SPECTROSCOPIC AND CHROMATOGRAPHIC APPROACH FOR THE CHEMICAL FINGERPRINTING OF *GELSEMIUM ELEGANS* VIA HPTLC, LC-MS, FT-IR AND NMR

NG CHIEW HOONG

UNIVERSITI SAINS MALAYSIA

2019

SPECTROSCOPIC AND CHROMATOGRAPHIC APPROACH FOR THE CHEMICAL FINGERPRINTING OF *GELSEMIUM ELEGANS* VIA HPTLC, LC-MS, FT-IR AND NMR

by

NG CHIEW HOONG

Thesis submitted in fulfillment of the requirements

for the degree of

Doctor of Philosophy

June 2019

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincerest gratitude to my supervisor, Dr. Yam Mun Fei, who has supported me throughout my project with his patience, advices and knowledge. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. He has inspired me to become an independent researcher and helped me realize the power of critical reasoning. He also demonstrated what a brilliant and hard-working scientist can accomplish. This work would not have been possible without his guidance, support and encouragement. Under his guidance I successfully overcame many difficulties and learned a lot.

My sincere thanks also go to my lab mates for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last four years. They have been more than helpful with their guidance and elaborate explanations on how to properly execute the methodology of the study. The experiences and knowledge I gained throughout the process of completing this research would prove invaluable to better equip me for the challenges which lie ahead.

I gratefully acknowledge all the lecturers and professionals especially Dr. Teh Chin Hoe who opened my horizons to the wonderful world of NMR spectroscopy and guided me constantly. Thanks for also teaching me how to perform the related instrument and analyzing the data during this project. I would also like to extend my sincere gratitude towards the students, lab assistants, and staffs of the Department of Pharmacology and Pharmaceutical Chemistry in School of Pharmaceutical Sciences, Centre of Herbal Characterization and Standardization (CHEST), Malaysian Institute

of Pharmaceuticals and Nutraceuticals (IPharm) and College of Pharmacy in Fujian University of Traditional Chinese Medicine. Thank you very much for guiding me in the usage of the laboratory instruments and the cooperation provided during my research.

Last but not least, I would also like to take this opportunity to sincerely acknowledge my family and friends for the support they provided me through my entire life and in particular, I must acknowledge my parents because without those love, encouragement and motivation, I would not have finished this thesis.

TABLE OF CONTENTS

Acknowledgement	ii
Table of Contents	iv
List of Tables	x
List of Figures	xii
List of Abbreviations	xvii
List of Symbols	xx
List of Appendices	xxii
Abstrak	xxiii
Abstract	xxv

CHAPTER 1 – INTRODUCTION

1.1 Overview of TCM	1
1.2 Importance of Chinese Herbal Medicines (CHM)	2
1.3 Issues of CHM	4
1.3.1 CHM Adulteration	4
1.3.1(a) Direct or deliberate adulteration	5
1.3.1(b) Indirect or unintentional adulteration	5
1.4 Quality Control of CHM	6
1.5 Analytical methods for quality control of herbal medicines	8
1.5.1 The 'Marker-based' Approach to Quality Control	8
1.5.2 The 'Fingerprinting' Approach to Quality Control	10
1.5.3 Spectroscopic Fingerprinting	12
1.5.3(a) Fourier Transform Infrared (FT-IR)	12
1.5.3(b) Nuclear Magnetic Resonance (NMR)	14
1.5.4 Chromatographic Fingerprinting	17
1.5.4(a) Ultra Performance Liquid Chromatography (UPLC)	17
1.5.4(b) High Performance Thin Layer Chromatography (HPTLC)	20
1.6 Multivariate Analysis	22
1.6.1 Unsupervised Method	23
1.6.2 Supervised Methods	24

1.7 Overview of *Gelsemium elegans* 25
1.8 Overview of the Reference Compounds 29
1.8.1 Gelsemine 29
1.8.2 Koumine 30
1.9 USP Standard 30
1.9.1 Sampling 31
1.9.1(a) Gross Sample 31
1.9.1(b) Laboratory Sample 31
1.9.1(c) Test Sample 31
1.9.2 Instruments 32
1.10 Problem statement and the research objectives 32

CHAPTER 2 – FINGERPRINTING OF *GELSEMIUM ELEGANS* AND QUANTIFICATION OF KOUMINE WITH HIGH PERFORMANCE THIN LAYER CHROMATOGRAPHY (HPTLC)

2.1 Introduction 33
2.2 Chemicals and instruments 34
2.3 Methodology 35
2.3.1 Preparation of reference solution and internal standard solution 35
2.3.2 Sample Preparation 35
2.3.3 Sample application 36
2.3.4 Chromatography 36
2.3.5 Plate evaluation 36
2.4 Method Development 37
2.4.1 Extraction Solvents Screening Method 37
2.4.2 Mobile Phase Development 38
2.4.3 Derivatizing Reagent 38
2.4.4 Method Validation 39
2.5 Results and Discussion 40
2.5.1 Method Development 40
2.5.1(a) Extraction Solvents Screening Method 40

2.5.1(b) Mobile Phase Development 46
2.5.1(c) Variations of the chloroform-methanol-water ratio 50
2.5.1(d) Derivatizing Reagent 52
2.5.2 Fingerprint of *Gelsemium elegans* 56
2.5.3 Method Validation 60
2.5.3(a) Precision 60
2.5.3(b) Reproducibility 60
2.5.3(c) Working range, Limit of Detection (LOD) and Limit of Quantification (LOQ) 60
2.5.3(d) Robustness 63
2.5.3(d)(i) Volume of Developing Solvent 63
2.5.3(d)(ii) Equilibration Time 63
2.5.3(d)(iii) Dosage Speed of Sample Applicator 64
2.5.3(e) Specificity 64
2.5.3(f) Recovery studies (Accuracy) 65
2.5.4 Quantification of *Gelsemium elegans* 66
2.6 Conclusion 73

CHAPTER 3 – FINGERPRINTING OF *GELSEMIUM ELEGANS* WITH UPLC-PDA-QDA AND QUANTIFICATION OF GELSEMINE AND KOUMINE WITH UPLC-ESI-MS/MS

3.1 Introduction 74
3.2 Chemicals and instruments 75
3.3 Methodology 75
3.3.1 Preparation of reference solution and internal standard solution 75
3.3.2 Plant materials and sample preparation 76
3.3.3 UPLC-PDA-QDa conditions 77
3.3.4 Method Development for UPLC-PDA-QDa 78
3.3.4 (a) Optimization of the UPLC Conditions 78
3.3.4(a)(i) Optimization of the preparation methods for the sample solution 78

3.3.4(a)(ii) Optimization of the chromatographic conditions 79
3.3.4(a)(iii) Optimization of the column temperature and detection wavelength 79
3.3.5 UPLC-ESI-MS/MS conditions 79
3.3.6 UPLC-ESI-MS/MS method development 80
3.3.7 Method Validation for UPLC-ESI-MS/MS 81
3.3.7(a) Linearity and detection limit 81
3.3.7(b) Precision 81
3.3.7(c) Recovery and matrix effect 81
3.3.7(d) Stability 82
3.3.7(e) Repeatability 82
3.3.8 Principal Component Analysis (PCA) 82
3.4 Results and Discussion 83
3.4.1 Method Development for UPLC-PDA-QDa 83
3.4.1(a) Optimization of the preparation methods for the sample solution 83
3.4.1(b) Optimization of the chromatographic conditions 85
3.4.1(c) Optimization of the column temperature 87
3.4.1(d) Optimization of the detection wavelength 88
3.4.2 Fingerprint of *Gelsemium elegans* with UPLC-PDA-QDa 89
3.4.3 UPLC-ESI-MS/MS method development 93
3.4.3(a) Optimization of the chromatographic conditions 93
3.4.3(b) Optimization of the mass detector 96
3.4.4 Method Validation for UPLC-ESI-MS/MS 102
3.4.4(a) Linearity and detection limit 102
3.4.4(b) Precision 103
3.4.4(c) Recovery and matrix effect 104
3.4.4(d) Stability 104
3.4.4(e) Repeatability 105
3.4.5 Quantification of *Gelsemium elegans* with UPLC-ESI-MS/MS 105

3.4.6 Principal Component Analysis (PCA) 113
3.5 Conclusion 116

CHAPTER 4 – APPLICATION OF MID-INFARED SPECTROSCOPY WITH MULTIVARIATE ANALYSIS FOR THE DISCRIMINATION OF *GELSEMIUM ELEGANS*

4.1 Introduction 118
4.2 Chemicals and Instruments 119
4.3 Methodology 120
4.3.1 Samples and Materials 120
4.3.2 Procedure of FT-IR spectral acquisition 120
4.3.3 Reproducibility of the Infrared Spectra 120
4.3.4 Data Processing 121
4.4 Results and Discussion 121
4.4.1 Differentiation by FT-IR spectra 121
4.4.2 Differentiation by second derivative IR spectra 125
4.4.3 Differentiation by 2D-IR spectra 128
4.4.4 Reproducibility of the Infrared Spectra 133
4.4.5 Combination of FT-IR spectra and chemometrics method for discrimination of different parts of *Gelsemium elegans* plant 135
4.5 Conclusion 138

CHAPTER 5 – APPLICATION OF NUCLEAR MAGNETIC RESONANCE WITH MULTIVARIATE ANALYSIS FOR THE DISCRIMINATION OF *GELSEMIUM ELEGANS*

5.1 Introduction 139
5.2 Chemicals and Instruments 140
5.3 Methodology 140
5.3.1 Samples and Materials 140
5.3.2 Procedure of FT-IR spectral acquisition 141
5.3.3 NMR data reduction and preprocessing 141
5.3.4 Validation of the NMR Spectra 142
5.4 Results and Discussion 142

5.4.1 General Experimental Considerations 142
5.4.2 Validation of the NMR Spectra 143
5.4.3 Fingerprint of *Gelsemium elegans* 144
5.4.4 Principal component analysis of 1H NMR spectra 149
5.5 Conclusion 154

CHAPTER 6 – SUMMARY AND CONCLUSION

6.1 Summary of the Comparison of the Fingerprint of the Different Parts of *Gelsemium elegans* via HPTLC, UPLC-PDA-QDa, FT-IR and NMR 155
6.2 Summary of the Comparison of the Qualification of Reference Compounds in the Different Parts of *Gelsemium elegans* via HPTLC, UPLC-PDA-QDa and UPLC-ESI-MS/MS 155
6.3 Summary of the Comparison of the Quantitative Amount of Koumine in the Different Parts of *Gelsemium elegans* via HPTLC and UPLC-ESI-MS/MS 156
6.4 Summary of the Comparison of Chemometrics Analysis of *Gelsemium elegans* via UPLC-PDA, FT-IR and NMR 157
6.5 Conclusion 158
6.6 Future Work Recommendation 160

REFERENCES 161

APPENDICES

LIST OF PUBLICATIONS AND PRESENTATIONS

LIST OF TABLES

		Page
Table 2.1	*Rf* value of koumine with chloroform-methanol-water in different development ratios.	51
Table 2.2	Relative Standard Deviation expressed in percentages of known concentrations of koumine in intra-day precision and inter-day reproducibility tests.	60
Table 2.3	Equation and R^2 for the calibration curves of the 28 plates.	62
Table 2.4	Changes in the fingerprint analysis of the tested fractions as function of volume of developing solvent.	63
Table 2.5	Changes in the fingerprint analysis of the tested fractions as function of equilibration time of gas phase and stationary phase.	63
Table 2.6	Changes in the fingerprint analysis of the tested fractions as function of dosage speed of sample applicator.	64
Table 2.7	Results of Recovery Tests for koumine (n=3)	65
Table 2.8	The mean amount of koumine in *Gelsemium elegans* sample and their respective RSD and correlation coefficient with pure koumine.	67
Table 3.1	The molecular formula, retention time, MRM transition, cone voltage, collision energy and ESI of gelsemine and koumine.	97
Table 3.2	Calibration ranges, regression equation, correlation of determination, LOD, LOQ and RSD % obtained for the regression lines.	103
Table 3.3	Precision of gelsemine and koumine.	103
Table 3.4	Results of Recovery Tests for gelsemine and koumine (n=3).	104
Table 3.5	Stability of gelsemine and koumine.	104
Table 3.6	Repeatability of gelsemine and koumine.	105
Table 3.7	The average amount and RSD % of gelsemine and koumine detected in the samples.	107

Table 4.1	Peak assignments on the conventional FT-IR spectra of the different parts of *Gelsemium elegans.*	122
Table 4.2	The intermaterial distances of the SIMCA model of different parts of *Gelsemium elegans* samples.	138
Table 4.3	The recognition and rejection rates of the SIMCA model of different parts of *Gelsemium elegans* samples.	138

LIST OF FIGURES

		Page
Figure 1.1	The whole plant of *Gelsemium elegans.*	26
Figure 1.2	The stem of *Gelsemium elegans.*	26
Figure 1.3	The root of *Gelsemium elegans.*	27
Figure 1.4	The leaf of *Gelsemium elegans.*	27
Figure 1.5	Structure of gelsemine.	29
Figure 1.6	Structure of koumine.	30
Figure 2.1	Initial screening on *Gelsemium elegans* extracts with: Mobile phase: (a, b) toluene-ethyl acetate (95:5); (c, d) chloroform-methanol-water (70:30:4); (e, f) ethyl acetate-acetic acid-formic acid-water (100:11:11:27); (g, h) acetonitrile-water-formic acid (30:8:2); (i, j) butanol-acetic acid-water (7:1:2).	41
Figure 2.2	Method development for *Gelsemium elegans* in optimizing the mobile phase.	48
Figure 2.3	Method development for *Gelsemium elegans* in optimizing the mobile phase.	49
Figure 2.4	Comparison of mobile phase chloroform-methanol-water ratio: (a) 60:10:1; (b) 50:10: 1; (c) 40:10:1; (d) 30:10:1; (e) 20:10:1.	51
Figure 2.5	Comparison of mobile phase chloroform-methanol-water ratio: (a) 60:10:1; (b) 50:10: 1; (c) 40:10:1; (d) 30:10:1; (e) 20:10:1.	51
Figure 2.6	HPTLC plates after derivatization with different reagents: (a) 2,4-dinitrophenylhydrazine under UV 366 nm; (b) Dragendorff's reagent under UV 366 nm; (c) Dragendorff's reagent under white light; (d) Iodine spray solution under UV 366 nm; (e) Iodine spray solution under white light; (f) Sulphuric acid reagent under UV 366 nm; (g) Sulphuric acid reagent under white light; (h) Vanillin reagent under UV 366 nm; (i) Vanillin reagent under white light.	52
Figure 2.7	The HPTLC plate observed under: (a) UV 254 nm; and (b) UV 366 nm.	57

Figure 2.8	Calibration curves obtained from the graph of peak area versus concentration of koumine (µg/ml).	61
Figure 2.9	UV spectra: Spectrum 1: root extract; spectrum 2: leaf extract; spectrum 3: stem extract; spectrum 4: koumine.	65
Figure 2.10	HPTLC chromatograms: Track K: koumine; track 1: stem from Fu Jian province; track 2: stem from Guang Xi province; track 3: root from Fu Jian province; track 4: root from Guang Xi province; track 5: leaf from Fu Jian province; track 6: leaf from Guang Xi province, with retardation factor (Rf) scale at both sides.	66
Figure 2.11	Average amount of koumine in stem, root and leaf of *Gelsemium elegans*.	70
Figure 2.12	Average amount of koumine for different parts of *Gelsemium elegans* in Fu Jian Province.	71
Figure 2.13	Average amount of koumine for different parts of *Gelsemium elegans* in Guang Xi Province.	72
Figure 3.1	Different extraction solvents for *Gelsemium elegans*: (a) 50 % methanol with 0.1 % formic acid; (b) 50 % methanol; (c) 100 % methanol; (d) 70 % ethanol.	84
Figure 3.2	Different extraction time for *Gelsemium elegans* in 50 % methanol with 0.1 % formic acid: (a) 10 min; (b) 20 min; (c) 30 min; (d) 40 min.	85
Figure 3.3	Different mobile phase compositions for the extraction of *Gelsemium elegans*: (a) acetone-water with 0.1 % formic acid; (b) acetone-water; (c) methanol-water.	86
Figure 3.4	Different mobile phase flow rate for the extraction of *Gelsemium elegans*: (a) 0.10 ml/min; (b) 0.20 ml/min; (c) 0.30 ml/ min; (d) 0.40 ml/ min.	86
Figure 3.5	Different column temperature for the mixed standard solutions: (a) 20 ^{o}C; (b) 30 ^{o}C; (c) 40 ^{o}C.	87
Figure 3.6	Different detection wavelengths for *Gelsemium elegans*: (a) 254 nm; (b) 260 nm.	88
Figure 3.7	UPLC Chromatogram of gelsemine and koumine.	89
Figure 3.8	UPLC Chromatogram of *Gelsemium elegans* stem : (a) Fu Jian; (b) Guang Xi.	90
Figure 3.9	UPLC Chromatogram of *Gelsemium elegans* root : (a) Fu Jian; (b) Guang Xi.	90

Figure 3.10	UPLC Chromatogram of *Gelsemium elegans* leaf : (a) Fu Jian; (b) Guang Xi.	91
Figure 3.11	SIR of (a) gelsemine and (b) koumine for (i) Pure Compound; (ii) Stem; (iii) Leaf; (iv) Root.	91
Figure 3.12	MRM chromatogram of: (a) gelsemine; (b) koumine; (c) febrifugine with the chromatographic condition according to Chapter 2.24.	94
Figure 3.13	MRM chromatogram of: (a) gelsemine; (b) koumine; (c) febrifugine after the first modification.	95
Figure 3.14	MRM chromatogram of: (a) gelsemine; (b) koumine; (c) febrifugine with the chromatographic condition according to Chapter 2.25.	96
Figure 3.15	Total ion chromatogram of gelsemine characterized in *Gelsemium elegans* by UPLC-ESI-MS/MS.	98
Figure 3.16	Total ion chromatogram of koumine characterized in *Gelsemium elegans* by UPLC-ESI-MS/MS.	98
Figure 3.17	Parent-Daughter MRM chromatogram of gelsemine: (a) 323.15→236.15; (b) 323.15→70.1; (c) 323.15.	99
Figure 3.18	Parent-Daughter MRM chromatogram of koumine: (a) 307.1→180.1; (b) 307.1→167.1; (c) 307.1→70.05; (d) 307.1.	100
Figure 3.19	Representative MRM chromatograms of gelsemine in different parts of *Gelsemium elegans*: (a) Stem; (b) Root; (c) Leaf.	101
Figure 3.20	Representative MRM chromatograms of koumine in different parts of *Gelsemium elegans*: (a) Stem; (b) Root; (c) Leaf.	102
Figure 3.21	Calibration curve with line equation and R^2 obtained from the graph of peak area versus concentration of gelsemine (ng/ml).	106
Figure 3.22	Calibration curve with line equation and R^2 obtained from the graph of peak area versus concentration of koumine (ng/ml).	106
Figure 3.23	Average amount of gelsemine and koumine for different parts of *Gelsemium elegans.*	110
Figure 3.24	Average amount of gelsemine and koumine for different parts of *Gelsemium elegans* in Fu Jian Province.	112

Figure 3.25	Average amount of gelsemine and koumine for different parts of *Gelsemium elegans* in Guang Xi Province.	113
Figure 3.26	Principal component analysis scores plot for the discrimination of the different parts of *Gelsemium elegans.*	115
Figure 4.1	Comparison of FT-IR results for different parts of *Gelsemium elegans*: (a) stem from Fu Jian province; (b) stem from Guang Xi province; (c) root from Fu Jian province; (d) root from Guang Xi province; (e) leaf from Fu Jian province; (f) leaf from Guang Xi province.	124
Figure 4.2	Comparison of SD-IR results for different parts of *Gelsemium elegans*: (a) stem; (b) root; (c) leaf.	127
Figure 4.3	The 2D-correlation IR spectra of each part of the *Gelsemium elegans* plant in the range of 1250 - 850 cm^{-1}: (a) stem; (b) root; (c) leaf.	130
Figure 4.4	The 2D-correlation IR spectra of each part of the *Gelsemium elegans* plant in the range of 1750 - 1160 cm^{-1}: (a) stem; (b) root; (c) leaf.	132
Figure 4.5	Infrared Spectra of GW001-S repeatedly scanned five times.	134
Figure 4.6	Infrared Spectra of GW001-S with five different pellets.	134
Figure 4.7	Principal component analysis scores plot for the discrimination of the different parts of *Gelsemium elegans.*	136
Figure 5.1	PCA score plots of the original and duplicated samples.	144
Figure 5.2	1H NMR results for different parts of *Gelsemium elegans*: (a) stem from Fu Jian province; (b) stem from Guang Xi province; (c) root from Fu Jian province; (d) root from Guang Xi province; (e) leaf from Fu Jian province; (f) leaf from Guang Xi province.	145
Figure 5.3	1H NMR results for different parts of *Gelsemium elegans* in the zone of 0.50 – 1.70 ppm: (a) stem from Fu Jian province; (b) stem from Guang Xi province; (c) root from Fu Jian province; (d) root from Guang Xi province; (e) leaf from Fu Jian province; (f) leaf from Guang Xi province.	147
Figure 5.4	1H NMR results for different parts of *Gelsemium elegans* in the zone from 2.50 – 5.00 ppm: (a) stem from Fu Jian province; (b) stem from Guang Xi province; (c) root from Fu Jian province; (d) root from Guang Xi province; (e) leaf from Fu Jian province; (f) leaf from Guang Xi province.	148

Figure 5.5	Hotelling's T2 plot of the extracts of the different parts of *Gelsemium elegans*. The results are based on the 95% confidence intervals. Black, blue and green stand for stem, root and leaf, respectively.	150
Figure 5.6	(a) PCA score plots and (b) PCA loading plots of different parts of *Gelsemium elegans* without scaling.	151
Figure 5.7	(a) PCA score plots and (b) PCA loading plots of different parts of *Gelsemium elegans* with Pareto scaling.	153
Figure 6.1	Correlation of the koumine contents determined by UPLC-ESI-MS/MS versus HPTLC-UV.	157

LIST OF ABBREVIATIONS

2D-IR	Two-dimensional correlation infrared
2D-NMR	Two-dimensional NMR
AA	Acetic acid
AU	Intensity of Absorbance
AHP	American Herbal Pharmacopoeia
APCI	Atmospheric Pressure Chemical Ionization
BEH	Ethylene Bridged Hybrid
CD_3OD	Tetradeuteromethanol
CH_2Cl_2	Dichloromethane
$CHCl_3$	Chloroform
CHM	Chinese Herbal Medicine
ChP	Chinese Pharmacopoeia
COSY	Correlation Spectroscopy
DCM	Dichloromethane
DTGS	Deuterated Tri-Glycine Sulfate
ESI	Electrospray ionization
EtOAc	Ethyl Acetate
EU	European Union
FA	Formic acid
FDA	US Food and Drug Administration
FT-IR	Fourier transform infrared
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High Performance Liquid Chromatography
HPTLC	High performance thin layer chromatography
HSQC	Heteronuclear Single Quantum Coherence
IR	Infrared
KBr	Potassium bromide
LC	Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
LC-MS/MS	Liquid Chromatography-Mass Spectrometry - Mass Spectrometry
LD_{50}	Semi-Lethal Dose

LOD	Limit of Detection
LOQ	Limit of Quantification
MeOH	Methanol
MRM	Multiple Reaction Monitoring
NMR	Nuclear magnetic resonance
PA	Peak area
PC	Principal Components
PCA	Principal Components Analysis
PDA	Photodiode Array
PLS-DA	Partial Least Squares-Discriminant Analysis
PLS-R	PLS-regression
PTFE	Polytetrafluoroethylene
QC	Quality control
R	Correlation coefficient
R^2	Correlation of determination
R_f	Retardation factor
RSD	Relative Standard Deviation
S	Slope
SD-IR	Second derivative infrared
SE	Standard error of low level concentration
SEM	Standard error of mean
SFDA	China State Food and Drug Administration
Signal-to-Noise	S/N
SIMCA	Soft Independent Modeling by Class Analogy
SIR	Selected Ion Recording
SNR	Signal-to-noise ratio
SNV	Standard normal variate
SOP	Standard operating procedure
TBME	Methyl tert-butyl ether
TCM	Traditional Chinese Medicine
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMS	Tetramethylsilane

Tol	Toluene
UK	United Kingdom
UPLC	Ultra-performance liquid chromatography
UPLC-ESI-MS/MS	Ultra Performance Liquid Chromatography-Electrospray Tandem Mass Spectrometry
UPLC-MS	Ultra Performance Liquid Chromatography-Mass Spectrometry
UPLC-MS/MS	Ultra Performance Liquid Chromatography-Tandem mass Spectrometry
UPLC-PDA-MS	Ultra Performance Liquid Chromatography- Photodiode Array-Mass Spectrometry
UPLC-PDA-QDa	Ultra Performance Liquid Chromatography- Photodiode Array-Mass Spectrometry
USA	United States of America
USM	Universiti Sains Malaysia
USPC	U.S. Pharmacopoeia Convention
USP	United States Pharmacopoeia
UV	Ultraviolet
W	Water
WHO	World Health Organization

LIST OF SYMBOLS

%	Percentage
°C	Degree Celsius
°C/min	Degree Celsius per minute
µg/kg	Microgram per kilogram
µg/mg	Microgram per milligram
µg/ml	Microgram per liter
µl	Microliter
AU	Absorbance unit
cm	Centimeter
cm^{-1}	Reciprocal centimeter
g	Gram
h	Hour
Hz	Hertz
K	Kelvin
kV	kilovolt
L/h	Liter per hour
mg	Milligram
mg/ml	Milligram per liter
MHz	Megahertz
min	Minute
ml/min	Milliliter per minute
mm	Millimeter
mm/s	Millimeter per second
ms	Millisecond
ng/ml	Nanogram per milliliter
nl/s	Nanoliter per second
nm	Nanometer
ppm	Parts per million
psi	Pound-force per square inch
rpm	Rotation per minute
s	Second

Secs	Seconds
V	Volt
v/v	Volume per volume
µm	Micrometer

LIST OF APPENDICES

Appendix A Conditions and Procedures of the Four Analytical Instruments Based On U.S. Pharmacopeia

Appendix B Sample name, Part, Province, City and Specific Location of the 67 *Gelsemium elegans* Samples.

Appendix C Pre-Viva Certificate

PENDEKATAN KROMATOGRAFIK DAN SPEKTROSKOPIK UNTUK CAP JARI KIMIA *GELSEMIUM ELEGANS* DENGAN MENGGUNAKAN HPTLC, LC-MS, FT-IR DAN NMR

ABSTRAK

Cap jari kimia biasanya digunakan untuk melakukan pengesahan dan pengenalan ubat-ubatan herba Cina (CHM). Kebelakangan ini, kaedah cap jari kimia telah digunakan secara meluas untuk kawalan mutu ubat-ubatan herba Cina. Pengesahan cap jari kimia adalah satu kaedah yang menilai ciri-ciri bahan dengan menggunakan satu atau lebih daripada satu teknik pengenalan. Oleh itu, dalam kajian ini, cap jari spektroskopi dan kromatografi *Gelsemium elegans* dikaji dengan menggunakan empat instrumen analitikal, iaitu pengenalan HPTLC, LC-MS, FT-IR tiga-langkah dan NMR. Tujuan kajian ini adalah untuk membandingkan kaedah analitik yang berbeza pada *Gelsemium elegans* dari daerah Fu Jian dan Guang Xi sebagai objek kajian untuk membezakan antara batang, daun dan akar kerana setiap bahagian mengandungi jumlah alkaloid indole yang berbeza yang menentukan tahap keracunannya. Pembezaan antara bahagian-bahagian tersebut adalah sukar kerana mereka datang dari spesies yang sama dan berkongsi sifat-sifat yang serupa dengan sebatian aktif. Walau bagaimanapun, tiga bahagian yang berbeza dari *Gelsemium elegans* berjaya dibezakan oleh empat instrumen. Selain itu, dua kompleks aktif yang terdapat dalam *Gelsemium elegans,* gelsemine dan koumine dikaji secara kualitatif dengan HPTLC dan LC-PDA-QDa. Tambahan pula, LC-MS / MS juga digunakan untuk penentuan kuantitatif gelsemine dan koumine dalam *Gelsemium elegans*. Kandungan koumine juga digunakan untuk perbandingan kuantiti dalam LC-MS / MS oleh HPTLC. Kedua-dua kaedah menunjukkan ketepatan, pemulihan, kestabilan dan

kebolehulangan yang baik. Kandungan koumine dalam *Gelsemium elegans* yang diukur menggunakan HPTLC dan LC-MS / MS menunjukkan korelasi positif dengan korelasi penentuan 0.8998, membuktikan bahawa hasil daripada kedua-dua kaedah boleh disemak silang. Selain itu, untuk mendapatkan perbandingan yang sempurna, analisis komponen utama (PCA) dilakukan berdasarkan data yang diperoleh melalui FT-IR, NMR dan LC-PDA untuk mendiskriminasikan ketiga-tiga bahagian tersebut. Dengan kemajuan teknologi komputer, kaedah kemometrik telah menjadi alat utama di kalangan komuniti saintifik untuk keputusan analisis yang lebih cepat dan masa pembangunan produk yang lebih pendek. Oleh itu, penggunaan kemometrik dalam bidang ubatan adalah penting dan diperlukan. Di antara pelbagai kaedah kemometrik, PCA adalah teknik pengenalan corak yang tidak diselia yang paling sering digunakan untuk mengendalikan data multivariat tanpa pengetahuan terlebih dahulu mengenai sampel yang dipelajari. Sebaliknya, kaedah teknik pengiktirafan corak yang diselia, Pemodelan Lembut Kelas Analogi (SIMCA) juga dilakukan pada data FT-IR. Melalui kajian ini, ketiga-tiga bahagian *Gelsemium elegans* berjaya dikenalpasti dan didiskriminasi melalui cap jari spektroskopi dan kromatografi, PCA dan SIMCA daripada instrumen yang dikaji. Kesimpulannya, diskriminasi bahagian-bahagian yang berbeza dari *Gelsemium elegans* harus dilakukan daripada pelbagai sudut termasuk, cap jari, kuantifikasi dan analisis kemometrik, untuk memberikan hasil yang lebih muktamad.

SPECTROSCOPIC AND CHROMATOGRAPHIC APPROACH FOR THE CHEMICAL FINGERPRINTING OF *GELSEMIUM ELEGANS* VIA HPTLC, LC-MS, FT-IR AND NMR

ABSTRACT

Chemical fingerprints are commonly used to perform the authentication and identification of Chinese herbal medicines (CHM). During the last few years, the fingerprint method has been developed for quality control of Chinese herbal medicines. Fingerprinting is a method that evaluates the characteristic pattern of the ingredients using one or more identification techniques. Therefore, in this study, the spectroscopic and chromatographic fingerprints of *Gelsemium elegans* were studied by using four analytical instruments, namely the HPTLC, LC-MS, FT-IR tri-step identification, and NMR. The purpose of this study was to compare different analytical methods on *Gelsemium elegans* from Fu Jian and Guang Xi province as the object of study to distinguish between the stem, leaf and root as they contained different amounts of indole alkaloid that contributes to its toxicity. The differentiation between the different parts was difficult as they came from the same species and shared similar properties and active compounds. However, the three different parts of *Gelsemium elegans* was successfully distinguished by the four instruments. Besides that, two abundant active compound present in *Gelsemium elegans*, gelsemine and koumine was qualitatively studied in HPTLC and LC-PDA-QDa. Furthermore, LC-MS/MS conditions were also developed for quantitative determination of gelsemine and koumine in *Gelsemium elegans*. The content of Koumine was also used to cross check the quantity in LC-MS/MS by HPTLC. Both methods showed good precision, recovery, stability and repeatability. The content of koumine in *Gelsemium elegans* measured by HPTLC and

UPLC-MS/MS showed positive correlation with the correlation of determination of 0.8998, proving that the results from both methods can be cross checked. Other than that, to have a well-resolved comparison, principal component analysis (PCA) was performed based on the data obtained by FT-IR, NMR and UPLC-PDA to discriminate the three parts. With the advancement of computer technology, chemometrics methods have become a leading tool among the scientific communities towards faster result analysis and shorter product development time. Therefore, the application of chemometrics in the field of medicinal plants is crucial and necessary. Among the variety of chemometrics methods, PCA is an unsupervised pattern recognition technique that is most often used for handling multivariate data without prior knowledge about the studied samples. On the other hand, supervised pattern recognition technique method, Soft Independent Modelling of Class Analogy (SIMCA) was also performed on the FT-IR data. Through this study, all three parts of *Gelsemium elegans* were successfully identified and discriminated through the spectroscopic and chromatographic fingerprint, PCA and SIMCA of the instruments studied. It can be concluded that the discrimination of different parts of *Gelsemium elegans* should be performed from various angles including fingerprints, quantification and chemometric analysis, in order to provide a more conclusive outcome.

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Overview of TCM

According to the definition given by the World Health Organization (WHO), traditional medicine can be defined as: "the sum total of the knowledge, skills, and practices based on the theories, beliefs, and experiences indigenous to different cultures, whether explicable or not, used in the maintenance of health as well as in the prevention, diagnosis, improvement or treatment of physical and mental illness." (World Health Organization, 2013)

In this specific circumstance, Traditional Chinese Medicine (TCM) practices include medications such as Chinese Herbal Medicine (CHM), acupuncture, dietary treatments, and both Tui na and Shiatsu massage. Qigong and Taijiquan are likewise intently connected with TCM (Guimaraes, 2006). Like other conventional medicine, traditional medicine is the most well-known healing method of all and for the convenience of discussion, the term "Traditional Chinese Medicines" (TCM) will be used interchangeably with Chinese Herbal Medicines (CHM) in the subsequent parts of this thesis.

TCM depends on ancient Chinese logic (Wang et al., 2013). TCM adopts a comprehensive strategy in treating the person with customized treatment in view of the idea of "Disorder Differentiation". The essential speculations of TCM were derived from the Chinese rationality of Yin-Yang and Five Elements, and its central ideas, for example, the Zang-fu (viscera) idea, Qi (fundamental vitality), and meridians still cannot plainly explained in scientific terms. (Fung and Linn, 2015)

TCM treats diseases in a holistic manner and centres around the patient rather than the disease and regularly uses formulae that contains a few types of raw herbs customized to a person's condition in view of subjective diagnosis techniques (Wang et al., 2013). In the 1950s', modern TCM was systematized under the People's Republic of China. Previously, CHM was only practiced within family ancestry framework. (Guimaraes, 2006)

Traditional medical practices are very much recorded, with a portion of the legitimate medicinal books being more than 2,000 years of age. For instance, the *Compendium of Materia Medica* gathered by Li Shizhen of the Ming Dynasty shows 1,892 sorts of herbs and 10,000 prescriptions (Sahoo, 2012).

In 2015, Youyou Tu, who won the Nobel Prize for Physiology or Medicine, was the first science Nobel Prize granted to a China-based researcher. Tu's disclosure of artemisin in which saved millions of lives, was established in ancient Chinese herbal medicine. It has brought TCM to the cutting edge of the worldwide research network's consideration (Tu, 2011).

1.2 Importance of Chinese Herbal Medicines (CHM)

TCM plays an important role in Asia's health care system. It depends on natural products and has been in a crucial part in health safety for a few thousand years (Liu et al., 2011). Herbal medicines are naturally occuring; plant-derived substances with negligible or no mechanical processing that have been utilized to treat health issues within local healing systems (Tilburt and Kaptchuk, 2008). In certain customs, materials of inorganic or animal source may likewise be utilized (World Health Organization, 2000). As indicated by WHO, herbal medicines are characterized as herbs, herbal materials, herbal preparations and finished herbal products, which

contain as active ingredients parts of vegetation, or other herb materials, or their combinations (Liu, 2011).

CHM which is comprised of a specific combination of various components has been growing to be progressively prominent as a multi-component medication remedy (Ma et al., 2012). Herbal products have likewise been applied worldwide for many centuries as part of regular medications (Goodarzi et al., 2013).

CHM is now significantly well-known in both developing and developed nations. WHO quotes that 80% of the world's human population depends on CHM for their health care needs (Zhang et al., 2014). As indicated by WHO, TCM makes up about 30 to 50 % of total therapeutic intake in China. Besides that, bureaus of traditional medications have been created in almost all hospitals, including those hospitals that offer Western treatments. There are more than 500,000 professionals of TCM in China, and even specialists in provincial areas have some basic knowledge of acupuncture and therapeutic herbal products. In this way, TCM signifies an important source of medicinal services for the Chinese society and its effects cannot be disregarded.

The universal status of TCM has been upgraded by discovering that the Chinese herbal product, *Artemisia annua*, which has been utilized for more than 2,000 years, is impressively effective against resistant malaria. There is certainly hope that it could represent a significant step of progress in tackling the disease, specifically as modern medication development in this field has been considered to be lacking. WHO is currently supporting specialized medical studies with the Chinese *Artemisia annua* for the benefit of African nations influenced by malaria. WHO has perceived the worthiness of TCM and is currently dealing with China (Sahoo, 2012).

While TCM is extensively applied in Asian populations, for example, China, Taiwan, Hong Kong, and Singapore, numerous non-Asian countries have, in recent decades, also acknowledged the huge restorative capability of this customary practice and have been experiencing great things about TCM to be able to give patients an extra alternative in their health management. Although regular medicine is utilized generally in Western countries, however the consumption of CHM is also increasing, predominantly on account of the side effects or inefficacy of current medications (Liu, 2011). In 1991, a TCM medical centre was opened in KoEtzting, Germany, whereby trained TCM health care professionals from China implemented treatment in line with the traditional practice (Melchart et al., 1999). Regarding interim data from 2nd WHO TRM global study as of 11 June 2012, 80% of the 129 members claim that they now accept the utilization of acupuncture as a cure modality (World Health Organization, 2014).

1.3 Issues of CHM

1.3.1 CHM Adulteration

The adulteration of CHM is a major problem in herbal industry and it has caused a significant impact in the industrial use of natural products. An example of adulterated CHM is often a combination of the components of the original herb and the adulterant with a motive of increment in revenue (Chen et al., 2015; Kamboj and Saluja, 2012). Besides that, adulteration might be characterized as blending or replacing the initial drug material with different, substandard, spoiled, ineffective other parts of the same or of different herbs or the use of unsafe chemicals or medicines which do not comply with the official principles (Ansari, 2003; Kokate and Gokhale, 2004; Kamboj and Saluja, 2012).

Nevertheless, some vendors will include synthetic drugs in the formulation of their products marketed as herbal medicine in order to boost the efficiency of their products (Liang et al., 2006). Adulteration of CHM with illegal synthetic drugs is a common issue, which will probably cause serious negative effects (Bogusz et al., 2006). Even though by far almost all of the adulteration situations do not cause a public health risk, a few conditions have lead to actual or potential public health threats (Johnson, 2014). There are cases of medical problems caused by the man-made drugs in CHM reported. These incidences demonstrated the significance of discovering the occurrence of any adulterants in CHM to ensure the safety of the patients (Lau et al., 2003).

The adulteration practice abuses the laws of various countries, as the formulations authorized are different from the genuine structures. This practice can result in abnormal impact on the human body, either because of the pharmaceutical itself or its interaction with various components introduced into the formulation (Carvalho et al., 2010).

1.3.1(a) Direct or deliberate adulteration

Direct adulteration is where herbal medication is fully or partly replaced with other products on purpose. Due to the morphological similarity to the original herb, a variety of substandard herbs are being used as substitutes. This practice can usually be found when it involves expensive medication. For example, beeswax was replaced with coloured paraffin wax (Kamboj and Saluja, 2012).

1.3.1(b) Indirect or unintentional adulteration

Unintentional adulteration may in some cases happen without bad intentions of the producer or provider. Sometimes with the lack of appropriate methods for assessment, an authentic medicine partly or completely without containing the active

materials may be sold. Physical sources, developing conditions, processing, and capacity are some general causes that affect the grade of CHM (Ansari, 2003; Kokate, 2004).

Some herbal adulterations may be due to the carelessness of herbal collectors. Collection of other herbs by ignorance, because they look similar in appearance and colour, may lead to adulteration. Non-removal of unwanted parts such as cork from ginger rhizome is also a type of unintentional adulteration (Kamboj and Saluja, 2012).

1.4 Quality Control of CHM

TCM has become more popular in the developed nations for being natural therefore people usually assume that they are inherently safe (Goodarzi et al., 2013). There is insufficient unified, systematic legislation for evaluating the safe practices of TCM and ensuring the standard of TCM products (Fong et al., 2006; Tang et al., 1999; Barrett, 2004; Brinker, 2009; Chadwick et al., 2006, Ernst, 2006; Tyler, 2004). With the expanded consumption of TCM, the safety, standard and effectiveness of these medications is a very important matter for health regulators and medical researchers (Lau et al., 2003). The quality control of CHM is more complicated than that of western medications considering the myriad of synthetic compounds comprised in CHM and the associations of various phytochemicals adding to the curative impact (Liang et al., 2004; Gong et al., 2003). The source, developing condition, collection time and handling technique are important factors that will affect the safety of TCMs. Hence, the quality control of CHM has remained the focus of numerous studies and is complicated in CHM research (Zhu et al., 2017).

A few nations have issued rules for the quality standards for CHM. For instance, the Chinese Pharmacopoeia Commission compiled the “National Drug Standards Work Manual” (Chinese Pharmacopoeia Commission, 2013), used for research,

improvement and application of medication standards (Chen et al., 2017a). WHO is keen in regards to herbal medicine and has been active in making systems, rules and principles of natural pharmaceuticals (World Health Organization, 2000). WHO had also released several rules and acts regarding the safe and acceptable use of herbal pharmaceuticals in the year 2004.

Internationally, there have been purposeful efforts to screen the quality and guide the growing business of herbal medicines. Health experts and governments of different countries have been active and enthusiastic in providing standardized botanical prescriptions. The United States Congress has fuelled fast development in the nutraceutical market with a section on the Dietary Supplement Health and Education Act in 1994. The US Food and Drug Administration (FDA) has lately allocated the International Conference on Harmonization guidance Common Technical Document addressing concerns regarding standard of medicines that likewise includes herbals (Patwardhan, 2005). The National Centre for Complementary and Alternative Medicine has been initiated as the United States Federal Government's lead agency for scientific research in this field of medicine. Its central goal is to study correlative and elective healing practices in relation to thorough research, support sophisticated research, train specialists, disperse data to the people on the modalities that work and clarify the scientific rationale underlying disclosures (Cooper, 2005; Gavaghan, 1994). In 1820, the U.S. Pharmacopoeia (USP) Convention was founded to set standards that help ensure the grade and advantages of medicines and foods. The Herbal Medicines Compendium (HMC) was also published by USP to provide standards for herbal ingredients utilized in herbal drugs. Standards are expressed primarily in monographs (U.S. Pharmacopoeia Convention, 2013).

However, there are no nationwide TCM standards or rules for TCM medicine clinical trials, and evidence-based TCM medicine product trials and research are still required. In view of the huge dissimilarities in the requirements of TCM practitioners, the standard of TCM education must to be strengthened, and therefore administration and supervision of TCM institutions have to be managed (Sahoo, 2012).

1.5 Analytical methods for quality control of herbal medicines

The whole production process from the raw materials to final products of herbal medicines, need to be controlled (Bogusz, 2006). The China State Food and Drug Administration (SFDA) enforced a regulation in 2004 on the quality control of herbal injection in which chromatographic or spectrometric fingerprinting tests are obligatory for quality checking throughout the whole preparation process which include the standardization of crude material to the final item analysis (State Food and Drug Administration, 2000). The main aim of carrying out quality control is to make sure the authenticity of CHM which includes the identification the correct species from the adulterants. Different recognition methods have been utilized in the identification item of *Chinese Pharmacopoeia* (2015 edition) and that includes the morphology, microscopy, fingerprint, characteristic chromatogram, and DNA sequencing (Yang et al., 2017; Guo et al., 2015).

A combination of analytical methods should be used for the quality control of herbal medicines. In this study, both 'marker-based' and 'fingerprint' quality control were applied but the main focus will be on the 'fingerprint' quality control.

1.5.1 The 'Marker-based' Approach to Quality Control

TCM formulations are usually manufactured from numerous herbs. Thus, for the standardization of each formulation for manufacturers to create regular products, a marker compound will be chosen for each plant to be contained in the formulation.

Marker compounds are chemically identified compounds or groups of compounds that are useful to control batch-to-batch regularity of the completed item irrespective of their therapeutic activity. Marker compounds can be split into two types, which are the active constituent marker and the analytical marker. The active constituent marker is a compound or group of compounds that contributes to therapeutic action, whereas the analytical marker is a compound or several compounds that will not contribute to the therapeutic activity but is limited to analytical purposes only, which was utilized as part of this research (Health Canada, 2015).

The marker strategy focuses on the exact qualitative and quantitative measurements of a single synthetic substance. However, this method has some issues on its application in CHM. This is because, there are almost hundreds or even thousands of structurally and artificially varied entities with inadequately characterized composition in plants (Stone, 2008; Xie and Leung, 2009; Politi et al., 2009; Wolfender, 2009; Schmidt et al., 2008; Holmes et al., 2006; Ratcliffe and Scachar-Hill, 2005). The proof of limited chemical markers accessible for identification is observed in the field of CHMs.

Moreover, these markers might not exactly be unique to specific CHM. For example, chlorogenic acid can be used as the marker for *Flos Lonicerae*, *Flos Chrysanthemi*, and *Herba Saussureae Involucratae* while the biomarker of *Radix Angelicae Sinensis* could likewise be found in *Radix Ligustici chuanxiong* (Li et al., 2008, Mok and Chau, 2006).

Therefore, in this study, the marker chosen for the identification of the herbal medicines were properly selected and the specific markers that are present only in the herb were used.

1.5.2 The 'Fingerprinting' Approach to Quality Control

The fingerprinting of CHM for quality control is suggested by many administrative bodies, such as WHO, EU, FDA and SFDA due to the weakness of the marker-based quality analysis for CHM (World Health Organization, 2000; European Agency for the Evaluation of Medicinal Products, 2006; U.S. Department of Health and Human Services, 2004; State Drug Administration of China, 2000). In this approach, the medicinal herb is examined as an individual moiety with its elements being universally characterized irrespective of their medical relevance (Mok and Chau, 2006; Li et al., 2008). Adulteration through deliberate spiking of the few known markers can also be prevented as large quantities of components are considered in the quality assessment (Xie et al., 2006).

Generally, the fingerprint of a herb can be indicated by means of molecular biological profile, such as DNA patterns or a range of physicochemical reactions as shown in chromatograms and spectra. In any case, the intensive use of the DNA approach is bound by the thermal stability of the biomolecules. Steaming, frying or boiling at high temperature is frequently used in the preparation of TCM decoctions and genetic data can be easily damaged under such severe conditions (Bensky et al., 1986). The differences between chemical fingerprinting and DNA fingerprinting is shown in Table 1.1. In this study, the chemical fingerprinting approach was applied as it was more suitable for our study.

Table 1.1: Differences between chemical fingerprinting and DNA fingerprinting

Chemical Fingerprint	DNA Fingerprint
Difference between living plant and plant drug	No difference between plant and plant drug
Identification of raw material but also finished products	Identification of raw material only
Dried or processed materials can be used	Only fresh samples can be used
Differences in plant parts	Same for all plant parts

A chemical fingerprint of CHM is, by definition, "a pattern of the extract of some common chemical components which are pharmacologically active or possess certain chemical characteristics" (Liang et al., 2004; S.D.A.O.China, 2002; Ong, 2002; Xie, 2001). This chemical profile should be featured by the fundamental attributions of 'integrity' and 'fuzziness' or 'sameness' and 'differences' so as to chemically represent the CHM investigated (Liang et al., 2004; Xie, 2001; Welsh et al., 1996). In this way, the chemical fingerprints could effectively perform the authentication and identification of the CHM ('integrity') even if the amount of the characteristic constituents are marginally different for the same CHM ('fuzziness') and the chemical fingerprints can also successfully demonstrate the 'sameness' and the 'differences' between several samples (Liang et al., 2004; Xie, 2001;Valentao, 1999).

The verification and identification of all of the pharmacologically active compounds (Schaneberg et al., 2003), and their delegate quantity or concentration in the herb can be determined through a unique fingerprint profile. Moreover, fingerprint profiles of various samples with great separation can exhibit their similarities and distinguish them from closely related kinds (Ni et al., 2008).

1.5.3 Spectroscopic Fingerprinting

Spectroscopic methods produce results with short examination time and minimal sample preparation. The reproducibility is also very high even when the samples are analysed with different instruments and by different labs (Santos et al., 2017). Besides that, the process will not destroy the sample which allows sample recovery for further examination if needed. Spectral fingerprinting can be utilized to either group or segregate between samples or to evaluate certain compounds. However, spectroscopic instruments are insensitive as compared to chromatographic instrument such as LC-MS. Thus, compounds present in millimolar or micromolar concentrations cannot be practically detected directly (Chatham and Blackband, 2001).

1.5.3(a) Fourier Transform Infrared (FT-IR)

Among a number of quality control methods, infrared (IR) spectroscopy has been playing a more critical part in the testing for adulteration of natural products (Chen et al., 2015). As a direct, non-destructive, and label-free analytical method, IR spectroscopy can simultaneously distinguish the structures of organic and inorganic substances in complex mixtures (Sun et al., 2011).

That plant samples can be assessed specifically without prior extraction or labelling is the primary reason why IR spectroscopy can be considered as a basic, rapid, and green technique for the adulteration testing of plant materials (Chen et al., 2015). The repeatability and the reproducibility of the IR spectra measurements are great and most significantly, the interpretation of IR spectra is speedy and inexpensive (Sun et al., 2011).

Working principle of FT-IR spectra involve the study of interactions between matter and electromagnetic fields in the IR region. In this spectral region, the electromagnetic waves couple with the molecular vibrations. Molecules are excited to

a higher vibrational state by absorbing IR radiation. The IR frequency when absorbed would interact with the molecule at a certain frequency (Sun et al., 2011).

IR normally works in the range of 4000-400 cm^{-1}. Along with the innovation of Fourier transform infrared spectroscopy, scan time can be as short as 0.1 to 1 s. According to the bond vibration of different components within the herbal combination, the sophisticated IR spectra can be used as a unique fingerprint of the sample (Zou et al., 2005). The most straightforward type of assessment can be performed through comparing the peak shape, positions and intensities. Adulterant or substituent can be recognised through visual review, correlation coefficients or various chemometric procedures (Dong et al., 2002; Liu et al., 2008). To intensify the difference in slight changes and increase the quality of the spectral data, data treatment such as derivative of IR spectra can be applied (Liu et al., 2008; Hua et al., 2003).

Besides that, two-dimensional correlation infrared (2D-IR) can be employed to conquer some bottlenecks encountered by classic IR spectroscopy. For example, 2D-IR has higher spectral quality for discrimination of complicated mixture of herbal medications (Sun et al., 2011). This is appropriate to resolve the overlapping of absorption peaks of different components in IR spectroscopy. It will improve the specificity of IR spectroscopy to differentiate between real and adulterated samples (Chen et al., 2015).

Zhang et al., 2014 and Yan et al., 2016 applied infrared spectroscopic tri-step identification approach in their research to discriminate similar-looking plants, *Pterocarpus santalinus* and *Dalbergia louvelii* and *Lonicerae japonicae Flos* (LJF) and *Lonicerae Flos* (LF), respectively. In both studies, two different plants were successfully distinguished by using less amount of samples with IR technology.

Moreover, Chen et al. (2018) used tri-step FT-IR to discriminate between wild, cultivated and tissue cultivated *Anoectochilus roxburghii* plant.

In 2010, Zhang and team evaluated the different grades of ginseng using tri-step FT-IR analysis. Different cultivation types were plainly distinguished by the evaluation of the end fibrous root of ginseng with FT-IR and 2D-IR. The ginsengs with different growth years were also discriminated by the tri-step infrared fingerprint identification method as well. The results suggested that tri-step FT-IR is a rapid, straightforward and effective technique for the identification of the fingerprint characters and the analysis of different standards of ginseng. It may play a crucial part in the recognition of the source, the authentication of the real and adulterated products, and the discrimination of the high and low quality products for TCM. (Zhang et al., 2010)

1.5.3(b) Nuclear Magnetic Resonance (NMR)

Nuclear magnetic resonance (NMR) spectroscopy can be known as the most effective structural examination approach (Marston and Hostettmann, 2009). It has become an attractive elective platform for metabolite fingerprinting because of its non-selective nature, easier sample preparation and high resolution spectra (Tarachiwin et al., 2008; Lee et al., 2009; Kim et al., 2005; Liang et al., 2009; Bilia et al., 2002; Choi et al., 2005; Daykin et al., 2005; Belton et al., 1998). Since all proton-bearing types at reasonable concentration are discovered, an unbiased profile with significant structural information can be obtained in short examination time (Politi et al., 2009; Roos et al., 2004; Bailey et al., 2002; Tarachiwin et al., 2008; Daykin et al., 2005; Qin et al., 2009; Ward et al., 2003; Kim et al., 2007).

NMR spectroscopy has turned out to be progressively vital in food science (Alberti et al., 2002), both as a fingerprinting approach (Krishnan et al., 2005) and as

quantitative examination tool (Malz and Jancke, 2005; Caligiani, 2007). This progress in the advancement of NMR techniques in food characterization and control is mainly due to the simple preparation of samples, the speed of examination and the possibility of gaining structural information in a complicated food matrix (Caligiani et al., 2010). NMR spectroscopy is high-throughput, taking just a few minutes for every sample, (Lenz and Wilson, 2007) has generally low per-sample cost, and does not require a prior knowledge (Beckonert et al., 2007; Reo, 2002) of what metabolites to study (Nicholson and Lindon, 2008) since it yields a spectrum of every detectable metabolite. Besides that, it is a powerful tool for discriminating between sets of related samples and it identifies the most critical ranges of the spectrum for advanced examination (Krishnan et al., 2005).

NMR spectroscopy also provides a huge scope of metabolites per measurement and provides data on the chemical structure, chemical condition, dynamic atomic movements, and molecular interactions between metabolites (Lenz and Wilson, 2007). NMR spectroscopy is nondestructive, so it can be employed to samples before they are subjected to further destructive analysis (Kosmides et al., 2013). Above all, the intrinsic high reproducibility of NMR spectrum across instruments permits the development of a solid spectroscopic databank (Keun et al., 2002; Verpoorte et al., 2007).

The principle behind NMR is that many nuclei have spin and all nuclei are electrically charged. If an external magnetic field is applied, an energy transfer is possible between the base energy to a higher energy level. The energy transfer takes place at a wavelength that corresponds to radio frequencies and when the spin returns to its base level, energy is emitted at the same frequency. The signal that matches this

transfer is measured in many ways and processed in order to yield an NMR spectrum for the nucleus concerned (Shin et al., 2007).

The initial studies on the utilization of NMR in fingerprinting of plant extracts can be found as early as 1980s by Kubeczka and Formácek (Kubeczka and Formacek, 1982). Numerous studies have reported metabolic profiling of natural products with NMR and multivariate analysis for quality control (Holmes et al., 2006; van der Kooy et al., 2009; Yang et al., 2006; Shin et al., 2007; van der Kooy et al., 2008; Himmelreich et al., 2003).

Kim et al. (2005) utilized an NMR-based metabolic fingerprinting approach to distinguish three distinctive Ephedra species (Ephedra sinica, Ephedra intermedia, and Ephedra equisetina) for their quality control. Moreover, Yang et al. (2006) distinguished potential reference compounds for the quality control of Panax ginseng using proton (1 H) and two-dimensional NMR metabolomics approaches. They also categorized four types of ginseng roots for the productive screening procedure with Soft Independent Modelling by Class Analogy (SIMCA) and Principal Components Analysis (PCA). Rapid and sensitive 1H-NMR-based metabolomic profiling of deuterated methanol-D_2O buffer extracts of transgenic tomato flesh was performed by Ben Akal-Ben Fatma et al. (2012). Kim et al. (2010a) also used this approach for the discriminatory examination of 11 South American Ilex species. Other researches using NMR-based techniques have been applied for the examination of metabolites including natural products like sweet warmwood (*Artemisia annua*), gingko (*Gingko biloba*) leaves, and scutellaria (*Scutellaria baicalensis*) root (Choi et al., 2003; Kang et al., 2008a; Van der Kooy et al., 2008).

1.5.4 Chromatographic Fingerprinting

In recent years, chromatographic fingerprinting has gained increasing attention and been universally acknowledged as an achievable means for the quality control of CHM (Li et al., 2007). Chromatographic fingerprint is a chromatogram that represents the chemical characteristics of herbal medicine. Generally, samples with the same unique fingerprint will have comparative properties. Hence, chromatographic fingerprinting can determine the identity, authenticity, and consistency of herbal medicines (Fan et al., 2006). Chromatographic fingerprinting is also able to recognise a specific herb and differentiate it from closely related species (Li et al., 2007). In view of the idea of phytoequivalence, the chromatographic fingerprints of CHM could be used for addressing the problem of quality control of herbal medicines (Liang et al., 2004; Andola et al., 2010). However, the interlaboratory repeatability and reproducibility for chromatographic instruments was rather low (Bogusz and Wu, 1991).

1.5.4(a) Ultra Performance Liquid Chromatography (UPLC)

Ultra-performance liquid chromatography (UPLC) is equipped for resolving complicated mixtures of polar and non-polar substances, and has turned into the technique of choice for the qualitative and quantitative evaluation of CHM extracts and products. UPLC is generally used for fingerprinting as it can successfully isolate the distinctive constituents of the extract and can provide both qualitative and quantitative information (Loescher et al., 2014). UPLC is easy to operate with completely automatable approach with high resolution, selectivity and sensitivity.

One of the major features of Liquid Chromatography (LC) is the capability to hyphenate with different detectors, the main ones being the photodiode array (PDA) and mass spectrometer (MS) detectors (Tistaert et al., 2011). The use of liquid

chromatography-mass spectrometry (LC-MS) and liquid chromatography-mass spectrometry - mass spectrometry (LC-MS/MS) has speedily increased within the last few years. In the past decade, the improvement of straightforward, dependable, LC-MS interfaces, especially electrospray (ESI) and atmospheric pressure chemical ionization (APCI), has spurred the advancement and acknowledgment of LC-MS methods, in a way that, today, there are numerous laboratories that regularly utilized LC-MS as the essential analytical technique (Hayen and Karst, 2003; Niessen, 2003). The upside of these techniques is that after each substance is being eluted, it will be captured by the mass spectrometer and provides an instantaneous molecular ion and/or major mass fragment that allow positive recognition of the eluting "peak" (Gail et al., 2001).

UPLC involves a column packed with the porous medium made of a granular solid material (stationary phase), such as polymers and silica, where the sample is injected and the solvent (mobile phase) passes to transport the sample. When a sample is injected, it is adsorbed on the stationary phase, and the solvent passes through the column to separate the compounds one by one, based on their relative affinity to the packing materials and the solvent. The component with the most affinity to the stationary phase is the last to separate. This is because high affinity corresponds to more time to travel to the end of the column (Alina, 2019).

After the individual components in a mixture are separated, the components will be transferred to the MS detector. The components will then ionized and the ions will be separated on the basis of their mass/charge ratio. The separated ions are then directed to a photo or electron multiplier tube detector, which identifies and quantifies each ion (Kang et al., 2012).

Tandem MS (MS/MS) analysis is the ability of the analyzer to separate different molecular ions, generate fragment ions from a selected ion, and then mass measure the fragmented ions. The fragmented ions are used to for structural determination of original molecular ions. Typically, MS/MS experiments are performed by collision of a selected ion with inert gas molecules such as argon or helium, and the resulting fragments are mass analysed and quantified (Kang et al., 2012).

UPLC offers clear points of interest over standard HPLC in peak quality, analytical efficiency, solvent consumption and sensitivity. It is extensively utilized for the separation and study of complicated systems (Chen et al., 2017a). Nováková et al. (2006）and Wren et al. (2006) demonstrated that compared with the HPLC, the UPLC permits shorter examination time up to nine times, and the separation effectiveness is much higher.

Among the coupled analytical techniques utilized for these applications, the combination of LC with ESI and triple quadrupole analyzers is the most regularly used. This approach has been efficiently utilized to determine pesticides in fruits (Wong et al., 2010), vegetables (Chung and Chan, 2010), wines (Economou et al., 2009), milk (Dagnac et al., 2009), or meat (Carretero, Blasco and Pico, 2008), for example. With LC-MS/MS, values of LOQs as low as few μg/kg are generally reached; also, the examinations are relatively rapid. For instance, 58 antibiotics were examined in milk in less than 15 min (Gaugain-Juhel et al., 2009). Besides that, the selective identification of the triple quadrupole analyzers permits the precise examination of incompletely separated compounds. In fact, in nearly a similar analysis time, which is around 14 min, 191 pesticide residues were identified from various natural products (Wong et al., 2010a). In order to accelerate these separations, these days, short columns

with smaller molecule diameters are also utilized, to create separations at ultra-high pressures in UPLC. With this equipment, more than 100 pesticides were identified in strawberry samples in less than 5 min (Taylor et al., 2008).

At the moment, UPLC has been greatly implemented in pharmacopoeias. In the ChP (2010 edition), the first application of UPLC was used for the evaluation and fingerprint examination of Danshen Dropping Pill (Chinese Pharmacopoeia Commission, 2010). The ChP (2015 edition) (Chinese Pharmacopoeia Commission, 2015) widened the UPLC application to 5 monographs for multi-component determination. By the end of 2016, there have been 57 monographs using UPLC techniques in the United States Pharmacopoeia (USP) (The United States Pharmacopeia Convention, 2015).

1.5.4(b) High Performance Thin Layer Chromatography (HPTLC)

Thin layer chromatography (TLC) is a straightforward versatile technique applied in pharmaceutical research for both qualitative and quantitative assessment of chemical constituents. TLC is the only chromatographic approach that presents the outcome as a photo-like image with all identified bands regarded as a grouping of dark, coloured or fluorescent zones. Examination time is relatively short, and numerous samples can be easily analysed side by side on a single plate (Reich and Schibli, 2007). The visible outcome and simplicity of TLC technique allow inexperienced analyst to easily run the chromatographic procedure. Additionally, simple presentation of separated components by UV light makes TLC a method of choice for simple and rapid evaluation (Ashutosh, 2005). The same developed TLC plate can be visualized and analysed repeatedly with different light sources (Kumar et al., 2008; Rumalla et al., 2008; Tian et al., 2009).

The principle of TLC is the distribution of a compound between a solid phase applied to a glass or plastic plate and a liquid mobile phase that is moving over the solid phase. In TLC, the sample components will migrate for a certain time. This way, separation is based on variable migration distance and fixed time. In TLC, capillary forces drive the mobile phase and Rf value is used to express the relative retention.

Lately, High Performance Thin Layer Chromatography (HPTLC) seems to have a renaissance of interest for some chromatographic experts. It has emerged as a preferred technique to traditional TLC, as unique plates and computerised applicator are used (Di et al., 2003). High performance plates have smaller particle size distribution, allowing higher separation power and greater level of sensitivity (Reich and Schibli, 2007). In addition to stationary and mobile phases, HPTLC also takes advantage of a gas phase (Reich and Schibli, 2007).

When comparing with established TLC, HPTLC provides better resolution, and in this way limits errors that may arise in manual TLC methods (Pozharitskaya et al., 2006). HPTLC approach has been recognized as an excellent device for regular fingerprint analysis due to its speed, convenience (Sunita and Abhishek, 2008; Rathi et al., 2008; Kamboj and Saluja, 2013) and great post-chromatography visualization (Pereira et al., 2004). This is especially vital for screening and examination of crude materials as well as for process control during production (Reich and Schibli, 2007).

Quantification can also be done based on the evaluation of chromatograph with reference compound analysed on the same plate (Zlatkis and Kaiser, 2011). Another advantage of HPTLC is the repeated scanning of the chromatogram with the same or different conditions (Ankli et al., 2008). In addition, detection of contaminant and quantitative determination of reference compounds in sample are generally applied (Ankli et al., 2008; Jadhav et al., 2007).

European Pharmacopoeia, American Herbal Pharmacopoeia (AHP), United States Pharmacopoeia (USP) and Pharmacopoeia of the People's Republic of China have additionally acknowledged HPTLC together with scanning densitometry as the general chromatographic analysis method for CHM (AHP, 2009; Reich and Widmer, 2009).

Guan et al. (2011) did a study on fingerprints for HPTLC of garlic from different geographical populations. A HPTLC fingerprinting study was also carried out by Chothani et al., 2010 to identify, isolate, and quantify the marker in the root, stem, and leaf of Ruellia tuberosa (Acanthaceae). Moreover, Locher et al. 2017 explored HPTLC fingerprinting for the authentication of honey, such as Eucalyptus marginata, Corymbia calophyla, and Leptospermum scoparium. The findings the study by Pedan et al., 2018 highlight the use of a HPTLC as a rapid tool to examine quality parameters in cocoa proanthocyanidins and cocoa processing. The method can be also used as a future quality control tool for rapid resolution for chocolate production.

1.6 Multivariate Analysis

Chemometrics can be characterized as the study of extracting chemically relevant data from multivariate information by utilizing statistical techniques to decrease the dimensionality of the dataset. It offers an instrument to graphically summarise the analytical information to reveal connections amongst samples and to detect characteristic patterns that can be used to distinguish a specific material (Ahad and Nissar, 2017).

To deal with the huge dataset of the metabolomic examination of hundreds of samples and each spectrum comprising hundreds of signals, reasonable chemometric techniques are required (Kim et al., 2010). Quality control of herbal extracts is a challenging attempt because they usually contain various phytochemicals (Kim et al.,

2016). Besides, multivariate statistical analyses are expected to decrease the complexity of the information from metabolic profiling and facilitate the detection of the pattern of changes related to the ecological or hereditary factors in metabolite compositions. Principal Component Analysis (PCA), partial least squares-discriminant analysis (PLS-DA), and PLS-regression (PLS-R) are widely used multivariate data analysis methods (Lee et al., 2017).

1.6.1 Unsupervised Method

As an initial step, an exploratory examination is done in order to research the normal relations between the samples (Ahad and Nissar, 2017). In this manner, unsupervised methods—those in which no prior information of class membership is expected —are a typical initial step in data examination. The most widely recognized unsupervised technique utilized is principal component analysis (PCA), a calculation that reduces a high-dimensional data set to a small number of dimensions that clarify as much of the variation in the data as possible.

This overview provides information not only on the groups of observations, patterns, and exceptions but also on the connections between the observations (Eriksson et al., 2006). PCA is resolved using the score plot that shows the statistical differences between the groups and the loading plot that shows the compounds that are responsible for the contrasts between the groups (Lee et al., 2017).

In a PCA model, the original variables are transformed in new uncorrelated factors that emerge from the straight combination of the original factors: the principal components (PCs). Various PCs are separated in sequence with every essential part representing the maximum of the residual fluctuation in the information. The PCs extraction stops when most of the variance in the original data (typically around 90%) is clarified. The new set of PCs defines therefore a new space where the contribution

of every unique factor to every PC can be easily represented and the connections between the original samples highlighted (Ahad and Nissar, 2017).

While PCA is great at identifying groups and outliers, its outcomes frequently are utilized just in coordinating future examination since its precision can be enhanced by utilizing supervised methods. Although PCs depict the largest portion of the variation, this variation may not closely correspond to the division between classes (Bylesjo et al., 2006). In this way, PCA usually is used only as an initial step to help in the improvement of a model using better-classifying supervised methods (Kosmides et al., 2013).

1.6.2 Supervised Methods

PCA has a cut-off to validate statistical models because it cannot assign the class membership of unknown test samples (Kang et al., 2008a, 2008b). When potential biomarkers have been discovered, supervised methods can be used to expand the separation amongst classes and recognize the most robust biomarkers (Mao et al., 2008). Because supervised methods use data on class membership, they are much better at creating classifiers and predicting where a sample falls with respect to those already classified (Nicholson et al., 2008). Commonly used supervised methods include partial least squares with discriminate analysis (PLS-DA), and SIMCA (Mao et al., 2008).

SIMCA is a classification method for samples based on their similarities to principal component models (Lonni et al., 2003). A principal component is fitted independently to each class in the training set and multivariate confidence envelopes are developed around each class model containing the information points. The strategy is utilized to classify samples into predefined classifications, to provide outlier detection, and to predict external properties (Soares et al., 2008). SIMCA is a soft