Joubert, 2010).

Additional unidentified phenolic compounds

Several additional compounds could not be identified based on the available data, although UV-Vis and MS data point to their phenolic nature. These compounds together with their relevant UV, MS and MS/MS data are listed in Table 3.2. Further research is required to determine the identity of these molecules. Compounds **8** and **18** may possibly be naringenin derivatives, but due to the lack of MS/MS data in literature, they could not be identified.

Cyclopia maculata

In this discussion of *C. maculata* phenolics, references will be made to Figs. 3.3 and 3.4, and Table 3.3. The following compounds identified in *C. maculata* correspond to those confirmed or tentatively identified in *C. subternata*: maclurin-3-*C*- β -glucoside (**3**), iriflophenone-3-*C*-glucoside (**6**), mangiferin (**9**), isomangiferin (**10**), apigenin-6,8-di-*C*-glucoside (**11**), two eriodictyol-*O*-glucosides (**12,13**), 3-hydroxyphloretin-3',5'-di-*C*-hexoside (**14**), eriodictyol-7-*O*-rutinoside (**15**), luteolin-7-*O*-rutinoside (**16**), phloretin-3',5'-di-*C*-glucoside (**17**) and hesperidin (**19**). These compounds showed identical UV, MS and MS/MS characteristics as outlined above for *C. subternata* (note though that retention times vary slightly as a result of differences in gradient profiles). The following discussion will focus only on compounds uniquely identified in *C. maculata*.

Xanthones

Compounds **21** and **22** were characterised by molecular ions at m/z 437 ($[M-H]^-$) and proposed molecular formula of $C_{19}H_{18}O_{12}$ at retention times of 8.84 min and 10.20 min, respectively. Both displayed major fragment ions at m/z 317 ($[M-H-120]^-$) and m/z 347 ($[M-H-90]^-$), indicating *C*-glycosidic structures. Furthermore, compounds **21** and **22** differ from isomers **9** and **10** by 16 amu and have similar UV-spectra. These data suggest hydroxylated derivatives of mangiferin and isomangiferin, respectively. The principal MS^E fragments confirm that the additional hydroxyl group is attached to the aglycone moiety in both cases. According to their relative RP-LC elution order, compounds **21** and **22** are presumably the hydroxylated derivatives of mangiferin and isomangiferin, respectively. They eluted slightly later than compounds **9** and **10**, respectively, which was surprising in view of the additional OH group. Hydroxylation most likely occurs at positions 1 or 4 (indicated by R₃ or R₄) on the A ring (Fig. 3.4), although this could not be confirmed based on the available MS data.

Unidentified

Compound **20**, with proposed molecular formula of $C_{40}H_{39}O_{22}$ ($[M-H]^-$), had a retention time of 6.26 min and m/z of 871 $[M-H]^-$. This compound seems to be a derivative of mangiferin (**9**) or isomangiferin (**10**), since it contains identical fragments to these compounds at m/z 421 ((iso)mangiferin molecular ion), m/z 331 ($[M-H-90]^-$) and m/z 301 ($[M-H-120]^-$). The much higher molecular weight of compound **20** appears to indicate a dimeric derivative of (iso)mangiferin (Wang *et al.*, 1994; Isaka *et al.*, 2005), although the lower RP-LC retention time is not consistent with this, and rather points to a highly glycosylated derivative. High molecular fragments also indicates the presence of numerous *C*-glycosidic groups, for example m/z 751 ($[M-H-120]^-$), m/z 691 ($[M-H-120-60]^-$), m/z 601 ($[M-H-120-60-90]^-$) or ($[M-H-120-150]^-$), m/z 571 ($[M-H-120-60-120]^-$) and m/z 421 ($[M-H-120-60-120-150]^-$) (Fig. 3.6). It would be interesting to investigate this compound further, as no structure can be proposed with certainty based on current available data.

Table 3.3 – UV-Vis, LC-MS and LC-MS^E characteristics of phenolic compounds identified in freeze-dried aqueous extracts of green and fermented *C. maculata*

Peak ^a	Mode	t _R (min)	Accurate mass ^b	λ _{max} (nm)	Error (ppm)	Proposed molecular formula	Fragments	Phenolic compound
3	+	3.38	425.1047	235, 280,	-1.6	C ₁₉ H ₂₁ O ₁₁	407, 329, 231, 219, 195* , 177, 165	Maclurin-3- <i>C</i> -β-glucoside
	-		423.0927	316	0.0	C ₁₉ H ₁₉ O ₁₁	423, 333, 303* , 223, 193, 165, 151, 125, 109	
6	+	5.75	409.1126	293	-2.2	C ₁₉ H ₂₁ O ₁₀	391, 373, 355, 337, 325, 313, 289, 271, 259, 231, 219, 195* , 186, 177, 165	Iriflophenone-3- <i>C</i> -β-glucoside
	-		407.0978		-0.2	C ₁₉ H ₁₉ O ₁₀	407, 317, 287* , 245, 193, 165, 125, 117	
20	+	6.26	873.2100	263, 276,	1.3	C ₄₀ H ₄₁ O ₂₂	693, 577, 561* , 289, 257, 186, 145	Unidentified
	-		871.1915	319, 376	2.1	C ₄₀ H ₃₉ O ₂₂	871, 751, 691, 557, 539, 421* , 417, 331, 301, 269, 175, 152	
9	+	8.55	423.0930	239, 257,	0.7	C ₁₉ H ₁₉ O ₁₁	423* , 405, 387, 369, 339, 327, 303, 273, 257, 186	Mangiferin
	-		421.0768	317, 366	-0.7	C ₁₉ H ₁₇ O ₁₁	421* , 403, 331, 301, 271, 259	
21	+	8.84	439.0878	237, 259,	0.2	C ₁₉ H ₁₉ O ₁₂	411, 319, 289* , 273, 241	Hydroxymangiferin
	-		437.0716	327	-0.9	C ₁₉ H ₁₇ O ₁₂	437* , 421, 347, 317, 301, 288, 271, 175, 137	
10	+	9.05	423.0931	239, 255,	0.9	C ₁₉ H ₁₉ O ₁₁	423* , 405, 387, 369, 357, 327, 303, 273, 186	Isomangiferin
	-		421.0767	316, 365	-0.9	C ₁₉ H ₁₇ O ₁₁	421* , 331, 301, 271, 258	
11	+	9.24	595.1661	235, 271, 329	0.3	C ₂₇ H ₃₁ O ₁₅	595* , 577, 559, 543, 457, 395, 379, 337, 325, 295, 273, 145, 130	Apigenin-6,8-di- <i>C</i> -glucoside [Vicenin-2]
	-		593.1498		-1.3	C ₂₇ H ₂₉ O ₁₅	593* , 503, 473, 421, 383, 353, 331, 325, 301, 297	
12	+	9.61	451.1235	281	-1.1	C ₂₁ H ₂₃ O ₁₁	289* , 163, 153	Eriodictyol- <i>O</i> -glucoside
	-		449.1080		0.9	C ₂₁ H ₂₁ O ₁₁	449, 433, 407, 287* , 271, 151, 135	
22	+	10.20	439.0879	236, 255,	0.5	C ₁₉ H ₁₉ O ₁₂	439, 419, 319, 289, 273*	Hydroxyisomangiferin
	-		437.0727	272, 323	-1.6	C ₁₉ H ₁₇ O ₁₂	437* , 433, 347, 317, 289, 271, 175, 125	
13	+	10.39	451.1256	282	3.5	C ₂₁ H ₂₃ O ₁₁	289* , 163, 153	Eriodictyol- <i>O</i> -glucoside
	-		449.1087		0.7	C ₂₁ H ₂₁ O ₁₁	449, 433, 287* , 271, 151, 135	
14	+	11.78	615.1906	279	-3.1	C ₂₇ H ₃₅ O ₁₆	597, 579* , 495, 477, 447, 411, 381, 289, 273, 259, 205, 165, 123	3-Hydroxyphloretin-3',5'-di- <i>C</i> - hexoside
	-		613.1779		1.6	C ₂₇ H ₃₃ O ₁₆	613* , 493, 475, 433, 403, 373, 363, 331, 252, 209, 135	
15	+	12.70	597.1813	281	-1.0	C ₂₇ H ₃₃ O ₁₅	597, 487, 435, 417, 289* , 273, 153	Eriodictyol-7- <i>O</i> -rutinoside
	-		595.1660		-0.5	C ₂₇ H ₃₁ O ₁₅	595, 459, 433, 287* , 271, 151, 135	[Eriocitrin]
16	+	13.76	595.1661	240, 278,	-0.3	C ₂₇ H ₃₁ O ₁₅	595, 487, 449, 419, 287* , 273	Luteolin-7- <i>O</i> -rutinoside
	-		593.1519		2.2	C ₂₇ H ₂₉ O ₁₅	593* , 285	[Scolymoside]
17	+	14.18	599.1974	280	-0.3	C ₂₇ H ₃₅ O ₁₅	575* , 545, 527, 489, 461, 435, 419, 365	Phloretin-3',5'-di- <i>C</i> -glucoside
	-		597.1826		1.2	C ₂₇ H ₃₃ O ₁₅	597*, 477, 417, 387, 357* , 315, 209	
19	+	16.73	611.1979	283	0.5	C ₂₈ H ₃₅ O ₁₅	611, 449, 431, 413, 345, 303* , 195, 177, 153	Hesperitin-7- <i>O</i> -rutinoside
	-		609.1814		-0.8	C ₂₈ H ₃₃ O ₁₅	609, 301*	[Hesperedin]

^aPeak numbers correspond to numbered peaks in Figure 3.3; ^bAccurate mass determined experimentally; *Ion with highest relative intensity

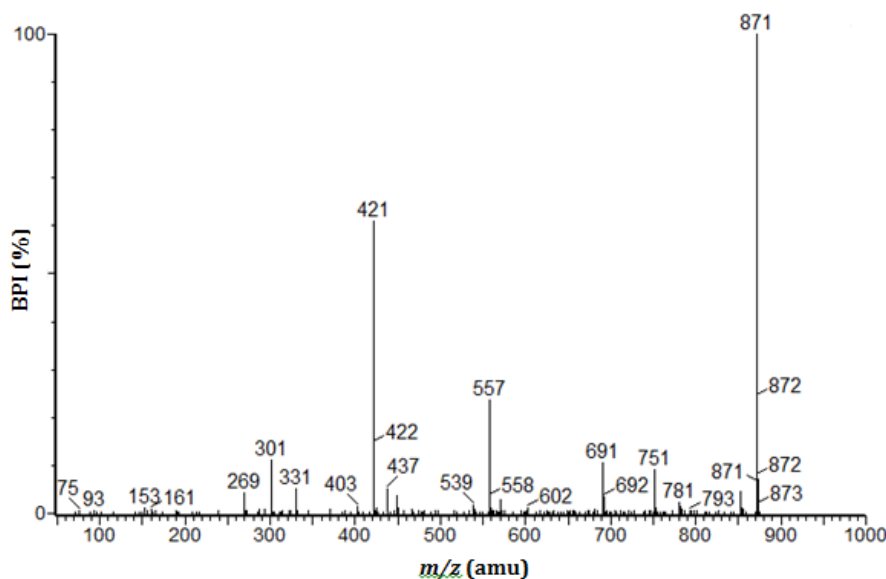


Figure 3.6 – LC-MS/MS spectrum of unidentified compound **20** obtained in negative ionisation mode.

3.4.3 Method validations

Validation of both methods was performed in order to determine their suitability for the routine analysis of fermented and green *C. subternata* and *C. maculata* phenolics. It should be noted that iriflophenone-3-*C*- β -glucoside and maclurin-3-*C*- β -glucoside could not be quantified in fermented extracts due to their substantially lower levels compared to green extracts, resulting in poor resolution with other minor compounds.

Specificity

The methods were deemed specific for the quantified peaks, as their UV-Vis and MS spectra matched those of authentic standards or values published earlier as described in the previous section.

Linearity

The linearity of the standard calibration curves, as determined with the two HPLC methods, were excellent (Tables 3.4 and 3.5). All y-intercepts were fairly low and the Pearson’s product moment correlation coefficients were all equal to 1.000.

Table 3.4 – Calibration curves for phenolic standards using the method for *C. subternata* extracts

Compound	Regression equation	<i>r</i>
Mangiferin	$y = 2089.4x + 7.3$	1.000
Aspalathin	$y = 2351.8x + 1.0$	1.000
Eriocitrin	$y = 1611.6x - 0.3$	1.000
Hesperidin	$y = 1782.1x - 1.8$	1.000
Luteolin	$y = 2781.8x - 4.0$	1.000

Stability (in the presence of ascorbic acid as antioxidant)

Ascorbic acid was added to standard calibration mixtures and samples, as it proved to be crucial to ensure stability of phenolic compounds, especially dihydrochalcones, in rooibos infusions (Joubert *et al.*, 2012). The

Table 3.5 – Calibration curves for phenolic standards using the method for *C. maculata* extracts

Compound	Regression equation	<i>r</i>
Mangiferin	$y = 1874.3x + 16.1$	1.000
Isomangiferin	$y = 2039.0x - 0.7$	1.000
Eriocitrin	$y = 1602.2x + 4.3$	1.000
Hesperidin	$y = 1772.9x + 5.8$	1.000

stability of the compounds in both the standard calibration mixtures and the reconstituted *C. subternata* extract was good when evaluated over a period of 26 h. As indicated in Table 3.6, the % change for all the compounds in *C. subternata* ranged from -4.4% to 4.0%. The exception was 3-hydroxyphloretin-3',5'-di-*C*-hexoside, which increased by 5.5% in the unfermented extract and 9.1% in the fermented extract. This could possibly be due to the co-elution of degradation products, or the instability of this dihydrochalcone, and a higher concentration of ascorbic acid could be needed. RSD values for all these compounds were less than 4% over the 26 h period. Therefore, no substantial decrease occurred in the levels of any of the compounds in *C. subternata* extracts or the calibration mixture.

Table 3.6 – Compound stability in standard calibration mixtures and reconstituted freeze-dried aqueous extracts of green and fermented *C. subternata* extracts over a 26 h period (injected in *ca.* 3 h intervals)

Compound	%RSD (% change)			
	Calibration mixture 2 μ L	Calibration mixture 15 μ L	Unfermented extract	Fermented extract
Mangiferin	0.6 (1.4)	0.2 (0.4)	0.9 (-2.7)	1.9 (-1.7)
Aspalathin	0.8 (-0.9)	0.4 (-1.1)	-	-
Eriocitrin	0.4 (0.4)	0.1 (0.2)	0.8 (-0.2)	0.7 (-2.3)
Hesperidin	1.1 (-2.1)	0.6 (-1.8)	1.2 (-3.3)	1.0 (-3.5)
Luteolin	0.9 (-1.3)	0.4 (-1.3)	-	-
Isomangiferin	-	-	0.9 (-1.8)	1.5 (-4.4)
Iriflophenone-3- <i>C</i> - β - glucoside	-	-	1.2 (-2.3)	-
3-Hydroxyphloretin- 3',5'-di- <i>C</i> -hexoside	-	-	2.0 (5.5)	2.9 (9.1)
Scolymoside	-	-	0.9 (-2.3)	0.7 (-1.7)
Phloretin-3',5'-di- <i>C</i> - β -glucoside	-	-	0.7 (-2.0)	0.6 (-1.2)
Iriflophenone-di- <i>O,C</i> - hexoside	-	-	1.5 (-3.6)	1.4 (-1.9)
Maclurin-3- <i>C</i> - β - glucoside	-	-	3.9 (4.0)	-

%RSD = relative standard deviation

Good stability of the compounds in the standard calibration mixtures and reconstituted *C. maculata* extracts analysed over a period of 24 h was also observed (Table 3.7). For *C. maculata*, the % change for all the compounds ranged from -4.0% to 3.3%. The %RSD for the compounds were less than 3% over the 24 h period. No substantial decrease occurred for any of the *C. maculata* compounds in the extracts or calibration mixture.

Precision

The intra- and inter-day precision of both methods were very good. The %RSD for all the compounds in the standard calibration mixture for *C. subternata* and its reconstituted extracts were $\leq 2\%$ (Table 3.8). Similarly, the %RSD for all the compounds in the standard calibration mixture for *C. maculata* and its reconstituted extracts were less than 5% (Table 3.9).

Table 3.7 – Compound stability in standard calibration mixtures and reconstituted freeze-dried aqueous extracts of green and fermented *C. maculata* over a 24 h period (injected in *ca.* 3 h intervals)

Compound	%RSD (% change)			
	Calibration mixture	Calibration mixture	Unfermented	Fermented
	2 μ L	15 μ L	extract	extract
Mangiferin	0.3 (0.7)	0.2 (0.2)	0.3 (-0.9)	0.4 (-1.3)
Isomangiferin	1.0 (1.1)	0.2 (-0.3)	0.4 (-0.6)	0.4 (-0.1)
Eriocitrin	0.6 (1.1)	0.3 (0.6)	2.8 (0.2)	2.9 (3.3)
Hesperidin	0.4 (1.2)	0.2 (0.3)	0.7 (-1.6)	1.0 (-0.2)
Iriflophenone-3- <i>C</i> - β -glucoside	-	-	0.6 (1.4)	-
Maclurin-3- <i>C</i> - β -glucoside	-	-	2.3 (-4.0)	-

%RSD = relative standard deviation

3.4.4 Differences in phenolic composition of green honeybush seedlings

Green extracts from *C. subternata* (n = 64) and *C. maculata* (n = 40) were analysed for their quantitative phenolic composition. The quantitative results are summarised in Fig. 3.7. More compounds could be quantified in *C. subternata* than in *C. maculata* extracts. Consequently, mangiferin, isomangiferin, scolymoside, iriflophenone-3-*C*-glucoside, iriflophenone-di-*O,C*-hexoside, maclurin-3-*C*- β -glucoside, 3-hydroxyphloretin-3',5'-di-*C*-hexoside, eriocitrin, phloretin-3',5'-di-*C*- β -glucoside and hesperidin were quantified in green *C. subternata* extracts, compared to mangiferin, isomangiferin, iriflophenone-3-*C*-glucoside, maclurin-3-*C*- β -glucoside, eriocitrin and hesperidin in the green *C. maculata* extracts.

It is clear from Fig. 3.7 that the two major xanthones, mangiferin and isomangiferin, and the flavanone, hesperidin, are present in much larger amounts in *C. maculata* compared to *C. subternata*. Phloretin-3',5'-di-*C*- β -glucoside was present at higher levels in *C. subternata*, and in fact could not be quantified in *C. maculata* extracts. Scolymoside, iriflophenone-3-*C*-glucoside and hesperidin displayed the largest quantitative variation in *C. subternata*, while iriflophenone-3-*C*-glucoside and eriocitrin displayed the largest quantitative variation in *C. maculata*. Maclurin-3-*C*- β -glucoside was present in very small amounts in both species, especially *C. subternata* and is therefore not regarded as important for quantification in *C. subternata* during routine analysis.

When comparing the data from the *C. subternata* extracts analysed here to previous data presented in De Beer & Joubert (2010) and Joubert *et al.* (2008b), a large variation is observed for compounds mangiferin and isomangiferin. In De Beer & Joubert (2010), mangiferin and isomangiferin quantities in green *C. subternata* were reported as 2.7% and 0.9%, compared to 0.9% and 0.3%, respectively, when analysed by the new method, which could indicate previous overestimation of these compounds, although mangiferin levels as high as 2.2% was observed for the present sample set. Furthermore, aqueous extracts of green *C. maculata* were quantified here for the first time. The mangiferin content observed in the green *C. maculata* extracts is not only higher than that observed for *C. subternata*, but also compared to values reported for *C. intermedia* and *C. sessiflora* extracts. However, *Cyclopia genistoides* remains the *Cyclopia* species with the highest quantity of mangiferin (Joubert *et al.*, 2008b; De Beer & Joubert, 2010).

These examples express the large variation in phenolic content observed on both an intra- and inter-species level. Considering the extracts analysed were all prepared from different seedlings, the variation on an intra-species level has implications for extract standardisation and quality control.

Table 3.8 – Intra- and inter-day precision (%RSD) for individual phenolic compounds as determined using standard calibration mixtures, and reconstituted freeze-dried aqueous extracts of green and fermented *C. subternata*

Compound	Intra-day (n = 6/day)			Inter-day (n = 3)
	Day 1	Day 2	Day 3	
<i>Calibration mixture (2 µL)</i>				
Mangiferin	0.5	0.3	0.2	1.6
Luteolin	0.3	0.6	0.5	1.2
Hesperidin	0.8	0.6	0.8	0.8
Eriocitrin	0.6	0.4	0.4	0.7
Aspalathin	0.7	0.3	0.6	1.4
<i>Calibration mixture (15 µL)</i>				
Mangiferin	0.1	0.1	0.1	1.6
Luteolin	0.1	0.1	0.4	0.8
Hesperidin	0.1	0.2	0.4	0.1
Eriocitrin	0.2	0.1	0.2	0.6
Aspalathin	0.0	0.1	0.3	1.1
<i>Unfermented C. subternata</i>				
Iriflophenone-3- <i>C</i> -β-glucoside	0.2	0.4	0.1	0.3
Mangiferin	0.1	0.1	0.1	0.1
Isomangiferin	0.1	0.1	0.1	0.2
3-Hydroxyphloretin-3',5'-di- <i>C</i> -hexoside	1.1	1.3	0.7	1.6
Eriocitrin	1.4	1.6	1.2	2.0
Scolymoside	0.1	0.1	0.2	0.3
Phloretin-3',5'-di- <i>C</i> -β-glucoside	0.1	0.2	0.2	0.1
Hesperidin	0.3	0.2	0.4	0.3
Iriflophenone-di- <i>O,C</i> -hexoside	0.6	1.1	0.2	0.9
Maclurin-3- <i>C</i> -β-hexoside	0.9	1.4	1.1	0.8
<i>Fermented C. subternata</i>				
Mangiferin	0.2	0.4	0.7	1.5
Isomangiferin	0.2	0.1	0.1	1.2
3-Hydroxyphloretin-3',5'-di- <i>C</i> -hexoside	2.0	1.7	1.8	1.6
Eriocitrin	0.5	0.6	0.2	1.2
Scolymoside	0.2	0.3	1.0	0.4
Phloretin-3',5'-di- <i>C</i> -β-glucoside	0.2	0.8	1.0	0.4
Hesperidin	0.4	0.3	0.3	0.7
Iriflophenone-di- <i>O,C</i> -hexoside	0.5	0.3	0.3	0.2

Table 3.9 – Intra- and inter-day precision (%RSD) for individual phenolic compounds as determined using standard calibration mixtures, and reconstituted freeze-dried green and fermented *C. maculata* extracts

Compound	Intra-day (n = 6/day)			Inter-day (n = 3)
	Day 1	Day 2	Day 3	
<i>Calibration mixture (2 µL)</i>				
Mangiferin	0.5	0.5	0.5	1.6
Isomangiferin	1.3	1.5	1.4	1.1
Hesperidin	0.4	0.4	0.3	0.2
Eriocitrin	0.4	0.4	0.3	0.8
<i>Calibration mixture (15 µL)</i>				
Mangiferin	0.1	0.1	0.2	1.7
Isomangiferin	0.2	0.2	0.3	1.5
Hesperidin	0.1	0.0	0.2	0.2
Eriocitrin	0.1	0.1	0.2	1.0
<i>Unfermented C. maculata</i>				
Iriflophenone-3- <i>C</i> - β -glucoside	4.9	1.5	4.2	2.7
Mangiferin	0.1	0.1	0.1	0.3
Isomangiferin	0.1	0.1	0.1	0.3
Eriocitrin	4.0	3.3	2.3	3.4
Hesperidin	0.2	0.6	0.6	0.9
Maclurin-3- <i>C</i> - β -glucoside	2.2	2.0	3.1	4.8
<i>Fermented C. maculata</i>				
Mangiferin	0.2	0.1	0.3	0.6
Isomangiferin	0.4	0.4	0.3	1.2
Eriocitrin	1.3	2.7	3.9	3.4
Hesperidin	0.8	0.5	0.1	2.2

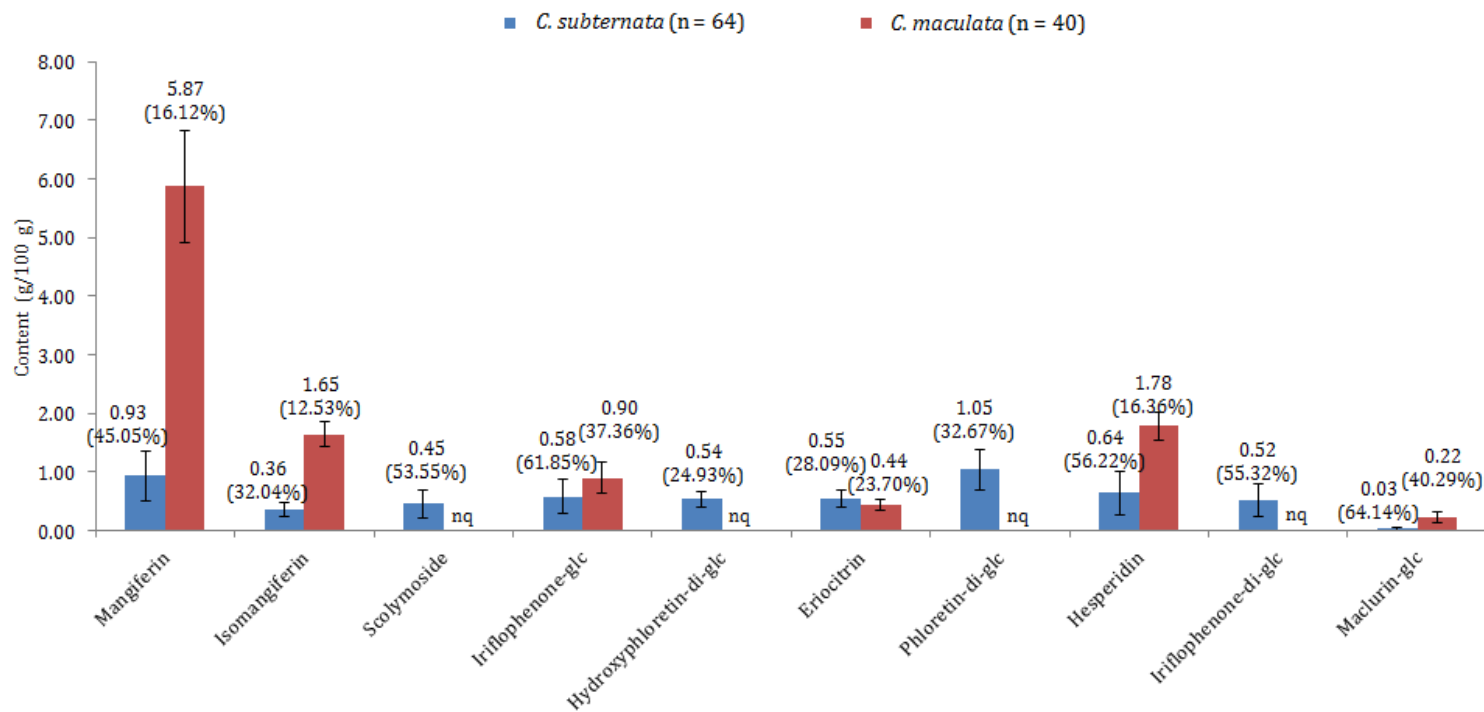


Figure 3.7 – Phenolic composition (g/100 g dried extract) of freeze-dried aqueous extracts of green *C. subternata* and *C. maculata* [Bars indicate the mean values \pm standard deviation. Average values are indicated on top of bars with relative standard deviation in brackets; n.q., compounds not quantified; glc, glucoside].

3.5 Conclusions

HPLC-DAD methods were successfully developed for the quantification of the major phenolic compounds present in two *Cyclopia* species, *C. subternata* and *C. maculata*. Several new compounds were tentatively identified for the first time in *Cyclopia* species by means of LC-ESI-MS. Both methods were validated and deemed reliable for routine HPLC quantitative analysis of green and fermented *C. subternata* and *C. maculata* extracts. These methods were employed for the analysis of a large number of hot water extracts of green *C. subternata* and *C. maculata* seedlings. The data obtained indicate the large variation in phenolic composition that exists between these two *Cyclopia* species, as well as intra-species variation. Furthermore, the separation of the major phenolic compounds of both species also contribute to improved chemical profiles (chromatographic fingerprints), which could be used in future for quality control analysis and determining authenticity of extracts. These methods will also serve as a basis for further research such as sensory characterisation, fermentation optimisation, bioactivity studies and studies on the role of environmental effects on the phenolic compounds, as well as further fingerprint analysis.

References

- Abad-García, B., Berrueta, L.A., Garmón-Lobato, S., Gallo, B. & Vicente, F. (2009). A general analytical strategy for the characterization of phenolic compounds in fruit juices by high-performance liquid chromatography with diode array detection coupled to electron ionization and triple quadrupole mass spectrometry. *Journal of Chromatography A*, **1216**, 5398-5415.
- Beelders, T., Kalili, K.M., Joubert, E., de Beer, D. & de Villiers, A.J. (2012a). Comprehensive two-dimensional liquid chromatographic analysis of rooibos (*Aspalathus linearis*) phenolics. *Journal of Separation Science*, **35**, 1808-1820.
- Beelders, T., Sigge, G.O., Joubert, E., de Beer, D. & de Villiers, A. (2012b). Erratum to "Kinetic optimisation of the reversed phase liquid chromatographic separation of rooibos tea (*Aspalathus linearis*) phenolics on conventional high performance liquid chromatographic instrumentation". [J. Chromatogr. A 1219 (2012) 128-139]. *Journal of Chromatography A*, **1241**, 128.
- Beelders, T., Sigge, G.O., Joubert, E., de Beer, D. & de Villiers, A. (2012c). Kinetic optimisation of the reversed phase liquid chromatographic separation of rooibos tea (*Aspalathus linearis*) phenolics on conventional high performance liquid chromatographic instrumentation. *Journal of Chromatography A*, **1219**, 128-139.
- Beltrán-Debón R., Rull, A., Rodríguez-Sanabria, F., Iswaldi, I., Herranz-López, M., Aragonès, G., Camps, J., Alonso-Villaverde, C., Menéndez, J.A., Micol, V., Segura-Carretero, A. & Joven, J. (2011). Continuous administration of polyphenols from aqueous rooibos (*Aspalathus linearis*) extract ameliorates dietary-induced metabolic disturbances in hyperlipidemic mice. *Phytomedicine*, **18**, 414-424.
- De Beer, D. & Joubert E. (2010). Development of HPLC method for *Cyclopia subternata* phenolic compound analysis and application to other *Cyclopia* spp. *Journal of Food Composition and Analysis*, **23**, 289-297.
- De Beer, D., Jerz, G., Joubert, E., Wray, V. & Winterhalter, P. (2009). Isolation of isomangiferin from honeybush (*Cyclopia subternata*) using high-speed counter-current chromatography and high-performance liquid chromatography. *Journal of Chromatography A*, **52**, 5391-5395.
- Du Toit, J. Joubert, E. & Britz, T.J. (1998). Honeybush tea, a rediscovered indigenous South African herbal tea. *Journal of Sustainable Agriculture*, **12**, 67-84.
- Dudhia, Z., Louw, J., Muller, C., Joubert, E., de Beer, D., Kinnear, C. & Pheiffer, C. (2013). *Cyclopia maculata* and *Cyclopia subternata* (honeybush tea) inhibits adipogenesis in 3T3-L1 pre-adepocytes. *Phytomedicine*, **20**, 401-408.
- Ferreira, D., Kamara, B.I., Brandt, E.V. & Joubert, E. (1998). Phenolic compounds from *Cyclopia intermedia* (honeybush tea). *Journal of Agricultural and Food Chemistry*, **46**, 3406-4310.
- Isaka, M., Palasarn, S., Kocharin, K. & Saenboonrueng, J. (2005). A cytotoxic xanthone dimer from the entomopathogenic fungus *Aschersonia* sp. BCC 8401. *Journal of Natural Products*, **68**, 945-946.
- Joubert, E., Beelders, T., De Beer, D., Malherbe, C.J., De Villiers, A.J. & Sigge, G.O. (2012). Variation in phenolic content and antioxidant activity of fermented rooibos herbal tea infusions: Role of production season and quality grade. *Journal of Agricultural and Food Chemistry*, **60**, 9171-9179.
- Joubert, E., Gelderblom, W.C.A., Louw, A. & De Beer, D. (2008a). South African herbal teas: *Aspalathus linearis*, *Cyclopia* spp. and *Athrixia phylloides* - A review. *Journal of Ethnopharmacology*, **119**, 374-412.
- Joubert, E., Joubert, M.E., Bester, C., De Beer, D. & De Lange, J.H. (2011). Honeybush (*Cyclopia* spp.): From local cottage industry to global markets - The catalytic and supporting role of research. *South*

African Journal of Botany, **77**, 877-907.

- Joubert, E., Otto, F., Grüner, S. & Weinreich, B. (2003). Reversed-phase HPLC determination of mangiferin, isomangiferin and hesperidin in *Cyclopia* and the effect of harvesting date on the phenolic composition of *C. genistoides*. *European Food and Research Technology*, **216**, 270-273.
- Joubert, E., Richards, E.S., Van der Merwe, J.D., Manley, M. & Gelderblom, W.C.A. (2008b). Effects of species variation and processing on phenolic composition and in vitro antioxidant activity of aqueous extracts of *Cyclopia* spp. (honeybush tea). *Journal of Agricultural and Food Chemistry*, **56**, 954-963.
- Kamara, B.I., Brand, D.J., Brandt, E.V. & Joubert, E. (2004). Phenolic metabolites from honeybush tea (*Cyclopia subternata*). *Journal of Agricultural and Food Chemistry*, **52**, 5291-5395.
- Kamara, B.I., Brandt, E.V., Ferreira, D. & Joubert, E. (2003). Polyphenols from honeybush tea (*Cyclopia intermedia*). *Journal of Agricultural and Food Chemistry*, **51**, 3874-3979.
- Kokotkiewicz, A., Luczkiewicz, M., Sowinski, P., Glod, D., Gorynski, K. & Bucinski, A. (2012). Isolation and structure elucidation of phenolic compounds from *Cyclopia subternata* Vogel (honeybush) intact plant and *in vitro* cultures. *Food Chemistry*, **133**, 1373-1382.
- Kokotkiewicz, A., Luczkiewicz, M., Pawlowska, J., Luczkiewicz, P., Sowinski, P., Witkowski, J., Bryl, E. & Bucinski, A. (2013). Isolation of xanthone and benzophenone derivatives from *Cyclopia genistoides* (L.) Vent. (honeybush) and their pro-apoptotic activity in synoviocytes from patients with rheumatoid arthritis. *Fitoterapia*, **90**, 199-208.
- Mose Larsen, P., Fey, S.J., Louw, J. & Joubert, E. (2008). An anti-diabetic extract of honeybush. World Intellectual Property Organization WO. 2008/110552A2.
- Muller, C.J.F., Joubert, E., Gabuza, K., De Beer, D., Fey, S.J. & Louw, J. (2011). Assessment of the antidiabetic potential of an aqueous extract of honeybush (*Cyclopia intermedia*) in streptozotocin and obese resistant Wistar rats. In: *Phytochemicals - bioactivities and impact on health*. Pp. 313-332. Rijeka: Intech.
- Pan, J., Zhang, S., Yan, L., Tai, J., Xiao, Q., Zou, K., Zhou, Y. & Wy, J. (2008). Separation of flavanone enantiomers and flavanone diastereomers from *Balanophora involucretrata* Hook. f. by capillary electrophoresis and reversed-phase high-performance liquid chromatography on a C18 column. *Journal of Chromatography A*, **1185**, 117-129.
- Tanaka, T., Sueyasu, T., Nonaka, G.I., Nishioka, I. (1984). Tannins and related compounds. XXI. Isolation and characterization of galloyl and *p*-hydroxybenzoyl esters of benzophenone and xanthone *C*-glucosides from *Mangifera indica* L. *Chemical and Pharmaceutical Bulletin*, **32**, 2676-2686.
- Wang, J., Hou, C. & Lu. Y. (1994). Swertifrancheside, an HIV-reverse transcriptase inhibitor and the first flavone-xanthone dimer, from *Swertia franchetiana*. *Journal of Natural Products*, **57**, 211-217.
- Zhang, Y., Han, L., Ge, D., Lui, X., Liu, E., Wu, C., Gao, X. & Wang, T. (2013). Isolation, structural elucidation, MS profiling, and evaluation of triglyceride accumulation inhibitory effects of benzophenone *C*-glucosides from leaves of *Mangifera indica* L. *Journal of Agricultural and Food Chemistry*, **61**, 1884-1895.

Chapter 4

HPLC chromatographic fingerprint analyses of *Cyclopia subternata* and *Cyclopia maculata* extracts: Similarity to biologically active extracts and identification of potentially active compounds

Part of this chapter has been presented as a poster at AfroData 2012, the 2nd African-European Conference on Chemometrics (19-23 November 2012):

Schulze, A., Joubert, E., Manley, M., De Villiers, A. & De Beer, D. Similarity analysis of aqueous extracts from *Cyclopia subternata* (honeybush) seedlings using chromatographic fingerprints.

4.1 Abstract

The use of similarity analysis and principal component analysis (PCA) on chromatographic fingerprints of a large number of randomly selected plant extracts of two *Cyclopia* species was investigated. *Cyclopia subternata* (n = 64) and *C. maculata* (n = 50) extracts were analysed by high-performance liquid chromatography (HPLC) to obtain the chromatographic fingerprints of each sample. An extract of each species that displayed biological activity in previous studies was included to determine which of the extracts from the two large sample sets were most similar to these active extracts. The active extracts of *C. subternata* (n = 1) and *C. maculata* (n = 1) demonstrated anti-obesity properties in a cell model, while the *C. subternata* extract also proved effective in lowering elevated blood glucose levels in streptozotocin-induced diabetic Wistar rats and diet-induced diabetic Vervet monkeys. Similarity analysis and PCA were only employed after extensive pre-treatment of the chromatographic data. Several of the *C. subternata* extracts correlated well ($r > 0.95$; n = 15) with its biologically active extract according to similarity analysis, and clustered together in PCA score plots. For *C. maculata* all the extracts were extremely similar to the active extract based on similarity analysis ($r > 0.97$; n = 50). It was therefore necessary to increase the threshold. Many of the samples with $r > 0.99$ also clustered with the active extract in the PCA score plots (n = 10). It was clear that for both data analysis techniques the contribution of the small UV-Vis peaks were dominated by the large peaks in the chromatogram. Therefore, very similar clusterings were observed in PCA score plots of the quantitative data. From the PCA loading line plots it was concluded that mangiferin and isomangiferin are not the major contributing compounds to the activity of the *C. subternata* extract. The clustering of highly correlated samples were due to the remaining compounds, which are considered to have a stronger relation to the activity. For *C. maculata*, with a less complex chromatogram, it seemed as though mangiferin and isomangiferin were associated with the clustering of the activity, or were inversely associated with hesperidin. These contradictory results suggest that more samples with proven activity are required for the identification of the compound(s) responsible for the activity. It should, however, be verified in cell and/or animal models, that the samples indicated as being similar to the respective active extracts do display activity. These methods should also be validated more stringently with a data set that includes more samples with proven activity. If these extracts do in fact display activity, similarity analysis in combination with PCA could be regarded as a much less expensive alternative to bio-screening of a large number of samples, or to reduce the number of samples required for cell-based testing.

4.2 Introduction

Herbal medicines have traditionally been used for the prevention and treatment of numerous human diseases. This practice used to be more popular in oriental countries, but in the past few decades, the use of herbal medicines gained popularity in Western countries. Chemical drugs are becoming more expensive and concerns about their adverse effects are increasing. Because herbal species have high pharmacological activity, low toxicity, affordability and accessibility, they are increasingly considered as alternatives to chemical drugs (Xu *et al.*, 2006; Newman & Cragg, 2007; Mander & Liu, 2010; Kingston, 2011; Tistaert *et al.*, 2011).

Quality control of herbal products and medicines is imperative as the effectiveness and quality of these products depend on the concentrations of their active ingredients. Differences that occur in the concentration of the active ingredients may also influence patient and consumer safety. These differences can be due to a number of factors such as the time of harvest, cultivation conditions, storage, drying and extraction procedure, as well as deliberate or coincidental adulterations (Anon., 2002; Liang *et al.*, 2006; Shin *et al.*, 2007; Tisteart *et al.*, 2011). New regulations (effective from 14 December 2012) by the European Food Safety Authority (EFSA) states that all nutraceuticals, functional food and drinks, and dietary supplements must comply to general health claim regulations (EC1924/2006, Article 13), which requires scientific evidence and authorisation to justify claims (Anon., 2007).

Standard quantitative analysis of one or more major or active components in a herbal product does not represent its quality accurately as its efficacy usually results from multiple compounds (Liang *et al.*, 2009) and/or synergistic effects between constituents (Williamson, 2001; Wagner & Ulrich-Merzenich, 2009; Gertsch, 2011). In many cases the active compounds in herbal medicines are not specifically known to the producers and analysts. In these cases marker compounds which do not necessarily represent activity, or compounds unique to the plant are used for quality control. Due to herbal products having such complex chemical profiles, minor differences between strongly related compounds could affect the health of individuals; even when these cannot be observed by normal quantitative analysis (Tistaert *et al.*, 2011). Approaches like the use of marker compounds might not always be suitable for quality control of herbal products, as they may lack unique chemical compounds and miss synergistic effects between numerous constituents (Gertsch, 2011; Tistaert *et al.*, 2011).

To account for these numerous limitations, the World Health Organisation (WHO) accepted fingerprint analysis for the evaluation of natural products, stating that 'if the identification of an active principle is not possible, it should be sufficient to identify a characteristic substance or mixture of substances e.g. "chromatographic fingerprint" to ensure consistent quality of the preparation' (Anon., 1991). A fingerprint is defined as a characteristic profile which chemically represents the composition of the sample. Fingerprints can be obtained by chromatographic methods, which are regarded to be more informative than those from spectroscopic methods. Chromatography can spread the information of the complex sample over a time axis, revealing the individual compounds as well as the underlying information of smaller peaks that are usually not quantified or taken into account (Tistaert *et al.*, 2011).

Screening for active components in herbal medicines is very important for the pharmaceutical discovery process (identification of new products with pharmaceutical activity or potential), the explanation of the remedial mechanism of the products, as well as for quality control. The more traditional methods such as animal models that are used for screening active compounds are expensive and time-consuming, and require many samples. Recently high throughput screening, using molecular and cell models, have been developed to handle the large number of compounds generated by combinatorial chemistry and have found application in analysis

of active extracts and fractions (Mishra *et al.*, 2008; Tu *et al.*, 2010). Advances in chromatography leading to improved chromatographic fingerprints, combined with pre-treatment of data, allow an alternative approach, i.e. similarity analysis. In the latter case fingerprints of standardised extracts (or active extracts) are used as references (Gad *et al.*, 2013).

In order to apply multivariate data analysis to chromatographic fingerprints, pre-treatment has to be applied to the raw chromatograms. Pre-treatment techniques generally employed include warping to align peaks (Tomasi *et al.*, 2004), removing background noise (Eilers, 2004) and normalisation (Hendriks *et al.*, 2005). Applying the correct pre-treatment methods enables one to obtain the proper information from the data.

Similarity analysis, a type of unsupervised data analysis, can be done by evaluating Pearson's product-moment correlation coefficients (r) between samples (Alaerts *et al.*, 2010). By definition, the correlation coefficient between two chemical fingerprints is 'equal to the scalar product of the normed mean-centred fingerprints' (Alaerts *et al.*, 2010). The closer the r value is to 1, the higher the similarity between the fingerprints. In this type of evaluation, samples are compared one by one (or in pairs when duplicates are included) and the (dis)similarity between fingerprints can successfully be analysed (Alaerts *et al.*, 2012).

The objectives of this study were to determine the (dis)similarities between the chromatograms of a large number of *Cyclopia subternata* and *C. maculata* extracts and two active extracts; one from each species. Similarity analysis and principal component analysis (PCA) were used to determine these (dis)similarities and to identify the compounds responsible for the largest potential contribution to the activity. The use of similarity analysis and PCA were also evaluated to determine whether these two methods could be employed as preliminary screening methods before cell model testing. If efficient, these methods could be used for the economical selection of plants to propagate for the production of more plant material for testing and bioactivity confirmation.

4.3 Experimental

4.3.1 Plant material and extract preparation

Cyclopia subternata seedling plants ($n = 64$) of the same age were harvested from a commercial plantation, established on Kanetberg Flora in the Barrydale district, South Africa. The harvested shoots (stems and leaves) were dried at 40 °C in a temperature-controlled drying tunnel with forced air circulation to *ca.* 8-10% moisture content and ground with a Retsch mill (1 mm sieve; Retsch GmbH, Haan, Germany). These samples represent green *C. subternata*.

In the case of *Cyclopia maculata*, two sets of plant material were used. Set one comprised the shoots of seedlings ($n = 40$; 2-year-old) grown in pots at ARC Infruitec-Nietvoorbij, Stellenbosch, South Africa. The seeds for the pot trial were collected from a natural population, growing in marsh land on Welgedacht farm near Riversdale, South Africa. The seedlings were topped after 6 months to stimulate sprouting. The regrowth (top 30 cm) was harvested and consisted of thin stems and leaves. These harvested shoots (material from each pot represented a sample) were dried and milled as described for *C. subternata*. These samples represent green *C. maculata* from the seedling pot trial.

Set two of *C. maculata* ($n = 10$) comprised the shoots of seedling plants of indeterminate age, harvested from the same natural population, from which the seeds were collected for the first set of samples. In this case shoots from two to three bushes made up a batch (sample). Coarse stems (thicker than *ca.* 10-13 mm) mainly

without leaves were removed, and the remaining plant material containing the thinner stems and leaves was cut to a small size (≤ 3 mm) ("tea bag cut"). The plant material was dried in a drying tunnel at 40 °C for 6 h, followed by sieving for 90 sec through a 1.4 mm (12 mesh) sieve using a SMC Mini-sifter (JM Quality Services, Cape Town, South Africa). Sieving removed most of the pieces of stem, resulting in the majority of the plant material used for the preparation of the extracts to be leaves. This plant material represents green wild-harvested *C. maculata* samples.

Aqueous extracts were prepared from all samples as described in Chapter 3, section 3.3.3.

Biologically active (i.e. active) extracts of *C. subternata* (n = 1) and *C. maculata* (n = 1) were also included in the sample sets. The active *C. subternata* extract was prepared from plants harvested at Kanetberg Flora. The plant material, processed as green honeybush herbal tea ("tea bag cut"), was obtained from a commercial tea processor in Mosselbay, South Africa. The plant material (ca. 400 kg) was batch extracted on an industrial scale with a percolator type extraction vessel using preheated (93 °C) purified water and the extract circulating for 35 min at a 1:10 (m/v) ratio (similar to laboratory extraction employed for *C. maculata*). The extract was drained, centrifuged, concentrated under vacuum, and then vacuum-dried at 40 °C for 24 h as described by Dudhia *et al.* (2013).

The active *C. maculata* extract was prepared from ca. 25 kg plant material harvested from a natural population on a farm near Riviersonderend, Overberg, South Africa. The thick stems were removed before cutting into smaller pieces ("tea bag cut"), and dried as described for the other *C. maculata* sample sets. The dried plant material was sieved (>8 mesh) to further remove coarse stem material. Multiple batches of aqueous extracts were prepared as described in Chapter 3 (section 3.3.3) but on a larger scale (150 g in 1500 mL) and pooled before freeze-drying.

The *C. subternata* extract has been shown to be effective in lowering elevated blood glucose levels in streptozotocin-induced diabetic Wistar rats and diet-induced diabetic Vervet monkeys (Mose Larsen *et al.*, 2008). Both the *C. subternata* and *C. maculata* extract demonstrated anti-obesity properties in a cell model, i.e. inhibition of adipogenesis in 3T3-L1 pre-adipocytes (Dudhia *et al.*, 2013).

4.3.2 HPLC analysis of *C. subternata* and *C. maculata* extracts

Extracts were prepared for high performance liquid chromatography (HPLC) analysis as described in Chapter 3 (section 3.3.5). The injection volumes for the *C. subternata* and *C. maculata* extracts were 15 and 10 μ L, respectively. Separation was achieved at 30 °C on a Gemini-NX C18 (150 \times 4.6 mm; 3 μ m; 110Å) column, using 2% acetic acid (A) and acetonitrile (B) as mobile phases and a flow rate of 1.0 mL/min. The mobile phase gradient for *C. subternata* was as follows: 0-2 min (8% B), 2-27 min (8%-38% B), 27-28 min (38%-50% B), 28-29 min (50% B), 29-30 (50%-8% B), 30-40 min (8% B). The mobile phase gradient for *C. maculata* was as follows: 0-2 min (8% B), 2-31 min (8%-38% B), 31-32 min (38%-50% B), 32-33 min (50% B), 33-34 (50%-8% B), 34-44 min (8% B). UV-Vis spectra were recorded from 200-550 nm. Calibration curves were prepared and individual phenolic compounds quantified as described in Chapter 3 (section 3.3.5).

4.3.3 Pre-treatment of chromatograms

The data handling and pre-treatment process is illustrated in Fig. 4.1. The recorded absorbance matrices for the duplicate injections of each extract (wavelengths \times time) of *C. subternata* (n = 130) and *C. maculata* (n = 102) were exported from ChemStation Software (Agilent) as *.csv files for the wavelength range 220 to 500 nm

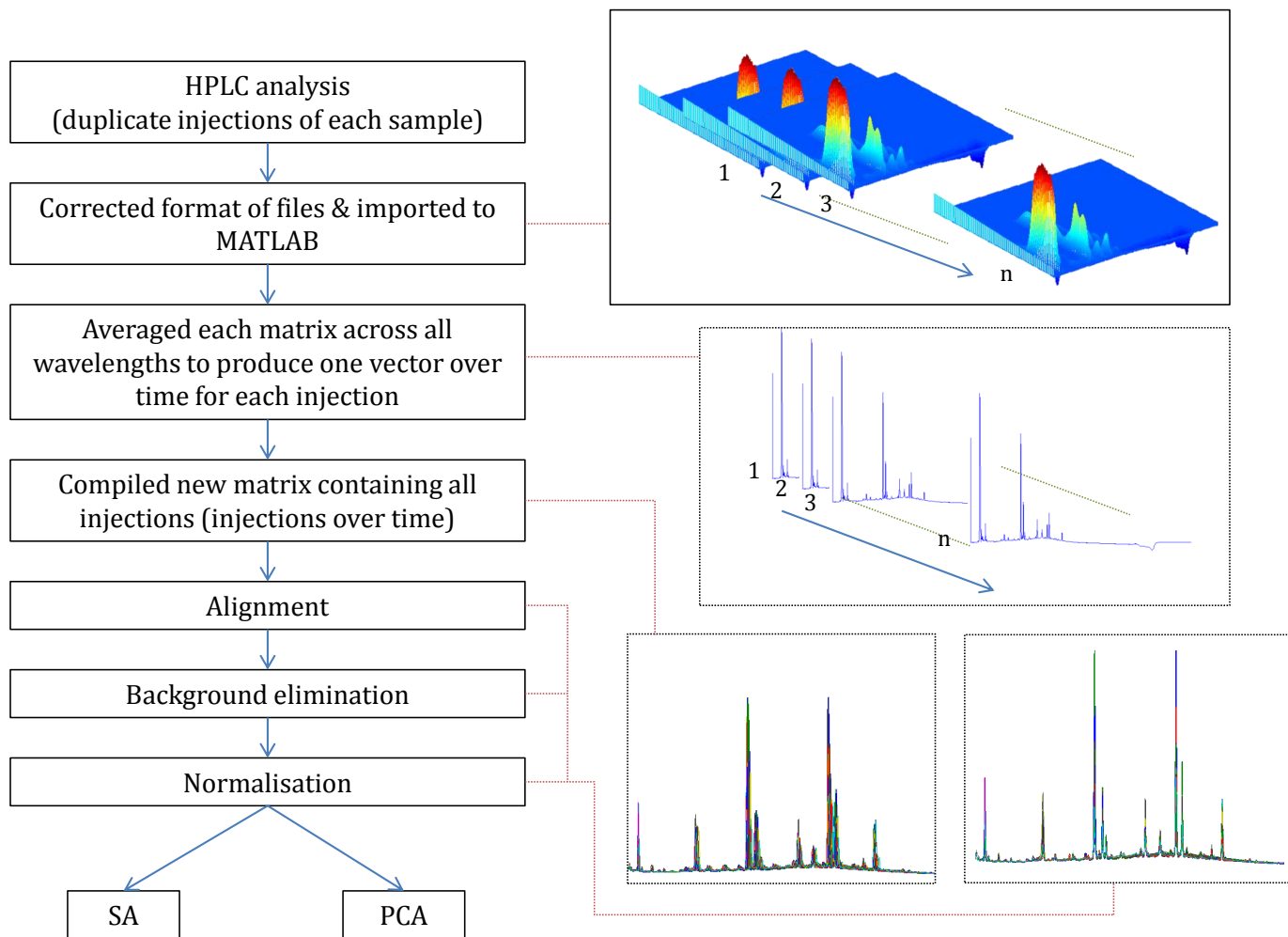


Figure 4.1 – Schematic representation of data handling and pre-treatment [SA = similarity analysis; PCA = principal component analysis].

(2 nm bandwidth) and 0 to 40 min using a ChemStation macro, export3D. The *.csv files were each changed to comma delimited format in Microsoft Excel (2010). Thereafter, the adapted *.csv files were imported into MATLAB R2012b (The MathWorks, Natick, MA, USA). All matrices contained 12000 time (data) points and were averaged to individual absorbance vectors (absorbance \times time). These new vectors were arranged into a new matrix, one for each species, with dimensions of 130×12000 and 102×12000 , respectively (duplicate extracts \times time).

Peak alignment was conducted by means of correlation optimised warping (COW) (Tomasi *et al.*, 2004) using MATLAB code freely available from www.models.life.ku.dk/DTW_COW, followed by background elimination by means of the penalised asymmetric least squares (PALS) method (Eilers, 2004). Normalisation (Hendriks *et al.*, 2005) to a constant sum was then performed. This was achieved by dividing each signal value for each sample by the sum of all the signal values for that specific sample. The time axes of the matrices were shortened to remove irrelevant data. The final dimensions of the new matrices for *C. subternata* were 130×5370 and 102×5381 for *C. maculata*.

4.3.4 Similarity analysis

Similarity analysis of the extracts of each species was conducted by determining the correlation coefficients between fingerprints of the extracts and the active extract. The similarity analysis conducted in this study was based on the Pearson product-moment correlation coefficient, r (Alaerts *et al.*, 2012).

$$r(\mathbf{x}_1, \mathbf{x}_2) = \frac{\text{cov}(\mathbf{x}_1, \mathbf{x}_2)}{s_{x1}s_{x2}} = \frac{\sum_{j=1}^q (x_{1j} - \bar{x}_1)(x_{2j} - \bar{x}_2)}{\sqrt{\sum_{j=1}^q (x_{1j} - \bar{x}_1)^2 \sum_{j=1}^q (x_{2j} - \bar{x}_2)^2}} = \frac{(\mathbf{x}_1 - \bar{x}_1)(\mathbf{x}_2 - \bar{x}_2)}{|\mathbf{x}_1 - \bar{x}_1| |\mathbf{x}_2 - \bar{x}_2|} \quad (4.3.1)$$

This coefficient was calculated between each pair of fingerprints, \mathbf{x}_i , where

$i = 1, 2, \dots, p$, samples;

each fingerprint was composed of measurements at $j = 1, 2, \dots, q$, time points;

\mathbf{x}_1 and \mathbf{x}_2 being the fingerprints or vectors that were considered;

x_{1j} and x_{2j} being the absorbance that was measured for each fingerprint at the j th time point;

\bar{x}_1 and \bar{x}_2 the means of the absorbance of each fingerprint;

cov the covariance of the fingerprints; and

s_{xi} the standard deviation.

4.3.5 Principal component analysis

PCA was conducted on the same data as for similarity analysis to determine whether any relevant information could be extracted from the data.

$$\mathbf{X} = \mathbf{TP}^t + \mathbf{E} \quad (4.3.2)$$

The matrix of data was decomposed into a product of two matrices; where

\mathbf{T} is the score matrix;

\mathbf{P} is the loading matrix; \mathbf{P}^t is the transpose of \mathbf{P} ; and

\mathbf{E} is the matrix of residuals.

4.4 Results and discussion

4.4.1 Pre-treatment

The chromatograms of the *C. subternata* and *C. maculata* data sets before any pre-treatment are shown in Fig. 4.2. From these figures it is clear that pre-treatment of the raw chromatograms was essential to align peaks due to extreme shifts in peak retention times. These shifts would have been due to column aging and general experimental variation (Alaerts *et al.*, 2010). Retention time shifts were larger for the *C. maculata* data set than the *C. subternata* data set.

Pre-treatment firstly consisted of peak alignment which was achieved by means of correlation optimised warping (COW). COW was the alignment method of choice, as it is an efficient and popular method among chemometricians (Tistaert *et al.*, 2011; Alaerts *et al.*, 2012; Lucio-Gutiérrez *et al.*, 2012) and freely available. COW consists of piecewise linear alignment of chromatograms towards a target chromatogram by means of stretching and compressing of segments using dynamic programming (Tomasi *et al.*, 2004). The target chromatogram was selected based on having the highest correlation with all the chromatograms in the sample set (Daszykowski & Walczak, 2007). Warping entails a trial-and-error based selection of the best input parameters to obtain the best alignment towards a target chromatogram (Tomasi *et al.*, 2004). For the alignment of these data sets, it was determined that the optimal number of segments were between 60 and 100, while the range for segment length was optimal between 20 and 30 data points.

After peak alignment, the background was removed from the data sets, in order to further improve the relevance of the data for analysis (Eilers, 2004). The lambda (λ) parameter was set to 10^4 for smoothing (usually between 10^2 and 10^9), as it was deemed optimal after visual inspection of the results (Eilers & Boelens, 2005). The first derivative of the signals was used to determine the signal smoothness. To account for differing HPLC injection volumes due to small inevitable instrumental variations and experimental weighing errors during sample preparation, normalisation was applied to each vector (injection) to scale them to the same unit. Fig. 4.3 shows the chromatograms after acceptable pre-treatment was achieved.

It should be noted that these chromatograms were shortened after pre-treatment of the data, by removing the elution time representing the dead volume and the ascorbic acid peak from the start, and the time required for the re-equilibration of the column from the end of the chromatogram. The removal of these chromatogram sections were to exclude all information that would not represent relevant data. Consequently, the dimensions of the new matrices were 130×5370 and 102×5381 for *C. subternata* and *C. maculata*, respectively.

4.4.2 Similarity analysis

The similarity of the fingerprints of the sample extracts compared to that of the active extract was evaluated by calculating r for each extract against the active extract. The closer r is to 1, the higher the correlation and the more similar the two fingerprints. It is a generally accepted concept that if two fingerprints are similar, the samples are assumed to have the same properties (such as biological activity) (Fan *et al.*, 2006; Alaerts *et al.*, 2010).

The correlation coefficients calculated can be visualised by means of a heat or colour map for easier interpretation. These heat maps are usually plotted as $p \times p$ (sample \times sample) matrices, as indicated in Fig. 4.4. The heat map is interpreted by means of the colour bar on the right of each image, which assigns colour to, in this case, the correlation value. Dark red represents good correlation and thus very similar fingerprints,

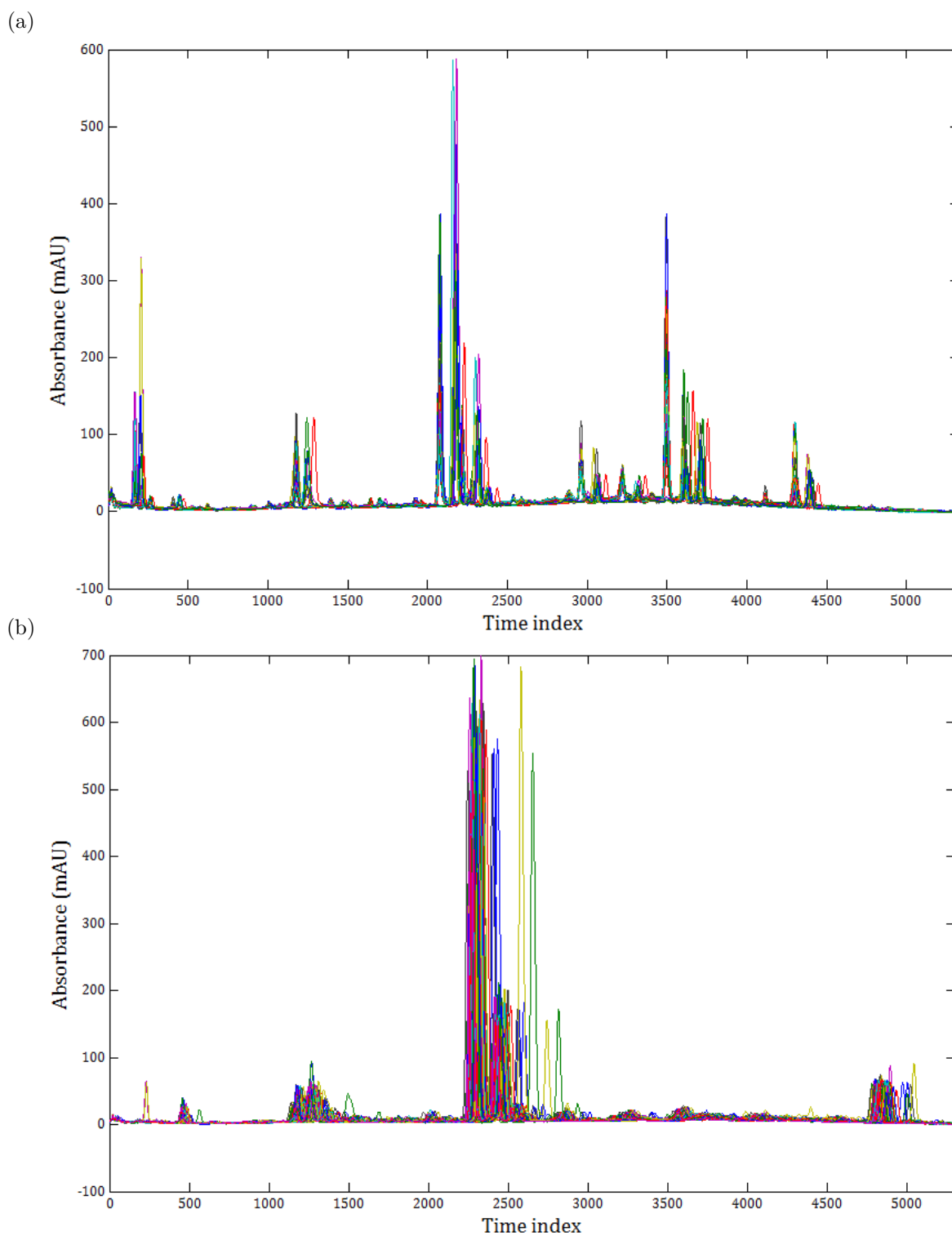


Figure 4.2 – (a) *C. subternata* and (b) *C. maculata* chromatograms before pre-treatment of data.

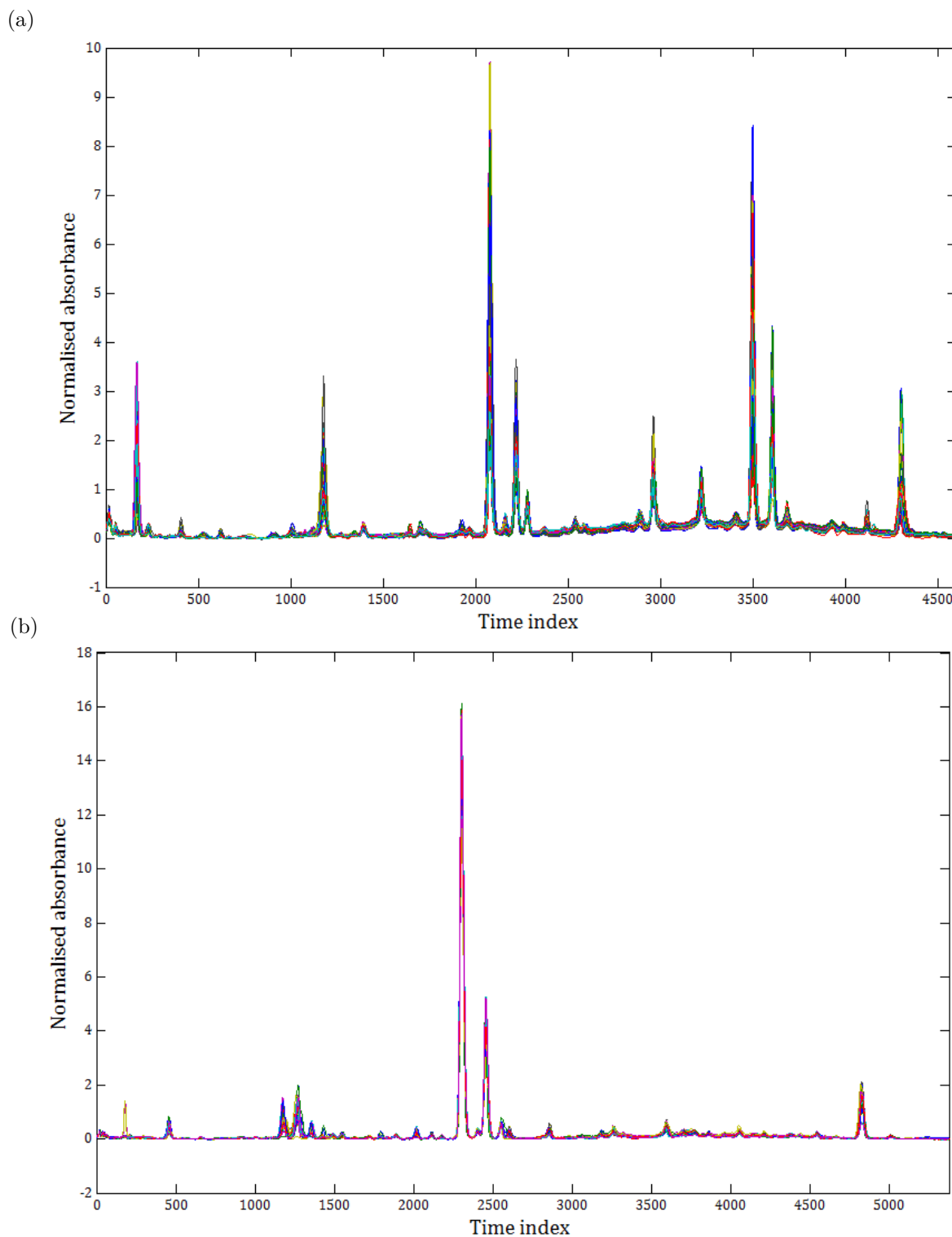


Figure 4.3 – (a) *C. subternata* and (b) *C. maculata* chromatograms after pre-treatment of data.

while blue and green represents lower correlation, indicating lower similarity. By pre-treating the data set, an improvement in correlation coefficients could be observed, as indicated in Figs. 4.4 and 4.5, representing the heat maps before and after pre-treatment for the two species, respectively. It appears as though there are less samples being compared in the pre-treated data, however, this is due to duplicate extracts that are now correctly shown as more similar. It also indicates that pre-treatment was effective in aligning the peaks and removing noise, as fingerprints of duplicate injections are expected to have an extremely high correlation.

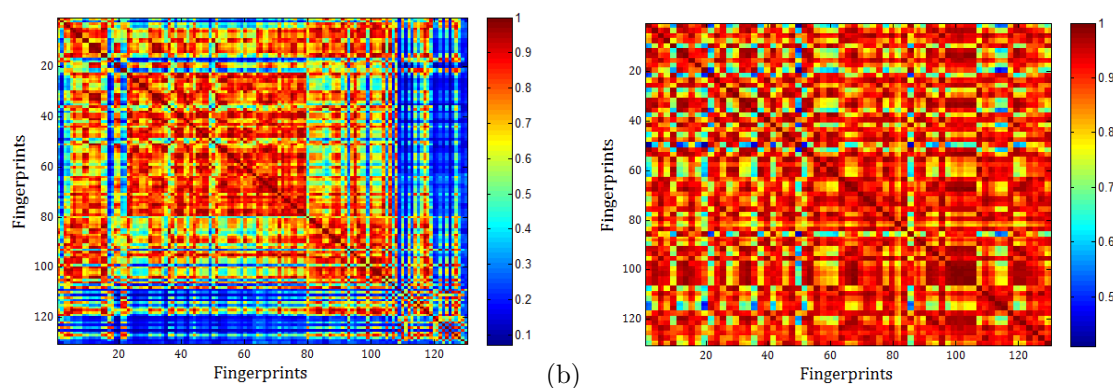


Figure 4.4 – Heat maps of correlation coefficients between samples of *C. subternata* (a) before and (b) after fingerprint data pre-treatment (duplicate injections are shown as separate measurements).

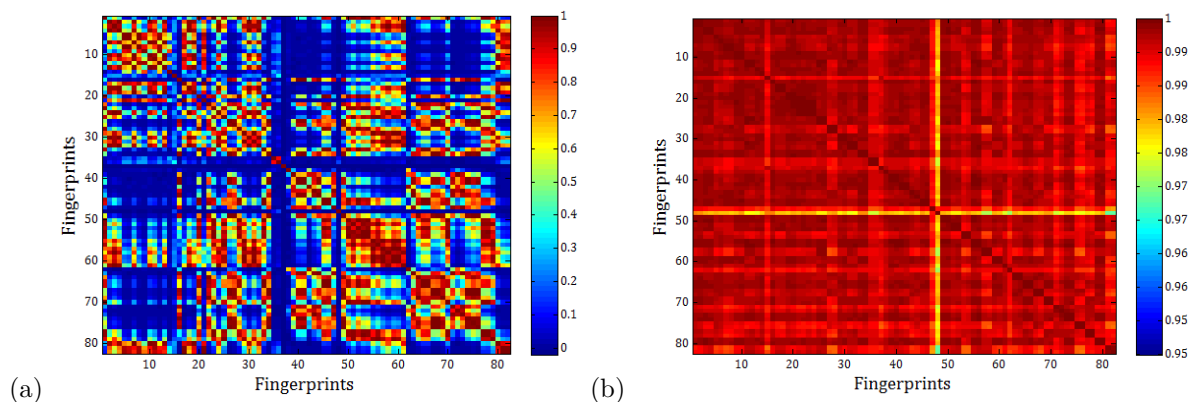


Figure 4.5 – Heat maps of correlation coefficients between samples of *C. maculata* (a) before and (b) after fingerprint data pre-treatment (duplicate injections are shown as separate measurements).

For the purpose of this study, the focus will only be on the correlation of the chromatograms of the randomly selected extracts to those of the active extracts of the two species. Fig. 4.6 thus depicts the correlations in a much smaller heat map, representing only the correlations between the *C. subternata* extracts and the duplicates of the *C. subternata* active extract. A large variation in the quantified phenolic compounds was observed for *C. subternata* extracts, as reported in Chapter 3 (section 3.4.4). Therefore, an even larger variation is expected between extracts when examining the entire fingerprint. Correlation coefficients varied from *ca.* 0.65 to 1.00. The threshold was selected as 0.95 and duplicate fingerprints that both presented an average r value greater than 0.95 against the duplicate fingerprints of the active extract were considered to have similar profiles to the latter. For *C. subternata*, $n = 15$ extracts were found to be similar to the active extract, based on r greater than 0.95. The assumption is thus made that these extracts could be biologically active, and merit further investigation using cell-based assays.

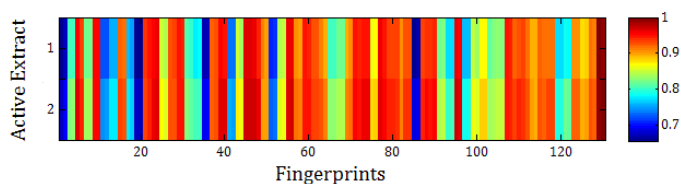


Figure 4.6 – Heat map of correlation coefficients of *C. subternata* extracts against the bioactive *C. subternata* extract (duplicate injections are shown as separate measurements).

To validate whether the fingerprints that showed a high correlation ($r > 0.95$) were similar to the active extract, the chromatogram of a randomly selected extract from the *C. subternata* data set with $r > 0.95$ was plotted against that of the bioactive extract for visual inspection (Fig. 4.7). Similarly, an extract that did not correlate well ($r < 0.70$) was also randomly selected from the data set and its chromatogram plotted against that of the active extract. It is clearly noted that the highly correlated extract displayed a more similar profile compared to the extract that correlated poorly to the bioactive extract.

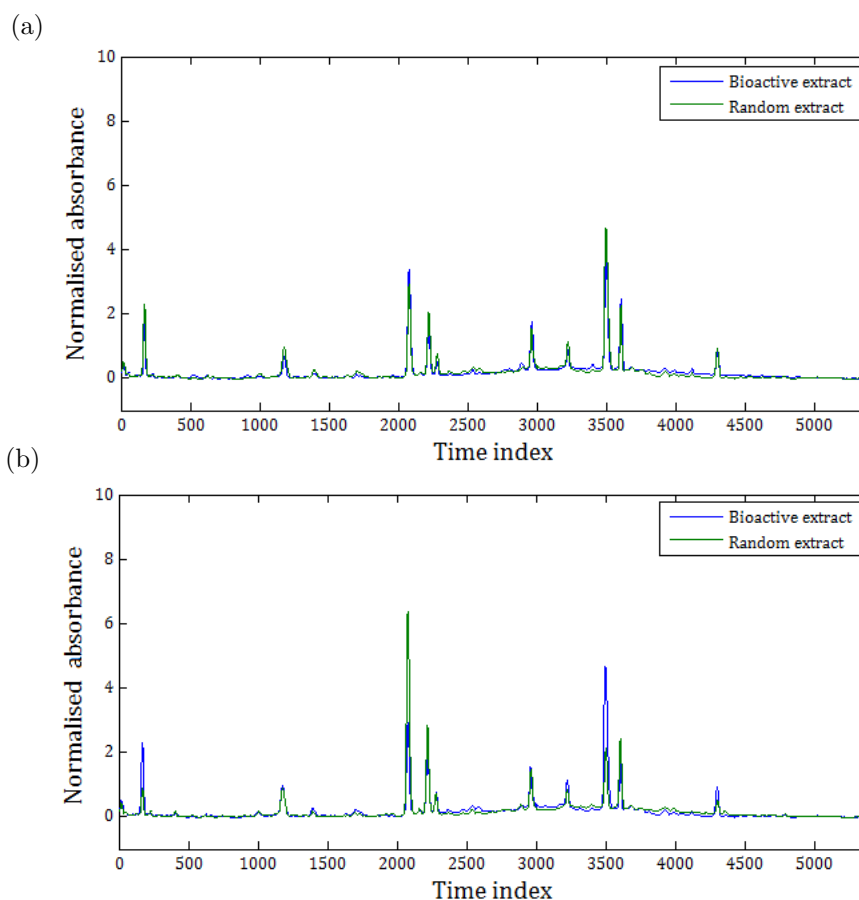


Figure 4.7 – Comparison of chromatograms of an extract that correlated (a) well ($r > 0.95$) and an extract that correlated (b) poorly ($r < 0.70$) with the chromatogram of the active *C. subternata* extract.

For the first set of the *C. maculata* extracts and the *C. maculata* active extract ($n = 41$), more similar fingerprints were observed than was the case for the *C. subternata* extracts. The r values for all the sample extracts against the active extract were extremely good, ranging from *ca.* 0.97 to 1.00 (Fig. 4.8). This might be due to the less complex fingerprint of this specific *Cyclopia* species. The chromatogram comprised of only two large peaks, three smaller peaks, and many minor peaks distributed over the rest of the fingerprint. *Cyclopia maculata* also did not display as large a variation as *C. subternata* in terms of its quantitative data (fewer peaks and lower

%RSDs, Chapter 3, section 3.4.4). When using the same threshold as for *C. subternata*, these data indicate that all the samples compare well with the extract and could probably be biologically active. The threshold could alternatively be increased for *C. maculata* to $r = 0.99$, to be more selective. With the higher threshold, $n = 35$ (out of 40) were indicated as active. The selected extracts should however be tested for activity in cell and/or animal models to confirm the observations.

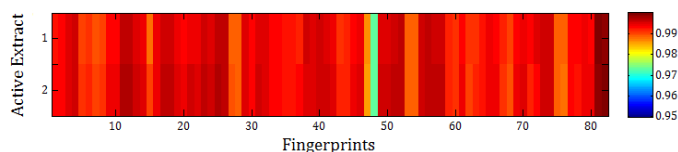


Figure 4.8 – Heat map of correlation coefficients of *C. maculata* extracts against a bioactive *C. maculata* extract (duplicate injections for each sample are included) ($n = 82$; set 1).

To confirm these extreme similarities, an additional set of *C. maculata* seedling samples ($n = 10$; set 2) were investigated. Although the plants originated from the same wild population which provided the seeds for set 1 of *C. maculata*, the plants of set 2 grew in marshy conditions, while irrigation was controlled during the pot trial used for set 1 samples. These extracts were analysed and the fingerprints pre-treated in the same way as sample set 1 of *C. maculata*. Similarity analysis was conducted on the combined data set, and a new heat map generated (Fig. 4.9). The correlations between the extracts of set 2 and set 1 combined and the active extract once again ranged from 0.97 to 1.0. These high correlations between different samples of this species indicate that there is less variation between the different seedling plants than observed for *C. subternata*. When using the higher threshold ($r > 0.99$), all the extracts from wild-harvested samples ($n = 10$) were similar to the active extract, resulting in $n = 45$ extracts in the combined sample set that were similar to the active extract. Extracts from seedlings of the same age grown under standardised conditions were very similar to extracts from the wild-harvested plants of indeterminate, and probably different ages, although a larger variation was expected due to quantitative differences between these sets of extracts (Appendix A, Tables A.4 and A.5).

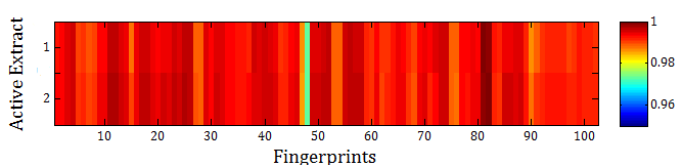


Figure 4.9 – Heat map of correlation coefficients of additional wild harvested *C. maculata* extracts together with the previous extracts against the bioactive *C. maculata* extract (duplicate injections are shown as separate measurements) (sets 1 and 2; $n = 102$).

4.4.3 Principal component analysis

PCA was conducted on the data sets of the two *Cyclopia* species after mean centering of the fingerprints. PCA is an unsupervised multivariate data analysis technique that reduces the dimensions of the data and recognises linear patterns using the most significant information, or the largest possible variances. The transformed data can then be plotted as score plots to investigate the relationships between samples, and as loading plots to interpret the influence of the variables (chemical composition). Biplots are used to display the scores and

loadings simultaneously, allowing for an easier interpretation of the data (Wold *et al.*, 1987; Abdi & Williams, 2010).

Cyclopia subternata

The PCA score values in the score plots were labelled according to their r values. PCA allowed for good separation between samples with low and high r values on the scores plots (Fig. 4.10). PC1 divided the samples at approximately $r = 0.90$, with $r > 0.90$ having high negative scores. According to the loadings of PC1 (Fig. 4.11) and PC2 (Fig. 4.12), the top right quadrant is associated with high mangiferin and isomangiferin contents, while the left top quadrant is associated with high iriflophenone-3-*C*-glucoside, iriflophenone-di-*O,C*-glucoside and scolymoside contents. The bottom left quadrant is represented by positive loadings for phloretin-3',5'-di-*C*-glucoside, eriocitrin and hesperidin contents. Therefore it can be seen that variation in the principal components are dominated by the large peaks present in the chromatogram which represents the majority of the variation. From the loadings plot of PC1, it is evident that the active sample and the high correlating samples are clustered together, because of the influence of the compounds iriflophenone-di-*O,C*-hexoside, iriflophenone-3-*C*-glucoside, scolymoside and phloretin-3',5'-di-*C*-glucoside, as indicated in the negative direction of PC1 (Fig. 4.11). Samples that did not correlate well were associated with mangiferin and isomangiferin, in the positive direction of PC1. This does not necessarily suggest that mangiferin and isomangiferin are not related to the activity, but rather that the remaining peaks have a larger influence. When examining further PCs up to PC5, no better clustering was observed for the high correlating samples. Based on the chromatographic fingerprints, PCA and similarity analysis delivered similar results in terms of high correlating samples also clustering together in PCA. This data distribution for complete chromatographic fingerprints coincide with that of the quantitative data (Fig. 4.13), which also indicates that the contribution of the smaller peaks were suppressed by those of larger peaks. This is to be expected as the larger peaks do represent the highest variability extracted by PCA (Wold *et al.*, 1987).

Cyclopia maculata

The similarity analysis of *C. maculata* samples resulted in all samples of the two sets ($n = 100$; duplicates) correlating well ($r > 0.97$) with the active extract. Although the variation obtained for *C. maculata* (set 1 and 2) was less than for *C. subternata* when quantifying selected peaks (Chapter 3, section 3.4.4), these high r values were investigated further.

PCA was conducted on the combined data set ($n = 102$; duplicates). Based on the first two PCs of the *C. maculata* fingerprints, the wild-harvested samples ($n = 10$; nrs 83-102) were clearly separated from the rest of the samples in the bottom right quadrant (Fig. 4.14). As indicated by the loadings (Figs. 4.15 and 4.16), this clustering of the wild-harvested samples is greatly associated with mangiferin and isomangiferin, while the samples from set 1 are associated with the remaining compounds. Quantitative data (Appendix A, Tables A.4 and A.5) suggests that the wild-harvested samples only contain approximately half of the hesperidin, maclurin-3-*C*-glucoside and iriflophenone-3-*C*-glucoside amounts present in extracts of set 1 samples. Therefore, this could be the reason for the separation from the rest of the samples in the scores plot of PC1 vs. PC2. This is supported by the biplot of the quantitative data (Fig. 4.17, nrs 42-51). The lower hesperidin content is attributed to the sieving process described in section 4.3.1, allowing for most of the stems to be removed from the samples. Analysis of *C. subternata* showed that hesperidin is mostly concentrated in the stems (Appendix

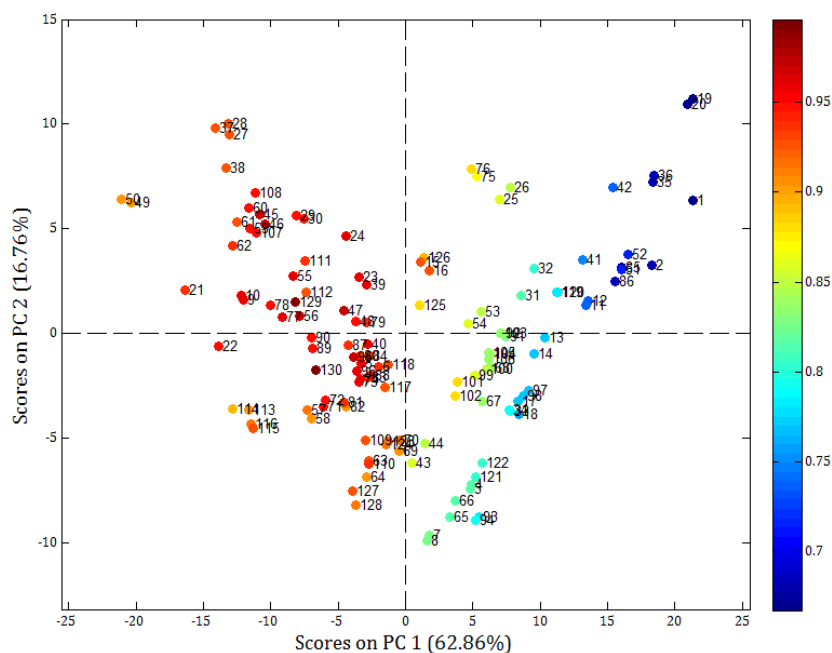


Figure 4.10 – PCA score plot of PC1 vs. PC2 for fingerprints of *C. subternata* samples with r values indicated by colour bar ($n = 130$; duplicates) [129 and 130 = duplicates of active extract].

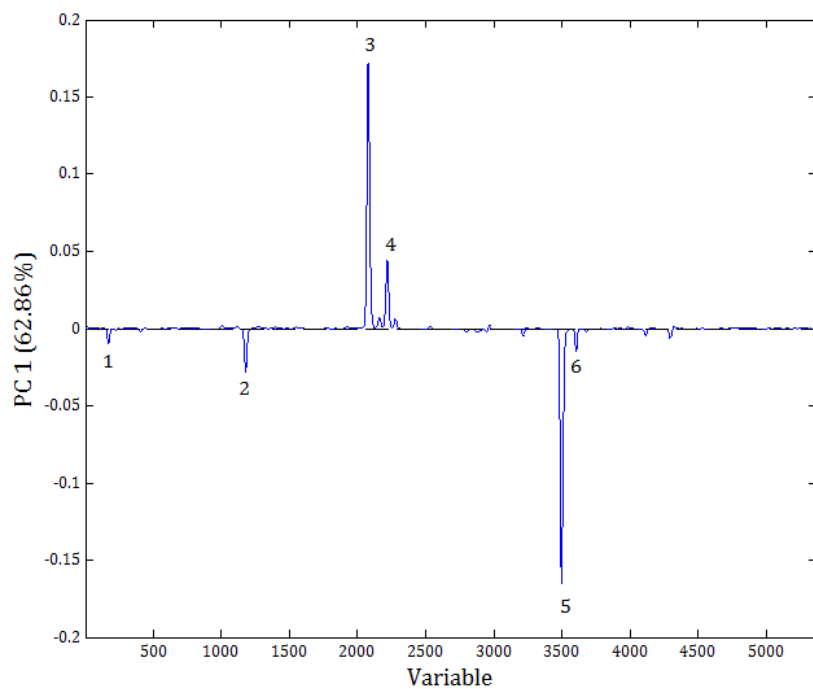


Figure 4.11 – Loadings of PC1 for fingerprints of *C. subternata* samples ($n = 130$; duplicates) [1 = iriflophenone-di-*O,C*-hexoside; 2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = scolymoside; 6 = phloretin-3',5'-di-*C*-glucoside].

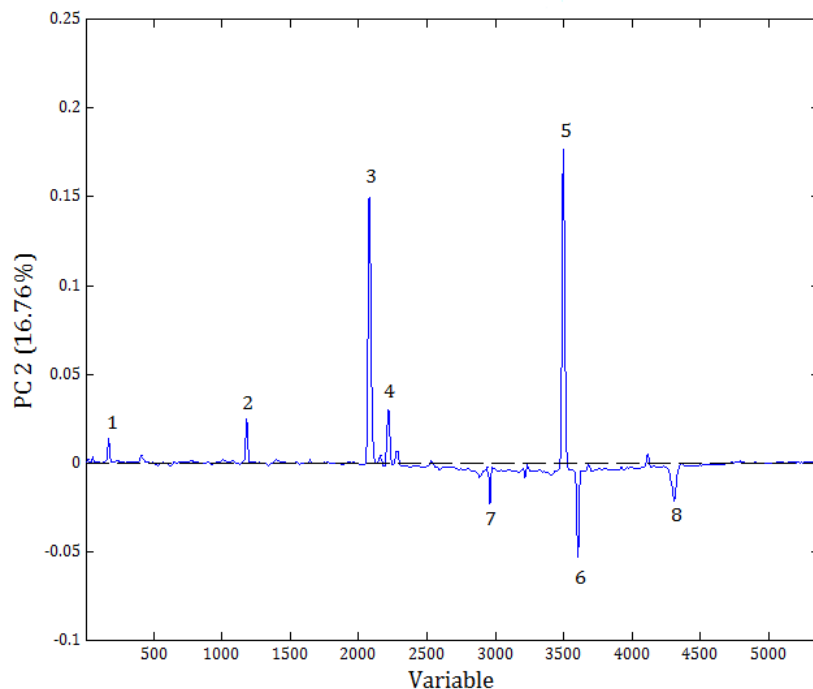


Figure 4.12 – Loadings of PC2 for fingerprints of *C. subternata* samples (n = 130; duplicates) [1 = iriflophenone-di-*O,C*-hexoside; 2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = scolymoside; 6 = phloretin-3',5'-di-*C*-glucoside; 7 = eriocitrin; 8 = hesperidin].

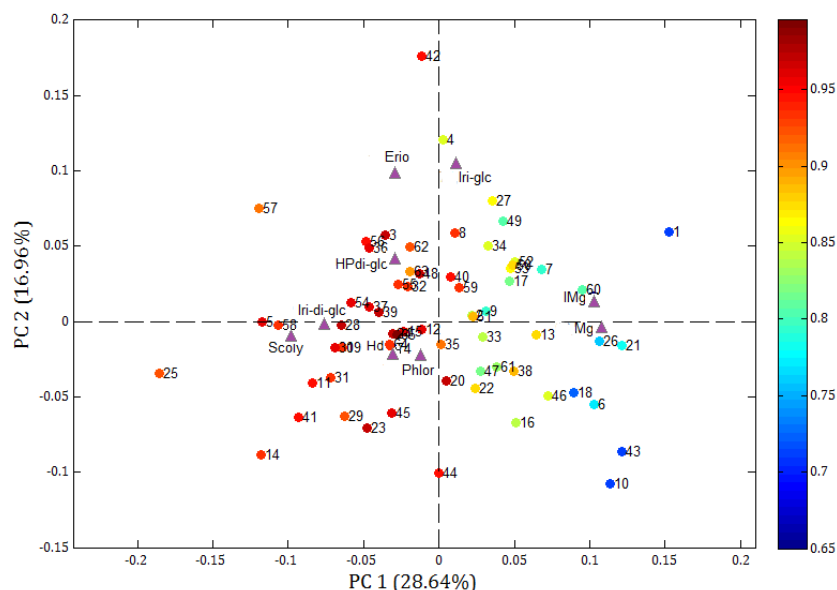


Figure 4.13 – PCA biplot of PC1 vs. PC2 scores and loadings of the quantitative data for *C. subternata* samples with *r* values indicated by colour bar (n = 65; no duplicates) [65 = active extract; Mg = mangiferin; IMg = isomangiferin; Iri-glc = iriflophenone-3-*C*-glucoside; Iri-di-glc = iriflophenone-di-*O,C*-hexoside; Phlor = phloretin-3',5'-di-*C*-glucoside; Erio = eriocitrin; scol = scolymoside; HPdi-glc = 3-hydroxyphloretin-3',5'-di-*C*-hexoside; Hd = hesperidin].

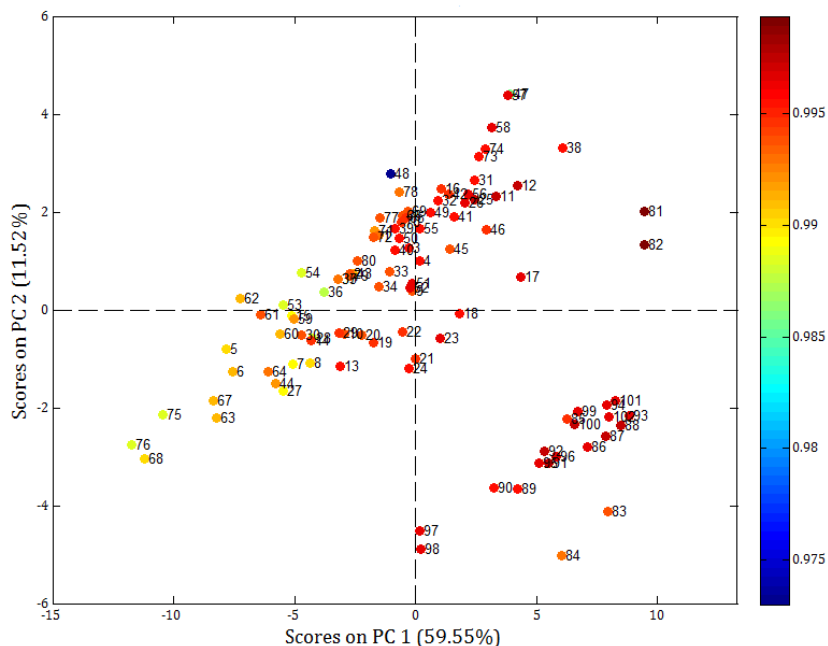


Figure 4.14 – PCA score plot of PC1 vs. PC2 for fingerprints of *C. maculata* samples (n = 102) [81 and 82 = duplicates of active extract].

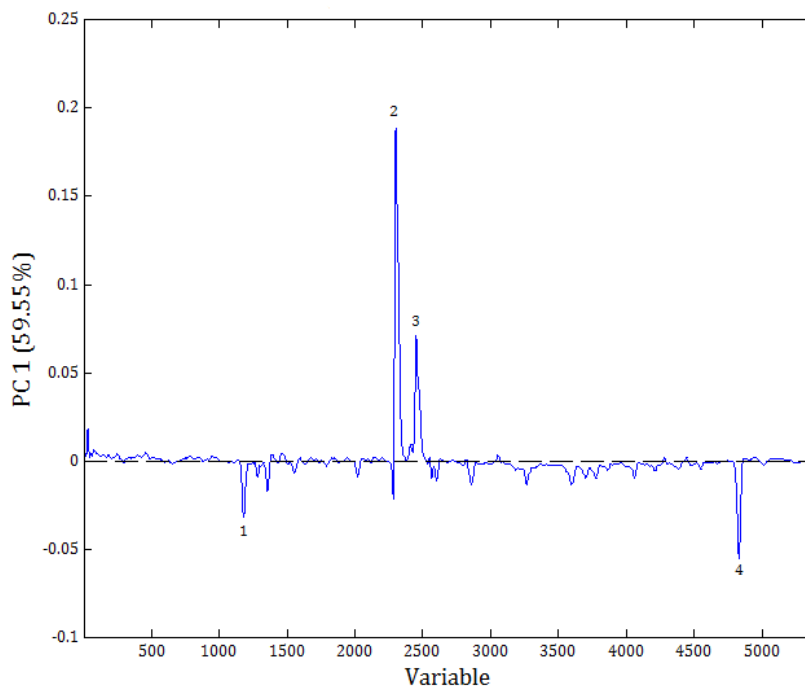


Figure 4.15 – Loadings of PC1 for fingerprints of *C. maculata* samples (n = 102) [1 = iriflophenone-3-*C*-glucoside; 2 = mangiferin; 3 = isomangiferin; 4 = hesperidin].

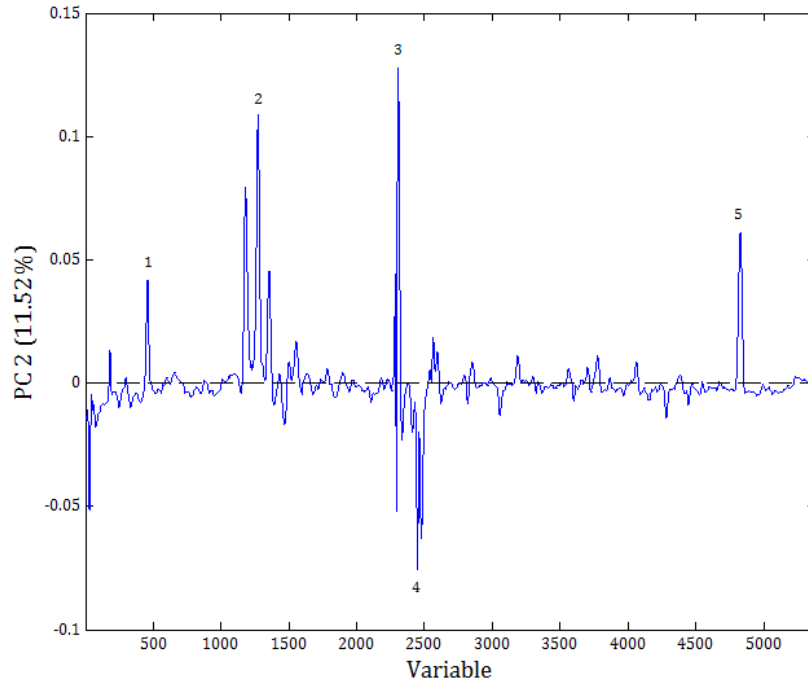


Figure 4.16 – Loadings of PC2 for fingerprints of *C. maculata* samples (n = 102) [1 = maclurin-3-*C*-glucoside; 2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = hesperidin].

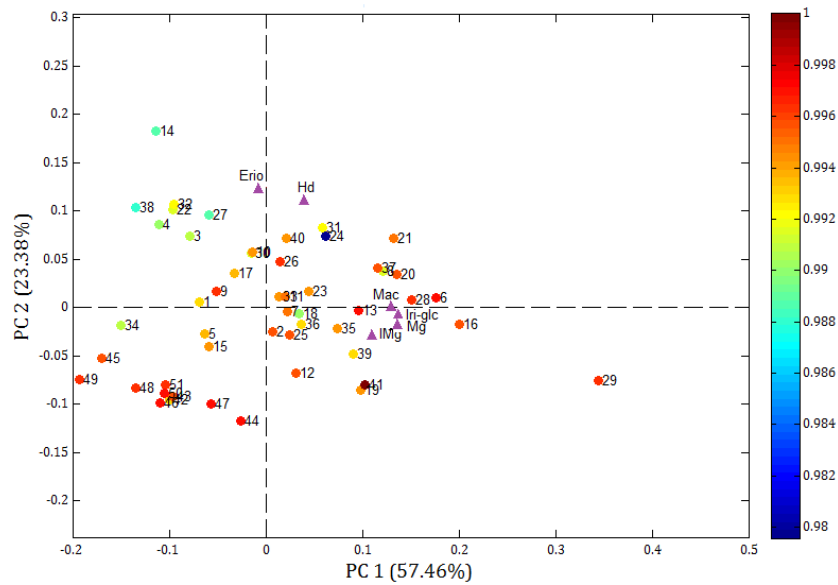


Figure 4.17 – PCA biplot containing PC1 vs. PC2 scores and loadings for the quantitative data of *C. maculata* (n = 51; no duplicates) [41 = active extract; Mg = mangiferin; IMg = isomangiferin; Iri-glc = iriflophenone-3-*C*-glucoside; Eriol = eriocitrin; Hd = hesperidin; Mac = maclurin-3-*C*-glucoside].

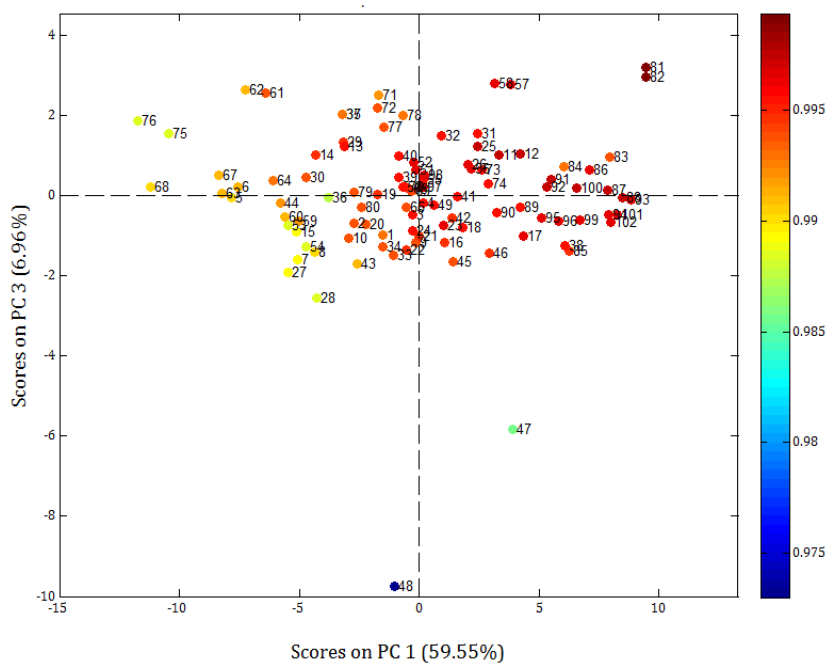


Figure 4.18 – PCA score plots of PC1 vs. PC3 for fingerprints of *C. maculata* samples (n = 102) [81 and 82 = duplicates of active extract].

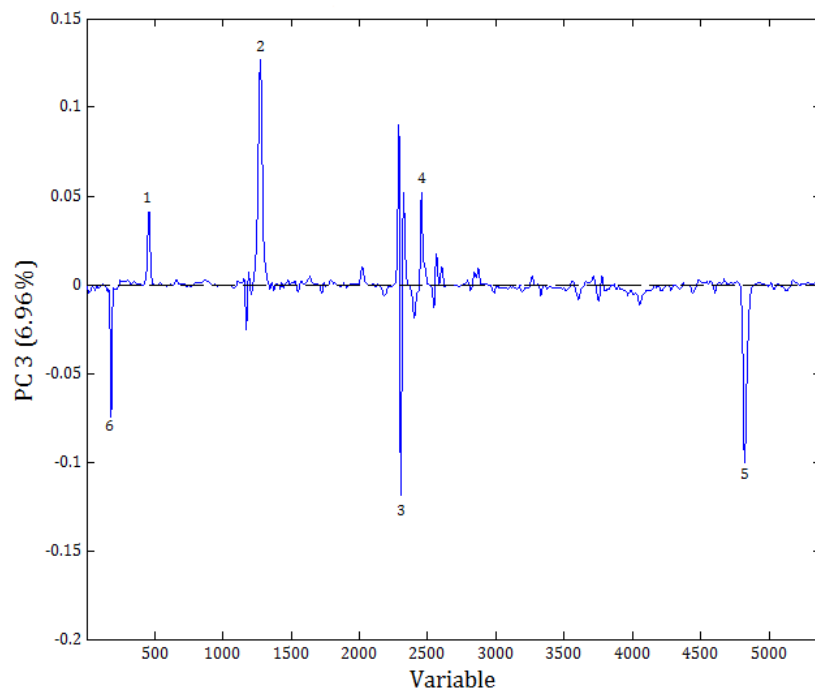


Figure 4.19 – Loading plot of PC3 for fingerprints of *C. maculata* samples (n = 102; duplicates) [1 = maclurin-3-*C*-glucoside; 2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = hesperidin; 6 = extra peak].

A, Tables A.2 and A.3). However, this does not explain the lower iriflophenone-3-*C*-glucoside and maclurin-3-*C*-glucoside concentrations, as leaves and stems contained similar concentrations of the latter compounds. The fingerprint data were more useful in separating the two sets of data for *C. maculata* analysed by PCA. This distinction was not indicated by similarity analysis.

When considering PC3 of the fingerprint data ($n = 102$) of the *C. maculata* extracts, more information was revealed and sample duplicates nrs 47 and 48 were separated from the other samples (Fig. 4.18). On further investigation (Fig. 4.19 and Appendix B), it was determined that this sample has a high concentration of an extra peak that none of the other samples contained. This important characteristic was not detected by the similarity analysis. According to quantitative data, this sample also had the highest quantity of hesperidin. Although PC3 separates the duplicate fingerprints of the ‘atypical’ sample from the rest of the samples, the r values seemed to still only be separated in the direction of PC1, and therefore due to mangiferin and isomangiferin. Another possible interpretation is that they were separated due to hesperidin content. PCs up to PC5 did not reveal any further information of interest. The threshold may have to be increased to $r > 0.99$ for *C. maculata*, considering that similarity analysis with a threshold of $r > 0.99$ eliminated such a sample with an extra peak. It is therefore assumed that the samples with $r > 0.99$ that clustered with the active sample have similar activity. These samples have to be tested in a similar manner as the bioactive sample to confirm their activity.

4.5 Conclusions

Similarity analysis was found to be a more effective tool when combining the data analysis with PCA. It was determined that for both multivariate data analysis techniques, the larger peaks dominated the results. For *C. subternata* $n = 15$ samples had $r > 0.95$. According to PCA it seems as though mangiferin and isomangiferin do not contribute greatly to the activity of the extract and that the remaining compounds have a greater association with the activity. For *C. maculata*, the threshold had to be increased to $r > 0.99$, as all the samples had $r > 0.95$. PCA was proven an essential aid to the similarity analysis process as it can highlight small variations in the fingerprints. Due to the lower complexity of *C. maculata* chromatograms, it seemed as though mangiferin and isomangiferin might have a large association with the activity, or that hesperidin may be the opposite driver for separation in PCA score plots. In both cases, samples with r values above the threshold that clustered near the active extract in PCA are assumed to have the same activity and should be tested in cell and/or animal models to confirm this hypothesis. If these extracts do encompass the same activity, similarity analysis in combination with PCA can be employed as a method to select future extracts for further biological screening. In this manner, the number of samples to be screened in cell and animal models can be substantially reduced. This screening method based on fingerprint analysis could therefore be applied in quality control, for example when these extracts are used for nutraceutical products. This would also aid in identifying which plants should be propagated vegetatively in order to produce more plant material that also encompasses the sought-after biological activity. For future work, it is recommended that more bioactive samples be obtained to have a better representation of the bioactive fingerprint. This will aid in the detection and verification process of the compounds responsible for the activity, as results obtained thus far are somewhat contradictory in terms of mangiferin and isomangiferin. With more active samples, other multivariate data analysis techniques, such as partial least squares discriminant analysis (PLS-DA), could then be employed to investigate the activity further.

References

- Abdi, H & Williams, L.J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**, 1-40.
- Anonymous (1991). Guidelines for the assessment of herbal medicines, World Health Organisation, Geneva, Switzerland. http://whqlibdoc.who.int/hq/1991/WHO_TRM_91.4.pdf. (February 2013).
- Anonymous (2002). Traditional Medicine Strategy 2002-2005, World Health Organisation, Geneva, Switzerland. http://whqlibdoc.who.int/hq/2002/who_edm_trm_2002.1.pdf. (February 2013).
- Anonymous (2007). Corrigendum to Regulation (EC) No 1924/2006 of the European Parliament and of the council of 20 December 2006 on nutrition and health claims made on foods. *Official Journal of the European Union*, Article 13.
- Alaerts, G., Dejaegher, B., Smeyers-Berbeke, J. & Vander Heyden, Y. (2010). Recent developments in chromatographic fingerprints from herbal products: Set-up and data analysis. *Combinatorial Chemistry and High Throughput Screening*, **13**, 900-922.
- Alaerts, G., van Erps, J., Pieters, S., Dumarey, M., van Nederkassel, A.M., Goodarzi, M., Smeyers-Verbeke, J. & Vander Heyden, Y. (2012). Similarity analyses of chromatographic fingerprints as tools for identification and quality control of green tea. *Journal of Chromatography B*, **910**, 61-70.
- Daszykowski, M. & Walczak, B. (2007). Target selection for alignment of chromatographic signals obtained using monochannel detectors. *Journal of Chromatography A*, **1176**, 1-11.
- Dudhia, Z., Louw, J., Muller, C., Joubert, E., de Beer, D., Kinnear, C. & Pfeiffer, C. (2013). *Cyclopia maculata* and *Cyclopia subternata* (honeybush tea) inhibits adipogenesis in 3T3-L1 pre-adipocytes. *Phytomedicine*, **20**, 401-408.
- Eilers, P.H.C. (2004). Parametric time warping. *Analytical Chemistry*, **76**, 404-411.
- Eilers, P.H.C. & Boelens, H.F.M. (2005). Baseline correction with asymmetric least squares smoothing. [WWW document]. URL http://www.science.uva.nl/~hboelens/publications/draftpub/Eilers_2005.pdf. (March 2013).
- Fan, X., Cheng, Y., Ye, Z., Lin, R. & Qian, Z. (2006). Multiple chromatographic fingerprinting and its application to the quality control of herbal medicines. *Analytica Chimica Acta*, **555**, 217-224.
- Gad, H.A., El-Ahmady, S.H., Abou-Shoer, M.I. & Al-Azizi, M.M. (2013). Application of chemometrics in authentication of herbal medicines: A review. *Phytochemical Analysis*, **24**, 1-24.
- Gertsch, J. (2011). Botanical drugs, synergy, and network pharmacology: Forth and back to intelligent mixtures. *Planta Medica*, **77**, 1086-1098.
- Hendriks, M.M.W.B., Cruz-Juarez, L., De Bont, D. & Hall, R.D. (2005). Preprocessing and exploratory analysis of chromatographic profiles of plant extracts. *Analytica Chimica Acta*, **545**, 53-64.
- Kingston, D.G.I. (2011). Modern natural products drug discovery and its relevance to biodiversity conservation, *Journal of Natural Products*, **74**, 496-511.
- Liang, Q., Qu, K., Luo, G. & Wang, Y. (2006). Rapid and reliable determination of illegal adulterant in herbal medicines and dietary supplements by LC/MS/MS. *Journal of Pharmaceutical and Biomedical Analysis*, **40**, 305-311.
- Liang, X., Jin, Y., Wang, Y. Jin, G., Fu, Q. & Xiao, Y. (2009). Qualitative and quantitative analysis in quality control of traditional Chinese medicines. *Journal of Chromatography A*, **1216**, 2033-2044.

- Lucio-Gutiérrez, J.R., Garza-Juárez, A., Coello, J., Maspoch, S., Salazar-Cavazos, M.L., Salazar-Aranda, R. & Waksman de Torres, N. (2012). Multi-wavelength high-performance liquid chromatographic fingerprints and chemometrics to predict the antioxidant activity of *Turnera diffusa* as part of its quality control. *Journal of Chromatography A*, **1235**, 68-76.
- Mander, L. & Liu, H.W. (2010). *Comprehensive Natural Products II: Chemistry and Biology*. Amsterdam: Elsevier.
- Mishra, K.P., Ganju, L., Sairam, M., Banerjee, P.K. & Sawhney, R.C. (2008). A review of high throughput technology for screening of natural products. *Biomedicine and Pharmacotherapy*, **62**, 94-98.
- Mose Larsen, P., Fey, S.J., Louw, J. & Joubert, E. (2008). An anti-diabetic extract of honeybush. World Intellectual Property Organization WO. 2008/110552A2.
- Newman, D.J. & Cragg, G.M. (2007). Natural products as sources of new drugs over the last 25 years. *Journal of Natural Products*, **70**, 461-477.
- Shin, Y., Liu, R.H., Nock, J.F., Holliday, D. & Watkins, C.B. (2007). Temperature and relative humidity effects on quality, total ascorbic acid, phenolics and flavonoid concentrations, and antioxidant activity of strawberry. *Postharvest Biology and Technology*, **45**, 349-357.
- Tomasi, G., van der Berg, F. & Andersson, C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics*, **18**, 231-241.
- Tistaert, C., Dejaegher, B. & Vander Heyden, Y. (2011). Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Analytica Chimica Acta*, **690**, 148-161.
- Tu, Y., Jeffries, C., Ruan, H., Nelson, C., Smithson, D., Shelat, A.A., Brown, K.A., Li, X., Hester, J.P., Smillie, T., Khan, I.A., Walker, L., Guy, K. & Yan, B. (2010). Automated high-throughput system to fractionate plant natural products for drug discovery. *Journal of Natural Products*, **73**, 751-754.
- Wagner, H. & Ulrich-Merzenich, G. (2009). Synergy research: Approaching a new generation of phytopharmaceuticals, *Phytomedicine*, **16**, 97-110.
- Williamson, E.M. (2001). Synergy and other interactions in phytomedicine. *Phytomedicine*, **8**, 401-409.
- Wold, S., Esbensen, K & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **2**, 37-52.
- Xu, C., Liang, Y., Chau, F. & Vander Heyden, Y. (2006). Pretreatment of chromatographic fingerprints for quality control of herbal medicines. *Journal of Chromatography A*, **1134**, 253-259.

Chapter 5

Predicting the total antioxidant capacity of *Cyclopia* spp. extracts using HPLC chromatographic fingerprints

5.1 Abstract

The prediction of the total antioxidant capacity (TAC) of aqueous extracts of *C. subternata* and *C. maculata* from their entire chromatographic fingerprints, as determined by high performance liquid chromatography (HPLC), was investigated. This approach does not require identification of the peaks. Chromatograms were evaluated to determine whether specific regions or peaks were responsible for the predicted TAC value that could therefore be used as markers. Such markers could consequently eliminate the need for physical analysis of TAC and thus evaluation of the TAC of extracts based on the HPLC data. TAC values of *C. subternata* extracts were determined, using the 2,2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging and oxygen radical absorbance capacity (ORAC) assays. *Cyclopia maculata* extracts were only analysed by the DPPH radical scavenging method. The HPLC chromatographic fingerprints of the extracts of the two *Cyclopia* species were pre-treated using alignment, background removal and normalisation, and outliers in each set were detected and eliminated. Prediction models were then constructed using partial least squares (PLS) regression. The prediction models performed satisfactorily with cross-validation and prediction errors ranging from 6.8% to 8.3% and 5.3% to 7.2%, respectively. These errors were within the expected precision for the assays. According to the variables of importance for projections (VIP) scores, the entire chromatogram appears to be important for the prediction of TAC. Several peaks in the chromatograms were recognised as very important for the construction of the model according to their large positive regression coefficients. A number of peaks with very low UV-Vis signals and high regression coefficients were found. These should be identified in future research by mass spectrometry. Furthermore, for the models constructed from TAC_{DPPH} data, isomangiferin and apigenin-6,8-di-*C*-glucoside had large positive regression coefficients, whereas isomangiferin had a large positive regression coefficient in the models constructed from TAC_{ORAC} data. These two compounds should be investigated further in terms of their relationship and contribution to the TAC of the extracts.

5.2 Introduction

Oxidative stress is recognised to be the cause of numerous diseases and health disorders in humans (Adly, 2010). Therefore, antioxidant-rich diets are important to promote the health of individuals (Vincent & Taylor, 2006; De la Iglesias *et al.*, 2013). As a result, plant extracts with high levels of antioxidants are increasingly used as nutraceuticals and/or food ingredients. Quality control of these extracts usually includes total antioxidant capacity (TAC) as quality indicator (Bell & Ou, 2007; Ninfali *et al.*, 2009). Studies relating dietary TAC with beneficial effects *in vivo* (Espín *et al.*, 2007; Puchau *et al.*, 2009; Hermsdorff *et al.*, 2011; Bahadoran *et al.*, 2012) support the use of this characteristic not only for quality control purposes, but also as a tool to assess health benefits of cumulative antioxidant capacity from food intake (Hermsdorff *et al.*, 2011).

Cyclopia spp. (honeybush) contain a variety of phenolic compounds, including mangiferin, shown to have antioxidant capacity in various tests (Mizra *et al.*, 2013). Polyphenol enrichment increased the TAC values of honeybush extracts (De Beer *et al.*, 2011). The compounds and concentrations vary across species and plants (Ferreira *et al.*, 1998; Kamara *et al.*, 2004; De Beer & Joubert, 2010; De Beer *et al.*, 2012; Kokotkiewicz *et al.*, 2012), making them very diverse. Wild-harvesting and harvesting of different genotypes (seedlings) contribute to the large variation. For nutraceuticals, this is not ideal, as production of a standard extract with set or known antioxidant capacity (quality control) is problematic.

The total antioxidant capacity (TAC) can be determined quantitatively by a variety of methods. These include assays such as the 2,2-diphenyl-1-picrylhydrazyl (DPPH) (Brand-Williams *et al.*, 1995), total phenols (TP) (Singleton & Rossi, 1965), oxygen radical absorbance capacity (ORAC) (Cao *et al.*, 1993) and ferric reducing antioxidant power (FRAP) (Benzie & Strain, 1996) assays, which have all been adapted to microplate format for high throughput (Huang *et al.*, 2002; Arthur *et al.*, 2011).

Alternative methods have been investigated to avoid the need for the chemical measurement of antioxidant capacity in large sample sets (Van Nederkassel *et al.*, 2005). These include the use of chemical fingerprinting. Fingerprints are a more complete representation of chemical data of the product, compared to quantitative data, which generally only represents a fraction of the chemical information of the product. The World Health Organisation (WHO) has accepted fingerprint analysis of herbal products for quality control purposes (Anon., 1991), and recent years have seen an increase in scientific literature in this regard, largely as a result of the application of fingerprint analysis to traditional Chinese medicine (Jiang *et al.*, 2010; Liang *et al.*, 2010; Yang *et al.*, 2011). High performance liquid chromatographic (HPLC) fingerprints are more popular compared to gas chromatography (GC) and capillary electrophoresis (CE) fingerprints (Jiang *et al.*, 2010) and would be more applicable when phenolic compounds are mainly responsible for antioxidant activity. This lead to the use of multivariate data analysis to determine or predict the TAC of samples by constructing a data model (Van Nederkassel *et al.*, 2005).

There have been a few successful attempts to predict the antioxidant activity or capacity of herbal products using multivariate calibration techniques in combination with chromatographic fingerprints. Van Nederkassel *et al.* (2005) developed and reported the first fast strategy for the determination of the total antioxidant capacity of Chinese green tea, using partial least squares (PLS) and uninformative variable elimination PLS (UVE-PLS) with more or less similar results. Similar studies followed on green tea (Daszykowski *et al.*, 2007; Dumarey *et al.*, 2008), *Mallotus* species (Nguyen Hoai *et al.*, 2009) and *Turnera diffusa* (Damiana) (Garza-Juárez *et al.*, 2011; Lucio-Gutiérrez *et al.*, 2012). Several of these papers suggested using multivariate regression as a better alternative to the chemical measurement of antioxidant capacity of herbal extracts. Although this may be the

case for HPLC methods using monolithic columns, resulting in a very short analysis time for each sample (Van Nederkassel *et al.*, 2005), employing HPLC methods with longer analysis times may be impractical and more expensive.

The objective of this study was to predict the TAC of aqueous extracts of two *Cyclopia* species from HPLC chromatographic fingerprints. Extensive pre-treatment of the data and the multivariate regressions technique, PLS, were used for model construction. Another goal was to determine whether specific compound(s) (peaks) or region(s) on the chromatogram contributed significantly to the prediction and could be used as a marker(s) for the prediction of TAC of the two *Cyclopia* species.

5.3 Experimental

5.3.1 Plant material and HPLC analysis

Cyclopia subternata plants (n = 64; same age), established from seeds collected in the wild, were harvested from a commercial plantation at Kanetberg Flora in the Barrydale district, South Africa. The harvested shoots (stems and leaves) were dried at 40 °C in a temperature-controlled drying tunnel with forced air circulation to ca. 8-10% moisture content and ground with a Retsch mill (1 mm sieve; Retsch GmbH, Haan, Germany). These samples represented green *C. subternata*.

The *Cyclopia maculata* sample set consisted of the shoots of seedlings (n = 40; 2-year-old) grown in pots at ARC Infruitec-Nietvoorbij, Stellenbosch, South Africa. The seeds for the pot trial were collected from a natural population, near Riversdale, South Africa. The seedlings were topped after 6 months to stimulate sprouting. The regrowth (top 30 cm) was harvested and comprised thin stems and leaves. The harvested shoots (material from each pot represented a sample) were dried and milled as described for *C. subternata*. These samples represented green *C. maculata*.

Aqueous extracts were prepared from all samples as described in Chapter 3 (section 3.3.3). HPLC analyses were conducted as described in Chapter 4 (section 4.3.2).

5.3.2 Pre-treatment of fingerprints

Handling of data and pre-treatment of chromatographic fingerprints were as described in Chapter 4 (section 4.3.3) and included techniques such as alignment of peaks by means of correlation optimised warping (COW), background removal and normalisation. Additional data pre-treatment was required for this study, which comprised the averaging of duplicate chromatograms (from duplicate HPLC injections) and the splitting of data sets into calibration and validation subsets to allow the validation of the prediction model (section 5.3.4.2). Fig. 5.1 summarises the required pre-treatment of chromatographic data as employed for this study.

5.3.3 Total antioxidant capacity assays

Two assays were conducted to determine the TAC of *Cyclopia* extracts. The TAC_{DPPH} value represents the 2,2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging activity of the samples. This microplate assay was conducted in triplicate as described by Arthur *et al.* (2011). The TAC_{DPPH} values are expressed as the average of the triplicate measurements, in μ moles Trolox equivalents (TE)/g extract. The DPPH assay was conducted on both *C. subternata* and *C. maculata* extracts.

The ORAC assay, adapted for a 96-well microplate fluorescence reader using automated dispensing of 2,2'-azobis (2-amidinopropane) dihydrochloride (AAPH), determines the TAC against a peroxy radical that is generated by thermal decomposition of AAPH (Huang *et al.*, 2002). The TAC_{ORAC} values also represent the average of triplicate measurements and are expressed as $\mu\text{moles TE/g}$ extract. The ORAC assay was conducted only on the *C. subternata* extracts.

The precision of the TAC_{DPPH} assay was determined using three randomly chosen *Cyclopia subternata* extracts. For each sample, the same extract was analysed six times in triplicate on one microplate. The relative standard deviation (%RSD) was calculated for the six repetitions, and represented the intra-plate precision. This procedure was repeated for six different plates. The relative standard deviation (%RSD) was calculated between the six plates and represented the inter-plate precision. A %RSD < 5% is acceptable for the DPPH assay. According to Huang *et al.* (2002), intra- and inter-day precision values < 15% are acceptable for the ORAC assay.

5.3.4 Data analysis

5.3.4.1 Data exploration and outlier detection

Histograms were used to detect extreme outliers in the TAC_{DPPH} and TAC_{ORAC} data sets. The Shapiro-Wilk test (Shapiro & Wilk, 1965) was also conducted on the TAC data sets of both *Cyclopia* species to determine if the data had a normal distribution. Data were regarded as normally distributed when $p > 0.05$ values were obtained.

Principal component analysis (PCA) was used for the exploration of the chromatographic fingerprint data of both species, as described in Chapter 4 (section 4.3.5) (Wold *et al.*, 1983).

5.3.4.2 Calibration and validation subset division

The chromatographic and corresponding TAC data sets of the two species were divided into two groups, respectively, to represent a calibration and validation set for each. Samples were first sorted from low to high TAC values. Samples distributed uniformly along the wide range of y values were then selected for the calibration set in order to have a good representation of data from which to construct the prediction model. In each case roughly 75% of the data set was used for calibration of the model.

5.3.4.3 Partial least squares

Partial least squares (PLS) regression is used to find the inner (linear) relationship or latent variables (LVs) that maximise the amount of variation explained between independent (\mathbf{X}) and dependent (\mathbf{y}) variables. The relationship represents the model, and allows one to predict new \mathbf{y} variables from new \mathbf{X} data. PLS is one of the most popular multivariate regression methods. This regression method involves the simultaneous modelling of both \mathbf{X} and \mathbf{y} by using the variance of \mathbf{X} and the covariance between \mathbf{X} and \mathbf{y} (Martens & Naes, 1989; Wold *et al.*, 2001).

In this study, the \mathbf{X} matrix was composed of the pre-treated chromatographic fingerprints, and the \mathbf{y} vector was constructed from the TAC values obtained from the two assays, respectively. The PLS regression model can be explained by the following equation(s):

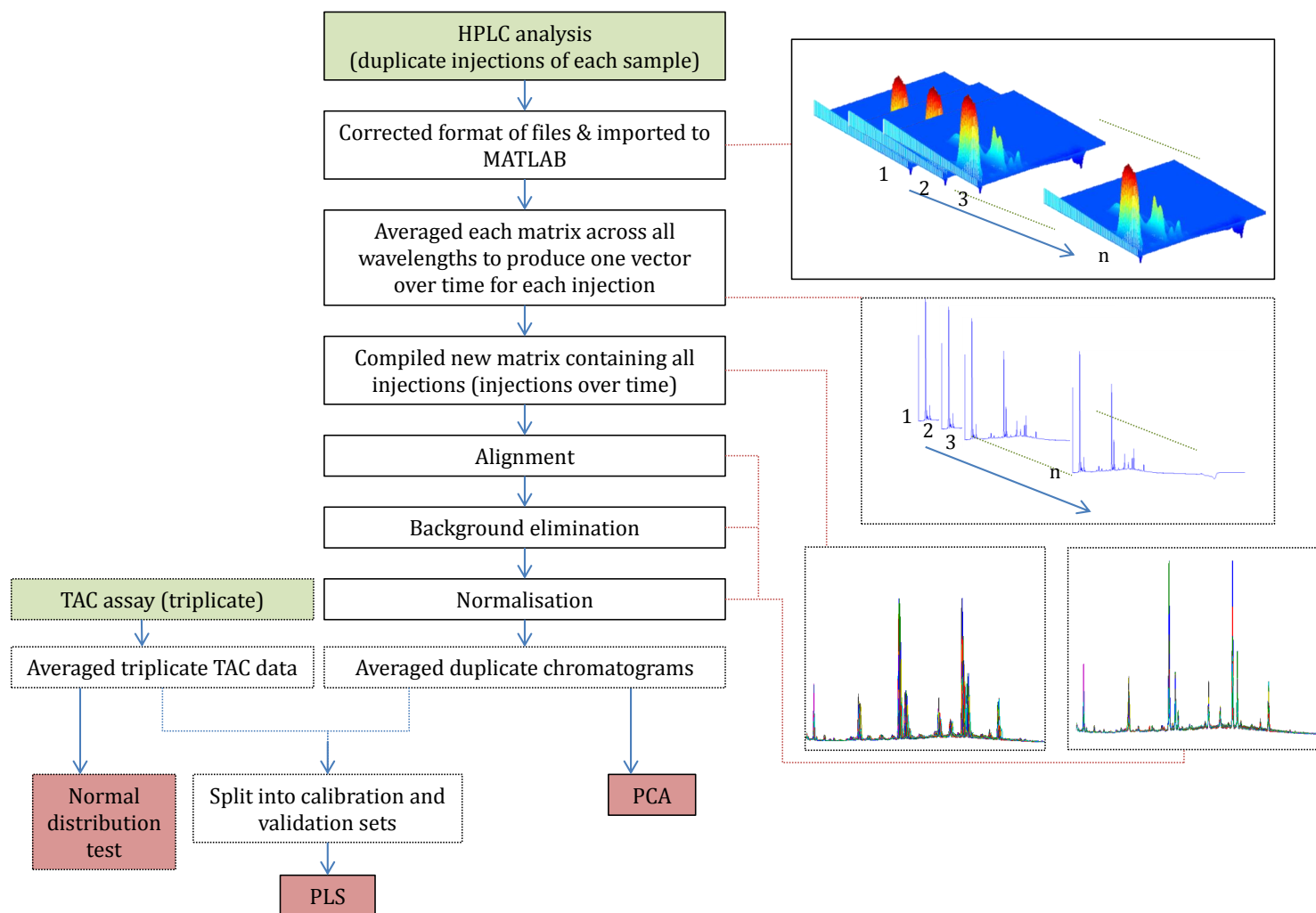


Figure 5.1 – Schematic representation of data handling and pre-treatment [TAC = total antioxidant capacity; PCA = principal component analysis; PLS = partial least squares].

$$\mathbf{X} = \mathbf{TP}^t + \mathbf{E} \quad (5.3.1)$$

$$\mathbf{y} = \mathbf{TP}^t\mathbf{b} + \mathbf{f} \quad (5.3.2)$$

where \mathbf{T} is the score matrix for \mathbf{X} and \mathbf{y} ;
 \mathbf{P} the loading matrix of \mathbf{X} on \mathbf{T} ;
 \mathbf{E} the residual matrix of \mathbf{X} ;
 \mathbf{b} the regression coefficients vector ($\mathbf{b} = \mathbf{Pq}$);
 \mathbf{q} the loading vector of \mathbf{y} on \mathbf{T} ; and
 \mathbf{f} the residual vector of \mathbf{y} .

From these equations it can be seen that the regression coefficients (\mathbf{b}) can be used to determine the contribution of the original variables to the final model (Dumarey *et al.*, 2008).

Cross-validation (CV) was used to evaluate the number of LVs to include in the model. This entails the construction of models from the calibration set using a certain percentage of samples (depending on the CV method used) and then using the remaining samples of the calibration set to test the prediction capability of the model. This is then expressed as the root mean square error of CV (RMSECV). The error is expressed as the difference between the actual (observed) and predicted values. This CV procedure is repeated numerous times, and the error values obtained for all of the samples are averaged for the given model complexity (LVs). In this study, the leave-one-out cross-validation (LOO-CV) method was used. During LOO-CV, each object of the data set is left out once, while the remaining objects of the calibration set are used to build the PLS model. The RMSECV (Equation 5.3.3) can be calculated from the cross-validation results.

$$\text{RMSECV}(f) = \sqrt{\sum_{i=1}^N \frac{(\hat{y}_{CV,i} - y_i)^2}{N}} \quad (5.3.3)$$

where y_i is the measured response of the i th sample;
 $\hat{y}_{CV,i}$ is a predicted response from a calibration equation obtained for the data without the i th sample;
 N is the number of calibration samples; and
 f denotes the number of latent factors, i.e. the model complexity.

The PLS model can also be validated with an independent test set (the selected validation set) for which the RMSEP is computed as:

$$\text{RMSEP}(f) = \sqrt{\sum_{i=1}^{N_t} \frac{(\hat{y}_i^t - y_i^t)^2}{N_t}} \quad (5.3.4)$$

where N_t is the number of test samples; and
 \hat{y}_i^t and y_i^t are the respective predicted and measured response values for test sample, i .

The suitability of the constructed PLS model is evaluated using the RMSECV and RMSEP values compared to the expected precision of the predicted \mathbf{y} value. The optimal model complexity (LV) is determined by selecting the simplest model which results in a low RMSECV value, while maintaining a relatively low RMSEP value.

5.4 Results and discussion

5.4.1 Pre-treatment of chromatographic data

The pre-treatment applied to chromatographic data was similar to that described in detail in Chapter 4 (section 4.4.1). For this study, however, averaging of all chromatograms from duplicate injections was done to obtain single vectors for each extract. To build a prediction model with PLS regression from these chromatographic signals, normalisation was required as the antioxidant capacity assays were calculated and expressed as normalised values (amount per soluble solids) for each sample. The final dimensions of the chromatographic data were 64×5370 for *C. subternata* and 40×5381 for *C. maculata*.

5.4.2 Data exploration and outlier detection

Before the model could be constructed, outlier detection was performed for the TAC and pre-treated fingerprint data sets of both species. The Shapiro-Wilk test for normal distribution was applied to the TAC data sets, together with the use of histograms. The histograms showed the data were normally distributed (Fig. 5.2), with the DPPH values for *C. subternata* slightly skewed to the left. The Shapiro-Wilk test confirmed that the TAC_{DPPH} and TAC_{ORAC} values were normally distributed ($p > 0.05$), and subsequently no outliers were detected based on the TAC values.

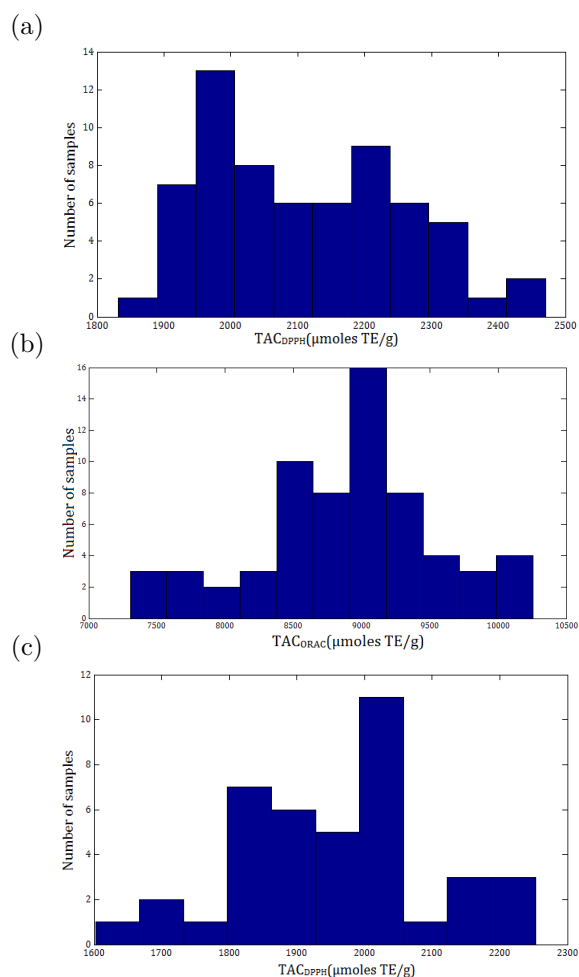


Figure 5.2 – Histograms showing the distribution of TAC values from (a) DPPH and (b) ORAC assays for *C. subternata* and (c) DPPH assay for *C. maculata*. Histograms of *C. subternata* and *C. maculata* data comprised 10 and 11 bins, respectively [TE = Trolox equivalents].

The pre-treated chromatographic fingerprint data were tested for outliers by means of PCA. No outliers were observed in PC1 vs. PC2. In the scores plot of PC2 vs. PC3 for the chromatograms of the *C. subternata* extracts, the only outlier detected, was sample 41 in the direction of PC3 (Fig. 5.3). After careful examination of the fingerprint, this extract was considered an extreme, with very large peaks, rather than an outlier and consequently not removed.

For *C. maculata*, the scores plot of PC2 vs. PC3 (Fig. 5.4) indicated that extract 24 could be considered an outlier as it had extreme score values for both PC2 and PC3. After careful scrutiny of the chromatogram, extract 24 was considered ‘atypical’ compared to the rest of the sample set, as it had an extra peak not present in the other samples (Addendum B, Figs. B.1, B.2 and B.3). Sample 24 was therefore discarded, as building a model which includes only one such sample could be detrimental for accurate prediction. No further samples of the *C. maculata* data set were considered outliers (Fig. 5.4).

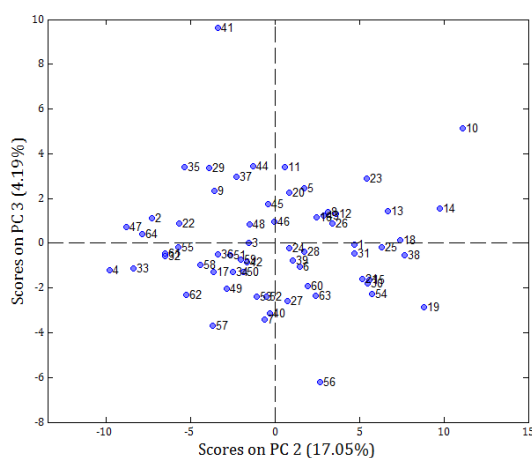


Figure 5.3 – PCA scores plot of PC2 vs. PC3 for chromatograms of *C. subternata*

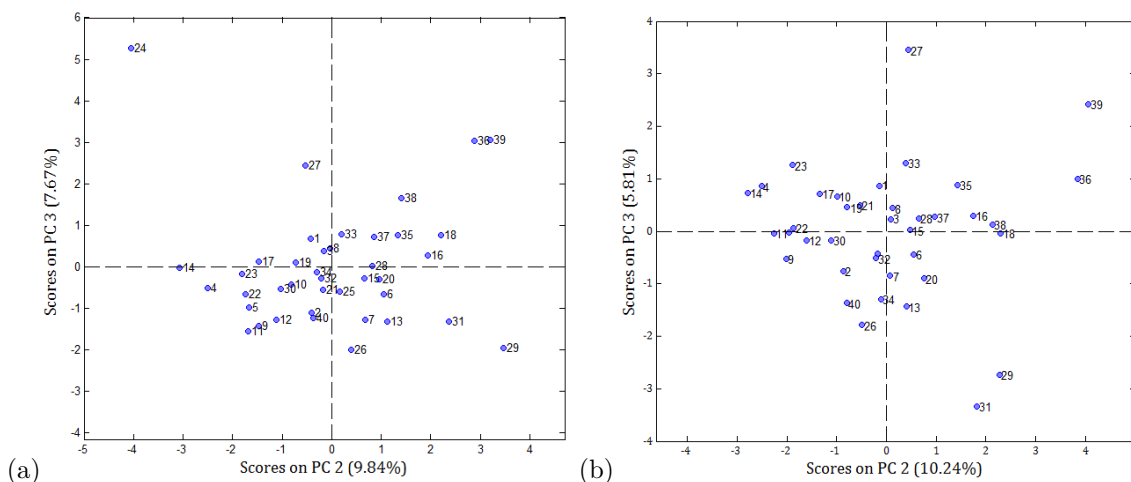


Figure 5.4 – PCA score plots of PC2 vs. PC3 for chromatograms of *C. maculata* (a) before and (b) after removal of the outlier (sample 24).

5.4.3 Calibration and validation subset division

For *C. subternata* $n = 16$ samples were selected for the validation set, with $n = 48$ samples comprising the calibration set. For *C. maculata*, $n = 29$ samples were selected for the calibration set, and $n = 10$ for the validation set as one sample had already been removed.

5.4.4 Partial least squares models

The first objective was to model the data from the TAC assays, using the calibration chromatographic data sets of the two *Cyclopi*a species. In the case of *C. subternata*, TAC_{DPPH} and TAC_{ORAC} data sets were available, while only TAC_{DPPH} data were available for *C. maculata*, requiring three separate models. The optimal model complexity of each of the three constructed models were determined from the RMSECV values obtained during the LOO-CV procedure and the RMSEP values obtained during the prediction of the external validation set. The optimal model complexity for each model was chosen by determining the LV with a combination of low RMSECV value, while still maintaining a relatively simple model and low RMSEP value (Van Nederkassel *et al.*, 2005; Verboven *et al.*, 2012). RMSE values obtained during the modelling processes are illustrated in Figs. 5.5, 5.7 and 5.9. Prediction models were plotted as scatter plots of measured vs. predicted TAC results in Figs. 5.6, 5.8 and 5.10. The RMSE values for the models constructed with the optimal model complexities are summarised in Table 5.1.

Table 5.1 – RMSE values of PLS models for *C. subternata* (TAC_{DPPH} and TAC_{ORAC}) and *C. maculata* (TAC_{DPPH})

Extracts from	TAC	LVs	RMSEC	RMSECV	RMSEP
<i>Cyclopi</i> a <i>subternata</i>	DPPH	5	110 (5.2%)	166 (7.9%)	111 (5.3%)
	ORAC	5	455 (5.1%)	741 (8.3%)	605 (6.8%)
<i>Cyclopi</i> a <i>maculata</i>	DPPH	2	88 (4.5%)	135 (6.8%)	141 (7.2%)

TAC = total antioxidant capacity; LVs = latent variables; RMSEC = root mean square error of calibration; RMSECV = root mean square error of cross-validation; RMSEP = root mean square error of prediction; values in brackets are the RMSE values expressed as a percentage.

The RMSE values obtained are comparable to TAC prediction models of other herbal species reported in literature, which ranged between 6.7% and 12.5% (Van Nederkassel *et al.*, 2005). The error values obtained during these model constructions are also in the range of precision values for the microplate TAC assays themselves. For the TAC_{DPPH} assay an %RSD < 5% is acceptable, as experimentally determined in this study, while an %RSD < 15% is acceptable for the TAC_{ORAC} assay (Huang *et al.*, 2002). %RSD values obtained for TAC assays were within these limits and therefore these prediction models can be considered acceptable.

To identify compounds or regions on the chromatograms that had large contributions to the modelled TAC of the extracts, the loadings, variables of importance for projection (VIP) scores and regression coefficients were plotted as a function of time for each model (Figs. 5.11, 5.12 and 5.13).

Relatively large regression coefficients are an indication of important variables for the modelling of \mathbf{y} (the TAC values), while important variables for the modelling of \mathbf{X} (matrix of fingerprints) are indicated by large loadings (Wold *et al.*, 2001). The VIP scores are regarded as a tool that summarises the important variables for projection. Loadings change with changes in model complexities, whereas the regression coefficients and VIP

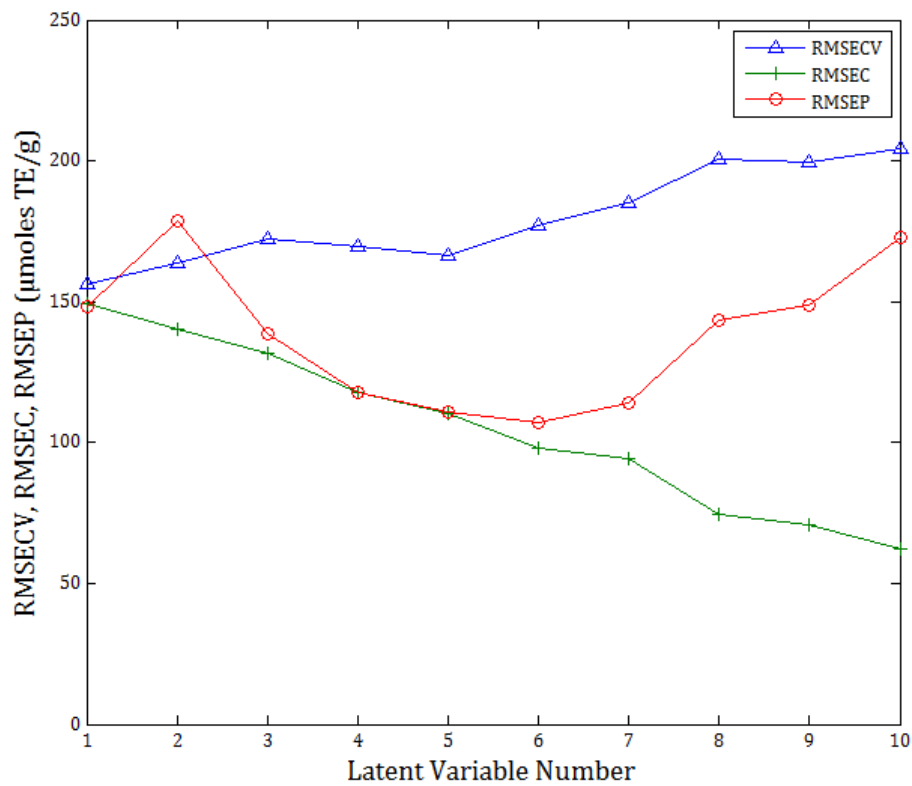


Figure 5.5 – RMSE values for models with varying complexity to predict TAC_{DPPH} of *C. subternata* extracts

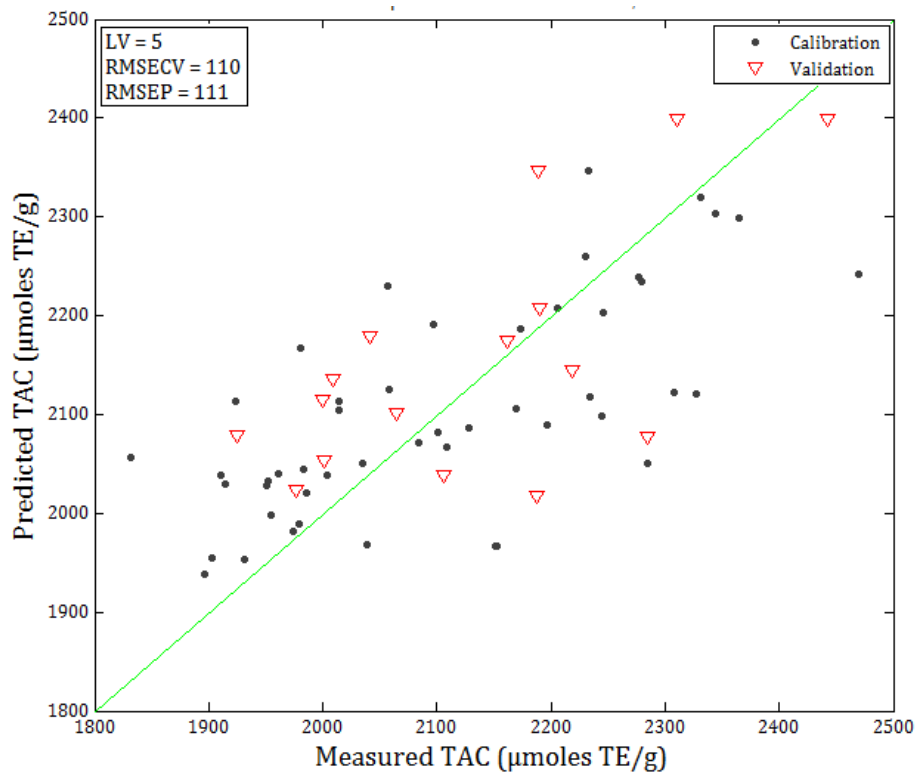


Figure 5.6 – Measured vs. predicted TAC_{DPPH} of *C. subternata* extracts

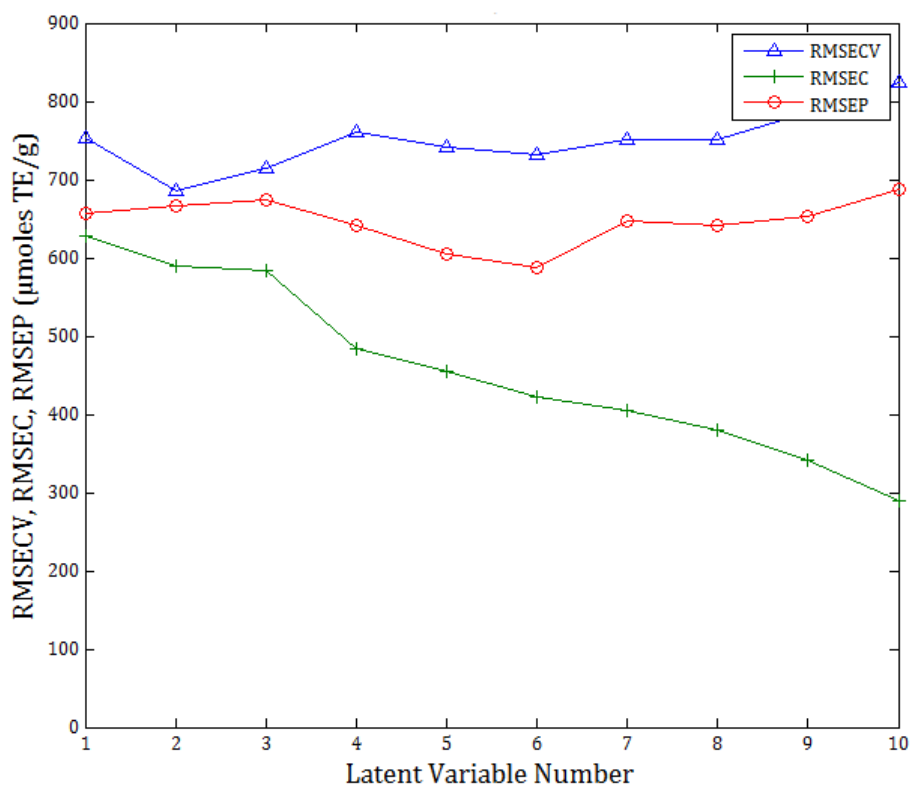


Figure 5.7 – RMSE values for models with varying complexity to predict TAC_{ORAC} of *C. subternata* extracts

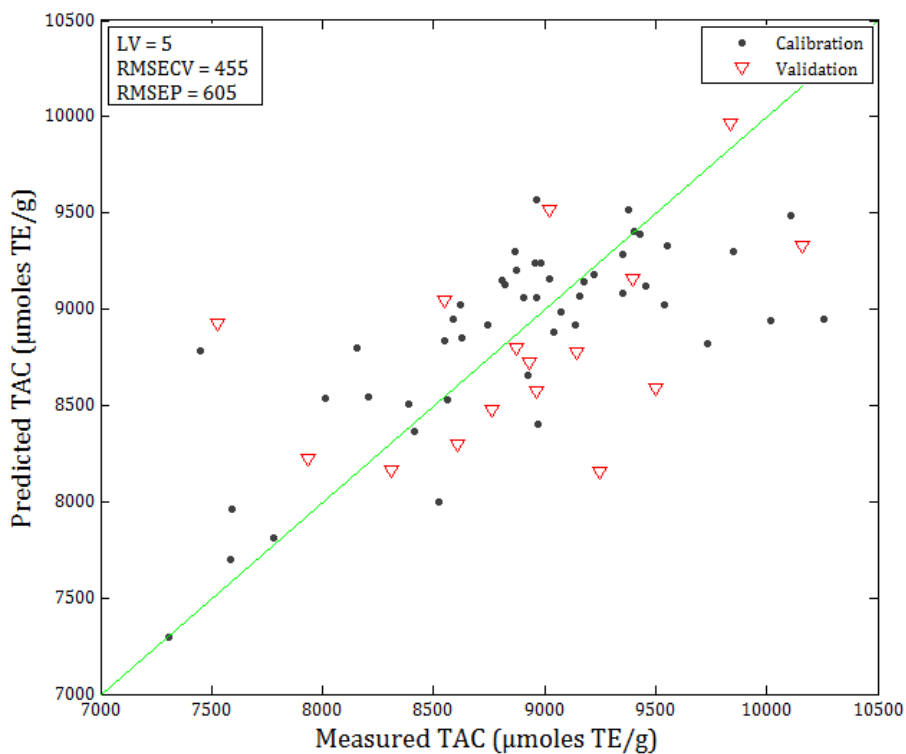


Figure 5.8 – Measured vs. predicted TAC_{ORAC} of *C. subternata* extracts

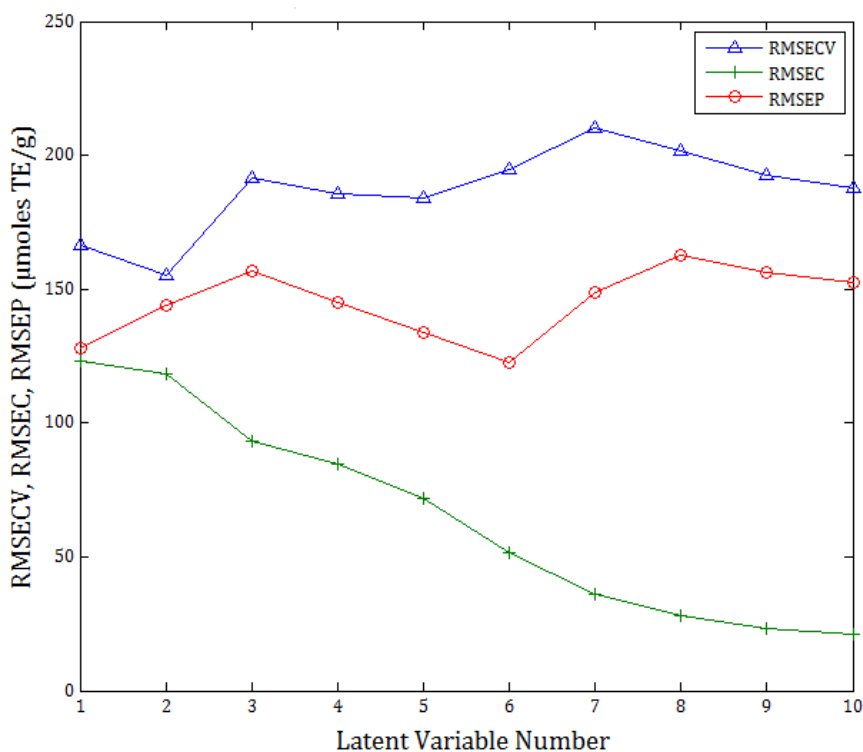


Figure 5.9 – RMSE values for models with varying complexity to predict TAC_{DPPH} of *C. maculata* extracts

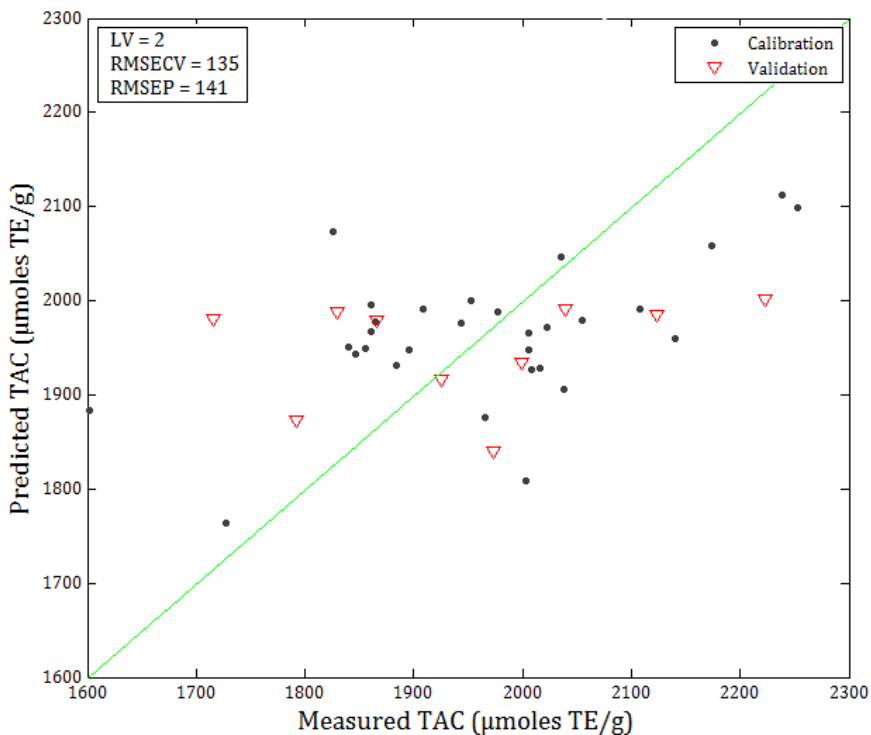


Figure 5.10 – Measured vs. predicted TAC_{DPPH} of *C. maculata* extracts

scores are more reliable and constant. Therefore the VIP scores and regression coefficients were examined in detail.

VIP scores are generally interpreted as follows: variables are plotted as the magnitude of importance on the y-axis, and chromatographic peaks with VIP scores > 1 are considered important. If it is preferred to reduce or increase the number of important variables, the threshold can be increased or lowered (Chong & Jun, 2005; Lucio-Gutiérrez *et al.*, 2012). The VIP scores for the three models indicated that most of the major peaks on the UV-Vis chromatograms are considered important for the prediction of TAC (VIP > 1).

To determine the nature of the contribution of the compounds indicated by the VIP scores, the regression coefficients were interpreted. Depending on the type of TAC assay conducted, the interpretation of regression coefficients can vary. For instance, when the values decrease with an increase in antioxidant capacity (for example EC₅₀ values (Chen *et al.*, 2013) which represent the concentration necessary to obtain an antioxidant effect equal to 50% scavenging in DPPH assay), the regression coefficients displayed as negative values have a positive contribution. In this study, the TAC_{DPPH} and TAC_{ORAC} values increased with an increase in antioxidant activity, therefore positive regressions coincide with an increase in TAC.

Several peaks with high VIP scores also displayed large positive regression coefficients. There appears to be some consistency between the different models, with the same peaks, namely isomangiferin and apigenin-6,8-di-*C*-glucoside, contributing positively to the two DPPH models, and isomangiferin contributing positively to the ORAC model. This could indicate that these two compounds are very good antioxidants, or have a positive correlation with the TAC of the extracts. Mangiferin, the isomer of isomangiferin, and hesperidin, on the other hand, had negative regression coefficients, irrespective of the model. Although mangiferin and hesperidin are good antioxidants (Wilmsen *et al.*, 2005; Matkowski *et al.*, 2013), this could be indicative of an underlying phenomenon (latent variable) responsible for the negative association between these compounds and the antioxidant capacity. A possible explanation could be the inter relationship of the biosynthesis of phenolic compounds and the formation of compounds more active as antioxidants.

For TAC_{DPPH} of *C. subternata*, 3-hydroxyphloretin-3',5'-di-*C*-hexoside and eriocitrin had high VIP scores and positive regression coefficients, while iriflophenone-di-*O,C*-hexoside, scolymoside and phloretin-3',5'-di-*C*-glucoside had negative regression coefficients and therefore negative contributions towards the TAC. As for the other two models, all of the aforementioned compounds had negative contributions. There were also cases where some of the compounds had simultaneous positive and negative regression coefficients in the same model, such as iriflophenone-3-*C*-glucoside (Figs. 5.11 and 5.12). This was most likely caused by orthogonal variation present in \mathbf{X} (variation in \mathbf{X} that is unrelated to \mathbf{y}) (Trygg & Wold, 2002). This could also be due to overfitting of the model, however, as the model complexities were relatively low, the former is more probable. It is also common for some compounds known not to have antioxidative properties to show positive regression coefficients (Dumarey *et al.*, 2008). In these cases, it can be concluded that the compound is either co-eluting with an antioxidant compound, or that there is an underlying correlation between the concentration of that compound and the antioxidant capacity, which coincidentally increases with the increase in its concentration. Another possibility is that the compound might have a positive correlation with another antioxidant compound. In this study, co-elution should not have been a contributing factor, as the HPLC-DAD analysis method was developed to provide good resolution between compounds and peak purities were confirmed during method validation (Chapter 3, section 3.4.3).

Interestingly, there were peaks with high VIP scores that had very small responses in the UV-Vis chroma-

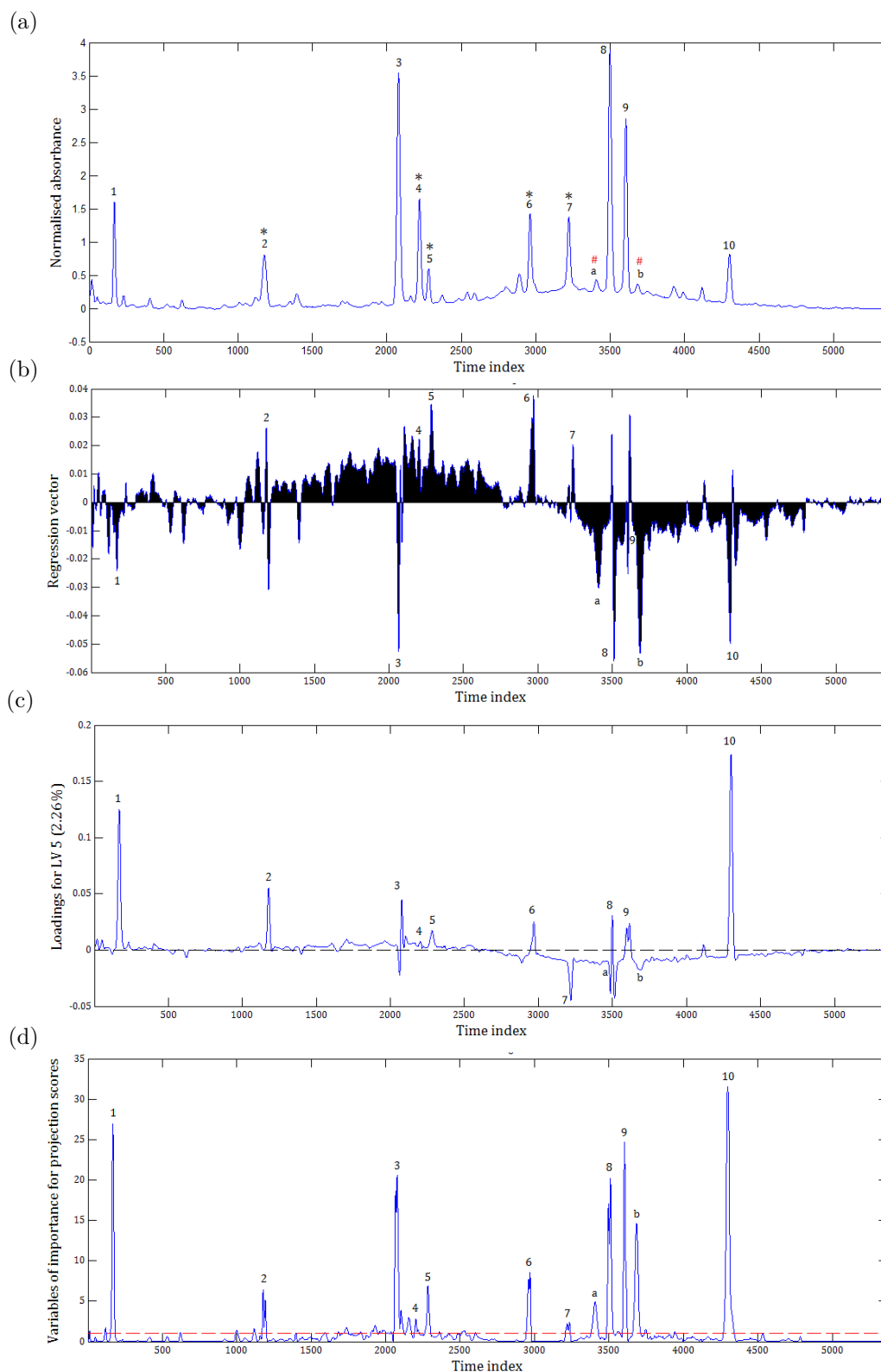


Figure 5.11 – Compounds of importance when predicting TAC_{DPPH} of *C. subternata* extracts shown on a (a) representative chromatogram, as evaluated using (b) regression coefficients, (c) loading scores and the (d) variables of importance for projection (VIP) scores [1 = iriflophenone-di-*O,C*-hexoside; 2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = apigenin-6,8-di-*C*-glucoside; 6 = 3-hydroxyphloretin-3',5'-di-*C*-hexoside; 7 = eriocitrin; 8 = scolymoside; 9 = phloretin-3',5'-di-*C*-glucoside; 10 = hesperidin; a = b = unidentified; * = compounds with positive regression and high variables of importance for projection scores; # = compounds of interest with low UV-Vis signals].

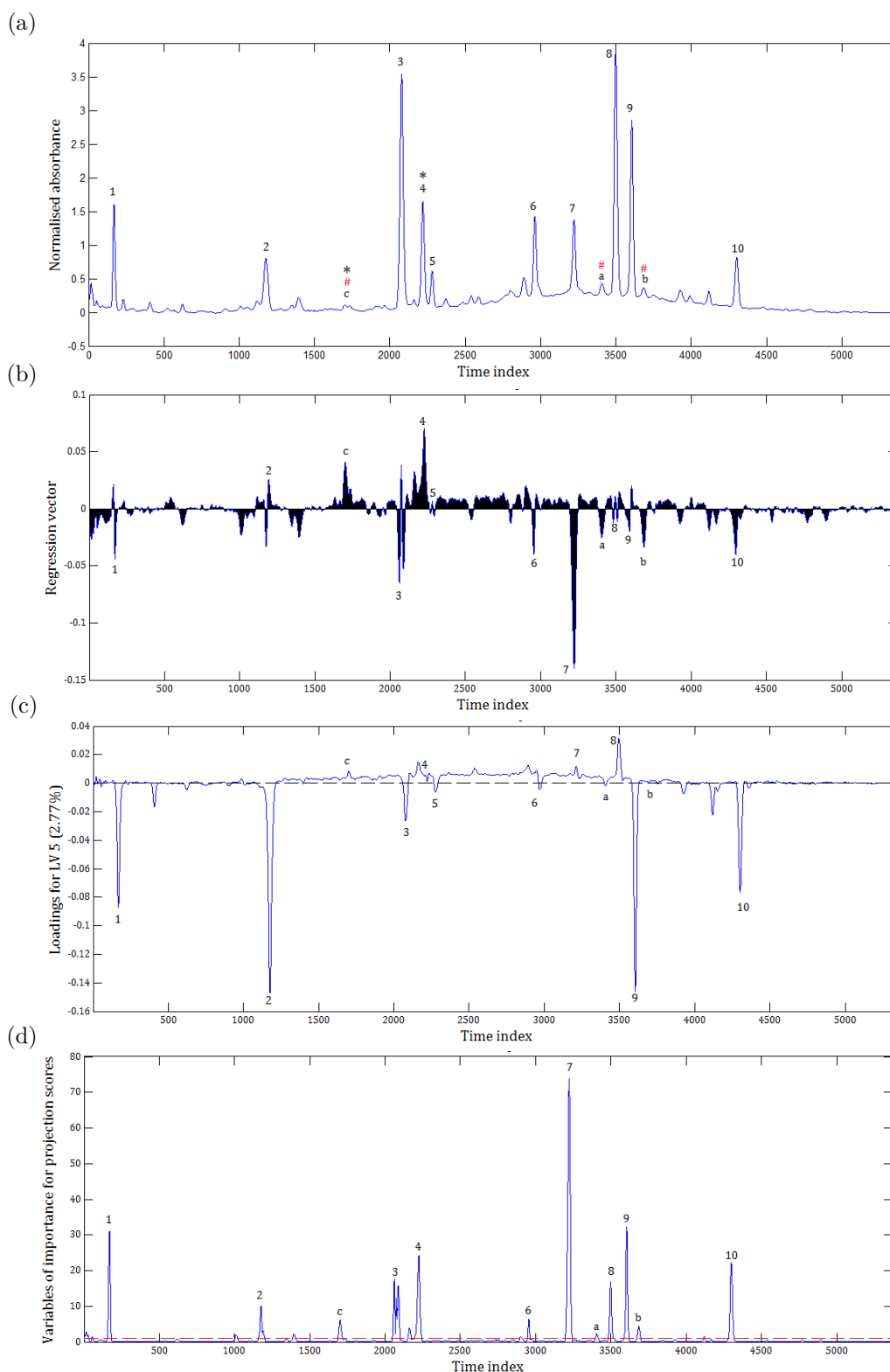


Figure 5.12 – Compounds of importance when predicting TAC_{ORAC} of *C. subternata* extracts shown on a (a) representative chromatogram, as evaluated using (b) regression coefficients, (c) loading scores and the (d) variables of importance for projection (VIP) scores [1 = iriflophenone-di-*O*,*C*-hexoside; 2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = apigenin-6,8-di-*C*-glucoside; 6 = 3-hydroxyphloretin-3',5'-di-*C*-hexoside; 7 = eriocitrin; 8 = scolymoside; 9 = phloretin-3',5'-di-*C*-glucoside; 10 = hesperidin; a = b = c = unidentified; * = compounds with positive regression and high variables of importance for projection scores; # = compounds of interest with low UV-Vis signals].

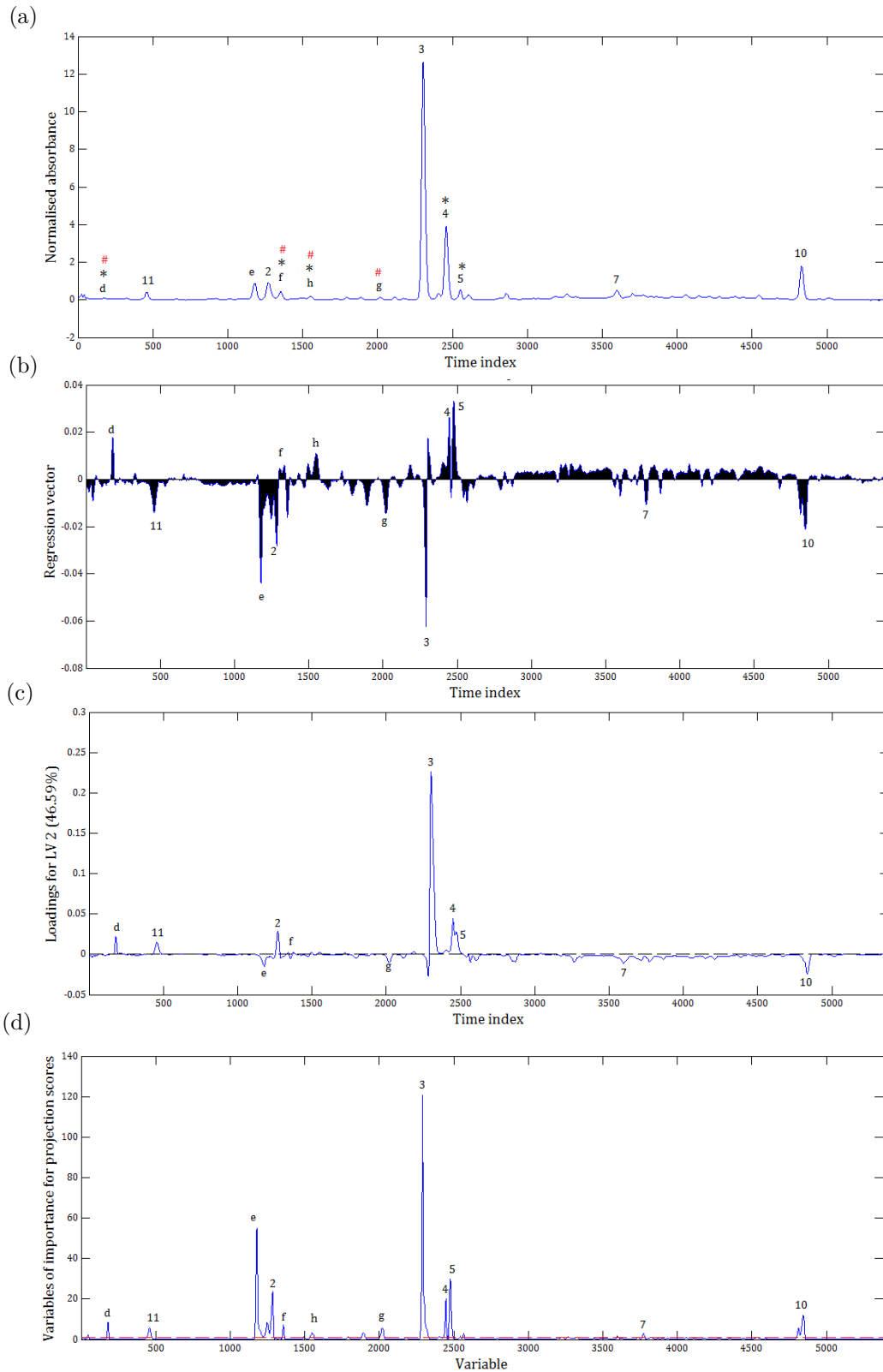


Figure 5.13 – Compounds of importance when predicting TAC_{DPPH} of *C. maculata* extracts shown on a (a) representative chromatogram, as evaluated using (b) regression coefficients, (c) loading scores and the (d) variables of importance for projection (VIP) scores [2 = iriflophenone-3-*C*-glucoside; 3 = mangiferin; 4 = isomangiferin; 5 = apigenin-6,8-di-*C*-glucoside; 7 = eriocitrin; 10 = hesperidin; 11 = maclurin-3-*C*-glucoside; d = e = f = g = h = unidentified; * = compounds with positive regression and high variables of importance for projection scores; # = compounds of interest with low UV-Vis signals].

tograms (peaks **a** - **h** in Figs. 5.11, 5.12 and 5.13). This could be an indication of compounds that contribute greatly to the prediction of TAC. Compounds **a**, **b**, **e**, **g** and **h** exhibited positive regression coefficients, while **d** and **f** had negative regression coefficients. These compounds should be investigated further, with respect to their identities, and to determine their relationship with the TAC of these extracts.

From the data it was clear that the correlation of the chromatographic fingerprints to TAC values of the extracts cannot be reduced to one or two marker compounds to predict TAC. The complex nature of the effect of the compounds on the TAC does not make it possible at this stage. However, compounds of interest have been determined that require further investigation. These compounds exhibited positive regression coefficients and have been marked with black star symbols (*) on the chromatograms. For *C. subternata* these were: iriflophenone-3-*C*-glucoside, isomangiferin, apigenin-6,8-di-*C*-glucoside, eriocitrin and 3-hydroxyphloretin-3',5'-di-*C*-hexoside for TAC_{DPPH}; and isomangiferin and unidentified compound **c** for TAC_{ORAC}. For *C. maculata* these were isomangiferin, apigenin-6,8-di-*C*-glucoside and unidentified compounds **d**, **f** and **h** in TAC_{DPPH}. Unidentified peaks that were also shown to have large contributing factors towards the TAC prediction, while displaying small chromatographic peaks, have been marked with red hash symbols (#). It is proposed that the latter compounds should be identified using techniques such as mass spectrometry as they could have either interesting antioxidative properties, or have good correlations with the TAC of these extracts.

5.5 Conclusions

PLS models were constructed in this study for the prediction of TAC of the extracts of two *Cyclopia* species, using two different antioxidant assays. The root mean squared errors obtained were comparable to those of the relative standard deviation expected when conducting the precision tests on the two microplate assays, as well as other antioxidant prediction models reported in literature. Therefore, the models could be used to predict the total antioxidant capacity from new HPLC chromatographic fingerprints of the same *Cyclopia* species, with a reasonable error. Although no marker compounds could be identified for rapid estimation of TAC of *Cyclopia* samples, many compounds of interest have been detected for further investigation. Based on VIP scores it appears as though the majority of the chromatographic peaks are important for the prediction models. Regions of higher importance in the chromatograms were determined through a combination of VIP scores and positive regression coefficients. Unidentified minor peaks exhibited high VIP scores and large positive regression coefficient are of interest. It is recommended that these compounds should be identified by mass spectrometry. Using PLS to predict the total antioxidant capacity of *Cyclopia* species was shown to be an adequate tool to determine the compounds of interest mostly responsible in the relationship between the chromatographic fingerprints of the extracts and their TAC values.

References

- Adly, A.A.M. (2010). Oxidative stress and disease: An updated review. *Research Journal of Immunology*, **3**, 129-145.
- Anonymous (1991). Guidelines for the assessment of herbal medicines, World Health Organisation, Geneva, Switzerland. http://whqlibdoc.who.int/hq/1991/WHO_TRM_91.4.pdf (February 2013).
- Arthur, H., Joubert, E., De Beer, D., Malherbe, C.J. & Witthuhn, R.C. (2011). Phenylethanoid glycosides as major antioxidants in *Lippia multiflora* herbal infusion and their stability during steam pasteurisation of plant material. *Food Chemistry*, **127**, 581-588.
- Bahadoran, Z., Golzarand, M., Mirmiran, P., Shive, N. & Azizi, F. (2012). Dietary total antioxidant capacity and the occurrence of metabolic syndrome and its components after 3-year follow-up in adults: Tehran lipid and glucose study. *Nutrition and Metabolism*, **9**(70), doi:10.1186/1743-7075-9-70.
- Bell, D. & Ou, B. (2007). The branding of ORAC: where will it lead? *Nutraceuticals World*, **10**(3), 72-74.
- Benzie, I.F.F. & Strain, J.J. (1996). The ferric reducing ability of plasma (FRAP) as a measure of 'antioxidant power': The FRAP assay. *Analytical Biochemistry*, **239**, 70-76.
- Brand-Williams, W., Cuvelier, M.E. & Berset, C. (1995). Use of free radical method to evaluate antioxidant activity. *Lebensmittel-Wissenschaft und Technologie*, **28**, 25-30.
- Cao, G.H., Alessio, H.M. & Cutler, R.G. (1993). Oxygen-radical absorbance capacity assay for antioxidants. *Free Radical Biology and Medicine*, **14**, 303-311.
- Chen, Z., Bertin, R. & Froldi, G. (2013). EC₅₀ estimation of antioxidant activity in DPPH• assay using several statistical programs. *Food Chemistry*, **138**, 414-420.
- Chong, I. & Jun, C. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, **78**, 103-112.
- Daszykowski, M., Vander Heyden, Y. & Walczak, B. (2007). Robust partial least squares model for prediction of green tea antioxidant capacity from chromatograms. *Journal of Chromatography A*, **1176**, 12-18.
- De Beer, D. & Joubert, E. (2010). Development of HPLC method for *Cyclopia subternata* phenolic compound analysis and application to other *Cyclopia* spp. *Journal of Food Composition and Analysis*, **23**, 289-297.
- De Beer, D., Joubert, E., Malherbe, C.J. & Brand, D.J. (2011). Use of countercurrent chromatography during isolation of 6-hydroxyluteolin-7-O-β-glucoside, a major antioxidant of *Athrixia phylicoides*. *Journal of Chromatography A*, **1218**, 6179-6186.
- De Beer, D., Schulze, A.E., Joubert, E., De Villiers, A., Malherbe, C.J. & Stander, M.A. (2012). Food ingredient extracts of *Cyclopia subternata* (honeybush) variation in phenolic composition and antioxidant capacity. *Molecules*, **17**, 14602-14624.
- De la Iglesia, R., Lopez-Legarrea, P., Celada, P., Sánchez-Muniz, F.J., Martínez, J.A. & Zulet, M.A. (2013). Beneficial effects of the RESMENA dietary pattern on oxidative stress in patients suffering from metabolic syndrome with hyperglycemia are associated to dietary TAC and fruit consumption. *International Journal of Molecular Science*, **14**, 6903-6919.
- Dumarey, M., van Nederkassel, A.M., Deconinck, E. & Vander Heyden, Y. (2008). Exploration of linear multivariate calibration techniques to predict the total antioxidant capacity of green tea from chromatographic fingerprints. *Journal of Chromatography A*, **1192**, 81-88.
- Espín, J.C., García-Conesa, M.T. & Tomás-Barberán, F.A. (2007). Nutraceuticals: Fact or fiction. *Phytochemistry*, **68**, 2986-3008.

- Ferreira, D., Kamara, B.I., Brandt, E.V. & Joubert, E. (1998). Phenolic compounds from *Cyclopia intermedia* (honeybush tea). *Journal of Agricultural and Food Chemistry*, **46**, 3406-4310.
- Garza-Juárez, A., Salazar-Cavazos, M.L., Salazar-Aranda, R., Pérez-Meseguer, J. & Waksman de Torres, N. (2011). Correlation between chromatographic fingerprint and antioxidant activity of *Turnera diffusa* (Dami-ana). *Planta Medica*, **77**, 958-963.
- Hermesdorff, H.H.M., Puchau, B., Volp, A.C.P., Barbosa, K.B.F., Bressan, J., Zulet, M.A. & Mertínez, J.A. (2011). Dietary total antioxidant capacity is inversely related to central adiposity as well as to metabolomic and oxidative stress markers in healthy young adults. *Nutrition and Metabolism*, **8**, 59-66.
- Huang, D., Ou, B., Hampsch-Woodill, M., Flanagan, J.A., Prior, R.L. (2002). High-throughput assay of oxygen radical absorbance capacity (ORAC) using a multichannel liquid handling system coupled with a microplate fluorescence reader in 96-well format. *Journal of Agricultural and Food Chemistry*, **50**, 4437-4444.
- Jiang, Y., David, B., Tu, P. & Barbin, Y. (2010). Recent analytical approaches on quality control of traditional Chinese medicines - A review. *Analytica Chimica Acta*, **657**, 9-18.
- Kamara, B.I., Brand, D.J., Brandt, E.V. & Joubert, E. (2004). Phenolic metabolites from honeybush tea (*Cyclopia subternata*). *Journal of Agricultural and Food Chemistry*, **52**, 5291-5395.
- Kokotkiewicz, A., Luczkiewicz, M., Sowinski, P., Glod, D., Gorynski, K. & Bucinski, A. (2012). Isolation and structure elucidation of phenolic compounds from *Cyclopia subternata* Vogel (honeybush) intact plant and *in vitro* cultures. *Food Chemistry*, **133**, 1373-1382.
- Liang, Y., Xie, P. & Chau, F. (2010). Chromatographic fingerprinting and related chemometric techniques for quality control of traditional Chinese medicine. *Journal of Separation Science*, **33**, 410-421.
- Lucio-Gutiérrez, J.R., Garza-Juárez, A., Coello, J., MasPOCH, S., Salazar-Cavazos, M.L., Salazar-Aranda, R. & Waksman de Torres, N. (2012). Multi-wavelength high-performance liquid chromatographic fingerprints and chemometrics to predict the antioxidant activity of *Turnera diffusa* as part of its quality control. *Journal of Chromatography A*, **1235**, 68-76.
- Martens, H. & Naes, T. (1989). *Multivariate Calibration*. Great Britain: John Wiley & Sons Ltd.
- Matkowski, A., Kuśś, P., Góralaska, E. & Woźniak, D. (2013). Mangiferin - a bioactive xanthonoid, not only from mango and not just antioxidant. *Mini-Reviews in Medicinal Chemistry*, **13**, 439-455.
- Mizra, R.H., Chi, N. & Chi, Y. (2013). Therapeutic potential of the natural product mangiferin in metabolic syndrome. *Journal of Nutritional Therapeutics*, **2**, 74-79.
- Nguyen Hoai, N., Dejaegher, B., Tistaert, C., Nguyen Thi Hong, V., Rivière, C., Chataigné, G., Phan Van, K., Chau Van, M., Quentin-Leclercq, J. & Vander Heyden, Y. (2009). Development of HPLC fingerprints for *Mallotus* species extracts and evaluation of peaks responsible for their antioxidant activity. *Journal of Pharmaceutical and Biomedical Analysis*, **50**, 753-763.
- Ninfali, P., Gennari, L. & Biagotti, E. (2009). Improvement in botanical standardisation of commercial freeze-dried herbal extracts by using the combination of antioxidant capacity and constituent marker concentrations. *Journal of AOAC International*, **92**, 797-805.
- Puchau, B., Zulet, M.A., de Echavarri, A.G., Hermesdorff, H.H.M. & Martinez, J.A. (2009). Dietary total antioxidant capacity: A novel indicator of diet quality in healthy young adults. *Journal of the American College of Nutrition*, **28**, 648-656.
- Shapiro, S.S. & Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, **52**, 591-611.

- Singleton, V.L. & Rossi, J.A. (1965). Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *American Journal of Enology and Viticulture*, **16**, 144-158.
- Trygg, J. & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, **16**, 119-128.
- Van Nederkassel, A.M., Daszykowski, M., Massart, D.L. & Vander Heyden, Y. (2005). Prediction of total green tea antioxidant capacity from chromatograms by multivariate modelling. *Journal of Chromatography A*, **1096**, 177-186.
- Verboven, S., Hubert, M. & Goos, P. (2012). Robust preprocessing and model selection for spectral data. *Journal of Chemometrics*, **26**, 282-289.
- Vincent, H.K. & Taylor, A.G. (2006). Biomarkers and potential mechanisms of obesity-induced oxidant stress in humans. *International Journal of Obesity*, **30**, 400-418.
- Wilmsen, P.K., Spada, D.S. & Salvador, M. (2005). Antioxidant activity of the flavonoid hesperidin in chemical and biological systems. *Journal of Agricultural Food Chemistry*, **53**, 4757-4761.
- Wold, S., Albano, C., Dunn, W.J., Esbensen, K., Hellberg, S., Johansson, E. & Sjöström, M. (1983). Pattern recognition: Finding and using patterns in multivariate data. In: *Food Research and Data Analysis* (edited by Martens, H. & Russwurm, H.). Pp. 147-188. London: Applied Science.
- Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: a basic tool for chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**, 109-130.
- Yang, D., An, Y., Jiang, X., Tang, D., Gao, Y., Zhao, H. & Wu, X. (2011). Development of a novel method combining HPLC fingerprint and multi-ingredients quantitative analysis for quality evaluation of traditional chinese medicine preparation. *Talanta*, **85**, 885-890.

Chapter 6

General Discussion and Conclusions

The importance of quality control of herbal products aimed at the nutraceutical and food ingredient markets is increasing due to a changing international regulatory environment (Anon., 2007; Anon., 2013). As nutraceuticals contain substances or chemical compounds that are responsible for promoting health in individuals, the confirmation of activity, or concentrations of the bioactive compounds responsible for the activity at the required level, should be monitored as guarantee for the consumer. Without specific methods to analyse the activity of herbal products, this control remains a challenge. Currently, marker compounds are often used (Tistaert *et al.*, 2011). In certain cases these compounds may only have been chosen as they were unique to a specific herbal product and they may not be the (only) compound(s) responsible for the specific activity. This approach also does not take into account synergistic effects, which should be considered during quality control (Williamson, 2001; Wagner & Ulrich-Merzenich, 2009; Gertsch, 2011). Honeybush tea extract, produced from *Cyclopia* species, has the potential to be used for nutraceutical products due to its high phenolic content, which is assumed to be responsible for bioactivity (Garg *et al.*, 2001; Wauthoz *et al.*, 2007; Joubert *et al.*, 2011). Several *Cyclopia* species are either commercialised or have commercial potential (Joubert *et al.*, 2011).

Qualitative and quantitative variation in phenolic composition between *Cyclopia* species (Ferreira *et al.*, 1998; Joubert *et al.*, 2003; Kamara *et al.*, 2003, 2004; De Beer & Joubert, 2010; Kokotkiewicz *et al.*, 2012, 2013) highlights the need for more stringent quality control. Thus far, total polyphenol content and total antioxidant activity are used for extract standardisation (Joubert *et al.*, 2011). Honeybush products (tea and extracts) are not usually species-specific due to the demand for honeybush exceeding its supply. Consequently, this also has a large effect on honeybush compositional variation.

The World Health Organisation (WHO) has accepted fingerprint analysis for the quality evaluation of natural products (Anon., 1991). It is however only during the last decade that HPLC fingerprint analysis has increasingly been investigated and evaluated as an alternative approach to marker(s) and general quantitative quality control procedures. HPLC fingerprints are potentially more effective than quantitative data as a fingerprint represents a more comprehensive chemical profile of a sample. Fingerprints can therefore in principle account for synergistic effects, leading to more effective quality control (Tistaert *et al.*, 2011).

In order to use honeybush extracts for condition-specific nutraceutical products, they should ideally be screened for the relevant bioactivity. Traditional methods for quality control such as cell and/or animal models are time-consuming and expensive; more rapid screening methods are therefore needed for quality control purposes (Mishra *et al.*, 2008; Tu *et al.*, 2010).

In the present study the focus was on *C. subternata*, a commercialised species, and *C. maculata*, which is being investigated for its commercial potential. While the phenolic composition of *C. subternata* has been investigated (Kamara *et al.*, 2004; De Beer & Joubert, 2010; Kokotkiewicz *et al.*, 2012), very little information is available on the phenolic composition of *C. maculata* (Joubert *et al.*, 2003).

HPLC methods had to be developed for *C. subternata* and *C. maculata* to overcome various limitations of the previously employed method (De Beer & Joubert, 2010). Separate HPLC methods were required for each of the species to enable quantification of a large number of compounds, to improve accuracy, and to produce chromatographic fingerprints suitable for comprehensive analysis. Two HPLC-DAD methods were developed to allow comprehensive characterisation of honeybush phenolics. These methods were also suitable for routine analysis of aqueous extracts of green and fermented *C. subternata* and *C. maculata*. The focus was on aqueous extracts since these are mainly produced by industry. HPLC methods were developed using a methodical approach, changing one parameter at a time, and using the conditions of the previously employed quantification

method for *C. subternata* as a starting point (De Beer & Joubert, 2010). A 3 μm particle size Gemini-NX column provided the best potential for optimal resolution in the retention window where mangiferin and isomangiferin elute. This region was regarded as the most problematic due to a number of unidentified compounds in close proximity, with more co-elution observed on the other columns evaluated. The Gemini-NX column maintained more acceptable pressure stability than a Kinetex superficially porous column throughout method development. The selected column is also stable at extreme mobile phase pH's, and has a longer expected lifetime. The final mobile phase composition and temperature were determined to be a mixture of acetonitrile and 2% acetic acid and 30 °C. These parameters were identical for the methods developed for each species, but a steeper gradient was used for *C. subternata*. This was used to obtain better resolution in the retention time window where mangiferin and isomangiferin eluted without affecting the separation of the remaining compounds. This resulted in total analysis times of 40 and 44 min for *C. subternata* and *C. maculata*, respectively, with good resolution of all the major phenolic compounds. Furthermore, the difficult integration process presented by the previous method as a result of a background 'hump', believed to be caused by uncharacterised polymers, was resolved by the increased gradient length. Compounds such as eriocitrin, scolymoside, phloretin-3',5'-di-*C*-glucoside and 3-hydroxyphloretin-3',5'-di-*C*-hexoside were better separated and the polymeric 'hump' was spread out, resulting in a simpler and more effective integration process. The method for *C. subternata* was therefore improved significantly and, as a result, 10 compounds could be quantified, compared to 9 compounds with the previous method (De Beer & Joubert, 2010). In the case of *C. maculata*, the developed HPLC method is the first reported species-specific routine HPLC-DAD method for the analysis of its extracts.

LC-MS and LC-MS/MS analyses in both positive and negative ionisation modes enabled confirmation of the identities of several phenolic compounds with the aid of authentic standards and literature, and the tentative identification of several additional compounds in *Cyclopia* species for the first time. In *C. subternata* the presence of the xanthenes mangiferin and isomangiferin (Joubert *et al.*, 2003), the flavanones hesperidin (Ferreira *et al.*, 1998; Kamara *et al.*, 2004) and eriocitrin (Kamara *et al.*, 2004), the flavone scolymoside (Kamara *et al.*, 2004), the dihydrochalcone, phloretin-3',5'-di-*C*-glucoside (Kokotkiewicz *et al.*, 2012) and the benzophenone iriflophenone-3-*C*-glucoside (Kokotkiewicz *et al.*, 2012), were confirmed. Furthermore, the benzophenones maclurin-3-*C*-glucoside and iriflophenone-di-*O,C*-hexoside, the flavone apigenin-6,8-di-*C*-glucoside, the flavanones eriodictyol-*C*-glucoside and (*R*)- and (*S*)-eriodictyol-di-*C*-glucoside, and the dihydrochalcone 3-hydroxyphloretin-3',5'-di-*C*-hexoside, were tentatively identified for the first time in *C. subternata*. Among them, iriflophenone-di-*O,C*-hexoside, apigenin-6,8-di-*C*-glucoside, hydroxyphloretin-3',5'-di-*C*-hexoside and eriodictyol-di-*C*-glucoside were reported for the first time in *Cyclopia*, while maclurin-3-*C*-glucoside was recently shown to be present in *C. genistoides* (Kokotkiewicz *et al.*, 2013). For extracts of *C. maculata*, on the other hand, it should be noted that only mangiferin, isomangiferin and hesperidin have previously been identified (Joubert *et al.*, 2003) and their presence were confirmed during this study. The following compounds were reported for the first time in *C. maculata*: maclurin-3-*C*-glucoside, apigenin-6,8-di-*C*-glucoside, eriodictyol-*O*-glucoside, 3-hydroxyphloretin-3',5'-di-*C*-hexoside, eriocitrin, scolymoside, iriflophenone-3-*C*-glucoside and phloretin-3',5'-di-*C*-glucoside, along with the tentative identification of the xanthenes, hydroxymangiferin and hydroxyisomangiferin. These xanthenes were reported for the first time in a *Cyclopia* species.

Interestingly, a compound with m/z 613 $[\text{M-H}]^-$, detected in both species, has previously been identified in *Aspalathus linearis* (rooibos) (Beelders *et al.*, 2012). Retention time, UV-Vis, MS and MS/MS data indicated that this compound was a *C*-5'-hexosyl derivative of aspalathin (3-hydroxyphloretin-3',5'-di-*C*-hexoside). It

should be noted that aspalathin (2',3,4,4',6'-pentahydroxy-3'-*C*- β -D-glucopyranosyldihydrochalcone), unique to rooibos tea, was not observed in any of the *C. subternata* or *C. maculata* extracts analysed. Some similarities between *Cyclopia* species and *A. linearis* could however be expected, as they both originate from the legume family Fabaceae (Joubert *et al.*, 2011). Another similarity included the detection of apigenin-6,8-di-*C*-glucoside (Vicenin-2) in *Cyclopia*, that has previously been identified in *A. linearis* (Beelders *et al.*, 2012).

Compounds in the *C. subternata* and *C. maculata* extracts that could not be identified based on the available UV-Vis, MS and MS/MS data, should be isolated and identified for further characterisation of these species. Furthermore, compounds that were tentatively identified based on the obtained data should also be confirmed by isolating them and performing structural elucidation by means of nuclear magnetic resonance (NMR) or alternatively methods such as capillary NMR (Glauser *et al.*, 2008) or on-line LC-UV-solid-phase extraction-NMR (Exarchou *et al.*, 2006).

In order to employ the developed HPLC-DAD methods for the routine analysis of aqueous extracts from fermented and green *C. subternata* and *C. maculata*, the reliability of the methods had to be determined. Method validation was therefore performed for both HPLC methods. Iriflophenone-3-*C*- β -glucoside and maclurin-3-*C*- β -glucoside could not be quantified in fermented *C. maculata* and *C. subternata* extracts due to low quantities which resulted in poor resolution, and was therefore not included in the validation. Both methods were deemed specific for the quantified peaks, as their UV-Vis and MS spectra matched those of authentic standards or literature values (De Beer & Joubert, 2010; Beelders *et al.*, 2012; Kokotkiewicz *et al.*, 2012, 2013). The linearity of the two methods was excellent, with correlation coefficients all equal to 1.000 and fairly low y-intercepts. The stability of all the compounds over a period of 24-26 h was good. The % change of all compounds was between -5% and 5%, except for 3-hydroxyphloretin-3',5'-di-*C*-hexoside which was slightly out of the acceptable range in *C. subternata* (5.5% in the green extract and 9.1% in the fermented extract). In a previous study, the dihydrochalcone, aspalathin, was found to be unstable in the absence of ascorbic acid (Joubert *et al.*, 2012). Although ascorbic acid was added to these extracts before analysis, the appropriate concentration should be determined in order to avoid the degradation of 3-hydroxyphloretin-3',5'-di-*C*-hexoside, presumably during storage or sample preparation. Furthermore, the %RSD for the stability of all the compounds was acceptable and below 5%. For the precision of the extracts, intra- and inter-day precisions were determined and both methods had RSD values below 5%. Therefore these HPLC-DAD methods were deemed reliable, and could be used for the routine analysis of green and fermented *C. subternata* and *C. maculata* extracts.

Following validation, the methods were employed for the quantitative analysis of a large number of aqueous extracts from green *Cyclopia subternata* (n = 64) and *C. maculata* (n = 40). More compounds were quantified in *C. subternata* than in *C. maculata*, due to higher levels of several compounds in extracts from the former species. As a result 10 compounds, namely maclurin-3-*C*-glucoside, iriflophenone-di-*O,C*-hexoside, iriflophenone-3-*C*-glucoside, mangiferin, isomangiferin, scolymoside, eriocitrin, phloretin-3',5'-di-*C*-hexoside, 3-hydroxyphloretin-3',5'-di-*C*-hexoside and hesperidin could be quantified in extracts of green *C. subternata*, compared to 6 compounds in extracts of green *C. maculata*, namely mangiferin, isomangiferin, maclurin-3-*C*-glucoside, iriflophenone-3-*C*-glucoside, eriocitrin and hesperidin. The results showed large intra-species variation, as %RSD values obtained between seedlings of the same species ranged between 24.93% and 64.14% for *C. subternata* and 12.49% to 40.29% for *C. maculata*. Large inter-species differences were also observed between the two species for the same compounds. For example, mangiferin quantities in *C. subternata* were 0.935 g/100 g on average, compared to 5.874 g/100 g in *C. maculata* extracts. Due to the very low (only traces) concentration

of maclurin-3-*C*- β -glucoside in green *C. subternata*, quantification of this compound is not important in future extracts. The variation observed on inter- and intra-species levels, support previous data (Joubert *et al.*, 2003; Joubert *et al.*, 2008; De Beer & Joubert, 2010), and is an important factor to consider for quality control.

The efficient separation of a large number of phenolic compounds provided by the new HPLC methods for *C. subternata* and *C. maculata* contributed to improved chemical profiles or fingerprints of extracts of these species. These fingerprints met criteria for fingerprint analysis such as good resolution and stability, and reduced sample complexity (Alaerts *et al.*, 2010). The chromatographic fingerprints of each extract could thus serve as the raw data for, among others, similarity analysis. However, prior to any multivariate data analysis of the chromatographic fingerprints, they had to be submitted to suitable pre-treatment. This proved to be a time-consuming process, as each chromatogram matrix (absorbance values across wavelengths vs. time) had to be exported manually from ChemStation software, transformed into a format MATLAB could recognise (comma delimited *.csv file) and imported into MATLAB. This procedure was repeated separately for each injection. Given this time-consuming process, a need exists for an automated exportation process or macro to simplify this step. Pre-treatment of the data also included averaging of each fingerprint data matrix over wavelengths to produce a single vector chromatographic fingerprint for each injection, alignment of chromatograms to ensure that data from corresponding peaks were located in the correct columns (Tomasi *et al.*, 2004), elimination of the background noise and baseline variation (Eilers, 2004) and finally normalisation to correct for small, inevitable instrumental variations and analyst weighing errors (Hendriks *et al.*, 2005). Alignment by correlation optimised warping (COW) was chosen as this technique is popular among researchers, sufficient for alignment of chromatograms, and the script is freely available (Tomasi *et al.*, 2004). However, it was a labour-intensive process, due to the lack of fully automated methods. Fingerprints were aligned, one fingerprint at a time, to a target fingerprint. Although automated methods such as developed by Daszykowski *et al.* (2010) and *icoshift* (Tomasi *et al.*, 2011) are available, these still require after-alignment-inspection to determine whether the automated alignment was sufficient, and manual warping of poorly aligned chromatograms.

For the purpose of similarity analysis, active extracts of both *C. subternata* and *C. maculata* were chosen to serve as references for the respective sample sets. The *C. subternata* extract has been proven effective in lowering elevated blood glucose levels in streptozotocin-induced diabetic Wistar rats and diet-induced diabetic Vervet monkeys (Mose Larsen *et al.*, 2008). Both the *C. subternata* and *C. maculata* extracts demonstrated anti-obesity properties in a cell model, i.e. inhibition of adipogenesis in 3T3-L1 pre-adipocytes (Dudhia *et al.*, 2013). These two extracts were referred to as active.

Similarity analysis, a type of unsupervised data analysis technique, comprised calculating the Pearson's product-moment correlation coefficients (r) between each extract and the reference (active) extract. The objective was to determine which of the chromatographic fingerprints of the newly analysed extracts were most similar to the fingerprint of the active extract of the corresponding *Cyclopia* species, and consequently might possess the same "activity". The prediction of activity in extracts is required for effective quality control of future extracts to be used in nutraceutical products in order to ensure the specific quality of a product.

For *C. subternata*, $n = 15$ extracts out of $n = 64$ were considered similar to the active extract based on their r values which were greater than 0.95 (the selected threshold). The r values of the remaining extracts ranged between 0.65 and 0.95. For *C. maculata* all of the extracts ($n = 40$), including an additional wild-harvested *C. maculata* sample set ($n = 10$), had $r > 0.97$, indicating that they all were potentially "active". This apparent similarity was investigated further for both species by means of principal component analysis (PCA) (Wold *et*

al., 1983), another unsupervised multivariate data analysis technique.

PCA of *C. subternata* provided separation of samples with a similar distribution as was observed for similarity analysis based on r . Extracts that correlated well with the active extract clustered around the active extract, and extracts that correlated poorly were far from the active extract on the PCA score plots. The same observation was made when conducting PCA on the quantitative peak data of *C. subternata*. This was to be expected as the quantified compounds represent the largest variability in the data-set extracted by PCA (Wold *et al.*, 1983). When considering the loadings, clustering of the active extracts were in the direction of compounds iriflophenone-di-*O,C*-hexoside, iriflophenone-3-*C*-glucoside, scolymoside and phloretin-3',5'-di-*C*-glucoside, while extracts with low r values clustered with mangiferin and isomangiferin.

For *C. maculata*, the wild-harvested extracts were immediately separated from the seedlings (grown in pots) on the PCA score plots, due to the lower hesperidin content of the former extracts, probably as a result of a lower stem content in the plant material. De Beer *et al.* (2012) showed that the hesperidin content of *C. subternata* leaves is lower than that of the stems. Furthermore, clustering with the active extract, according to the loadings, were caused by mangiferin and isomangiferin. The clustering of extracts with the active sample enabled a separation of data at an r value of 0.99 and therefore the threshold was increased to 0.99. This resulted in $n = 45$ out of $n = 50$ extracts having $r > 0.99$, and considered potentially "active".

The PCA score plots also allowed the detection of an outlier in the *C. maculata* fingerprint data. After careful scrutiny of the chromatogram, this 'atypical' extract was found to contain an extra peak at an early retention time that none of the other extracts possessed, which was also displayed on the corresponding loadings plot. This outlier was not identified by similarity analysis based on r alone. Due to its retention time and UV-Vis spectrum, this compound is presumed to be iriflophenone-di-*O,C*-hexoside, although this should be confirmed by LC-MS. This compound was not observed in the extract used for HPLC method development and validation, exemplifying the importance of selecting a representative sample for method development. A more representative sample is often only identified after the analysis of numerous samples. Re-validating the method with such a sample will result in better characterisation and quantification abilities.

From these data it was clear that fingerprint data were more effective in detecting differences between extracts than quantitative data. This demonstrates the importance of using more comprehensive representations of extract chemical composition such as fingerprint data. It would be advisable to test extracts identified as potentially "active" to confirm their activity; this could serve as a further validation of fingerprint analysis. For *C. maculata*, where many extracts with $r > 0.99$ were found, it is recommended that extracts that clustered closest to the active extract should be used for validation. If the selected extracts do encompass the same activity, similarity analysis could be used in combination with PCA to select extracts for further biological screening. This may lead to a substantial decrease in the number of samples to be screened in rat and cell models. This method could also aid in the selection of plants to propagate vegetatively for the production of plant material with the specific biological activity. From the perspective of quality control and authentication, PCA is a more effective tool than similarity analysis based purely on r , since the former highlights very small variations in the fingerprint data. The general premise that similar properties can be expected for large r valued fingerprints (Fan *et al.*, 2006; Alaerts *et al.*, 2010) is questioned, as smaller but distinct differences observed in the chromatograms were not detected by similarity analysis. Moreover, it is recommended that, for future investigation, more bioactive samples should be obtained to have a better representation of the bioactive fingerprint. Other multivariate data analysis techniques, such as partial least squares discriminant analysis

(PLS-DA), might be able to determine the underlying pattern or direction in the multivariate chromatographic space most strongly related to the biological activity. Extracts that specifically do not exhibit these activities should also be included in further studies for validation purposes. These recommendations could also aid in the detection and verification of the compounds responsible for the activity, as the results obtained thus far were somewhat contradictory in terms of mangiferin and isomangiferin (active *C. subternata* extracts clustered away from mangiferin and isomangiferin, as opposed to that of *C. maculata*, which was clustered in the direction of mangiferin and isomangiferin). This could then further improve the screening process.

Furthermore, the investigation of the relationship between the total antioxidant capacity (TAC) of extracts or food products and the prevention of oxidative stress has received increasing attention (Espín *et al.*, 2007; Puchau *et al.*, 2009; Hermsdorff *et al.*, 2011; Badhadoran *et al.*, 2012). As the TAC is used as a measure of quality of plant extracts (Bell & Ou, 2007; Ninfali *et al.*, 2009), the ability to identify the phenolic compounds in a herbal product that have the largest influence on the antioxidant activity would contribute greatly to more simplistic and rapid quality control measures. Fingerprints have also been used in other studies aimed at identifying compounds responsible for an activity (antioxidant and/or biological) by incorporating multivariate data analysis techniques (Nguyen Hoai *et al.*, 2009; Kvalheim *et al.*, 2011; Tistaert *et al.*, 2011).

TAC data for honeybush extracts were obtained with the oxygen radical absorbance capacity (ORAC) and 2,2-diphenyl-1-picrylhydrazyl (DPPH) assays in the present study. Data of both assays were available for the *C. subternata* extracts, while only DPPH data were available for *C. maculata* extracts. The aim was to build models that would be able to predict the TAC of new extracts based on their fingerprint data. This research was aimed at identifying compounds (peaks) in the chromatograms that exhibited large positive contributions towards the prediction of TAC.

In order to achieve these goals, additional pre-treatment of the chromatographic fingerprint data was required. This included averaging of duplicate injections of each sample, and dividing the chromatographic fingerprints and corresponding TAC values from each set, into calibration and validation sub-sets.

Models were successfully constructed using partial least squares (PLS) regression, with root mean square error values ranging from 4.5% to 8.3%. These error percentages were comparable to previous work of similar nature (Van Nederkassel *et al.*, 2005; Daszykowski *et al.*, 2007). The models were able to predict the TAC values of new extracts, as validated with an external validation data set. Identification of compounds that contributed most to the prediction of TAC proved to be a difficult task. Variables of importance for projection (VIP) scores indicated that the majority of the chromatographic peaks contributed to the construction of the prediction models. However, the regression coefficients suggested that isomangiferin and apigenin-6,8-di-*C*-glucoside were considered important for predicting the TAC in the TAC_{DPPH} models of both species, while isomangiferin was considered important for the prediction of TAC_{ORAC} of *C. subternata*. Other compounds of interest, which presented small peaks in the HPLC-DAD chromatograms and yet had high VIP scores and positive regression coefficients, were highlighted for future work. These compounds should be identified with the use of more extensive mass spectrometric techniques, as they could not be successfully identified during this study.

To conclude, two HPLC-DAD methods were successfully developed and validated for the routine analysis and quantification of green and fermented *C. subternata* and *C. maculata* extracts. LC-MS and -MS/MS analysis allowed the identification of several compounds, including the tentative identification of new compounds in these *Cyclopia* species. The developed HPLC methods were applied to numerous extracts of unfermented (green) plant material from both species to obtain quantitative data. The first quantitative data for several *C. maculata*

compounds, including maclurin-3-*C*-glucoside, eriocitrin, and iriflophenone-3-*C*-glucoside were recorded here. These methods could serve as the basis for future research on bioactivity, sensory characterisation of taste-impact compounds, process optimisation, and the effect of environmental conditions on phenolic composition, among others. Furthermore, these methods provided suitable fingerprints for further multivariate data analysis which was applied to the data after extensive pre-treatment of the chromatographic fingerprints. Similarity analysis of the extracts using a bioactive extract as reference resulted in the identification of several extracts as potentially "active". However, confirmation on the activity of these selected extracts in appropriate cell and/or animal models are required. If this screening tool proves to be successful, it can also be applied to other species or herbal products, as the suggested screening method does not require identification or quantification of compounds, and consequently does not require authentic standards. Using fingerprints to predict the antioxidant capacity of *C. subternata* and *C. maculata* extracts was also successful, although the multivariate data analysis techniques could not identify marker compounds. As this study also forms part of a long-term investigation to develop cost-effective quality control and rapid screening methods, the information presented here could form the basis of further research. This should include confirmation of the compounds responsible for specific health promoting activities in *Cyclopia*, as well as investigation of the relevance of similar fingerprint analysis on other *Cyclopia* species.

References

- Anonymous (1991). Guidelines for the assessment of herbal medicines, World Health Organisation, Geneva, Switzerland. http://whqlibdoc.who.int/hq/1991/WHO_TRM_91.4.pdf (February 2013).
- Anonymous (2007). Corrigendum to Regulation (EC) No 1924/2006 of the European Parliament and of the council of 20 December 2006 on nutrition and health claims made on foods. *Official Journal of the European Union*, Article 13.
- Anonymous (2013). Dietary Supplements. U.S. Food and Drug Administration. [WWW document]. URL <http://www.fde.gov/Food/DietarySupplements/default.htm>. August 2013.
- Alaerts, G., Dejaegher, B., Smeyers-Berbeke, J. & Vander Heyden, Y. (2010). Recent developments in chromatographic fingerprints from herbal products: Set-up and data analysis. *Combinatorial Chemistry and High Throughput Screening*, **13**, 900-922.
- Bahadoran, Z., Golzarand, M., Mirmiran, P., Shive, N. & Azizi, F. (2012). Dietary total antioxidant capacity and the occurrence of metabolic syndrome and its components after 3-year follow-up in adults: Tehran lipid and glucose study. *Nutrition and Metabolism*, **9**, 70-78.
- Beelders, T., Sigge, G.O., Joubert, E., de Beer, D. & de Villiers, A. (2012). Kinetic optimisation of the reversed phase liquid chromatographic separation of rooibos tea (*Aspalathus linearis*) phenolics on conventional high performance liquid chromatographic instrumentation. *Journal of Chromatography A*, **1219**, 128-139.
- Bell, D. & Ou, B. (2007). The branding of ORAC: where will it lead? *Nutraceuticals World*, **10**(3), 72-74.
- Daszykowski, M., Vander Heyden, Y., Boucon, C. & Walczak, B. (2010). Automated alignment of one-dimensional chromatographic fingerprints. *Journal of Chromatography A*, **1217**, 6127-6133.
- Daszykowski, M., Vander Heyden, Y. & Walczak, B. (2007). Robust partial least squares model for prediction of green tea antioxidant capacity from chromatograms. *Journal of Chromatography A*, **1176**, 12-18.
- De Beer, D. & Joubert E. (2010). Development of HPLC method for *Cyclopia subternata* phenolic compound analysis and application to other *Cyclopia* spp. *Journal of Food Composition and Analysis*, **23**, 289-297.
- De Beer, D., Schulze, A.E., Joubert, E., De Villiers, A., Malherbe, C.J. & Stander, M.A. (2012). Food ingredients extracts of *Cyclopia subternata* (honeybush): Variation in phenolic composition and antioxidant capacity. *Molecules*, **17**, 14602-14624.
- Dudhia, Z., Louw, J., Muller, C., Joubert, E., de Beer, D., Kinnear, C. & Pheiffer, C. (2013). *Cyclopia maculata* and *Cyclopia subternata* (honeybush tea) inhibits adipogenesis in 3T3-L1 pre-adipocytes. *Phytomedicine*, **20**, 401-408.
- Eilers, P.H.C. (2004). Parametric time warping. *Analytical Chemistry*, **76**, 404-411.
- Espín, J.C., García-Conesa, M.T. & Tomás-Barberán, F.A. (2007). Nutraceuticals: Fact or fiction. *Phytochemistry*, **68**, 2986-3008.
- Exarchou, V., Fiamegos, Y.C., van Beek, T.A., Nanos, C. & Vervoort, J. (2006). Hyphenated chromatographic techniques for the rapid screening and identification of antioxidants in methanolic extracts of pharmaceutically used plants. *Journal of Chromatography A*, **1112**, 293-302.
- Fan, X., Cheng, Y., Ye, Z., Lin, R. & Qian, Z. (2006). Multiple chromatographic fingerprinting and its application to the quality control of herbal medicines. *Analytica Chimica Acta*, **555**, 217-224.
- Ferreira, D., Kamara, B.I., Brandt, E.V. & Joubert, E. (1998). Phenolic compounds from *Cyclopia intermedia* (honeybush tea). *Journal of Agricultural and Food Chemistry*, **46**, 3406-4310.

- Garg, A., Garg, S., Zaneveld, L.J.D. & Singla, A.K. (2001). Chemistry and pharmacology of the citrus bioflavonoid hesperidin. *Phytotherapy Research*, **15**, 655-669.
- Gertsch, J. (2011). Botanical drugs, synergy, and network pharmacology: Forth and back to intelligent mixtures. *Planta Medica*, **77**, 1086-1098.
- Glauser, G., Guillarme, D., Grata, E., Bocchard, J., Thiocone, A., Carrupt, P., Veuthey, J., Rudaz, S. & Wolfender, J. (2008). Optimized liquid chromatography-mass spectrometry approach for the isolation of minor stress biomarkers in plant extracts and their identification by capillary nuclear magnetic resonance. *Journal of Chromatography A*, **1180**, 90-98.
- Hendriks, M.M.W.B., Cruz-Juarez, L., De Bont, D. & Hall, R.D. (2005). Preprocessing and exploratory analysis of chromatographic profiles of plant extracts. *Analytica Chimica Acta*, **545**, 53-64.
- Hermesdorff, H.H.M., Puchau, B., Volp, A.C.P., Barbosa, K.B.F., Bressan, J., Zulet, M.A. & Martínez, J.A. (2011). Dietary total antioxidant capacity is inversely related to central adiposity as well as to metabolomic and oxidative stress markers in healthy young adults. *Nutrition and Metabolism*, **8**(59), doi:10.1186/1743-7075-8-59.
- Joubert, E., Beelders, T., De Beer, D., Malherbe, C.J., De Villiers, A. & Sigge, G.O. (2012). Variation in phenolic content and antioxidant activity of fermented rooibos herbal tea infusions: Role of production season and quality grade. *Journal of Agricultural and Food Chemistry*, **60**, 9171-9179.
- Joubert, E., Joubert, M.E., Bester, C., De Beer, D. & De Lange, J.H. (2011). Honeybush (*Cyclopia* spp.): From local cottage industry to global markets - The catalytic and supporting role of research. *South African Journal of Botany*, **77**, 877-907.
- Joubert, E., Otto, F., Grüner, S. & Weinreich, B. (2003). Reversed-phase HPLC determination of mangiferin, isomangiferin and hesperidin in *Cyclopia* and the effect of harvesting date on the phenolic composition of *C. genistoides*. *European Food and Research Technology*, **216**, 270-273.
- Joubert, E., Richards, E.S., Van der Merwe, J.D., De Beer, D. Manley, M. & Gelderblom, W.C.A. (2008). Effect of species variation and processing on phenolic composition and in vitro antioxidant activity of aqueous extracts of *Cyclopia* spp. (honeybush tea). *Journal of Agricultural and Food Chemistry*, **56**, 954-963.
- Kamara, B.I., Brand, D.J., Brandt, E.V. & Joubert, E. (2004). Phenolic metabolites from honeybush tea (*Cyclopia subternata*). *Journal of Agricultural and Food Chemistry*, **52**, 5291-5395.
- Kamara, B.I., Brandt, E.V., Ferreira, D. & Joubert, E. (2003). Polyphenols from honeybush tea (*Cyclopia intermedia*). *Journal of Agricultural and Food Chemistry*, **51**, 3874-3979.
- Kokotkiewicz, A., Luczkiewicz, M., Sowinski, P., Glod, D., Gorynski, K. & Bucinski, A. (2012). Isolation and structure elucidation of phenolic compounds from *Cyclopia subternata* Vogel (honeybush) intact plant and *in vitro* cultures. *Food Chemistry*, **133**, 1373-1382.
- Kokotkiewicz, A., Luczkiewicz, M., Pawlowska, J., Luczkiewicz, P., Sowinski, P., Witkowski, J., Bryl, E. & Bucinski, A. (2013). Isolation of xanthone and benzophenone derivatives from *Cyclopia genistoides* (L.) Vent. (honeybush) and their pro-apoptotic activity in synoviocytes from patients with rheumatoid arthritis. *Fitoterapia*, **90**, 199-208.
- Kvalheim, O.M., Chan, H., Benzie, I.F.F., Szetto, Y., Hing-chung Tzang, A., Kam-wah Mok, D. & Chau, F. (2011). Chromatographic profiling and multivariate analysis for screening and quantifying the contributions from individual components to the bioactive signature in natural products. *Chemometric and Intelligent Laboratory Systems*, **107**, 98-105.

- Mishra, K.P., Ganju, L., Sairam, M., Banerjee, P.K. & Sawhney, R.C. (2008). A review of high throughput technology for screening of natural products. *Biomedicine and Pharmacotherapy*, **62**, 94-98.
- Mose Larsen, P., Fey, S.J., Louw, J. & Joubert, E. (2008). An anti-diabetic extract of honeybush. World Intellectual Property Organization WO. 2008/110552A2.
- Nguyen Hoai, N., Dejaegher, B., Tistaert, C., Nguyen Thi Hong, V., Rivière, C., Chataigné, G., Phan Van, K., Chau Van, M., Quentin-Leclercq, J. & Vander Heyden, Y. (2009). Development of HPLC fingerprints for *Mallotus* species extracts and evaluation of peaks responsible for their antioxidant activity. *Journal of Pharmaceutical and Biomedical Analysis*, **50**, 753-763.
- Ninfali, P., Gennari, L. & Biagotti, E. (2009). Improvement in botanical standardisation of commercial freeze-dried herbal extracts by using the combination of antioxidant capacity and constituent marker concentrations. *Journal of AOAC International*, **92**, 797-805.
- Puchau, B., Zulet, M.A., de Echavarri, A.G., Hermsdorff, H.H.M. & Martinez, J.A. (2009). Dietary total antioxidant capacity: A novel indicator of diet quality in healthy young adults. *Journal of the American College of Nutrition*, **28**, 648-656.
- Tistaert, C., Dejaegher, B. & Vander Heyden, Y. (2011). Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Analytica Chimica Acta*, **690**, 148-161.
- Tomasi, G., van der Berg, F. & Andersson, C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics*, **18**, 231-241.
- Tomasi, G., Savaroni, F. & Engelsen, S.B. (2011). *icoshift*: An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, **1218**, 7832-7840.
- Tu, Y., Jeffries, C., Ruan, H., Nelson, C., Smithson, D., Shelat, A.A., Brown, K.A., Li, X., Hester, J.P., Smillie, T., Khan, I.A., Walker, L., Guy, K. & Yan, B. (2010). Automated high-throughput system to fractionate plant natural products for drug discovery. *Journal of Natural Products*, **73**, 751-754.
- Van Nederkassel, A.M., Daszykowski, M., Massart, D.L. & Vander Heyden, Y. (2005). Prediction of total green tea antioxidant capacity from chromatograms by multivariate modeling. *Journal of Chromatography A*, **1096**, 177-186.
- Wagner, H. & Ulrich-Merzenich, G. (2009). Synergy research: Approaching a new generation of phytopharmaceuticals, *Phytomedicine*, **16**, 97-110.
- Wauthoz, N., Balde, A., Balde, E.S., Van Damme, M. & Duez, P. (2007). Ethnopharmacology of *Mangifera indica* L. bark and pharmacological studies of its main C-glucosylxanthone, mangiferin. *International Journal of Biomedical and Pharmaceutical Sciences*, **1**, 112-119.
- Wold, S., Albano, C., Dunn, W.J., Esbensen, K., Hellberg, S., Johansson, E. & Sjöström, M. (1983). Pattern recognition: Finding and using patterns in multivariate data. In: *Food Research and Data Analysis* (edited by Martens, H. & Russwurm, H.). Pp. 147-188. London: Applied Science.
- Williamson, E.M. (2001). Synergy and other interactions in phytomedicine. *Phytomedicine*, **8**, 401-409.

Appendix A

Quantitative data of *Cyclopia subternata* and *Cyclopia maculata*

Table A.1 – Quantitative data from extracts of unfermented (green) *C. subternata* stems and leaves (n = 64)

Compound	g/100 g soluble solids				%RSD
	Average	Minimum	Maximum	Std dev	
Mangiferin	0.935	0.397	2.174	0.421	45.054
Isomangiferin	0.358	0.149	0.688	0.115	32.037
Scolymoside	0.455	0.115	1.269	0.244	53.554
Eriocitrin	0.550	0.265	0.954	0.154	28.091
Hesperidin	0.641	0.281	2.118	0.380	56.220
Iriflophenone-3- <i>C</i> - β -glucoside	0.583	0.120	1.938	0.360	61.847
Maclurin-3- <i>C</i> - β -glucoside	0.034	traces	0.117	0.022	64.139
Iriflophenone-di- <i>O,C</i> -hexoside	0.519	0.129	1.772	0.287	55.319
3-Hydroxyphloretin-3',5'-di- <i>C</i> -hexoside	0.536	0.256	1.092	0.134	24.930
Phloretin-3',5'-di- <i>C</i> - β -glucoside	1.047	0.273	2.183	0.342	32.674

%RSD = relative standard deviation

Std dev = standard deviation

Table A.2 – Quantitative data from extracts of unfermented (green) *C. subternata* stems (n = 10)

Compound	g/100 g soluble solids				%RSD
	Average	Minimum	Maximum	Std dev	
Mangiferin	0.373	0.188	0.558	0.138	36.906
Isomangiferin	0.156	0.099	0.223	0.042	27.178
Scolymoside	0.391	0.152	0.826	0.242	61.966
Eriocitrin	0.517	0.344	0.744	0.127	24.512
Hesperidin	1.559	1.164	1.893	0.289	18.534
Iriflophenone-3- <i>C</i> - β -glucoside	0.359	0.226	0.576	0.097	26.914
Maclurin-3- <i>C</i> - β -glucoside	0.001	0.000	0.013	0.004	316.228
Iriflophenone-di- <i>O,C</i> -hexoside	0.556	0.184	1.119	0.268	48.267
3-Hydroxyphloretin-3',5'-di- <i>C</i> -hexoside	0.493	0.311	0.597	0.101	20.489
Phloretin-3',5'-di- <i>C</i> - β -glucoside	1.243	0.878	1.525	0.237	19.064

%RSD = relative standard deviation

Std dev = standard deviation

Table A.3 – Quantitative data from extracts of unfermented (green) *C. subternata* leaves (n = 10)

Compound	g/100 g soluble solids				%RSD
	Average	Minimum	Maximum	Std dev	
Mangiferin	0.817	0.313	1.396	0.359	43.943
Isomangiferin	0.342	0.177	0.489	0.104	30.402
Scolymoside	0.812	0.264	1.443	0.454	55.852
Eriocitrin	0.633	0.422	1.003	0.177	27.927
Hesperidin	0.504	0.147	1.517	0.495	98.253
Iriflophenone-3- <i>C</i> - β -glucoside	0.552	0.330	0.795	0.165	29.812
Maclurin-3- <i>C</i> - β -glucoside	0.031	0.017	0.058	0.014	0.495
Iriflophenone-di- <i>O,C</i> -hexoside	0.610	0.178	1.168	0.324	53.165
3-Hydroxyphloretin-3',5'-di- <i>C</i> -hexoside	0.432	0.332	0.561	0.073	16.991
Phloretin-3',5'-di- <i>C</i> - β -glucoside	0.899	0.631	1.364	0.252	28.049

%RSD = relative standard deviation

Std dev = standard deviation

Table A.4 – Quantitative data from extracts of unfermented (green) *C. maculata* (pot trial) stems and leaves (n = 40)

Compound	g/100 g soluble solids				
	Average	Minimum	Maximum	Std dev	%RSD
Mangiferin	5.874	4.024	8.091	0.959	16.326
Isomangiferin	1.646	1.225	2.131	0.205	12.485
Eriocitrin	0.436	0.246	0.690	0.105	24.013
Hesperidin	1.782	1.442	2.686	0.230	12.897
Iriflophenone-3- <i>C</i> - β -glucoside	0.890	0.361	2.125	0.335	37.641
Maclurin-3- <i>C</i> - β -glucoside	0.221	0.094	0.500	0.089	40.291

%RSD = relative standard deviation
Std dev = standard deviation

Table A.5 – Quantitative data from extracts of unfermented (green) *C. maculata* (wild-harvested) stems and leaves (n = 10)

Compound	g/100 g soluble solids				
	Average	Minimum	Maximum	Std dev	%RSD
Mangiferin	5.059	4.011	6.043	0.599	11.846
Isomangiferin	1.454	1.213	1.1713	0.141	9.671
Eriocitrin	0.334	0.307	0.384	0.022	6.595
Hesperidin	0.892	0.746	0.987	0.085	9.564
Iriflophenone-3- <i>C</i> - β -glucoside	0.542	0.448	0.673	0.077	14.184
Maclurin-3- <i>C</i> - β -glucoside	0.115	0.079	0.144	0.020	17.750

%RSD = relative standard deviation
Std dev = standard deviation

Appendix B

Typical and 'atypical' *C. maculata* extract

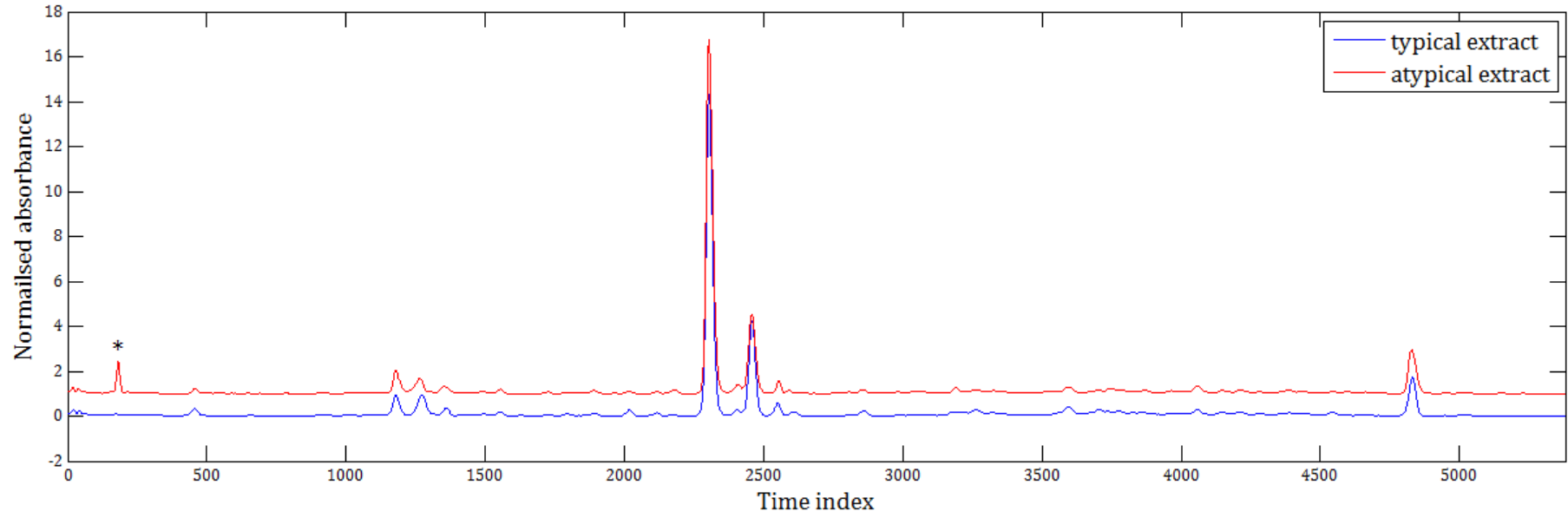


Figure B.1 – Chromatographic fingerprints of a typical and the ‘atypical’ *C. maculata* extract (sample 24) as used for data analysis; extra peak indicated with black star symbol (*).

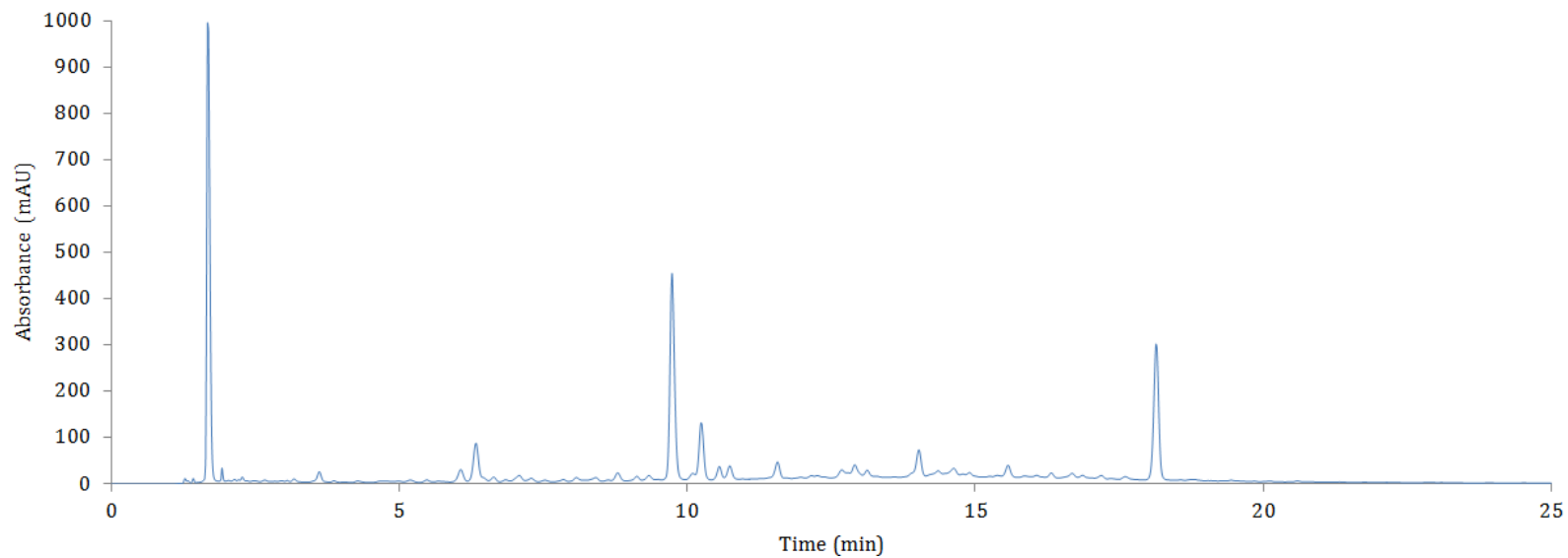


Figure B.2 – Chromatogram of a typical *C. maculata* extract at 288 nm.

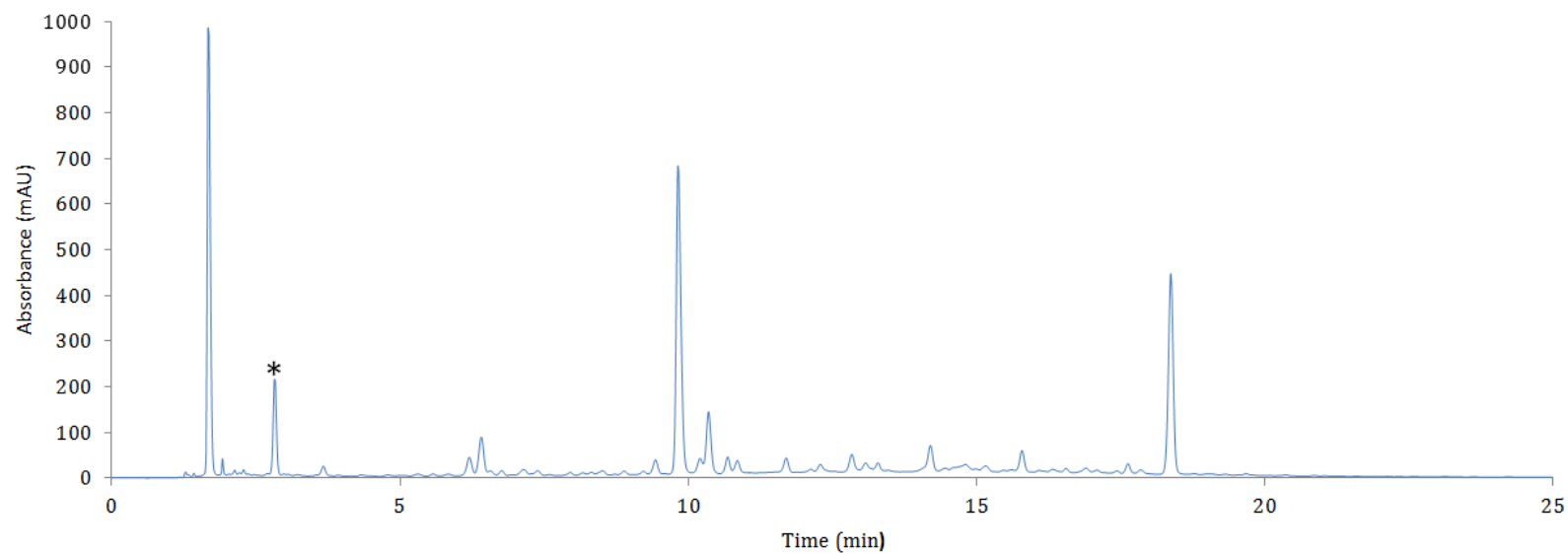


Figure B.3 – Chromatogram of the 'atypical' *C. maculata* extract (sample 24) at 288 nm; extra peak indicated with black star symbol (*).

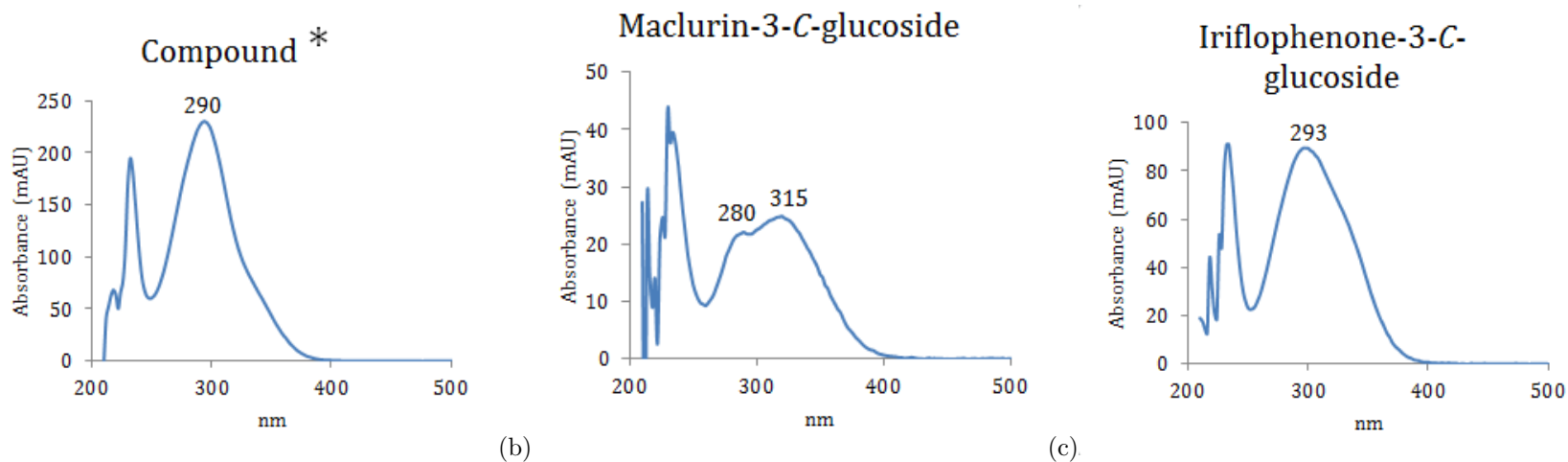


Figure B.4 – UV-spectra of benzophenones (a) compound * (extra peak, presumably iriflophenone-di-*O,C*-hexoside), (b) maclurin-3-*C*-glucoside and (c) iriflophenone-3-*C*-glucoside from the 'atypical' *C. maculata* extract (sample 24).

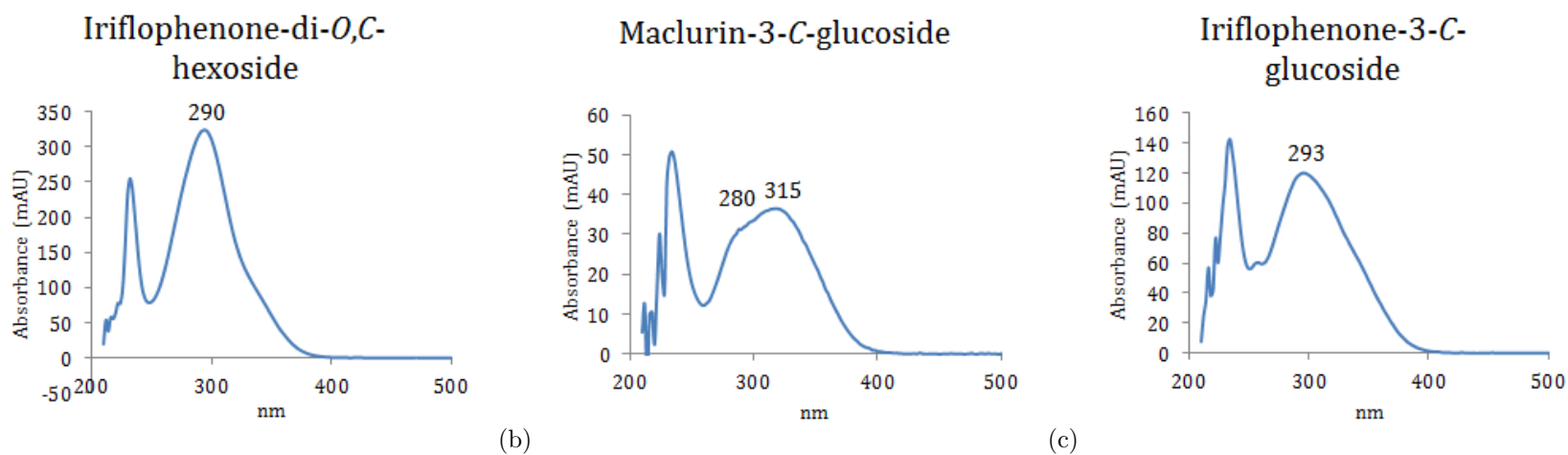


Figure B.5 – UV-spectra of benzophenones (a) iriflophenone-di-*O,C*-hexoside, (b) maclurin-3-*C*-glucoside and (c) iriflophenone-3-*C*-glucoside from a *C. subternata* extract.