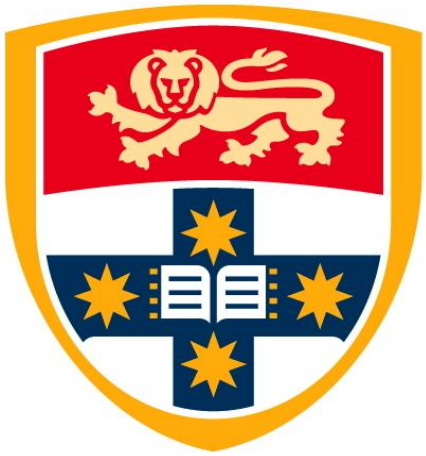
THE UNIVERSITY OF
SYDNEY

The Quality Control of Puerariae Lobatae Radix and Puerariae Thomsonii Radix

Ka Ho Wong



A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Pharmacy

The University of Sydney

NSW, Australia

May, 2014

Declaration

Unless where due acknowledgement has been given, the work contained in this thesis is the result of original investigations by the author carried out in the Faculty of Pharmacy, The University of Sydney, New South Wales, Australia, under the primary supervision of Professor Kelvin Chan, Dr George Li, Dr Valentina Razmovski-Naumovski and Dr Kong Li. I hereby declare that this thesis has not been previously submitted to any university or institution for a higher degree.

_____

Ka Ho Wong

May 2014

Acknowledgements

I would like to especially thank my primary supervisor, Professor Kelvin Chan, for giving me this opportunity to pursue a PhD project under his guidance. Without his support and encouragement, my PhD experience would not have been as enjoyable as it was. I would like to thank my co-supervisors Dr George Li, Dr Valentina Razmovski-Naumovski and Dr Kong Li for their continued support, advice, patience and encouragement during the course of my degree.

I would like to express my gratitude to Dr Samiuela Lee and Mr Jarryd Pearson from the National Institute of Complementary Medicine, University of Western Sydney, for their technical support during the UPLC work. Thank you to Dr Leila Hejazi from the Mass Spectrometry Facility, School of Medicine, University of Western Sydney for her expert technical assistance in the UPLC-MS/MS analyses.

I would like to thank Ms Sanaz Maleki from the Histopathology Laboratory, Discipline of Pathology, Sydney Medical School, The University of Sydney for her expert technical assistance on the microscopic slide preparation. I am also grateful to Dr Elizabeth Carter from the Vibrational Spectroscopy Facility, School of Chemistry, The University of Sydney and Dr Richard Wuhrer from the Advanced Materials Characterisation Facility, University of Western Sydney for their training and technical support during the Raman analysis.

I would also like to acknowledge and thank Dr Antony Kam for his time and patience in teaching me the cell culture technique. I would like to extend my

thanks to the past and present members of the laboratory and other technical staff: Dr Eshaifol Omar, Dr Jun-Lae Cho, Dr Sunny Kim, Dr Yiming Li, Mr Ali Alqahtani and Mr Benjamin Kimble.

Last but not least, I would like to express my deepest thanks to my beloved parents, my sister, family and friends for all their continual support and encouragement during my years of studying in Australia.

Table of Contents

List of Figures

List of Tables

List of Publications

1. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. Comparing morphological, chemical and anti-diabetic characteristics of Puerariae Lobatae Radix and Puerariae Thomsonii Radix. (**Under review**)

2. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. Prediction of total phenolic content and antioxidant capacities in *Pueraria* species using Raman spectroscopy coupled with partial least squares regression analysis. (**Under second review**)

3. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. (2014) Differentiating Puerariae Lobatae Radix and Puerariae Thomsonii Radix using HPTLC coupled with multivariate classification analyses. *J Pharm Biomed Anal.,* 95 11-19.

4. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. (2013). Differentiation of Pueraria lobata and Pueraria thomsonii using partial least squares discriminant analysis (PLS-DA). *J Pharm Biomed Anal.*, 84 5-13.

5. Alqahtani, A., Hamid, K., Kam, A., **Wong, K. H.**, Razmovski-Naumovski, V., Chan, K., Li, K.M., Groundwater, P. W. & Li, G. Q. (2013). The pentacyclic triterpenoids in herbal medicines and their pharmacological activities in diabetes and diabetic complications. *Curr Med Chem.,* 20 908-931.

6. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. (2012) A comparative study on the chemical profiles and cytoprotective effects of the roots of Pueraria lobata and Pueraria thomsonii. *J Pharm Pharm Sci.*, 16 174. (**Abstract**).

7. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. (2012) Chemical profiling of Pueraria lobata and Pueraria thomsonii using total phenol, flavonoid, carbohydrate and antioxidant assays. *Planta Med*., 78 68. (**Abstract**).

8. Li, G., Kam, A., **Wong, K. H.**, Zhou, X., Eshaifol, A., Li, K. M., Razmovski-Naumovski, V. & Chan, K. (2012). Herbal medicines for the management of diabetes. *Diabetes: An old disease, a new insight.* Landes Bioscience. (**Book Chapter**).

9. Li, G., **Wong, K. H.**, Kam, A., Zhou, X., Eshaifol, A. O., Alqahtani, A., Li, M. K., Razmovski-Naumovski, V. & Chan, K. (2012). Anti-inflammatory nutraceuticals and herbal medicines for the management of metabolic syndrome. *Chronic Inflammation: Molecular Pathophysiology, Nutritional and Therapeutic Interventions*. CRC Press. (**Book Chapter**).

10. Chan, H., Poon, J., Poon, K., Chan, K., **Wong, K. H.** & Sze, D. (2012) Evaluation of herbal chromatographic fingerprint identification systems. Bioinformatics and Biomedicine Workshops, *IEEE International Conference Proceeding*. 418 – 423. (**Conference proceeding**).

11. **Wong, K. H.**, Li, G. Q., Li, K. M., Razmovski-Naumovski, V. & Chan, K. (2011). Kudzu root: Traditional uses and potential medicinal benefits in diabetes and cardiovascular diseases. *J Ethnopharmacol.,* 13 584-607.

List of Conference Abstracts

1. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. Application of ATR-FTIR coupled with multivariate analyses to discriminate Puerariae Lobatae Radix and Puerariae Thomsonii Radix. 13th Consortium for Globalization of Chinese Medicine, 27th – 29th Aug. 2014, Beijing, China. (**Poster**)

2. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Qualitative analysis of Puerariae Lobata Radix granules using partial least squares discriminant analysis (PLS-DA).* 2nd China Studies Centre Annual Conference, 29th Nov. 2013, Sydney, Australia. (**Oral**)

3. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Differentiating Puerariae Lobatae Radix and Puerariae Thomsonii Radix using ATR-FTIR coupled with multivariate analysis*. Medicinal Chemistry Symposium, 25th Nov. 2013, Sydney, Australia. (**Poster**)

4. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Qualitative analysis of Puerariae Lobatae Radix granules using partial least squares discriminant analysis.* 12th Consortium for Globalization of Chinese Medicine, 26th – 29th Aug. 2013, Graz, Austria. (**Poster**)

5. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Identification of potential cytoprotective constituents from Gegen (Puerariae Lobatae Radix).* 11th Consortium for Globalization of Chinese Medicine, 21st – 23rd Aug. 2012, Macau, China. (**Poster)**

6. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Chemometrics evaluation of aqueous extracts from kudzu root and its commercial granules.* Good Practice in Traditional Chinese Medicine Research Congress, 30th April 2012, Leiden, Netherland. (**Poster**)

7. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Optimisation of ultrasound-assisted extraction of kudzu (Pueraria lobata) root using response surface methodology*. Integrative Medicine Symposium, 21st – 23rd Oct. 2011, Sydney, Australia. (**Poster**)

8. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Quality standardization of Gegen.* 10th Consortium for Globalization of Chinese Medicine, 26th – 28th Aug. 2011, Shanghai, China. (**Poster**)

9. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Authentication of kudzuvine root based on pharmacognostic characteristics*. Australasian Pharmaceutical Science Association Annual Conference, 6th – 9th Dec. 2010, Brisbane, Australia. (**Oral**)

10. **Wong, K. H.**, Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. *Studies on macroscopic, microscopic and chromatographic identification of Pueraria lobata.* Bosch Institute Young Investigators Symposium, 30th Nov. 2010, Sydney, Australia. (**Poster**)

List of Abbreviations

ABTS	2,2’-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
ANOVA	Analysis of variance
Ara	Arabinose
Bcl-2	B-cell lymphoma-2
BP	British Pharmacopoeia
Cal	Calibration set
CART	Classification and regression tree
CMM	Chinese Materia Medica
COW	Correlation optimised warping
CUPRAC	Cupric reducing antioxidant capacity
CV	Cross-validation set
DE	Dried extract
DM	Dried material
DPPH	2,2-Diphenyl-1-picrylhydrazyl
EP	European Pharmacopoeia
ER	Error rate
EtOH	Ethanol
FN	False negative
FP	False positive
g	Gram
GA	Genetic algorithm
GAE	Gallic acid equivalent

Gal	Galactose
Glu	Glucose
H	Hydrogen
HKCMMS	Hong Kong Chinese Materia Medica Standards
HPLC	High-performance liquid chromatography
hr	Hour
I.D.	Internal diameter
JP	Japanese Pharmacopoeia
K-S	Kennard-Stone algorithm
KNN	K-nearest neighbors
L	Litre
LDA	Linear discriminant analysis
LOD	Limit of detection
LOOCV	Leave-one-out cross-validation
LOQ	Limit of quantification
LV	Latent variable
M	Molar
Me	Methyl
mg	Milligram
min	Minute
MIR	Mid-infrared
mL	Millilitre
mM	Millimolar
MS	Mass spectrometry
MSC	Multiplicative scatter correlation

MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium bromide
N/A	Not available
NER	Non-error rate
NIR	Near-infrared
NMR	Nuclear magnetic resonance
OH	Hydroxyl
OMe	Methoxyl
PC	Principal component
PCA-DA	Principal component analysis-discriminant analysis
PCA	Principal component analysis
PLR	Puerariae Lobatae Radix, the root of *Pueraria lobata* (Willd.) Ohwi (Fabaceae)
PLS	Partial least squares
PLS-DA	Partial least squares-discriminant analysis
PLSR	Partial least squares regression
PPRC	Pharmacopoeia of the People's Republic of China
PTR	Puerariae Thomsonii Radix, the root of Pueraria thomsonii Benth. (Fabaceae)
PVDF	Polyvinylidene fluoride
QE	Quercetin equivalent
r^2 cal	Regression coefficient of calibration set
r^2 cv	Regression coefficient of cross-validation set
r^2 val	Regression coefficient of validation set
RCT	Randomised controlled trial

RE	Relative error
Rf	Retention factor
Rha	Rhamnose
RMSE	Root mean square error
RMSEC	Root mean square error of calibration
RMSECV	Root mean square error of cross-validation
RMSEV	Root mean square error of validation
ROS	Reactive oxidation species
RPD	Ratio of performance to deviation
RSD	Relative standard deviation
SD	Standard deviation
SIMCA	Soft independent modelling of class analogy
SNV	Standard normal variate
SPA	Successive projection algorithm
TCM	Traditional Chinese medicine
TEAC	Trolox equivalent antioxidant capacity
TFC	Total flavonoid content
TLC	Thin-layer chromatography
TN	True negative
TP	True positive
TPC	Total phenolic content
U.S.A.	United States of America
UPLC	Ultra-performance liquid chromatography
Val	Validation set
WHO	World Health Organization

µg	Microgram
µL	Microliter
µM	Micromolar

List of Symbols

A	Number of calculated principal components
Abs	Absorbance
b	Regression coefficients
^{o}C	Degree Celsius
C_{cal}	Concentration of chemical calculated in the testing sample using the calibration curve
$C_{extract}$	Original concentration of the testing sample
c	Velocity of light
$conc$	Concentration
cov	Covariance
E	Residue matrices of the X matrix
E_o	Initial energy
E_v	Energy change
e_{ij}^{i}	Residual between object n and principal component
F	Residue matrices of the X matrix
f	Frequency
g	Number of sample class
I	Intensity in the sample spectrum
I_c	Number of sample in the calibration set
I_o	Intensity in the background spectrum
I_p	Number of sample in the validation set
k	Number of latent variable

l	Pathlength
m	Number of variable
N_i	Number of data point in a sample class
n	Number of samples
P	Loading of the X matrix
P^T	Transposed of loading matrix (A x n)
Q	Loading of the Y matrix
s_n^K	Residual standard deviation of object n
s_0^K	Residual standard deviation of an unknown object n_K
T	Score of the X matrix
t	Transposition
U	Score of the Y matrix
v	Wavenumber
W^*	X-weights matrix (m x k)
X	Fingerprint X matrix (n x m)
$\bar{x}$	Mean of the whole data set
$\bar{x}_i$	Class mean
Y	Response Y matrix (n x 1)
y_i	Measured value of sample i
$\hat{y}_i$	Calculated value of sample i
$\bar{y}$	Mean of the reference measurement results
y_{ic}	Measured value of sample i in the calibration set
$\hat{y}_{ic}$	Calculated value of sample i in the calibration set
$\hat{y}_{icv}$	Calculated value of sample i in the cross validation set

y_{ip}	Measured value of sample i in the validation set
$\hat{y}_{ip}$	Calculated value of sample i in the validation set
ε	Molar absorptivity
λ	Wavelength

Summary

Puerariae Lobatae Radix (PLR), the root of *Pueraria lobata* (Willd.) Ohwi, is a traditional Chinese medicine (TCM) used for the treatment of diabetes and cardiovascular diseases. Puerariae Thomsonii Radix (PTR), the root of *P. thomsonii* Benth., is a closely related species of PLR. In clinical practice, PLR and PTR are used interchangeably. The confusing nomenclature and identical descriptions of PLR and PTR in the pharmacopoeias and monographs has made it difficult to differentiate the species. Therefore, the aim of the study was to differentiate PLR from PTR using qualitative and quantitative methods and various analytical instruments in coupled with chemometrics analysis. The microscopic features of these two herbs were statistically analysed and confirmed using colorimetric assays which quantified the amount of starch and dietary fibre present (Chapter 2). PLR was found to have larger fibre bundles (PLR: 32.6800 ± 2.8780; PTR: 16.5900 ± 0.9982; $p<0.01$) and xylem vessels (PLR: 0.1390 ± 0.0184 mm; PTR: 0.0471 ± 0.0109 mm; $p<0.01$), and less total starch content ($p<0.001$) and total dietary fibre content (PLR: 4.2886 ± 0.3466 g/100 g DM; PTR: 12.4148 ± 0.4541 g/100 g DM; $p<0.001$) as compared to PTR. It was found that the diameter of starch granules was insignificant between these two species ($p>0.05$) and should not be recommended as a criterion for differentiating PLR from PTR.

Isoflavonoids in PLR and PTR are believed to be associated with their chemical activities. However, no correlation has been made between the amount of chemicals present in PLR and PTR and their respective antioxidant capacity. Since extracts are used in clinical practice, the major chemical constituents were

quantified and identified using ultra-performance liquid chromatography (UPLC)-mass spectrometry (Chapter 3). The amount of puerarin, daidzin and genistein in PLR was 11.95, 8.64 and 1.38 times, respectively, greater than those in PTR. Genistin was found in PLR and was absent in PTR. In addition, the colorimetric assays revealed that the total flavonoid content (TFC) and DPPH antioxidant capacity of PLR was 4.42 times and 4.91 times, respectively, greater than PTR. The high correlation between TFC and DPPH activity suggested that the antioxidant capacity of PLR and PTR is influenced by the amount and type of chemicals (such as puerarin) present in the plant.

Puerarin is the major chemical consistent and is recommended as the chemical marker for the authentication of PLR and PTR. However, the use of a single chemical marker does not completely reflect the complexity of a herb. To examine this issue, partial least squares-discriminant analysis (PLS-DA) models based on the entire UPLC chromatographic fingerprint were constructed and compared to puerarin alone (Chapter 4). The results revealed that the PLS-DA model using the entire chromatographic fingerprint achieved a 100% correct species classification rate in the validation set, whereas a 86.6% correct species classification rate was achieved using puerarin, suggesting that the use of puerarin alone was not sufficient to differentiate PLR from PTR. The established PLS-DA model using the entire fingerprint was subsequently applied to analyse the raw material used in manufacturing PLR granules. It was found that 4 out of the 17 brands of granules used PTR as the raw material, but labeled as PLR.

To investigate a simpler technique of acquiring reliable chromatographic fingerprints of PLR and PTR, high performance thin-layer chromatography (HPTLC) was compared to UPLC (Chapter 5). The results demonstrated that HPTLC classification models, including K-nearest neighbors, PLS-DA, principal component analysis-discriminant analysis, support vector machine-discriminant analysis, were as effective as the UPLC models in differentiating PLR from PTR, and achieved a 100% correct species classification rate with zero error rate.

As an alternative to chromatographic analyses, Raman spectroscopic fingerprints of PLR and PTR were correlated to the total phenolic content, ABTS antioxidant capacity and cupric reducing antioxidant capacity using partial least squares (PLS) regression algorithm (Chapter 6). The high correlation coefficient and ratio of performance to deviation values indicated that the PLS regression models were accurate and robust in predicting the TPC and antioxidant capacities in both PLR and PTR.

The existing literature has outlined biological activities of PLR including anti-diabetic, anti-hypertensive and anti-inflammatory effects. However, it has not been shown whether the chemical differences between these two species could impact on their respective pharmacological actions. Thus, enzymatic diabetic assays and *in vitro* endothelial and cancer cellular models were employed to investigate the pharmacological differences between PLR and PTR (Chapter 7). The results revealed that PLR had a superior inhibitory effect on the activity of α-amylase and α-glucosidase, cyto-protective effect against hydrogen peroxide-induced cell death

on human endothelial EA.hy926 cells and cytotoxic effect on human prostate cancer PC3 cells as compared to PTR.

In summary, this thesis revealed that PLR is morphologically, chemically and pharmacologically different to PTR and the pharmacopoeias should be reviewed to reflect this. The use of PLR and PTR interchangeably in clinical practice should be reviewed. UPLC and simple analytical instruments such as HPTLC and Raman spectroscopy presenting the whole chemical profile combined with chemometrics can effectively differentiate PLR from PTR. It is anticipated that this thesis will provide further insight into the comprehensive nature of the quality control of herbal medicines, particularly the application of chemometrics to discriminate large sample sizes and reduce analysis time. The methods outlined in this thesis can be readily adopted by the pharmaceutical and herbal industry for discerning herbal materials. This will ultimately improve the quality, safety and efficacy of herbal products for consumers.

Chapter One

Introduction

The materials presented in this chapter were extracted from **Wong, K. H.,** Li, G. Q., Li, K. M., Razmovski-Naumovski, V. & Chan, K. (2011). Traditional uses and potential medicinal benefits in diabetes and cardiovascular diseases. *J Ethnopharmacol.*, 13 584-607.

1.1 Puerariae Lobatae Radix and Pueraria Thomsonii Radix

1.1.1 Traditional Chinese medicine

Traditional Chinese medicine (TCM) is an integral part of Chinese culture, with the fundamental concept based on the Five Elements theory (Fire, Earth, Metal, Water and Wood) and the principle of Yin and Yang. Ancient Chinese believed that all natural phenomena could be divided into Yin and Yang, which are the two opposite, complementary and exchangeable aspects of nature. Everything in the universe comprises of the Five Elements, and the universe is constantly changing towards a dynamic balance or harmony (Jiuzhang and Lei, 2010; Zhu and Wang, 2011).

In TCM, an individual is considered 'healthy' when all the Five Elements and Yin and Yang are in balance (Cheng, 2000; Lu et al., 2004). Balance is considered to be the complex interplay between body, mind and the environment, while disease is a consequence of imbalances in Yin and Yang or flow of blood and Qi, which is caused by a variety of external or internal factors (Wang and Zhu, 2011). Each disease may have a set of identical symptoms among different patients, but the cause of imbalance of the same disease can be quite different. TCM approach recognises the uniqueness of each individual and the necessity to develop a personalised therapy to obtain an optimal homeostasis between human and nature, based on a multi-component and multi-targeted treatment (Chan, 1995; Chan, 2005; Jiang, 2005).

TCM treatments aim to expel or suppress the disturbance and restore the balance and harmony in the body. Herbal medicine is one of the most common techniques

used in clinical practice. (Jiuzhang and Lei, 2010; Liao, 2011; Zhu and Wang, 2011). The dosage forms used in Chinese herbal medicine comprise of decoctions, granules, pills, powder, tinctures, lotions and pastes. The most common one is the use of decoction, which is made by extracting the herbs in boiling water (Jiuzhang and Lei, 2010). According to TCM theory, a herb is rarely used alone and is normally combined with other herbs and presented as a formula, which comprises of 2 to 30 herbs (Liao, 2011; Zhu and Wang, 2011).

The holistic approach of TCM can also be applied to explaining the pharmacological effect of a single herb, as it is believed that the therapeutic effect is due to the presence of all the chemical constituents within a herb which might act synergistically, addictively or antagonistically (Chan, 1995; Wang et al., 2009).

1.1.2 Puerariae Radix

Puerariae Radix is a Chinese herbal medicine and has been used for treating diabetes and cardiovascular diseases. It is native to East Asia and has been utilised as a food source, fodder and medicine for thousands of years. Puerariae Radix is widely distributed around the world and is also known as kudzuvine, kudsu, wa yaka, aka, nepalem, Japanese arrowroot, kuzu (Japanese), kudzu comun (Spanish), vigne japonaise (French) and kopoubohne (German) (EPPO, 2007). Puerariae Radix powder and extract are sold in the United States of America (U.S.A.), United Kingdom and Australia as a supplement. Puerariae Radix is often used in combination with other herbs (e.g. *Salvia miltiorrhiza*, *Bacopa monnieri*, *Ginkgo biloba*, *Silybum marianum* and *Salix alba*) for relieving hangover, fever and flu, improving liver function, enhancing the detoxification

processes, regulating cardiac functions and aiding weight loss (Amazon, 2011; Healthstore, 2011; TGA, 2011).

1.1.2.1 Botanical descriptions of Puerariae Radix

Puerariae Radix refers to the root of *Pueraria lobata* (Willd.) Ohwi and *P. thomsonii* Benth.. These plants are climbing, semi-woody and perennial vines, with hairy rusty-brown stems. They form large fleshy root tubers up to 2 m long, 18 – 45 cm in diameter and can reach a depth of 1 – 5 m (EPPO, 2007). Leaves are arranged alternately along the stem with three leaflets (Banks, 2008). Flowers are pea-like, pink to purple with yellow in the centers. They are highly fragrant with sweet grape-like scent and are borne in long hanging panicles at nodes on the stems. Fruits are brown, flattened, hairy seedpods containing 3 – 10 reddish brown, ovoid to ellipsoid, slightly flattened seeds (Csurhes, 2008).

Figure 1.1 Whole plant of *Pueraria lobata* (Willd.) Ohwi (Chang and Ho, 2009).

1.1.2.2 Taxonomy and nomenclature of Puerariae Radix

Pueraria species belongs to the Fabaceae family, Papilionoideae subfamily, Phaseoleae tribe, Glycininae subtribe and *Pueraria* genus. Seventeen recognised species were described by Van der Maesen under the *Pueraria* genus, however, the taxonomy of plants in this genus is unclear, with multiple synonyms and multiple varieties within the species (Van der Maesen, 1985). Three varieties have been illustrated within *Pueraria montana*, including *P. montana* var. *lobata* (Willd.) Maesen & SM Almeida, *P. montana* var. *montana* (Loureiro) Maesen, and *P. montana* var. *chinensis* (Bentham) Maesen & SM Almeida (equivalent to *P. thomsonii* Benth.) (EPPO, 2007). It is suggested that the morphological characteristics of these three varieties are highly variable.

However, it has been considered that these three varieties suggested by Van der Maesen are actually comprised of two or three species (Ohashi et al., 1988; Van der Maesen, 1985, 2002). *Pueraria montana* var. *lobata* (Willd.) Maesen & SM Almeida is the official name employed by the International Union for Conservation of Nature and Department of Agriculture in the U.S.A., yet *Pueraria lobata* (Willd.) Ohwi currently remains a legitimate name and is used in various pharmacopoeias, monographs and publications (Csurhes, 2008; IUCN, 2005). Recent DNA sequencing analyses revealed that most of the *Pueraria* species grown in the U.S.A. are *P. lobata* (Willd.) Ohwi and are genetically close to the *P. lobata* (Willd.) Ohwi from China. However, high genetic differentiation was observed among *P. lobata* (Willd.) Ohwi*, P. thomsonii* Benth. and *P. montana,* suggesting that these plants are three distinctive species within the *Pueraria* genus (Jewett et al., 2003; Sun et al., 2005; Sun et al., 2006). The *Pueraria* species grown in Australia and

Europe are reported to be *P. lobata* (Willd.) Ohwi by various authorities and organisations (Banks, 2008; Csurhes, 2008; EPPO, 2007).

The Pharmacopoeia of the People's Republic of China (PPRC) is the official pharmacopoeia used in China for the authentication of herbal medicines and pharmaceutical products. "Radix Puerariae" is the original pharmaceutical Latin name used to collectively describe the roots of *P. lobata* (Willd.) Ohwi and *P. thomsonii* Benth (Table 1.1). "Radix Puerariae" was the name used to describe both species in the first edition (1965) of the PPRC until the seventh edition at 2000. From the eighth edition (2005), the Pharmacopoeia's Commission decided to separate "Radix Puerariae" into two monographs namely, Radix Puerariae Lobatae and Radix Puerariae Thomsonii (PPRC, 1997, 2000, 2005). In the latest edition of the PPRC (2010), the Commission rearranged the order of the pharmaceutical Latin name, and the plants are now named Puerariae Lobatae Radix and Puerariae Thomsonii Radix (PPRC, 2010a, b). The first part of the pharmaceutical Latin name (Puerariae) refers to the genus of the herb, whereas the last part (Radix) refers to the part of the herb used. If there are more than one species used as a herbal medicine, the species name is included (Lobatae or Thomsonii).

In addition to its PPRC description, Puerariae Radix is included in various herbal pharmacopoeias around the world. Table 1.2 illustrates the nomenclature of Puerariae Radix in the latest edition of the European Pharmacopoeia (EP), British Pharmacopoeia (BP), Hong Kong Chinese Materia Medica Standards (HKCMMS) monograph and Japanese Pharmacopoeia (JP). For the BP, EP and HKCMMS

two distinct monographs describe the roots of *P. lobata* (Willd.) Ohwi and *P. thomsonii* Benth. However, only the root of *P. lobata* (Willd.) Ohwi is listed as a monograph in the JP.

Thus, in this thesis, Puerariae Radix refers to the two species that are normally used in TCM practice, Puerariae Lobatae Radix and Puerariae Thomsonii Radix. Puerariae Lobatae Radix represents the root of *P. lobata* (Willd.) Ohwi, which is equivalent to *P. montana* var. *lobata* (Willd.) Maesen & SM Almeida, whilst Puerariae Thomsonii Radix represents the root of *P. thomsonii* Benth., which is equivalent to *P. montana* var. *chinensis* (Bentham) Maesen & SM Almeida.

Table 1.1 Summary of the nomenclature of Puerariae Radix in various editions of the PPRC (PPRC, 1997, 2000, 2005, 2010a, b).

	Prior to 2005	2005		PPRC 2010	
Pharmaceutical Latin name	"Radix Puerariae"	Radix Puerariae Lobatae	Radix Pueraria Thomsonii	Puerariae Lobatae Radix	Pueraria Thomsonii Radix
Species	Dried root of *P. lobata* (Willd.) Ohwi or *P. thomsonii* Benth.	Dried root of *P. lobata* (Willd.) Ohwi	Dried root of *P. thomsonii* Benth.	Dried root of *P. lobata* (Willd.) Ohwi	Dried root of *P. thomsonii* Benth.
Common name	*P. lobata*: Gegen or Kudzuvine Root *P. thomsonii*: Starchy Puerariae Radix	Gegen or Yege	Fenge or Thomson Kudzuvine Root	Gegen or Yege	Fenge or Thomson Kudzuvine Root

Table 1.2 Summary of the nomenclature of Puerariae Radix in various pharmacopoeias (BP, 2014a, b; EP, 2012a, b; HKCMMS, 2010a, b; JP, 2011).

	EP/BP		HKCMMS		JP
Pharmaceutical Latin name	Puerariae Lobatae Radix	Puerariae Thomsonii Radix	Radix Puerariae Lobatae	Radix Puerariae Thomsonii	Puerariae Radix
Species	Dried root of *P. lobata* (Willd.) Ohwi	Dried root of *P. thomsonii* Benth.	Dried root of *P. lobata* (Willd.) Ohwi	Dried root of *P. thomsonii* Benth.	Dried root of *P. lobata* (Willd.) Ohwi
Common name	Kudzuvine Root	Thomson Kudzuvine Root	Gegen	Fenge	Pueraria root

1.1.3 History, traditional and current uses of Puerariae Radix

1.1.3.1 Puerariae Lobatae Radix (PLR)

Puerariae Lobatae Radix (PLR) is one of the earliest medicinal herbs employed in ancient China (Prasain et al., 2003). It is also known as Gegen (葛根), Yege (野葛) or kudzuvine root. The oldest Chinese written reference of the use of PLR is described in the *Classic of Poetry (Shih Ching)* dated between 1000 BC and 500 BC. PLR was first documented as a medicinal herb in the *Divine Husbandman's Classic of Chinese Materia Medica (Shen Nong Ben Cao Jing)* during the Western Han Dynasty (206 BC – 8 AD) for the relief of fever, diarrhoea and vomiting (Yan et al., 2004). The herb is mentioned in the classic medical book *Treatise on Fevers (Shang Han Lun)* in 200 AD, which described the use of Gegen Tang, a decoction prepared from PLR, for the treatment of neck stiffness, lack of perspiration and aversion to air drafts (Fang, 1980). In 600 AD, it was used as an anti-intoxication agent to treat alcohol-related problems, and was recommended as an anti-dipostropic agent by Li Shi-Zhen in 1200 AD (Keung and Vallee, 1998). In TCM practice, the herb is described as cool in nature, sweet and acrid in taste and used to promote Spleen Yang to arrest diarrhoea and encourage the production of body fluid.

During the Jomon period (10, 000 BC – 400 BC), *P. lobata* (Willd.) Ohwi. was introduced into Japan from China and was used for fences, baskets, trunks, fishing nets and clothes. PLR was also included in the Japanese Pharmacopoeia for the treatment of various ailments and disorders during the Edo period (1600 AD – 1867 AD) (Bodner and Hymowitz, 2002). *P. lobata* (Willd.) Ohwi. was first introduced into the U.S.A. in 1876 as a garden ornamental plant (Everst et al.,

1991; Pieters, 1932; Stevens, 1976). However, *P. lobata* (Willd.) Ohwi. soon became a problem as it invaded and replaced native vegetation and road, natural and industrial forest sites (Csurhes, 2008; Tabor and Susoot, 1941). It has been listed as one of the 100 world's worst invasive species and has been naturalised in the U.S.A., South America, South Africa, Australia and Europe (IUCN, 2005). It is now abundant in the southeastern states of the U.S.A. and is recognised as a noxious weed in five states, including Florida, Kansas, Oregon, Pennsylvania and West Virginia (Csurhes, 2008; Everst et al., 1991; US Congress, 1993). In Australia, *P. lobata* (Willd.) Ohwi. exists as small infestations scattered along the east coast of Queensland and north-east coast of New South Wales (Banks, 2008; Turnbull, 2004). Specific legislations were established to control the introduction, storage, supply and transportation of *P. lobata* (Willd.) Ohwi. in New South Wales and Queensland (DPI&F, 2008; NWO, 2008).

1.1.3.2 Puerariae Thomsonii Radix (PTR)

Puerariae Thomsonii Radix (PTR), also known as Fengen (粉葛) or thomson kudzuvine root, is a medicinal and edible plant used extensively in southern Chinese cuisines (Chen et al., 2006). *P. thomsonii* Benth. is widely distributed and cultivated in the northeastern and southeastern provinces of China. The first documentation of PTR cultivation dated back to the Jin Dynasty near the current Hengfeng County of the Jiangxi province. PTR is regarded as an imperial tribute to the Emperor since the Song Dynasty (Qu et al., 2011). Currently, the cultivation of PTR in three regions including the Hengfeng County in Jiangxi Province and Foshan and Shaoguan City in Guangdong Province of China is registered as a Geographical Indication and is protected by the Intellectual Property Law in China

(Qu et al., 2011). PTR is used for cooking soup (胡蘿蔔玉米粉葛) by mixing with corn, carrot and pork tenderloin and for cooking a dinner dish (粉葛燜五花肉) made by braising PTR with pork belly or chicken (Xinshipu, 2014).

1.1.4 Clinical applications of PLR and PTR

The most common dosage form is decoction pieces, while granules and tincture are less common. The PLR and PTR decoction pieces are often used in combination with other herbs as a formula in clinical practice. For example, Gegen Tang is a TCM formula consisting of seven herbs, including PLR, Cinnamomi Cortex, Zizyphi Fructus, Poeoniae Radix, Ephedrae Herba, Zingiberis Rhizoma and Glyrrhizae Radix for treating fever and the common cold.

1.1.4.1 Cardiovascular diseases

In a randomised control trial (RCT), the effect of Salviae Miltiorrhizae Radix and PLR extract on vascular function and structure was investigated (Tam et al., 2009). One hundred patients (mean age 58 ± 8 years) with documented coronary artery disease were divided into treatment and placebo group. In the treatment group, six capsules, each containing 500 mg water extract of Salviae Miltiorrhizae Radix (350 mg) and PLR (150 mg), were given daily for 24 weeks. At the end of the period, there were a mild decrease in plasma low-density lipoprotein, improvement in brachial flow-mediated dilation and carotid intima-media thickness in the treatment group. The study was followed by an optional open-label trial for six months. Patients in the treatment group (n=45) received 1.5 g water extract of Salviae Miltiorrhizae Radix (1.05 g) and PLR (450 mg) daily. Further improvement

in both brachial flow-mediated dilation and carotid intima-media thickness were observed in the treatment group as compared to the placebo group (n=47) (Tam et al., 2009).

The efficacy of PLR on the management of hypertension was evaluated among 69 patients with Stage I/II hypertension. Patients were randomly divided into treatment and control group. In the treatment group (n=46), patients were given 200 mL Gegen Qushi decoction, containing PLR, Codonopsis Radix, Atractylodes Radix, Wolfiporia Radix, Massa Medicata Fermentata, Amomi Fructus, Zingiberis Rhizoma, Aucklandiae Radix and Citri Reticulatae Pericarpium. The decoction was administrated for two times per day for four weeks. Patients in the control group (n=23) were treated with the compound anti-hypertensive tablet (0.1 mg reserpine, 12.5 mg dihydralazine, 12.5 mg hydrochlorothiazide and 3 mg chlordiazepoxide) twice daily. The results revealed that the decoction significantly reduced serum total cholesterol and total triglyceride level and improved hypertensive symptoms such as chest tightness, fatigue, tiredness and drowsiness (Liu, 2004). Similar results were observed in another trial; the administration of Gegen Tianma Gouteng decoction twice daily for four weeks considerably improved general well-being and relieved hypertensive symptoms (Yu, 2007).

1.1.4.2 Cerebrovascular diseases

A RCT was carried out among 108 patients after cerebral infarction. Both treatment (n=58, mean age 61.30 ± 7.53 years) and control (n=50, mean age 60.93 ± 7.86 years) groups were treated with intravenous infusion of citicolin 500 mg or troxerutin 250 mg daily for four weeks. Patients in the treatment group were

given a water extract dose of 100 mL Gegen Qinlian decoction (PLR, Scutellariae Radix, Coptidis Rhizoma, Glycyrrhizae Radix and Zingiberis Rhizoma) twice daily. At the end of the study, there was a significant reduction of the plasma fibrinogen and fibrin D-dimer levels and improvement of cognitive function as compared to baseline. The plasma fibrinogen level in the treatment group was noticeably lower than that in the control group (Xu and Wu, 2009).

1.1.4.3 Type 2 diabetes mellitus

In a RCT, the anti-diabetic effect of Gegen Qinlian decoction combined with insulin treatment on type 2 diabetic patients with dampness-heat syndrome was investigated. The treatment group (n=14) was given Gegen Qinlian decoction and insulin, whereas the conventional group was given insulin alone. It was found that the daily average insulin dosage used in the treatment group was significantly lower as compared to the conventional group (Zeng et al., 2006). In a retrospective study, the effect of Gegen Qinlian decoction at different doses on type 2 diabetic patients was examined. The patients were randomly divided into three groups and were treated with low- (n=20; 100 mL), medium- (n=19; 300mL) and high-dose (n=15; 500 mL) of Gegen Qinlian decoction once daily. After three months of treatment, there was a significant reduction in the fasting blood glucose, postprandial blood glucose and haemoglobin A1c levels as compared to the baseline in all three treatment groups. The diabetic symptoms such as thirst and night urination were significantly improved after treatment (Zeng et al., 2006).

1.1.5 Chemical analysis of PLR and PTR

1.1.5.1 Flavonoids

Phenolic compounds or polyphenols are by-products of the secondary metabolism of plants. Collectively, phenolic compounds or polyphenols are naturally occurring compounds that possess an aromatic ring bearing one or more hydroxyl substituents, including functional derivatives such as esters, methyl ethers and glycosides (Crozier et al., 2009). They are widely dispersed throughout the plant kingdom, with more than 8,000 phenolic structures reported in the literature (Quideau et al., 2011).

Flavonoids are the most common and widely distributed phenolic compounds, which comprise of 15 carbon atoms in the skeleton as shown in Figure 1.2. They are arranged as diphenylpropanes (C_6-C_3-C_6) and consist of two aromatic rings connected by three carbon atoms and are usually in the form of an oxygenated heterocycle (Crozier et al., 2009). The aromatic A ring is derived from either resorcinol or phloroglucinol synthesised in the acetate/malonate pathway, whereas the aromatic B ring is derived from phenylalanine in the shikimate pathway (Bravo, 1998; Crozier et al., 2009). The variation in substitution patterns in the flavonoid skeleton gives rise to a variety of flavonoid sub-classes, including flavonols, flavones, flavan-3-ol, anthocyanidins, flavanones, isoflavones, dihydroflavonols, flavan-3,4-diols, chalcones, dihydrochalcones and aurones.

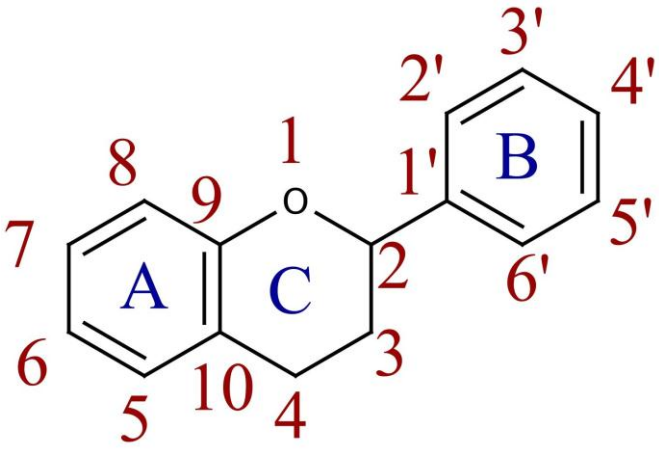


Figure 1.2 Basic skeleton and numbering system of flavonoids.

1.1.5.2 Isoflavonoids

Isoflavonoids are one of the major sub-classes within flavonoids and account for about one-fifth of the total flavonoids reported in the literature (Veitch, 2007, 2009, 2013). They are primarily distributed in the subfamily Fabaceae of the family Leguminosae, but they are also found in some other angiosperm families such as Rutaceae, Mysticaceae, Iridaceae and Rosaceae (Reynaud et al., 2005). As shown in Figure 1.3, the major difference that separates isoflavonoids from its counterpart flavonoids is the position of the aromatic B ring, which is attached to the C-3 position instead of the C-2 position as in flavonoids (M. Boland and M. X. Donnelly, 1998). Hydroxylation, methoxylation or methylenedioxy substitution, prenylation (addition of hydrophobic molecules) and glycosidation are the major processes involved in the formation of different sub-classes of isoflavonoids. Glucose is the dominant sugar moiety. Other minor sugar moieties include rhamnose, galactose, xylose, apiose, arabinose and cymarose. The aglycone is the most common form of isoflavonoid in plants, whereas glycosidic derivatives are less common (Falcone Ferreyra et al., 2012).

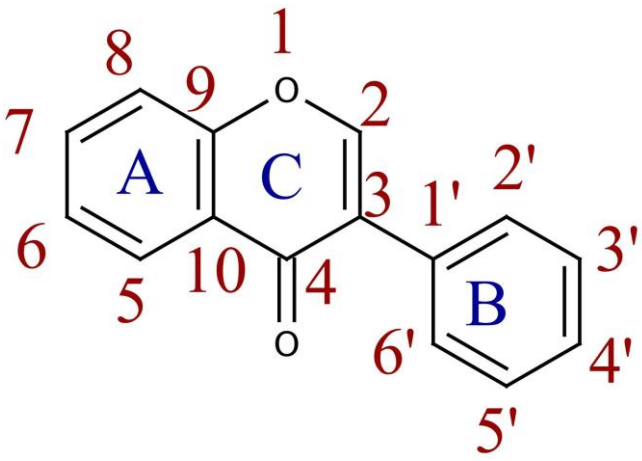


Figure 1.3 Basic skeleton and numbering system of isoflavonoids.

1.1.5.3 Triterpenoids

Triterpenoids are widely distributed in the plant and marine animal kingdoms. The most common form found in plants is the glycosylate triterpenoid, which is also known as saponin. The basic structure of saponins includes a hydrophobic triterpenoid moiety (sapogenin) and one or more hydrophilic sugar moieties (Connolly and Hill, 2008, 2010). Triterpenoids are comprised of six isoprene units $(C_5H_8)_6$ synthesised from either the cytosolic mevalonic acid or the plastidial methylerythritol phosphate pathway. The triperpenoids can be broadly divided into three major sub-classes, including acyclic, tetracylic and pentacyclic. The pentacyclic triterpenoids are the most abundant, in which oleanane-type pentacyclic triterpenoids account for over 50% of the triterpenoids found in leguminous plants (Hill and Connolly, 2011, 2012, 2013). The basic chemical skeleton of oleanane-type pentacyclic triterpenoids is presented in Figure 1.4. Oleanane-type pentacyclic triterpenoids consist of five six-membered carbon rings, namely A, B, C, D and E ring. Rings A/B, B/C and C/D are generally *trans*-linked, whereas rings D/E are *cis*-linked (Alqahtani et al., 2013). Hydroxylation commonly

occurs at C-2, C-3, C-7, C-11, C-15, C-17 and C-22 positions. Triterpenoid glycosides are formed by the addition of one or more sugar moieties to these hydroxyl groups. Common sugar moieties include glucose, rhamnose, galactose, arabinose and glucuronic acid (Vincken et al., 2007).

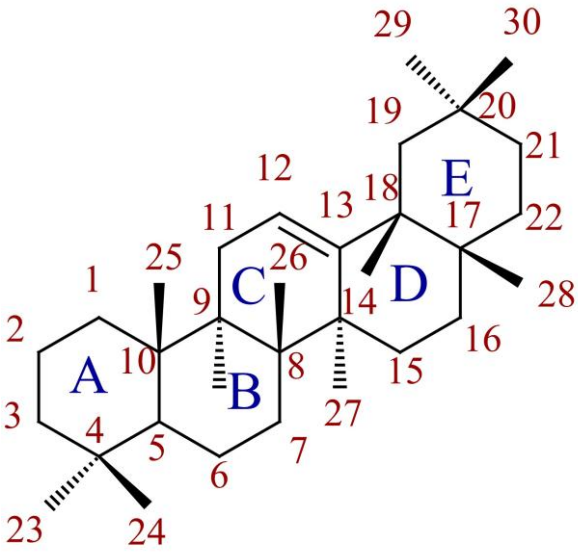


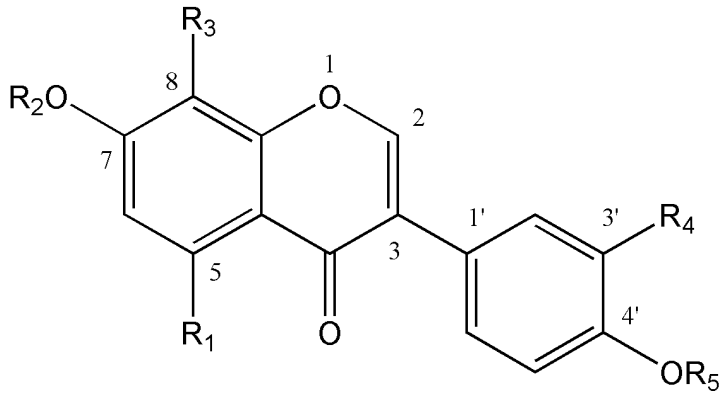
Figure 1.4 Basic skeleton and numbering system of oleanane-type pentacyclic triterpenoids.

1.1.5.4 Major chemical constituents in PLR and PTR

PLR and PTR are a rich source of phenolic compounds, including isoflavonoids, isoflavonoid glycosides, coumarins, puerarols, 2-butenolides and their derivatives. Figure 1.5 illustrates the chemical structures of isoflavonoids and isoflavonoid glycosides isolated from PLR and PTR, while Figure 1.6 illustrates the chemical structures of other minor phenolic compounds. Isoflavonoids and their glycosides are the major constituents (Rong et al., 2002). Puerarin (7-Hydroxy-3-(4-hydroxyphenyl)-8-[(3R,4R,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]chromen-4-one) is the most abundant secondary metabolite and was first isolated and identified from PLR in the early 1950's (Shibata et al., 1959). Genistin

(5-hydroxy-3-(4-hydroxyphenyl)-7-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-4-one) and daidzin (3-(4-hydroxyphenyl)-7-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-4-one) and their aglycones, genistein and daidzein, respectively, were originally detected in soybeans (Walz, 1931), whereas other components such as formononetin and biochanin are widely distributed in various leguminous plants (Price and Fenwick, 1985).

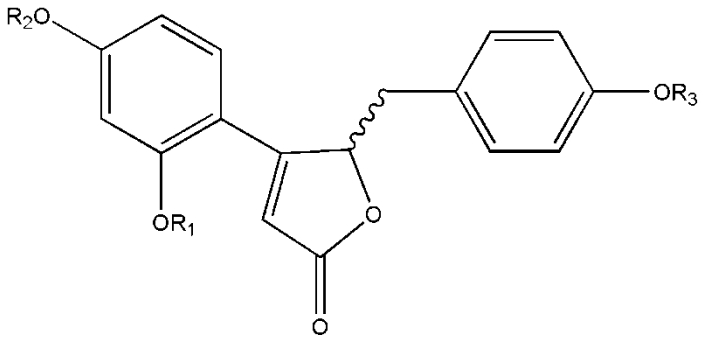
Other prominent chemicals detected in PLR and PTR include triterpenoids and triterpenoid glycosides are shown in Figure 1.7. Oleanene-type triterpenes and triterpenoid glycosides have been identified from PLR, namely kudzusaponin, kudzusapogenol and soyasapogenol. Soyaspogenol A is the aglycone of kudzusaponins SA_{1-4} and soyasaponin A_3, whereas the glycosylation of aglycone kudzusapogenol A forms kudzusaponins A_{1-5} (Rong et al., 2002). Furthermore, PLR and PTR's slightly sweet and mild fruity-winey odour is attributed to the presence of volatile components such as methyl palmitate (42.2%), methyl stearate (5.2%), 2-methoxyethyl acetate (4.8%), acetyl carbinol (4.5%) and butanoic acid (4.1%) (Miyazawa and Kameoka, 1988). There are trace amounts of minor constituents, including 5-methylhydrantoin, tuberosin, choline chloride, acetylcholine chloride, D-mannitol, glycerol 1-monotetracosanoate (Lin et al., 2005), eicosanoic acid, hexadecanoic acid, tetracosanoid acid-2, 3-dihydroxypropyl ester (Wang et al., 2007), diacetonamine and D-(+)-pinitol (Nakamoto et al., 1975).



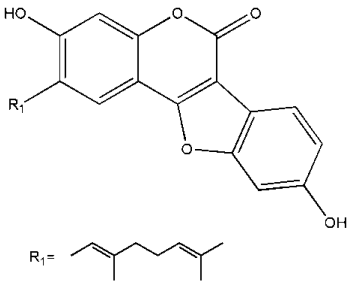
	R_1	R_2	R_3	R_4	R_5
Puerarin	H	H	Glu	H	H
Daidzein	H	H	H	H	H
Daidzin	H	Glu	H	H	H
Genistein	OH	H	H	H	H
Genistin	OH	Glu	H	H	H
Formononetin	H	H	H	H	Me
Biochanin A	OH	H	H	H	Me
Puerarin-4'-O-D-glucoside	H	H	Glu	H	Glu
3'-Hydroxypuerarin	H	H	Glu	OH	H
3'-Hydroxypuerarin-4'-O-deoxyhexoside	H	H	Glu	OH	E
3'-Hydroxy-4'-O-β-D-glucosylpuerarin	H	H	Glu	OH	Glu
3'-Methoxypuerarin	H	H	Glu	OMe	H
4'-Methoxypuerarin	H	H	Glu	H	OMe
6"-O-D-xylosylpuerarin	H	H	Glu^6-xyl	H	H
6"-O-Malonyl ester of puerarin	H	H	A	H	H
Daidzein-7-O-methyl ether	H	Me	H	H	H
3'-Methoxydaidzin	H	Glu	H	OMe	H
3'-Methoxydaidzein-7-O-methyl ether	H	Me	H	OMe	H
4'-Glucosyldaidzin	H	Glu	H	H	Glu
6"-O-Malonyl ester of daidzin	H	A	H	H	H
8-[α-D-Glucopyranosyl-(1→6)-β-D-glucopyranosyl] daidzein	H	H	B	H	H
6"-O-Malonyl ester of genistin	OH	A	H	H	H
Formononetin 8-C-[β-D-apiofuranosyl-(1→6)]-β-D-glucopyranoside	H	H	C	H	Me
Formononetin 8-C-[β-D-xylopyranosyl-(1→6)]-β-D-glucopyranoside	H	H	D	H	Me

Figure 1.5 Chemical structures of isoflavonoids and isoflavonoid glycosides reported from PLR and PTR (Harada and Ueno, 1975; Hirakura et al., 1997; Kinjo et

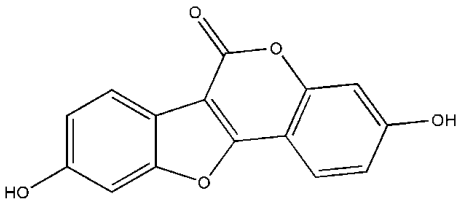
al., 1987; Murakami et al., 1960; Nguyen et al., 2009; Prasain et al., 2007; Rong et al., 1998; Shibata et al., 1959; Sun et al., 2008). Glu: glucose; H: hydrogen; Me: methyl; OH: hydroxyl; OMe: methoxy; A: 6”-O-manlonyl-Glu; B: α-D-glucopyranosyl-(1→6)-β-D-glucopyranosyl; C: β-D-apiofuranosyl-(1→6)]-β-D-glucopyranoside; D: β-D-xylopyranosyl-(1→6)]-β-D-glucopyranoside; E: Deoxyhexosyl



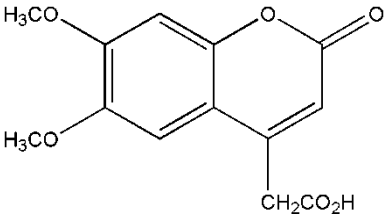
	R_1	R_2	R_3
Pueroside A	Glu^6-Rha	H	H
Pueroside B	Glu	Me	Glu
(±)-Puerol B 2-O-glucopyranoside	Glu	Me	H
Kuzubutenolide A	Glu	H	H



Puerarol

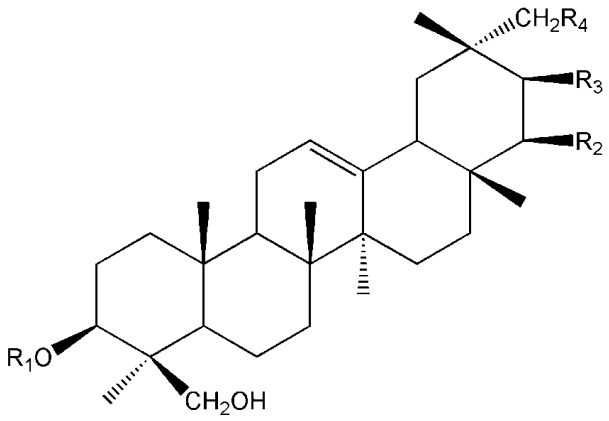


Coumestrol



6,7-Dimethoxycoumarin

Figure 1.6 Chemical structures of minor phenols reported from PLR and PTR (Hirakura et al., 1997; Kim et al., 2008; Lin et al., 2005; Nohara et al., 1993). Glu: glucose; H: hydrogen; Me: methyl; Rha: rhamnose.



	R_1	R_2	R_3	R_4
Kudzusapogenol A	H	OH	OH	H
Kudzusapogenol C	H	H	OH	H
Kudzusaponin A_1	Glu A^2-ara^2-rha	o-xyl	OH	OH
Kudzusaponin A_2	Glu A^2-gal	OH	OH	OH
Kudzusaponin A_3	Glu A^2-gal^2-rha	OH	OH	OH
Kudzusaponin A_4	Glu A^2-gal	OH	OH	OH
Kudzusaponin A_5	Glu A^2-glu	OH	OH	OH
Kudzusaponin SA_1	Glu A^2-gal	OH	OH	H
Kudzusaponin SA_2	Glu A^2-gal	o-ara	OH	H
Kudzusaponin SA_3	Glu A^2- gal^2-rha	o-ara	OH	H
Kudzusaponin SA_4	Glu A^2-glu A	o-ara	OH	H
Kudzusaponin SB_1	Glu A^2- gal^2-rha	o-ara	H	H
Kudzusaponin C_1	Glu A^2- gal^2-rha	H	OH	H
Soyasapogenol A	H	OH	OH	H
Soyasapogenol B	H	OH	H	H
Soyasaponinl A_3	Glu A^2- gal^2-rha	OH	OH	H
Soyasaponinl I	Glu A^2- gal^2-rha	OH	H	H

Figure 1.7 Chemical structures of triterpenoids and triterpenoid glycosides reported from PLR and PTR (Arao et al., 1995; Arao et al., 1997; Kinjo et al., 1987). Ara: arabinose; Gal: galactose; Glu A: glucuronic acid; H: hydrogen; OH: hydroxyl; Rha: rhamnose.

1.1.5.5 Impact of seasonal and geographical variations on major constituents

Traditionally, PLR is harvested during winter and autumn, as the ancient Chinese believed that the contents of active components were greater compared with other seasons. Chen et al. (2007) demonstrated that the total content of seven major isoflavonoids (3'-hydroxypuerarin, puerarin, 3'-methoxypuerarin, daidzin, genistin, formononetin-7-glucoside and daidzein) was lowest during April to August in the northern hemisphere (flowering period), with the content gradually increasing from October to December during fruiting period, and finally reaching a maximum peak in January of the following year when the plant is in the dormancy period (Chen et al., 2007). In addition, it is suggested that the best time to harvest PLR is around three years of growth, as the total amount of isoflavonoids is the highest (Chen et al., 2007). PLR from the north and north-eastern regions of China often has a relatively higher total isoflavonoid content as compared to those from the southern parts of China (Bodner and Hymowitz, 2002).

1.2 Quality control of herbal products

Herbal medicines and their preparations have been used around the world for centuries to prevent and cure diseases. According to a recent study, up to 80% of the population in Asia, Africa and Latin-America uses herbal medicines to meet their health care needs (World Health Organisation, 2002). Another study revealed that approximately two-thirds of Australians were using complementary medicines, which include traditional herbal medicines, vitamin and minerals to maintain general health (NICM, 2014; NPS, 2008). However, the lack of regulations and

legislations has prompted the global health care authorities to improve guidelines on the quality, safety and efficacy of herbal products (Uzuner et al., 2012).

The quality control of herbal materials is regarded as one of the most important procedures in the manufacturing process of herbal products as it could potentially impact on the safety and efficacy of the final products (Jiang et al., 2010). For example, the substitution of Stephaniae Tetrandrae Radix with Aristolochiae Fangchi Radix in a weight loss herbal product marketed at Belgium during 1990 to 1992 caused over 100 cases of irreversible nephropathy, due to the presence of the renal toxin aristolochic acid from Aristolochiae Fangchi Radix (Lewis, 2001; Zhu and Woerdenbag, 1995).

As herbal materials are acquired from natural resources, their quality and the amount of active constituents can be affected by various factors such as climate, soil condition, altitude, harvest time, drying process and storage condition (Tistaert et al., 2011a). To assess the quality of herbal products, the gold standard is to quantify the amount of one or several chemical compounds using analytical instruments such as high-performance liquid chromatography (HPLC) or thin-layer chromatography (TLC) (Alaerts et al., 2010). In most cases, the single compound approach has been found to be insufficient in differentiating herbs with similar appearance and/or similar major chemical compounds. For instance, chlorogenic acid was unable to differentiate Lonicerae Flos from Chrysanthemi Indici Flos and oleanolic acid was unable to differentiate Ligustri Lucidi Fructus, Achyranthes Bidentatae Radix and Clematidis Radix (Liang et al., 2009). Therefore, it is necessary to develop robust methods that uncover the complex nature of herbs.

To address this, the chemical 'fingerprint', which is defined as the characteristic profile or pattern that defines the complex chemical composition of a herb, has been adopted (Goodarzi et al., 2013). In 1991, the World Health Organisation (WHO) published the first edition of the *'Guidelines for the Assessment of Herbal Medicines'*, emphasising the use of the chemical fingerprint as a tool for assessing the quality, safety and efficacy of herbal products (Alaerts et al., 2010; Tistaert et al., 2011a). Since then, the Chinese State Food and Drug Administration (State Food and Drug Administration, 2000), the Food and Drug Administration of the U.S.A. (Food and Drug Administration of the United States of America, 2004) and the European Medicines Agency (European Medicines Agency, 2006, 2008, 2011) have adopted fingerprint analysis for evaluating the quality of herbal medicines and their preparations.

1.2.1 Analytical techniques for quality control

There are many analytical instruments that can be used to obtain the chemical fingerprint of a herb and these are divided into two major categories: chromatography and spectroscopy. A chromatographic fingerprint can be obtained from TLC, high-performance thin-layer chromatography (HPTLC), HPLC, ultra-performance liquid chromatography (UPLC), gas chromatography and capillary electrophoresis. The spectroscopic fingerprint can be obtained from mid-infrared (MIR) spectroscopy, near-infrared (NIR) spectroscopy, Raman spectroscopy, mass spectroscopy (MS) and nuclear magnetic resonance (NMR) spectroscopy (Alaerts et al., 2010; Gad et al., 2013; Tistaert et al., 2011a).

1.2.1.1 Liquid chromatography

1.2.1.1.1 High- and ultra-performance liquid chromatography

Chromatography is a differential migration process where sample components are distributed between a stationary and a mobile phase (Lough and Wainer, 1995). The type of chromatography is based on the physical state of the stationary and mobile phases. The most common method used in separating chemical constituents in herbal medicine is the liquid-solid column chromatography that features a liquid mobile phase, which slowly filters down through the solid stationary phase (Snyder et al., 2011). HPLC is a type of liquid chromatography in which the liquid mobile phase is mechanically pumped through a column that contains the solid stationary phase. A typical HPLC comprises of a mobile phase reservoir, a high-pressure pump, an injection port for introducing the sample, a stainless steel column containing packing material and a detector, and is normally connected to a computer for data acquisition and interpretation (Lough and Wainer, 1995). The most common type of HPLC used for herbal analysis is the reversed-phase HPLC, which consists of a column packed with uniform particles of silica whose surface has a covalently-bonded coating of long hydrocarbon chain such as octyldecyl silica. Chemical constituents from the sample mixture are separated depending on their different interaction with the adsorbent stationary phase. In reversed-phase HPLC, polar molecules are eluted at the beginning of the analysis, whereas less polar molecules are eluted later and hence, have a longer retention time. The efficacy of separating chemical components within a sample depends on several factors including flow rate, particle size, column diameter, solvent polarity and the distribution constant between the liquid and solid phase. In general, the identification of small molecular weight chemicals such as

phenolic compounds is based on ultraviolet (UV)-visible diode array detector (DAD) with multiple wavelengths between 200 and 800 nm (Lough and Wainer, 1995; Yang et al., 2010). HPLC can be hyphenated with other spectral instruments such as MS and NMR for the identification of individual chemical constituents. However, the major pitfalls of using HPLC are the relatively long analytical time and large organic solvent consumption (Sherma, 2012).

In comparison, UPLC is a collective term describing the liquid chromatography technology that employs a column packed with smaller particles and/or higher flow rates for increased speed, with superior resolution and sensitivity (Naushad and Khan, 2014; Xu, 2013). As compared to conventional HPLC (column diameter: 2.1 – 4.6 mm; particle size: 3 – 5 μm), the use of UPLC with a column of particle size of 1.7 μm and diameter of 2.1 mm significantly decreases the solvent consumption, analytical run time and injection volume, and enhances resolution and sensitivity (Grumbach et al., 2009; Naushad and Khan, 2014; Snyder et al., 2011; Xu, 2013). Although UPLC could significantly reduce the above problems, this technique is accompanied with high instrumental and maintenance cost and requires highly specialised skills.

1.2.1.1.2 Thin-layer chromatography

TLC is one of the oldest liquid chromatography techniques and has been widely used for the chemical analysis of herbal medicines and is one of the fundamental techniques for the authentication of herbs described in the various pharmacopoeias. TLC is a form of planar chromatographic technique whereby the stationary phase is spread on a flat and planar surface (Wall, 2005). Silica gel is

the most commonly used stationary phase in TLC analysis, and is supported by a sheet of aluminum or a glass plate, while the mobile phase (consisting of a mixture of solvents in a defined ratio) migrates up the stationary phase. Since the different chemical constituents in the sample mixture have different interactions with the mobile and stationary phase, separation takes place via capillary action (Srivastava, 2010). The separation of chemical constituents within a herbal mixture is governed by several factors including the partition coefficients, retention factor (Rf), capacity factor of an individual constituent on the plate, selectivity of the mobile and stationary phase to the solute, as well as the plate height that the solvent travels (Spangenberg et al., 2011; Srivastava, 2010). As most of the organic compounds are colourless, they will need to be made visible on the plate. The most common method is to spray the plate with a chemical reagent that reacts with the particular chemicals and detect under UV light (Fried and Sherma, 1996, 1999).

The modern HPTLC method is generally referred to as the application of using a small particle-sized HPTLC plate, automated or semi-automated sample injector and high-resolution densitometry (Srivastava, 2010). An automated sample injector applies the sample in a narrow band with equal volume, which enhances the separation between chemicals and minimises the injection variability. As compared to TLC, the relatively smaller particle size (TLC: 10 – 12 µm; HPTLC: 5 – 6 µm) and thinner plate thickness (TLC: 250 µm; HPTLC: 100 µm) in HPTLC has been found to significantly reduce the development time and solvent consumption (Fried and Sherma, 1996; Srivastava, 2010; Wall, 2005). As compared to HPLC,

the advantages of using HPTLC are its simplicity, versatility and the analysis of multiple samples simultaneously.

1.2.1.2 Vibrational spectroscopy

Vibrational spectroscopy, comprising of near-infrared (NIR), mid-infrared (MIR) and Raman spectroscopy, measures the dynamic vibration transitions in molecules (Chalmers et al., 2012). The fundamental units used in describing a spectrum include wavenumber (cm^{-1}) and wavelength (nm). The choice of unit depends on the type of spectrometer used, *e.g.* dispersive or Fourier transform. The conversion between wavenumber and wavelength can be expressed as follows (Larkin, 2011):

$$v = \frac{1}{\lambda} = \frac{f}{c}$$

where v denotes as the wavenumber, λ represents the wavelength, f represents the frequency and c represents the velocity of light.

1.2.1.2.1 Mid- or near-infrared spectroscopy

Infrared spectroscopy is based on the absorption of electromagnetic radiation. The infrared spectrum can be divided into three major spectral regions, including NIR (4000 – 12500 cm^{-1}), MIR (400 – 4000 cm^{-1}) and far-infrared (10 – 400 cm^{-1}) (Sun, 2009).

The MIR spectrum measures the absorbance of light at the vibrational and rotational frequencies of the atoms within a molecule, whilst NIR spectrum records the overtone and combination of fundamental vibrations (Larkin, 2011). The

molecule is considered as NIR active if there is a change in dipole moment of the molecule. A dipole refers to the product of charges, either positive or negative, and distance between atoms within a molecule. In general, R-H groups such as O-H, N-H, C-H and S-H bonds have the strongest overtone, as they are high in dipole moment (Smith, 2011; Workman and Weyer, 2007). However, one of the major disadvantages of using NIR is that the peaks in the spectrum overlap and are relatively broad as compared to MIR, which makes it difficult to interpret the molecular structure (Ozaki et al., 2006; Smith, 2011; Sun, 2009; Workman and Weyer, 2007).

MIR spectrum is often used for structure identification of organic compound because the absorption peak is caused by the fundamental vibration of a specific function group (Chalmers et al., 2012; Larkin, 2011; McCreery, 2005; Ozaki et al., 2006; Sun, 2009). However, one of the primary pitfalls of employing MIR is the high sensitivity of the water molecule, which restricts the type of samples to be analysed and requires sample preparation prior to analysis (*e.g.* dehydration) (Sasic and Ekins, 2008; Schrader, 2008; Smith, 2011; Workman and Weyer, 2007).

1.2.1.2.2 Raman spectroscopy

Raman spectroscopy is based on the scattering phenomenon of electromagnetic radiation as a result of energy exchange during molecular vibrations. The light scattering effect can be broadly divided into two categories: elastic and in-elastic scattering (McCreery, 2005). Figure 1.8 illustrates the differences between Raman, MIR and NIR spectroscopy during energy transition. With irradiation of the

monochromatic light onto a sample, photons from the light source collide with the sample molecules at an initial energy of E_o and then scatter in all dimensions. If the scattered light carries the same energy as in the incident light ($E = E_o$), this phenomenon is called elastic scattering, and is sometimes known as Rayleigh scattering. On the contrary, if there is an energy change (E_v) between the scattered and incident light, this process is called in-elastic scattering, and is also known as Raman scattering (Larkin, 2011). It is important to note that Raman scattering is a rare phenomenon. Most of the light will be scattered without any changes in the energy, whilst only 10^{-8} of the light will be scattered in-elastically. In Raman scattering, the scattered light can have either a lower energy (anti-stroke Raman scattering; $E = E_o - E_v$) or a higher energy (stroke Raman scattering; $E = E_o + E_v$) level as compared to incident light (Chalmers et al., 2012; Larkin, 2011). Typically, intense laser beams from a wide spectral range can be used as the excitation source in Raman spectroscopy. Argon ion laser (488.0 nm and 514.5 nm), krypton ion laser (530.9 and 647.1 nm) and neodymium-doped yttrium aluminum garnet laser (1064 nm) are the most common (McCreery, 2005).

A Raman spectrum displays the relationship between the intensity of scattered light and the Raman shift. The Raman shift refers to the frequency difference between the incident and scattered light and is usually expressed as wavenumber (Larkin, 2011). The criterion used to determine whether a molecule is Raman active depends on its polarisability during molecular vibration. The polarisability of a molecule is defined as the degree of distortion of the election cloud in a molecule affected by an electromagnetic field. Molecules with a symmetric shape are one of the basic requirements to obtain an intense peak in a Raman spectrum. As a rule

of thumb, molecules with functional groups such as -C-halogen, -C-S-, -C=C-, -C=S-, -N=N- and -S-H- tend to provide a more intense peak as they exhibit greater polarised changes as compared to asymmetric molecules such as water (Chalmers et al., 2012; Larkin, 2011; McCreery, 2005).

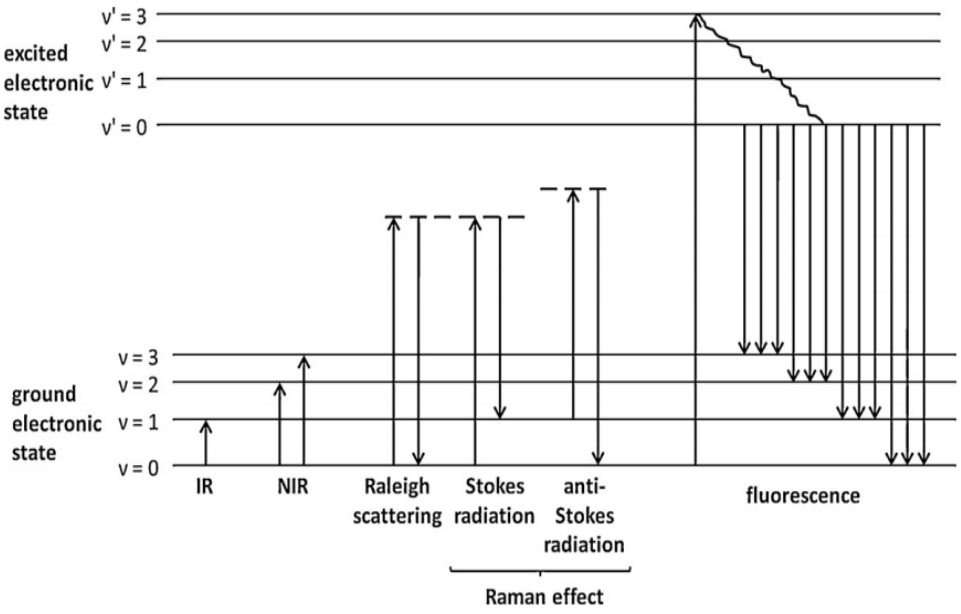


Figure 1.8 Comparison of the energy transition of Raman, MIR and NIR spectroscopy (De Beer et al., 2011).

One of the major advantages of adopting Raman spectroscopy in the quality control of herbal medicine relates to the high penetration power. The high-energy laser allows Raman spectroscopy to analyse samples through glass and plastic, which can be extremely useful for on-line monitoring of production procedures (Sasic and Ekins, 2008). Thus, minimal sample preparation is required. In addition, the water molecule is poorly sensitive to Raman spectroscopy due to its asymmetric shape and this enables the analysis of fresh, water-rich and *in-situ* samples (McCreery, 2005; Schrader, 2008). However, as Raman scattering is a weak signal, it can be affected significantly by other environment factors. One of

the most common interferences in analysing herbal materials is the auto-fluorescence of molecules such as DNA and protein (Chalmers et al., 2012; Larkin, 2011; Sasic and Ekins, 2008). Fluorescence is emitted when a molecule decays from its excited electronic state back to the ground electronic state. This phenomenon can be minimised by employing a laser with a longer wavelength. (Sasic and Ekins, 2008; Schrader, 2008; Sun, 2009). A summary comparing the characteristics of Raman, MIR and NIR are illustrated in Table 1.3.

Table 1.3 Summary of the Raman, MIR and NIR spectroscopy (Chalmers et al., 2012; Larkin, 2011).

	Raman	MIR	NIR
Ease of sample preparation			
Liquid	Very simple	Very difficult	Fair
Powder	Very simple	Simple	Simple
Gas	Very simple	Simple	Simple
Fingerprint	Excellent	Excellent	Very good
Best vibration	Symmetric	Asymmetric	Overtone
Quantitative	Good	Good	Excellent

1.2.2 Colorimetric assay

Colorimetry measures the absorbance intensity of UV-visible light when it passes through the test sample, and is calculated according to the Beer-Lambert's Law as follows (Ozaki et al., 2006; Sun, 2009):

$$A = \varepsilon lc = log\left(\frac{I_o}{I}\right)$$

where A represents absorbance, ε represents molar absorptivity, l represents pathlength, c represents concentration, I_o represents intensity in the background spectrum and I represents intensity in the sample spectrum.

In a colorimetric assay, a reference standard is normally used to generate a calibration curve (Liu, 2011). The results of the tested sample is expressed as an equivalent dose to the reference standards and is calculated as follows:

$$equivalent\ dose = \frac{C_{cal}}{C_{extract}} \times dilution\ factor \times extraction\ yield\ (\%)$$

where C_{cal} represents the concentration of chemical calculated in the testing sample using the calibration curve and $C_{extract}$ represents the original concentration of the testing sample.

This approach provides general chemical information related to the type of chemical families within the herb (Liu, 2011; World Health Organisation, 2011). Common colorimetric assays used in herbal medicines include the determination of total flavonoid content, total phenolic content, antioxidant capacity and *in-vitro* cell viability.

1.2.2.1 Total flavonoid content assay

The total flavonoid content (TFC) assay measures the total quantity of flavonoids, including flavones, flavonols, flavanones and isoflavones, present in the sample. This assay uses aluminium chloride in a potassium acetate solution. Aluminium chloride reacts with the keto group at the C-4 position and either the hydroxyl groups at C-3 or C-5 position of flavonoids to form a stable complex (Chang et al., 2002). This stable complex is yellowish-orange in colour and has a maximum absorbance at 415 nm. Quercetin is commonly used as the reference compound and the TFC of the sample is expressed as quercetin equivalent (QE) dose (Meda et al., 2005).

1.2.2.2 Total phenolic content assay

The total phenolic content (TPC) colorimetric assay is a popular analytical method used to quantify the total amount of phenolic compounds in herbal medicines and food products. This assay uses the Folin-Ciocalteu reagent, which is a mixture of phosphotungstic acid ($H_2PW_{12}O_{40}$) and phosphomolybdic acid ($H_3PMo_{12}O_{40}$). In an alkaline condition, the phenolic compounds oxidise the Folin-Ciocalteu reagent and generate a mixture of blue oxides of tungsten (W_8O_{23}) and molybdenum (Mo_8O_{23}) (Ainsworth and Gillespie, 2007). The blue-coloured oxide mixture has a maximum absorbance at 760 nm. The intensity of the absorbance is proportional to the total quantity of phenolic compounds in the sample. Phenolic compounds such as gallic acid are commonly used to generate calibration curve. However, the formation of blue oxides can be affected by air oxidation after the sample is in alkaline condition, and therefore the Folin-Ciocalteu reagent needs to be added before the alkaline reagent (Ainsworth and Gillespie, 2007; Prior et al., 2005).

1.2.2.3 2,2-Diphenyl-1-picrylhydrazyl (DPPH) radical scavenging capacity assay

The 2,2-diphenyl-1-picrylhydrazyl (DPPH) assay measures the scavenging ability of chemical compounds for the stable DPPH• free radical (Brand-Williams et al., 1995). The DPPH• free radical (due to the delocalisation of the spare electron throughout the molecule) is purple in colour and has a maximal absorbance at 520 nm. When DPPH reacts with a hydrogen donor, the free radical reduces to a colourless stable form. As a result, the free radical scavenging capacity of the sample is estimated by measuring the decolorisation of the mixture at 520 nm. Trolox, a water-soluble form of vitamin E, is commonly used as a reference

compound and the scavenging capacity of the sample is expressed as Trolox equivalent antioxidant capacity (TEAC) (Brand-Williams et al., 1995; Molyneux, 2004).

1.2.2.4 2,2'-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) radical scavenging capacity assay

The 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) measures the ability of chemical compounds in scavenging the ABTS radical cation (Re et al., 1999; Thaipong et al., 2006). The production of the ABTS radical cation involves the oxidation of ABTS with potassium persulfate or manganese dioxide. The cation is bluish-green in colour and has a maximal absorbance at 736 nm. Similar to the DPPH assay where $ABTS^{+}\bullet$ reacts with a hydrogen donor, the cation is reduced to a colourless stable form. As a result, the radical scavenging capacity of the sample is estimated by measuring the decolourisation of the mixture at 736 nm. Trolox is commonly used as a standard compound and the scavenging capacity of the sample is expressed as TEAC (Re et al., 1999).

1.2.2.5 Cupric reducing antioxidant capacity (CUPRAC) assay

The cupric reducing antioxidant capacity (CUPRAC) assay measures the reducing capacity of the chemicals in a redox reaction. In this assay, the ability of the sample in reducing cupric-neocuprine complex to cuprous-neocuprine complex in an alkaline condition is assessed (Apak et al., 2004). Cuprous-neocuprine complex is yellowish-orange in colour and has a maximal absorbance at 450 nm. Thus, the absorbance intensity of the mixture at 450 nm is proportional to the amount of complex produced, which is used to estimate the cupric ion reducing

capacity of the sample. Gallic acid is commonly used as the reference compound and the CUPRAC in the sample is expressed as gallic acid equivalent (GAE) dose (Apak et al., 2007; Apak et al., 2004).

1.2.2.6 Cell viability assay

The cell viability assay uses the reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) for measuring the viability of cells in a cellular model. The MTT enters the cells and passes into the mitochondria where it is reduced to an insoluble coloured formazan complex by mitochondrial succinate dehydrogenase (Mosmann, 1983). This complex is dark purple in colour and has a maximum absorbance at 550 nm. As a result, the amount of insoluble formazan formed reflects the metabolic status of the mitochondria, which can be used as an indicator for cell viability (Berridge and Tan, 1993). The insoluble formazan is normally solubilised with organic solvents such as isopropanol and dimethyl sulfoxide. The cell viability is calculated as follows:

$$cell\ viability\ (\%) = \frac{absorbance\ of\ formazan\ in\ treatment\ group}{absorbance\ of\ formazan\ in\ non-treatment\ group} \times 100\%$$

1.2.3 Chemometrics

Although it is possible to visually differentiate the chemical fingerprint differences between an authenticated and an unknown sample, the process is rather subjective and may vary between different investigators and laboratories. Furthermore, it is time-consuming especially if a large quantity of samples needs to be analysed (Gad et al., 2013; Tistaert et al., 2011a). To solve this problem, statistical analysis in the form of chemometrics is introduced. Chemometrics is defined as (Massart et al., 1997; Wold and Sjostrom, 1977):

> *"The science of relating measurements made on a chemical system or process to the state of the system via application of mathematical or statistical methods."*

Chemometrics covers a wide range of techniques including the optimisation of experimental parameters, design of experiments, signal processing, pattern recognition, calibration and regression modelling and estimation of structure-property relationship. In the case of the chemical fingerprint analysis, pattern recognition and calibration and regression modelling are the two most common techniques employed. As compared to traditional analytical method, which consists of a few independent variables, the chemical fingerprint analysis normally involves a large number of independent and/or dependent variables. Therefore, instead of univariate analysis, multivariate analysis is usually involved in interpreting the relationship between the dependent and independent variables in the chemical fingerprint analysis (Massart et al., 1997; Otto, 2007; Varmuza and Filzmoser, 2009; Wold and Sjostrom, 1977).

1.2.3.1 Data pre-processing

Prior to any multivariate analysis, pre-processing techniques are applied to remove undesired variability from the raw data due to analytical instrument/operator differences, and to improve the quality and interpretability of the data by removing redundant variables (Brereton, 2003; Gemperline, 2006).

For chromatographic fingerprint analysis, the most important procedure is to ensure that all the peaks in the chromatograms from different samples are aligned at their respective retention time. In chromatographic analysis, individual peaks can shift between injections due to variations in mobile phase composition, column ageing, instrumental instability and operator handling (Brereton, 2003; Gemperline, 2006; Joshi, 2012). Many approaches have been suggested to solve the peak alignment issue, including correlation optimised warping (COW), target peak alignment, fuzzy warping, dynamic time warping, parametric time warping, semi-parametric time warping and peak alignment by a genetic algorithm (Alaerts et al., 2010; Gad et al., 2013; Tistaert et al., 2011a). Among them, COW is the most common method and has been proven to be effective in correcting the retention time shift. COW is a piecewise data pre-processing algorithm which involves the alignment of a sample chromatogram in the form of a vector with a reference chromatogram (Skov et al., 2006; Tomasi et al., 2004). The process is achieved by the combination of stretching or compressing of sample segments using linear interpolation. One of the major advantages of using COW in the chemical fingerprint alignment is its ability to preserve the peak area and shape after alignment. However, it is important to note that three parameters including the target chromatogram, slack number and the number of segments need to be

determined and optimised before the application of COW for peak alignment (Skov et al., 2006; Tomasi et al., 2004).

For spectroscopic fingerprint analysis, however, the undesired variability of the raw data are mainly due to light scattering and differences in the effective path length, especially when performing NIR spectroscopic analysis (Mark and Workman Jr., 2007). Light scattering is a complex phenomenon and could be linear or non-linear. There are three major types of scattering effects. The first one is a simple baseline shift due to the additive effect. The second type of scattering effect called multiplicative effect, scales the entire spectrum by a given factor. The last type contributes to the wavelength-dependent baseline shift, in which the degree of baseline shift varies with wavelength (Mark and Workman Jr., 2007; Rinnan et al., 2009). To minimise the variability due to the light scattering effect, several commonly used pre-processing algorithms namely Savitzky-Golay polynomial derivatives, multiplicative scatter correction, extended multiplicative signal correction, standard normal variate and de-trending can be employed (Joshi, 2012; Mark and Workman Jr., 2007).

Apart from the instrumental specific pre-processing techniques, there are several algorithms that are commonly applied to the raw data to remove variability unrelated to the property of interest (Joshi, 2012; Mark and Workman Jr., 2007). Normalisation is a row-wise algorithm which makes overall intensity scales comparable across samples, by dividing each variable with the sum of the absolute value of all the variables for a given sample (Afseth et al., 2006). In contrast, standard normal variate (SNV) is a weighted normalisation method which

divides the variables with the standard deviation of all variables for a given sample. Apart from the row-wise methods, the matrix can be pre-processed by column-wise algorithms such as mean centering, median centering class centering, auto-scaling, poisson scaling, pareto scaling and variance scaling. In most cases, one of these techniques will be the last algorithm to be applied to the data matrix and is an essential step in many pattern recognition algorithms such as principal component analysis (PCA) and partial least square (PLS) analysis. The most popular centering method is mean centering, which involves the subtraction of variables with the mean value of each column (Tistaert et al., 2011a).

In general, data pre-processing is needed prior to the application of multivariate analysis. When analysing a chemical fingerprint, it is advisable that one should try different pre-processing techniques and their combinations to determine which approach generates the best result by comparing with the subsequent model performance (Massart et al., 1997; McLachlan, 2004; Otto, 2007; Varmuza and Filzmoser, 2009).

1.2.3.2 Variable selection

Variable selection selects a subset of variables that is most representative of the relationship between the independent and response variable matrices, particularly if the number of observations (samples) is much smaller than the number of independent variables (X matrix) or if the data matrix contains much redundant data (Balabin and Smirnov, 2011).

1.2.3.2.1 Successive projection algorithm

Successive projection algorithm (SPA) is a deterministic algorithm which does not involve random factors. The aim of SPA is to minimise collinearity and the propagation of measurement noise in the calibration set. Initially, a maximum number of variables are pre-defined. The variables are projected onto an orthogonal sub-space. The vector with the highest projection value is selected and employed as the new starting vector. The aforementioned steps are repeated until the iteration reaches the maximum number of variables. For each iteration, the choice of orthogonal sub-space is made in such a way that only non-collinear variables are selected. The variable subsets are used to establish multivariate linear regression models and their root mean square of leave-one-out cross-validation (RMSECV) are calculated. The variable subset with the smallest RMSECV is selected (Araujo et al., 2001).

1.2.3.2.2 Genetic algorithm

Genetic algorithm (GA) is a probabilistic search algorithm inspired by Darwin's evolution theory, which incorporates biological concepts such as inheritance, natural selection, chromosome crossover and mutation. First, an initial population is randomly generated based on the uniform distribution of the data matrix and coded as a number of chromosomes, which refers to a legal solution to the predefined problem and is composed of a string of genes. Next, the chromosomes are divided into pairs of subsets and crossover to generate daughter chromosomes, which will replace the parent chromosome in the new generation. Mutation is then carried out with a user-defined mutation rate. Once the new population is generated, each chromosome is subjected to multivariate regression

analysis and its fittest value is calculated. The algorithm is repeated until the termination condition is fulfilled (Broadhurst et al., 1997).

1.2.3.3 Unsupervised pattern recognition

Pattern recognition analysis can be broadly divided into two major categories, namely unsupervised and supervised pattern recognition analysis (McLachlan, 2004). Figure 1.9 illustrates a schematic diagram of the composition of unsupervised and supervised pattern recognition analysis. In an unsupervised pattern recognition analysis, the chemical fingerprints are rearranged into a two-dimensional data matrix X (m x n). In the X matrix, m indicates the number of fingerprints (observations) included, whereas n represents the signal intensity at a certain time point (chromatogram) or wavenumber (spectrum). The number of data points depends on the resolution of the analytical instrument and the number of data points acquired per unit of time or wavenumber (Brereton, 2003; McLachlan, 2004). On the contrary, supervised pattern recognition analysis utilises the information from both X and Y matrices. The Y matrix refers to the response(s) such as the total amount of chemical component and pharmacological activity of samples (McLachlan, 2004; Varmuza and Filzmoser, 2009). The supervised pattern recognition can be broadly divided into three major categories, including exploratory data analysis, similarity analysis and clustering analysis.

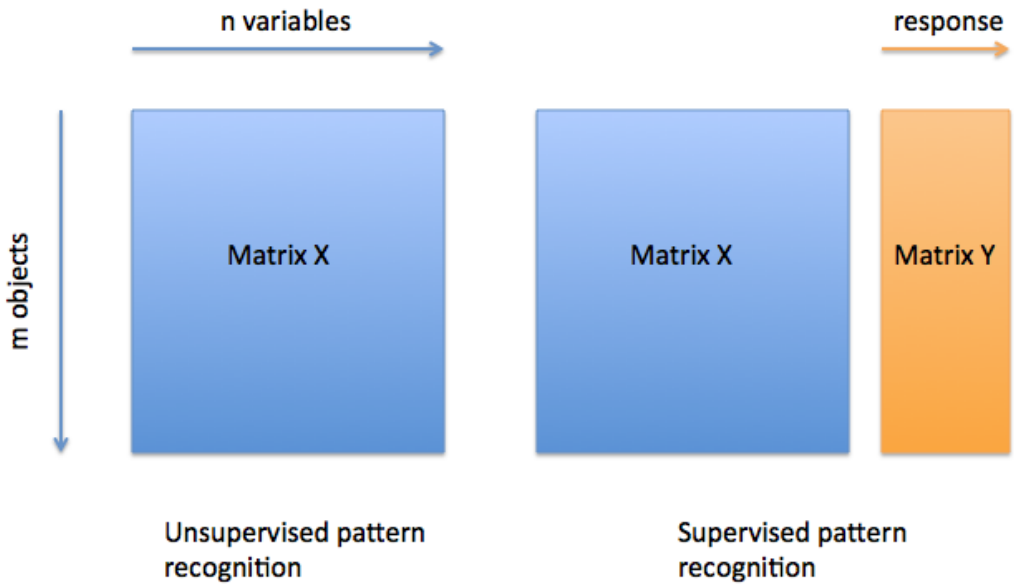


Figure 1.9 Schematic diagram of the composition of unsupervised and supervised pattern recognition analysis.

1.2.3.3.1 Exploratory data analysis

There are several exploratory data analyses available, such as PCA, robust PCA, projection pursuit and factor analysis (Massart et al., 1997). Among them, PCA is the most common unsupervised pattern recognition method and is used as a preliminary step for separating a pool of samples into different clusters (species) and identifying outliers within the cluster. PCA is a variable reduction algorithm that allows the visualisation of the data in a low dimensional space (Otto, 2007). This is achieved by projecting the variables in the X matrix from the original data space onto a new dimensional space called principal component (PC). Each PC is a linear combination of the original variables and each successive PC is orthogonal to the previous PC. As a result, the X matrix is decomposed into a number of PCs that describe the maximum amount of variance between variables in the data. The first PC accounts for the maximum covariance between variables.

The second PC is unrelated to the first PC and accounts for the next maximum covariance. This procedure continues until it reaches the pre-defined maximum number of PC, and the remaining unexplained data are expressed as residual (E) (Brereton, 2003).

The decomposition of the X matrix by PCA includes two matrices, scores and loadings matrix. The scores are the new coordinates of the samples in the PC. In the case of a herb, each point represents a single chemical fingerprint in the scores plot, and the distribution of the points on the scores plot can be used to reveal patterns, trends, clusters and outliers of the data. In other words, the points that are located close to each other have similar chemical characteristics, whereas the points that are located far from each other have dissimilar or different chemical characteristics (Brereton, 2003; Otto, 2007; Varmuza and Filzmoser, 2009). Loadings describe the correlation coefficients between the original variable and the PCs. In a loading plot, the loading value indicates the importance of a variable in a model during decomposition and can be used to determine which variable is important in the separation clusters or groups. The PCA model can be described in the formula as follows (Brereton, 2003; Otto, 2007; Varmuza and Filzmoser, 2009):

$$X = TP^T + E$$

where X is the chemical fingerprints matrix (m x n), T is the matrix of scores (T x A), with A representing the number of calculated PCs, P is the loading matrix, P^T is the transposed of loading matrix (A x n) and E is the residual matrix.

1.2.3.3.2 Similarity analysis

Similarity analysis is one of the conventional methods in unsupervised pattern recognition and has been used to investigate the relationship between the chemical fingerprints of herbs (Goodarzi et al., 2013). The Chinese Pharmacopoeia Committee has developed a specialised software named 'Similarity evaluation system for chromatographic fingerprints of TCM' to evaluate the similarity between fingerprints based on the Pearson's correlation coefficient (r) as follows (Ganzera, 2009):

$$r(x_1, y_2) = \frac{cov\ (x_1, x_2)}{s_{x1} s_{x2}}$$

$$= \frac{\sum_{i=1}^{n}(x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{\sqrt{\sum_{i=1}^{n}(x_{1i} - \bar{x}_{1i})^2 \sum_{i=1}^{n}(x_{2i} - \bar{x}_2)^2}}$$

where (i = 1, 2, 3, … n), x_{1i} and x_{2i} are the ith elements of x_1 (fingerprint 1) and x_2 (fingerprint 2), n is the number of variables in the fingerprints, $cov\ (x_1, x_2)$ is the covariance of the vectors x_1 and x_2 and s_{x1} and s_{x2} are the standard deviation of x_1 and x_2.

1.2.3.3.3 Clustering analysis

There are two types of clustering analysis: non-hierarchical and hierarchical analysis (Massart et al., 1997). One of the examples in non-hierarchical clustering analysis is the fuzzy clustering analysis, which pre-defines a number of set points, and all samples are then assigned with a point value. The clustering is based on the distance between the samples and the set points or inclusion within a prescribed distance to a specific seed point (Chen and Wang, 1999). In hierarchical analysis, the samples are grouped into clusters and expressed as a tree structure called a dendrogram. The clustering creation can be divided into

agglomerative and non-agglomerative approaches (Johnson, 1967). In an agglomerative approach, each sample is considered as an individual object and is grouped together based on their similarity. In a non-agglomerative approach, all the samples are first grouped into a large cluster, followed by division into smaller clusters based on their dissimilarity (Romesburg, 2004). Two parameters, linkage and similarity function, need to be determined prior to clustering analysis. The examples of a linkage function include single linkage, complete linkage, Ward's method and centroid method. The similarity between samples can be calculated using common algorithms such as Euclidean, Mahalanobis, Minkowski and Chebychev distance (Johnson, 1967; Romesburg, 2004). Among them Euclidean distance is the most common algorithm employed and the formula is expressed as follows:

$$Euclidean\ distance = \sqrt{\sum_{i=1}^{n} (x_i - v_i)^2}$$

where i = 1, 2, 3, … n, $(x_1, x_2, \ldots x_n)$ and $(v_1, v_2, \ldots v_n)$ are the two sets of n measurements from chemical fingerprints (Massart et al., 1997).

1.2.3.4 Supervised pattern recognition

In general, supervised pattern recognition refers to a group multivariate analyses that utilise a calibration or training set with prior knowledge about the classes or properties of interests in the X matrix to construct a calibration model. The model's performance is evaluated by comparing the predictability and/or accuracy of the true class of the validation set before applying the mode to unknown samples (Tistaert et al., 2011a). The calibration model can be used as a classification (discriminant) or regression model. In a classification mode, the Y matrix

represents the class memberships of the samples in X matrix. The Y matrix is qualitative and discrete and normally coded in vectors (Otto, 2007; Varmuza and Filzmoser, 2009). For example, class 1 coded as 0, class 2 coded as 1, … class n coded as n-1. On the contrary, the Y matrix in a regression model indicates the properties of the samples. However, it is important to note that most of the supervised pattern recognition algorithms, except soft independent modelling of class analogies, can be applied for the purpose of both classification and regression (Gad et al., 2013; Joshi, 2012).

In addition, the supervised pattern recognition algorithms can be categorised as parametric or non-parametric and linear or non-linear algorithm. Linear algorithms include linear discriminant analysis, whilst support vector machines and artificial neural network are considered as non-linear algorithms (Brereton, 2003; Gemperline, 2006; Mark and Workman Jr., 2007; Otto, 2007). In parametric algorithms such as linear discriminant analysis and partial least squares analysis, statistical parameters of the normal distribution of the sample from X matrix are used in the decision rules and *vice versa* for the non-parametric algorithms. Examples of non-parametric algorithms include K-nearest neighbors and classification and regression tree (Brereton, 2003; Gemperline, 2006; Mark and Workman Jr., 2007; Otto, 2007).

1.2.3.4.1 Linear discriminant analysis

Linear discriminant analysis (LDA), also known as Fisher's linear discriminant analysis, is a variable reduction algorithm similar to PCA (Welling, 2005). In the LDA model, the data are projected to a lower-dimensional vector space (latent

variable) such that the ratio of the between-class variance is maximised while the ratio of the within-class variance is minimised. The latent variable obtained in LDA is a linear combination of the original variables from the data set and hence, the discriminant boundaries are also linear (Lachenbruch, 1975). Unlike the PC in PCA that aims to maximise the number of explained variables in a lower dimensional space, the latent variable in LDA aims to maximise the separation among the given classes (Figure 1.10) (Lachenbruch, 1975; Massart et al., 1997).

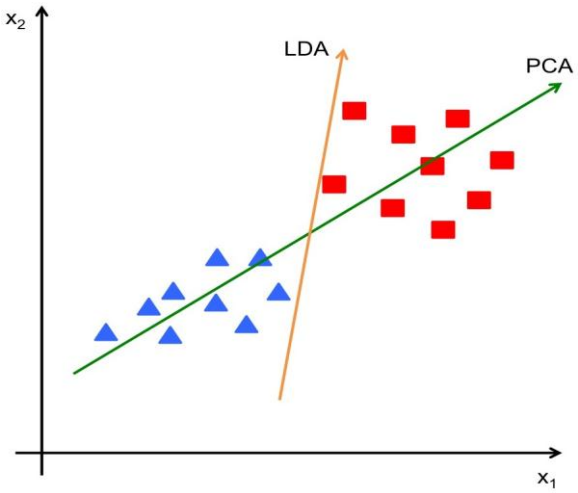


Figure 1.10 Difference in projection dimension between PCA and LDA.

The first step in determining the optimal latent variable is to maximise the separation between classes. The data matrix is decomposed into two scatter matrices, namely between-class (S_w) and within-class matrix (S_b) (Varmuza and Filzmoser, 2009). The mean of each sample class ($\bar{x}_i$) is defined as:

$$S_b = \sum_{i=1}^{g} N_i(\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

where N_i is the number of data point in a sample class, g is the number of sample class, $\bar{x}_i$ is the class mean, $\bar{x}$ is the mean of the whole data set and T indicates transposition (Varmuza and Filzmoser, 2009).

$$S_w = \sum_{i=1}^{g} (N_i - 1) \sum i$$

where $\sum i$ is the covariance matrix

Since the covariance matrix relates to $(N_i - 1)$ and it has to be greater than zero, the maximum number of data points that a LDA model can handle must be smaller than $(N_i - g)$. In other words, the number of samples in the data set must be greater than the number of variables from each sample (Otto, 2007). To address this criterion, the aforementioned variable selection methods can be applied to the data set. This will allow the selection of variables that are informative in discriminating the classes (Alaerts et al., 2010; Gad et al., 2013).

1.2.3.4.2 Partial least squares

Another approach to solving the limitation of number of variables raised by LDA is to use partial least squares-discriminant analysis (PLS-DA). PLS, also known as projection to the latent structures, is a supervised pattern recognition method which reduces the dimensionality of the data matrix similar to PCA and projects variables onto a newly formed dimensional space called latent variable (LV) (Gemperline, 2006; Mark and Workman Jr., 2007). The purpose of PLS is to find the least number of LVs in the X matrix that describe the maximum variation between variables and, simultaneously, have maximal correlation to the variables in the response Y matrix. One of the major differences that distinguishes PLS from PCA is that PLS utilises both X and Y matrices and hence, maximises the covariance between them. As in PCA, the PCs explain the variance between variables within the X matrix, and have little or no relevance to the response Y

matrix. PLS is often used in analysing chemical fingerprint data as it can deal with multicollinearity amongst the X matrix and in a condition where the number of samples/observations is smaller than the number of X variables (Geladi and Kowalski, 1986; Rosipal and Krämer, 2006). PLS model is constructed by arranging the X and Y matrices as follows:

$$X = TP^T + E$$

$$Y = UQ^T + F$$

where X (n x m) is the chemical fingerprint matrix, with m = number of variable in each spectra and n = number of samples; T and P are the score and loading of the X matrix, respectively; Y (n x 1) is the response matrix; U and Q are the score and loading of the Y matrix, respectively; E and F are the residue matrices of the X and Y matrix, respectively. The score value is defined as the projection of the data on a weighted vector W, while the loading value refers to the contribution to the score value to the matrix.

It can be observed that the decomposition of the X and Y matrices are inter-related. The loading values of X matrix are calculated from the score values of Y matrix, whereas the loading values of matrix Y are determined from the score values of X matrix. As a result, the PLS forms a bidiagonal matrix during decomposition, whilst a diagonal matrix is generated from PCA. PLS algorithm can be used as both classification and regression (Massart et al., 1997; Varmuza and Filzmoser, 2009). In PLS regression, the coded vector Y matrix is replaced by experimental data. The regression coefficients of PLSR are calculated as follows:

$$T = XW^*$$

$$W^* = W(P^TW)^{-1}$$

$b = W^* Q^T = W(P^T W)^{-1} Q^T$

where W^* is the (m x k) matrix of X-weights that represents the common LV space (n x k) relating the X and Y matrices; k is the number of LV; b is the regression coefficient.

Finally, the predicted response value (y^*) is determined by the k number of LVs as follows:

$y^* = XW_k b$

Unlike unsupervised pattern recognition methods, PLS-DA maximises the separation between pre-defined classes, rather than explaining the variations within a data set. In comparison to PCA-discriminant analysis, PLS-DA aims to maximise the inter-class variance rather than capturing the variations within each class. PLS-DA calculates the probability of each sample belonging to each class, and then the y predicted value from each sample is fitted to the Gaussian distributed model from the validation set. A probability density function is established for each class based on the mean and standard deviation of their y predicted value. The classification threshold is defined as the y predicted value at which the probability of both classes is equal. The class of unknown sample is assigned by its y predicted value and the classification threshold of each model.

1.2.3.4.3 K-nearest neighbors

K-nearest neighbors (KNN) is a non-parametric classification method that does not form any mathematical assumption on the calibration set distribution. This approach utilises the distance between samples in the p-space as its criterion

(Coomans and Massart, 1982a, b, c). Euclidean distance is the most common metric algorithm used to measure the nearness between samples. An unknown sample is categorised by its proximity to the K samples that have been placed in pre-defined categories in the calibration set (Zadeh, 1965). Figure 1.11 illustrates the concept of the classification of samples in KNN. In this diagram, each point represents a sample, and the sign (positive or negative) represents the class. Firstly, a point of interest (e.g. x_q) is chosen from the X matrix and a boundary is formed depending on the pre-defined K value. In this case, the K value is defined as 5. Therefore, five other samples with the closest proximity are selected. The class of the point of interest is based on the majority class of the nearest neighbors.

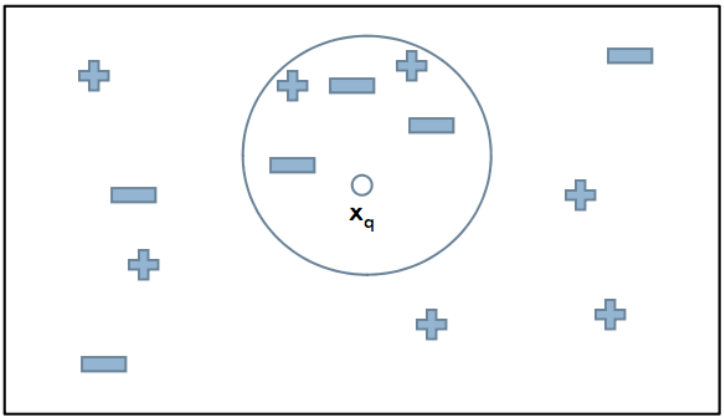


Figure 1.11 Schematic diagram of the concept of sample classification in KNN.

The determination of K value is one of the most important procedures required before analysis. If the K value is too small, the model will be unstable as there is not enough information to classify new samples. On the other hand, if the K value is too large, the model will contain bias, as some of the neighbors used to make predictions will no longer be similar to the one being predicted. To solve this

problem, leave-one-out cross-validation (LOOCV) can be used to select the optimal number of the nearest neighbors (K value) (Beckonert et al., 2003; Ni et al., 2009; Roggo et al., 2007).

1.2.3.4.4 Classification and regression tree

Classification and regression tree (CART) is a non-parametric classification algorithm that involves splitting the data into mutually exclusive sub-classes to form a decision tree in a binary recursive manner (Breiman et al., 1984). The construction starts from the root, which represents all the samples in X matrix. The root splits into two nodes called parent node, and each parent node further splits into two nodes called child node. This binary division procedure is repeated until all subjects in the terminal node are as pure or homogenous as possible. The node without a child node is called terminal node. The CART procedure consists of three major steps. Initially, the aforementioned binary partition process forms a complete decision tree, which is also known as the maximal tree. Secondly, an overgrown tree is pruned so that a series of less complex trees are formed. This pruning step is essential to minimise over-fitting. Lastly, the optimal tree size is determined by the best pruned sub-tree, which is chosen based on the LOOCV (Breiman et al., 1984). The splitting criterion is calculated by the Gini index as follows (Milanovic, 1997; Zang et al., 2011):

$$Gini = \sum_{j=1}^{k} p_{lj}\left(1 - p_{lj}\right) = \sum_{j=1}^{k} \frac{n_{lj}}{n_l}\left(1 - \frac{n_{lj}}{n_l}\right)$$

where k is the number of class, n_l is the number of objects in node l and n_{lj} is the number of objects from class j present in the node l.

If the dependent variables (response Y matrix) are categorical, a classification tree is generated, whilst a continuous variable will produce a regression tree. In general, the splitted sub-classes are more homogeneous with respect to the original data matrix. This approach estimates the discriminating surface directly without losing the degree of freedom of each class (Breiman et al., 1984).

1.2.3.4.5 Soft independent modelling of class analogy

Soft independent modelling of class analogy (SIMCA) minimises assumptions on the linearity of relationships between samples and pre-defined classes. One of the major differences that separates SIMCA from other classification techniques is the term 'soft', which means that an unknown sample can be classified as multiple classes rather than restricted to a discrete (non-overlapped) class (Wold and Sjostrom, 1977). To establish a SIMCA classification model, a PCA model is built for each class separately. Unlike PCA-DA which describes the overall variance in the data, the SIMCA model comprises of a collection of PCA models which define the characteristics of each class (Frank and Lanteri, 1989). The optimal number of PCs is determined by LOOCV. To classify an unknown sample, its matrix is projected to each established PCA model and the residual distance is calculated. The residue PC is used to create boundaries around each class. The residual standard deviation (s_k^K) of an object k to a class K in the calibration set can be expressed as (De Maesschalck et al., 1999; Wold and Sjostrom, 1977):

$$s_k^K = \sqrt{\frac{\sum_{j=1}^{m}(e_{kj}^k)^2}{m - p_k}}$$

where e_{ij}^i is the residual between object k and the PC along variable j and $(m - p_k)$ is the degree of freedom

The residual standard deviation (s_0^K) of an unknown object n_K from the validation set is calculated as:

$$s_0^K = \sqrt{\frac{\sum_{j=1}^{m}\sum_{i=1}^{nK}(e_{ij}^k)^2}{(n_K - p_K - 1)(m - p_K)}}$$

where $(n_K - p_K - 1)$ and $(m - p_K)$ are the degrees of freedom

An unknown sample is categorised by comparing its residual distance to the mean residual variance of the PCA model from each class.

1.2.3.4.6 Support vector machine-discriminant analysis

Support vector machine-discriminant analysis (SVM-DA) represents samples as points in a space by generating a hyperplane boundary between the two classes. The hyperplane maximises the distance between classes and is unrelated to the probabilistic distribution of the samples in the calibration set. The idea of a SVM model is to determine the optimal hyperplane that has a maximum margin between the two separated classes. It can be applied to both linear and non-linear data sets and to separate multiple classes. However, this approach projects points to a higher dimensional space by a suitable kernel function. The class of an unknown sample is assigned by mapping it as a point onto the same dimensional space of the calibration set, and then determining the side it belongs using the v-support vector classification (Vapnik, 1999).

1.2.3.5 Validation of a model

One of the crucial procedures in pattern recognition analysis is the validation of the established calibration model. It is essential to ensure that the calibration model is

able to accurately predict the class or property of interest of an unknown sample. In order to investigate the performance of a model, the X matrix is first divided into a calibration and a validation set (Gemperline, 2006; Massart et al., 1997; Otto, 2007). The calibration set builds a calibration model, whilst the validation set tests the accuracy and robustness of the established model. The validation set is completely independent of the model establishment procedures such as variable selection, parameters optimisation and the determination of optimal PCs or LVs. Both calibration and validation sets comprise of samples from each representative class (Brereton, 2003; Varmuza and Filzmoser, 2009).

As a rule of thumb, approximately 60% of the total amount of samples is classified as the calibration set and the remaining 40% is used as the validation set. To avoid bias in subset division, several types of sample selection algorithms such as Kennard-Stone (K-S), Duplex algorithm and random selection can be employed (Esbensen and Geladi, 2010; Kennard and Stone, 1969; Snee, 1977). Among them, K-S algorithm is the most popular method in pattern recognition analysis. K-S selects a subset of samples from a large population, which provides uniform and diverse coverage over the data matrix. According to the K-S algorithm, it assumes that if two samples share similar characteristics, then the distance between them is short, whilst two samples that are far apart have low similarity. Euclidean distance is the most common geometric algorithm used in the determination of distance of the samples in K-S algorithm. The algorithm selects an object that is closest to the mean of the data matrix and adds it to a subset. Then, the dissimilarity between the remaining objects within the original data matrix and the object(s) in the subset is calculated. Subsequently, the object with the highest

dissimilarity from the original data matrix is selected and added into the subset. These procedures are repeated until the number of objects in the subset reaches a pre-defined limit (Esbensen and Geladi, 2010; Kennard and Stone, 1969).

1.2.3.5.1 Root mean square errors and regression coefficient

To validate the predictability and stability of the calibration model, LOOCV is adopted. LOOCV is a special case of k-fold cross validation in which the k value is equal to the number of objects. Figure 1.12 illustrates the concept of LOOCV. In LOOCV, the model performance is determined by removing one sample from the calibration set at a time and using it as a validation set (Esbensen and Geladi, 2010; Kohavi, 1995).

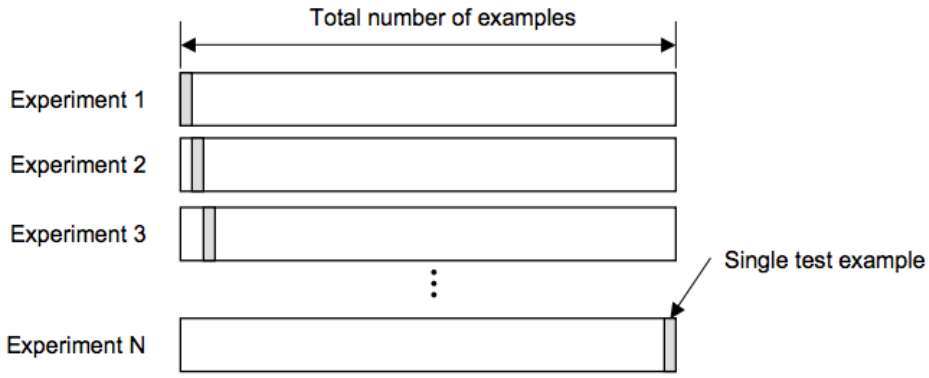


Figure 1.12 Schematic diagram of the leave-one-out cross-validation.

To access the model performance, the root mean square error of calibration (RMSEC), root mean square error of leave-one-out cross-validation (RMSECV) and root mean square error of validation (RMSEV) can be calculated as follows (Chen et al., 2008b; Wu et al., 2010):

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{Ic}(\hat{y}_{ic} - y_{ic})^2}{I_c}}$$

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{Ic}(\hat{y}_{icv} - y_{ic})^2}{I_c}}$$

$$RMSEV = \sqrt{\frac{\sum_{i=1}^{Ip}(\hat{y}_{ip} - y_{ip})^2}{I_p}}$$

where $\hat{y}_{ic}$, $\hat{y}_{icv}$ and $\hat{y}_{ip}$ are the calculated value of sample i in the calibration, cross validation and validation set, respectively, y_{ic} and y_{ip} are the measured value of sample i in the calibration and validation set, respectively, and I_c and I_p are the number of samples in the calibration and validation set, respectively.

In a regression model, the regression coefficient of the calibration (r^2 cal), cross-validation (r^2 CV) and validation set (r^2 val) are calculated as follows (Massart et al., 1997; Wold and Sjostrom, 1977):

$$r = \sqrt{1 - \frac{\sum_{n=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{n=1}^{n}(y_i - \bar{y})^2}}$$

where $\hat{y}_i$ and y_i are the calculated and measured value of sample i, $\bar{y}$ is the mean of the reference measurement results and n is the number of samples.

1.2.3.5.2 Confusion matrix

A confusion matrix displays the performance of a classification model in a contingency table (Todeschini et al., 2008). Figure 1.13 illustrates a classic binary confusion matrix. Each column represents the predicted class (identity) of the sample, whereas each row represents the true class of the sample.

		Predicted class	
		Class 1	Class 2
True class	Class 1	**True positive**	*False positive*
	Class 2	*False negative*	**True negative**

Bold: main diagonal; Italic: off diagonal

Figure 1.13 Classic binary confusion matrix.

The matrix on the main diagonal refers to the number of correct class assignations, while off diagonal refers to the classification error. True positive (TP) is the number of class 1 correctly classified as class 1. True negative (TN) is the number of class 2 correctly classified as class 2. On the contrary, false negative (FN) is the number of class 1 misidentified as class 2, while false positive (FP) is the number of class 2 misidentified as class 1. There are four major parameters commonly used to assess the classification model's performance, including sensitivity, specificity, non-error rate (NER) and error rate. Sensitivity refers to the model's ability to correctly identify samples belonging to that particular class. Specificity represents the model's ability to reject sample of all other classes (Todeschini et al., 2008). NER measures the quality of the classification model, while the error rate evaluates the efficiency of the model. These parameters are calculated as follows:

$$sensitivity = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{FP + TN}$$

$$NER = \frac{sensitivity + specificity}{number\ of\ class}$$

$$Error\ rate\ = (1 - NER)$$

1.2.3.5.3 Ratio of performance to deviation

Ratio of performance to deviation (RPD) is normally used to assess the predictability of a model. RPD measures the accuracy of the calibration model by dividing the standard deviation of the experimental values by the standard error of the estimated values obtained from the validation set. The higher the RPD value, the greater the predictability of a model (Williams and Sobering, 1993).

1.3 Current issues in the quality control of PLR and PTR

Quality control is an essential procedure to ensure the quality, safety and efficacy of a herb and its products. Table 1.4 illustrates the extraction method, mobile phase, stationary phase and other HPLC chromatography parameters for the detection of chemical marker(s) from the PPRC (PPRC, 2010a, b), EP (EP, 2012a, b), BP (BP, 2014a, b), JP (JP, 2011) and HKCMMS monograph (HKCMMS, 2010a, b).

Table 1.4 Summary of the extraction method and chromatographic parameters for the authentication of PLR and PTR in various pharmacopoeias and monographs (BP, 2014a, b; EP, 2012a, b; HKCMMS, 2010a, b; JP, 2011; PPRC, 2010a, b).

	PPRC	EP/BP	HKCMMS	JP
Extraction method	Reflux for 30 mins	Reflux for 30 mins	Sonicate for 30 mins	Reflux for 30 mins
Solvent	30% ethanol	30% ethanol	70% ethanol	50% methanol
Mobile phase	A: water; B: methanol	A: 0.1% acetic acid in water; B: acetonitrile	A: 0.1% formic acid in water; B: acetonitrile	A: 0.05 M sodium dihydrogen phosphate in water; B: acetonitrile
	N/A	0 – 16.5 min: A 90 – 71%	0 – 40.0 min: A 90 – 65%	N/A
Detection wavelength (nm)	250	260	250	250
Column	Octadecylsilanised silica gel	Octadecylsilanised silica gel	Octadecylsilanised silica gel	Octadecylsilanised silica gel
Dimensions	N/A	5 µm particle size I.D. 4.6 mm; L 15 cm	5 µm particle size I.D. 4.6 mm; L 25 cm	5 µm particle size I.D. 4.6 mm; L 15 cm
Flow rate (mL/min)	N/A	3	1	N/A
Chemical marker(s)	Puerarin	Puerarin and total isoflavonoids	Puerarin	Puerarin
Content (w/w)	PLR ≥ 2.4% PTR ≥ 0.3%	PLR ≥ 6.5% total isoflavonoids with ≥ 45% puerarin PTR ≥ 0.4% total isoflavonoids with ≥ 55% puerarin	PLR ≥ 2.6% PTR ≥ 0.16%	PLR ≥ 2.0%

N/A: not available.

Puerarin is used as the chemical marker in PPRC, JP and HKCMMS for the authentication of PLR and PTR due to its high abundance among other chemical constituents (HKCMMS, 2010a, b; JP, 2011; PPRC, 2010a, b). It was observed that the extraction method, extraction solvent and chromatographic parameters varied among monographs. These factors could significantly alter the amount of the extracted chemical. As a result, the amount of puerarin content for the authentication of PLR ranges from 2.0% to 2.9% (w/w), whereas for the authentication of PTR, the puerarin content ranges from 0.16 to 0.30% (w/w) (BP, 2014a, b; EP, 2012a, b; HKCMMS, 2010a, b; JP, 2011; PPRC, 2010a, b). In addition to puerarin, the amount of total isoflavonoids including puerarin, 3'-hydroxypuerarin, 3'-methoxypuerarin, 6"-O-D-xylosylpuerarin and daidzin is recommended by EP and BP (BP, 2014a, b; EP, 2012a, b). This approach might provide additional information on the chemical differences between PLR and PTR, which is currently missing in other monographs. However, the reason why these chemical markers were chosen is not disclosed. Furthermore, 3'-hydroxypuerarin, 3'-methoxypuerarin and 6"-O-D-xylosylpuerarin are not available as authentic reference standard and hence, identification of these compounds may need hyphenated instruments such as HPLC-mass spectrometry. The lack of standardisation between monographs might cause confusion in differentiating PLR from PTR. Thus, this could explain why these two herbs are used interchangeably in clinical practice.

1.4 Rationale and objectives

In clinical practice, PLR and PTR are substituted for each other even though it is recognised that these two herbs are different. This would have an impact on patient dosages and the efficacy of the treatment. The current monographs need to be updated and standardised to reflect the macroscopic, microscopic and chemical differences (both qualitative and quantitative) between PLR and PTR.

Although numerous studies have attempted to differentiate PLR from PTR, these studies mainly focused on the quantification of individual compounds such as puerarin, which does not completely represent the complex chemical profile of PLR and PTR. Conventional quantification processes using HPLC are time-consuming and are difficult to adopt as a daily, on-line procedure during the manufacturing process to monitor the quality of a herbal product. Thus, simpler methods using HPTLC and Raman spectroscopy need to be examined and compared.

Chemometrics has been shown to be a useful tool for the authentication of herbal medicines and differentiation of closely related species and is recommended by the WHO and various pharmaceutical organisations around the world. However, the application of chemometrics on the quality control of PLR and PTR is currently lacking.

Furthermore, it has been observed that most of the *in vitro* studies focus on the pharmacological activity of individual compounds from PLR and PTR. There are limited studies investigating the pharmacological activity of the extracts and

whether the chemical differences between the two species could impact on their respective pharmacological activity. Therefore, the aim of the study was to differentiate PLR from PTR using qualitative and quantitative methods and various analytical instruments in coupled with chemometrics analysis.

The specific objectives of this thesis were to:

1. Provide an up-to-date and comprehensive study on the macroscopic and microscopic characteristics of PLR and PTR.

2. Characterise the phytochemical profile of PLR and PTR by quantifying the major chemical constituents and comparing the total phenolic content and antioxidant capacity.

3. Develop a PLS-DA model for the differentiation of PLR and PTR using UPLC chromatographic fingerprints.

4. Develop multivariate classification models for the differentiation of PLR and PTR using HPTLC chromatographic fingerprints and comparing their model's performance and productivity with the models generated from the UPLC chromatographic fingerprints.

5. Develop PLS regression models for the prediction of TPC and antioxidant capacity using Raman spectroscopy.

6. Elucidate whether the chemical differences between PLR and PTR influence the *in vitro* pharmacological activity in models pertaining to the inhibition of carbohydrate hydrolysis enzymes α-amylase and α-glucosidase, the cytoprotective effect against stimuli on human endothelial EA.hy926 cells and the cytotoxicity on human prostate cancer PC3 cells.

Therefore, in this thesis, PLR and PTR were differentiated using qualitative and quantitative methods. Rapid and accuracy classification models were developed by combining different analytical instruments and variable selection algorithms with the aim of reducing the analytical time, while at the same time preserving the predictability and accuracy of the model. With the aid of chemometrics, in particular supervised classification model, and supported by bioactivity data, it is hypothesised that the two species can be rapidly differentiated in an objective and systematic manner.

Chapter Two

Morphological identification of Puerariae Lobatae Radix and Puerariae Thomsonii Radix

2.1 Introduction

Correct species identification is regarded as one of the fundamental procedures for ensuring the quality, safety and efficacy of a medicinal herb (Upton et al., 2011). The substitution and misidentification of medicinal herbs readily occurs in clinical practice, especially if they share similar macroscopic features or nomenclature, including analogous species names. This can lead to dire consequences especially if the substituted herb is toxic when used inappropriately (Nortier et al., 2000).

Morphological traits are the fundamental criteria for authenticating a medicinal herb, and whenever possible, should be carried out before any further analyses are undertaken (Liu, 2011). These techniques have been recommended as the first step for assessing the identity of a medicinal herb by the World Health Organisation (World Health Organisation, 2011), Pharmacopoeia of the People's Republic of China (PPRC) (PPRC, 2010a, b), European Pharmacopoeia (EP) (EP, 2012a, b), British Pharmacopoeia (BP) (BP, 2014a, b), Japanese Pharmacopoeia (JP) (JP, 2011) and various monographs around the world (Blumethal et al., 1999; HKCMMS, 2010a, b; Upton et al., 2011). The procedure involves the comparison of the morphological traits between an unknown sample and an authenticated reference. In general, the morphological traits of a medicinal herb can be divided into two major groups, macroscopic and microscopic characteristics (Liu, 2011; Upton et al., 2011). The macroscopic characteristics refer to the shape, size, colour, smell and texture of a medicinal herb. On the other hand, microscopic characteristics refer to the features of cell types and organelles when observed with the aid of a microscope (Ruzin, 1999).

In the case of PLR and PTR, the misinterpretation is related to the confusing taxonomical nomenclature and monograph descriptions in various pharmacopoeias. Many TCM practitioners and herbal dispensers think that these two species are the same and hence, are used interchangeably in clinical practice (Wong et al., 2011). Table 2.1 to 2.4 summarises the macroscopic and microscopic characteristics described in the PPRC, EP, BP, JP and Hong Kong Chinese Materia Medica Standards (HKCMMS) monograph (BP, 2014a, b; EP, 2012a, b; HKCMMS, 2010a, b; JP, 2011; PPRC, 2010a, b). PLR and PTR are separated into two monographs in the latest versions of the PPRC, EP, BP and HKCMMS, whereas in the JP only PLR is noted. In JP, “Puerariae Radix” refers to the root of *P. lobata* (Willd.) Ohwi while the root of *P. thomsonii* Benth. is not considered as a medicinal herb (JP, 2011). In general, the monographs in the BP are direct copies of the EP (BP, 2014a, b; EP, 2012a, b). Clearly, most of the monographs are based on the descriptions in the PPRC and hence, PPRC is considered as the primary focus in this study. Any changes in the PPRC could significantly impact on the descriptions of other monographs.

Since the inclusion of Puerariae Radix as a medicinal herb in the first edition of the PPRC (1965), the macroscopic and microscopic characteristics described in the content have not been changed considerably. Prior to the 2005 edition, Puerariae Radix referred to the roots of both *Pueraria lobata* (Willd.) Ohwi and *P. thomsonii* Benth. The macroscopic section is separated into two parts to describe the two species, whilst the microscopic features are the same (PPRC, 1997, 2000). The accumulating scientific evidence suggests that these two species have different

chemical and DNA profile. In 2005, the Commission of the PPRC separated the two species into two distinct monographs. Nonetheless, the contents of the monographs are almost identical. The only major difference is the content of the chemical marker puerarin. The content of puerarin in PLR should not be less than 2.4% (w/w), whereas in PTR the puerarin content should not be less than 0.3% (w/w) (PPRC, 2005).

Since PLR and PTR are obtained from two different species, their morphological and microscopic characteristics should not be assumed to be identical. However, without detailed descriptions and photographic illustrations in the PPRC, it becomes a difficult task to correctly differentiate PLR and PTR. In addition, morphological identification is rather subjective and the results may vary between different investigators and laboratories. Therefore, it is necessary to update and measure the morphological characteristics in an objective manner.

To provide additional information to support the results obtained from the morphological characteristics, two colorimetric methods, which determine the total content of starch and dietary fibre, are introduced. This approach would provide quantitative data supporting the morphological features observed under the microscope, and is relatively fast compared to the preparation of microscopic slides preparation. These two assays have been widely used in the field of food chemistry and have been applied to determine the starch and dietary fibre content in grains, crops and fruits (Englyst and Hudson, 1987; Rose et al., 1991). However, their application in herbal medicine is limited (Englyst and Hudson, 1987; Laurentin and Edwards, 2003; Rose et al., 1991).

Therefore, the aims of this chapter were to compare the macroscopic and microscopic characteristics of PLR and PTR and examine whether the quantitative colorimetric assays are useful as an authentication criterion. Comprehensive descriptions on the morphological traits of these two species using quantification techniques, in addition to individual observation, will promote the correct species to be used in clinical practice.

Table 2.1 Comparison of the macroscopic characteristics of PLR and PTR in different editions of PPRC (PPRC, 1997, 2000, 2005, 2010a, b).

	PPRC prior to 2005	PPRC after 2005	
	PLR and PTR	PLR	PTR
Shape	**PLR:** Small square pieces; Longitudinally rectangular thick slices, length 5 – 35 cm, thickness: 0.5 – 1 cm **PTR:** Cylindrical, subfusiform or semi-cylindrical, length 12 – 15 cm, diameter 4 – 8 cm; Longitudinally or obliquely cut thick slices	Small square pieces; Longitudinally rectangular thick slices, length 5 – 35 cm, thickness: 0.5 – 1 cm	Cylindrical, subfusiform or semi-cylindrical, length 12 – 15 cm, diameter 4 – 8 cm Longitudinally or obliquely cut thick slices, varying in size
Colour	**PLR:** Outer bark: pale brown with longitudinal wrinkles and rough; Inner: yellowish-white and indistinct striations **PTR:** Outer bark removed; Unpeeled: greyish-brown; External: pale brown	Outer bark: pale brown with longitudinal wrinkles and rough; Inner: yellowish-white and indistinct striations	Outer bark removed; Unpeeled: greyish-brown; External: yellowish-white or pale brown
Texture	**PLR:** Pliable and strongly fibrous **PTR:** Heavy, hard and starchy	Pliable and strongly fibrous	Heavy, hard and starchy
Smell	**PLR:** Slightly odour **PTR:** N/A	Slightly odour	Same as PLR
Taste	**PLR:** Slightly sweet **PTR:** N/A	Slightly sweet	Same as PLR
Other features	**PLR & PTR:** Collected in autumn and winter	Collected in autumn and winter	Same as PLR

Table 2.2 Comparison of the macroscopic characteristics of PLR and PTR in various pharmacopoeias (BP, 2014a, b; EP, 2012a, b; HKCMMS, 2010a, b; JP, 2011).

	EP/BP		HKCMMS		JP
	PLR	PTR	PLR	PTR	PLR
Shape	Small and square pieces; Thick and rectangular slices, length: 5 – 35 cm, thickness: 0.5 – 1 cm	Cylindrical, subfusiform or semi-cylindrical, length 12 – 15 cm, diameter 4 – 8 cm; Longitudinally or obliquely cut thick slices, varying in size	Small irregular cube; Thick and longitudinally cut slices, length, 0.4 – 36.2 cm, width: 0.4 – 10 cm, thickness: 0.2 – 5.4 cm	Longitudinally or obliquely; cylindrical, subfusiform or conical, length 0.4 – 39.3 cm, diameter, 27 – 84 cm	Small pieces of irregular hexagons: 0.5 cm cube longitudinally plate-like pieces, length: 20 – 30 cm, width: 5 – 10 cm, thickness: 1 cm
Colour	Outer bark: pale brown; Inner: yellowish-white and indistinct striations	Outer bark removed; External: yellowish-white or pale brown	External: brown, with longitudinal wrinkles; Inner: pale yellowish-brown, concentric annular rings	Outer bark: pale brown; External: yellowish-white	Greyish-yellow or greyish-white
Texture	Strongly fibrous	Heavy, hard and starchy	Tough, pliable and strongly fibrous	Heavy, hard and starchy	Extremely fibrous
Smell	N/A	N/A	Odourless	Same as PLR	N/A
Taste	N/A	N/A	Slightly sweet	Same as PLR	Slightly sweet followed by bitterness
Other features	N/A	N/A	Collected in autumn and winter	Same as PLR	N/A

Table 2.3 Comparison of the microscopic characteristics of PLR and PTR in different editions of PPRC (PPRC, 1997, 2000, 2005, 2010a, b).

	PPRC prior to 2005	PPRC after 2005	
	PLR and PTR	PLR	PTR
Xylem vessels	N/A	N/A	N/A
Fibre bundles	Thickened and lignified fibre, surrounded by a calcium prism sheath	Thickened and lignified fibre, surrounded by a calcium prism sheath	Same as PLR
Starch granules	Numerous, simple or 2 – 20 compound	Simple or 2 – 20 compound	Numerous
Shape	Spheroidal, semi-rounded or polygonal with a hilum pointed, cleft or stellate	Spheroidal, semi-rounded or polygonal with a hilum pointed, cleft or stellate	Same as PLR
Size (diameter)	3 – 37 μm	3 – 37 μm	8 – 15 μm
Others	Border pitted vessels relatively large Pits hexagonal or elliptical, arranged very densely	Border pitted vessels relatively large Pits hexagonal or elliptical, arranged very densely	Same as PLR
Presence of photos or diagrams	N/A	N/A	N/A

Table 2.4 Comparison of the macroscopic characteristics of PLR and PTR in various pharmacopoeias (BP, 2014a, b; EP, 2012a, b; HKCMMS, 2010a, b; JP, 2011).

	EP/BP		HKCMMS		JP
	PLR	PTR	PLR	PTR	PLR
Xylem vessels	N/A	N/A	Relatively large, numerous, densely and alternately arranged with fibre bundles	Same as PLR	Distinct vessels and xylem fibres
Fibre bundles	Thick-walled lignified fibre, surrounded by a calcium prism sheath	Same as PLR	Thicken and lignified wall and surrounded by a calcium prism sheath	Same as PLR	Accompanied by crystal cells in phloem
Starch granules	Numerous, simple or 2 – 20 compound	Same as PLR	Numerous, simple or 2 – 12 compound	Numerous, mostly 2 – 15 compound, rarely simple	Numerous, paranchyma, polygonal simple grains, or 2- to 3-compound grains
Shape	Spheroidal, semi-rounded or polygonal with a pointed, cleft or stellate hilum	Same as PLR	Spherical, hemi-spherical or ellipsoidal with dotted, cleft or asteroidal hilum	Spheroidal or sub-globose	Hilum or cleft in the center with straie
Size (diameter)	15 μm	Same as PLR	Single: 2 – 53 μm Compound: 8 – 87 μm	Single: 4 – 38 μm Compound: 8 – 78 μm	2 – 18 μm, mostly 8 – 12 μm
Others	N/A	N/A	Narrow cortex and cork board with several rows	Same as PLR	N/A
Presence of photos or diagrams	N/A	N/A	Both macroscopic and microscopic diagrams were present	Same as PLR	N/A

N/A: not available

2.2 Methods and materials

2.2.1 Solvents and chemical reagents

Analytical grade methanol, chloroform, ethyl acetate, absolute ethanol (99.5% w/w), glacial acetic acid and acetone were purchased from Thermo Fisher Scientific (VIC, Australia). Deionised water was purified by Siemens Ultra Clean Series water purification system (Siemens Water Technologies, NSW, Australia). All other chemicals and solvents were of analytical grade and were obtained from Sigma-Aldrich unless otherwise stated (NSW, Australia).

2.2.2 Herbal samples

A total of 46 samples were purchased from various herbal pharmacies located in Australia, China and the U.S.A., and were authenticated by Dr George Li from the Faculty of Pharmacy, The University of Sydney, Australia. The authentication criteria were based on the macroscopic (Table 2.1) and microscopic characteristics (Table 2.3) and further confirmed by quantifying the puerarin content using ultra-performance liquid chromatography (UPLC) in Chapter 3 (p.98). Twenty-three samples were authenticated as PLR, whereas 23 samples were authenticated as PTR (Table 2.5). Voucher specimens were kept at the Faculty of Pharmacy, The University of Sydney. Principal component analysis (PCA) coupled with UPLC chromatographic fingerprints obtained was used to detect outliers. Four outliers were detected based on the Hotelling T^2 versus Q residuals and scores plot from the PCA Chapter 4 (p.126) and were removed from the subsequent analysis.

Table 2.5 Sample code and place of location of PLR and PTR.

PLR		PTR	
Sample code	Purchase location	Sample code	Purchase location
PLR1	Sydney, Australia	PTR1	Sydney, Australia
PLR2	Nanning, China	PTR2	Sydney, Australia
PLR3	Henan, China	PTR3	Guangzhou, China
PLR4	Sydney, Australia	PTR4	Sydney, Australia
PLR5	Hong Kong, China	PTR5	Sydney, Australia
PLR6	Hong Kong, China	PTR6	Sydney, Australia
PLR7	Hong Kong, China	PTR7	Sydney, Australia
PLR8	Guangzhou, China	PTR8	Sydney, Australia
PLR9	Guangzhou, China	PTR9	Sichuan, China
PLR10	Auhui, China	PTR10	Hong Kong, China
PLR11	Auhui, China	PTR11	Hong Kong, China
PLR12	Auhui, China	PTR12	Hong Kong, China
PLR13	Auhui, China	PTR13	Macau, China
PLR14	Auhui, China	PTR14	Hong Kong, China
PLR15	Auhui, China	PTR15	Melbourne, Australia
PLR16	Auhui, China	PTR16	Shanghai, China
PLR17	Auhui, China	PTR17	Shanghai, China
PLR18	Auhui, China	PTR18	Guangzhou, China
PLR19	Hebei, China	PTR19	Guangzhou, China
PLR20	Tianjin, China	PTR20	Pennsylvania, U.S.A.
PLR21	Zhejiang, China	PTR21	Sydney, Australia
PLR22	Nanning, China	PTR22	Auhui, China
PLR23	Beijing, China	PTR23	Hunan, China

2.2.3 Microscopic characterisation

As the preparation of microscopic slides is a relatively time-consuming and labour intensive procedure, eight samples, including four PLR and four PTR, were randomly selected for microscopic characterisation. The histological preparations were performed at the Discipline of Pathology, University of Sydney, Australia. Each sample was fixed by immersing in a solution with formalin, glacial acetic acid and 70% ethanol (5:5:9 v/v/v) for two days. The fixed sample was dehydrated by passing through a series of ethanol and water mixture of increasing ethanol strength as follows: 70% ethanol (5 mins); 80% ethanol (5 mins); 90% ethanol (5 mins); 95% ethanol (5 mins); 100% ethanol (5 mins); 100% ethanol (10 mins), and subsequently immersed in 100% xylene for 15 mins. The dehydrated sample was embedded in paraffin wax in accordance to the procedures described previously (Ruzin, 1999). The embedded sample was sectioned by a microtome and was subsequently stained with safranin-fast green solution. An optical microscope (Olympus DP70) equipped with a digital camera was used to acquire the microscopic features of each sample. To validate the key authentication parameters, the values of various cells and tissues were obtained by taking at least 100 measurements from three different random views for each sample.

2.2.4 Total starch content assay

The total starch content assay measures the starch polymer present within the cells. The powdered sample (50 mg) was placed in an oven at 100 ^{o}C for 1 hour and was then mixed with 5 mL of solution contained methanol, chloroform and water (12:5:3 v/v/v). The mixture was placed in a sonication bath for 10 mins and was centrifuged at 4,500 rpm for 5 mins. The supernatant was carefully removed

with a pipette. To ensure removal of the chemicals which could interfere with the glucose measurement, the above extraction steps were repeated at least 3 times or until the supernatant became colourless. The residue was dried in an oven at 50 ^{o}C for 4 hours. The dried residue was macerated with 15 mL 35% perchloric acid and was shaken continuously for 1 hour to extract and solubilise the starch granules. The mixture was filtered through a Whatman No.1 filter paper (Rose et al., 1991). Concentrated sulfuric acid (5 mL) was added to hydrolysis the remaining starch. The amount of glucose presented in the filtrate was determined by the anthrone-sulfuric acid colorimetric assay. Briefly, an anthrone solution was made by dissolving 0.1146 g anthrone in 50 mL concentrated sulfuric acid and 20 mL deionised water on ice. The filtrate (100 µL) was added to a test tube, then mixed with 1 mL anthrone solution and boiled in a water bath for 10 mins. The mixture was allowed to cool down to room temperature for 15 mins. The mixture (200 µL) was transferred to a 96 well microplate and measured at 655 nm (Laurentin and Edwards, 2003) using a Bio-Rad Model 680 microplate reader (Bio-Rad Laboratories, NSW, Australia). The experiment was carried out in triplicates. Glucose (1 – 100 µg/mL) was used as the positive control and to generate a calibration curve. The results were calculated and expressed as milligrams of starch per gram of dried mass (DM) using:

$$y = \frac{y_g \times df \times v \times f}{dw}$$

where y is the mg of starch/ g of DM, y_g is the glucose concentration, df is the dilution factor, v is the original volume of starch extract (15 mL), f is the starch hydrolysis factor (0.9), which refers to the conversion factor that contributes to the hydrolysis from starch to glucose, and dw is the original dry weight of the samples (100 mg).

2.2.5 Total dietary fibre content assay

The total dietary fibre content colorimetric assay measures the non-starch polysaccharides within the plant sample. The powdered sample (100 mg) was placed in an oven at 100 ^{o}C for 1 hour and was then mixed with 2 mL dimethyl sulfoxide for 10 mins. The mixture was placed in a boiling water bath for 1 hour to disperse and hydrolyse starch granules. Without cooling, the mixture was combined with 8 mL of 0.1 M sodium acetate buffer (pH 5.2) and incubated at 40 ^{o}C for 5 mins. α-Amylase (5 mL; 5 U/mL) and pullulanase (0.1 mL; 2.5 U/mL) solutions were added and the mixture was incubated at 42 ^{o}C for 16 hours. Subsequently, 40 mL absolute ethanol was added to the mixture at room temperature and left for 1 hour and was centrifuged at 1500 g for 10 mins. The supernatant was removed and the residue was washed twice with 50 mL of 85% ethanol, and then 40 mL of acetone. The residue was placed in an oven at 65 ^{o}C for 1 hour. The dried residue was macerated with 2 mL of 12 M sulfuric acid and incubated at 35 ^{o}C for 1 hour, which was then mixed with 22 mL of water and placed in a boiling water bath for 2 hours. This mixture was considered as the hydrolysate of the dietary fibre. The total dietary fibre content is expressed as the sum of the amount of uronic acid and reducing sugar present in the mixture after hydroxylation (Englyst and Hudson, 1987).

To determine the amount of uronic acid, hydrolysate (0.3 mL) was mixed with 0.3 mL sodium chloride/boric acid solution, which was prepared by dissolving 2 g of sodium chloride and 3 g of boric acid in 100 mL deionised water. Then, the mixture was mixed with 5 mL concentrated sulfuric acid and incubated at 70 ^{o}C for 40 mins. Dimethylphenol solution was prepared by dissolving 0.1 g of dimethylphenol

in 100 mL glacial acetic acid. The mixture was allowed to cool at room temperature for 15 mins and was mixed with 0.2 mL dimethylphenol solution. The mixture was incubated at room temperature for 15 mins and 200 µL was transferred to a 96 well microplate and measured at 450 nm. Heparin was used as the positive control and to generate the calibration curve (Englyst and Hudson, 1987).

To determine the total reducing sugar, dinitrosalicylate solution was prepared by dissolving 1 g of 3,5-dinitrosaliylate, 1.6 g of sodium hydroxide and 30 g of sodium potassium tartrate in 100 mL of deionised water. The hydrolysate (1 mL) was mixed with 2 mL of dinitrosalicylate solution and boiled in a water bath for 10 mins. The mixture was allowed to cool down at room temperature for 15 mins and 200 µL of mixture was transferred to a 96 well microplate and read at 530 nm. Glucose was used as the positive control and to generate calibration curve. The total dietary fibre content was calculated as the sum of the uronic acid and total reducing sugar content as previously described (Englyst and Hudson, 1987). The experiment was carried out in triplicates, and the results were expressed as milligrams of starch per 100 g DM.

2.2.6 Data analysis

The statistical analyses were performed on GraphPad Prism version 6.01 (GraphPad Software, CA, U.S.A.). The data was analysed using either one way analysis of variance (ANOVA) or Student's t-test.

2.3 Results and discussion

2.3.1 Morphological characteristics

2.3.1.1 Macroscopic characteristics

Figure 2.1 and 2.2 illustrate the macroscopic characteristics of PLR and PTR, respectively. As shown in Figure 2.1, the PLR samples were previously cut into small irregular cube with variable sizes. Two out of 23 samples were cut longitudinally. The outer surface was dark brown in colour with longitudinal wrinkles. The inner surface was yellowish-brown with numerous small pores and several distinctive concentric annular rings. The texture was tough, and porous with a slightly sweet odour and taste.

In contrast, most of the PTR samples purchased were cut longitudinally into thin slices as per pharmacopoeias and monographs (Table 2.1 and 2.2), with only 6 out of 23 samples in a cubical shape (Figure 2.2). Most of the samples were peeled and there were 5 out of 23 samples with outer skin. The peeled outer surface is light brown, while the inner surface is yellowish white with several distinctive concentric annular rings. The texture is pliable and starchy with a slightly sweet odour and taste.

An additional feature that was observed in this study and is normally absent in the monographs is the presence of abundant xylem vessels (porous texture) in PLR. However, the macroscopic results alone could not differentiate between PLR and PTR using these macroscopic characteristics. In addition, the original shape and size of the PLR and PTR samples varied significantly. If authentication is based on these criteria, PLR10 and PLR23 would be classified as PTR as they are in slices,

whereas PTR9, PTR12, PTR13, PTR20, PTR22 and PTR23 would be classified as PLR as they are in cubical shape.

It is important to note that smell and taste is one of the crucial features in determining the nature of a herb according to TCM theory. Although, these characteristics were mentioned in the PPRC, JP and HKCMMS, they were absent in the EP and BP. Therefore, the EP and BP may need to be revised to incorporate these aspects.

Figure 2.1 Morphological characteristics of PLR (Photos taken by Ka H. Wong).

Figure 2.2 Morphological characteristics of PTR (Photos taken by Ka H. Wong).

2.3.1.2 Microscopic characteristics

In this study, the microscopic characteristics were significantly different between PLR and PTR, which were further supported by the results obtained from the colorimetric assays. Figure 2.3 shows the microscopic characteristics of fibre bundle and calcium oxalate crystals in PLR and PTR. It can be observed that fibre bundles were the dominant cell type in PLR. The fibre bundles were thick-walled, lignified and surrounded by parenchyma cells with calcium oxalate crystals. They were mainly located at the cortex, sometimes positioned around the vessels. Table

2.6 illustrates the microscopic characteristics, total starch content and total dietary fibre content of PLR and PTR. In general, PLR has dramatically larger fibre bundles (area: 0.0075 ± 0.0003 mm^2; $p<0.01$) associated with a greater number of fibres per each bundle (32.6800 ± 2.8780; $p<0.01$) as compared to PTR (area: 0.0025 ± 0.0002 mm^2; number of fibre per bundle: 16.5900 ± 0.9982).

Dietary fibre is defined as the polysaccharides and lignin that are not digested or absorbed in the human small intestine. There are two common methods to determine the total dietary fibre content, including the use of digestive enzymes coupled with high-performance liquid chromatography-mass spectrometry and chemical hydroxylation coupled with colorimetric assay (Van Soest et al., 1991). In this study, the colorimetric approach was chosen as it utilises a simple chemical procedure as compared to the sophisticated analytical instruments. The quantitative colorimetric assay demonstrated that the mean of the total dietary fibre content in the PLR powdered samples (4.2886 ± 0.3466 g/ 100 g DM) was significantly lower ($p<0.001$) than that in the PTR samples (12.4148 ± 0.4541 g/ 100 g DM). The lower total dietary fibre content obtained in the PLR samples was related to the large and densely distributed xylem vessels which occupied most of the space within the cortex and hence, reduced the population of other cell types. Additionally, a higher total dietary fibre content was observed in the PTR samples, which further supports its application as a food source in Chinese cuisine.

Table 2.6 Microscopic characteristics, total starch content and total dietary fibre content of PLR (n=4) and PTR (n=4).

	PLR	PTR
Xylem diameter (mm)	0.1390 ± 0.0184**	0.0471 ± 0.0109
Xylem area (mm^2)	0.2222 ± 0.0042**	0.0037 ± 0.0015
Number of fibres per bundle	32.6800 ± 2.8780**	16.5900 ± 0.9982
Fibre bundle area (mm^2)	0.0075 ± 0.0003**	0.0025 ± 0.0002
Starch diameter (μm)	5.2830 ± 0.5495	4.6790 ± 0.4455
Total starch content (mg starch/ g DM)	0.5288 ± 1.2559***	76.7954 ± 2.9905
Total dietary fibre content (g/ 100 g DM)	4.2886 ± 0.3466***	12.4148 ± 0.4541

mg starch/ g DM: milligrams of starch per grams of dried mass; g/ 100 g DM: grams of dietary fibre per 100 grams of dried mass p-value < 0.01; *** p < 0.001.**

Figure 2.4 illustrates the microscopic characteristics of xylem vessels in PLR and PTR. It was found that the xylem vessels were elliptical, thick-walled and arranged either in radial multiples or solitary for PLR. In addition, it can be observed that PLR had significantly larger xylem vessels (diameter: 0.1390 ± 0.0184 mm; area: 0.2222 ± 0.0042 mm^2; $p<0.01$) distributed densely in the cortex and stele as compared to PTR (diameter: 0.0471 ± 0.0109 mm; area: 0.0037 ± 0.0015 mm^2). The high xylem vessel distribution explains the porous texture of PLR.

Figure 2.5 illustrates the microscopic characteristics of starch granules in PLR and PTR. It was observed that starch granules were abundant and distributed in both PLR and PTR. In PLR, the granules were spherical, hemispherical or ellipsoidal, with a mean diameter of 5.2830 ± 0.5495 μm. They were mostly singular, and at times, existed as a compounded granule consisting of 2 to 10 starch granules. In

contrast, the starch granules in PTR were frequently found to be a compounded granule, which comprised of 2 to 20 individual granules. They were in the shape of a sphere or ellipse, with a mean diameter of 4.6790 ± 0.4455 μm. These results revealed that the diameter of starch granules was similar between the two species ($p>0.05$) and may not be a suitable criterion for differentiating them. In contrast, the mean total starch content in PTR (76.7954 ± 2.9905 mg starch/ g DM; $p<0.001$) was at least seven times greater than that in PLR (10.5288 ± 1.2559 mg starch/ g DM). The quantification of the starch content was in agreement with the macroscopic and microscopic characteristics observed. Thus, the high total starch content explained the starchy texture of PTR and was found to be a suitable parameter for the differentiation of PLR and PTR.

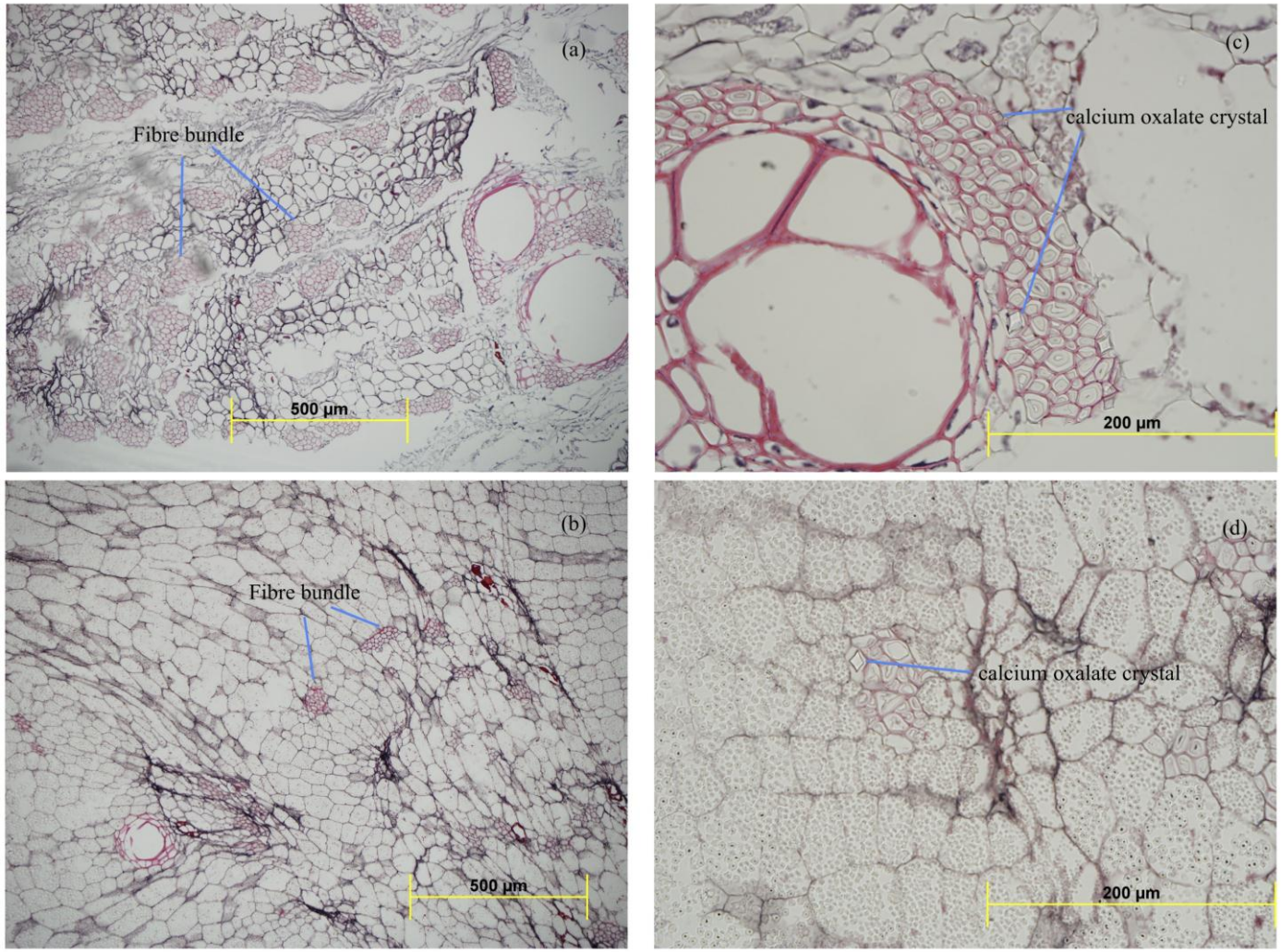


Figure 2.3 Microscopic characteristics of fibre bundle in (a) PLR and (b) PTR at magnification of 10x and calcium oxalate crystal in (c) PLR and (d) PTR at magnification of 40x (objective lens 10x).

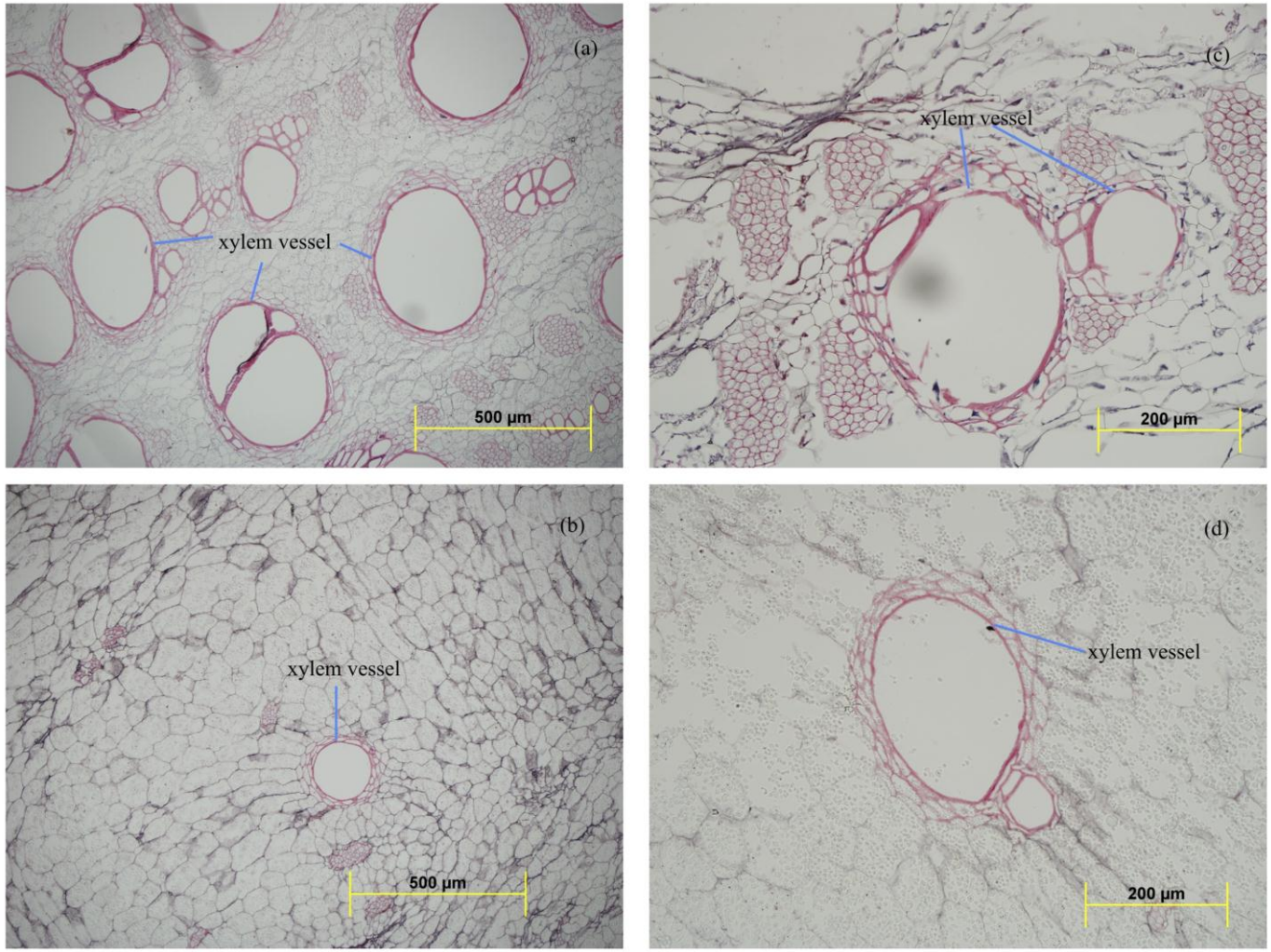


Figure 2.4 Microscopic characteristics of xylem vessel in (a) PLR and (b) PTR at magnification of 10x and xylem vessel in (c) PLR and (d) PTR at magnification of 20x (objective lens 10x).

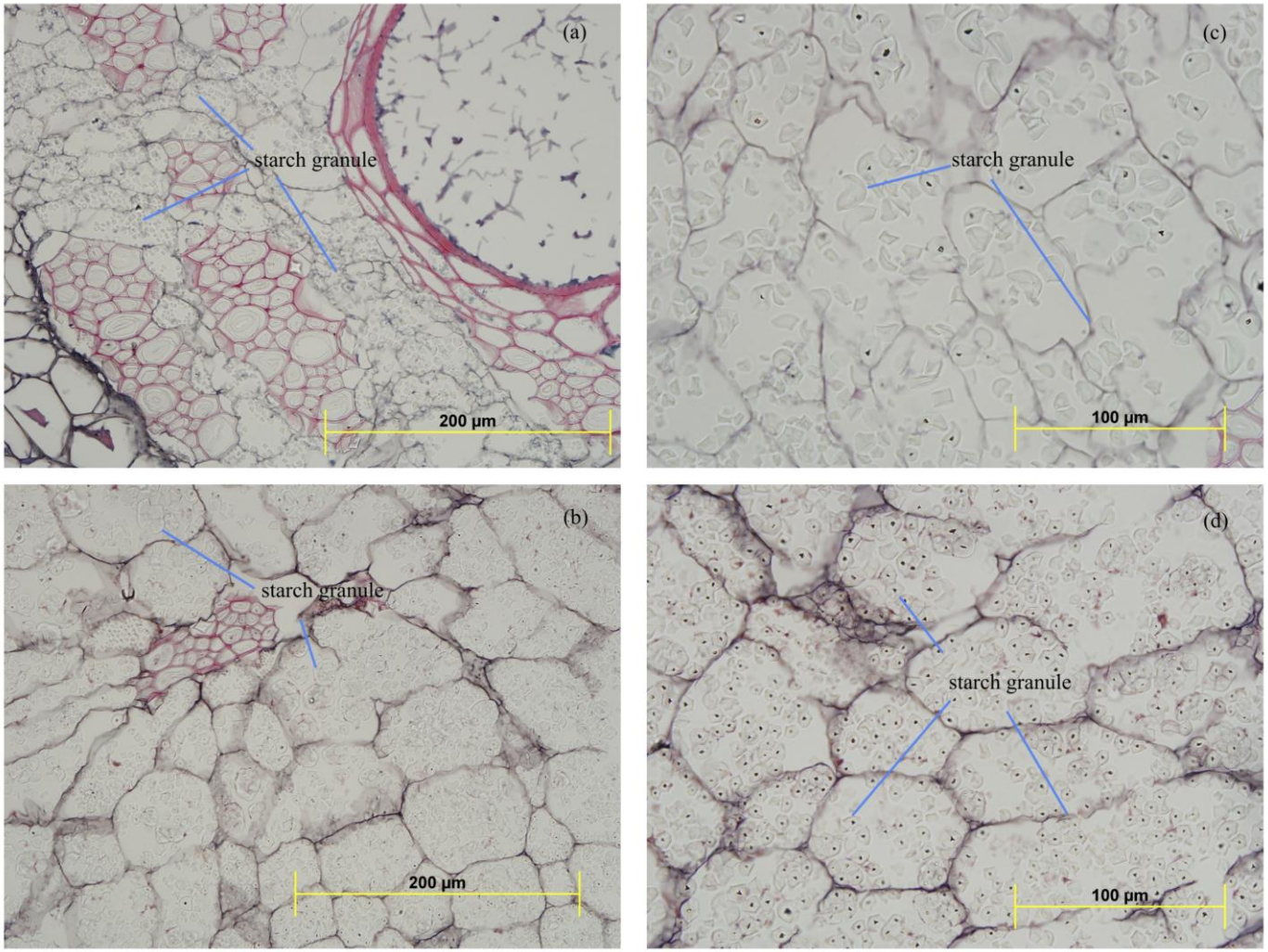


Figure 2.5 Microscopic characteristics of starch granule in (a) PLR and (b) PTR at magnification of 40x and starch granule in (c) PLR and (d) PTR at magnification of 60x (objective lens 10x).

2.4 Conclusion

This is the first extensive investigation to demonstrate the macroscopic and microscopic differences between PLR and PTR. In addition, these morphological differences were further supported by the results obtained from the two colorimetric assays. It is anticipated that the combined methods of macroscopy, microscopy and colorimetry will contribute towards a detailed understanding of the morphological characteristics of PLR and PTR. The results will provide important information for the quality control of these two species and subsequently, the use of the correct species in clinical practice. Therefore, it is proposed that these techniques are employed as part of the criteria for the authentication of PLR and PTR in the pharmacopoeias and monographs.

Chapter Three

Comparing the chemical profiles of Puerariae Lobatae Radix and Puerariae Thomsonii Radix

3.1 Introduction

The chemical profile of the herb is defined by the types of chemical constituents within the herb. One approach that provides a preliminary indication of the types of chemical families present or absent in the herb is to carry out several qualitative experiments (Table 3.1). These qualitative tests are based on treating the sample with a succession of reagents, which only react with a particular group of chemical family (Raaman, 2006). These tests have been applied to identify the presence of triterpenoids, tannins, starch and proteins in Thespesia Populneae Radix (Patil et al., 2012), and saponins, glycosides, phytosteroids and phenolic compounds in Costus Speciosus Rhizoma (Verma and Khosa, 2012). The experimental procedures are simple and convenient and the results are generated instantly.

Another approach for analysing the chemical profile of a herb is to use colorimetric assays, which takes into account of all the chemical constituents from a specific chemical family. The most commonly used colorimetric method is quantifying the total amount of phenolics and flavonoids in a herbal sample (Chang et al., 2002). Although PLR has been shown to have a different chemical profile as compared to PTR, there is no study on quantifying the total flavonoid content using colorimetry and investigating how these chemical differences between PLR and PTR could impact of their respective antioxidant capacity. In a previous study (Zhang et al., 2011a), the 2,2-diphenyl-1-picrylhydrazyl (DPPH) antioxidant capacity of PLR and PTR was assessed by HPLC coupled with flow injection chemiluminescence. Hierarchical clustering analysis demonstrated that the DPPH-active chromatographic fingerprint of PLR was significantly different from PTR. However, this approach is time-consuming and requires special instruments, which restricts

its application as a screening method. On the contrary, the adaptation of colorimetry with ELISA 96 well plate is rapid and convenient and hence, can be applied as a high-throughput screening method. Therefore, the investigation on the total flavonoid content (TFC) (Chapter 1 p.33) and DPPH antioxidant capacity (Chapter 1 p.34) of PLR and PTR using colorimetry could provide further insight on the chemical profiles of these two species and might potentially be used as a quality control method.

The quantification of major chemical constituent puerarin using high–performance liquid chromatography (HPLC) is recommended by various pharmacopoeias and monographs for the authentication of PLR and PTR (BP, 2014a; EP, 2012a, b; JP, 2011; PPRC, 2010a, b). Ultra-performance liquid chromatography (UPLC) refers to the use of a sub-2 micro particle size with narrow bore column operated under high pressure (≥1000 bar) (Grumbach et al., 2009; Lough and Wainer, 1995; Naushad and Khan, 2014). A recent study demonstrated that UPLC significantly decreased the analytical run time 10-fold and solvent consumption 20-fold as compared to conventional HPLC in analysing ginsenosides from Ginseng Radix (Yang et al., 2010). UPLC has been applied for the quality control of Salviae Miltiorrhizae Radix (Zhong et al., 2009), Panax Quinquefolii Radix (Zhang et al., 2011b), Coptidis Rhizoma (Kong et al., 2009) and Magnoliae Officinalis Cortex (Wang et al., 2010). However, the application of UPLC in comparing the chemical profile between PLR and PTR has not yet been explored.

As compared to the vast amount of literature which focuses on the isolation and identification of chemical constituents from PLR and PTR, there are only a few

studies focusing on comparing and quantifying the chemical components of these two species. Therefore, the aim of this chapter was to compare the chemical characteristics of PLR and PTR (1) qualitatively using established chemical tests and (2) quantitatively using a newly established UPLC method (verified by UPLC-MS). In this study, the TFC of PLR and PTR was quantified and correlated to their respective DPPH activity. In addition to the results obtained from the morphological characteristics in Chapter 2, the qualitative and quantitative analysis of the chemical constituents will provide definitive information on the identity of the samples. The validated UPLC condition will be subsequently used in obtaining chromatographic fingerprints in Chapter 4 and 5.

3.2 Methods and materials

3.2.1 Solvents and chemical reagents

HPLC-graded acetonitrile were purchased from Thermo Fisher Scientific (VIC, Australia). Puerarin (>98%), daidzin (>99%), daidzein (>99%), genistin (>98%) and genistein (>99%) were obtained from Tauto Biotech (Shanghai, China). All other chemicals and solvents were of analytical grade and were obtained from Sigma-Aldrich unless otherwise stated (NSW, Australia).

3.2.2 Herbal samples

The details were mentioned in Chapter 2, section 2.2.2 (p.75).

3.2.3 Qualitative chemical tests

The powdered PLR and PTR samples were extracted using solvents with increasing polarity as follows: hexane, dichloromethane, ethyl acetate, ethanol and

water. The mixture was filtered through a funnel with a Whatman No.1 filter paper. The filtrate were then subjected to ten common qualitative chemical tests for the identification of alkaloids, carbohydrates, anthraquinone glycosides, saponin, proteins, amino acids, phytosterols, flavonoids, indoles and tannins as previously described (Raaman, 2006) (Table 3.1). Briefly, the extract was mixed with the testing reagent (2 – 3 mL) and the colour change was recorded. A chemical family was regarded as "present" if the colour change was the same as the positive control, whereas no colour change was regarded as "absent" of such a chemical.

3.2.4 Herbal samples extraction - sonication

The herbal materials were dried at 40 ^{o}C for 24 hours and were kept in a desiccator prior to use. Each sample was ground into a fine powder using a pulveriser and was sieved by a No.188 (177 µm) sieve. The powdered sample was extracted by mixing with absolute ethanol (1g: 50 mL) and was then placed in a sonication bath at a frequency of 40 Hz for 30 mins. The mixture was filtered through a Buchner funnel using a Whatman No.1 filter paper under vacuum. The ethanolic filtrate was evaporated at 40 ^{o}C under vacuum for 2 hours using a rotary evaporator. Then, the dried ethanolic extract was stored in a glass scintillation vial at −20 ^{o}C prior to further analysis.

Table 3.1 Summary of the experimental procedures of common qualitative tests (Raaman, 2006).

	Reagent preparation	Procedures	Positive control	Indication
Alkaloid				
Dragendroff's test	Boil 5.2 g $Bi_2O_2CO_3$ + 4 g NaI in 50 mL acetic acid; Filtrate and mix with 160 mL ethyl acetate and 1 mL water	Mix 5 mL extract + 2 mL reagent	Berberine	Orange colour mixture
Wagner's test	Dissolve 1.27 g I_2 + 3 g NaI in 5 mL water; Add water to 100 mL	Mix 5 mL extract + 3 drops reagent	Berberine	Reddish-brown precipitate
Carbohydrates				
Fehling's test	Dissolve 6.932 g $CuSO_4$ in 100 mL water; Dissolve 34.6 g $KNaC_4H_4O_6$ + 10 g NaOH in 100 mL water; Mix the two solution in equal volume	Mix 1 mL extract + 1 mL reagent	Starch	Red precipitate
Benedict's test	Boil 17.3 g $Na_3C_6H_5O_7$ + 10 g Na_2CO_3 in 80 mL water; Mix with 1.73 g $CuSO_4$	Mix 1 mL extract + 1 mL reagent; Boil for 5 mins	Starch	Orange precipitate
Molish's test	Dissolve 10 g α-naphthol in 100 mL absolute ethanol	Mix 1 mL extract + 0.2 mL reagent; Add 1 mL H_2SO_4	Starch	Purple ring
Anthraquinone glycosides				
Borntrager's test	Boil 10 mL extract with 10 mL HCl for 2 hrs	Mix 2 mL extract + 3 mL $CHCl_3$; Remove aqueous layer and add 10% NH_3	Emodin	Reddish-orange-coloured mixture

Saponins				
Froth test		Dissolve 10 mL extract in 10 mL water; Shake for 5 mins	Total saponins	Stable foam
Proteins				
Biuret's test	Dissolve 10 g NaOH in 100 mL water; Dissolve 1 g $CuSO_4$ in 100 mL water	Mix 2 mL extract + 2 mL NaOH; Add 3 drops CuSO4	Gelatin	Purplish-pink-coloured mixture
Amino acids				
Ninhydrin test	Dissolve 5 g ninhydrin in 100 mL acetone	Mix 2 mL extract + 2 mL reagent	Amino acid	Purple-coloured mixture
Steroids				
Salkowski's test		Mix 1 mL extract + 2 mL $CHCl_3$ + 2 mL H_2SO_4	Steroid	Deep red-coloured mixture
Flavonoids				
Shinoda's test		Mix 1 mL extract + 2 mL 95% EtOH + 5 drops HCl + 0.3 g magnesium	Rutin	Pink-coloured mixture
Sodium hydroxide test	Dissolve 10 g NaOH in 100 mL water	Mix 1 mL extract + 1 mL reagent + 5 drops HCl	Rutin	Pink-coloured mixture
Indoles				
Ehrlich's test	Dissolve 1 g p-dimethylaminobenzaldehyde in 50 mL absolute ethanol + 50 mL HCl	Mix 1 mL extract + 1 mL reagent	Hydroxytryptamine	Change from yellow to colourless
Tannins				
Vanillin hydrochloric acid test	Dissolve 4.8 g vanillin in 60 mL methanol + 30 mL HCl + 10 mL water	Mix 1 mL extract + 1 mL reagent	Catechin	Reddish-pink-coloured mixture

3.2.5 UPLC condition

The UPLC analyses was performed on a Waters Acquity UPLC® H series (Waters, MA, U.S.A.) equipped with Waters Empower 3 Chromatography Data software as described in my previous study (Wong et al., 2013). Acetic acid (0.3%) in water was used as the aqueous phase (solvent A), while 0.3% acetic acid in acetonitrile was used as the organic phase (solvent B). The flow rate was set at 0.2 mL/min and the binary elution is shown in Table 3.2 (Wong et al., 2013). The system was connected to a Waters BEH C18 column (1.7 µm, 2.1 mm × 150 mm) with a Waters BEH C18 guard column. The injection volume was 0.5 µL, while the detection wavelength was set at 254 nm. For the UPLC analysis, the PLR and PTR ethanolic extracts were redissolved in absolute ethanol (PLR: 1 mg/mL; PTR: 5 mg/mL) and filtered through a 0.22 µm polyvinylidene fluoride (PVDF) syringe filter.

Table 3.2 UPLC gradient profile for analysis of chemical constituents of PLR and PTR.

Retention time (min)	Solvent A (%)	Solvent B (%)
0	90	10
7	75	25
8	72	28
12	68	32
13	60	40
15	60	40
16	90	10
20	90	10

3.2.6 UPLC method validation

The method was validated according to the International Conference on Harmonisation guideline (ICH, 2005). The stock standard solutions of the five standard compounds, including puerarin, daidzin, genistin, daidzein and genistein were prepared and diluted with methanol to appropriate concentrations for the construction of the calibration curves. Each calibration curve was established by running the authentic standard compound at ten different concentrations (0.2 – 800 μg/mL) in triplicate. The calibration curve was obtained by plotting mean peak area versus the concentration of each reference compound. The limit of detection (LOD) and limit of quantification (LOQ) were defined as 3 and 10 times of the signal-to-noise ratio, respectively. The precision and accuracy of the chromatographic method were established by analysing five replicates of quality control samples at low, medium and high concentrations (puerarin and daidzin: 160, 320, 640 μg/mL; genistin, daidzein and genistein: 100, 200, 400 μg/mL). To determine the intra-day precision and accuracy for each standard compound, the quality control samples were examined six times within one day. The inter-day precision and accuracy were determined by analysing the quality control samples on three consecutive days, in which each sample was injected six times daily. The precision was expressed as the relative standard deviation (RSD (%)) of repeated measurements, whereas accuracy was expressed as the relative error (RE (%)), which was calculated using the formula: RE (%) = [(mean of observed concentration – spiked concentration)/ spiked concentration] × 100.

3.2.7 Identification of chemical constituents using UPLC-MS/MS

The mass spectrometry analysis was carried on a similar Waters Acquity series consisting of a H class binary solvent manager, an Acquity sample manager-FTN, an Acquity DAD detector and a Xevo tandem triple quadropole mass spectrometer equipped with a Zspray ESI interface (UPLC-MS/MS; Waters, MA, U.S.A.). Standard autotune was applied to adjust the MS parameters for full mass scan range prior to analysis. ESI positive mode was chosen to detect the ion fragmentation throughout the experiment. Nitrogen was used as a nebulising gas with a temperature of 300 ^{o}C and at a flow rate of 1000 L/hr. The capillary voltage was set at 3.5 kV with a cone voltage of 85 V. All analyses were performed with a mass precision of 0.5 atomic mass units (amu). In multiple reaction monitoring (MRM) mode, the collision energy for each standard was optimised using direction infusion with increasing energy from 5 – 50 eV. The optimal collision energy for puerarin, daidzein, genistin, daidzin and genistin was set at 20, 30, 30, 20 and 15 eV, respectively.

3.2.8 Total flavonoid content

Total flavonoid content (TFC) was measured as described previously (Bao et al., 2005). Briefly, 0.5 mL of diluted ethanolic extract was mixed with 2.0 mL of deionised water and 0.15 mL of sodium nitrite (5% w/v). After 5 min, 0.15 mL of aluminium chloride (10% w/v) was added. Five minutes later, 1 mL of sodium hydroxide (1M) was added to the mixture and mixed thoroughly. The reaction mixture was kept for another 15 min at room temperature. The absorbance was measured at 415 nm and was expressed as gram of quercetin equivalents (QE)/100 g of dried extract (DE).

3.2.9 2,2-Diphenyl-2-picrylhydrazyl (DPPH) free radical scavenging assay

The DPPH free radical scavenging activity was carried out as previously described (Cheng et al., 2006). The DPPH stock solution was prepared by dissolving 24 mg in 100 mL methanol. The stock solution was further diluted with methanol until the absorbance of the solution reached 1.0 ± 0.05 Abs at 550 nm. The DPPH working solution was prepared daily. Subsequently, 10 µL of diluted ethanolic extract was mixed with 200 µL of DPPH solution. The mixture was incubated in the dark at room temperature for 30 mins. The absorbance was measured at 550 nm and was expressed as milligram trolox equivalent antioxidant capacity (TEAC)/100 g of DE.

3.2.10 Data analysis

The statistical analyses were performed on GraphPad Prism version 6.01 (GraphPad Software, CA, U.S.A.). The data was analysed using either one way analysis of variance (ANOVA) or Student's t-test.

3.3 Results and discussion

3.3.1 Qualitative phytochemical characterisation

As shown in Table 3.3, PLR and PTR shared similar types of chemical families such as flavonoids, saponins and carbohydrates. Flavonoids were present in the ethyl acetate, ethanol and water extracts. Carbohydrates were found in both the ethanol and water extract. The presence of carbohydrates was in good agreement with the colorimetric assay results obtained in Chapter 2 (Table 2.6 p.85), which suggested that starch was the main type of carbohydrate in both species. On the contrary, alkaloids, anthraquinone glycosides, protein, amino acids, phytosterols, insoles and tannins were not detected in any of the solvent extracts for both species.

Traces of saponins were detected in the water extracts of both PLR and PTR, but were absent in the non-polar solvent extracts. Previous studies suggested that oleanene-type triterpene glycosides are the major type of saponins in PLR and PTR, which are soluble in water and methanol under heat. The yield of the total flavonoid fraction was three times greater than that of the total saponin fraction, suggesting the dominance of the flavonoid compounds (Arao et al., 1996; Arao et al., 1997).

To determine which solvent was more suitable for extracting chemical constituents from PLR and PTR, several factors needed to be considered. Since it has been found that flavonoids, in particular isoflavonoids, are the major chemical component and is believed to contribute to the pharmacological activities of PLR and PTR, solvents with high polarity such as ethanol and water should be chosen.

However, the use of water as the extraction solvent would extract more carbohydrates as compared to ethanol. From previous experience, it has been observed that carbohydrates are more likely to block the PVDF filter during filtration and the pre-column during chromatographic analysis. Additionally, the use of ethanol provides a complete chemical profile as it extracts both polar and non-polar compounds and is easier to concentrate down as compared to water. Therefore, ethanol was chosen as the extracting solvent and this extract would be used in the subsequent analyses.

3.3.2 UPLC method validation

The LOD, LOQ and calibration curve parameters of puerarin, daidzin, genistin, daidzein and genistein are illustrated in Table 3.4. The LOD and LOQ of the standard compounds were less than 0.08 and 0.27 μg/mL, respectively. These results were similar to a previously published HPLC method (Du et al., 2010). Correlation coefficients ($r^2 \geq 0.9995$) with wide linear ranges (0.2 – 800 μg/mL) were obtained, which indicated that the relationships between the reference compounds and peak area were highly correlated. In addition, this wide linear range is suitable for analysing samples with a variable amount of chemicals. The intra- and inter-day precisions and accuracies are summarised in Table 3.5. The average RSDs of the intra-day precision at low, medium and high concentrations were 0.37%, 0.17% and 0.65%, respectively, whereas the average inter-day precision was 2.69%, 1.59% and 1.14%, respectively. Furthermore, the average intra-day accuracies at low, medium and high concentration were 1.54%, 2.97% and 2.56%, respectively, whereas the average inter-day accuracies were 2.29%, 3.89% and 4.34%, respectively.

Table 3.3 Qualitative chemical characteristics of various solvent fractions extracted from PLR and PTR.

	Puerariae Lobatae Radix					Puerariae Thomsonii Radix				
	Hexane	Dichloromethane	Ethyl acetate	Ethanol	Water	Hexane	Dichloromethane	Ethyl acetate	Ethanol	Water
Alkaloids										
Dragendroff	-	-	-	-	-	-	-	-	-	-
Wagner	-	-	-	-	-	-	-	-	-	-
Carbohydrates										
Fehling	-	-	-	+	+	-	-	-	+	+
Benedict	-	-	-	+	+	-	-	-	+	+
Molish	-	-	-	+	+	-	-	-	+	+
Anthraquinone glycosides										
Borntrager	-	-	-	-	-	-	-	-	-	-
Saponins										
Froth	-	-	-	-	+	-	-	-	-	+
Proteins										
Biuret	-	-	-	-	-	-	-	-	-	-
Amino acids										
Ninhydrin	-	-	-	-	-	-	-	-	-	-
Phytosterols										
Salkowski	-	-	-	-	-	-	-	-	-	-
Flavonoids										
Shinoda	-	-	+	+	+	-	-	-	+	+
NaOH	-	-	+	+	+	-	-	-	+	+
Indoles										
Ehrlish	-	-	-	-	-	-	-	-	-	-
Tannins										
Vanillin-HCl	-	-	-	-	-	-	-	-	-	-

These results revealed that the developed chromatographic method had good accuracy and repeatability and hence can be applied to analyse the samples in the subsequent experiments.

Table 3.4 Limit of detection (LOD), limit of quantification (LOQ) and calibration curve parameters of five standard compounds.

	Linear range (μg/mL)	Correlation coefficient (r^2)	Slope	y-intercept	LOD (μg/mL)	LOQ (μg/mL)
Puerarin	0.2 – 800	0.9995	1.055×10^7	2.084×10^4	0.04	0.13
Daidzin	1 – 800	0.9996	1.039×10^7	2.568×10^4	0.08	0.27
Genistin	1 – 600	0.9995	1.257×10^7	1.687×10^4	0.07	0.23
Daidzein	1 – 600	0.9995	1.535×10^7	3.860×10^4	0.06	0.20
Genistein	0.2 – 500	0.9996	1.964×10^7	4.650×10^3	0.05	0.17

Table 3.5 Validation of the intra- and inter-day precisions and accuracies of five standard compounds at low, medium and high concentration levels.

Compounds	Spiked concentration (μg/mL)	Intra-day (n = 6)			Inter-day (n = 18)		
		Observed concentration (μg/mL) [a]	Precision RSD (%) [b]	Accuracy (%) [c]	Observed concentration (μg/mL) [a]	Precision RSD (%) [b]	Accuracy (%) [c]
Puerarin	160	163.985 ± 0.510	0.307	2.490	164.856 ± 4.777	2.863	3.035
	320	331.499 ± 0.469	0.141	3.593	334.001 ± 5.344	1.591	4.375
	640	665.703 ± 4.431	0.664	4.016	677.678 ± 8.193	1.205	5.887
Daidzin	160	150.871 ± 0.600	0.391	-5.705	151.861 ± 4.292	2.781	-5.087
	320	306.030 ± 0.478	0.155	-4.366	308.348 ± 5.118	1.647	-3.641
	640	611.741 ± 3.914	0.637	-4.416	622.155 ± 7.365	1.179	-2.788
Genistin	100	105.176 ± 0.406	0.381	5.176	105.835 ± 2.858	2.666	5.835
	200	213.231 ± 0.326	0.152	6.615	215.003 ± 3.441	1.590	7.502
	400	425.478 ± 2.823	0.661	6.370	433.123 ± 5.275	1.214	8.281
Daidzein	100	101.404 ± 0.391	0.377	1.404	102.369 ± 2.583	2.463	2.369
	200	206.980 ± 0.327	0.156	3.490	209.447 ± 3.039	1.434	4.724
	400	414.831 ± 2.736	0.656	3.708	423.727 ± 4.410	1.035	5.932
Genistein	100	104.318 ± 0.430	0.411	4.318	105.289 ± 2.804	2.657	5.289
	200	211.009 ± 0.462	0.219	5.504	212.966 ± 3.543	1.662	6.483
	400	412.496 ± 2.565	0.621	3.124	417.631 ± 4.348	1.041	4.408

[a]Mean ± standard deviation (SD)

[b]Relative standard deviation (RSD) % = (SD/mean) × 100

[c]Accuracy % = [(mean observed concentration – spiked concentration) / spiked concentration] x 100

3.3.3 Quantification of chemical content by UPLC

Using UPLC, mean chromatograms of the PLR and PTR and the amount of the major chemical constituents are illustrated in Figure 3.1 and 3.2, respectively. The mean amount of puerarin (peak 2), daidzin (peak 5), genistin (peak 6), daidzein (peak 7) and genistein (peak 8) was found to be 60.69 ± 4.51, 11.83 ± 1.10, 3.00 ± 0.35, 2.31 ± 0.25 and 0.99 ± 0.05 mg/g DM, respectively in PLR. As anticipated, the overall chemical content in PTR was significantly lower than that in PLR. In the PTR samples, the mean amount of puerarin, daidzin, daidzein and genistein was found to be 5.08 ± 0.49, 1.37 ± 0.35, 4.73 ± 0.75 and 0.72 ± 0.06 mg/g DM, respectively. Genistin was not detected in the PTR ethanol extracts.

Furthermore, the UPLC results revealed that puerarin at 6.24 min was the dominant chemical constituent in both species and was higher in the PLR samples (6.07 ± 0.05% (w/w)) as compared to the PTR samples (0.51 ± 0.05% (w/w)). This result indicates that all the samples met the minimal requirement of puerarin content illustrated in the PPRC, EP, BP and HKCMMS (Table 1.4 p.64). Indeed, it was observed that peak 1 at 4.75 min, peak 3 at 6.17 min and peak 4 at 6.36 min were only found in PLR and were absent in PTR.

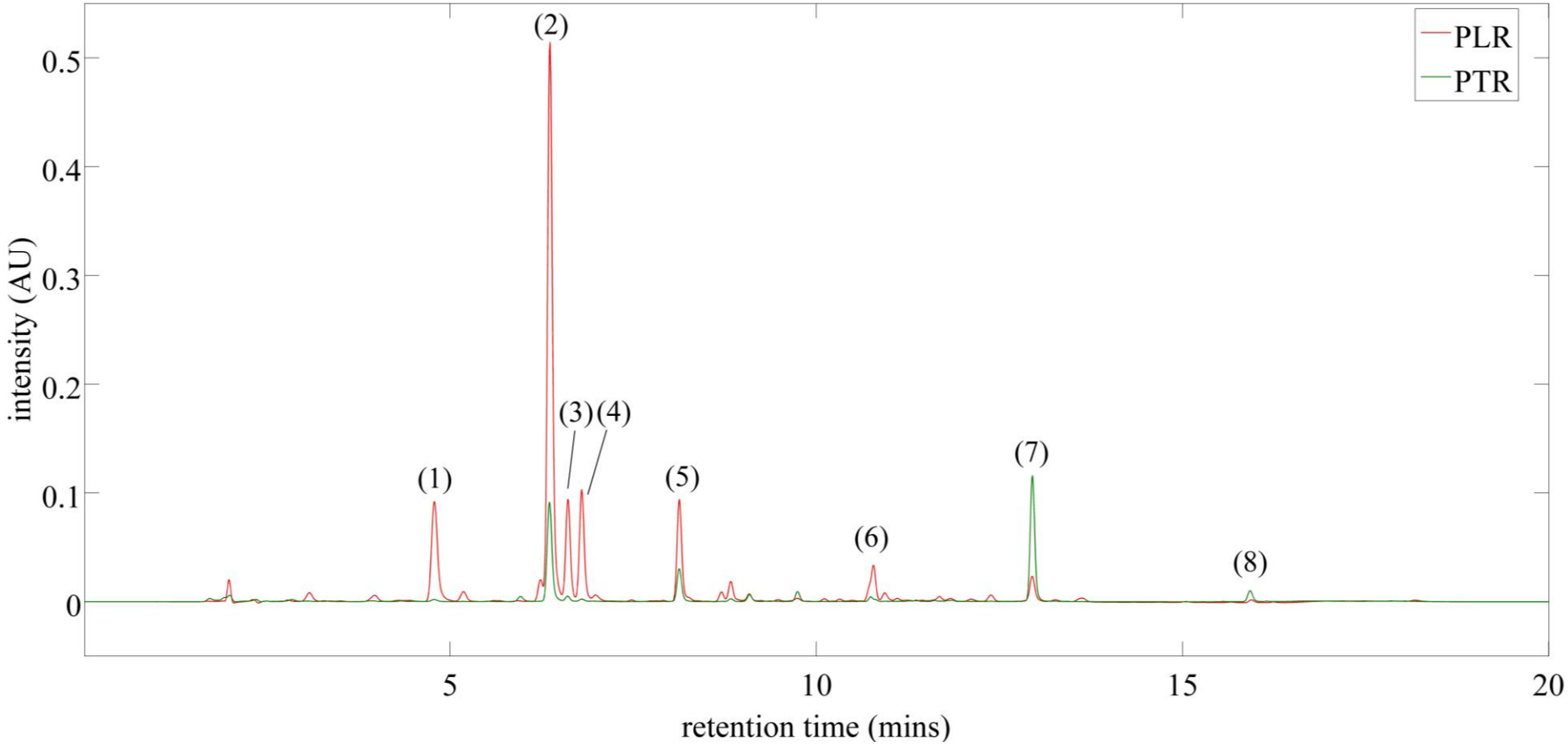


Figure 3.1 Mean UPLC chromatograms of 22 PLR (red) and 20 PTR (green) samples. (1) Unknown; (2) Puerarin; (3) Unknown; (4) Unknown; (5) Daidzin; (6) Genistin; (7) Daidzein; (8) Genistein.

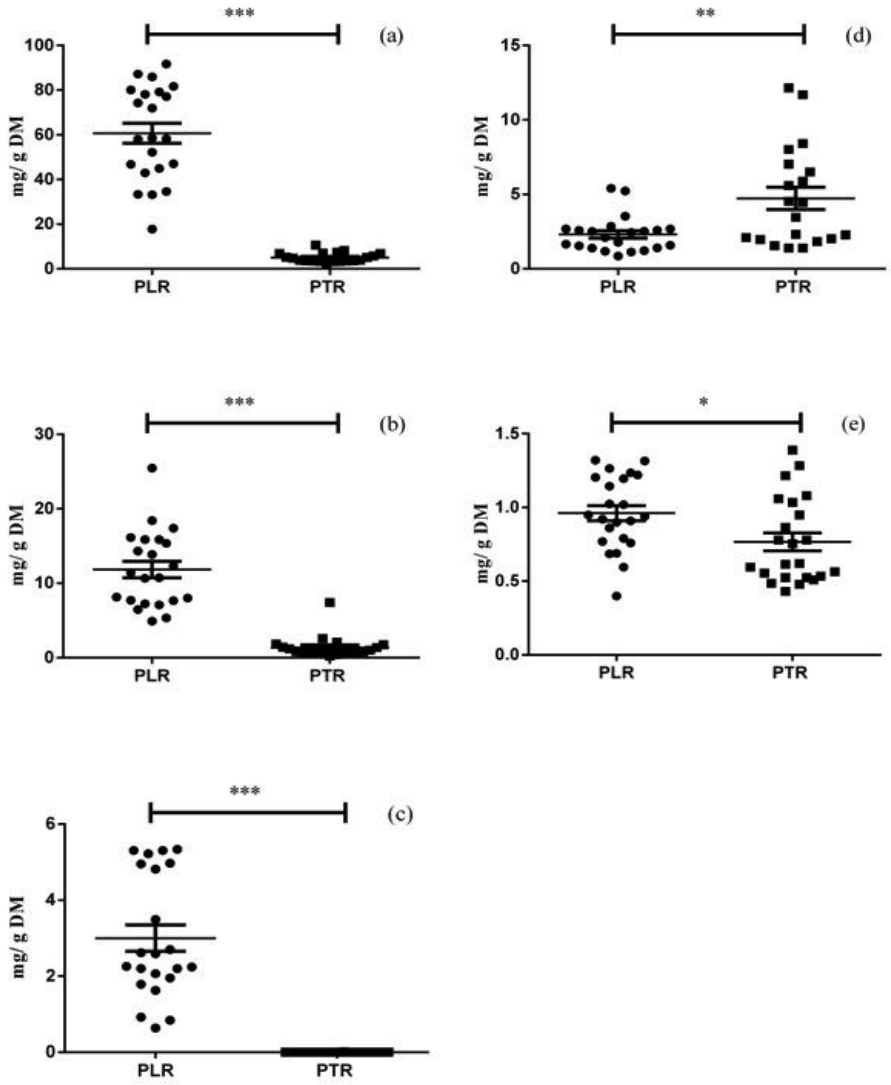


Figure 3.2 Quantification of major chemical constituents from 22 PLR and 20 PTR samples using UPLC. (a) Puerarin; (b) Daidzin; (c) Genistin; (d) Daidzein; (e) Genistein; mg/g DM: milligrams of chemical per gram of dried mass; * p < 0.05; ** p < 0.01; *** p < 0.001. (n = 42) Experiment was performed in triplicate.

3.3.4 Identification of the major chemical markers using UPLC-MS/MS

The identities of the major chemical components were identified by comparing the retention time and UV spectrum (Appendix I p.230) of the reference standards and were confirmed by comparing with the fragmentation pathway (Appendix II p. 232)

described in previous studies (Li et al., 2007; Lin et al., 2005). Preliminary runs were performed to compare the two different ionisation modes, ESI positive and ESI negative. Higher signal intensity and enhanced fragmentation were obtained with ESI positive mode and hence, it was chosen in the subsequent UPLC-MS/MS analysis. The peaks at 6.01, 7.37, 9.44, 12.87 and 16.21 min were identified as puerarin (417.10 m/z), daidzin (416.97 m/z), genistin (432.91 m/z), daidzein (255.07 m/z) and genistein (270.82 m/z), respectively. Both puerarin and daidzin share the same molecular mass (416.38 g/mol) and formula ($C_{21}H_{20}O_9$). What differentiates them is the position of the attached glucose. In puerarin, the glucose is attached to the C-8 position of the aromatic A ring, forming a C-glucosidic bond. Since the C-glucosidic bond is much more rigid than the O-glucosidic bond, a different fragmentation pathway was observed (Prasain et al., 2003; Zhang et al., 2010). Firstly, a fragment of $C_4H_8O_4$ was lost, resulting in a cation $[M+H-C_4H_8O_4]^+$ with a mass of 297.07 m/z. The glucose remnant further disintegrated into a $[M+H-C_4H_8O_4\text{-}CHO]^+$ ion, which contributed to the loss of a CHO fragment with a mass of 29 m/z (Figure 3.3) (Li et al., 2007; Lin et al., 2005).

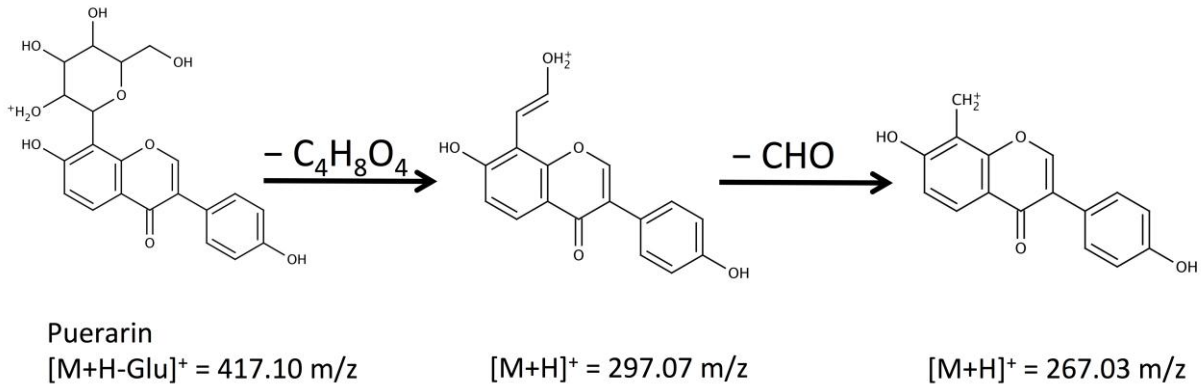


Figure 3.3 Fragmentation pathway of puerarin in ESI positive ion mode.

By comparing the molecular weight and fragmentation pathway with literature, peaks at 4.75 (peak 1) and 6.36 (peak 4) min were identified as 3'-hydroxypuerarin (433.04 m/z) and 3'-methoxypuerarin (447.15 m/z), respectively. These two compounds share the chemical structure as puerarin, except with the addition of either a hydroxyl (-OH) or methoxy ($-OCH_3$) group on the C-3 position of the aromatic B ring (Figure 3.4). These two compounds followed the same fragmentation pathway as puerarin, except that for each individual ion fragment, the mass is higher by 16 and 30 m/z, respectively (Lin et al., 2005; Prasain et al., 2003).

The peak at 6.17 (peak 3) min was identified as 6"-O-D-xylosylpuerarin, which generated a $[M+H]^+$ ion at 549.18 m/z. Interestingly, the chemical structure of 6"-O-D-xylosylpuerarin is similar to that of puerarin, with an addition of an xylose moiety attached to the glucose moiety of puerarin. After the detachment of the xylose moiety, a fragment cation $[M+H\text{-}xyl]^+$ with a mass of 417.82 m/z was generated, which is the same as the $[M+H]^+$ ion of puerarin. Hence, this fragment ion followed the same fragmentation pathway as puerarin (Li et al., 2007; Prasain et al., 2007).

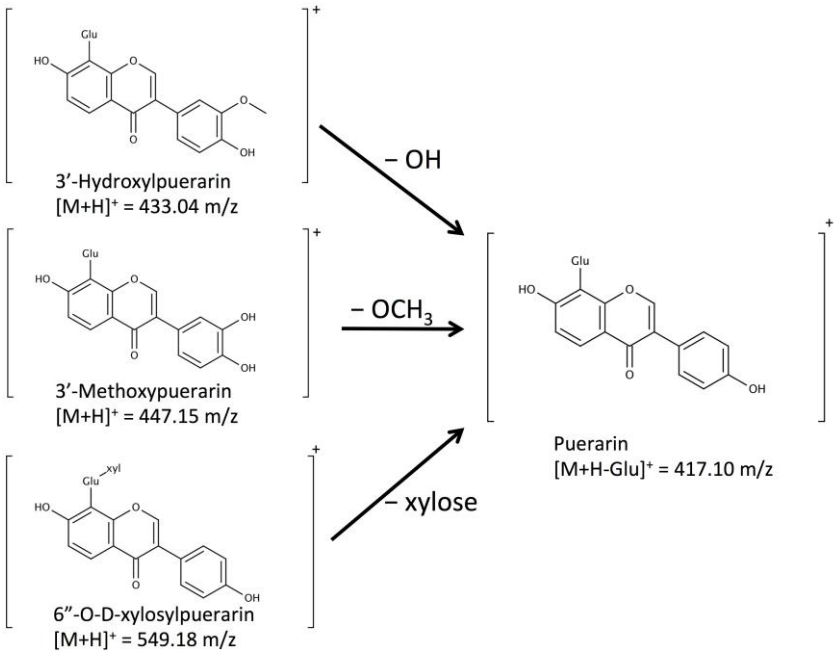


Figure 3.4 Fragmentation pathway of 3'-hydroxypuerarin, 3'-methoxypuerarin and 6"-O-D-oxylosylpuerarin to puerarin in ESI positive ion mode.

3.3.5 Total flavonoid content and antioxidant capacity

Figure 3.5 illustrates the TFC and DPPH free radical scavenging capacity of the ethanolic extracts of PLR and PTR. The results showed that the PLR ethanolic extracts had significantly greater TFC ($p<0.001$) and DPPH antioxidant capacity ($p<0.001$) as compared to PTR. For PLR, the mean values of the TFC and DPPH assay were 108.28 ± 2.49 g QE/ 100 g DE and 133.70 ± 3.10 µM TEAC/ 100 g DE, respectively. For the PTR ethanolic extracts, the mean TFC and DPPH values were 24.51 ± 2.84 g QE/ 100 g DE and 27.24 ± 1.80 µM TEAC/ 100 g DE, respectively. The mean values of TFC correlated well with the mean values of the DPPH activity with a correlation coefficient of 0.9430 (Figure 3.5c).

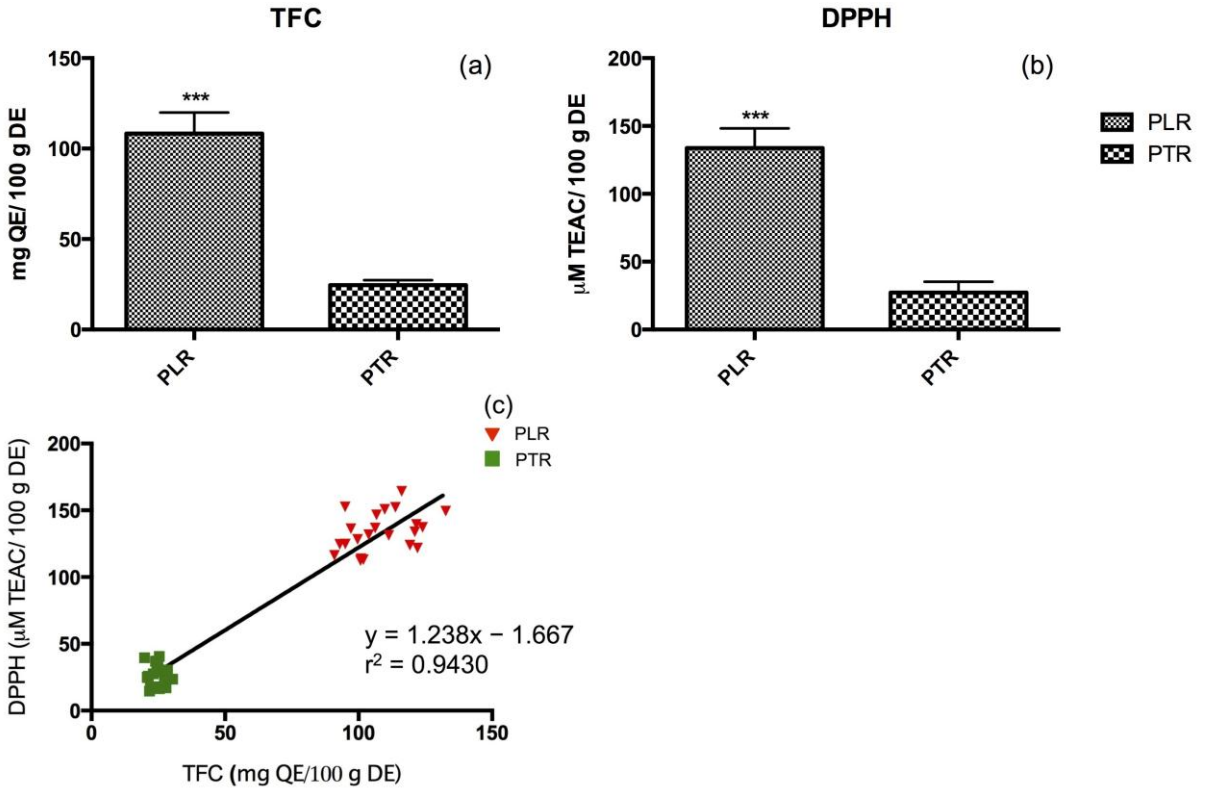


Figure 3.5 Correlation between total flavonoid content and DPPH antioxidant capacity. (a) Total flavonoid content assay; (b) DPPH free radical scavenging assay; (c) Correlation between total flavonoid content and DPPH capacity of PLR and PTR; mg GAE/ 100 g DE: mg gallic acid equivalent per 100 g of dried extract; µM TEAC/ 100 g DM: µM trolox equivalent antioxidant capacity (TEAC)/100 g of dried extract; (n = 42) *** p < 0.001. Experiment was performed in triplicate.

The chemical profiles of PLR and PTR were summarised in Table 3.6. Isoflavonoids have been shown to possess strong reducing power and can act as an electron donor in a redox reaction. The chemical colorimetric assays revealed that the TFC in the PLR ethanolic extracts were significantly greater than in the PTR ethanolic extracts, which was in a good agreement with the results obtained from the UPLC quantification. Additionally, the DPPH antioxidant assay demonstrated that the PLR ethanolic extracts possessed superior free radical scavenging activity as compared

to the PTR ethanolic extracts. These results were correlated to their differences in the chemical contents obtained from the TFC colorimetric assay and were in good agreement with previous studies, which suggested that the chemical content contributed to the antioxidant activity of a herb. The higher the content of phenolic and flavonoid compounds, the greater the free radical scavenging capacity (Kam et al., 2013; Malencic et al., 2007). This phenomenon explained the superior radical scavenging capacity of PLR as it has a significantly greater amount of TFC compared to PTR. In another study (Jiang et al., 2005), HPLC and peroxyl radical 2,2'-azo-bis-(2-amindinopropane) dihydrochloride were used to assess the chemical content and antioxidant capacity, respectively of PLR and PTR. The results showed that the amount of puerarin, daidzin and daidzein in the PLR water extract was five, three and three times, respectively greater than in the PTR water extract. Furthermore, the free radical scavenging capacity of the PLR water extract was five times greater than in the PTR water extract (Jiang et al., 2005). Such reducing capacity is beneficial to various diseases as it may rebalance the cellular ROS and hence, offset the oxidative stress. The higher antioxidant capacities in PLR may also related to the presence of 3'-hydroxypuerarin and 3'-methoxypuerarin, which are absent in PTR. 3'-Hydroxypuerarin and 3'-methoxypuerarin have been reported to possess antioxidant property against peroxynitrite, nitric oxide, superoxide anion and DPPH free radical (Jin et al., 2012).

Table 3.6 Summary of the chemical profiles of PLR and PTR.

	PLR	PTR
Chemicals types detected in qualitative assays	Flavonoids, saponins, carbohydrates	Flavonoids, saponins, carbohydrates
Chemicals detected in UPLC	Puerarin, daidzin, genistin, daidzein, genistein	Puerarin, daidzin, daidzein, genistein
TPC	High	Low
DPPH	High	Low

3.4 Conclusion

In this study, the qualitative and quantitative chemical assays confirmed that flavonoids, in particular isoflavonoids, were the major chemical constituents in both PLR and PTR. To support these results, a rapid UPLC method for the quantification of chemical constituents in PLR and PTR was developed and validated. Puerarin was the major chemical constituent in both species. Puerarin, daidzin and daidzein are the common chemicals found in both species. The amount of 3'-hydroxypuerarin, 3'-methoxypuerarin, 6"-O-D-oxylsylpuerarin and genistin in PLR were significantly greater than PTR. Furthermore, the PLR ethanolic extracts has been shown to possess significantly higher amount of total phenolic content and DPPH scavenging capacity as compared to the PTR ethanolic extracts. The results obtained from this study provide further information in authenticating the species of the samples and the confirmed species identity will be used as one of the criteria to build the classification models in subsequent analyses.

Chapter Four

Differentiation of Puerariae Lobatae Radix and Puerariae Thomsonii Radix using ultra-performance liquid chromatography coupled with partial least squares-discriminant analysis

The materials presented in this chapter were extracted from **Wong, K. H.,** Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. (2013). Differentiation of Pueraria lobata and Pueraria thomsonii using partial least squares discriminant analysis (PLS-DA). *J Pharm Biomed Anal.*, 84 5-13.

4.1 Introduction

Single or a few chemical markers are commonly used as a tool for the quality control of herbal materials and their products. However, there are obvious problems with this approach as the chemical marker may not be the most pharmacologically active component of the herb. It may be present in other species or herb families and may not be available as a reference standard. For the authentication of PLR and PTR, the major chemical component puerarin is used as the chemical marker. Puerarin alone does not represent the complexity of PLR and PTR. Puerarin also exists in other species within the *Pueraria* genus, including *Pueraria omeiensis* and *P. phaseolides* and its content varies vary due to harvest season, habitat, rainfall, latitude, soil conditions and other factors (Wong et al., 2011). Therefore, it is necessary to examine whether the use of puerarin alone is sufficient to differentiate PLR from PTR.

To address the issues of using a single chemical marker, chemical fingerprint analysis is introduced, which defines a unique pattern of a herb and reflects the presence of multiple chemical constituents. This approach follows the fundamental holistic theory of traditional Chinese medicines as all chemicals may contribute to the quality, safety and efficacy of a herbal medicine as therapeutic effects are based on the additive and/or synergistic interactions of numerous components (Tistaert et al., 2011a). Another major advantage of using the entire chromatographic fingerprint is that calibration curves and the quantification of chemical marker(s) are not required (Liang et al., 2010). In this study, chromatographic fingerprints of PLR and PTR based on the UPLC condition established in Chapter 3 are investigated.

The complexity of the whole fingerprint of the herb makes it difficult to assess the subtle differences between samples manually. Thus, multivariate statistical analysis combined with the fingerprint pattern has been widely used for metabolic profiling and the characterisation of various plants, food and biological tissues (Arvanitoyannis and Vlachos, 2009; Gad et al., 2013; Madsen et al., 2010). Several multivariate analytic methods have been applied in the quality control of herbal medicines (Alaerts et al., 2010; Berrueta et al., 2007; Tistaert et al., 2011a). Zhang and his colleagues (Zhang et al., 2011a) used HPLC in conjunction with flow injection chemiluminescence to correlate chromatographic peaks with their corresponding antioxidant activity. Similarity analysis was unable to differentiate PLR from PTR, whereas two distinct clusters of the species were obtained in hierarchical clustering analysis. In another study, similarity analysis failed to determine the geographical origins of PLR using the HPLC chromatographic fingerprints (Zhao et al., 2011). Furthermore, Lau and his colleagues (Lau et al., 2009) used near-infrared (NIR) spectroscopy to authenticate PLR and PTR. Linear discriminant analysis and soft independent modelling of class analogy models were constructed and achieved a 100% classification rate. However, unlike liquid chromatographic analyses, the use of NIR spectroscopy alone was unable to identify important chemical components, which could differentiate the two species.

It can be observed that the failure of most of the previous studies may be related to an inappropriate choice of analytical measurement and/or multivariate analysis. The limited samples size could also impact on the quality of the data obtained. In this study, partial least squares-discriminant analysis (PLS-DA) was employed as it determines the maximal covariance between X and Y matrix, which is lacking in

supervised pattern recognition analysis such as similarity analysis and hierarchical clustering analysis. This technique has been successfully applied in the differentiation of Citri Reticulatae Pericarpium from Citri Aurantii Pericarpium (Tistaert et al., 2011b), Scutellariae Radix from six different geographical origins (Li et al., 2011b) and Phyllanthus Niruri Herba and Phyllanthus Tenellus Herba from Phyllanthus Urinaria Herba, Phyllanthus Caroliniensis Herba, Phyllanthus Amarus Herba and Phyllanthus Stipulates Herba (Martins et al., 2011).

The authentication of raw herbal materials used in manufacturing herbal products is particularly important if there are one or more similar species available in the market as with *Pueraria* species. Herbal granules are now becoming one of the major dosage forms for delivering herbs in Asian and Western countries due to their ease of administration, transport and storage. Granules are formulated by mixing the concentrated water extract of a herb, combined with other excipients such as starch, dextrin, lactose and soluble fibre and presenting as a dried powder (Luo et al., 2012). Unlike orthodox prescription medicines, herbal granules are less strictly regulated and hence, their formulations and qualities may vary from different manufacturers. This could have a major impact on their quality, safety and efficacy.

To the best of my knowledge, there is no publication focusing on the differentiation of PLR and PTR using PLS-DA and no studies have investigated the quality of commercial PLR granules available in the market. Therefore, the aims of this chapter were to differentiate PLR from its related species PTR and to examine the raw herbal material used in manufacturing PLR granules using PLS-DA. In this

study, the UPLC chromatograms of 46 raw PLR and PTR and 17 granule products were evaluated. There are limited studies employing the entire chromatographic fingerprint for the quality control of PLR and PTR and this was compared to using the single chemical marker, puerarin. In hindsight, it is important to investigate the raw materials used in the manufacturing of herbal products so that consumers are confident with the products they are taking.

4.2 Materials and methods

4.2.1 Solvent, standard compounds and chemical reagents

HPLC-graded acetonitrile and glacial acetic acid were obtained from Thermo Fisher Scientific (VIC, Australia). Deionised water was purified by a Milli-Q water purification system from Millipore (MA, U.S.A.). Puerarin (>98%), daidzin (>99%), daidzein (>99%), genistin (>98%) and genistein (>99%) were purchased from Tauto Biotech (Shanghai, China). All other chemicals were purchased from Ajax Finechem (NSW, Australia) unless otherwise stated.

4.2.2 Herbal samples

Forty-six dried PLR and PTR (Chapter 2 p.75) and 17 commercial PLR granules were purchased from herbal pharmacies in various regions of China, Australia and U.S.A. (Table 4.1). Voucher specimens were deposited at the Faculty of Pharmacy, The University of Sydney, Australia.

Table 4.1 Sample code and production location of commercial PLR granules.

Sample code	Production location
G1	Nanning, China
G2	Taichung, Taiwan
G3	Taoyuan, Taiwan
G4	Yilan, Taiwan
G5	Guangdong, China
G6	Taichung, Taiwan
G7	Taichung, Taiwan
G8	Pintong, Taiwan
G9	Guangzhou, China
G10	Sichuan, China
G11	Guangzhou, China
G12	Beijing, China
G13	Guangdong, China
G14	Jiangsu, China
G15	Beijing, China
G16	Guangdong, China
G17	Beijing, China

4.2.3 Herbal sample extraction - reflux

The herbal samples were dried at 40 ^{o}C for 24 hours before use. Each dried sample was ground to fine powder using a pulveriser and passed through a No. 180 (177 μm) sieve. To mimic the granules extraction procedure, the sample was extracted with deionised water. Each sample (10 g) was accurately weighed and macerated with 400 mL of deionised water. The mixture was extracted under reflux at 95 ^{o}C for 3 hours. The mixture was allowed to cool to room temperature and centrifuged at 10,000 g for 10 mins. The supernatant was filtered through Whatman No.1 filter paper under vacuum. The filtrate was evaporated at 40 ^{o}C and subsequently lyophilised. The dried residue was considered as the water extract. The water extracts and commercial granules were further extracted with

methanol (1:40 g/mL) in a sonication bath for 30 mins. Granules are prepared by freeze- or spray-drying the concentrated decoction in combination with other pharmaceutical excipients such as dextran and starch powder. This step eliminated starch and other hydrophilic excipients in granules. This procedure was repeated three times and the filtrates were combined. The final solution was evaporated and lyophilised. The dried residue was stored in a glass scintillation vial at -20 ^{o}C prior to analysis. During chromatographic analysis, the methanolic extract was re-dissolved in methanol (PLR and granules: 1 mg/mL; PTR: 5 mg/mL) and filtrated through a 0.22 μm syringe filter.

4.2.4 UPLC measurement

The details of the UPLC conditions are outlined in Chapter 3 (p.98). The chromatograms were documented by Waters Empower 3 Chromatography Data software at a frequency of 0.5 s/data point with a total of 2401 data points per sample.

4.2.5 Chemometrics data processing

Several pre-processing techniques were applied on the raw chromatographic matrix to improve the interpretability of the discriminant models. Correlation optimised warping (COW) removed retention time shift and the reference chromatogram, slack number and the number of segments were optimised as proposed by Skov et al. (Chapter 1 p.38) (Skov et al., 2006). The aligned chromatographic fingerprints were subjected to baseline removal, followed by standard normal variate (SNV) (Chapter 1 p.39) and column centering (Chapter 1 p.40). The pre-processed chromatographic fingerprints were then divided into a

calibration and a validation set using Kennard-Stone (K-S) algorithm (Chapter 1 p.56). Furthermore, variable selection algorithms such as genetic algorithm (GA) (Chapter 1 p.41) and successive projection algorithm (SPA) (Chapter 1 p.41) selected important variables from the pre-processed chromatographic fingerprints. In SPA, the maximum number of selected variables was set at 26. The GA routine was performed on 100 generations with a population size of 30 in each generation. Mutation rate and convergence were set at 1% and 50%, respectively. Partial lease squares (PLS) regression with 10 latent variables was chosen to validate the model, associated with a full random cross-validation of five splits. The selection of the optimal number of latent variables was based on the RMSECV from the LOOCV. The algorithm was repeated ten times. Apart from using automatic algorithm, variables were selected using PLS-DA loading plot presented by Tistaert et al. (Tistaert et al., 2011b). In addition, the use of the single chemical marker puerarin, proposed by the PPRC, was compared to the entire chromatographic matrix. Principal component analysis (PCA) (Chapter 1 p.43) was used to detect the presence of outliers. PLS-DA (Chapter 1 p.49) models using the entire chromatographic fingerprint, variables selected by GA, SPA, loading plot and puerarin were established and their model's performance such as root mean square values and correlation coefficient were compared. Lastly, the established PLS-DA models were applied to assess the quality of commercial PLR granules. The models were validated using root mean square errors and regression coefficient (Chapter 1 p.57). Similarity analysis (Chapter 1 p.45) based on the mean chromatograms was used to assess the similarity of the chromatographic characteristics between granules, PLR and PTR.

4.2.6 Software

Data processing and modelling were performed on MATLAB R2012b (The MathWorks, MA, U.S.A.). The algorithms for PCA, PLS-DA and GA were from the PLS toolbox (Eigenvector Research Incorporated, WA, U.S.A.). The K-S and COW algorithms were written on MATLAB m-file as previously described (Daszykowski et al., 2002; Skov et al., 2006). The SPA toolbox 1.0 was developed by Araujo et al. (Araujo et al., 2001). The similarity was assessed by Similarity Evaluation System for Chromatographic Fingerprints of TCM (2004A version, Chinese Pharmacopoeia Commission, China). All other statistical analyses were performed on GraphPad Prism 5 (GraphPad Software, CA, U.S.A.).

4.3 Results and discussion

4.3.1 Data pre-processing

4.3.1.1 Peak alignment

Several data pre-processing techniques were performed on the raw data matrix to reduce noise and inconsistency as well as to improve the data quality. Figure 4.1 shows the chromatograms of 46 methanolic extracts before and after peak alignment, baseline subtraction and concentration correction. The diagrams revealed that the retention time shifts within chromatograms were perfectly corrected after treatment. The reference chromatogram, segment length and slack number were optimised according to the automated selection approach proposed by Skov et al. (Skov et al., 2006). In this experiment, PTR1 was chosen as the reference chromatogram, whilst the segment length and slack number were determined as 49 and 2, respectively.

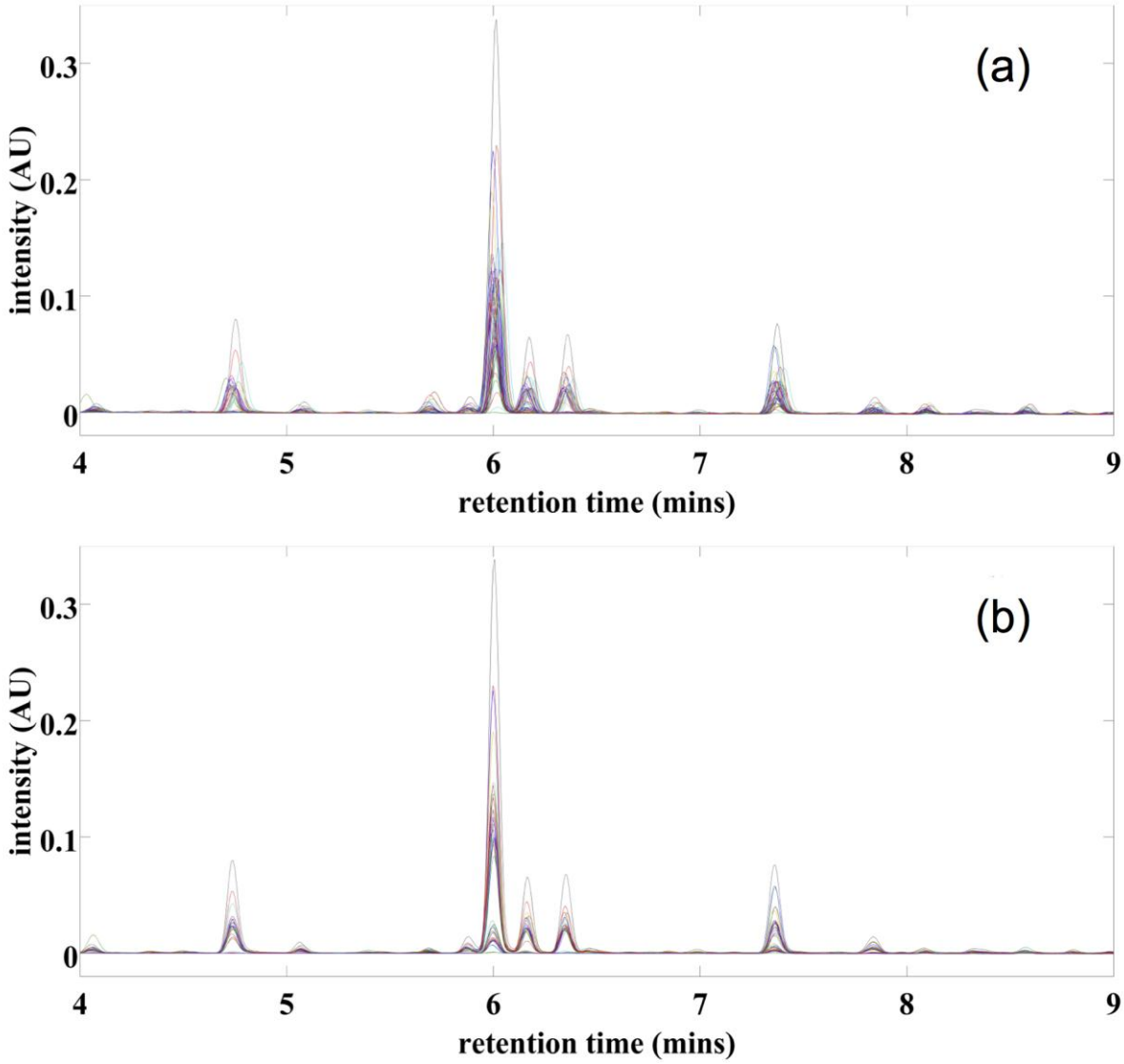


Figure 4.1 UPLC chromatograms of 23 PLR and 23 PTR methanolic extracts between retention time 4 and 9 min. (a) Raw data and (b) after peak alignment, baseline subtraction and concentration correction.

4.3.1.2 Outliers detection

Prior to constructing a multivariate calibration model, it was important to identify possible anomalous samples since such samples could affect the quality of the model and therefore, should be removed beforehand. To identify possible outliers, PCA was performed on the chromatographic data matrix, which was pre-processed with COW, baseline subtraction, concentration correction, SNV and column centering. The determination of outliers was assessed by Hotelling's T^2 versus Q residuals plot and score plot as shown in Figures 4.2a, 4.2c and 4.2e. Samples with high values of Hotelling's T^2 and Q residuals and located in an extreme position in a score plot were considered as outliers. After preliminary PCA analyses, four outliers (PLR9 and PTR5, PTR13 and PTR23) were found and removed for the subsequently analysis. The occurrence of such outliers may be due to variations in the plant source such as growth habitat, soil condition, climate and harvest season. Hence, the data set was reduced to 42 samples, consisting of 22 PLR and 20 PTR. Figure 4.2b, 4.2d and 4.2f shows the PC1-PC2 scores plots of the pre-processed data before and after the four outliers were eliminated. It can be observed that the outliers were located far from the major cluster (Figure 4.4b). In contrast, after removing the outliers, PLR and PTR were well separated into two distinct clusters (Figure 4.4f). The PCA model was constructed using two principal components, which explained 79.02% of the original information of the variables.

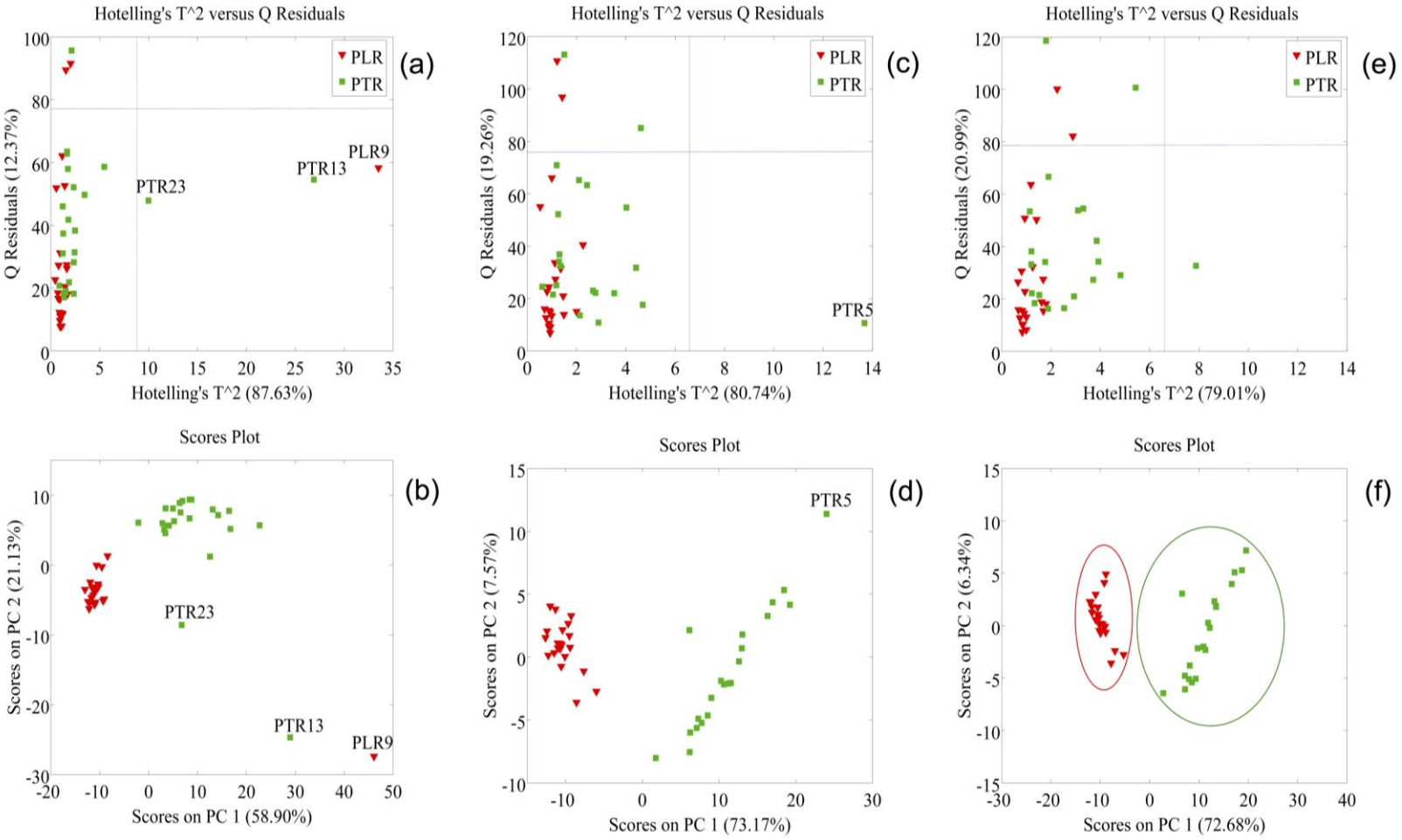


Figure 4.2 Pre-processed chromatographic data of the 46 samples. (a) Hotelling's T^2 versus Q residuals and (b) score plot with PLR9, PTR13 and PTR23 excluded (c) Hotelling's T^2 versus Q residuals and (d) score plot with PLR9, PTR5, PTR13 and PTR23 excluded (e) Hotelling's T^2 versus Q residuals and (f) score plot.

4.3.2 Variables selection algorithms

4.3.2.1 Variables selected by GA and SPA

Recent studies have demonstrated that variable selection in association with PLS can improve the model's predictability and reduce its complexity (Ebrahimi-Najafabadi et al., 2012; Silva et al., 2012). In this chapter, GA and SPA were compared. To avoid over-fitting and minimising the search domain in GA, the original 2401 variables from chromatograms were reduced in two steps. The initial 180 and the last 242 data points were eliminated due to lack of information. Variables from these regions were located before solvent front and during equilibrium time. The remaining 1980 data points were reduced to 66 intervals. Each interval consisted of 30 variables, which was equivalent to average peak width from the chromatograms. Auto-scaling was chosen as pre-treatment method before performing GA analysis. After 100 generations with 10 repeated cycles, 8 intervals (240 variables) were selected based on their fitness values and are illustrated in Figure 4.3. In contrast to GA, all the variables from the chromatographic data matrix were used when performing SPA. From the SPA results, 26 variables were selected and are shown in Figure 4.4.

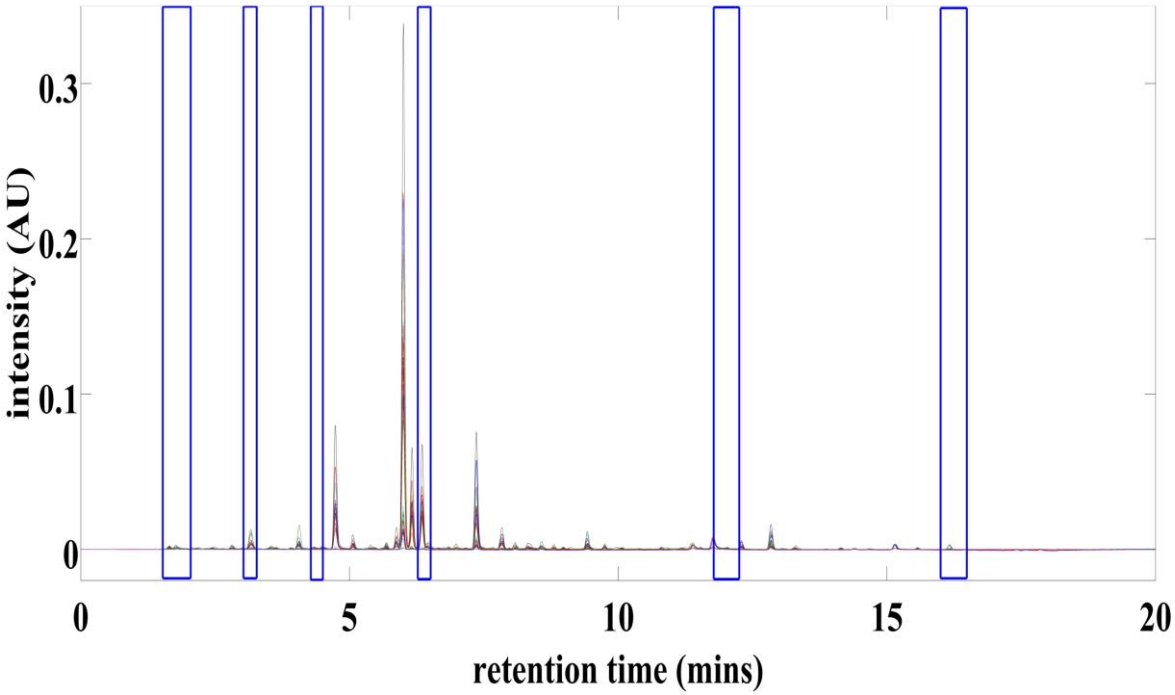


Figure 4.3 Pre-processed chromatograms of 22 PLR and 20 PTR samples with variables selected by genetic algorithm (GA).

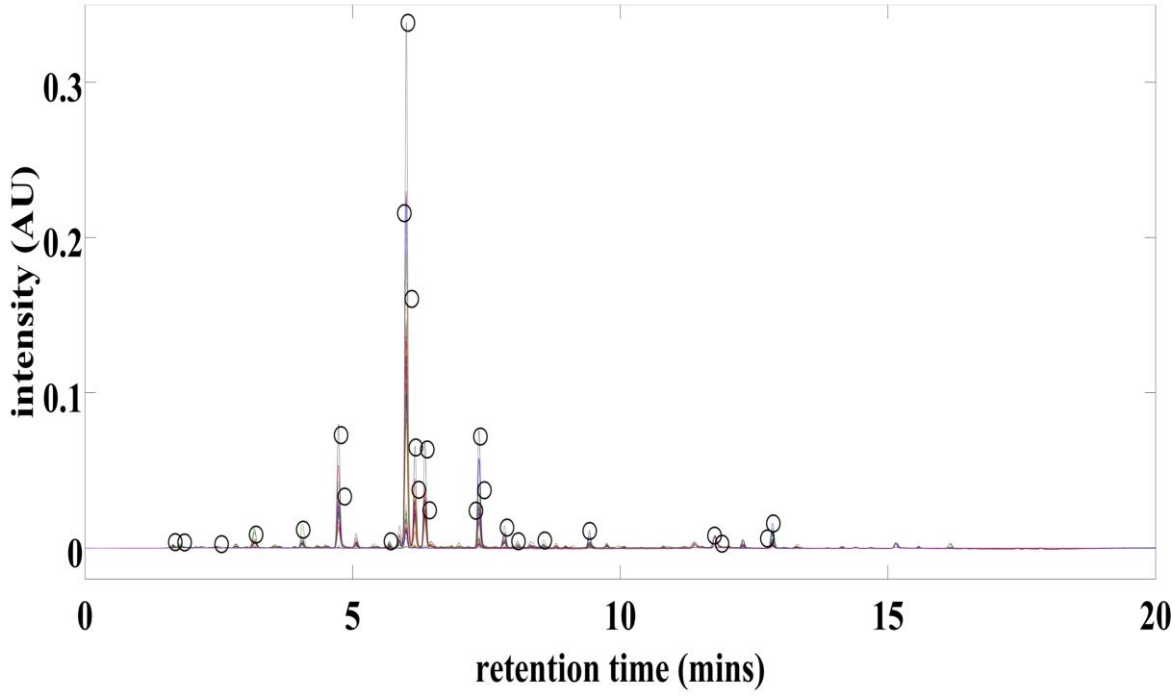


Figure 4.4 Pre-processed chromatograms of 22 PLR and 20 PTR samples with variables selected by successive projection algorithm (SPA).

4.3.2.2 Variables selected by PLS-DA X-loading plots

The pre-processed chromatograms of 22 PLR, 20 PTR, five reference standards and X-loading plots, which were generated from the PLS-DA model of the entire chromatographic matrix in the calibration set, are shown in Figure 4.5. The X-loading plot is useful for detecting important variable/peaks. Using this approach, Tistaert and his colleagues (Tistaert et al., 2011b) had successfully identified potential chemical markers for the authentication of Citri Reticulatae Pericarpium. The higher the loading score of a variable in the X-loading plot, the more likely it can differentiate PLR from PTR. By comparing the latent variable (LV) 1 and LV2 X-loading plots with the PLR and PTR chromatograms, four regions (155 variables) with large loading scores were selected. These regions consisted of six dominant peaks, in which three chemical constituents were identified, including puerarin at 6.01 min, daidzin at 7.37 min and daidzein at 12.87 min. The identities of these peaks were confirmed by comparing their retention times and MS spectra with the standard compounds. The UPLC-MS results from Chapter 3 (p.109) revealed that the three remaining peaks at 4.75, 6.17 and 6.36 min were 3'-hydroxypuerarin (Hirakura et al., 1997), 6"-O-D-xylosylpuerarin (Du et al., 2010) and 3'-methoxypuerarin (Hirakura et al., 1997), respectively.

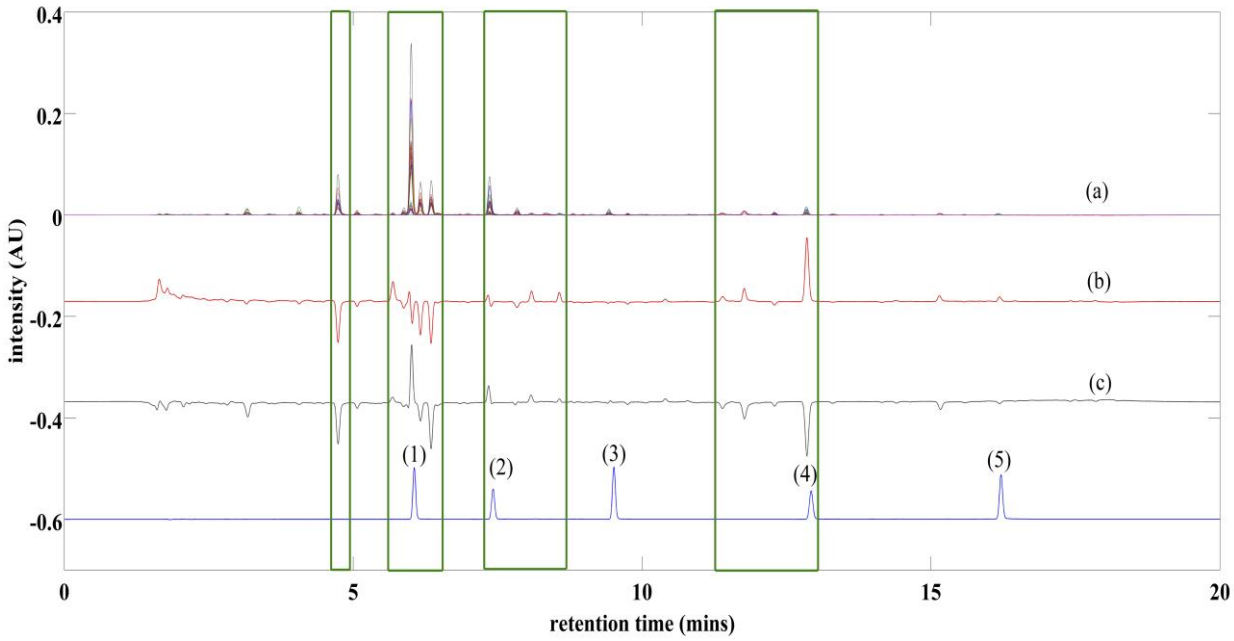


Figure 4.5 Pre-processed chromatograms of PLR and PTR with variables selected by X-loading plots. (a) 22 PLR and 22 PTR samples; (b) Latent variable (LV) 1 loading plot; (c) LV2 loading plot; Five references: (1) Puerarin; (2) Daidzin; (3) Genistin; (4) Daidzein; (5) Genistein.

4.3.2.3 Variables selected by puerarin

To evaluate a single chemical marker to authenticate PLR and PTR as proposed by the PPRC, the peak corresponding to puerarin (21 variables) was also adopted in the subsequent PLS-DA analysis. By comparing the ranges of variables selected, it was revealed that the variables selected by SPA and six selected peaks from X-loading plots were mainly located between 4.6 to 13 min, which corresponds to the majority of the compounds detected. On the other hand, variables selected by GA focused on the regions before 4.6 min and after 13 min, with only four intervals from GA overlapping with the regions selected by SPA and X-loading plots.

4.3.3 PLS-DA classification

4.3.3.1 Calibration set

PLS-DA discriminant models were constructed using the entire chromatographic matrix and variables selected from GA, SPA, X-loading plots and puerarin. A model was established from 28 samples (14 samples from each species), which was selected in accordance to the K-S algorithm, while the remaining fifteen samples (PLR: 9 samples; PTR: 6 samples) were considered as an external validation set. To avoid model over-fitting, the number of optimal LVs for each calibration model was determined by the leave-one-out cross-validation method. Statistical performances of the models were evaluated in terms of the root mean square error of calibration (RMSEC), root mean square error of cross-validation (RMSECV), root mean square error of validation (RMSEV) and correlation coefficient as shown in Table 4.2. In general, a good model should have low root mean square error vales, a high correlation coefficient and small differences between RMSECV and RMSEV.

In this study, the model using the entire chromatographic matrix as the reference was compared to the models established by the different approaches (Table 4.2). Comparing the entire chromatographic matrix and puerarin alone, the former model provided superior model's stability, simplicity and performance. In addition, the possibility of preserving the characteristics of the chromatographic fingerprints (multiple components), but at the same time, reducing the amount of information that needed to be analysed was examined. Herein, the model performances of GA, SPA and X-loadings plots were compared to that of the entire chromatographic matrix. The model established by GA exhibited the lowest root

mean square error values (RMSE: 0.0610; RMSECV: 0.0655; RMSEV: 0.0422) and the highest correlation coefficient (r^2 cal: 0.9851; r^2 CV: 0.9828; r^2 val: 0.9991). The models based on SPA and X-loading plots provided comparable root mean square error values and correlation coefficient in comparison to that of the entire chromatographic matrix. These results revealed that the use of variables selection could generate comparable or even superior results when compared to the use of the entire chromatographic fingerprint.

Table 4.2 Comparison of the statistical performance of various variable selection algorithms on the PLS-DA models.

Method	LV(s)	RMSEC	RMSECV	RMSEV	r^2 cal	r^2 CV	r^2 val
Entire matrix	2	0.1183	0.1452	0.0854	0.9440	0.9159	0.9938
GA	1	0.0610	0.0655	0.0422	0.9851	0.9828	0.9991
SPA	2	0.1298	0.1606	0.1593	0.9326	0.8973	0.9898
X-loading	3	0.1220	0.1700	0.1206	0.9405	0.8864	0.9812
Puerarin	3	0.2893	0.3725	0.2049	0.6653	0.4868	0.9158

LV: optimal number of latent variables used; RMSEC: root mean square error of calibration; RMSECV: root mean square error of leave-one-out cross-validation; RMSEV: root mean square error of validation; r^2 cal: correlation coefficient of the calibration set; r^2 CV: correlation coefficient of the cross-validation set; r^2 val: correlation coefficient of the validation set.

4.3.3.2 Validation set

To evaluate the performance of the established models on predicting species differences, discriminant analyses were performed using 14 samples (8 PLR and 6 PTR) from the validation set. Table 4.3 illustrates the species (class) classification results of the various PLS-DA models. For the models established by the entire

chromatographic matrix, GA, SPA and X-loading plots achieved a 100% correct prediction rate in the validation set. In contrast, the model using puerarin alone misidentified two samples with a correct prediction rate of 86.67%. Therefore, there is a nearly 14% of chance for incorrect authentication of unknown *Pueraria* samples when the puerarin model is selected. The results demonstrate that the model established with the use of puerarin as the single marker for quality control of PLR and PTR is unsatisfactory.

The variables selected by GA yielded the best classification model, with lower model complexity and root mean square error values as well as higher correlation coefficient as compared to the model based on the entire chromatographic matrix. However, they were located in the regions that contained relatively small peaks and generally could not be quantified. The aim of GA is to determine variables that can generate the best performance in the PLS-DA model, but these variables might not correspond to the compounds that can be detected in UPLC. Previous studies demonstrated that GA provided a superior performance in enhancing root mean square values and predictive ability for ascertaining the addition of vegetable oil in olive oil (Ruiz-Samblas et al., 2012) and barley in coffee (Ebrahimi-Najafabadi et al., 2012). In this study, GA was inadequate in identifying potential chemical makers for the authentication of PLR and PTR. In contrast, the model based on the six selected peaks from the X-loading plots provided similar correlation coefficient and root mean square error values in the calibration set and the same classification ability in the validation set as compared to the models using the entire chromatographic matrix. These results suggested that these six compounds are important factors determining the model's performance and

hence, can be used as potential chemical markers for authenticating PLR and discriminating it from its related species, PTR. Additionally, variables selected using SPA shared similar pattern and were overlapped with those from the X-loading plots, revealing that this approach could provide additional information to support the variables/peaks obtained from the X-loading plots.

Table 4.3 Confusion matrices of various variable selection algorithms from the validation set.

		Entire matrix		GA		SPA		X-loading		Puerarin	
True class		Predicted class		Predicted class		Predicted class		Predicted class		Predicted class	
	N	PLR	PTR	PLR	PTR	PLR	PTR	PLR	PTR	PLR	PTR
PLR	9	9	-	9	-	9	-	9	-	8	1
PTR	6	-	6	-	6	-	6	-	6	1	5

N: number of samples employed from each species in the validation set.

4.3.4 Classification of commercial granules

Table 4.4 illustrates the classification results of 17 PLR granules using the PLS-DA models established above. Prior to analysis, variables from the chromatographic matrix of granules were pre-processed and selected as aforementioned. All the granules tested in this study were labelled PLR as the raw material. According to the results, the herbal materials of four brands of granules, including two brands manufactured in Taiwan (G3 and G6) and two brands manufactured in China (G9 and G17) were classified as PTR, and the remaining 13 were classified as PLR, based on the models from the entire chromatographic matrix, SPA, GA and six selected peaks from the X-loading plots. Figure 4.6 illustrates the chromatograms obtained from 22 PLR, 20 PTR and the four alleged mislabelled granules. It can be observed that the chromatograms of the four granule products shared similar chromatographic characteristics with PTR, rather than PLR. In addition to the visual inspection, the chromatograms were evaluated and compared using similarity analysis. In Table 4.5, the similarity values of G3, G6, G9 and G17 were 0.977, 0.977, 0.970 and 0.981, respectively when compared to PTR (0.952 ± 0.038), and the similarity values of G3, G6, G9 and G17 were 0.929, 0.916, 0.901 and 0.924, respectively when compared to PLR (0.991 ± 0.007). The smaller differences between the similarity values of these four mislabeled granule products and the PTR extracts suggested that they are relatively similar to PTR rather than PLR. Based on this evidence, it can be concluded that these granules were manufactured using PTR as the raw herbal material and were mislabelled as PLR. The problem of substitution may relate to the confusing nomenclature and monographs for the two species as outlined in Chapter 1 and 2.

Table 4.4 Classification of commercial PLR granules using various PLS-DA models.

Samples	Entire matrix	GA	SPA	X-loading plots	Puerarin
G1	PLR	PLR	PLR	PLR	PLR
G2	PLR	PLR	PLR	PLR	PLR
G3	**PTR**	**PTR**	**PTR**	**PTR**	**PTR**
G4	PLR	PLR	PLR	PLR	PLR
G5	PLR	PLR	PLR	PLR	PLR
G6	**PTR**	**PTR**	**PTR**	**PTR**	**PTR**
G7	PLR	PLR	PLR	PLR	PLR
G8	PLR	PLR	PLR	PLR	PLR
G9	**PTR**	**PTR**	**PTR**	**PTR**	PLR
G10	PLR	PLR	PLR	PLR	PLR
G11	PLR	PLR	PLR	PLR	PLR
G12	PLR	PLR	PLR	PLR	PLR
G13	PLR	PLR	PLR	PLR	PLR
G14	PLR	PLR	PLR	PLR	PLR
G15	PLR	PLR	PLR	PLR	PLR
G16	PLR	PLR	PLR	PLR	PLR
G17	**PTR**	**PTR**	**PTR**	**PTR**	**PTR**

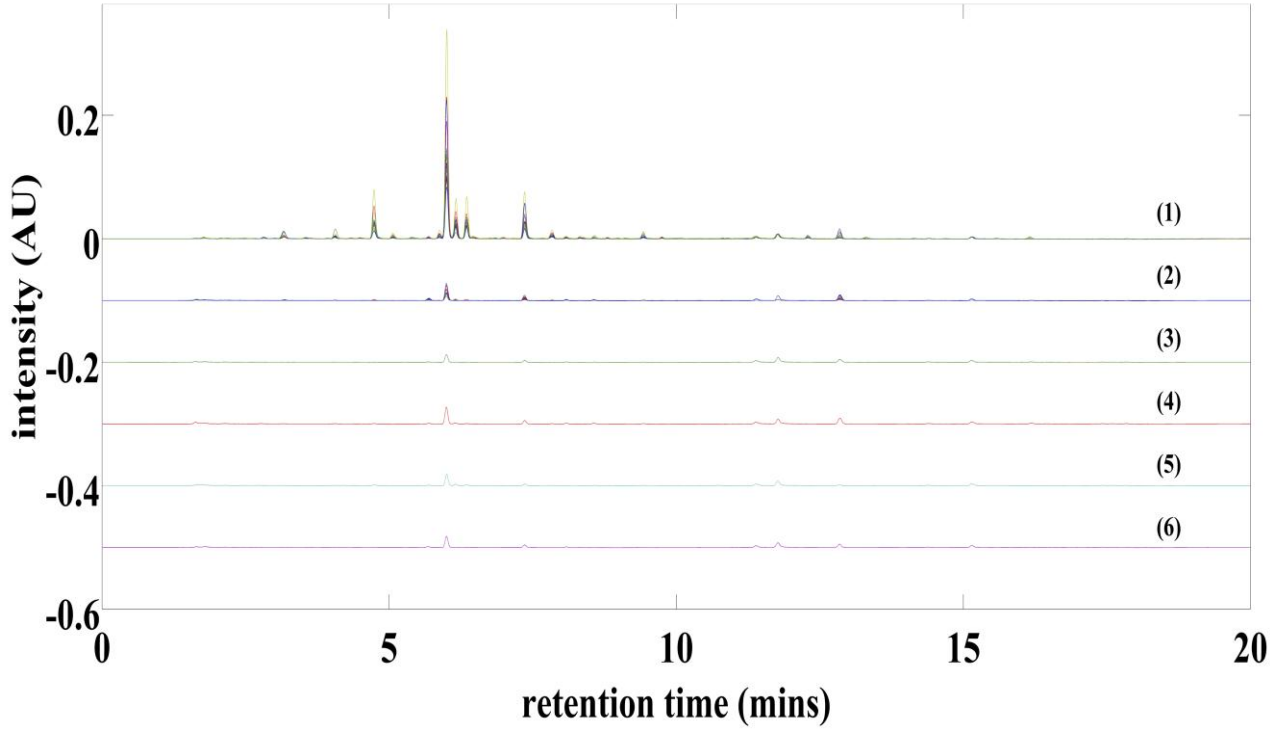


Figure 4.6 Pre-processed chromatograms of PLR and PTR samples and granules. (1) 22 aligned PLR samples; (2) 20 aligned PTR samples; (3) G3; (4) G6; (5) G9; (6) G17.

Table 4.5 Comparison of the similarity values between granules, PLR and PTR.

PLR	Similarity value*	Granules	Similarity value*	PTR	Similarity value*	Granules	Similarity value*
PLR1	0.98	**G1**	0.997	**PTR1**	0.959	**G1**	0.96
PLR2	0.975	**G2**	0.975	**PTR2**	0.969	**G2**	0.939
PLR3	0.992	**G3**	0.929	**PTR3**	0.97	**G3**	0.977
PLR4	0.998	**G4**	0.975	**PTR4**	0.959	**G4**	0.934
PLR5	0.995	**G5**	0.989	**PTR6**	0.878	**G5**	0.956
PLR6	0.998	**G6**	0.916	**PTR7**	0.978	**G6**	0.977
PLR7	0.988	**G7**	0.966	**PTR8**	0.864	**G7**	0.943
PLR8	0.98	**G8**	0.954	**PTR9**	0.926	**G8**	0.95
PLR10	0.994	**G9**	0.901	**PTR10**	0.968	**G9**	0.97
PLR11	0.991	**G10**	0.962	**PTR11**	0.983	**G10**	0.943
PLR12	0.994	**G11**	0.955	**PTR12**	0.951	**G11**	0.956
PLR13	0.99	**G12**	0.957	**PTR14**	0.947	**G12**	0.946
PLR14	0.997	**G13**	0.963	**PTR15**	0.962	**G13**	0.949
PLR15	0.997	**G14**	0.953	**PTR16**	0.911	**G14**	0.936
PLR16	0.99	**G15**	0.955	**PTR17**	0.996	**G15**	0.942
PLR17	0.997	**G16**	0.951	**PTR18**	0.978	**G16**	0.951
PLR18	0.997	**G17**	0.924	**PTR19**	0.982	**G17**	0.981
PLR19	0.998			**PTR20**	0.968		
PLR20	0.991			**PTR21**	0.991		
PLR21	0.991			**PTR22**	0.896		
PLR22	0.982						
PLR23	0.994						

***Similarity values of the granule samples were calculated by comparing the chromatographic characteristics with the mean chromatogram of PLR and PTR.**

In comparison, the results from the model using puerarin alone only partially agreed with other models. This model correctly identified three out of four brands of granules, which were manufactured using PTR as the raw herbal material. This result further confirmed that the use of puerarin alone was insufficient in authenticating PLR and its commercial products.

4.4 Conclusion

This study has demonstrated that UPLC PLS-DA, coupled with GA and SPA and X-loading plots, can be applied as a screening method for the quality control of PLR and PTR. The results showed that models developed from GA, SPA and six potential chemical markers could provide similar classification ability compared to the entire chromatographic matrix. Furthermore, the results suggest that puerarin alone is not sufficient to be used as the sole chemical marker and current regulations for the manufacturing, labelling and quality control of PLR products needs to be reassessed. In conclusion, this study indicates that PLS-DA is an effective technique for the quality control of PLR and PTR and other herbal materials used in the manufacture of granules and other dosage forms in the pharmaceutical industry.

Chapter Five
Differentiation of Puerariae Lobatae Radix and Puerariae Thomsonii Radix using high-performance thin-layer chromatography coupled with multivariate classification analyses

The materials presented in this chapter were extracted from **Wong, K. H.,** Razmovski-Naumovski, V., Li, G. Q., Li, K. M. & Chan, K. (2014) Differentiating Puerariae Lobatae Radix and Puerariae Thomsonii Radix using HPTLC coupled with multivariate classification analyses. *J Pharm Biomed Anal.,* 95 11-19.

5.1 Introduction

Thin-layer chromatography (TLC) is regarded as one of the fundamental authentication methods in various pharmacopoeias and monographs. Recent advances in TLC, especially the development of high performance thin-layer chromatography (HPTLC), have made it an attractive alternative to the more costly and sophisticated analytical methods used for the quality control of herbal medicines (Fried and Sherma, 1996, 1999; Spangenberg et al., 2011).

HPTLC offers several advantages over conventional analytical methods such as HPLC including reliability, simplicity, flexibility, fast analytical time, low running cost, low organic solvent consumption and the analyses of numerous samples simultaneously (Marston, 2011). In terms of PLR and PTR, there is only one study that employs HPTLC to investigate the chemical differences between PLR and PTR (Chen et al., 2006). In this study, two different solvent systems detecting the glycoside and aglycone fraction were developed and validated. The results revealed that the total chemical content in PLR was 21.58 times greater than in PTR. However, this study focused on the quantification and identification of a few chemical components such as puerarin, daidzein and genistein. From the results in Chapter 4, it has been shown that the use of a single or few chemical markers is insufficient to differentiate PLR from PTR. In addition, the HPTLC results were not compared to the other chromatographical methods.

HPTLC is based on the visualisation of the plate and comparison to a reference sample, which is rather subjective, and the results may vary between different investigators. Furthermore, it would be a difficult task to assess the differences

with larger sample sizes, and rather cumbersome to adopt as a daily routine procedure in industry. To solve this problem, multivariate statistical analysis named pattern recognition classification methods can be used. Pattern recognition classification algorithms, including both unsupervised and supervised, determine the class of samples based on the data matrix. This approach provides a systemic and objective way of analysing the HPTLC plate. One of the major advantages of adopting multivariate analysis for chromatographic fingerprints is that it scrutinises the subtle differences within the chromatogram. It does not rely on the comparison to a reference sample or compound, or quantifying a particular chemical compound(s), which can be difficult to discern for an unknown sample (Alaerts et al., 2010; Tistaert et al., 2011a). In addition, these classification algorithms have been applied to classify different descriptors in the field of food chemistry (Munck et al., 1998), forensic chemistry (Thanasoulias et al., 2003), environment chemistry (Mas et al., 2010), bioinformatics (Yetukuri et al., 2007), pharmaceutics (Roggo et al., 2007), metabonomics (Rousseau et al., 2008), proteomics (Lee et al., 2003), genomics (Eriksson et al., 2004) and so on. Recently, three comprehensive review articles (Alaerts et al., 2010; Gad et al., 2013; Tistaert et al., 2011a) were published which summarised the application of various classification algorithms for differentiating medicinal herbs of different species, age and geographical origins.

In literature, there are limited studies using TLC or HPTLC coupled with chemometrics for the quality control of herbal medicine (Gad et al., 2013; Tistaert et al., 2011a). This is mainly due to the difficulty of converting digitalised HPTLC plate into chromatographic fingerprints. For the first time, simple built-in functions

from the Matlab were combined with a series of pre-processing algorithms to generate HPTLC chromatographic fingerprints in this study.

Therefore, the aim of this study was to compare the classification accuracy of HPTLC multivariate analysis to UPLC in differentiating PLR from its closely related species PTR using a digitised HPTLC plate. This combined technique could provide a rapid and low-cost quality control method for the differentiation of species in the production process.

5.2 Methods and materials

5.2.1 Chemicals and solvents

Analytical grade methanol, chloroform, ethyl acetate, absolute ethanol (99.5% w/w), glacial acetic acid and HPLC-grade acetonitrile were purchased from Thermo Fisher Scientific (VIC, Australia). Sulfuric acid was obtained from Sigma-Aldrich (NSW, Australia). The details of standard compounds were described in Chapter 3 (p.94).

5.2.2 Herbal samples and extraction

The details of herbal samples and extraction were described in Chapter 2 (p.75) and Chapter 3 (p.95), respectively.

5.2.3 HPTLC measurement

The HPTLC measurement was performed as previously described with some modifications (Jiang et al., 2005). Briefly, the ethanolic extract was reconstituted in

absolute ethanol (PLR: 5 mg/mL; PTR: 5 mg/mL) and filtered through a 0.22 µm PVDF syringe filter. The filtered ethanolic extract was applied onto a 10 cm (width) x 10 cm (height) pre-coated silica gel 60 plate (Catalogue number: 105547; Merck, Darmstadt, Germany) using a CAMAG Linomat IV semi-automatic TLC sampler (Muttenz, Switzerland) equipped with WinCATS 1.2.3 software. Both 10 and 20 cm plates are recommended for HPTLC (CAMAG, 2014), however, a 10 cm width plate was chosen in this study as it was easier to handle and presented a more even coating when spraying reagent was applied. Nitrogen was used as a carrier gas. Application volume was set at 5 µL, with a bandwidth of 6 mm and an application speed of 125 nL/s, 10 mm from the edges, 4 mm apart and 15 mm from the bottom. The mobile phase (8 mL) consisting of chloroform-ethyl acetate-methanol-water (20:40:22:10 v/v) was added to a CAMAG twin-trough chamber. The chamber was saturated with the mobile phase vapour for 30 mins at ambient temperature and humidity. The pre-loaded plate was developed vertically from the lower edge to 8 cm. After development, the plate was air-dried for 5 mins and was sprayed with 10% sulfuric acid in methanol. The plate was then incubated in an oven at 105 ^{o}C for 5 mins until the colour of the polyphenols on the plate appeared (Chen et al., 2006). The plate was observed under UV light at 366 nm in a CAMAG TLC visualiser and was documented using a digital camera.

5.2.4 HPTLC data matrix pre-processing

Figure 5.1 summarises the procedures involved in the digitalisation and transformation of HTPLC plate images into chromatographic fingerprint matrices. The original HPTLC fluorescent images were RGB triplet matrices composed of red, green and blue colour channels. To reduce the dimensions of the data, the

images were converted to 8-bit monochromatic grayscale using Matlab R2012b (The MathWorks, MA, U.S.A.). The number of pixel per sample was set as 1450 pixels. The pixel intensity of each sample was measured using the “improfile” function from Matlab. The measurement was performed in triplicate and the mean of three measurements was used as the final pixel intensity. The mean pixel intensity of each sample was combined to form a chromatographic data matrix consisting of 42 rows (number of samples) and 1450 columns (number of pixels per sample). The matrix was smoothed using Savitzky-Golay filter with a filter width of 15 and a zero order polynomial (Bromba and Ziegler, 1981). The baseline offset was removed by automatic Whittaker algorithm at 100 λ (Eilers, 2003). Correlation optimized warping (COW) (Chapter 1 p.38) was employed to correct the inter- and intra-plate peak shift due to variations in mobile phase composition, humidity, temperature, operator handling and instrumental instability. The COW parameters were optimised using a systemic algorithm proposed by Skov et al. (Skov et al., 2006).

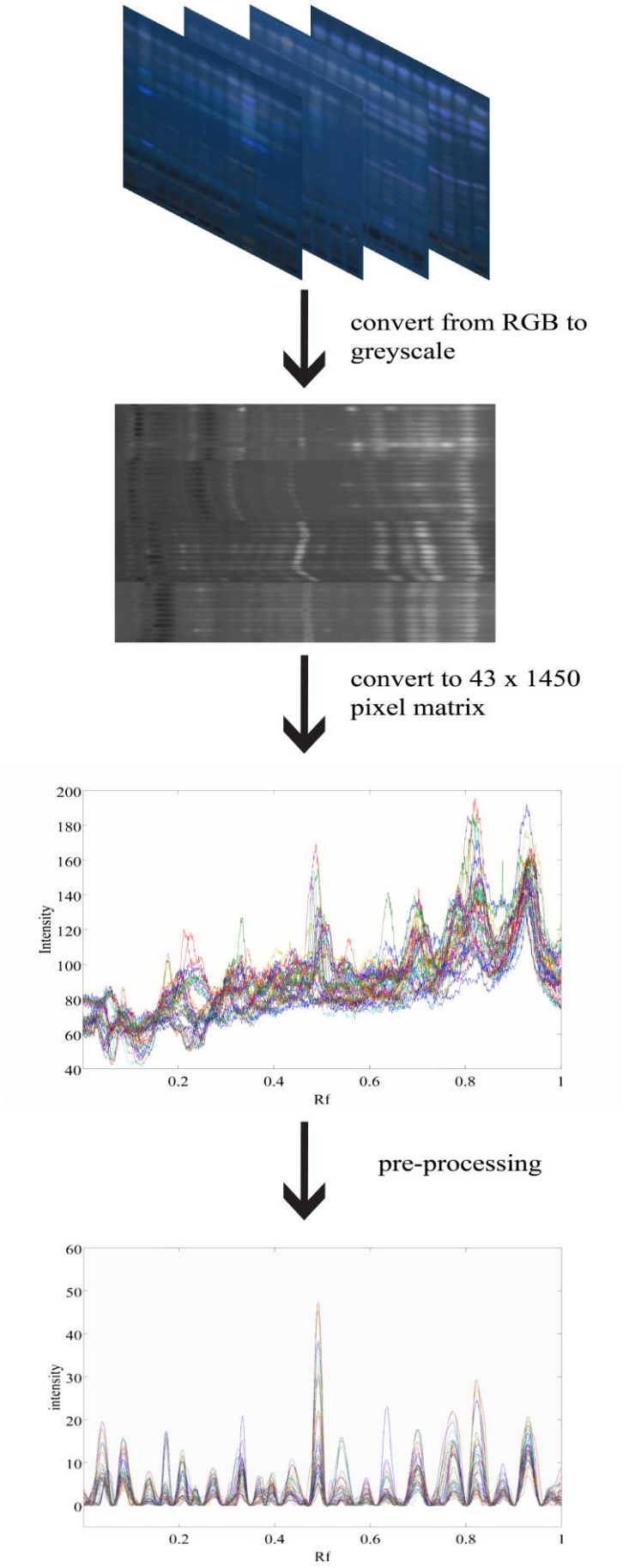


Figure 5.1 Conversion and pre-processing of HPTLC results prior to the chemometric analyses.

(a) Digitalised HPTLC RGB colour space images converted to greyscale.

(b) Data matrix consisting of 42 samples (22 PLR and 20 PTR), in which each sample has 1450 variables.

(c) Raw data matrix subjected to several pre-processing algorithms such as Savitzky-Golay smoothing, baseline removal and correlation optimised warping.

The chromatographic data matrix was separated into calibration and validation set using K-S random selection algorithm to minimise the selection bias (Chapter 1 p.56) (Kennard and Stone, 1969; Wu et al., 1996). Hence, the data was separated into two matrices: a calibration matrix containing 14 PLR and 14 PTR samples (28 in total), and a validation matrix consisting of 8 PLR and 6 PTR samples (14 in total). Additionally, several pre-processing techniques such as normalisation (Chapter 1 p.40), standard normal variate (SNV) (Chapter 1 p.39) and mean column centering (Chapter 1 p.40) were applied to the data matrices to enhance the signal-to-noise ratio and the interpretability of the classification models. The influences of various pre-processing methods on the partial least square-discriminant analysis (PLS-DA) models were evaluated and compared. The technique with the best species classification ability was selected for subsequent analyses.

5.2.5 UPLC measurement and data pre-processing

The UPLC measurement was performed as described in Chapter 3 (p.98) and the data was pre-processed as previously described in Chapter 4 (p.124).

5.2.6 Multivariate classification analyses

The pre-processed HPTLC and UPLC chromatographic fingerprints were subjected to seven multivariate classification algorithms, namely K-nearest neighbors (KNN; Chapter 1 p.51), classification and regression tree (CART; Chapter 1 p.53), successive projection algorithm-linear discriminant analysis (SPA-LDA; Chapter 1 p.47), principal component analysis-discriminant analysis (PCA-DA; Chapter 1 p.43), soft independent modelling of class analogy (SIMCA;

Chapter 1 p.54), PLS-DA (Chapter 1 p.49) and support vector machine-discriminant analysis (SVM-DA; Chapter 1 p.55). The calibration set was used to build the classification model, whereas the validation set was used to assess the accuracy of the established model. The models were compared using confusion matrix (Chapter 1 p.58).

5.2.7 Software

Data pre-processing and classification modelling were performed on PLS toolbox (Eigenvector Research Incorporated, WA, U.S.A.) and classification toolbox version 2.0 (Ballabio and Consonni, 2013) under Matlab environment. The details of SPA toolbox 1.0, K-S and COW algorithm m-files were described in Chapter 4 (p.124).

5.3 Results and discussion

5.3.1 Detection of outliers

PCA detected the presence of possible outliers and provided a general idea of the samples' distribution, as well as the relationship between classes. Figure 5.2 illustrates the Hotelling's T^2 versus Q residues plot and PCA scores plot obtained from the HPTLC chromatographic data matrix, which has been pre-processed with Savitzky-Golay smoothing, baseline removal and COW. Hotelling's T^2 reflects the variation in each sample within the model, whereas Q residuals indicate the difference between the original and projected data. A sample is considered as an outlier when it has high Hotelling's T^2 and Q residuals value and located at an extreme position in the scores plot. Such an outlier should be avoided and

eliminated before further multivariate analysis as it could dramatically affect the model's quality. Fortunately, no such outlier was observed from the data (Figure 5.2a) and hence all the samples obtained from the HPTLC analyses were used in the subsequent classification. As shown in Figure 5.2b, PLR and PTR were separated into two clusters, suggesting that these two species had distinct chromatographic characteristics.

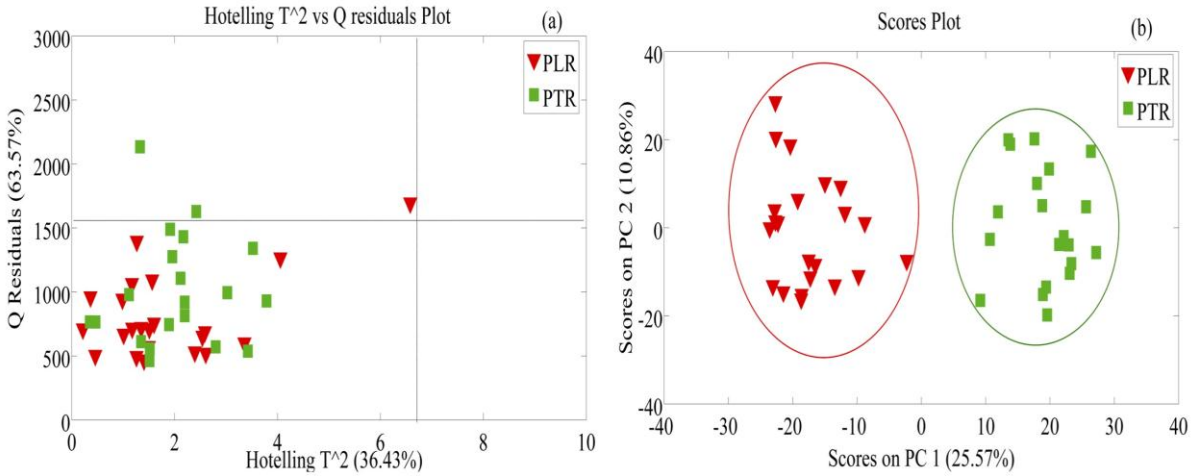


Figure 5.2 PCA plots obtained from pre-processed HPTLC chromatographic data matrix. (a) Hotelling's T^2 versus Q residuals plot, (----) line represents 95% confidence interval; (b) PC1-PC2 scores plot.

5.3.2 Optimisation of pre-processing methods

It is well known that the application of pre-processing algorithms prior to multivariate analysis can significantly improve the data quality and enhance the interpretability of the classifications models. However, there is no standard protocol on which pre-processing algorithms should be used, as each data set and analytical method is unique. As a result, it is necessary to determine the optimal pre-processing techniques before establishing a model.

PLS-DA compared and evaluated the effect of different pre-processing methods and their respective order on the model's performance. After the application of preliminary pre-processing methods such as smoothing, baseline removal and peak alignment, the data matrix was subjected to column centering, normalisation and SNV. The impact of various pre-processing methods on the PLS-DA model's performance was determined according to the root mean square of calibration (RMSEC), root mean square of cross validation (RMSECV), root mean square of validation (RMSEV) and correlation coefficient from leave-one-out cross-validation (LOOCV) as shown in Table 5.1. It was observed that all the pre-processing techniques were able to increase the correlation coefficient. However, the PLS-DA model pre-processed with SNV followed by column centering presented the highest correlation coefficient, the lowest complexity and the smallest root mean square error deviation between the calibration and validation set. It was, therefore, chosen as the optimal pre-processing method and was applied to the both HPTLC and UPLC chromatographic data matrices in the subsequent analysis.

5.3.3 SPA variable selection

The number of variables employed in establishing a LDA classification model needed to be smaller than the number of samples. Thus, SPA reduced the number of variables. The optimal variables number was evaluated by LOOCV and was determined as 25 and 23 for the HPTLC and UPLC matrix, respectively. Figure 5.3 and 5.4 shows the distribution of selected variables for the chromatographic fingerprints. For the HPTLC chromatograms, the selected variables were mainly located before Rf 0.3 and after Rf 0.7, whilst for UPLC, the variables were evenly distributed. Additionally, the peak at Rf 0.32 representing puerarin was the

predominant peak in the HPTLC chromatograms at 254 nm but not at 366 nm, suggesting that the intensity of puerarin diminished after derivatisation. Instead, an unknown peak at Rf 0.5 generated the highest intensity, which was absent at 254 nm. It is important to note that the intensity of a nominated peak may no longer be proportional to its respective concentration after derivatisation and this should be taken into account when quantifying chemical constituents in future studies.

Table 5.1 Comparison of the statistical performance of various pre-processing methods on the HPTLC calibration and validation set.

Pre-processing method(s)	LV(s)	RMSEC	RMSECV	RMSEV	Deviation between RMSEV and RMSEC (%)	r^2 cal	r^2 CV	r^2 val
None	2	0.2486	0.2659	0.3094	19.63	0.7547	0.7198	0.7654
Normalisation + mean centering	2	0.5119	0.5097	0.5333	4.01	0.9517	0.9190	0.9391
SNV + mean centering	2	0.5110	0.5109	0.5135	0.50	0.9556	0.9279	0.9401
Mean centering + normalisation	2	0.4805	0.4969	0.5094	5.68	0.9309	0.8914	0.9082
Mean centering + SNV	3	0.4564	0.4666	0.3908	-16.78	0.9137	0.8876	0.8749

LV: optimal number of latent variables; RMSEC: root mean square of calibration; RMSECV: root mean square of leave-one-out cross-validation; RMSEV: root mean square of validation; r^2 cal: correlation coefficient of calibration set; r^2 CV: correlation coefficient of leave-one-out cross-validation set; r^2 val: correlation coefficient of validation set. Deviation is calculated as [RMSEV – RMSEC)/RMSEV] × 100%.

5.3.4 Comparison of various classification models

The parameters from each classification model were optimised using LOOCV based on the calibration set and are summarised in Table 5.2. To evaluate the performance of the classification models, both the pre-processed HPTLC and UPLC chromatographic matrices were randomly separated into calibration (28 samples with 14 PLR and 14 PTR) and validation (14 samples with 8 PLR and 6 PTR) set using K-S algorithm. The calibration set was used to establish and train the classification model, whilst the validation set was employed in the final stage to evaluate the predictive ability of the calibrated model.

Table 5.2 Optimisation of parameters in various classification models

	HPTLC	UPLC
KNN		
Optimal K value	3	6
CART	none	none
SPA-LDA		
Optimal number of variables	25	23
PCA-DA		
Optimal PCs	3	2
SIMCA		
Optimal PCs for PLR	3	6
Optimal PCs for PTR	5	2
PLS-DA		
Optimal LVs	2	2
SVM-DA		
X-block compression	none	none
Probability estimation	zero	zero

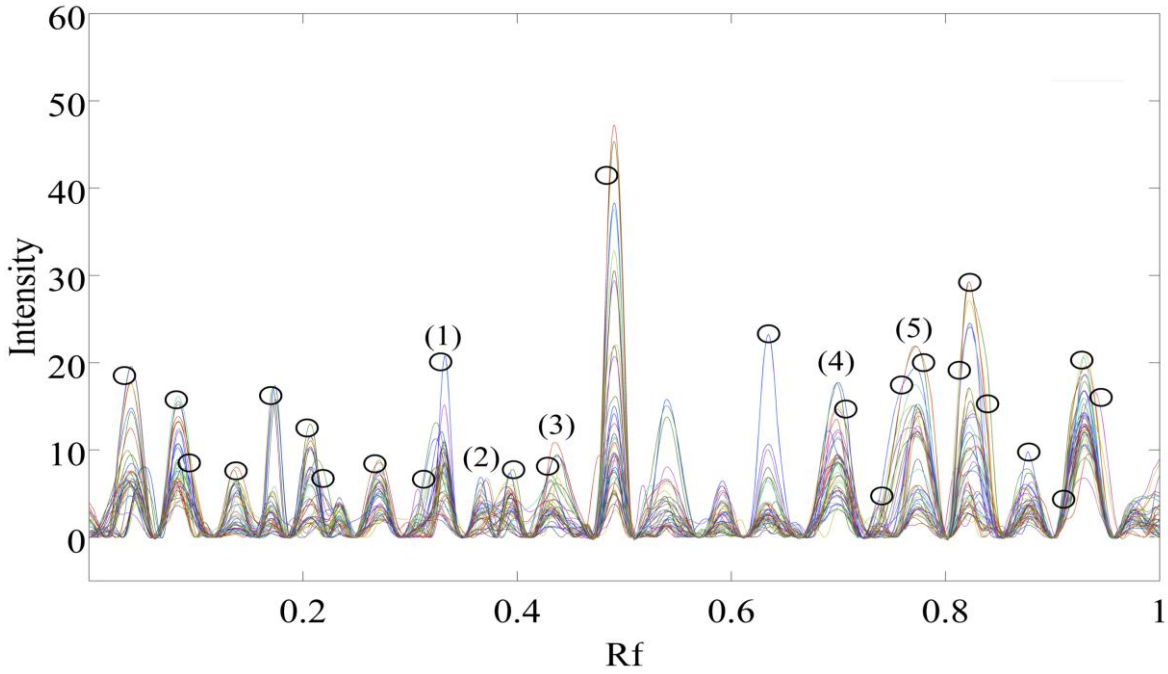


Figure 5.3 Comparison of the pre-processed HPTLC chromatograms (22 PLR and 20 PTR) with variables selected by successive projection algorithm. (1) Puerarin; (2) Daidzin; (3) Genistin; (4) Daidzein; (5) Genistein.

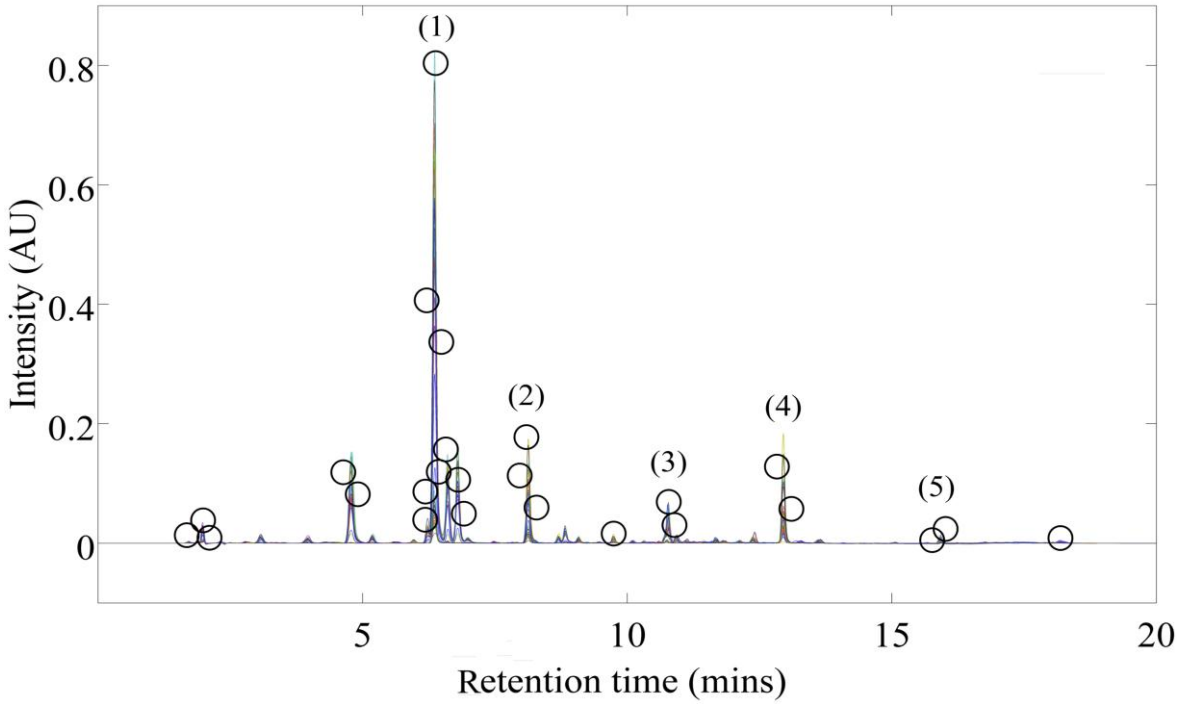


Figure 5.4 Comparison of the pre-processed UPLC chromatograms (22 PLR and 20 PTR) with variables selected by successive projection algorithm. (1) Puerarin; (2) Daidzin; (3) Genistin; (4) Daidzein; (5) Genistein.

The confusion matrices of the different classification techniques obtained from the pre-processed HPTLC and UPLC chromatographic fingerprints are demonstrated in Table 5.3 and 5.4, respectively. The LOOCV accuracy illustrated the interpretability of a model, whereas the model's predictive ability was represented by the accuracy of the validation set (Val). For the HPTLC data (Table 5.3), the classification models established by PCA-DA, PLS-DA and SVM-DA provided the greatest interpretability (100.00%) and predictability (100.00%) for both LOOCV and validation set, while the models from SPA-LDA (CV: 100.00%; Val: 92.86%), KNN (CV: 96.43%; Val: 100%) and CART (CV: 78.57%; Val: 64.29%) generated a moderate interpretability and predictability, with misidentification in either the LOOCV or validation set. In contrast, the SIMCA model delivered the worst model performance (CV: 80.00%; Val: 60.00%), with 13 (2 PLR and 11 PTR) and 4 (2 PLR and 2PTR) samples not assigned to the LOOCV and validation set, respectively.

It was observed that the models obtained from the UPLC data provided a relatively higher model's classification ability compared to the HPTLC data. Table 5.4 showed that KNN, PCA-DA, PLS-DA and SVM-DA classification models were given the highest interpretability (100.00%) and predictability (100.00%) in both LOOCV and validation set. Additionally, the models from SPA-LDA (CV: 96.43%; Val: 100%) and CART (CV: 96.43%; Val: 100%) generated a slightly lower interpretative ability, and misclassified a PTR sample as PLR. Similar to the results obtained from the HPTLC data, the model established using SIMCA had the least model performance (CV: 62.50%; Val: 75.00%), with 12 (8 PLR and 4 PTR) and

Table 5.3 Confusion matrices obtained from the validation set of the pre-processed HPTLC data.

		CV					Val				
		Predicted class					Predicted class				
	True class	N	PLR	PTR	NA	Accuracy (%)	N	PLR	PTR	NA	Accuracy (%)
SPA-LDA	PLR	14	14	−	−	100.00	8	7	1	−	92.86
	PTR	14	−	14	−		6	−	6	−	
CART	PLR	14	12	2	−	78.57	8	3	5	−	64.29
	PTR	14	4	10	−		6	−	6	−	
KNN	PLR	14	13	1	−	96.43	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	
SIMCA	PLR	14	12	−	2	80.00	8	6	−	2	60.00
	PTR	14	3	−	11		6	4	−	2	
PCA-DA	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	
PLS-DA	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	
SVM-DA	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	

CV: LOOCV set; Val: validation set; N: number of samples; NA: not assigned.

Table 5.4 Confusion matrices obtained from the validation set of the pre-processed UPLC data.

		CV					Val				
		Predicted class					Predicted class				
	True class	N	PLR	PTR	NA	Accuracy (%)	N	PLR	PTR	NA	Accuracy (%)
SPA-LDA	PLR	14	14	−	−	96.43	8	8	−	−	100.00
	PTR	14	1	13	−		6	−	6	−	
CART	PLR	14	14	−	−	96.43	8	8	−	−	100.00
	PTR	14	1	13	−		6	−	6	−	
KNN	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	
SIMCA	PLR	14	−	6	8	62.50	8	−	2	6	75.00
	PTR	14	−	10	4		6	−	6	−	
PCA-DA	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	
PLS-DA	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	
SVM-DA	PLR	14	14	−	−	100.00	8	8	−	−	100.00
	PTR	14	−	14	−		6	−	6	−	

CV: LOOCV set; Val: validation set; N: number of samples; NA: not assigned.

6 PLR samples not assigned to the LOOCV and validation set, respectively.

To further investigate the classification ability of the various classification algorithms, the error rate (ER), non-error rate (NER), specificity and sensitivity of each class were compared and evaluated and are shown in Table 5.5. The sensitivity was defined as the model's ability to correctly classify samples belonging to a specific class, while the ability to reject the samples of all other classes was expressed as specificity. Since PLR and PTR were the only two classes in this study, the specificity and sensitivity of the two classes were symmetrical. In other words, the specificity of PLR will be equivalent to the sensitivity of PTR, and vice versa. Therefore, only the specificity and sensitivity of PLR was described herein.

As shown in Table 5.5, KNN, PCA-DA, PLS-DA and SVM-DA for the HPTLC data gave a perfect score of one in both specificity and sensitivity, suggesting that the models were able to correctly classify all the samples with zero error rate in both classes. This minimal deviation between specificity and sensitivity indicated that there was no particular trend displayed by the models in recognising either PLR or PTR. However, the model established using CART provided a relatively low specificity value (0.55). This indicated that the model was more favourable in discriminating PTR rather than PLR. Furthermore, the small ER (0.23) and high NER (0.77) suggested that there was a higher chance of misidentifying PLR as PTR, rather than incorrectly identifying PTR as PLR. In agreement with the results obtained from Table 5.4, the SIMCA model was unable to classify any of the PTR samples in the validation set (ER: 0.7; NER: 0.3; specificity: 0; sensitivity: 0.6). For

the UPLC data, all the classification models, except SIMCA, were able to correctly identify samples from both classes in the validation set with a high specificity (1), a high sensitivity (1) and zero error rate. In comparison to the SIMCA model from HPTLC, the model from UPLC had zero sensitivity towards PLR, suggesting that it was unable to classify any of the PLR samples in the validation set (ER: 0.63; NER: 0.38; specificity: 0.75; sensitivity: 0). In terms of the preference of classification model, KNN and CART were the most favourable algorithms for differentiating of PLR and PTR as they required minimal data handling procedures and parameter optimisation and short calculation time. Nonetheless, if the focus of a study is to investigate the distribution of classes and the relationship between independent and response variables, PCA-DA, PLS-DA and SVM-DA are preferred, as these techniques provided detailed information of the data matrices from the score and loading plot.

Table 5.5 Classification parameters of the pre-processed HPTLC and UPLC data obtained from the validation set.

	HPTLC				UPLC			
			PLR				PLR	
	ER	NER	Specificity	Sensitivity	ER	NER	Specificity	Sensitivity
SPA-LDA	0.07	0.93	1	0.86	0	1	1	1
CART	0.23	0.77	0.55	1	0	1	1	1
KNN	0	1	1	1	0	1	1	1
SIMCA	0.7	0.3	0	0.6	0.63	0.38	0.75	0
PLS-DA	0	1	1	1	0	1	1	1
PCA-DA	0	1	1	1	0	1	1	1
SVM-DA	0	1	1	1	0	1	1	1

ER: error rate; NER: non-error rate.

Similar results were obtained from a previous study which compared the efficacy of seven classification methods on the quality control and identification of motor oil adulteration using NIR spectroscopy. The results demonstrated that the classification error rates of the models established from SVM-DA and KNN were significantly smaller than that from SIMCA and PLS-DA (Balabin et al., 2011). In another study (Martins et al., 2011), PLS-DA and KNN gave a 100% correct prediction rate in classifying six different *Phyllanthus* species using HPLC, while SIMCA had a 12% correct prediction rate with more than 80% of the samples not assigned. On the contrary, SIMCA and SVM-DA have a 100% correct classification rate in identifying rice seed samples from four different cultivars using NIR, whereas only 80% of the samples were correctly classified using PLS-DA and KNN model. It can be revealed that a classification method with high classification ability in one data set does not mean that it will achieve a similar result in another data set or analytical method as each data set is unique.

In conjunction with the results from Table 5.3 to 5.5, the data matrices obtained from UPLC generated a better classification performance in various classification models. More importantly, the results showed that the models established from the HPTLC data matrices provided comparable classification abilities to the UPLC data matrices. These results suggested that HPTLC can be applied as an alternative analytical method in the differentiation of PLR and PTR.

5.4 Conclusion

In this study, the classification accuracy of HPTLC coupled with seven multivariate classification methods was evaluated and compared to that of UPLC. The application of Matlab enabled information to be rapidly extracted from a digitalised HPTLC plate. The results indicate that chromatographic fingerprints obtained from HPTLC provide comparable classification ability to UPLC. In particular, KNN, PLS-DA, PCA-DA and SVM-DA from HPTLC achieved a 100% correct species classification rate, whilst SIMCA gave the lowest correct classification rate (60%). HPTLC coupled with multivariate classification methods provides a simple and user-friendly approach for the quality control and authentication of PLR and PTR. The application of this technique can be adopted as a template for authenticating other herbal materials.

Chapter Six

Prediction of total phenolic content and antioxidant capacities of Puerariae Lobatae Radix and Puerariae Thomsonii Radix using Raman spectroscopy coupled with partial least squares regression analysis

6.1 Introduction

PLR and PTR is an abundant source of polyphenols such as puerarin, daidzin and genistin, which have been found to exert anti-diabetic, anti-apoptotic, anti-hypertensive, anti-inflammatory and vasodilatory properties in various *in vitro* and *in vivo* studies (Wong et al., 2011). Recent studies have suggested that the beneficial pharmacological effects of polyphenols may be related to their high antioxidant capacity (Jin et al., 2012; Rice-evans et al., 1995). The prevention of oxidative stress damage is particularly important for the treatment of cardiovascular and cerebrovascular diseases (Procházková et al., 2011). Therefore, the determination of polyphenol content and antioxidant capacity in PLR and PTR is necessary to provide information on their potential beneficial effects and can be applied as one of the criteria for the differentiation and thus, the quality control of these two herbs.

Currently, there are two major conventional methods for the quantification of total phenolic content (TPC) within a herb. The first approach is to use Folin-Ciocalteu colorimetric method (Ainsworth and Gillespie, 2007). The second technique involves the use of sophisticated analytical instruments such as high-performance liquid chromatography (HPLC) or thin-layer chromatography to quantify individual chemical component (Gómez-Caravaca et al., 2013). To determine antioxidant capacity, colorimetric methods such as 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) (Re et al., 1999) and cupric reducing antioxidant capacity (CUPRAC) assays (Apak et al., 2004) are employed. However, these time-consuming techniques involve tedious extraction procedures, large organic solvent consumption and establishing calibration curve(s), which is impractical if numerous

samples are required to be analysed. As a result, the development of an alternative method for high throughput screening of herbal material is necessary.

To address this issue, spectroscopic fingerprint coupled with multivariate regression analysis has been proposed. The spectroscopic fingerprint of a herbal sample refers to a unique pattern resulting from the combination of spectral characteristics of all chemical components within a herb. As compared to conventional chromatographic analyses, the major advantage of adopting spectroscopy is that the analyses are rapid, non-destructive and require minimal sample preparation whilst taking into account the chemical components within a herb (Wartewig and Neubert, 2005).

The combination of spectroscopic fingerprint with multivariate regression analysis has been used in the field of pharmaceutics (Sacré et al., 2010), agriculture (Yu et al., 2009), food science (Rambla et al., 1997), forensic chemistry (Macleod and Matousek, 2008) and petrochemistry (Blanco and Villarroya, 2002; Breitkreitz et al., 2003). Recently, two comprehensive reviews have been published which summarise the application of the spectroscopic fingerprint in food science (Lu and Rasco, 2011) and pharmaceutics (Roggo et al., 2007). Near-infrared (NIR) spectroscopy coupled with partial least squares regression (PLSR) has been applied to predict TPC and antioxidant capacity in green tea (Zhang et al., 2004). Total polysaccharide and triterpenoid content in *Ganoderma* species was predicted using NIR spectroscopy coupled with PLSR (Chen et al., 2012). Attenuated total reflectance-fourier transform mid-infrared spectroscopy coupled

with PLSR predicted the total phenolic and flavonoid content and antioxidant capacity in Moscatel dessert wine (Silva et al., 2014).

NIR spectroscopy combined with partial lease squares-discriminant (PLS-DA) has effectively differentiated PLR from PTR (Lau et al., 2009). However, the spectral range measured by NIR spectroscopy (4,000 - 10,000 cm^{-1}) is located in the region responsible for broad overlapping overtone and the combination bands of fundamental vibrational modes, making it difficult to interpret specific functional groups (Rodriguez-Saona and Allendorf, 2011). In another study done by Chen and his colleagues (Chen et al., 2013), the one- and two-dimensional ^{1}H nuclear magnetic resonance (NMR) fingerprints of the PLR and PTR were investigated. The results demonstrated that PLS-DA coupled with NMR fingerprints is a beneficial tool for the authentication of PLR and PTR. However, the acquisition time of high resolution 2D NMR is longer than conventional vibrational spectroscopies and the high cost of the instrument and the requirement of highly-trained personnel would restrict its application as a screening method in industry.

Fortunately, recent advances in the development of Raman spectroscopy has made it become one of the fastest growing analytical instruments due to its simplicity, sensitivity and versatility (Das and Agrawal, 2011). Raman spectroscopy is able to analyse sample in different physical states (*e.g.* gas, liquid and solid) with minimal sample preparation. It has been employed for the determination of ethyl esters in soybean oil (Ghesti et al., 2006), α-tocopherol in vegetable (Feng et al., 2013) and omega-3 and omega-6 fatty acids in pork adipose tissue (Olsen et al., 2008).

Furthermore, the recent advances in Raman handheld devices have made it an affordable instrument which can potentially be applied as a quality control device for industry and on-site investigation (Markert et al., 2011). Recent studies have been demonstrated that a handheld Raman spectrometer can effectively differentiate extra virgin olive oil from soybean oil, corn oil and sunflower oil (Zhang et al., 2011c) and to predict the quality of olive oil based on the olive fruits (Guzman et al., 2012). However, the application of Raman spectroscopy in the analysis of herbal medicine is very limited, and therefore, the capability of this technique in predicting the chemical profile of PLR and PTR and other herbal species should be investigated.

To the best of my knowledge, there is no study combining multivariate regression analysis with the Raman spectra of PLR and PTR in the literature. Therefore, the aim of this chapter was to develop a rapid method for the prediction of TPC (Chapter 1 p.34) and antioxidant capacities (Chapter 1 p.35) using Raman spectroscopy in combination with PLSR. This novel approach could provide an alternative method in predicting the chemical properties of a herb, without physically performing the ‘wet’ laboratory experiments. This would be beneficial to industry for the quality control of their herbal products.

6.2 Methods and materials

6.2.1 Chemicals and solvents

HPLC grade methanol and absolute ethanol were purchased from Thermo Fisher Scientific (VIC, Australia). All other chemicals and solvents were of analytical grade and were obtained from Sigma-Aldrich (NSW, Australia) unless otherwise stated.

6.2.2 Herbal samples and extraction

The details of the herbal samples and extraction were described in Chapter 2 (p.75) and Chapter 3 (p.95), respectively.

6.2.3 Raman spectroscopy measurement

The measurement was carried out as described in previous studies with slight modifications (Daferera et al., 2002; Schrader et al., 2000). The spectra were acquired with a Bruker MultiRam Fourier transform Raman spectroscopy equipped with a diode pumped neodymium-doped yttrium aluminium garnet laser and a high throughput module using software OPUS version 6.5.92 (Karlsruhe, Germany). The excitation wavelength was set at 1064 nm and the scatter light was collected by a liquid-nitrogen cooled germanium detector. The powdered samples (10 mg) were loaded onto a black 96 well microplate with a transparent quartz bottom. One hundred scans were averaged for each sample (measured in duplicate). The laser power was set at 100 mW. The spectra were recorded at a resolution of 4 cm^{-1} with a spectral range between 50 cm^{-1} and 3600 cm^{-1} at room temperature. The run time for each sample in duplicate was 3.5 mins.

6.2.4 Colorimetric measurement

6.2.4.1 Total phenolic content

The TPC was determined according to the Folin-Ciocalteau method with slight modifications (Ainsworth and Gillespie, 2007). The diluted ethanolic extract (0.5 mL) was mixed with 2.5 mL of 0.2 N Folin-Ciocalteau's phenol reagent, which was pre-diluted 10 times with deionised water. Sodium carbonate (2 mL; 7.5% w/v) was added five minutes later, and the mixture was incubated at room temperature for 2 hours. The absorbance of the reaction mixture was measured at 760 nm using a Bio-Rad Model 680 microplate reader (Bio-Rad Laboratories, NSW, Australia). The results were expressed as gram of gallic acid equivalents (GAE)/100 g of dried mass (DM).

6.2.4.2 ABTS assay

The ABTS free radical scavenging activity was assessed as previously described (Re et al., 1999). The ABTS free radical solution was prepared by mixing 7 mM ABTS solution with 2.4 mM potassium persulfate (1:1 v/v) and was stored in the dark for at least 12 hours. The absorbance of the ABTS working solution was adjusted to 1.00 ± 0.005 at 760 nm prior to analysis. Subsequently, 10 µL of diluted ethanolic extract was mixed with 200 µL of ABTS working solution. The mixture was incubated in the dark at room temperature for 30 mins. The absorbance was measured at 760 nm and was expressed as milligram trolox equivalent antioxidant capacity (TEAC)/100 g of DM.

6.2.4.3 CUPRAC

The cupric ion reducing capacity was measured using the CUPRAC method proposed by Apak et al. (Apak et al., 2004). The CUPRAC working solution was prepared by mixing equal volumes of 10 mM copper (II) chloride solution, 7.5 mM neocuprine solution and 1M ammonium acetate buffer solution at pH 7.0. Subsequently, 10 µL of diluted ethanolic extract was mixed with 300 µL of the CUPRAC working solution. The mixture was incubated at room temperature for 30 mins and the absorbance was measured at 450 nm. The results were expressed as GAE/100 g of DM.

6.2.5 Data pre-processing

Principal component analysis (PCA) was used to detect outliers (Chapter 1 p.43). Prior to the construction of a PLSR model, the Raman spectra were pre-processed to improve the quality and interpretability of the data matrices (Tistaert et al., 2011a). The influences of standard normal variate (SNV) (Chapter 1 p.39), multiplicative scatter correlation (MSC) (Chapter 1 p.39) and mean centering (Chapter 1 p.40) on the PLSR model performance were evaluated and compared. The combination of pre-processing technique with the best predictability was chosen as the optimal method and was used in the subsequent calculations. The models were validated using root mean square error (RMSE), correlation coefficient (Chapter 1 p.57) and ratio of performance to deviation (Chapter 1 p.59).

6.2.6 Software

All the multivariate analyses and variable selection algorithms were performed on MATLAB R2012b (The MathWorks, MA, U.S.A.). PLS toolbox (Eigenvector

Research Incorporated, WA, U.S.A.) was used to calculate multivariate regression analyses. The K-S algorithm m-files were developed by Daszykowski et al. (Daszykowski et al., 2002).

6.3 Results and discussion

6.3.1 Detection of outliers

To provide a visual representation of the dominant patterns in the Raman spectra of PLR and PTR, the data matrix was analysed by PCA. Figure 6.1a illustrates the Hotelling's T^2 versus Q-residual plot of the samples after the spectra was decomposed by PCA. The Hotelling's T^2 represents the variation in each sample within the PCA model, whereas Q-residual measures how well the reduced dimensional data conforms to the model. A sample was considered as an outlier if Hotelling's T^2 and Q-residual value is high (Wong et al., 2013). As shown in Figure 6.1, no outlier was detected and hence, all the samples were used in the subsequent analysis. Figure 6.1b demonstrates the principle component (PC)1-PC2 scores plot of the Raman spectra. The PCA model was constructed using two PCs which represented 95.83% of the original information from the data matrix. As expected, two distinct clusters were formed in the scores plot, suggesting that the PLR samples had different spectroscopic characteristics as compared to the PTR samples.

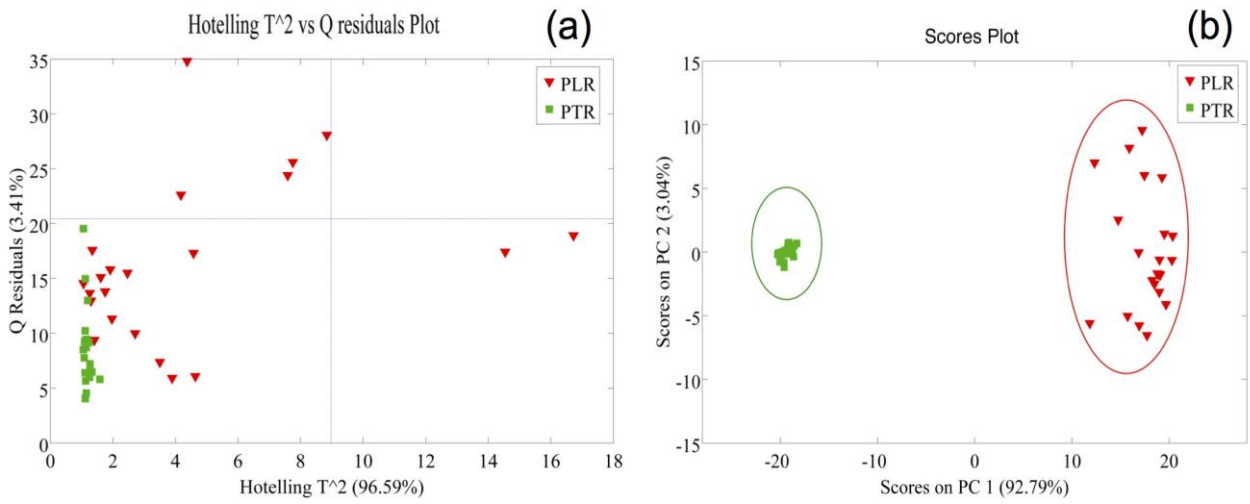


Figure 6.1 PCA plots obtained from pre-processed Raman spectroscopic data matrix. (a) Hotelling's T^2 versus Q residuals plot, (----) line represents 95% confidence interval; (b) PC1-PC2 scores plot.

6.3.2 Raman spectroscopic characteristics of PLR and PTR

The normalised Raman spectra of the samples (the average of 22 PLR and 20 PTR) in the range of 190-1900 cm^{-1} are shown in Figure 6.2. The spectral characteristics of PLR were significantly different from that of PTR, especially in the region of 400 and 1500 cm^{-1}. This region denotes the biochemical composition of carbohydrates and is regarded as a unique fingerprint for the assignment of starch. The spectral characteristics of PTR in this region shared a similar pattern with corn, potato and wheat starch reported in previous studies (Almeida et al., 2010; Kizil et al., 2002). The peaks at 440, 478 and 576 cm^{-1} were due to the skeletal vibrations of the pyranose in the glucose unit of starch (Kizil et al., 2002). The strong peak observed at 478 cm^{-1} reflects the degree of polymerisation in polysaccharides (Bulkin et al., 1987) and has been used to detect the presence of starch in pharmaceutical formulations (de Veij et al., 2009) and carrots (Baranska et al., 2005). The peak at 866 cm^{-1} was assigned to the deformation of C-H bond,

whereas the peak at 941 cm^{-1} was attributed to the C-O-C skeletal vibration of α-1,4-glycosidic linkages (Tu et al., 1979). The peak at 1050-1128 cm^{-1} corresponded to the combination of the stretching mode of C-O, C-C and C-O-H bonds. The peak at 1262 cm^{-1} originated from the deformation of the CH_2OH bond, while the peak at 1460 cm^{-1} was due to the stretching of the C-H bond (Kizil et al., 2002). It was observed that the peak intensities of PTR within this region were, in general, much greater than PLR, and this may be related to the amount of starch present within the samples. In Chapter 2 (p.85), the content of starch of PTR was more than 70% (w/w) of starch per dried mass, whereas PLR had less than 11% (w/w). Therefore, the starch spectral characteristics were the dominant pattern for PTR and correlated to the quantitative results.

A strong peak at 1626 cm^{-1} was observed for PLR and not for PTR. This was attributed to the C=C stretching vibration from the aromatic ring and is regarded as the characteristic peak for polyphenols (Schulz and Baranska, 2007) and has been observed for gallic acid (Calheiros et al., 2008) and curcumin (Baranska et al., 2004). This result was in good agreement with previous observations, which showed that the total flavonoid content in PLR was 4.42 times greater than that in PTR (Chapter 3 p.113).

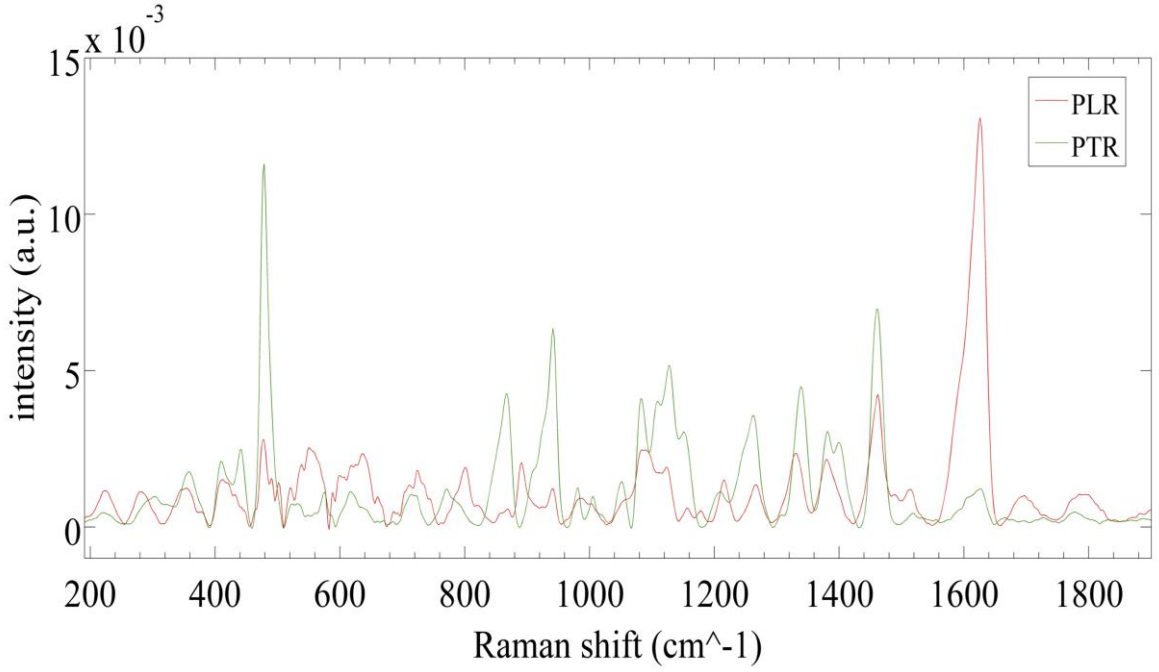


Figure 6.2 Pre-processed mean Raman spectra of 22 PLR and 20 PTR samples.

6.3.3 Quantitative analysis of TPC and antioxidant capacities

Table 6.1 illustrates the descriptive statistics of the TPC, ABTS activity and CUPRAC of the PLR and PTR ethanolic extracts. It is important to note that a wide range of variability presented in the data matrix is crucial for generating a robust and stable PLSR calibration model, whereas a narrow-ranged data matrix can negatively affect the stability and predictability of the model (Tistaert et al., 2011a). In general, the PLR ethanolic extracts had significantly greater TPC and antioxidant capacity than the PTR ethanolic extracts ($p<0.001$). For the PLR ethanolic extracts, TPC, ABTS and CUPRAC were 55.24 ± 10.47 GAE/100 g DM, 161.41 ±17.68 TEAC/100 g of DM and 104.77 ± 15.43 GAE/100 g of DM, respectively. For the PTR ethanolic extracts, TPC, ABTS and CUPRAC were 5.08 ± 2.17 GAE/100 g DM, 28.56 ± 3.58 TEAC/100 g of DM and 24.50 ± 2.84 GAE/100 g of DM, respectively. The relatively higher TPC found in PLR

contributes to a greater antioxidant capacity as compared to PTR. The TPC results obtained from this study were in good agreement with a previous study, which demonstrated that the TPC in PLR was six times more than PTR quantified by high performance liquid chromatography (Chen et al., 2008a). A study by Jiang and his colleagues (Jiang et al., 2005) showed that PLR aqueous extract was five times more effective than PTR aqueous extract in preventing the 2,2'-azo-bis-(2-amidinopropane) dihydrochloride free radical medicated haemolysis in rat red blood cells.

Table 6.1 Summary of the TPC, ABTS activity and CURRAC statistics of the tested samples.

		mean	SD	min	max
PLR	TPC*	55.24	10.47	42.95	74.32
	ABTS*	161.41	17.68	117.91	186.91
	CUPRAC*	104.77	15.43	64.52	136.40
PTR	TPC	5.08	2.17	2.12	10.59
	ABTS	28.56	3.58	22.61	35.66
	CUPRAC	24.50	2.84	19.89	30.33

SD: standard deviation; min: minimum value; max: maximum value.
*** Value obtained in PLR was significantly greater than that in PTR ($p<0.001$).**

6.3.4 Optimisation of pre-processing methods

To improve the data quality and interpretability of the Raman spectra, various pre-processing techniques were compared and evaluated. Preliminary PLSR models using different combinations of the aforementioned techniques were constructed. Their model performances were assessed by LOOCV and are summarised in Table 6.2. A desired model should have low complexity, low RMSE values and

high regression coefficient value (Rinnan et al., 2009; Tistaert et al., 2011a). Since the aim of the regression model is to predict the response Y matrix, ratio of performance to deviation (RPD) was adopted to assess the model's predictability. RPD measures the accuracy of the calibration model by dividing the standard deviation of the experimental values by the standard error of the estimated values obtained from the prediction set (Williams and Sobering, 1993). It has been proposed that a RPD value above 8.1 is classified as an 'excellent' regression model, whereas a between 6.5 and 8.0 is classified as a 'very good' regression model. A value between 5.0 and 6.4 is categorised as a 'good' model (Escuredo et al., 2013; Williams and Sobering, 1993). The higher the RPD value, the better the productivity of the model. It can be observed that the PLSR model pre-processed with normalisation followed by mean centering yielded the lowest RMSE values (RMSEC: 0.4476; RMSECV: 3.7654; RMSEV: 2.7927) and the highest regression coefficient (r^2 cal: 0.9992; r^2 cal: 0.9458; r^2 pred: 0.9721) and RPD value (9.84) among others. Therefore, it was chosen as the optimal pre-processing combination and was applied to the Raman spectra in the subsequent analysis.

Table 6.2 Comparison of the PLSR model parameters obtained from various pre-processing methods.

	LVs	RMSEC	RMSECV	RMSEV	r^2 cal	r^2 CV	r^2 val	RPD
None	2	4.7053	5.1339	4.1175	0.9149	0.8989	0.9488	3.60
Nor + MC	5	0.4476	3.7654	2.7927	0.9992	0.9458	0.9721	9.84
MSC + MC	7	0.3083	3.9347	3.1301	0.9996	0.9408	0.9646	8.91
1st Derivative (5) + MC	3	1.8575	4.6996	4.6415	0.9867	0.9155	0.9317	5.28
1st Derivative (9) + MC	4	1.3012	4.5935	5.5496	0.9935	0.9193	0.8936	4.82
1st Derivative (13) + MC	4	1.3080	4.6960	5.6316	0.9934	0.9161	0.8867	4.75
2nd Derivative (5) + MC	2	2.2277	4.7239	4.5150	0.9809	0.9153	0.9311	5.12
2nd Derivative (9) + MC	3	1.2440	4.4746	5.5727	0.9941	0.9232	0.8989	4.82
2nd Derivative (13) + MC	3	2.0631	4.5103	4.7412	0.9836	0.9222	0.9321	5.06

1st Derivative: First order Savitzky-Golay derivative with bracket represents the window number; 2nd Derivative: Second order Savitzky-Golay derivative; MSC: multiplicative scatter correction; MC: mean centering; LVs: latent variables; RMSEC: root mean square error of calibration; RMSECV: root mean square error of leave-one-out cross-validation; RMSEV: root mean square error of validation; r^2 cal: regression coefficient of the calibration set; r^2 CV: regression coefficient of the leave-one-out cross-validation set; r^2 val: regression coefficient of the validation set; Deviation: deviation between RSMEV and RMSEC; RPD: ratio of performance to deviation.

6.3.5 PLSR model performance of Raman spectra

6.3.5.1 Calibration model

By employing the optimal pre-processing techniques, three PLSR models were constructed which enabled the prediction of the TPC, ABTS and CUPRAC of PLR and PTR using the entire Raman spectra. The number of LVs employed for each model was assessed by the RMSECV generated from LOOCV of the calibration set. The LV with the lowest RMSEC was considered as the optimal number of LVs, and was determined as 5, 4 and 4 in the TPC, ABTS and CUPRAC model, respectively. Figure 6.3 illustrates the correlation between the values measured by the reference analytical methods and the values predicted by the PLSR models. A linear relationship was obtained for each model, suggesting that the predicted values were highly correlated with their respective values measured in the actual experiments. For the TPC model, the regression coefficient of calibration (r^2 cal) and leave-one-out cross-validation (r^2 CV) were determined as 0.9992 and 0.9458, respectively, with a RMSEC and RMSECV of 0.4476 and 3.7654, respectively. Comparatively, the regression coefficients obtained from the antioxidant capacity models were slightly lower than that from the TPC model. In the ABTS model, the r^2 cal and r^2 CV were calculated as 0.9972 and 0.9595, respectively, with a RMSEC and RMSECV of 2.0857 and 7.9320, respectively. In the CUPRAC model, the r^2 cal and r^2 CV were calculated as 0.9966 and 0.9305, respectively, with a RMSEC and RMSECV of 0.3250 and 1.4721, respectively.

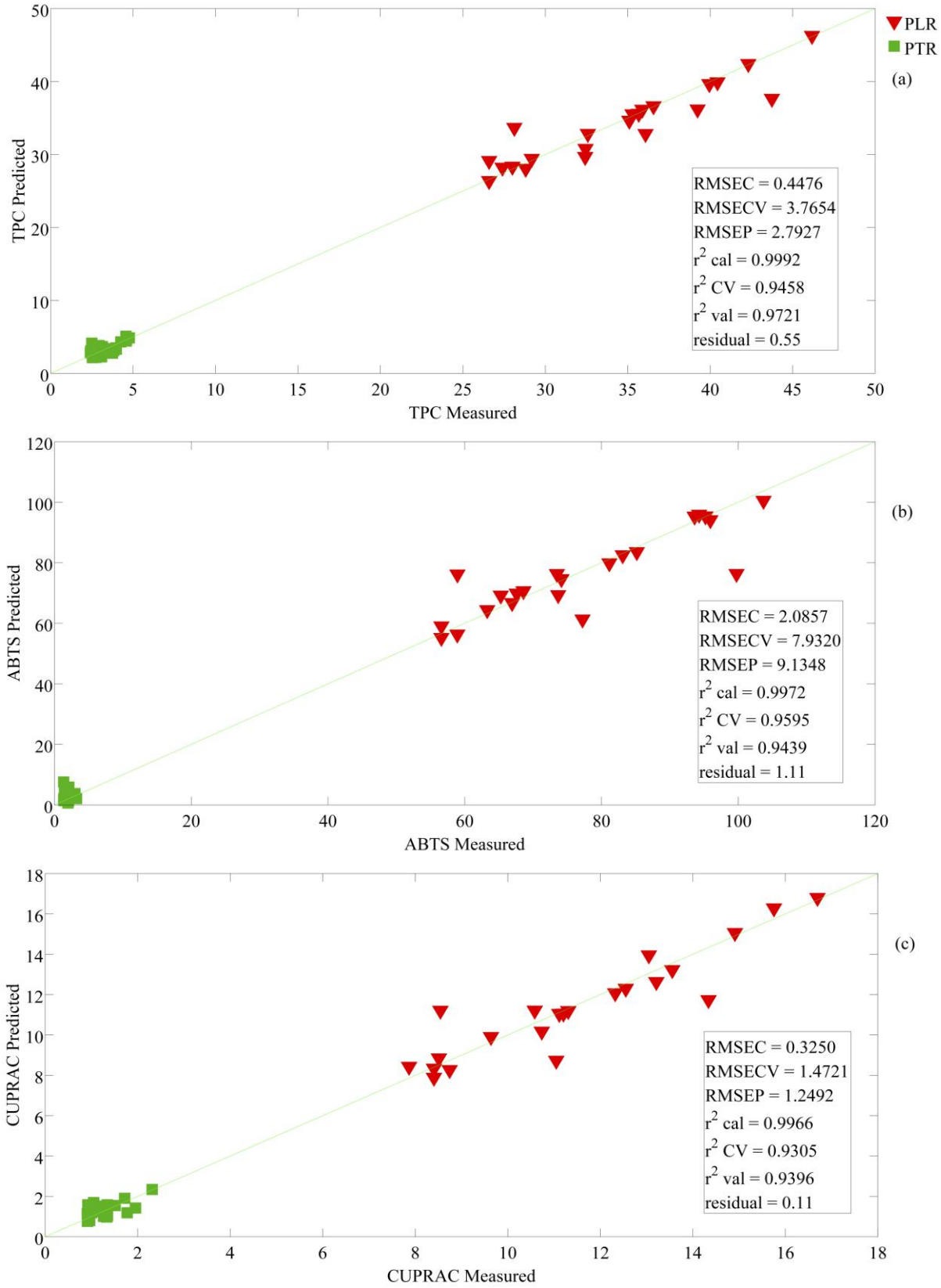


Figure 6.3 Comparison of the measured and predicted values using the calibration and validation PLSR model. (a) Total phenol content; (b) ABTS activity; (c) CUPRAC.

6.3.5.2 Validation model

The 14 samples (8 samples from PLR and 6 samples from PTR) selected by K-S algorithm were used as a validation set to assess the robustness of the established PLSR models. The TPC model yielded the highest regression coefficient of validation (r^2 val: 0.9721) with a RMSEV of 2.7927 in the validation set. Similar to the results from the calibration set, the regression coefficients obtained from the antioxidant capacity models were smaller than those of the TPC model. The r^2 val and RMSEV in the ABTS model were determined as 0.9439 and 9.1348, respectively. For the CUPRAC model, the r^2 val was calculated as 0.9396 with a RMSECV of 1.2492. Furthermore, the RPD values of the TPC, ABTS and CUPRAC model were calculated as 9.84, 7.11 and 7.13, respectively, suggesting that these models are either 'excellent' or 'very good' regression models in predicting the response Y matrix. The residual values of the three models from the validation set were calculated to assess the models' fitness in predicting new samples. The residual refers to the amount of variability from a dependent value that is unexplained in a regression analysis which represents the differences between the measured and predicted value. Consequently, the lower the residual value, the more accurate the regression model is. From the residual plot, the absolute mean residual values from the TPC, ABTS and CUPRAC models were determined as 0.55, 1.11 and 0.11, respectively, suggesting that these models provided similar estimated values comparable with those obtained from the reference analytical methods.

6.3.5.3 Loading plot

The loading plot can be used to identify important variables/peaks. Figure 6.4 illustrates the X-loading plot of the first LV from the TPC, ABTS and CUPRAC models in the calibration set. If a peak in the X-loading plot had a large loading score, this suggested that the corresponding peak in the Raman spectrum was important in predicting the TPC and antioxidant capacity of the samples. For Figure 6.4, the first LV for each model shared a similar spectral pattern in the X-loading plot. In addition, the peak at 1626 cm^{-1} possessed the highest intensity, indicating that the polyphenol characteristic peak was the most influential peak in estimating the response variables. In addition, eight dominant peaks were also observed in the range of 400 cm^{-1} to 1500 cm^{-1}, which is the spectral region responsible for the identification of starch. These results suggested that the presence of starch and polyphenols presence in the sample are important factors in differentiating PLR from PTR, which was in good agreement with the results from previous chapters.

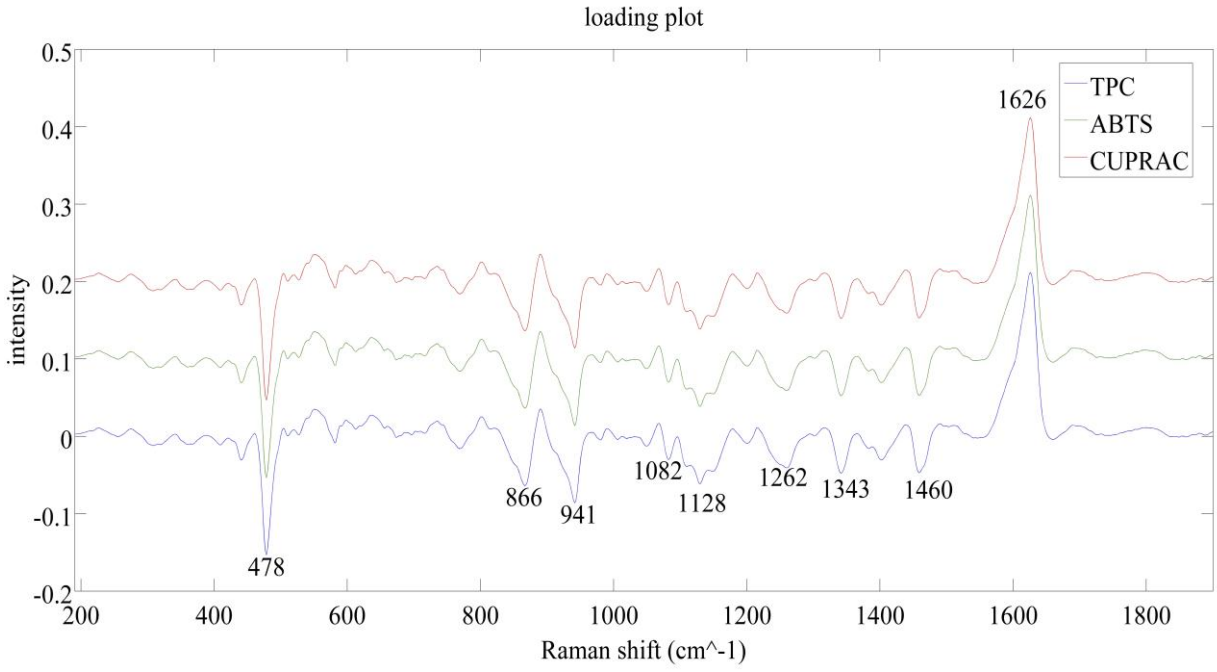


Figure 6.4 Loading plot of the PLSR model using full Raman spectra.

6.4 Conclusion

In this study, a rapid and non-destructive method coupling Raman spectroscopy with multivariate regression analysis PLSR for the prediction of TPC and antioxidant capacities was presented. High RMSE and regression coefficient values suggest that the proposed PLSR models can be used to predict the TPC and antioxidant capacity in PLR and PTR using Raman spectra. Raman spectroscopy coupled with PLSR is an efficient technique for the quality control application of PLR and PTR in industry, and can be potentially applied as a template for the prediction of major chemical families in herbal materials. With the practicality of portable Raman devices, these models can serve as a potential on-site or on-line screening method in the quality control of natural products.

Chapter Seven

Comparing the bioactivities of Puerariae Lobatae Radix and Puerariae Thomsonii Radix

7.1 Introduction

According to the PPRC, PLR and PTR are indicated for the treatment of diabetes and cardiovascular diseases. Although PLR and PTR are separated into two distinct monographs in the PPRC, their indications and recommended doses are identical (10 - 15 g) (Wong et al., 2011). From the results obtained from previous chapters, the chemical profile of PLR was significantly different to that of PTR. As a result, it is necessary to compare the *in vitro* pharmacological activities of PLR and PTR.

PLR and PTR are a rich source of antioxidants, in particular isoflavonoids (Jiang et al., 2005; Wong et al., 2011; Zhang et al., 2011a), and have been shown to possess anti-diabetic, anti-hypertensive and anti-hyperlipidaemia properties in several clinical trials (Liu et al., 2008; Liu, 2004; Lu, 2004; Tam et al., 2009).

The major chemical constituents from PLR and PTR such as puerarin, genistein and daidzein have been found to significantly decrease the baseline fasting plasma glucose, total cholesterol and insulin levels in several mice models (Choi et al., 2008; Hsu et al., 2003; Lee, 2006). One of the therapeutic approaches in the management of diabetes is to prevent postprandial hyperglycaemia (Bischoff, 1994). The absorption of glucose is delayed after a meal by inhibiting the activities of intestinal carbohydrate hydrolysis enzymes such as α-amylase and α-glucosidase. α-Amylase is a major endoglycosidase secreted by the salivary glands and pancreas, and catalyses the initial hydrolysis of starch into shorter oligosaccharides (Sales et al., 2012). Subsequently, α-glucosidase further digests the mixture and releases the monosaccharide α-glucose (Tadera et al., 2006). The

inhibition of these enzymes prolongs carbohydrate digestion time, causing a reduction in the rate of glucose absorption in the small intestine and, consequently, minimising the occurrence of the post-prandial serum glucose spike (Bischoff, 1994; Hanefeld, 1998).

The apoptosis of endothelial cells disturbs the integrity of the endothelial monolayer and is considered as the initial step in the pathogenesis of atherosclerosis, which leads to the development of various vascular disorders (Davigon and Ganz, 2004). Recent studies have reported that the rebalance of the cellular reactive oxygen species (ROS) equilibrium by antioxidants is beneficial in counteracting the endothelial injury. Puerarin has been found to effectively attenuate ROS-induced damage in various *in vitro* and *in vivo* studies (Jiang et al., 2003; Xiong et al., 2006).

The pharmacological activities of other major chemical compounds in PLR and PTR such as genistein and daidzein have been reported in the literature. Genistein and daidzein are widely distributed in the plant kingdom (Barnes, 2010), with many studies focusing on their potent anti-cancer and phytoestrogenic properties. Genistein and daidzein inhibited the growth of hormonal-related cancers such as prostate, breast and cervical cancer (Adjakly et al., 2013; de Lemos, 2001). Recent studies have suggested that the pre-treatment of puerarin effectively inhibited growth and induced apoptosis in breast (Lin et al., 2009) and colon cancer (Yu and Li, 2006). These findings suggest that PLR and PTR might be beneficial in the treatment of cancer therapy.

Most of the previous pharmacological studies focused on the activities of individual compounds in PLR and PTR, with a handful of studies investigating the pharmacological activities of PLR extract. The investigation of the extract provides a better simulation of TCM therapy. To the best of my knowledge, no study on the activity of PTR extract has been reported. Therefore, the pharmacological differences between PLR and PTR extracts were investigated in this study. Since it is believed that the chemical composition could significantly impact on the pharmacological activity of a herb, it is hypothesised that the pharmacological activity and/or potency of PLR should be different to PTR.

In particular, the potential beneficial effect of PLR and PTR extracts on the activity of carbohydrate hydrolysis enzymes α-amylase and α-glucosidase, cyto-protective effect against hydrogen peroxide (H_2O_2)-induced cell death on human endothelial EA.hy926 cells and cytotoxic effect on human prostate cancer PC3 cells have not yet been reported. This study will provide preliminary evidence to support the traditional use of PLR for treating diabetes and cardiovascular diseases and also generate information on whether the clinical recommended dose for the two species should be reviewed in the near future.

7.2 Methods and materials

7.2.1 Herbal sample extraction

The details of herbal samples extraction was described in Chapter 3 (p.95). Four samples (PLR4, PLR7, PLR15 and PLR23; PTR3, PTR6, PTR15 and PTR17) from each species with the highest total phenolic content, total flavonoid content and antioxidant capacities from Chapter 3 and Chapter 7 were chosen.

7.2.2 Porcine pancreatic α-amylase inhibition assay

The experiment was performed as previously described with minor modifications (Kam et al., 2012b). Briefly, the ethanolic extract (10 µL) was mixed with 20 mM sodium phosphate buffer (10 µL; pH 6.9), which contained 33.3 U/mL of porcine pancreatic α-amylase. The mixture was incubated at 25 ^{o}C for 10 mins. Subsequently, 10 µL of soluble potato starch solution (1.5% w/v) was added and the mixture was incubated at 25 ^{o}C for 10 mins. The enzymatic reaction was terminated by adding 96 mM dinitrosalicylic acid (20 µL) and boiled in a water bath for 15 mins. The cooled mixture was diluted with 150 µL of deionised water and was analysed at 550 nm. The final absorbance of the reaction mixture was calculated as the absorbance of the experimental reaction mixture minus the absorbance of the mixture without the addition of enzyme solution. Acarbose (1 µM – 8 mM) was used as the positive control. The concentration of sample required for inhibiting 50% of the enzymatic activity under the assay conditions was the IC_{50} value.

7.2.3 Rat intestinal α-glucosidase inhibition assay

The experiment was carried out in accordance to a previous study (Gao et al., 2007). To obtain an α-glucosidase working solution, 100 mg of rat intestinal acetone powder was dissolved in 3 mL of water and was shaken for 1 hour. The mixture was centrifuged at 4,500 rpm for 5 mins and the supernatant was collected. The ethanolic extract (10 µL) was mixed with 10 µL of α-glucosidase solution and the mixture was incubated at 25 ^{o}C for 5 mins. Subsequently, 10 µL of 40 mM sucrose solution was added and the mixture was incubated at 37 ^{o}C for 30 mins. The enzymatic reaction was terminated in a boiling water bath for 5 mins. The glucose content was determined by adding 200 µL of Glucose C2 (Wako Pure Chemical Industries, Osaka, Japan) colour reagent. The mixture was incubated at 37 ^{o}C for 10 mins and was read at 490 nm. The final absorbance of the reaction mixture was calculated as mentioned in 7.2.2 (p.186). Acarbose (0.61 µM – 500 µM) was used as the positive control.

7.2.4 Cell cultures

Human vascular endothelial cell (EA.hy926) and human prostate cell (PC3) were kindly donated by Dr Shanhong Ling (Monash University Central Clinical School, Australia) and Dr Qihan Dong (Department of Medicine, The University of Sydney, Australia), respectively. The cells were cultured in Dulbecco's modification Eagle's medium (DMEM)/ Ham's F12 supplemented with 15 nM of HEPES and L-glutamine, 100 U/mL of penicillin, 100 U/mL of streptomycin and 10% of foetal bovine serum (Life Technologies, Australia) at 37 ^{o}C under 5% carbon dioxide in humidified air. The PLR and PTR ethanolic extracts were dissolved in DMEM and

filtered with a 0.2 µm PVDF membrane. The filtered ethanolic stock solutions (5 mg/mL) were further diluted with DMEM to desired concentrations.

7.2.4.1 Cyto-protective effect against oxidative stress

EA.hy926 cells were seeded in 96-well plates at a density of 1.0 x 10^5 cell/mL and allowed to grow until confluent for 24 hours. The cells were pre-treated for 4 hours with 1, 10, 50, 100, 150, 200, 400, 500, 600 µg/mL of PLR and PTR ethanolic extracts, and then exposed to 0.4 mM of hydrogen peroxide (H_2O_2) for another 20 hours. The cell viability of the endothelial cells after exposure to oxidative stress was measured by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) dye reduction assay as previously described (Chapter 1 p.36) (Kam et al., 2012a). Briefly, MTT solution (5 mg/mL) was added to each well and the plates were incubated for 4 hours at 37 ^oC. After the medium was aspirated, the purple dye crystal formed inside the viable cells was dissolved in dimethyl sulfoxide. The optical density of each well was measured on a microplate reader at 550 nm. The experiments were performed in triplicates, and the optical density of formazan in the control was taken as 100% viability (Chapter 1 p.36). Gallic acid (20 µg/mL) was used as the positive control. DMEM without the addition of PLR and PTR ethanolic extract was used as the negative control. The experiment was performed in triplicates.

7.2.4.2 Prostate cancer cell growth inhibition assay

The cytotoxicity of the PLR and PTR ethanolic extracts on PC3 cells was investigated using a standard MTT colorimetric assay. PC3 cells were seeded at a density of 1 x 10^4 cell/mL onto a 96-well plate and incubated at 37 ^oC for 24 hours.

The cells were untreated (negative control) or treated with 10, 50, 100, 500, 1000, 2000, 3000, 4000, 5000 µg/mL of ethanolic extract and incubated at 37 ^{o}C for 72 hours. The cell viability of endothelial cells after treatment was assessed by the MTT colorimetric assay. Gallic acid (20 µg/mL) was used as the positive control. DMEM without the addition of PLR and PTR ethanolic extract was used as the negative control. The experiment was performed in triplicates.

7.2.5 Data analysis

The statistical analyses were performed on GraphPad Prism version 6.01 (GraphPad Software, CA, U.S.A.). The data was analysed using either one way analysis of variance (ANOVA) or Student's t-test.

7.3 Results and discussion

7.3.1 Porcine pancreatic α-amylase inhibition assay

The activities of the PLR and PTR ethanolic extracts on α-amylase are presented in Figure 7.1a. The ethanolic extracts of PLR and PTR demonstrated a concentration-dependent inhibitory effect on porcine pancreatic α-amylase activity. In the concentration range of 97.66 µg/mL to 250.00 mg/mL, the percentage of inhibition ranged from 23.54 ± 1.50% to 95.29 ± 0.69% for PLR; for PTR, the inhibition ranged from 29.45 ± 3.73% to 94.91 ± 1.28%. However, a significantly lower IC_{50} value (4.67 ± 0.68 mg/mL; $p<0.001$) was obtained for the PLR ethanolic extracts suggesting that it was more potent in inhibiting α-amylase activity as compared to the PTR ethanolic extracts (IC_{50}: 14.46 ± 0.73 mg/mL).

7.3.2 Rat intestinal α-glucosidase inhibition assay

As illustrated in Figure 7.1b, PLR and PTR ethanolic extracts inhibited rat intestinal α-glucosidase activity in a concentration-dependent manner. In the concentration range of 18.31 µg/mL to 150.00 mg/mL, the percentage of inhibition ranged from 3.45 ± 0.98% to 95.08 ± 0.49% for PLR; for PTR, the inhibition ranged from 15.58 ± 1.53% to 92.10 ± 1.97%. As for the α-amylase inhibition assay, the PLR ethanolic extracts showed superior inhibitory activity (IC_{50}: 2.70 ± 0.40 mg/mL; $p<0.001$) on α-glucosidase as compared to the PTR ethanolic extracts (IC_{50}: 11.40 ± 0.27 mg/mL). Comparing the IC_{50} values of the two inhibition assays, both PLR and PTR ethanolic extracts had a higher potency towards the inhibition of α-glucosidase activity than that of α-amylase activity.

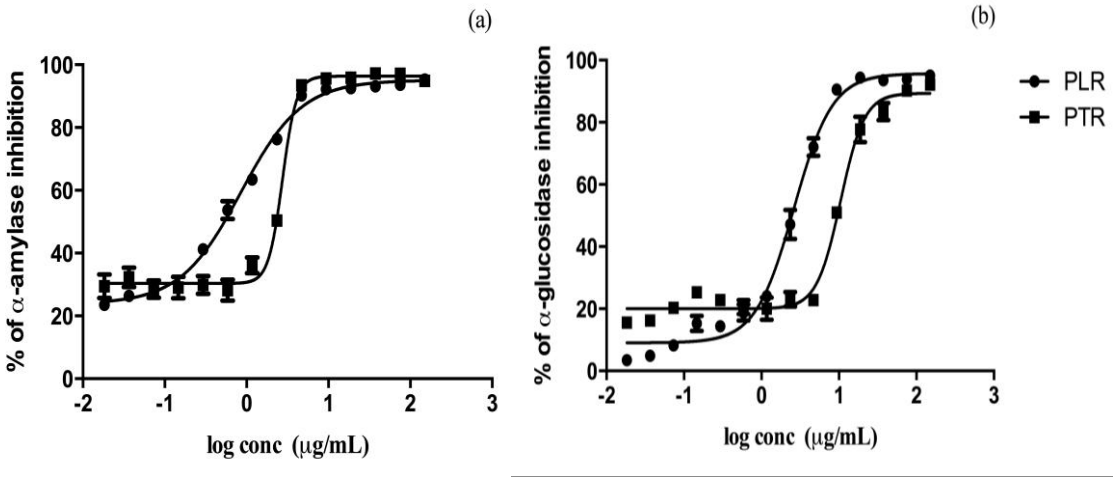


Figure 7.1 Effect of the PLR and PTR ethanolic extracts on the activity of carbohydrate hydrolysis enzymes. (a) Porcine pancreatic α-amylase (n = 8); (b) Rat intestinal α-glucosidase (n = 8). Experiment was performed in triplicate. Acarbose was used as the positive control.

PLR extract has previously been shown to have anti-diabetic activity (Prasain et al., 2012), while the potential anti-diabetic effect of PTR has not yet been examined. The anti-diabetic effect of PLR and PTR might be partly related to the

inhibition of α-amylase and α-glucosidase. A previous study has compared the inhibitory effect of 16 common polyphenols on porcine-originated α-amylase and rat small intestinal and yeast-originated α-glucosidase. Among them, isoflavonoids was one of the polyphenol groups that possessed potent inhibitory effects on all three types of carbohydrate hydrolysis enzymes (Tadera et al., 2006). This potent inhibitory effect exerted by isoflavonoids is related to the presence of hydroxyl groups on aromatic ring A and B of the isoflavonoid skeleton, especially at C-5 and C-7 position of ring A, which facilitates the binding to the active site of the enzymes (Xiao et al., 2013a; Xiao et al., 2013b).

7.3.3 Cytoprotective effect against oxidative stress

The effects of the PLR and PTR ethanolic extracts on human endothelial EA.hy926 cells against H_2O_2-induced cell death are presented in Figure 7.2. PLR ethanolic extracts significantly attenuated ($p<0.001$) the H_2O_2-induced cell injury and restored the cell viability as compared to the H_2O_2-treated group. The cytoprotective effect of the PLR extract was initiated at 200 µg/mL (60.49 ± 3.54%) and increased to 1000 µg/mL (96.00 ± 1.17%) in a concentration-dependent manner as compared to the H_2O_2-treated group (47.86 ± 2.46%). In contrast, the PTR ethanolic extracts had an insignificant protective effect on EA.hy926 cells against H_2O_2-induced apoptosis in the concentration range of 100 to 1000 µg/mL ($p>0.05$) as compared to the H_2O_2-treated group (48.36 ± 1.08%).

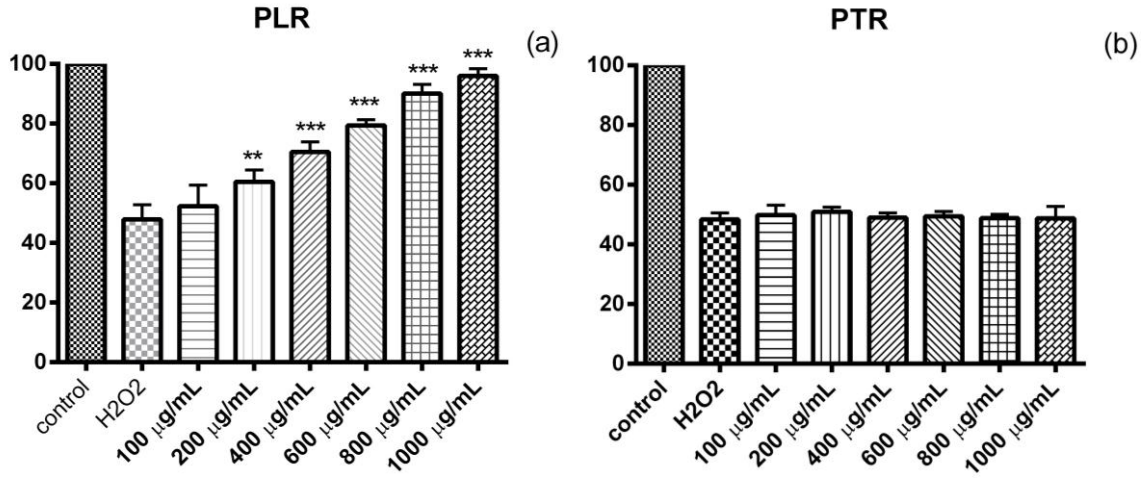


Figure 7.2 Effects of the (a) PLR and (b) PTR on H_2O_2-induced cell death on human endothelial EA.hy926 cells ** p < 0.01; *** p < 0.001 (n = 8). The experiment was performed in triplicate. Gallic acid was used as the positive control.

Oxidative stress induced by ROS, including H_2O_2, superoxide and peroxynitrite, is regarded as a key factor in endothelial cell dysfunction (Higashi et al., 2009). Excess H_2O_2 affects ROS generation and impairs cellular antioxidant defenses, resulting in an increase of ROS production (Pennathur and Heinecke, 2007). Excess ROS can damage the macromolecules such as nucleic acids, lipids and proteins and/or trigger apoptotic pathways. Therefore, the inhibition of this process could prevent the cellular injury, and subsequently halt the progression of various cardiovascular and cerebrovascular diseases (Ogita and Liao, 2004). Puerarin has been found to protect PC12 neuronal cells (Jiang et al., 2003) and rat pancreatic islets from Wistar rats (Xiong et al., 2006) against H_2O_2-induced apoptosis. Apart from puerarin, other major chemical constituents in PLR such as genistein and daidzein have also been demonstrated to possess protective effect against oxidative stress by restoring the activity and protein expression of

malondialdehyde and the decreased superoxide dismutase in streptozotocin-induced diabetic rats (Baluchnejadmojarad and Roghani, 2008; Roghani et al., 2013). Malondialdehyde is a highly active by-product from lipid peroxidation, whereas superoxide dismutase is an anti-oxidant defense metalloenzymes for the homeostasis of intracellular ROS. The activation of superoxide dismutase could restore the intracellular ROS imbalance and hence prevent cell injury (Roghani et al., 2013).

The results from Figure 7.2 indicate that PLR extracts, but not PTR extract, have the potential to protect EA.hy926 cells against injury induced by H_2O_2, which may be partly related to the high abundance of puerarin and the presence of certain unique chemical constituents such as 3'-methoxypuerarin. The UPLC results from Chapter 3 (p.107) illustrated that the amount of puerarin in PLR was 11.95 times greater than in PLT and 3'-methoxypuerarin was absent in PTR. In a previous study, the cytoprotective effect of 3'-methoxypuerarin against cerebral ischaemia/reperfusion injury in rats was investigated (Liu et al., 2010). 3'-Methoxypuerarin (60 mg/kg) was administrated intraperitoneally daily for 3 days after ischaemic insult. The results demonstrated that the number of surviving hippocampal neuron cells in the 3'-methoxypuerarin-treated group was significantly greater than the non-treatment control group. This cytoprotective effect was related to the β-adrenergic effect in combined with the reduction of PGI_2 release, which prevented cerebrovascular spasm and enhanced the vascular permeability and blood flow (Liu et al., 2010; Zhao et al., 2007).

7.3.4 Prostate cancer growth inhibition assay

The effects of the PLR and PTR ethanolic extracts on the cell viability of prostate cancer PC3 cells are illustrated in Figure 7.3. PLR and PTR dramatically diminished the PC3 cell growth in a concentration-dependent manner. A significantly higher IC_{50} value (833.19 ± 1.74 µg/mL; $p<0.001$) obtained from the PTR ethanolic extracts suggested that it was less potent in inhibiting PC3 cell growth as compared to the PLR ethanolic extracts (IC_{50}: 451.83 ± 1.12 µg/mL).

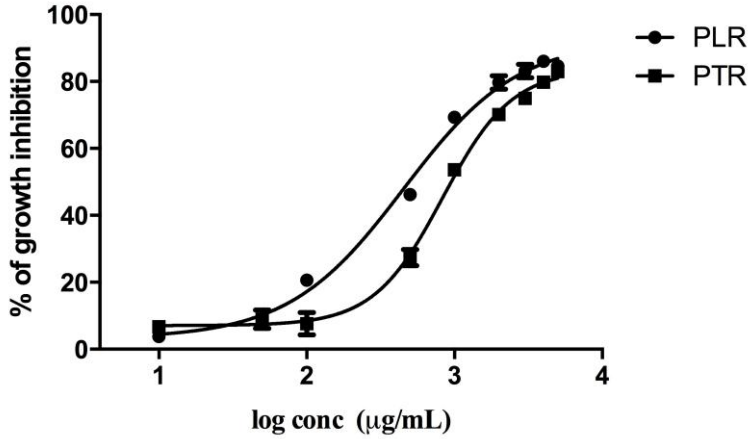


Figure 7.3 Effects of the PLR and PTR ethanolic extracts on the cell viability of human prostate cancer PC3 cells (n = 8). Experiment was performed in triplicate.

Puerarin has been shown to effectively inhibit growth and induce apoptosis in various cancer cell lines (Hien et al., 2010; Lin et al., 2009; Yu and Li, 2006). In a study done by Yu and Li (Yu and Li, 2006), puerarin pre-treatment (50 – 100 µM) for 72 hours effectively decreased the cell viability of HT-29 cells in a concentration-dependent manner. The cell death was triggered by the activation of intrinsic apoptotic pathway, via down-regulation of B-cell lymphoma-2 (Bcl-2) and the up-regulation of caspase-3 and Bcl-2-associated X protein expression. The high anticancer potency in PLR may also be related to the presence of minor

constituents such as daidzin, genistin and genistein, which have been shown to effectively inhibit the growth of prostate and ovarian cancer cell lines, via both intrinsic and extrinsic apoptotic pathways (Adjakly et al., 2013; de Lemos, 2001).

7.4 Conclusion

In this chapter, the pharmacological activities of PLR and PTR were evaluated and compared and this correlated to the chemical profile. The inferior potency of PTR in these preliminary studies suggests that it should not be used interchangeably with PLR in clinical practice. Clinical studies could provide additional information on whether the recommended therapeutic dose of PTR should be reviewed. The enzymatic and cyto-protective assays further verify the clinical use of PLR for the treatment and management of diabetes and cardiovascular diseases. Furthermore, the cytotoxic effect showed in this study reveals that PLR and PTR might be potentially useful as a functional food for the management of cancer.

Chapter Eight

Conclusion

8.1 General discussion and summary

The use of herbal medicines for the prevention and treatment of diseases has increased worldwide (World Health Organisation, 2002). Although there is much research supporting their potential role in mainstream medicine, their safety and efficacy remains inconclusive. One of the major reasons is the inconsistency of bioactive components within a herb which can be significantly affected by environmental factors and production processes. Another issue is the misidentification and substitution of herbal material. This could be due to incorrect nomenclature and taxonomical classification, similar descriptions in the monographs and identical indications and recommended doses for species from the same genus.

As a result, the quality control of herbal materials is considered as the most crucial step in the field of herbal medicine as it could dramatically impact on the safety and efficacy of the herbal products (Jiang et al., 2010; Liang et al., 2009). For Puerariae Lobatae Radix (PLR) and Puerariae Thomsonii Radix (PTR), it has been found that these two species are used interchangeably in the clinical practice (Wong et al., 2011). To date, there has not been an exhaustive study on the differences in their morphological, chemical profile and biological activity. Consequently, these two herbs might produce different clinical effectiveness.

In this thesis, four major differentiation procedures were examined:

1. The morphological characteristics of PLR and PTR.
2. The qualitative and quantitative analysis of the chemical profiles of PLR and PTR.

3. The use of chemometrics coupled with analytical instruments as an efficient and rapid method for differentiating PLR and PTR.
4. The biological activity of PLR and PTR.

Currently, the morphological characteristics of PLR and PTR have similar descriptions in the monographs. However, the macroscopic and microscopic characteristics such as the size of xylem vessels and the size and number of fibre bundles in PLR were observed to be significantly higher than PTR. Higher total dietary fibre content and total starch content was found in PTR, which correlated to the starchy texture and high starch abundance observed under the microscope.

The quantitative colorimetric assays revealed that the total flavonoid content and DPPH antioxidant capacity in PLR were significantly greater than in PTR. Chemical quantification using a newly developed ultra-performance liquid chromatography (UPLC) method revealed that isoflavonoids such as puerarin, daidzin, daidzein and genistein were the primary chemical constituents in both species, while only genistin, 3'-methoxypuerarin, 3'-hydroxypuerarin and 6"-O-D-xylosylpuerarin were present in PLR.

To show that the chromatographic fingerprint is superior to a single chemical marker, chromatographic fingerprint coupled with multivariate discriminant analysis was adopted. A partial least squares-discriminant analysis (PLS-DA) model constructed using the entire chromatographic fingerprint from a newly developed UPLC method achieved a 100% correct species classification rate, while the model established by puerarin alone achieved a correct species classification rate

of 86.6%. The model was applied to granule products and showed that four out seventeen granule products, which were labelled as PLR, were manufactured using PTR.

For the industry, maintaining expensive instruments and examining multiple samples is a challenge. In this study, a simple and effective HPTLC method was developed to directly extract information from a digitalised data of HPTLC into chromatographic fingerprints. As compared to a UPLC chemometrics approach to classify samples, HPTLC was equally effective in differentiating PLR from PTR, with a high specificity and sensitivity and zero error rate.

Another analytical method that can be used as an on-line method during different manufacturing processes and requires minimal sample preparation is the adaptation of spectroscopy. The current study has shown that Raman spectroscopy of the presence of starch and phenolic components in the two herbs was important factors and useful in differentiating PLR from PTR. The high correlation coefficient and ratio of performance to deviation indicated that the partial least squares regression models were accurate and robust in predicting the TPC and antioxidant capacities in both PLR and PTR.

To examine the influence of the chemical profile, pharmacological studies were undertaken. PLR extracts showed significantly greater inhibitory activity of α-amylase and α-glucosidase and the growth of prostate cancer PC3 cells in a concentration-dependent manner than PTR extracts. For the oxidative stress on endothelial EA.hy926 cells induced by H_2O_2, the cytoprotective effects were only

observed for the PLR extracts. The higher pharmacological potency observed in PLR is likely due to the relatively high puerarin, total phenolic and flavonoid content, and the presence of different chemical constituents such as genistin, 3'-methoxypuerarin and 3'-hydroxypuerarin which have been shown to possess anti-inflammatory, antioxidant and cytoprotective effects in various *in vitro* and *in vivo* models (Lee, 2006; Liu et al., 2010; Zhao et al., 2007).

In summary, PLR exhibited significantly different morphological, chemical and pharmacological properties as compared to PTR and thus, revision of the monographs is warranted. Simple analytical instruments such as HPTLC and Raman spectroscopy presenting the whole chemical profile combined with chemometrics can effectively differentiate PLR from PTR. The application of various chemometric techniques can be employed as a template for the quality control of PLR and PTR in the pharmaceutical industry. The differences in pharmacological activities between PLR and PTR suggest that the substitution of PLR with PTR should be reviewed. Further study on pharmacological activities and efficacy of PLR and PTR would provide much needed knowledge on these popular herbal medicines.

8.2 Limitations and future directions

Due to the limited time frame, only 42 samples were purchased and analysed in this thesis. The major acquired locations were restricted to China and Australia. Since *Pueraria* species have been naturalised in different countries around the world, the analysis of samples from various regions could provide a much complete picture about the influence of different continents on the morphological,

chemical and pharmacological characteristics of PLR and PTR (Banks, 2008; DPI&F, 2008; NWO, 2008; US Congress, 1993; Wong et al., 2011). In addition to the increase of sample size, the inclusion of other species within the *Pueraria* genus such as *P. mirifica*, *P. tuberosa* and *P. edulis* might generate a more robust and accurate model for the quality control of *Pueraria* species.

Since it is believed that isoflavonoids, which are sensitive to UV, are the major chemical components within PLR and PTR, UPLC coupled with diode array detector (DAD) recorded at 245 nm was used for the acquisition of chromatographic fingerprints (Jiang et al., 2005; Rong et al., 1998). However, this approach might ignore the presence of non-UV sensitive chemical constituents such as saponins (Park et al., 2000; Zhang and Qu, 2013), which might also contribute to the chemical and pharmacological differences of PLR and PTR. To solve this problem, DAD detector measured at multi-wavelength and/or hyphenated analytical instruments including UPLC-MS and UPLC-NMR can be used. Recent studies have demonstrated that chromatographic fingerprints measured at multiple wavelengths are useful in detecting different chemical compositions and can be employed for the quality control of herbal materials (Lucio-Gutiérrez et al., 2012a; Lucio-Gutiérrez et al., 2012b). PLS-DA model constructed using chromatographic fingerprints measured at four wavelengths (226, 254, 280 and 326 nm) has been shown to effectively differentiate Valerianae Officinalis Radix from other species within the *Veleriana* genus such as Valerianae Wallichii Radix and Valerianae Edulis Radix, with a 100% correct species classification rate (Lucio-Gutiérrez et al., 2012a).

In the preparation of TCM, water and sometimes Chinese white wine such as Kaoliang were used to extract chemical constituents of the herbs (Kastner, 2011). Nonetheless, in this study absolute ethanol was employed as it presented the highest TPC and antioxidant capacity. Therefore, to emulate conventional TCM preparation, water and/or water/ethanol mixture can be used in the future study. The ratio of water and ethanol used to extract chemical constituents from PLR and PTR can be optimised using a chemometrics technique, namely response surface methodology, which aims to explore an adequate functional relationship between a number of explanatory variables (experimental conditions, *e.g.* ethanol concentration, extraction time and temperature) and one or more response of interests such as total phenolic content and amount of bioactive chemical component (Bezerra et al., 2008). This approach has optimised ethanol concentration and extraction duration and temperature on Saposhinikovlae Radix (Li et al., 2011a), Astragali Radix (Xiao et al., 2008), Hypericum Perforatum Herba (Liu et al., 2000) and Morinda Citrifolia Fructus (Thoo et al., 2010).

In regard to the pharmacological activities of the current thesis, the experiments were limited to enzymatic and *in vitro* molecular assays. It is important to note that the enzymatic and cellular assays used in this study are non-specific and serve as a preliminary screening method. Despite PTR demonstrates inferior anti-diabetic, cyto-protective and cytotoxic effects as compared to PTR in current study, these results need to be re-examined in the *in vivo* model. The *in vitro* approach might not completely reflect the *in-vivo* process of metabolism, pharmacokinetic and pharmacology in the human body. As a result, further studies on the therapeutic effects of PLR and the impact of substituting PTR with PLR in clinical practice

should be investigated. Bioassay-guided fraction to isolate individual potential active constituents may provide information on the active constituent(s). Further investigation on the signaling pathways and mode of actions of the pharmacological effects exerted by PLR and PTR extracts and/or compounds are urgently warranted.

According to the TCM practice, a herb is seldom used on its own and is usually combined with other herbs and presented in a formula, which consists of at least two and up to 20-30 herbs (Ung et al., 2007). The most fundamental and simplest form in a formula is the herb pair, which comprises of two TCM herbs that complementary to each other. This approach is based on the Yin and Yang theory and is believed to enhance certain therapeutic effects by mutual enhancement, assistance and restraint (Wang et al., 2012). Therefore, instead of investigating the pharmacological effect of a single herb, the effect of herb pairs and formulae can be explored in a future study.

References

ADJAKLY, M., NGOLLO, M., BOITEUX, J. P., BIGNON, Y. J., GUY, L. & BERNARD-GALLON, D. 2013. Genistein and daidzein: different molecular effects on prostate cancer. *Anticancer Res.,* 33**,** 39-44.

AFSETH, N. K., SEGTNAN, V. H. & WOLD, J. P. 2006. Raman spectra of biological samples: A study of preprocessing methods. *Appl Spectrosc.,* 60**,** 1358-67.

AINSWORTH, E. A. & GILLESPIE, K. M. 2007. Estimation of total phenolic content and other oxidation substrates in plant tissues using Folin-Ciocalteu reagent. *Nat Protoc.,* 2**,** 875-7.

ALAERTS, G., DEJAEGHER, B., SMEYERS-VERBEKE, J. & VANDER HEYDEN, Y. 2010. Recent developments in chromatographic fingerprints from herbal products: set-up and data analysis. *Comb Chem High Throughput Screen.,* 13**,** 900-22.

ALMEIDA, M. R., ALVES, R. S., NASCIMBEM, L. B., STEPHANI, R., POPPI, R. J. & DE OLIVEIRA, L. F. 2010. Determination of amylose content in starch using Raman spectroscopy and multivariate calibration analysis. *Anal Bioanal Chem.,* 397**,** 2693-701.

ALQAHTANI, A., HAMID, K., KAM, A., WONG, K. H., ABDELHAK, Z., RAZMOVSKI-NAUMOVSKI, V., CHAN, K., LI, K. M., GROUNDWATER, P. W. & LI, G. Q. 2013. The pentacyclic triterpenoids in herbal medicines and their pharmacological activities in diabetes and diabetic complications. *Curr Med Chem.,* 20**,** 908-31.

AMAZON 2011. Kudzu. Amazon.com Inc. Accessed on 7 January 2011 http://www.amazon.com/s/ref=nb_sb_noss?url=search-alias%3Dhpc&field-keywords=kudzu&x=0&y=0.

APAK, R., GÜÇLÜ, K., DEMIRATA, B., ÖZYÜREK, M., ÇELIK, S. E., BEKTAŞOĞLU, B., BERKER, K. I. & ÖZYURT, D. 2007. Comparative evaluation of various total antioxidant capacity assays applied to phenolic compounds with the CUPRAC assay. *Molecules.,* 12**,** 1496-1547.

APAK, R., GUCLU, K., OZYUREK, M. & KARADEMIR, S. E. 2004. Novel total antioxidant capacity index for dietary polyphenols and vitamins C and E, using their cupric ion reducing capability in the presence of neocuproine: CUPRAC method. *J Agric Food Chem.,* 52**,** 7970-81.

ARAO, T., IDZU, T., KINJO, J., NOHARA, T. & ISOBE, R. 1996. Oleanene-Type Triterpene Glycosides from Puerariae Radix. III. Three New Saponins from *Pueraria thomsonii*. *Chem Pharm Bull (Tokyo).* 44**,** 1970-72.

ARAO, T., KINJO, J., NOHARA, T. & ISOBE, R. 1995. Oleanene-type triterpene glycosides from Puerariae Radix. II. Isolation of saponins and the application of tandem mass spectrometry to their structure determination. *Chemical & Pharmaceutical Bulletin,* 43**,** 1176-1179.

ARAO, T., KINJO, J., NOHARA, T. & ISOBE, R. 1997. Oleanene-type triterpene glycosides from puerariae radix. IV. Six new saponins from Pueraria lobata. *Chem Pharm Bull (Tokyo).* 45**,** 362-366.

ARAUJO, M. C. U., SALDANHA, T. C. B., GALVAO, R. K. H., YONEYAMA, T., CHAME, H. C. & VISANI, V. 2001. The successive projection algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometr Intell Lab Syst.,* 57**,** 65-73.

ARVANITOYANNIS, I. S. & VLACHOS, A. 2009. Maize authentication: quality control methods and multivariate analysis (chemometrics). *Crit Rev Food Sci Nutr.,* 49**,** 501-37.

BALABIN, R. M., SAFIEVA, R. Z. & LOMAKINA, E. I. 2011. Near-infrared (NIR) spectroscopy for motor oil classification: From discriminant analysis to support vector machines. *Microchem. J.,* 98**,** 121-128.

BALABIN, R. M. & SMIRNOV, S. V. 2011. Variable selection in near-infrared spectroscopy: benchmarking of feature selection methods on biodiesel data. *Anal Chem Acta.,* 692**,** 63-72.

BALLABIO, D. & CONSONNI, V. 2013. Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal Methods,* 5**,** 3790-98.

BALUCHNEJADMOJARAD, T. & ROGHANI, M. 2008. Chronic administration of genistein improves aortic reactivity of streptozotocin-diabetic rats: mode of action. *Vascul Pharmacol.,* 49**,** 1-5.

BANKS, M. 2008. Weed profile: Kudzu. Newsletter of the Land for Wildlife Program South East Queensland, Australia, Vol 2, pp.6.

BAO, J., CAI, Y., SUN, M., WANG, G. & CORKE, H. 2005. Anthocyanins, flavonols, and free radical scavenging activity of Chinese bayberry (Myrica rubra) extracts and their color properties and stability. *J Agric Food Chem.,* 53**,** 2327-32.

BARANSKA, M., SCHULZ, H., BARANSKI, R., NOTHNAGEL, T. & CHRISTENSEN, L. P. 2005. In situ simultaneous analysis of polyacetylenes, carotenoids and polysaccharides in carrot roots. *J Agric Food Chem.,* 53**,** 6565-71.

BARANSKA, M., SCHULZ, H., ROSCH, P., STREHLE, M. A. & POPP, J. 2004. Identification of secondary metabolites in medicinal and spice plants by NIR-FT-Raman microspectroscopic mapping. *Analyst.,* 129**,** 926-930.

BARNES, S. 2010. The biochemistry, chemistry and physiology of the isoflavones in soybeans and their food products. *Lymphat Res Biol.,* 8**,** 89-98.

BECKONERT, O., E. BOLLARD, M., EBBELS, T. M. D., KEUN, H. C., ANTTI, H., HOLMES, E., LINDON, J. C. & NICHOLSON, J. K. 2003. NMR-based metabonomic toxicity classification: hierarchical cluster analysis and k-nearest-neighbour approaches. *Anal Chem Acta.,* 490**,** 3-15.

BERRIDGE, M. V. & TAN, A. S. 1993. Characterization of the cellular reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): subcellular localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction. *Arch Biochem Biophys.,* 303**,** 474-82.

BERRUETA, L. A., ALONSO-SALCES, R. M. & HEBERGER, K. 2007. Supervised pattern recognition in food analysis. *J Chromatogr A.,* 1158**,** 196-214.

BEZERRA, M. A., SANTELLI, R. E., OLIVEIRA, E. P., VILLAR, L. S. & ESCALEIRA, L. A. 2008. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta.,* 76**,** 965-977.

BISCHOFF, H. 1994. Pharmacology of alpha-glucosidase inhibition. *Eur J Clin Invest.,* 24**,** 3-10.

BLANCO, M. & VILLARROYA, I. 2002. NIR spectroscopy: a rapid-response analytical tool. *TrAC Trend Anal Chem.,* 21**,** 240-250.

BLUMETHAL, M., BUSSE, W. R. & GOLDBERG, A. 1999. *The Complete German Commission E Monographs: Therapeutic Guid to Herbal Medicine,* Texas, USA, American Botanical Concil.

BODNER, C. C. & HYMOWITZ, T. 2002. Chapter 2: Ethnobotany of Pueraria species. *In:* KEUNG, W. M. (ed.) *Pueraria: The genus Pueraria.* CRC Press, USA, pp. 13.

BP 2014a. Volumn IV: Kudzuvine root. British Pharmacopoeia Online. Accessed on 29 March 2014 http://www.pharmacopoeia.co.uk/.

BP 2014b. Volumn IV: Thomson kudzuvine root. British Pharmacopoeia Online. Accessed on 29 March 2014 http://www.pharmacopoeia.co.uk/.

BRAND-WILLIAMS, W., CUVELIER, M. E. & BERSET, C. 1995. Use of a free radical method to evaluate antioxidant activity. *LWT-Food Science and Technology.,* 28**,** 25-30.

BRAVO, L. 1998. Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. *Nutr Rev.,* 56**,** 317-33.

BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A. & STONE, C. J. 1984. *Classificaiton and Regressioin Trees,* Chapman & Hall (Wadsworth, Inc.), New York.

BREITKREITZ, M. C., RAIMUNDO, J. I. M., ROHWEDDER, J. J. R., PASQUINI, C., DANTAS FILHO, H. A., JOSE, G. E. & ARAUJO, M. C. U. 2003.

Determination of total sulfur in diesel fuel employing NIR spectroscopy and multivariate calibration. *Analyst.,* 128**,** 1204-1207.

BRERETON, R. G. 2003. *Chemometrics: Data analysis for the laboratory and chemical plant,* John Wiley & Sons, West Sussex.

BROADHURST, D., GOODACRE, R., JONES, A., ROWLAND, J. J. & KELL, D. B. 1997. Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. *Anal Chem Acta.,* 348**,** 71-86.

BROMBA, M. U. A. & ZIEGLER, H. 1981. Application hints for Savitzky-Golay digital smoothing filters. *Anal Chem.,* 53**,** 1583-1586.

BULKIN, B. J., KWAK, Y. & DEA, I. C. M. 1987. Retrogradation kinetics of waxy-corn and potato starches; a rapid, Raman-spectroscopic study. *Carbohydrate Research.,* 160**,** 95-112.

CALHEIROS, R., MACHADO, N. F. L., FIUZA, S. M., GASPAR, A., GARRIDO, J., MILHAZES, N., BORGES, F. & MARQUES, M. P. M. 2008. Antioxidant phenolic esters with potential anticancer activity: A Raman spectroscopy study. *J Raman Spectrosc.,* 39**,** 95-107.

CAMAG. 2014. *TLC/HPTLC methods* [Online]. Accessed on 27 Jan 2014 http://www.camag.com/en/tlc_hptlc/camag_laboratory/methods.cfm

CHALMERS, J. M., EDWARDS, H. G. M. & HARGREAVES, M. D. 2012. *Infrared and Raman Spectroscopy in Forensic Science*, John Wiley & Sons, West Sussex, UK.

CHAN, K. 1995. Progress in traditional Chinese medicine. *Trends Pharmacol Sci.,* 16**,** 182-7.

CHAN, K. 2005. Chinese medicinal materials and their interface with Western medical concepts. *J Ethnopharmacol.,* 96**,** 1-18.

CHANG, C.-C., YANG, M. H., WEN, H. M. & CHERN, J. C. 2002. Estimation of total flavonoid content in propolis by two complementary colorimetric methods. *J Food Drug Anal.,* 10**,** 178-182.

CHANG, Y. S. & HO, Y. L. 2009. *Colour Illustration of Chinese Materia Medica,* The Committee on Chinese Medicine and Pharmacy, Department of Health, Taiwan, pp. 167.

CHEN, J., SONG, M., ZHANG, Y. & T.J., H. 2008a. Determination of pueraria isoflavones in Pueraria thosonii and Pueraria lobata by HPLC reference fingerprint chromatogram and Folin-phenol reagent colorimetry. *Chinese Journal of New Drug.,* 17**,** 1240-1246.

CHEN, M.-S. & WANG, S.-W. 1999. Fuzzy clustering analysis for optimizing fuzzy membership functions. *Fuzzy Sets and Systems.,* 103**,** 239-254.

CHEN, Q., ZHAO, J., LIU, M., CAI, J. & LIU, J. 2008b. Determination of total polyphenols content in green tea using FT-NIR spectroscopy and different PLS algorithms. *J Pharm Biomed Anal.,* 46**,** 568-573.

CHEN, S., YANG, D., CHEN, S., XU, H. & CHAN, A. S. 2007. Seasonal variations in the isoflavonoids of radix Puerariae. *Phytochem Anal.,* 18**,** 245-50.

CHEN, S. B., LIU, H. P., TIAN, R. T., YANG, D. J., CHEN, S. L., XU, H. X., CHAN, A. S. & XIE, P. S. 2006. High-performance thin-layer chromatographic fingerprints of isoflavonoids for distinguishing between Radix Puerariae Lobate and Radix Puerariae Thomsonii. *J Chromatogr A.,* 1121**,** 114-9.

CHEN, Y., XIE, M., ZHANG, H., WANG, Y., NIE, S. & LI, C. 2012. Quantification of total polysaccharides and triterpenoids in Ganoderma lucidum and Ganoderma atrum by near infrared spectroscopy and chemometrics. *Food Chem.,* 135**,** 268-275.

CHEN, Y. G., SONG, Y. L., WANG, Y., YUAN, Y. F., HUANG, X. J., YE, W. C., WANG, Y. T. & ZHANG, Q. W. 2013. Metabolic differentiations of Pueraria lobata and Pueraria thomsonii using H NMR spectroscopy and multivariate statistical analysis. *J Pharm Biomed Anal.,* 20**,** 00218-5.

CHENG, J. 2000. Review: Drug Therapy in Chinese Traditional Medicine. *J Clin Pharmacol.,* 40**,** 445-50.

CHENG, Z., MOORE, J. & YU, L. 2006. High-Throughput Relative DPPH Radical Scavenging Capacity Assay. *J Agric Food Chem.,* 54**,** 7429-7436.

CHOI, M. S., JUNG, U. J., YEO, J., KIM, M. J. & LEE, M. K. 2008. Genistein and daidzein prevent diabetes onset by elevating insulin level and altering hepatic gluconeogenic and lipogenic enzyme activities in non-obese diabetic (NOD) mice. *Diabetes Metab Res Rev.,* 24**,** 74-81.

CONNOLLY, J. D. & HILL, R. A. 2008. Triterpenoids. *Nat Prod Rep.,* 25**,** 794-830. .

CONNOLLY, J. D. & HILL, R. A. 2010. Triterpenoids. *Nat Prod Rep.,* 27**,** 79-132. .

COOMANS, D. & MASSART, D. L. 1982a. Alternative k-nearest neighbour rules in supervised pattern recognition : Part 1. k-Nearest neighbour classification by using alternative voting rules. *Anal Chem Acta.,* 136**,** 15-27.

COOMANS, D. & MASSART, D. L. 1982b. Alternative k-nearest neighbour rules in supervised pattern recognition : Part 2. Probabilistic classification on the basis of the kNN method modified for direct density estimation. *Anal Chem Acta.,* 138**,** 153-165.

COOMANS, D. & MASSART, D. L. 1982c. Alternative k-nearest neighbour rules in supervised pattern recognition : Part 3. Condensed nearest neighbour rules. *Anal Chem Acta.,* 138**,** 167-176.

CROZIER, A., JAGANATH, I. B. & CLIFFORD, M. N. 2009. Dietary phenolics: chemistry, bioavailability and effects on health. *Nat Prod Rep.,* 26**,** 1001-43.

CSURHES, S. 2008. Kudzu (*Pueraria montana* var. *lobata*) infestation on the Gold Coast. Department of Primary Industries and Fisheries. Australia: Queensland Government.

DAFERERA, D., PAPPAS, C., TARANTILIS, P. A. & POLISSIOU, M. 2002. Quantitative analysis of α-pinene and β-myrcene in mastic gum oil using FT-Raman spectroscopy. *Food Chem.,* 77**,** 511-515.

DAS, R. S. & AGRAWAL, Y. K. 2011. Raman spectroscopy: Recent advancements, techniques and applications. *Vib Spectrosc.,* 57**,** 163-176.

DASZYKOWSKI, M., WALCZAK, B. & MASSART, D. L. 2002. Representative subset selection. *Anal Chim Acta.,* 468**,** 91-103.

DAVIGON, J. & GANZ, P. 2004. Role of endothelial dysfunction in atherosclerosis. *Circulation.,* 15**,** 27-32.

DE BEER, T., BURGGRAEVE, A., FONTEYNE, M., SAERENS, L., REMON, J. P. & VERVAET, C. 2011. Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes. *Int J Pharm.,* 417**,** 32-47.

DE LEMOS, M. L. 2001. Effects of soy phytoestrogens genistein and daidzein on breast cancer growth. *Ann Pharmacother.,* 35**,** 1118-21.

DE MAESSCHALCK, R., CANDOLFI, A., MASSART, D. L. & HEUERDING, S. 1999. Decision criteria for soft independent modelling of class analogy applied to near infrared data. *Chemometr Intell Lab Syst.,* 47**,** 65-77.

DE VEIJ, M., VANDENABEELE, P., DE BEER, T., REMON, J. P. & MOENS, L. 2009. Reference database of Raman spectra of pharmaceutical excipients. *J Raman Spectrosc.,* 40**,** 297-307.

DPI&F 2008. Fact sheet: Declared plants of Queensland. Department of Primary Industries and Fisheries, Queensland, Australia.

DU, G., ZHAO, H. Y., ZHANG, Q. W., LI, G. H., YANG, F. Q., WANG, Y., LI, Y. C. & WANG, Y. T. 2010. A rapid method for simultaneous determination of 14 phenolic compounds in Radix Puerariae using microwave-assisted extraction and ultra high performance liquid chromatography coupled with diode array detection and time-of-flight mass spectrometry. *J Chromatogr A.,* 1217**,** 705-14.

EBRAHIMI-NAJAFABADI, H., LEARDI, R., OLIVERI, P., CHIARA CASOLINO, M., JALALI-HERAVI, M. & LANTERI, S. 2012. Detection of addition of barley to coffee using near infrared spectroscopy and chemometric techniques. *Talanta.,* 99**,** 175-9.

EILERS, P. H. 2003. A perfect smoother. *Anal Chem.,* 75**,** 3631-6.

ENGLYST, H. N. & HUDSON, G. J. 1987. Colorimetric method for routine measurement of dietary fibre as non-starch polysaccharides. A comparison with gas-liquid chromatography. *Food Chem.,* 24**,** 63-76.

EP 2012a. Kudzuvine root: Puerariae lobatae radix 2434. European Pharmacopoeia, European Directorate for the Quality of Medicine & HealthCare, Strasbourg, France.

EP 2012b. Thomson kudzuvine root: Puerariae thomsonii radix 24383. European Pharmacopoeia, European Directorate for the Quality of Medicine & HealthCare, Strasbourg, France.

EPPO 2007. Data sheets on quarantine pests: Pueraria lobata. European and Mediterranean Plant Protection Organization Bulletin, 37, 230-235.

ERIKSSON, L., ANTTI, H., GOTTFRIES, J., HOLMES, E., JOHANSSON, E., LINDGREN, F., LONG, I., LUNDSTEDT, T., TRYGG, J. & WOLD, S. 2004. Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Anal. Bioanal. Chem.,* 380**,** 419-429.

ESBENSEN, K. H. & GELADI, P. 2010. Principles of Proper Validation: use and abuse of re-sampling for validation. *J Chemometr.,* 24**,** 168-187.

ESCUREDO, O., SEIJO, M. C., SALVADOR, J. & GONZALEZ-MARTIN, M. I. 2013. Near infrared spectroscopy for prediction of antioxidant compounds in the honey. *Food Chem.,* 141**,** 3409-3414.

EUROPEAN MEDICINES AGENCY 2006. Guideline on quality of herbal medicinal products/ traditional herbal medicinal products. Accessed on 4 March 2014 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003370.pdf.

EUROPEAN MEDICINES AGENCY 2008. Guideline on quality of combination herbal medicinal product/ traditional herbal medicinal products. Accessed on 4 March 2014 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003286.pdf.

EUROPEAN MEDICINES AGENCY 2011. Guideline on specifications: test procedures and acceptance criteria for herbal substances, herbal preparations and herbal medicinal products/ traditional herbal medicinal products. Accessed on 4 March 2014 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/09/WC500113210.pdf.

EVERST, J. W., MILLER, J. H., BALL, D. M. & PATTERSON, M. G. 1991. *Kudzu in Alabama,* Alabama Cooperative Extension Service Circular ANR-65, Aubure, AL, pp.6.

FALCONE FERREYRA, M. L., RIUS, S. P. & CASATI, P. 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front Plant Sci.,* 3**,** 1-15.

FANG, Q. 1980. Some current study and research approaches relating to the use of plants in the traditional Chinese medicine. *J Ethnopharmacol.,* 2**,** 57-63.

FENG, S., GAO, F., CHEN, Z., GRANT, E., KITTS, D. D., WANG, S. & LU, X. 2013. Determination of alpha-tocopherol in vegetable oils using a molecularly imprinted polymers-surface-enhanced Raman spectroscopic biosensor. *J Agric Food Chem.,* 61**,** 10467-75.

FOOD AND DRUG ADMINISTRATION OF THE UNITED STATES OF AMERICA 2004. Guidance for Industry: Botanical Drug Products. U.S. Depart of Health and Human Services, Food and Drug Administration, Center for Drug Evaulation and Research. Accessed on 10 March http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070491.pdf.

FRANK, I. E. & LANTERI, S. 1989. Classification models: Discriminant analysis, SIMCA, CART. *Chemometr Intell Lab Syst.,* 5**,** 247-256.

FRIED, B. & SHERMA, J. 1996. *Practical Thin-Layer Chromatography: A Multidisciplinary Approach*, Taylor & Francis.

FRIED, B. & SHERMA, J. 1999. *Thin-Layer Chromatography, Revised And Expanded*, Taylor & Francis.

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. *Phytochem Anal.,* 24**,** 1-24.

GANZERA, M. 2009. Recent advancements and applications in the analysis of traditional Chinese medicines. *Planta Med.,* 75**,** 776-83.

GAO, H., HUANG, Y.-N., XU, P.-Y. & KAWABATA, J. 2007. Inhibitory effect on α-glucosidase by the fruits of Terminalia chebula Retz. *Food Chem.,* 105**,** 628-634.

GELADI, P. & KOWALSKI, B. R. 1986. Partial least-squares regression: a tutorial. *Anal Chem Acta.,* 185**,** 1-17.

GEMPERLINE, P. 2006. *Practical guide to chemometrics,* CRC Press, Boca Raton.

GHESTI, G., MACEDO, J., BRAGA, V., SOUZA, A. C. P., PARENTE, V. I., FIGUERÊDO, E., RESCK, I., DIAS, J. & DIAS, S. L. 2006. Application of raman spectroscopy to monitor and quantify ethyl esters in soybean oil transesterification. *J Am Oil Chem Soc.,* 83**,** 597-601.

GÓMEZ-CARAVACA, A. M., VERARDO, V., TOSELLI, M., SEGURA-CARRETERO, A., FERNÁNDEZ-GUTIÉRREZ, A. & CABONI, M. F. 2013. Determination of the Major Phenolic Compounds in Pomegranate Juices by HPLC–DAD–ESI-MS. *J Agric Food Chem.,* 61**,** 5328-5337.

GOODARZI, M., RUSSELL, P. J. & VANDER HEYDEN, Y. 2013. Similarity analyses of chromatographic herbal fingerprints: A review. *Anal Chem Acta.,* 804**,** 16-28.

GRUMBACH, E. S., ARSENAULT, J. C. & MCCABE, D. R. 2009. *Beginners Guide to UPLC: Ultra-Performance Liquid Chromatography*, Waters Corporation.

GUZMAN, E., BAETEN, V., PIERNA, J. A. & GARCIA-MESA, J. A. 2012. A portable Raman sensor for the rapid discrimination of olives according to fruit quality. *Talanta.,* 93**,** 94-98.

HANEFELD, M. 1998. The role of acarbose in the treatment of non-insulin-dependent diabetes mellitus. *J Diabetes Complications.,* 12**,** 228-37.

HARADA, M. & UENO, K. 1975. Pharmacological studies on pueraria root. I. Fractional extraction of pueraria root and identification of its pharmacological effects. *Chem Pharm Bull (Tokyo).* 23**,** 1798-1805.

HEALTHSTORE. 2011. *FSC Kudzu Capsule* [Online]. Healthstore.uk.com Accessed on 7 January 2011 http://www.healthstore.uk.com/c43206/kudzu.html.

HIEN, T. T., KIM, H. G., HAN, E. H., KANG, K. W. & JEONG, H. G. 2010. Molecular mechanism of suppression of MDR1 by puerarin from Pueraria lobata via NF-kappaB pathway and cAMP-responsive element transcriptional activity-dependent up-regulation of AMP-activated protein kinase in breast cancer MCF-7/adr cells. *Mol Nutr Food Res.,* 54**,** 918-28.

HIGASHI, Y., NOMA, K., YOSHIZUMI, M. & KIHARA, Y. 2009. Endothelial function and oxidative stress in cardiovascular diseases. *Circ J.,* 73**,** 411-8.

HILL, R. A. & CONNOLLY, J. D. 2011. Triterpenoids. *Nat Prod Rep.,* 28**,** 1087-117. .

HILL, R. A. & CONNOLLY, J. D. 2012. Triterpenoids. *Nat Prod Rep.,* 29**,** 780-818. .

HILL, R. A. & CONNOLLY, J. D. 2013. Triterpenoids. *Nat Prod Rep.,* 30**,** 1028-65. .

HIRAKURA, K., MORITA, M., NAKAJIMA, K., SUGAMA, K., TAKAGI, K., NIITSU, K., IKEYA, Y., MARUNO, M. & OKADA, M. 1997. Phenolic glucosides from the root of Pueraria lobata. *Phytochemistry.,* 46**,** 921-928.

HKCMMS 2010a. Radix Puerariae Lobatae. Chinese Medicine Division, Department of Health, HKSAR, volumn 3, p. 215-223.

HKCMMS 2010b. Radix Puerariae Thomsonii. Chinese Medicine Division, Department of Health, HKSAR, volumn 3, p. 225-233.

HSU, F. L., LIU, I. M., KUO, D. H., CHEN, W. C., SU, H. C. & CHENG, J. T. 2003. Antihyperglycemic effect of puerarin in streptozotocin-induced diabetic rats. *J Nat Prod.,* 66**,** 788-92.

ICH 2005. Validation of analytical procedure: text and methodology Q2 (R1). International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. Accessed on 6 March 2014 http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf.

IUCN. 2005. *Global Invasive Species Database: Puerarin montana var. lobata* [Online]. Accessed 20 August 2010 http://www.issg.org/database/species/ecology.asp?si=81&fr=1&sts=&lang=EN.

JEWETT, D. K., JIANG, C. J., BRITTIN, K. O., SUN, J. H. & TANG, J. 2003. Characterizing specimens of kudzu and related taxa with RAPD's. *Castanea,* 68**,** 254-260.

JIANG, B., LIU, J. H., BAO, Y. M. & AN, L. J. 2003. Hydrogen peroxide-induced apoptosis in pc12 cells and the protective effect of puerarin. *Cell Biol Int.,* 27**,** 1025-31.

JIANG, R.-W., LAU, K.-M., LAM, H.-M., YAM, W.-S., LEUNG, L.-K., CHOI, K.-L., WAYE, M. M. Y., MAK, T. C. W., WOO, K.-S. & FUNG, K.-P. 2005. A comparative study on aqueous root extracts of Pueraria thomsonii and Pueraria lobata by antioxidant assay and HPLC fingerprint analysis. *J Ethnopharmacol.,* 96**,** 133-138.

JIANG, W. Y. 2005. Therapeutic wisdom in traditional Chinese medicine: A prespective from modern science. *Trends Pharmacol Sci.,* 26**,** 558-63.

JIANG, Y., DAVID, B., TU, P. & BARBIN, Y. 2010. Recent analytical approaches in quality control of traditional Chinese medicines--a review. *Anal Chim Acta.,* 657**,** 9-18.

JIN, S., SON, Y., MIN, B.-S., JUNG, H. & CHOI, J. 2012. Anti-inflammatory and antioxidant activities of constituents isolated from Pueraria lobata roots. *Arch Pharm Res.,* 35**,** 823-837.

JIUZHANG, M. & LEI, G. 2010. *A General Introduction to Traditional Chinese Medicine*, Taylor & Francis.

JOHNSON, S. 1967. Hierarchical clustering schemes. *Psychometrika.,* 32**,** 241-254.

JOSHI, D. D. 2012. *Herbal drugs and fingerprints: evidence based herbal drugs*, Springer, India.

JP 2011. Kudzuvine root. Japanese Pharmacopoeia, Pharmaceuticals and Medical Devices Agency, Ministry of Health, Labour and Welfare, p. 1720-1721.

KAM, A., LI, K. M., RAZMOVSKI-NAUMOVSKI, V., NAMMI, S., CHAN, K. & LI, G. Q. 2012a. Combination of TNF-α, homocyteine and afenosine exacerbated cytotoxicity in human cardiovascular and cerebrovascular endothelial cells. *Cell Physiol Biochem.,* 30**,** 805-814.

KAM, A., LI, K. M., RAZMOVSKI-NAUMOVSKI, V., NAMMI, S., CHAN, K. & LI, G. Q. 2013. Variability of the Polyphenolic Content and Antioxidant Capacity of Methanolic Extracts of Pomegranate Peel. *Nat Prod Commun.,* 8**,** 707-710.

KAM, A., LI, K. M., RAZMOVSKI-NAUMOVSKI, V., NAMMI, S., SHI, J., CHAN, K. & LI, G. Q. 2012b. A Comparative Study on the Inhibitory Effects of Different Parts and Chemical Constituents of Pomegranate on alpha-Amylase and alpha-Glucosidase. *Phytother Res.,* 19.

KASTNER, J. 2011. *Chinese Nutrition Therapy: Dietetics in Traditional Chinese Medicine (TCM)*, Thieme.

KENNARD, R. W. & STONE, L. A. 1969. Computer Aided Design of Experiments. *Technometrics.,* 11**,** 137-148.

KEUNG, W. M. & VALLEE, B. L. 1998. Kudzu root: an ancient Chinese source of modern antidipsotropic agents. *Phytochemistry.,* 47**,** 499-506.

KIM, J. M., JANG, D. S., LEE, Y. M., KIM, Y. S. & KIM, J. S. 2008. Puerarol from the roots of Pueraria lobata inhibits the formation of advanced glyceration end products (AGEs) in vitro. *Nat Prod Sci.,* 14**,** 192-195.

KINJO, J. E., FURUSAWA, J. C., BABA, J., TAKESHITA, T., YAMASAKI, M. & NOHARA, T. 1987. Studies on the Constituents of Pueraria lobata. III. Isoflavonoids and Related Compounds in the Roots and the Voluble Stems. *Chem Pharm Bull (Tokyo).* 35**,** 4846-4850.

KIZIL, R., IRUDAYARAJ, J. & SEETHARAMAN, K. 2002. Characterization of irradiated starches by using FT-Raman and FTIR spectroscopy. *J Agric Food Chem.,* 50**,** 3912-8.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. 1995 1995. 1137-1145.

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 Coptidischinensis based on UPLC-PAD combined with chemometrics methods. *Phytomedicine.,* 16**,** 950-9.

LACHENBRUCH, P. A. 1975. *Discriminant analysis*, Wiley Online Library.

LARKIN, P. 2011. *Infrared and Raman Spectroscopy; Principles and Spectral Interpretation*, Elsevier Science.

LAU, C. C., CHAN, C. O., CHAU, F. T. & MOK, D. K. 2009. Rapid analysis of Radix puerariae by near-infrared spectroscopy. *J Chromatogr A.,* 1216**,** 2130-5.

LAURENTIN, A. & EDWARDS, C. A. 2003. A microtiter modification of the anthrone-sulfuric acid colorimetric assay for glucose-based carbohydrates. *Anal Biochem.,* 315**,** 143-5.

LEE, J. S. 2006. Effects of soy protein and genistein on blood glucose, antioxidant enzyme activities, and lipid profile in streptozotocin-induced diabetic rats. *Life Sci.,* 79**,** 1578-84.

LEE, K. R., LIN, X., PARK, D. C. & ESLAVA, S. 2003. Megavariate data analysis of mass spectrometric proteomics data using latent variable projection method. *Proteomics.,* 3**,** 1680-1686.

LEWIS, C. J. 2001. Aristolochi acid: letter to industry associations regarding safety concerns related to the use of botanical products containing aristolochic acid. US Fod and Drug Administration. Accessed on 4 March 2014 http://www.fda.gov/Food/RecallsOutbreaksEmergencies/SafetyAlertsAdvisories/ucm096374.htm.

LI, H., WAN, L., HASHI, Y. & CHEN, S. 2007. Fragmentation study of a 8-C-glycosyl isoflavone, puerarin, using electrospray ion trap time-of-flight mass spectrometry at high resolution. *Rapid Commun Mass Spectrom.,* 21**,** 2497-2504.

LI, W., WANG, Z., SUN, Y.-S., CHEN, L., HAN, L.-K. & ZHENG, Y.-N. 2011a. Application of response surface methodology to optimise ultrasonic-assisted extraction of four chromones in Radix Saposhnikoviae. *Phytochem Anal.,* 22**,** 313-321.

LI, W., XING, L. & QU, H. 2011b. Classification and quantification analysis of *Radix scutellariae* from different origins with near infrared diffuse reflaction spectroscopy. *Vib Spectrosc.,* 55**,** 58-64.

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. *J Chromatogr A.,* 1216**,** 2033-44.

LIANG, Y., XIE, P. & CHAU, F. 2010. Chromatographic fingerprinting and related chemometric techniques for quality control of traditional Chinese medicines. *J Sep Sci.,* 33**,** 410-21. .

LIAO, Y. 2011. *Traditional Chinese Medicine*, Cambridge University Press.

LIN, C. C., WU, C. I. & SHEU, S. J. 2005. Determination of 12 pueraria components by high-performance liquid chromatography-mass spectrometry. *J Sep Sci.,* 28**,** 1785-95.

LIN, Y. J., HOU, Y. C., LIN, C. H., HSU, Y. A., SHEU, J. J., LAI, C. H., CHEN, B. H., LEE CHAO, P. D., WAN, L. & TSAI, F. J. 2009. Puerariae radix isoflavones and their metabolites inhibit growth and induce apoptosis in breast cancer cells. *Biochem Biophys Res Commun.,* 378**,** 683-8. .

LIU, F. F., ANG, C. Y. W. & SPRINGER, D. 2000. Optimization of Extraction Conditions for Active Components in Hypericum perforatum Using Response Surface Methodology. *J Agric Food Chem.,* 48**,** 3364-3371.

LIU, L. H., LIU, F. Q. & TIAN, H. 2008. Clinical observation of puerarin decoction in treating hyperlipoproteinaemia. *Hebei Journal of Traditional Chinese Medicine.,* 30**,** 804-805.

LIU, W. J. H. 2011. *Traditional Herbal Medicine Research Methods: Identification, Analysis, Bioassay, and Pharmaceutical and Clinical Studies*, Wiley.

LIU, X. 2004. Clinical observation of Gegen Qushi hypertension treatment of 46 cases. *Chinese Journal of Traditional Medical Science and Technology.,* 11 5

LIU, Y., CHEN, X., LIU, K., CHAI, Q., LI, J., LIU, Z., SUN, D. & CHEN, L. 2010. Protection by 3'-methoxypuerarin of rat hippocampal neurons against ischemia/reperfusion injury. *Chin J Physiol.,* 53**,** 136-40.

LOUGH, W. J. & WAINER, I. W. 1995. *High Performance Liquid Chromatography: Fundamental Principles and Practice*, Taylor & Francis.

LU, A., JIA, H., XIAO, C. & LU, Q. 2004. Theory of traditional Chinese medicine and therapeutic method of diseases. *World J Gastroenterol.,* 10**,** 1854-1856.

LU, X. 2004. Clinical observation on the effect of Gegen Qushi decoction in treating hyperlipideamia. *Chinese Journal of Traditional Medical Science and Technology.,* 11**,** 12-14.

LU, X. & RASCO, B. A. 2011. Determination of Antioxidant Content and Antioxidant Activity in Foods using Infrared Spectroscopy and Chemometrics: A Review. *Crit Rev Food Sci Nutr.,* 52**,** 853-875.

LUCIO-GUTIÉRREZ, J. R., COELLO, J. & MASPOCH, S. 2012a. Enhanced chromatographic fingerprinting of herb materials by multi-wavelength selection and chemometrics. *Anal Chem Acta.,* 710**,** 40-49.

LUCIO-GUTIÉRREZ, J. R., GARZA-JUÁREZ, A., COELLO, J., MASPOCH, S., SALAZAR-CAVAZOS, M. L., SALAZAR-ARANDA, R. & WAKSMAN DE TORRES, N. 2012b. Multi-wavelength high-performance liquid chromatographic fingerprints and chemometrics to predict the antioxidant activity of Turnera diffusa as part of its quality control. *J Chromatogr A.,* 1235**,** 68-76.

LUO, H., LI, Q., FLOWER, A., LEWITH, G. & LIU, J. 2012. Comparison of effectiveness and safety between granules and decoction of Chinese herbal medicine: a systematic review of randomized clinical trials. *J Ethnopharmacol.,* 140**,** 555-67.

M. BOLAND, G. & M. X. DONNELLY, D. 1998. Isoflavonoids and related compounds. *Nat Prod Rep.,* 15**,** 241-260.

MACLEOD, N. & MATOUSEK, P. 2008. Emerging Non-invasive Raman Methods in Process Control and Forensic Applications. *Pharm Res,* 25**,** 2205-2215.

MADSEN, R., LUNDSTEDT, T. & TRYGG, J. 2010. Chemometrics in metabolomics--a review in human disease diagnosis. *Anal Chim Acta.,* 659**,** 23-33.

MALENCIC, D., POPOVIC, M. & MILADINOVIC, J. 2007. Phenolic content and antioxidant properties of soybean (Glycine max (L.) Merr.) seeds. *Molecules.,* 12**,** 576-81.

MARK, H. & WORKMAN JR., J. 2007. *Chemometrics in spectroscopy,* Elsevier, Amsterdam.

MARKERT, H., RING, J., CAMPBELL, N. & GRATES, K. 2011. A comparison of four commercially available portable Raman spectrometers. National Forensic Science Technology Center, FL, USA. Accessed on 29 April 2014 http://www.nfstc.org/?dl_id=214.

MARSTON, A. 2011. Thin-layer chromatography with biological detection in phytochemistry. *J Chromatogr A.,* 1218**,** 2676-83.

MARTINS, L. R., PEREIRA-FILHO, E. R. & CASS, Q. B. 2011. Chromatographic profiles of Phyllanthus aqueous extracts samples: a proposition of classification using chemometric models. *Anal Bioanal Chem.,* 400**,** 469-81. .

MAS, S., DE JUAN, A., TAULER, R., OLIVIERI, A. C. & ESCANDAR, G. M. 2010. Application of chemometric methods to environmental analysis of organic pollutants: A review. *Talanta.,* 80**,** 1052-1067.

MASSART, D. L., VANDEGINSTE, B. G. M., BUYDENS, L. M. C., DE JONG, S., LEWI, P. J. & SMEYERS-VERBEKE, J. 1997. *Handbook of chemometrics and qualimetrics,* Elsevier, Amsterdam.

MCCREERY, R. L. 2005. *Raman Spectroscopy for Chemical Analysis*, Wiley.

MCLACHLAN, G. 2004. *Discriminant analysis and statistical pattern recognition,* John Wiley & Sons, New Jersey.

MEDA, A., LAMIEN, C. E., ROMITO, M., MILLOGO, J. & NACOULMA, O. G. 2005. Determination of the total phenolic, flavonoid and proline contents in Burkina Fasan honey, as well as their radical scavenging activity. *Food Chem.,* 91**,** 571-577.

MILANOVIC, B. 1997. A simple way to calculate the Gini coefficient, and some implications. *Econ Lett.,* 56**,** 45-49.

MIYAZAWA, M. & KAMEOKA, H. 1988. Volatile flavor components of Puerariae Radix (*Pueraria lobata* Ohwi). *Agric Biol Chem.,* 52**,** 1053-1055.

MOLYNEUX, P. 2004. The use of the stable free radical diphenylpicrylhydrazyl (DPPH) for estimating antioxidant activity. *Songklanakarin J Sci Technol.,* 26**,** 211-219.

MOSMANN, T. 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods.,* 65**,** 55-63.

MUNCK, L., NØRGAARD, L., ENGELSEN, S. B., BRO, R. & ANDERSSON, C. A. 1998. Chemometrics in food science—a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemometr Intell Lab Syst.,* 44**,** 31-60.

MURAKAMI, T., NISHIKAWA, Y. & ANDO, T. 1960. Studies on the Constituents of Japanese and Chinese Crude Drugs. IV. On the Constituents of Pueraria Root. (2). *Chem Pharm Bull (Tokyo).* 8**,** 688-691.

NAKAMOTO, H., MIYAMURA, S., INDAD, K. & NAKAMURA, N. 1975. The study of the aqueous extrac of Puerariae Radix I. The preparation and the components of the active extract. *Yakugaku Zasshi Journal of the Pharmaceutical Society of Japan.,* 95**,** 1123-1127.

NAUSHAD, M. & KHAN, M. R. 2014. *Ultra Performance Liquid Chromatography Mass Spectrometry: Evaluation and Applications in Food Analysis*, Taylor & Francis.

NGUYEN, V. D., MIN, B. C., KYUNG, M. O., PARK, J. T., LEE, B. H., CHOI, C. H., SEO, N. S., KIM, Y. R., AHN, D. U., LEE, S. J., PARK, C. S., KIM, J. W. & PARK, K. H. 2009. Identification of a naturally-occurring 8-[alpha-D-glucopyranosyl-(1-->6)-beta-D-glucopyranosyl] daidzein from cultivated kudzu root. *Phytochem Anal.,* 20**,** 450-5.

NI, L.-J., ZHANG, L.-G., XIE, J. & LUO, J.-Q. 2009. Pattern recognition of Chinese flue-cured tobaccos by an improved and simplified K-nearest neighbors classification algorithm on near infrared spectra. *Anal Chem Acta.,* 633**,** 43-50.

NICM. 2014. *Understanding complementary medicine: facts and statistics* [Online]. The National Institute of Complementart Medicine, Australia, Accessed on 6 March 2014 http://www.nicm.edu.au/understanding-cm/facts-and-statistics.

NOHARA, T., KINJO, J., FURUSAWA, J., SAKAI, Y., INOUE, M., SHIRATAKI, Y., ISHIBASHI, Y., YOKOE, I. & KOMATSU, M. 1993. But-2-enolides from Pueraria lobata and revised structures of puerosides A, B and sophoroside A. *Phytochemistry.,* 33**,** 1207-1210.

NORTIER, J. L., MARTINEZ, M. C., SCHMEISER, H. H., ARLT, V. M., BIELER, C. A., PETEIN, M., DEPIERREUX, M. F., DE PAUW, L., ABRAMOWICZ, D., VEREERSTRAETEN, P. & VANHERWEGHEM, J. L. 2000. Urothelial carcinoma associated with the use of a Chinese herb (Aristolochia fangchi). *N Engl J Med.,* 342**,** 1686-92.

NPS 2008. Information use and needs of complementary medicines users. National Prescribing Servic Limited, Australia. Accessed on 6 March 2014 http://www.nicm.edu.au/images/stories/understanding/docs/complementary_medicines_report_-_consumers.pdf.

NWO 2008. Weed List 2008. Noxious Weeds Office, NSW, Australia.

OGITA, H. & LIAO, J. K. 2004. Endothelial Function and Oxidative Stress. *Endothelium,* 11**,** 123-132.

OHASHI, H., TATEISHI, Y., NEMOTO, T. & ENDO, Y. 1988. Taxonomic studies on the Leguminosae of Taiwan III. *Sci Rep Tohoku Univ Ser.,* 4**,** 131-248.

OLSEN, E. F., RUKKE, E. O., EGELANDSDAL, B. & ISAKSSON, T. 2008. Determination of omega-6 and omega-3 fatty acids in pork adipose tissue with nondestructive Raman and fourier transform infrared spectroscopy. *Appl Spectrosc.,* 62**,** 968-74.

OTTO, M. 2007. *Chemometrics: Statistics and computer application in analytical chemistry*, Wiley-VCH, Weinheim.

OZAKI, Y., MCCLURE, W. F. & CHRISTY, A. A. 2006. *Near-Infrared Spectroscopy in Food Science and Technology*, John Wiley & Sons, New Jersey, USA.

PARK, I. S., KANG, E. M. & KIM, N. 2000. High-performance liquid chromatographic analysis of saponin compounds in Bupleurum falcatum. *J Chromatogr Sci.,* 38**,** 229-33.

PATIL, P. S., VENKATANARAYANAN, R., ARGADE, P. D. & SHINDE, P. R. 2012. Assessment of pharmacognostic and phytochemical standards of Thespesia populnea (L.) root. *Asian Pac J Trop Med.,* 2**,** S1212-S1216.

PENNATHUR, S. & HEINECKE, J. W. 2007. Oxidative stress and endothelial dysfunction in vascular disease. *Curr Diab Rep.,* 7**,** 257-64.

PIETERS, A. J. 1932. Kudzu, a Forage Crop for the Southeast. United States Department of Agriculture, Leaflet 91, pp. 8.

PPRC 1997. Chinese Pharmacopoeia Commission, People's Medical Publishing House, Beijing, China, p. 164-165.

PPRC 2000. Chinese Pharmacopoeia Commission, People's Medical Publishing House, Beijing, China, p. 188-189.

PPRC 2005. *Chinese Pharmacopoeia Commission, People's Medical Publishing House, Beijing, China, p. 230-231.*

PPRC 2010a. *Puerariae Lobatae Radix Chinese Pharmacopoeia Commission, People's Medical Publishing House, Beijing, China, p. 272; 312-313.*

PPRC 2010b. Puerariae Thomsonii Radix Chinese Pharmacopoeia Commission, People's Medical Publishing House, Beijing, China, p. 312-313.

PRASAIN, J. K., JONES, K., KIRK, M., WILSON, L., SMITH-JOHNSON, M., WEAVER, C. & BARNES, S. 2003. Profiling and quantification of isoflavonoids in kudzu dietary supplements by high-performance liquid chromatography and electrospray ionization tandem mass spectrometry. *J Agric Food Chem.,* 51**,** 4213-8.

PRASAIN, J. K., PENG, N., RAJBHANDARI, R. & WYSS, J. M. 2012. The Chinese Pueraria root extract (Pueraria lobata) ameliorates impaired glucose and lipid metabolism in obese mice. *Phytomedicine.,* 20**,** 17-23.

PRASAIN, J. K., REPPERT, A., JONES, K., MOORE, D. R., 2ND, BARNES, S. & LILA, M. A. 2007. Identification of isoflavone glycosides in Pueraria lobata cultures by tandem mass spectrometry. *Phytochem Anal.,* 18**,** 50-9.

PRICE, K. R. & FENWICK, G. R. 1985. Naturally occuring estrogens in foods- a review. *Food Addit Contam.,* 2**,** 73-106.

PRIOR, R. L., WU, X. & SCHAICH, K. 2005. Standardized methods for the determination of antioxidant capacity and phenolics in foods and dietary supplements. *J Agric Food Chem.,* 53**,** 4290-302.

PROCHÁZKOVÁ, D., BOUŠOVÁ, I. & WILHELMOVÁ, N. 2011. Antioxidant and prooxidant properties of flavonoids. *Fitoterapia.,* 82**,** 513-523.

QU, F., SUN, Z. J., XIONG, W. Z., WANG, S., T. & SUN, Z. G. 2011. Thoughts on protection of geographical indication intellectual property of kudzu. *Acta Agriculturae Jiangxi,* 23**,** 172-175.

QUIDEAU, S., DEFFIEUX, D., DOUAT-CASASSUS, C. & POUYSEGU, L. 2011. Plant polyphenols: chemical properties, biological activities, and synthesis. *Angew Chem Int Ed Engl.,* 50**,** 586-621.

RAAMAN, N. 2006. *Phytochemical Techniques*, Delhi, New India Publishing Agency, pp. 19-24.

RAMBLA, F. J., GARRIGUES, S. & DE LA GUARDIA, M. 1997. PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. *Anal Chem Acta.,* 344**,** 41-53.

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 Radic Biol Med.,* 26**,** 1231-1237.

REYNAUD, J., GUILET, D., TERREUX, R., LUSSIGNOL, M. & WALCHSHOFER, N. 2005. Isoflavonoids in non-leguminous families: an update. *Nat Prod Rep.,* 22**,** 504-15.

RICE-EVANS, C. A., MILLER, N. J., BOLWELL, P. G., BRAMLEY, P. M. & PRIDHAM, J. B. 1995. The Relative Antioxidant Activities of Plant-Derived Polyphenolic Flavonoids. *Free Radic Res.,* 22**,** 375-383.

RINNAN, Å., BERG, F. V. D. & ENGELSEN, S. B. 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trend Anal Chem.,* 28**,** 1201-1222.

RODRIGUEZ-SAONA, L. E. & ALLENDORF, M. E. 2011. Use of FTIR for Rapid Authentication and Detection of Adulteration of Food. *Annu Rev Food Sci Tech.,* 2**,** 467-483.

ROGGO, Y., CHALUS, P., MAURER, L., LEMA-MARTINEZ, C., EDMOND, A. & JENT, N. 2007. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J Pharm Biomed Anal.,* 44**,** 683-700.

ROGHANI, M., MAHDAVI, M. V., JALAI-NADOUSHAN, M., T., B., NADERI, G., GROGHANI-DEHKORDI, F., JOGHATAEI, M. T. & KORD, M. 2013. Chronic administration of daidzein, a soybean isoflavone, improves endothelial dysfunction and attenuates oxdiative stress in streptozocin-induced diabetic rats. *Phytother Res.,* 27**,** 112-117.

ROMESBURG, H. C. 2004. *Cluster analysis for researchers,* Lulu Press, North Carolina.

RONG, H., KEUKELEIRE, D. D. & COOMAN, L. D. 2002. Chapter 5: Chemical constituents of Pueraria plants: identification and methods of analysis. *In:* KEUNG, W. M. (ed.) *Pueraria: The genus Pueraria.* CRC Press, USA, pp. 23-25.

RONG, H., STEVENS, J. F., DEINZER, M. L., COOMAN, L. D. & KUEUKELEIRE, D. D. 1998. Identification of isoflavones in the roots of Pueraria lobata. *Planta Med.,* 64**,** 620-627.

ROSE, R., ROSE, C. L., OMI, S. K., FORRY, K. R., DURALL, D. M. & BIGG, W. L. 1991. Starch determination by perchloric acid vs enzymes: evaluating the accuracy and precision of six colorimetric methods. *J Agric Food Chem.,* 39**,** 2-11.

ROSIPAL, R. & KRÄMER, N. 2006. Overview and Recent Advances in Partial Least Squares. *In:* SAUNDERS, C., GROBELNIK, M., GUNN, S. & SHAWE-TAYLOR, J. (eds.) *Subspace, Latent Structure and Feature Selection.* Springer Berlin Heidelberg.

ROUSSEAU, R., GOVAERTS, B., VERLEYSEN, M. & BOULANGER, B. 2008. Comparison of some chemometric tools for metabonomics biomarker identification. *Chemometr Intell Lab Syst.,* 91**,** 54-66.

RUIZ-SAMBLAS, C., MARINI, F., CUADROS-RODRIGUEZ, L. & GONZALEZ-CASADO, A. 2012. Quantification of blending of olive oils and edible vegetable oils by triacylglycerol fingerprint gas chromatography and

chemometric tools. *J Chromatogr B Analyt Technol Biomed Life Sci.,* 910**,** 71-7.

RUZIN, S. E. 1999. Plant Microtechnique and Microscopy, New York, USA, University Press, pp.33-67.

SACRÉ, P.-Y., DECONINCK, E., DE BEER, T., COURSELLE, P., VANCAUWENBERGHE, R., CHIAP, P., CROMMEN, J. & DE BEER, J. O. 2010. Comparison and combination of spectroscopic techniques for the detection of counterfeit medicines. *J Pharm Biomed Anal.,* 53**,** 445-453.

SALES, P. M., SOUZA, P. M., SIMEONI, L. A. & SILVEIRA, D. 2012. alpha-Amylase inhibitors: a review of raw material and isolated compounds from plant source. *J Pharm Pharm Sci.,* 15**,** 141-83.

SASIC, S. & EKINS, S. 2008. *Pharmaceutical Applications of Raman Spectroscopy*, Wiley.

SCHRADER, B. 2008. *Infrared and Raman Spectroscopy: Methods and Applications*, Wiley.

SCHRADER, B., SCHULZ, H., ANDREEV, G. N., KLUMP, H. H. & SAWATZKI, J. 2000. Non-destructive NIR-FT-Raman spectroscopy of plant and animal tissues, of food and works of art. *Talanta.,* 53**,** 35-45.

SCHULZ, H. & BARANSKA, M. 2007. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vib Spectrosc.,* 43**,** 13-25.

SHERMA, J. 2012. Biennial review of planar chromatography: 2009-2011. *J AOAC Int.,* 95**,** 992-1009.

SHIBATA, S., MURAKAMI, T., NISHIKAWA, Y. & HARADA , M. 1959. The Constituents of Pueraria Root. *Chem Pharm Bull (Tokyo).* 7**,** 134-136.

SILVA, A. C., PONTES, L. F., PIMENTEL, M. F. & PONTES, M. J. 2012. Detection of adulteration in hydrated ethyl alcohol fuel using infrared spectroscopy and supervised pattern recognition methods. *Talanta.,* 93**,** 129-34.

SILVA, S. D., FELICIANO, R. P., BOAS, L. V. & BRONZE, M. R. 2014. Application of FTIR-ATR to Moscatel dessert wines for prediction of total phenolic and flavonoid contents and antioxidant capacity. *Food Chem.,* 150**,** 489-493.

SKOV, T., VAN DEN BERG, F., TOMASI, G. & BRO, R. 2006. Automated alignment of chromatographic data. *J Chemometr.,* 20**,** 484-497.

SMITH, B. C. 2011. *Fundamentals of Fourier Transform Infrared Spectroscopy, Second Edition*, Taylor & Francis, Boca Raton, USA.

SNEE, R. D. 1977. Validation of Regression Models: Methods and Examples. *Technometrics.,* 19**,** 415-428.

SNYDER, L. R., KIRKLAND, J. J. & DOLAN, J. W. 2011. *Introduction to Modern Liquid Chromatography*, Wiley.

SPANGENBERG, B., POOLE, C. F. & WEINS, C. 2011. *Quantitative Thin-Layer Chromatography: A Practical Survey*, Springer.

SRIVASTAVA, M. M. 2010. *High-Performance Thin-Layer Chromatography (HPTLC)*, Springer.

STATE FOOD AND DRUG ADMINISTRATION 2000. Status Quo of Drug Supervision in China. accessed on 4 March 2014 http://former.sfda.gov.cn/cmsweb/webportal/W205/A64028182.html.

STEVENS, L. 1976. King Kong Kudzu, menace to the south. *Smithsonian.,* 7**,** 93-98.

SUN, D. W. 2009. *Infrared Spectroscopy for Food Quality Analysis and Control*, Elsevier Science.

SUN, J. H., LI, Z., JEWETT, D. K., BRITTIN, K. O., YE, W. H. & GE, X. 2005. Genetic diversity of Pueraria lobata (kudzu) and closely related taxa as revealed by inter-simple sequence repeat analysis. *Weed Research.,* 45**,** 255-260.

SUN, Y., SHAW, P. & FUNG, K. P. 2006. Molecular authentication of Radix Puerariae Lobatae and Radix Puerariae Thomsonii by ITS and 5s rRNA spacer sequencing. *Biol Pharm Bull.,* 30**,** 173-175.

SUN, Y. G., WANG, S. S., FENG, J. T., XUE, X. Y. & LIANG, X. M. 2008. Two new isoflavone glycosides from Pueraria lobata. *J Asian Nat Prod Res.,* 10**,** 729-33.

TABOR, P. & SUSOOT, A. W. 1941. Zero to thirty mile-a-minute seedlings. *Soil Conservation.,* 8**,** 61-65.

TADERA, K., MINAMI, Y., TAKAMATSU, K. & MATSUOKA, T. 2006. Inhibition of alpha-glucosidase and alpha-amylase by flavonoids. *J Nutr Sci Vitaminol (Tokyo).* 52**,** 149-53.

TAM, W. Y., CHOOK, P., QIAO, M., CHAN, L. T., CHAN, T. Y. K., POON, Y. K., FUNG, K. P., LEUNG, P. C. & WOO, K. S. 2009. The efficacy and tolerability of adjunctive alternative herbal medicine (Salvia miltiorrhiza and Pueraria lobata) on vascular function and structure in coronary patients. *J Altern Complement Med.,* 15**,** 415-421.

TGA. 2011. *Pueraria lobata* [Online]. eBS Australian Register of Therapeutic Goods Medicine. Accessed on 7 January 2011 https://http://www.ebs.tga.gov.au/ebs/ANZTPAR/PublicWeb.nsf/publicSearch?openAgent&id=AI~pueraria lobata~0~cuMedicines?OpenView.

THAIPONG, K., BOONPRAKOB, U., CROSBY, K., CISNEROS-ZEVALLOS, L. & HAWKINS BYRNE, D. 2006. Comparison of ABTS, DPPH, FRAP, and

ORAC assays for estimating antioxidant activity from guava fruit extracts. *J Food Compost Anal.,* 19**,** 669-675.

THANASOULIAS, N. C., PARISIS, N. A. & EVMIRIDIS, N. P. 2003. Multivariate chemometrics for the forensic discrimination of blue ball-point pen inks based on their Vis spectra. *Forensic Sci Int.,* 138**,** 75-84.

THOO, Y. Y., HO, S. K., LIANG, J. Y., HO, C. W. & TAN, C. P. 2010. Effects of binary solvent extraction system, extraction time and extraction temperature on phenolic antioxidants and antioxidant capacity from mengkudu (Morinda citrifolia). *Food Chem.,* 120**,** 290-295.

TISTAERT, C., DEJAEGHER, B. & HEYDEN, Y. V. 2011a. Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Anal Chem Acta.,* 690**,** 148-161.

TISTAERT, C., THIERRY, L., SZANDRACH, A., DEJAEGHER, B., FAN, G., FREDERICH, M. & VANDER HEYDEN, Y. 2011b. Quality control of Citri reticulatae pericarpium: exploratory analysis and discrimination. *Anal Chim Acta.,* 705**,** 111-22.

TODESCHINI, R., CONSONNI, V., MANNHOLD, R., KUBINYI, H. & TIMMERMAN, H. 2008. *Handbook of Molecular Descriptors*, Wiley.

TOMASI, G., VAN DEN BERG, F. & ANDERSSON, C. 2004. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *J Chemometr.,* 18**,** 231-241.

TU, A. T., LEE, J. & MILANOVICH, F. P. 1979. Laser-Raman spectroscopic study of cyclohexaamylose and related compounds: spectral analysis and structural implications. *Carbohydrate Research.,* 76**,** 239-244.

TURNBULL, I. 2004. Kudzu: Pueraria lobata Identification and Control. Bellingen Shire Council Andrew Storrie, NSW Agriculture, Sydney, Australia.

UNG, C. Y., LI, H., CAO, Z. W., LI, Y. X. & CHEN, Y. Z. 2007. Are herb-pairs of traditional Chinese medicine distinguishable from others? Pattern analysis and artificial intelligence classification study of traditionally defined herbal properties. *J Ethnopharmacol.,* 111**,** 371-7.

UPTON, R., GRAFF, A., JOLLIFFE, G., LANGER, R. & WILLIAMSON, E. 2011. *American Herbal Pharmacopoeia: Botanical Pharmacognosy - Microscopic Characterisation of Botanical Medicines,* New York, USA, CRC Press.

US CONGRESS 1993. Harmful, Non-indigenous Species in the United States. Office of Technology Assessment, United States, OTA-F-565.

UZUNER, H., BAUER, R., FAN, T. P., GUO, D. A., DIAS, A., EL-NEZAMI, H., EFFERTH, T., WILLIAMSON, E. M., HEINRICH, M., ROBINSON, N., HYLANDS, P. J., HENDRY, B. M., CHENG, Y. C. & XU, Q. 2012. Traditional Chinese medicine research in the post-genomic era: good

practice, priorities, challenges and opportunities. *J Ethnopharmacol.,* 140**,** 458-68.

VAN DER MAESEN, L. J. G. 1985. *Revision of the genus Pueraria DC With some notes on Teyleria backer (Leguminosae)*, Agric Univ Wageningen Paper 85-1.

VAN DER MAESEN, L. J. G. 2002. Chapter 1: Pueraria: botanical characteristics. *In:* KEUNG, W. M. (ed.) *Pueraria: The genus Pueraria.* CRC Press, USA, pp. 15.

VAN SOEST, P. J. V., ROBERTSON, J. B. & LEWIS, B. A. 1991. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J Dairy Sci.,* 74**,** 3583-3597.

VAPNIK, V. N. 1999. *The Nature of Statistical Learning Theory. 2nd Ed*, CA, USA, Springer.

VARMUZA, K. & FILZMOSER, P. 2009. *Introduction to multivariate statistical analysis in chemometrics,* CRC Press, Boca Raton.

VEITCH, N. C. 2007. Isoflavonoids of the leguminosae. *Nat Prod Rep.,* 24**,** 417-64.

VEITCH, N. C. 2009. Isoflavonoids of the leguminosae. *Nat Prod Rep.,* 26**,** 776-802.

VEITCH, N. C. 2013. Isoflavonoids of the leguminosae. *Nat Prod Rep.,* 30**,** 988-1027.

VERMA, N. & KHOSA, R. L. 2012. Development of standardization parameters of Costus speciosus rhizomes with special reference to its pharmacognostical and HPTLC studies. *Asian Pac J Trop Med.,* 2**,** S276-S283.

VINCKEN, J. P., HENG, L., DE GROOT, A. & GRUPPEN, H. 2007. Saponins, classification and occurrence in the plant kingdom. *Phytochemistry.,* 68**,** 275-97. .

WALL, P. E. 2005. *Thin Layer Chromatography: A Modern Practical Approach*, Royal Society of Chemistry.

WALZ, E. 1931. Isoflavon- and Saponin-glucosides in *Soja hispida*. *Justus Liebigs Annalen der Chemie.,* 498**,** 118-155.

WANG, H. & ZHU, B. 2011. *Diagnostics of Traditional Chinese Medicine*, Jessica Kingsley Publishers.

WANG, J., VAN DER HEIJDEN, R., SPRUIT, S., HANKERMEIER, T., CHAN, K., VAN DER GREEF, J., XU, G. & WANG, M. 2009. Quality and safety of Chinese herbal medicines guided by a systems biology perspective. *J Ethnopharmacol.,* 126**,** 31-41.

WANG, L., YUAN, K. & YU, W. W. 2010. Studies of UPLC fingerprint for the identification of Magnoliae officinalis cortex processed. *Pharmacogn Mag.,* 6**,** 83-8.

WANG, S., HU, Y., TAN, W., WU, X., CHEN, R., CAO, J., CHEN, M. & WANG, Y. 2012. Compatibility art of traditional Chinese medicine: from the perspective of herb pairs. *J Ethnopharmacol.,* 143**,** 412-23.

WANG, Y. Z., FENG, W. S., SHI, R. B. & LIU, B. 2007. A new chemical component of Pueraria lobata (Willd.) Ohwi. *Acta Pharmaceutica Sinica.,* 42**,** 964-7

WARTEWIG, S. & NEUBERT, R. H. H. 2005. Pharmaceutical applications of Mid-IR and Raman spectroscopy. *Adv Drug Delivery Rev.,* 57**,** 1144-1170.

WELLING, M. 2005. Fisher linear discriminant analysis. *Department of Computer Science, University of Toronto*.

WILLIAMS, P. C. & SOBERING, D. C. 1993. Comparison of commerical near infrared transmittance and reflectance instruments for analysis whole grains and seeds. *J Near Infrared Spec.,* 1**,** 25-32.

WOLD, S. & SJOSTROM, M. 1977. *Chemometrics: Theory and Application*, American Chemical Society Symposium Series, 52 pp 243-282.

WONG, K. H., LI, G. Q., LI, K. M., RAZMOVSKI-NAUMOVSKI, V. & CHAN, K. 2011. Kudzu root: Traditional uses and potential medicinal benefits in diabetes and cardiovascular diseases. *J Ethnopharmacol.,* 134**,** 584-607.

WONG, K. H., RAZMOVSKI-NAUMOVSKI, V., LI, K. M., LI, G. Q. & CHAN, K. 2013. Differentiation of Pueraria lobata and Pueraria thomsonii using partial least square discriminant analysis (PLS-DA). *J Pharm Biomed Anal.,* 84**,** 5-13.

WORKMAN, J. & WEYER, L. 2007. *Practical Guide to Interpretive Near-Infrared Spectroscopy*, Taylor & Francis, Boca Raton,USA.

WORLD HEALTH ORGANISATION 2002. WHO traditional medicine strategy 2002-2005. accessed on 4 March 2014 http://whqlibdoc.who.int/hq/2002/WHO_EDM_TRM_2002.1.pdf?ua=1.

WORLD HEALTH ORGANISATION 2011. Quality control methods for herbal materials. 2nd edition. http://apps.who.int/medicinedocs/documents/h1791e/h1791e.pdf.

WU, D., HE, Y., NIE, P., CAO, F. & BAO, Y. 2010. Hybrid variable selection in visible and near-infrared spectral analysis for non-invasive quality determination of grape juice. *Anal Chim Acta.,* 659**,** 229-37.

WU, W., WALCZAK, B., MASSART, D. L., HEUERDING, S., ERNI, F., LAST, I. R. & PREBBLE, K. A. 1996. Artificial neural networks in classification of NIR

spectral data: Design of the training set. *Chemometr Intell Lab Syst.,* 33**,** 35-46.

XIAO, J., KAI, G., YAMAMOTO, K. & CHEN, X. 2013a. Advance in dietary polyphenols as alpha-glucosidases inhibitors: a review on structure-activity relationship aspect. *Crit Rev Food Sci Nutr.,* 53**,** 818-36.

XIAO, J., NI, X., KAI, G. & CHEN, X. 2013b. A review on structure-activity relationship of dietary polyphenols inhibiting alpha-amylase. *Crit Rev Food Sci Nutr.,* 53**,** 497-506.

XIAO, W., HAN, L. & SHI, B. 2008. Optimization of Microwave‐Assisted Extraction of Flavonoid from Radix Astragali using Response Surface Methodology. *Separ Sci Technol.,* 43**,** 671-681.

XINSHIPU. 2014. *Cooking recipes of Gegen.* [Online]. Accessed on 3 March 2014 http://www.xinshipu.com/%E5%81%9A%E6%B3%95/%E8%91%9B%E6%A0%B9/.

XIONG, F. L., SUN, X. H., GAN, L., YANG, X. L. & XU, H. B. 2006. Puerarin protects rat pancreatic islets from damage by hydrogen peroxide. *Eur J Pharmacol.,* 529**,** 1-7.

XU, J. & WU, D. 2009. Clinical research of Gegen Qinlian combined with wuling powder in treating 58 cases of acute cerebral infarction. *Fujian Journal of traditional Chinese Medicine.,* 40**,** 17-18.

XU, Q. A. 2013. *Ultra-High Performance Liquid Chromatography and Its Applications*, Wiley.

YAN, B., WANG, D. Y., XING, D. M., DING, Y., WANG, R. F., LEI, F. & DU, L. J. 2004. The antidepressant effect of ethanol extract of radix puerariae in mice exposed to cerebral ischaemia reperfusion. *Pharmacol Biochem Behav.,* 78**,** 319-25.

YANG, G. L., YANG, L. W., LI, Y. X., CAO, H., ZHOU, W. L., FANG, Z. J., ZHOU, H. B., MO, J. L., XIAO, S. X. & LIN, H. R. 2010. Applications of ultra-performance liquid chromatography to traditional Chinese medicines. *J Chromatogr Sci.,* 48**,** 18-21.

YETUKURI, L., KATAJAMAA, M., MEDINA-GOMEZ, G., SEPPANEN-LAAKSO, T., VIDAL-PUIG, A. & ORESIC, M. 2007. Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. *BMC Sys Biol.,* 1**,** 12.

YU, H. Y., NIU, X. Y., LIN, H. J., YING, Y. B., LI, B. B. & PAN, X. X. 2009. A feasibility study on on-line determination of rice wine composition by Vis–NIR spectroscopy and least-squares support vector machines. *Food Chem.,* 113**,** 291-296.

YU, Q. 2007. The clinical research of improve symptom of hypertension with Gegen Tianma Gouteng decoction. *Guang Ming Journal of traditional Chinese Medicine.,* 22**,** 50-52

YU, Z. & LI, W. 2006. Induction of apoptosis by puerarin in colon cancer HT-29 cells. *Cancer Lett.,* 238**,** 53-60.

ZADEH, L. A. 1965. Fuzzy sets. *Information and Control.,* 8**,** 338-353.

ZANG, Q., KEIRE, D., BUHSE, L., WOOD, R., MITAL, D., HAQUE, S., SRINIVASAN, S., MOORE, C. V., NASR, M., AL-HAKIM, A., TREHY, M. & WELSH, W. 2011. Identification of heparin samples that contain impurities or contaminants by chemometric pattern recognition analysis of proton NMR spectral data. *Anal Bioanal Chem.,* 401**,** 939-955.

ZENG, Y. P., HUANG, Y. S. & HU, Y. G. 2006. Effect of gegen qinlian decoction combined with short-term intensive insulin treatment on patients with type 2 diabetes mellitus of dampness-heat syndrome. *Zhongguo Zhong Xi Yi Jie He Za Zhi,* 26**,** 514-6.

ZHANG, C. L., DING, X. P., HU, Z. F., WANG, X. T., CHEN, L. L., QI, J. & YU, B. Y. 2011a. Comparative study of Puerariae lobatae and Puerariae thomsonii by HPLC-diode array detection-flow injection-chemiluminescence coupled with HPLC-electrospray ionization-MS. *Chem Pharm Bull (Tokyo).* 59**,** 541-5.

ZHANG, C. Y., DONG, L., WANG, J. & CHEN, S. L. 2011b. Simultaneous determination of ten ginsenosides in panacis quinquefolii radix by ultra performance liquid chromatography and quality evaluation based on chemometric methods. *Pharmazie.,* 66**,** 553-9.

ZHANG, H. J., YANG, X. P. & WANG, K. W. 2010. Isolation of two new C-glucofuranosyl isoflavones from Pueraria lobata (Wild.) Ohwi with HPLC-MS guiding analysis. *J Asian Nat Prod Res.,* 12**,** 293-9.

ZHANG, J. & QU, F. 2013. Methods for Analysis of Triterpenoid Saponins. *In:* RAMAWAT, K. G. & MÉRILLON, J.-M. (eds.) *Natural Products.* Springer Berlin Heidelberg.

ZHANG, M. H., LUYPAERT, J., FERNANDEZ PIERNA, J. A., XU, Q. S. & MASSART, D. L. 2004. Determination of total antioxidant capacity in green tea by near-infrared spectroscopy and multivariate calibration. *Talanta.,* 62**,** 25-35.

ZHANG, X.-F., ZOU, M.-Q., QI, X.-H., LIU, F., ZHANG, C. & YIN, F. 2011c. Quantitative detection of adulterated olive oil by Raman spectroscopy and chemometrics. *J Raman Spectrosc.,* 42**,** 1784-1788.

ZHAO, C., CHAN, H. Y., YUAN, D., LIANG, Y., LAU, T. Y. & CHAU, F. T. 2011. Rapid simultaneous determination of major isoflavones of Pueraria lobata and discriminative analysis of its geographical origins by principal component analysis. *Phytochem Anal.,* 22**,** 503-8.

ZHAO, T., HAN, J., CHEN, Y., WAN, H. & BIE, X. 2007. The mechanism of 3-methoxy puerarin on decreasing the cerebral ischemia-reperfusion injury in rats. *Asia Pac J Clin Nutr.,* 16**,** 302-4.

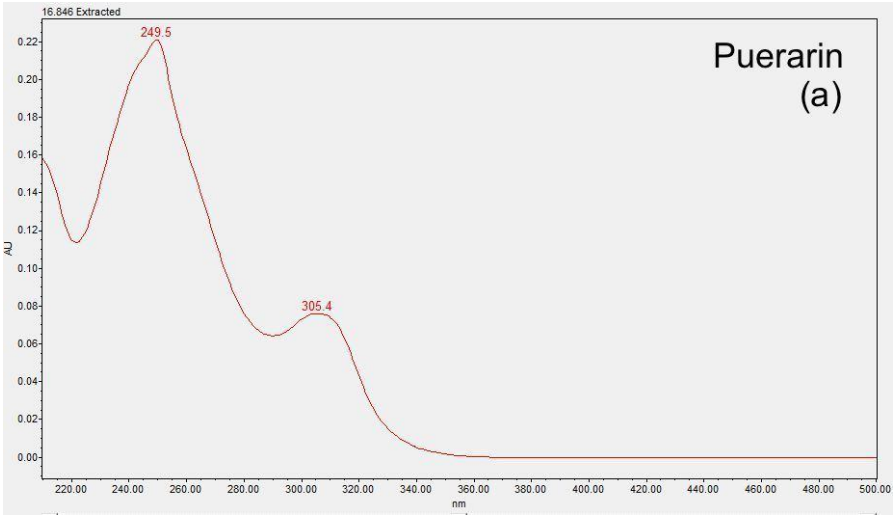
ZHONG, G.-X., LI, P., ZENG, L.-J., GUAN, J., LI, D.-Q. & LI, S.-P. 2009. Chemical Characteristics of Salvia miltiorrhiza (Danshen) Collected from Different Locations in China. *Journal of Agricultural and Food Chemistry,* 57**,** 6879-6887.

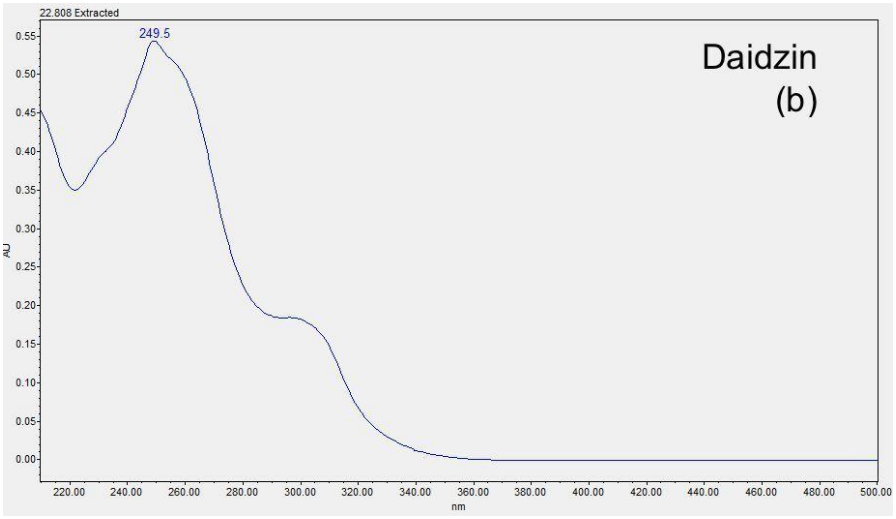
ZHU, B. & WANG, H. 2011. *Basic Theories of Traditional Chinese Medicine*, Jessica Kingsley Publishers.

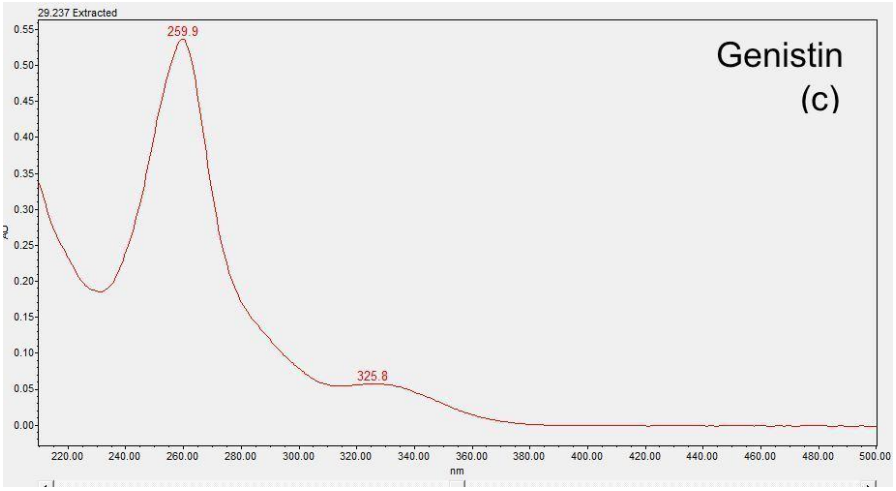
ZHU, Y.-P. & WOERDENBAG, H. 1995. Traditional Chinese herbal medicine. *Pharm World Sci,* 17**,** 103-112.

Appendices

Appendix I







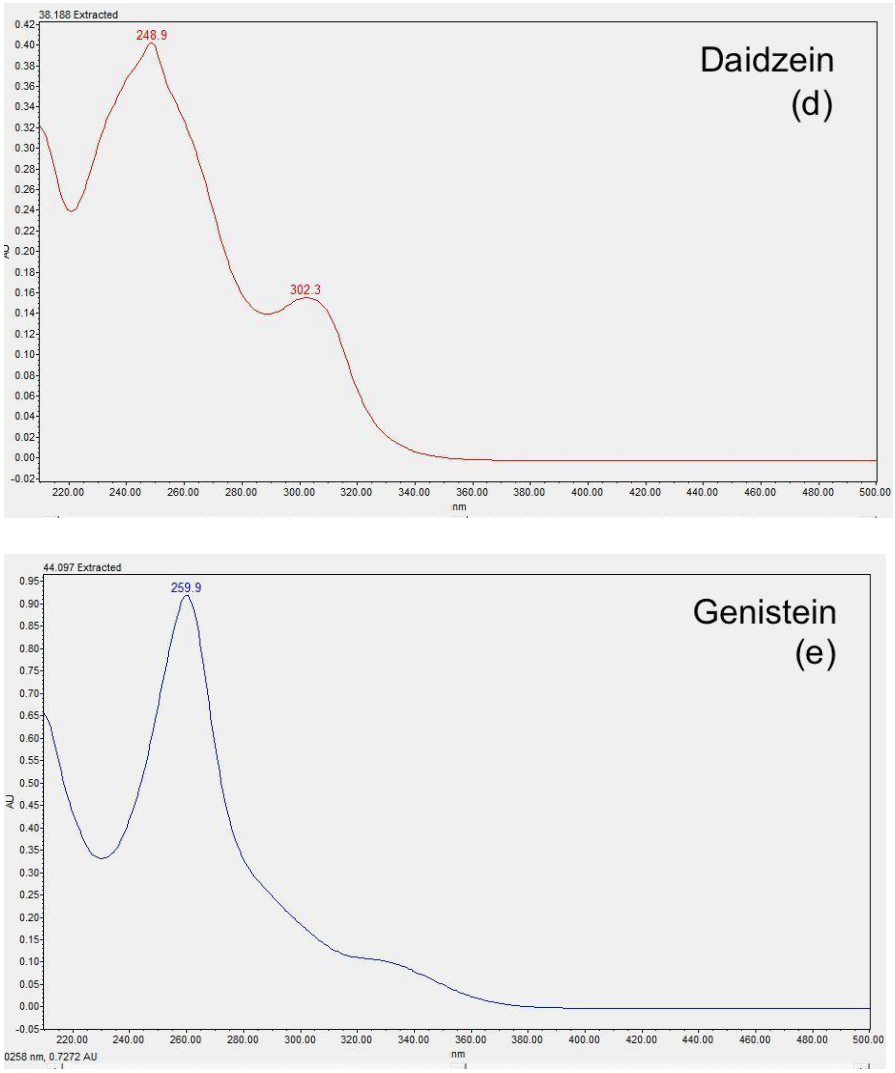
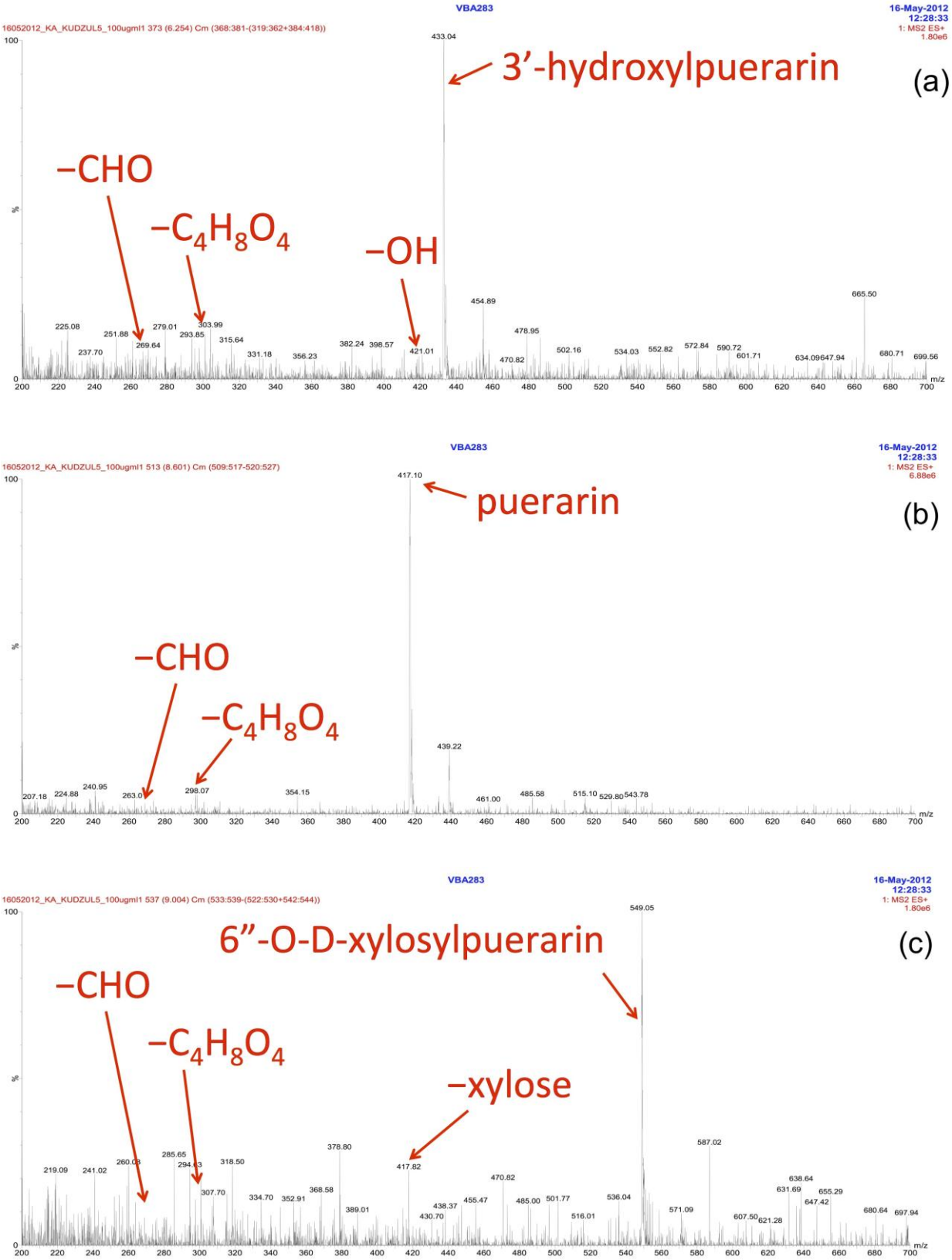
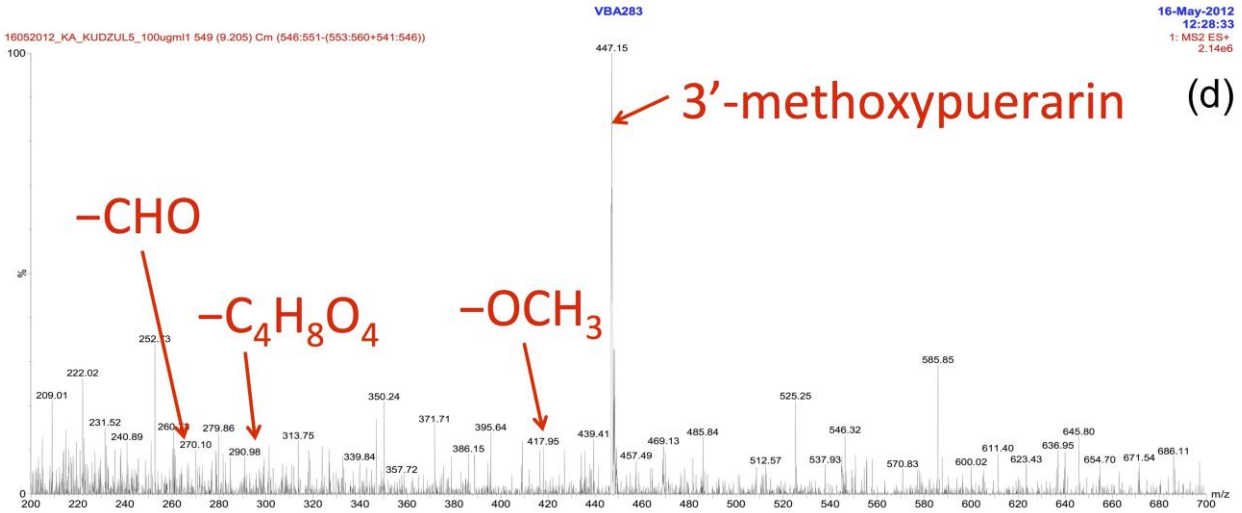
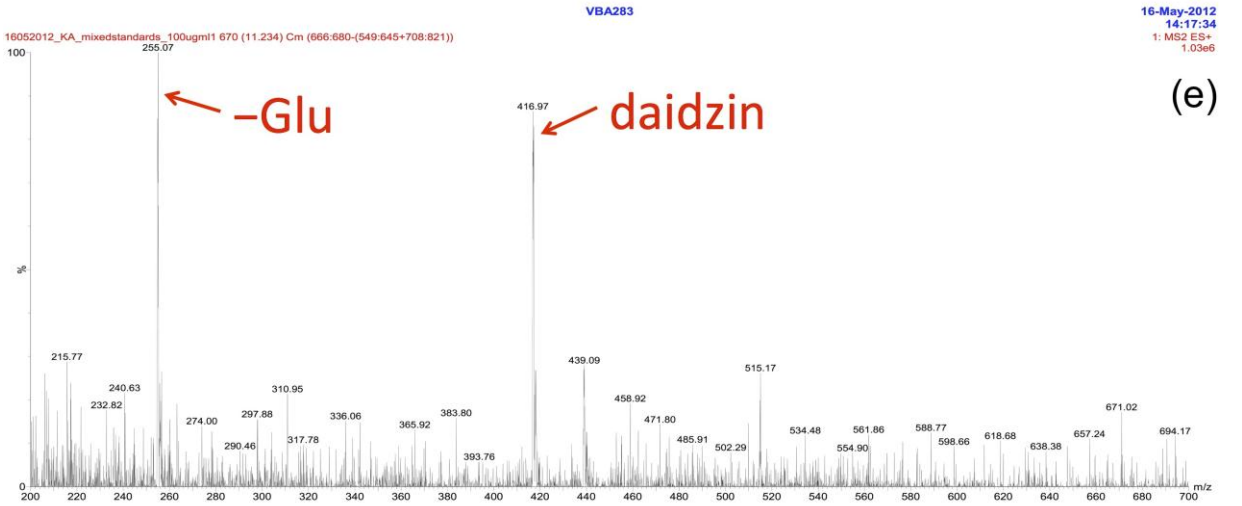


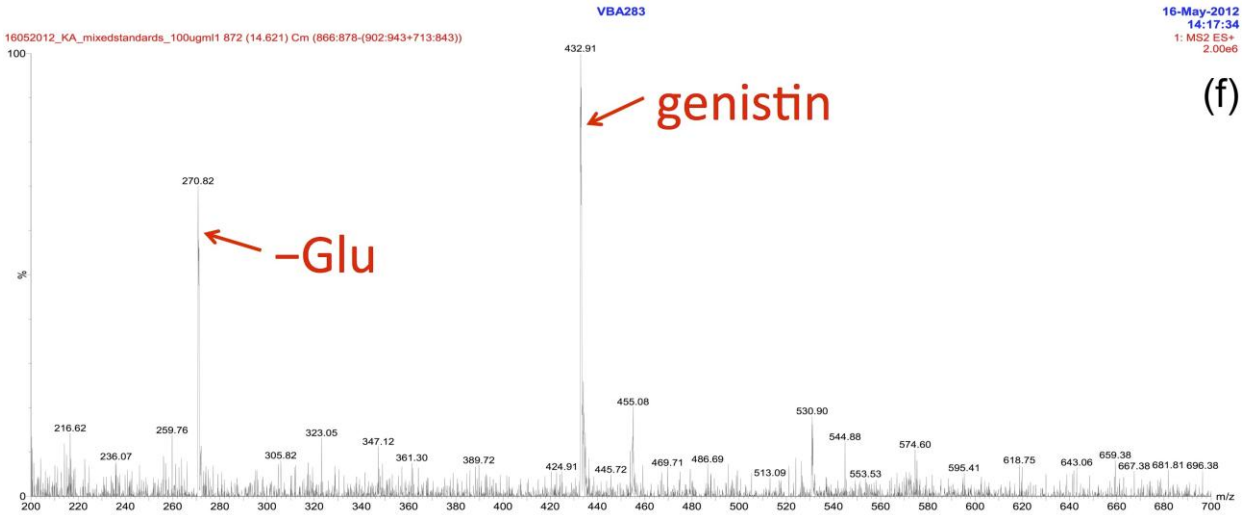
Figure A.3.1 Spectra of reference compounds detected at 245 nm using ultra-performance liquid chromatography coupled with photodiode array. (a) Puerarin; (b) Daidzin; (c) Genistin; (d) Daidzein; (e) Genistein.

Appendix II


VBA283
16-May-2012
12:28:33
1: MS2 ES+
1.80e6
16052012_KA_KUDZUL5_100ugml1 373 (6.254) Cm (368:381-(319:362+384:418))
3'-hydroxylpuerarin
(a)
–CHO
–C4H8O4
–OH
VBA283
16-May-2012
12:28:33
1: MS2 ES+
6.88e6
16052012_KA_KUDZUL5_100ugml1 513 (8.601) Cm (509:517-520:527)
puerarin
(b)
–CHO
–C4H8O4
VBA283
16-May-2012
12:28:33
1: MS2 ES+
1.80e6
16052012_KA_KUDZUL5_100ugml1 537 (9.004) Cm (533:539-(522:530+542:544))
6"-O-D-xylosylpuerarin
(c)
–CHO
–C4H8O4
–xylose


VBA283
16-May-2012
12:28:33
1: MS2 ES+
2.14e6
16052012_KA_KUDZUL5_100ugml1 549 (9.205) Cm (546:551-(553:560+541:546))
3'-methoxypuerarin
(d)
–CHO
$-C_4H_8O_4$
$-OCH_3$
m/z


VBA283
16-May-2012
14:17:34
1: MS2 ES+
1.03e6
16052012_KA_mixedstandards_100ugml1 670 (11.234) Cm (666:680-(549:645+708:821))
–Glu
daidzin
(e)
m/z


VBA283
16-May-2012
14:17:34
1: MS2 ES+
2.00e6
16052012_KA_mixedstandards_100ugml1 872 (14.621) Cm (866:878-(902:943+713:843))
genistin
(f)
–Glu
m/z

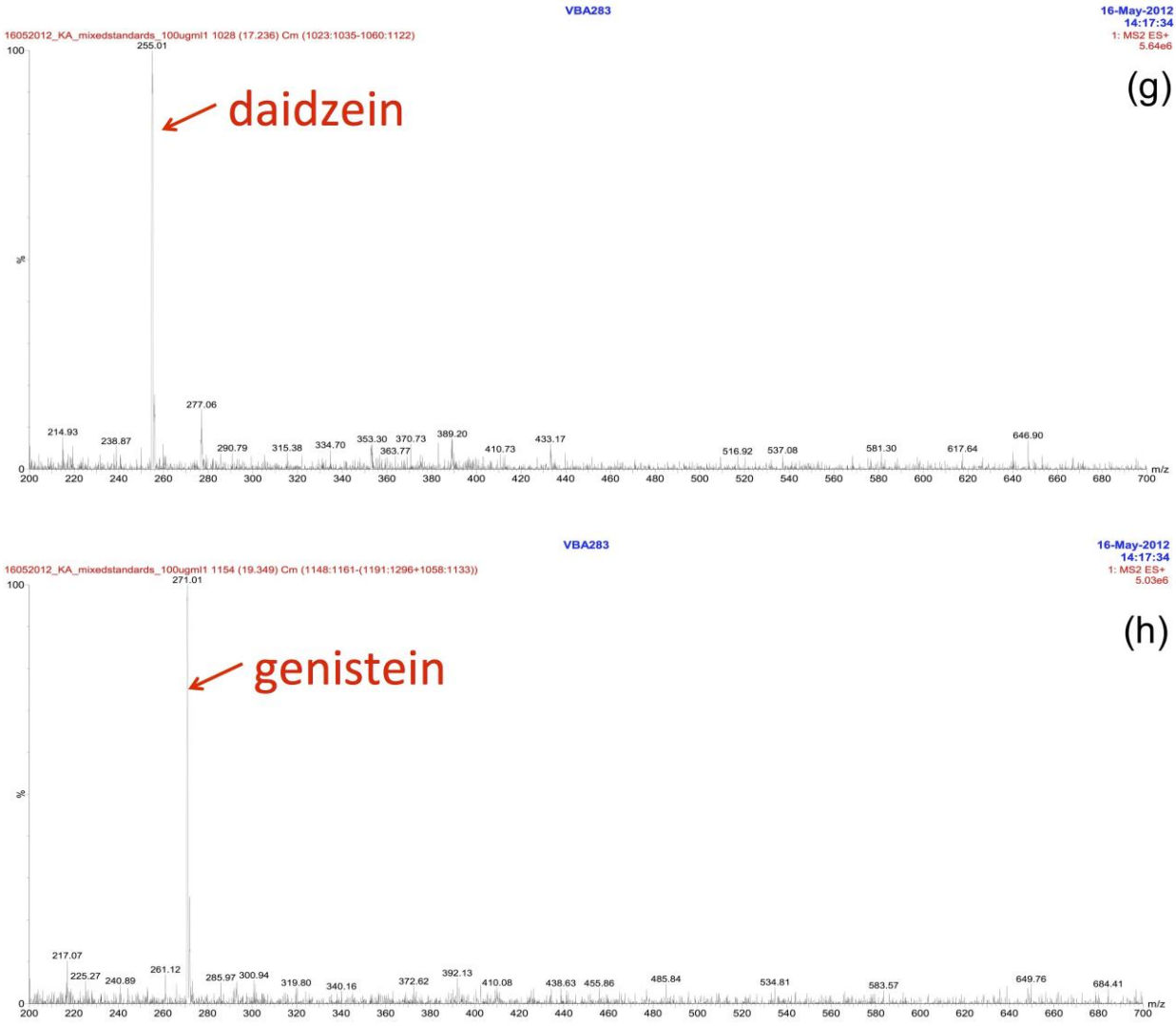


Figure A.3.2 Total ion chromatograms of chemical constituents detected in ESI positive ion mode using ultra-performance liquid chromatography-mass spectroscopy. (a) 3'-Hydroxypuerarin; (b) Puerarin; (c) 6'-O-D-xylosylpuerarin; (d) 3'-Methoxypuerarin; (e) Daidzin; (f) Genistin; (g) Daidzein; (h) Genistein.