Quantitative analysis and metabonomic study of phytoestrogens in Africans

by

Olumuyiwa Lateefah Ogegbo B.Sc. (Hons) M.Sc. AMRSC

Institute for Health Research and Policy

A thesis submitted in partial fulfilment of the requirements of London Metropolitan University for the degree of Doctor of Philosophy

May 2009

Abstract

Phytoestrogens have been widely studied in various parts of the world. However, little is known about the availability, absorption, metabolism and its biochemical effects in Africans. The principal aim of this research was to investigate the phytoestrogen content in six African staple foods, their metabolism and biochemical effects after soy-intervention in Africans.

With respect to the food investigations, sensitive and reliable analytical methods for the simultaneous quantification of seven phytoestrogens using GC-MS, LC-MS, UPLC-MS were developed. The phytoestrogens analysed include four isoflavones (formononetin, biochanin A, daidzein, genistein), one coumestan (coumestrol), and two lignans (secoisolariciresinol and matairesinol). The methods were applied to quantitate the phytoestrogen content in six African staple foods: beans, cassava, plantain, pumpkin leaves '*ugu*', rice and yam. The results showed that pumpkin leaves contained the highest amount of phytoestrogens analysed (5.66 – 12.68 μg/g) whilst rice contaned the least amount of phytoestrogens (0.58 – 1.17 μg/g). In addition, three phytoestrogenic compounds: 3,5,7- trihydroxy-2-(4-hydroxyphenyl)-4*H*-chromen-4-one, 4-(hydroxyl(4-(4-hydroxy-3-methoxy-benzyl)tetrahydrofuran-3-yl)-2-methoxyphenol and 3-(4-hydroxy-3-methoxybenzyl)-4-((4-hydroxy-3-methoxyphenyl)-(methoxy)methyl)di-hydrofuran-2(3*H*)-one, were found in pumpkin leaves for the first time.

The second part of this research was based on a soy – intervention, where African subjects; 25 UK-based and 25 Nigerian - based were recruited and fed 200 mL Alpro soy-milk (~20 mg phytoestrogen) daily over a five - day period. The urine samples collected over 2-weeks (inclusive of the 5 consumption days) were used as biomarkers media in assessing the phytoestrogen ingestion and metabolism. The biomarkers monitored and analysed were daidzein, genistein and their metabolites: equol and o-desmethylangolensin. The results showed that 35% of the selected population are 'good responders' to soy - phytoestrogens being equol and o-DMA producers. In addition, results showed metabolism was influenced by certain factors including location, gender and age.

^{1}H NMR-metabonomics was performed on the collected bio-fluids to investigate soy – phytoestrogens' effects on specific endogenous metabolites involved in major metabolic processes. Twenty-two endogenous metabolites were quantified using Chenomx® NMR Suite and statistically analysed using multivariate analyses: Principal Component Analysis (PCA) and (orthogonal-) Partial-Least Square - Discriminant

Analysis ((O-)PLS-DA). The results showed the levels of endogenous metabolites present in urine ranged from 5 μM to 15 mM with huge inter – individuality variances in acetone, citrate, creatinine, glycine, hippurate and trimethylamine-N-oxide. With the MVA, results showed high inter – individuality and inter – sampling variances based on the PCA plots, however with a definitive MVA: (O)-PLS-DA, results clearly showed metabolism to be mainly influenced by location and gender.

From the exploration of the soy - challenge and metabonomics results, correlation between soy - phytoestrogen, their metabolites and endogenous metabolites were investigated using Pearson product – moment correlation. The results showed that there were significant correlations ($p<0.05$) between soy - phytoestrogens and their metabolites as well as between the endogenous metabolites. On the other hand, no distinct correlation was observed between either the soy – phytoestrogens or their metabolites and endogenous metabolites.

Table of Contents

Abstract i

List of Figures viii

List of Tables xiii

Abbreviations and Symbols xv

Acknowledgements xvii

Dedication xix

Declaration xx

Phytoestrogens and their metabolites 1

1.1 Introduction 1

1.2 Chronological history of phytoestrogen research 1

1.3 Phytoestrogen occurrence 3

1.3.1 Definition of phytoestrogens 3

1.3.2 Classification of phytoestrogens 4

1.3.2.1 Isoflavones 7

1.3.2.2 Coumestans 9

1.3.2.3 Lignans 10

1.3.3 Chemical ecology of phytoestrogens in plants 13

1.3.4 Occurrence in food 14

1.4 Exposure to phytoestrogens 17

1.5 Phytoestrogen metabolism 18

1.5.1 Phytoestrogen metabolites 18

1.5.1.1 Intestinal microflora metabolism pathway of phytoestrogen 18

1.5.1.2 Metabolic fate of isoflavones in soy products 22

1.6 Phytoestrogen analyses 24

1.6.1 Extraction protocols 24

1.6.1.1 Food or plant samples 24

1.6.1.2 Biological samples 28

1.6.2 Analytical techniques: chromatography & spectrometry 28

1.6.2.1 Gas chromatography 30

1.6.2.2 Liquid chromatography 31

1.6.2.3 Hyphenated chromatographic techniques 33

1.6.2.4 Spectrometric techniques 34

1.6.3 Analysis of phytoestrogens 36

1.6.3.1 Chromatographic separation techniques 37
1.6.3.2 Non - chromatographic separation techniques 39
1.7 Phytoestrogen metabonomics 40
1.7.1 Introduction of metabonomics 40
1.7.2 Metabonomics in human nutrition 44
1.7.3 Biochemical effects of phytoestrogen on other metabolites 47
1.7.4 Metabonomics in physiological monitoring of human nutritional diet 47
1.7.5 Endogenous metabolites 48
1.8 Hypothesis 51
1.9 Aims and Objectives 52
Phytoestrogens in some African staple foods 54
2.1 Introduction 54
2.2 Results 56
2.2.1 Chromatographic analyses of phytoestrogens 56
2.2.2 Method validation for quantitative analyses 59
2.2.3 Quantitation of phytoestrogen contents in food samples 60
2.2.4 Identification of novel phytoestrogenic compounds in an African staple food 60
2.3 Discussion 63
2.3.1 Design of experiment: Samples and sample preparation 63
2.3.2 Chromatographic separation 64
2.3.3 Detection methods and conditions 65
2.3.4 Method validation 66
2.3.5 Quantitation of African food analysis 67
2.3.6 Identification of new phytoestrogenic compounds in pumpkin leaves '*ugu*' 70
2.4 Conclusions 75
2.5 Experimental 76
2.5.1 Chemicals, solvents & materials 76
2.5.1.1 Chemicals 76
2.5.1.2 Solvents 76
2.5.1.3 Materials: food samples 76
2.5.2 Instrumentation 77
2.5.2.1 General 77
2.5.2.2 GC-FID 77
2.5.2.3 GC-MS 77
2.5.2.4 LC-MS 78

2.5.3 Mass spectrometry analysis 79
2.5.3.1 Standard calibration 79
2.5.3.2 Sample analysis 79
2.5.4 Sample preparation 79
2.5.4.1 Extraction of phytoestrogens: Liquid (soxhlet) extraction 79
2.5.4.2 Enzymatic hydrolysis 80
2.5.4.3 Acid hydrolysis 80
2.5.4.4 Purification/ isolation of phytoestrogen aglycones 80
2.5.4.5 Derivatisation of samples prior GC analyses 80
Absorption and metabolism of soy - phytoestrogens 81
3.1 Introduction 81
3.2 Results 83
3.2.1 Chromatographic analyses of phytoestrogens and their metabolites 83
3.2.2 Method validation for quantitative analyses 84
3.2.3 Quantitation of phytoestrogens- Calibration plots 85
3.2.4 Recovery 86
3.2.5 General changes in phytoestrogen metabolic profile 86
3.2.6 Factors influencing soy metabolism 96
3.3 Discussion 104
3.3.1 Recruitment of subjects, sample collection 104
3.3.2 Sample clean-up comparison between LLE and SPE 105
3.3.3 Chromatographic separation – method development 105
3.3.4 Spectrometric separation and method development 106
3.3.5 Method validation: limits of detection and quantitation 106
3.3.6 Intervention diet sample analysis 106
3.3.7 LC-MS metabolic profiling 107
3.3.7.1 DAY 1: 1st sampling- baseline 'control' pre- intervention (PHASE I) 107
3.3.7.2 DAY 3: 2nd sampling- During interventional diet (PHASE II) 109
3.3.7.3 DAY 5: 3rd sampling- During interventional diet (PHASE III) 110
3.3.7.4 DAY 12: 4th sampling - Post diet intervention (PHASE IV) 110
3.3.7.5 Factors influencing soy metabolism 111
3.4 Conclusions 111
3.5 Experimental 113
3.5.1 Chemicals, solvents & materials 113
3.5.1.1 Chemicals/ standards 113

3.5.1.2 Solvents 113
3.5.1.3 Materials 113
3.5.2 Instrumentation 113
3.5.3 Design of experiment 114
3.5.3.1 Ethical approval 114
3.5.3.2 Subjects recruitment 115
3.5.3.3. Study preparations 116
3.5.3.4. Sample collection 117
3.5.4 Methodolgy 117
3.5.4.1 Extraction of phytoestrogens in Alpro® soymilk 117
3.5.4.2 Extraction of conjugated phytoestrogens in urine samples 117
3.5.4.3 Recovery- extraction of unconjugated phytoestrogens 118
3.5.4.4 Method development using LC-MS 118
3.5.4.5. Sample analysis on LC-MS-SIM 119
^{1}H-NMR-based metabonomics of soy - phytoestrogens. 120
4.1 Introduction 120
4.2 Results 123
4.2.1 General changes in spectra following intervention 123
4.2.2 Quantitation of endogenous metabolites 126
4.2.2.1 Phase I (Day 1) soy intervention: Pre-Soy/ Control 129
4.2.2.2 Comparative study of metabolic profile changes in intervention 133
4.2.3 Multivariate statistical analysis (MVA) of the whole cohort (UK and Nigeria) ... 136
4.2.3.1 Whole spectrum binning (standard binning techniques) 136
4.2.3.2 Proposed factors influencing metabolic changes following soy intervention 140
4.2.3.3 Target binning 149
4.3 Discussion 153
4.3.1 Quantitation of endogenous metabolites 153
4.3.2 Visual Interpretation and Chenomx® quantitation 153
4.3.3 Biochemistry of metabolites- overall biochemical effect of phytoestrogen 154
4.3.4 Multivariate analysis interpretation 155
4.3.4.1 Influencing factors in metabolic biochemical effects of soy-metabolism 155
4.3.4.2 Target profiling & binning 156
4.4 Conclusions 157
4.5 Experimental 158
4.5.1 Chemicals and materials 158

4.5.1.1 Chemicals 158
4.5.1.2 Material: descriptive information of interventional diet 158
4.5.2 Design of experiment: subjects and sample collection 158
4.5.3 Sample preparation 160
4.5.4 Analysis: ^{1}H-NMR spectroscopy 161
4.5.5 Identification and quantitation of endogenous metabolites 161
4.5.6 Data reduction (binning) prior multivariate statistical analysis 163
4.5.7 Multivariate statistical analysis (MVA) 164
Correlation-ship of metabolites 166
5.1 Introduction 166
5.2 Results & Discussion 167
5.2.1 Correlation-ship in soy - phytoestrogen metabolites 167
5.2.2 Pearson correlation in endogenous metabolites 174
5.2.3 Correlation-ship of soy - phytoestrogen metabolites and endogenous metabolites 178
5.3 Conclusions 181
5.4 Experimental 181
5.4.1 Statistical data analysis 181
Conclusions 182
References 184

List of Figures

FIGURE 1.1: Structural-function similarities of flavonoidal phytoestrogen with estradiol4

FIGURE 1.2: General classification & examples of phytoestrogens (adopted from (*84*)) 4

FIGURE 1.3: Structural similarities of five phytoestrogen groups 5

FIGURE 1.4: Structural similarities in flavonoids - isoflavone, coumestan and prenylflavonoids 6

FIGURE 1.5: Structural similarities in non-flavonoids - lignan and stilbene.....7

FIGURE 1.6: Isoflavone skeletal structure.....7

FIGURE 1.7: Structures of some examples of isoflavones isolated from various plants (*18, 46, 84, 91, 92*).....8

FIGURE 1.8: Coumestan skeletal structure.....9

FIGURE 1.9: Structures of some examples of coumestans isolated from plants (*84, 91*) 10

FIGURE 1.10: Dimerisation biosynthesis of lignan 10

FIGURE 1.11: Structures of some examples of lignans isolated from plants (*30, 96-99*) 12

FIGURE 1.12: Biosynthetic production of isoflavones (adopted from (*106*) 14

FIGURE 1.13: Gut/ Intestinal metabolism of different phytoestrogens.....19

FIGURE 1.14: Relationship between phytoestrogens groups and some examples 20

FIGURE 1.15: Structural similarities between some phytoestrogens, metabolites and human endogenous estrosfagen.....21

FIGURE 1.16: Schematic metabolic fate of daidzin (adopted from (*37, 38, 133*)).....23

FIGURE 1.17: Schematic diagram of an electrospray. 36

FIGURE 1.18: NMR/ Chenomx - based metabonomics flowchart.....44

FIGURE 2.1: Structures of the seven phytoestrogens analysed. 55

FIGURE 2.2: GC chromatograms of seven standards' TMS derivatives at 0.5 ng/µL each. 56

FIGURE 2.3: LC-MS (SIM and TIC) chromatograms of seven standards at 1 ng/µL each. 57

FIGURE 2.4: UPLC-MS (SIM and TIC) chromatograms at 1 ng/µL each.....58

FIGURE 2.5: GC-MS chromatogram and mass spectra of three new phytoestrogenic compounds in pumpkin leaves 'ugu' (inset). 62

FIGURE 2.6: Structures of novel phytoestrogenic compounds in pumpkin leaves 'ugu' 70

FIGURE 2.7: LC-MS-MS spectra of two new phytoestrogenic compounds found in pumpkin leaves 'ugu'. (a) Compound 1 and (b) compound 3.....71

FIGURE 2.8: Proposed Scheme I: Proposed fragmentation pattern of $[M-H]^-$ ion of compound 1. rDA: retro-Diels Alder reaction (retrocyclisation) 73

FIGURE 2.9: Scheme II- Proposed fragmentation pattern of $[M-H]^-$ ion of compound 3. *r rearrangement of electrons/ ions/ atoms.....74

FIGURE 3.1: Phytoestrogens and metabolites analysed in this study 82

FIGURE 3.2: LC-MS-SIM chromatogram of mixture of 11 standards at 100 pg/µL.....83

FIGURE 3.3: Typical LC-MS (TIC and SIM) overlaid chromatogram on a subject's urine.....87

FIGURE 3.4: Representative plot of six metabolites excretion levels in African subjects. 89

FIGURE 3.5: Metabolic profile changes of soy phytoestrogens and metabolites in the general cohort. 92

FIGURE 3.6: Comparison plots of changes in urinary excretion levels of phytoestrogen 93

FIGURE 3.7: Taxonomy of hypothesised factors influencing daidzein excretion during soy consumption. 96

FIGURE 3.8: Age - influence plot on daidzein excretion following location and gender classification....97

FIGURE 3.9: Taxonomy of hypothesised factors influencing genistein excretion (during soy consumption) 98

FIGURE 3.10: Age - influence plot on genistein excretion following location and gender classification....99

FIGURE 3.11: Taxonomy of hypothesised factors influencing equol production and excretion during soy-milk consumption (Phases II & III). 100

FIGURE 3.12: Age - influence plot on equol production and excretion following location and gender classification....101

FIGURE 3.13: Taxonomy of hypothesised factors influencing o-DMA production and excretion (during soy consumption)....102

FIGURE 3.14: Age - influence plot on o-DMA production following location and gender classification....103

FIGURE 4.1: Structures of the endogenous metabolites analysed in this study and reference compound (TSP-d_4)....122

FIGURE 4.2: a. Typical 1H-NMR spectrum (δ: 1.2- 3) region of a subject's urine.....123

FIGURE 4.3: ^{1}H-NMR spectra inter-sampling comparison of metabolic profile changes. 125

FIGURE 4.4: ^{1}H-NMR spectra inter-subject comparison of metabolic profile changes. 126

FIGURE 4.5: Annotations of chemical shift (predicted) in endogenous metabolites...128

FIGURE 4.6: Bar-Chart of all subjects' metabolites concentrations (normalised) excluding urea and creatinine.....131

FIGURE 4.7: Metabolic effects of soy - phytoestrogen on endogenous metabolites over two-weeks (expressed as % difference) excluding urea. 134

FIGURE 4.8: PCA-X (PCA-Class) plots of Phase I-4.....137

FIGURE 4.9: PLS-DA Plot of whole data excluding outliers.....138

FIGURE 4.10: O-PLS-DA Plots of whole data excluding outliers.....139

FIGURE 4.11: PCA-X comparison of location classification based on 2D-scoreplots.141

FIGURE 4.12: Supervised MVA (PLS) plots of Phase I location difference dataset (δ: 4.6 – 5.3 excluded).....142

FIGURE 4.13: Supervised MVA (O-PLS-DA) 3D- score plot of Phase I-IV of soy intervention comparison based on location.....143

FIGURE 4.14: Comparison of gender using three MVA -3D score plots (δ: 0-9).....145

FIGURE 4.15: Supervised MVA (O-PLS-DA) 2D- score plot of groupings based on location & gender. 146

FIGURE 4.16: Comparison of three MVA in 2D- score plots of age y-variable dataset (δ: 0-9) 147

FIGURE 4.17: Supervised MVA (O-PLS-DA) 2D- score plot of Phase I-IV of soy-intervention based on location & gender differences. 149

FIGURE 4.18: Unsupervised MVA (PCA-X) of Chenomx® profiled dataset.....150

FIGURE 4.19: Supervised MVA (PLS-DA) of Chenomx® profiled dataset.....151

FIGURE 4.20: Supervised MVA (O-PLS-DA) of Chenomx® profiled dataset.....152

FIGURE 4.21: Example of superimposition of Lorentzian peak shape models of reference on raw ^{1}H-NMR data. 162

FIGURE 4.22: Typical Chenomx® Profile page 163

FIGURE 5.1: Matrix-Plot of soy - phytoestrogens and their metabolites within UK cohort 168

FIGURE 5.2: Matrix-Plot of soy - phytoestrogens and their metabolites within NIGERIA cohort.....169

FIGURE 5.3: Matrix-Plot of soy - phytoestrogens and their metabolites within UK cohort- Day 12 (post-soy).....171

FIGURE 5.4: Matrix-Plot of o-DMA and daidzein within UK-female group- Day 12 (post-soy).....172

FIGURE 5.5: Matrix-Plot of soy - phytoestrogens and their metabolites within UK-female group- Day 12 (post-soy). 172

FIGURE 5.6: Matrix-Plot of o-DMA-daidzein within NIGERIA-female group- Day 12 (post-soy).....173

FIGURE 5.7: Matrix-Plot of some endogenous metabolites on Day 3.....177

FIGURE 5.8: Matrix Plots of two selected metabolite pairs.....179

FIGURE 5.9: Matrix plot of some metabolites correlation in NIGERIA cohort 180

FIGURE 5.10: Matrix-Plot of equol-creatine phosphate in UK cohort.....180

List of Tables

TABLE 1.1: Chronological key findings/ discoveries of phytoestrogen.....2

TABLE 1.2: Literature of some foods analysed in different locations. 16

TABLE 1.3: Summarised extraction protocols from literature.....26

TABLE 1. 4: Metaboniomics studies reported in literature 46

TABLE 2.1: Recovery of phytoestrogen standards analysed by GC-MS and LC-MS...59

TABLE 2.2: Evaluation of precision by intra-assay and inter-assay (SD values).....59

TABLE 2.3: Limits of detection (LOD) and quantitation (LOQ) of phytoestrogen standards 59

TABLE 2.4: Phytoestrogen levels in African foods and dehydrated soy as reference ..61

TABLE 3.1: Precision calculation using internal standards. 84

TABLE 3.2: a. Test for Method Linearity (Method A) 85

TABLE 3.3: Recovery of phytoestrogens and their metabolites in spiked 'blank'. 86

TABLE 3.4: Phytoestrogen levels in urine collected at different Phases of soy intervention 94

TABLE 4.1: Annotation of 1H-NMR signal chemical shift and multiplicity.....127

TABLE 4.2: Concentration (mM) levels of metabolites in location- and gender- grouped control samples (Phase I intervention) 132

TABLE 4.3: Metabolic changes following soy-milk intervention (Days 3-12).....135

TABLE 4.4: Summary of model 4.14 R^2 (cum) & Q^2 (cum) values 144

TABLE 4.5: Summary of model 4.15 R^2 (cum) & Q^2 (cum) values 148

TABLE 4.6: Summary of model Chenomx®- profiled dataset: R2 (cum) & Q2 (cum) values 152

TABLE 4.7: Characteristics of study groups 160

TABLE 5.1: Pearson correlation of Δ soy - phytoestrogens and their metabolites (General cohort): Day 3.....167

TABLE 5.2: Pearson correlation of Δ soy - phytoestrogens and their metabolites (Location Classified).....168

TABLE 5.3: Pearson correlation of Δ soy - phytoestrogens and their metabolites (General cohort): Day 12 (post-soy) 170

TABLE 5.4: Pearson correlation of Δ soy - phytoestrogens and their metabolites (Location Classified): Day 12 (post-soy).....170

TABLE 5.5: Pearson correlation of Δ soy - phytoestrogens and their metabolites (UK-Male group): Day 12 (post-soy) 173

TABLE 5.6: Pearson correlation (r- value; p-value) of endogenous metabolites- Day 3 175

TABLE 5.7: Pearson correlation of soy - and endogenous – metabolites on Day 3 ...178

Abbreviations and Symbols

<	Less than
>	Greater than
°C	Degree celsius
^{1}H-NMR	Proton nuclear magnetic resonance
AMIX	Analysis of mixtures
BSTFA	N,O- bis (trimethylsilyl) trifluoroacetamide
CE	Capillary electrophoresis
CoA	Coenzyme A
D_2O	Deuterated water
DAD	Diode array detector
DEAE	Diethyl amino ethyl
ER - β	Estrogen receptor - beta
ESI	Electrospray ionisation
EtOH	Ethanol
FID	Flame ionisation detector
GC	Gas chromatography
GC-MS	Gas chromatography - mass spectrometry
HPLC	High performance liquid chromatography
h	Hour
Hz	Hertz
ID-GC-MS	Isotopic dilution gas chromatography - mass spectrometry
LC	Liquid chromatography
LC-MS	Liquid chromatography - mass Spectrometry
LLE	Liquid - liquid extraction
LOD	Limit of detection
LOQ	Limit of quantitation
m/z	mass-to-charge ratio
MALDI	Matrix - assisted laser desorption ionisation
MeOH	Methanol
min	minute
MS	Mass spectrometry
MVA	Multivariate analysis
MW	Molecular weight
NMR	Nuclear magnetic resonance

o-DMA	ortho - Desmethylangolensin
O-PLS-DA	Orthogonal - partial least square - discriminant analysis
PCA	Principal component analysis
PLS	Partial least square
PLS-DA	Partial least square - discriminant analysis
p-value	Probability value
qTOF	Quadrupole – time - of - flight
r	Pearson correlation co-efficient
r^2	Co-efficient of determination
rDA	retro Diels Alder
RIA	Radio immunoassay
RP-HPLC	Reversed phase - high performance liquid chromatography
RSD	Relative standard deviation
SIM	Selective ion monitoring
SIMCA	Soft independent modelling of class analogy
SPE	Solid phase extraction
TIC	Total ion count
TMS	Trimethyl-silyl
TR-FIA	Time resolved – fluoroimmunoassay
TSP	Trimethylsilyl-2,2,3,3-tetradeuteropropionic acid
UPLC	Ultra - performance liquid chromatography
UV	Ultra - violet
y	year
δ (ppm)	chemical shift (in parts per million)
λ (nm)	wavelength (in nanometers)
cps	Counts per second
Δ	Change
SD	Standard deviation
TMA	Trimethyl-amine
TMAO	Trimethyl-amine-N-oxide

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor and mentor, Professor S.W.A. Bligh for suggesting the research project, her guidance, her suggestions, valuable discussion, welcoming of new ideas and encouragement throughout the completion of this research study. Many thanks are due to Professor C. Branford-White for his constructive advices and the financial support of my research project. I am also grateful to Dr D. Bhakta for her valuable ideas in some of my design of experiments. I am indebted to Dr K. White and Dr D. Spillane for their support through my programme; as well as Dr D. Green for his constant motivation and belief in my research programme and its completion. Sincere gratitude is expressed to Mr W. Dissanayake for his understanding and technical assistance in my research; also, Mr J. Crowder, Ms P. Haria, Mr A. Bashall, Mr S. Boyler and Mr J. Morgan for their technical advices and supports using their laboratory. I extend my warm thank you to Professor O. Familoni and Dr Ebuehi from University of Lagos for always being supportive of my academic progression (from the start of my B.Sc) and giving me the collaboration opportunity as well to Dr T. Asekun for her collaboration study support.

Undoubtedly, my exceptional thank-yous go to our former researchers - Maria, Elif & Farzad; and current researchers in the university: Yicheng, Xiao-Bei, Kevin, Ravi, Amara, Sumeet, Ephraim, Dimple, - especially Ed Welbeck '*my evil-twin*' for his assistance, motivation and presence in the laboratory. Many thanks to friends: Eric, Sharon, Ruth. Also, I would like to thank Dr N. Wardle, Ms M. Ali, Ms Taiwo, Ms Karlene and Ms Josie for their encouragement and enquiries about my research progress.

Personally, I would like to thank my family- Dad, Mum, Sis Shola, Lola, Bro. Femi, Bro. Gbenga, Mercedes and Semisi for their support and constant concern on my research-work. Again, gratitude to my *'twin sis'* - Bimbo Sanni for her encouragement and belief in me throughout my research programme. I am also

grateful to my new family- 'The Ogunjimi's'- Dad, Mum, Bola, sisters and brothers for their understanding and support during my research. Special and heart-felt thanks to Mr Ariyo Ogunjimi *'my other-half'* for his ever-present support, understanding and assistance, especially in proof-reading my thesis.

Dedication

I wish to dedicate my research work to ALMIGHTY ALLAH (subhana wa tala) for his mercies and guidance throughout this challenging research programme.

Declaration

I declare that whilst studying for the Doctorate in Chemistry at London Metropolitan University, I have not been registered for any other award at another university. The work undertaken for this degree has not been submitted elsewhere for any other award. The work contained within this submission is my own work and, to the best of my knowledge and belief, it contains no material previously published or written by another person, except where due acknowledgement has been made in the text.

Olumuyiwa Lateefah Ogegbo

May 2009

Chapter I

Phytoestrogens and their metabolites

1.1 Introduction

Increasing attention has been drawn to phytochemicals as it has shown that they have health protective properties. In actual fact, it was noticed that some populations consuming high level of polyphenols including phytoestrogen-rich diets have lower incidence of hormone-related cancers and cardiovascular diseases (*1-5*). Variances in the incidences of diseases and cancers over the world population may be accounted for by unavailable detection instruments, late diagnosis, treatment differences/ delays especially in under-developed and developing countries (*6, 7*). However, environmental factors- racial, lifestyle and dietary differences also play major roles in these incidence disparities (*8*). Supporting the investigation of environmental factors effects, several studies have been idealised on the East Asian population with the lowest incidence of hormonal cancers (*9*). Following these studies, it was established that the East Asian diet is soy-based, containing high concentration of phytoestrogens (*10-13*). This soy-rich diet was proposed to explain the lower risk and incidence of hormonal cancers in this population. Due to this, several studies have been carried out on this soy-rich diet for the identification and analysis of the vital compounds responsible for these effects. Conclusively, the results of these reports showed that one of the main groups of compounds in this diet are the phytoestrogens. Supporting this, previous studies have shown that phytoestrogens influence hormone production and metabolism (*1, 4, 11, 14-16*).

1.2 Chronological history of phytoestrogen research

In 1928, Baker and Robinson synthesised a novel phytoestrogenic compound, genistein (*17*). A few years later, isoflavones, which is a group of phytoestrogens, were found to be present in soy - bean meal (*Soja hispida*) by Waltz (*18*). The first isolation and identification of a metabolite of isoflavone, equol was found in the urine of pregnant mares' and reported by Marrian et al (*19*). Following Waltz's discovery, another group of phytoestrogenic compounds called lignans, were found and named by Harworth in 1941 (*20*). On the biological potential of phytoestrogens, in 1946, it was reported that phytoestrogens caused infertility by acting as contraceptive pills in red – clover

(phytoestrogen-rich) pasture grazed sheep (*21*). Since then, more biological studies on the metabolism and monitoring of phytoestrogen metabolites in animals (e.g. California quails, rats, cats, dogs, fowls, sheep, goats and monkeys) including humans have been studied (*22-39*). In 1980, Tang and Adams found that the isoflavone metabolite, equol and human estrogen, 17β-estradiol competed for the same cytoplasmic estrogen receptors (ER) (*40*). Thus, following this lead, some other phytoestrogens and their metabolites have been reported to exhibit both estrogenic and anti-estrogenic activities as a result of their binding affinity to uterine estrogen receptors (β-ER) (*40, 41*). Another key history of phytoestrogens is the first chromatographic analyses and characterisation of phytoestrogens by Naim et al (*42*) and West et al (*43*). TABLE 1.1 shows the summary of the key phytoestrogen discoveries and studies.

TABLE 1.1: Chronological key findings/ discoveries of phytoestrogen

Year	Researcher(s)	Findings	Ref.
1928	Baker and Robinson	Synthesis of genistein	(*17*)
1931	Waltz	Isoflavones in soy-bean meal (Soja hispida)	(*18*)
1932	Marrian et al.	Discovery and isolation of equol, a phytoestrogen metabolite, from mare's urine	(*19*)
1941	Harworth et al	Discovery of lignans, which is another type of phytoestrogen	(*20, 44*)
1946	Bennetts et al.	Infertility caused by red clover (rich in phytoestrogen) pasture fed to sheep	(*21*)
>1950	Klyne, Batterham et al.	Animal studies on metabolism of phytoestrogen	(*22-33*)
1970s	Tang et al, Shutt et. al.	Investigation of weak estrogenic activities of phytoestrogens	(*40, 41*)
1974, 1978	Naim. M. et al., West. et al.	First chromatographic analyses and characterisation of phytoestrogens in plants/ food- GLC and HPLC respectively	(*42, 43, 45*)
>1990	Adlercreutz. et al. Mazur et al Messina et al	Phytoestrogens in foods. Phytoestrogen metabolites in bio-fluids. Epidemiological studies	(*14, 16, 46-60*)
>2000	Liggins et al Lampe at al	Analysis on foods, biological samples, epidemiological, interventional studies	(*60-72*)

The last twenty years in phytoestrogen research, more focus has been drawn towards food analyses, dietary analysis, epidemiology, phytoestrogen biomarker analyses in bio-fluids, and more recently interventional studies to monitor the effects on human subjects (*14, 16, 46-73*). It is noteworthy that these studies have been carried out on selected populations with the East Asians and Europeans been the vastly studied population. Consequently, for nearly three decades following the fore - mentioned discoveries and studies listed in TABLE 1.1, research scientists have been *sceptically* aware of the potential importance of phytoestrogens on the human population regarding some of their health benefits on prevalent diseases. Conversely to phytoestrogen positive health benefits (*1-5*), negative effects or no effects have also been reported due to their antagonist estrogenic effects (*74, 75*).

1.3 Phytoestrogen occurrence

Phytoestrogens exist ubiquitously in plants as mainly glycoside conjugates (*76-79*). In nature, they usually occur as one, two or three of these main forms- glycosides, acetyl-glycosides or malonyl-glycosides. The pre-dominate glycoside of phytoestrogens are the β-glucoside, and in fewer cases, they occur in addition as acetyl-β-glucoside and 6'-O-malonyl-β-glucoside. However, in rare cases, there are α-monomers, where the sugar moiety is substituted on the C-4' position, oxygen bonds at the C-1' and carbon at the C-5' are in trans– conformation (*80, 81*).

1.3.1 Definition of phytoestrogens

The classic definition of phytoestrogen is a non-steroidal plant compound that is able to exert estrogenic effects (*82*). The growing interest on phytoestrogen research is based on the hypothesis that these compounds play important roles in the prevention of many hormone-related conditions (*83*).

Phytoestrogens are the new form of plant-derived *'dietary'* estrogens, which possess estrogen agonist or antagonist effects due to their structural similarities with the human hormone (17β-estradiol) based on two main suppositions (FIGURE 1.1) (*84*):

- Phytoestrogen ring A (red box) mimics estrogen ring A at receptor binding.
- Almost identical distance (blue arrow) of terminally opposite hydroxyl groups similar to those of estradiol

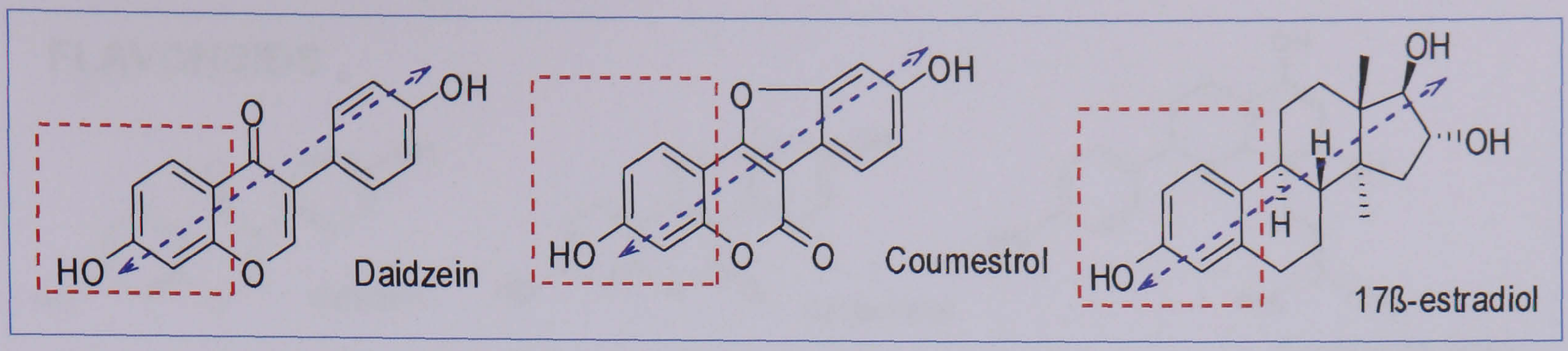


FIGURE 1.1: Structural-function similarities of flavonoidal phytoestrogen with estradiol

1.3.2 Classification of phytoestrogens

Phytoestrogens can be generally classified based on their structure similarities into two main classes, the flavonoids and the non-flavonoids as detailed in FIGURE 1.2. From this classification, isoflavones and coumestans are flavonoids possessing the highest estrogenic activity (*85-90*). Of this class, prenyl flavonoids have not been studied extensively being just newly discovered (*91*). With non-flavonoids, lignans are the popularly studied group having a higher estrogenic activity than stilbenes (*91*).

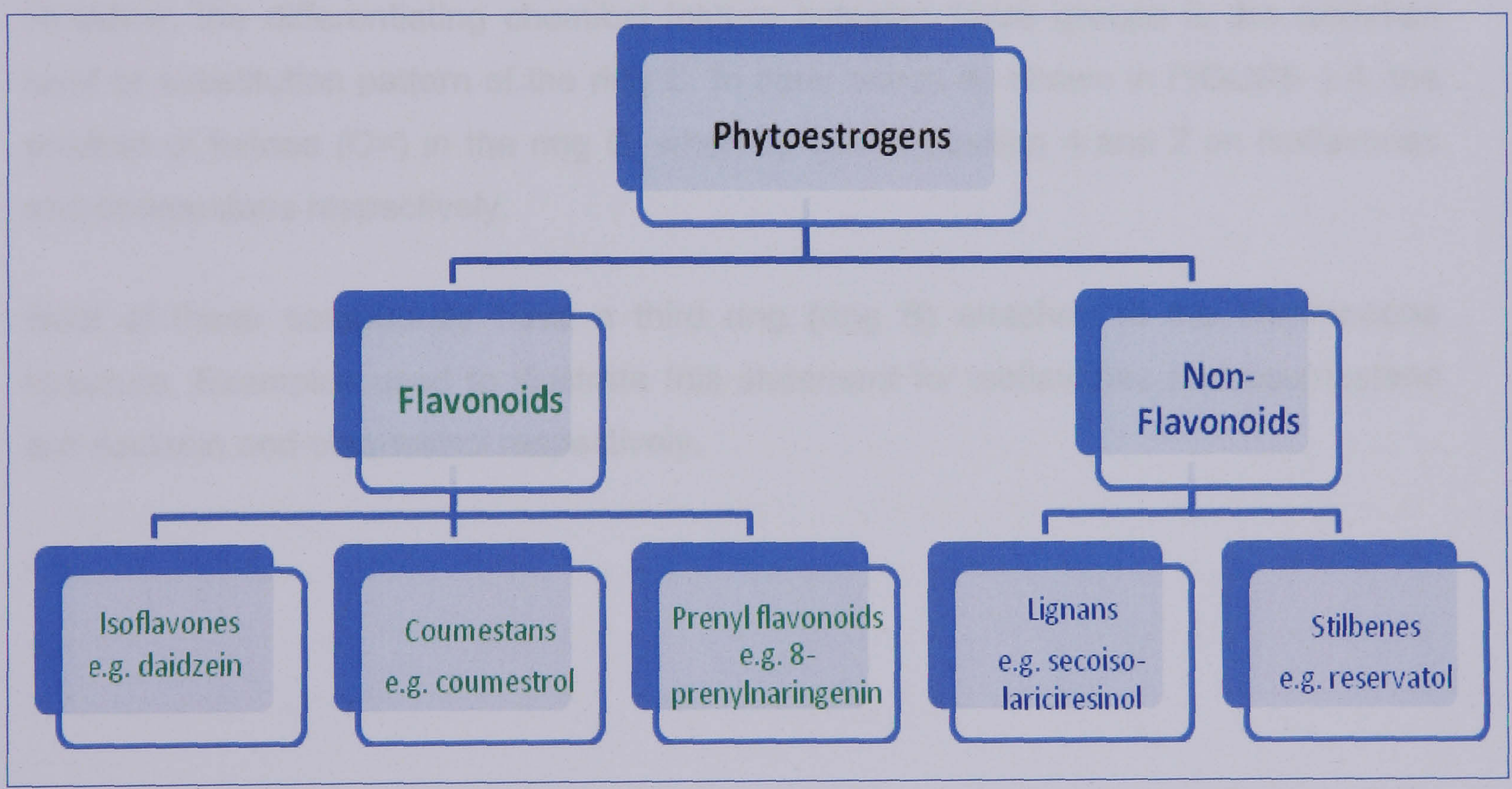


FIGURE 1.2: General classification & examples of phytoestrogens (adopted from (*84*))

FIGURE 1.3 shows the structural similarities and differences between the two major phytoestrogen classes- flavonoids and non-flavonoids with examples.

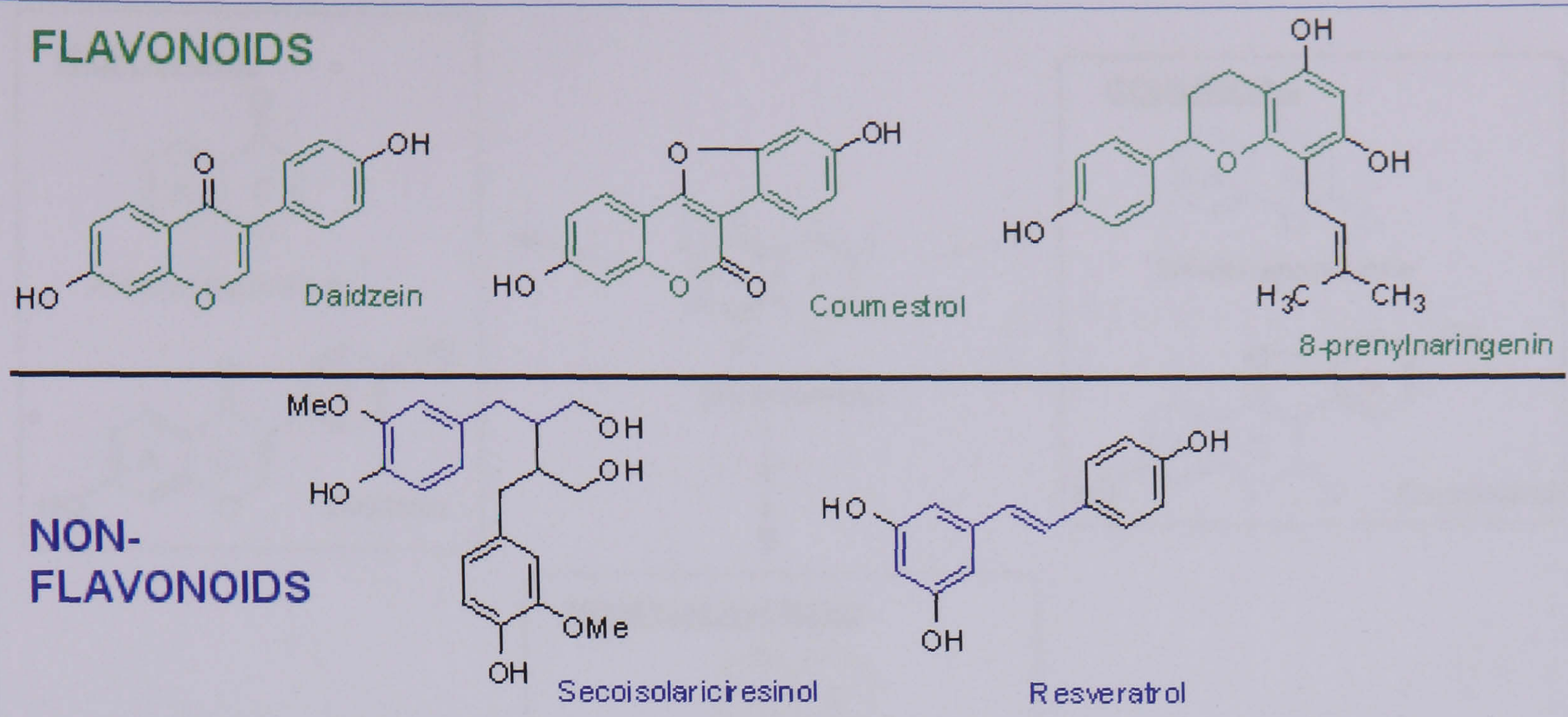


FIGURE 1.3: Structural similarities of five phytoestrogen groups

The most studied flavonoidal phytoestrogens are the isoflavones and coumestans, which are structurally similar, i.e. having a chromene skeletal frame (rings A and C). However, the differentiating chemical feature between these groups is the oxidation level or substitution pattern of the ring C. In other words as shown in FIGURE 1.4, the position of ketone (O=) in the ring C, whereby it is at position 4 and 2 on isoflavones and coumestans respectively.

Both of these compounds have a third ring (ring B) attached to the chromenone structure. Examples used to illustrate this statement for isoflavones and coumestans are daidzein and coumestrol respectively.

.

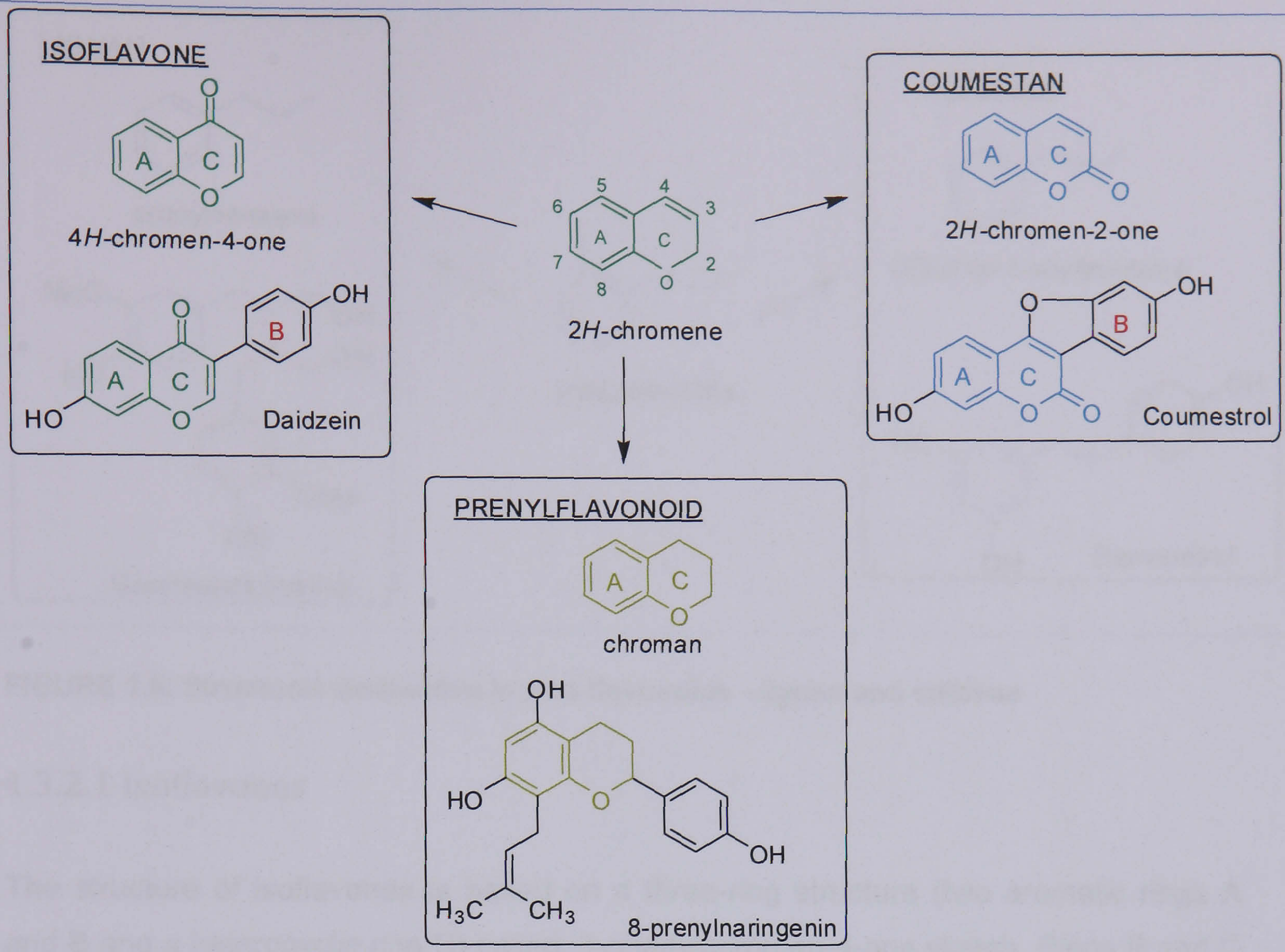


FIGURE 1.4: Structural similarities in flavonoids - isoflavone, coumestan and prenylflavonoids

Prenylflavonoids being a member of flavonoidal phytoestrogen are also chemically similar to the other two in this class. However, the difference between these groups is the saturation in ring C, whereby the skeletal frame is *chroman* in the prenylflavonoid in opposition to *chromenone* in isoflavones and coumestans (as shown in FIGURE 1.4). Also, it is observed that prenylflavonoids have ring B attached at C-2 of ring C, whereas isoflavones' or coumestans' ring B is attached at C-3. Examples include: 8-prenylnaringenin, 6-prenylnarigenin, xanthohumol, isoxanthohumol

With the non-flavonoidal phytoestrogen groups, lignans and stilbenes are slightly structurally similar based on their propylbenzene skeletal frame as shown in FIGURE 1.5.

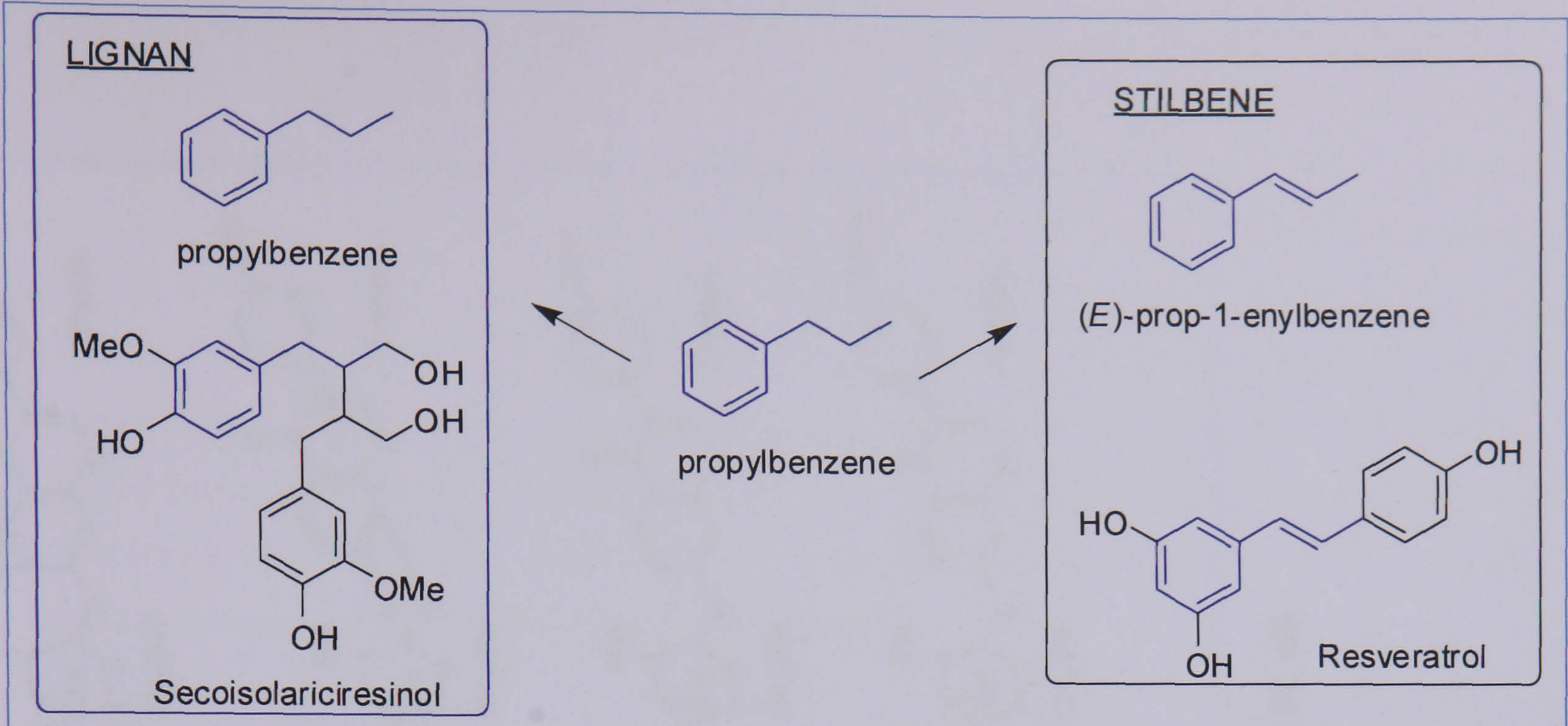


FIGURE 1.5: Structural similarities in non-flavonoids - lignan and stilbene

1.3.2.1 Isoflavones

The structure of isoflavones is based on a three-ring structure (two aromatic rings A and B and a heterocyclic ring C) called: 3-phenylchromen-4-one sketch. Rings B and C are singly bonded at C-3 on ring C and C-1' on ring B, forming a 6-3-6 carbon backbone. The general structure of isoflavones is as shown in FIGURE 1.6:

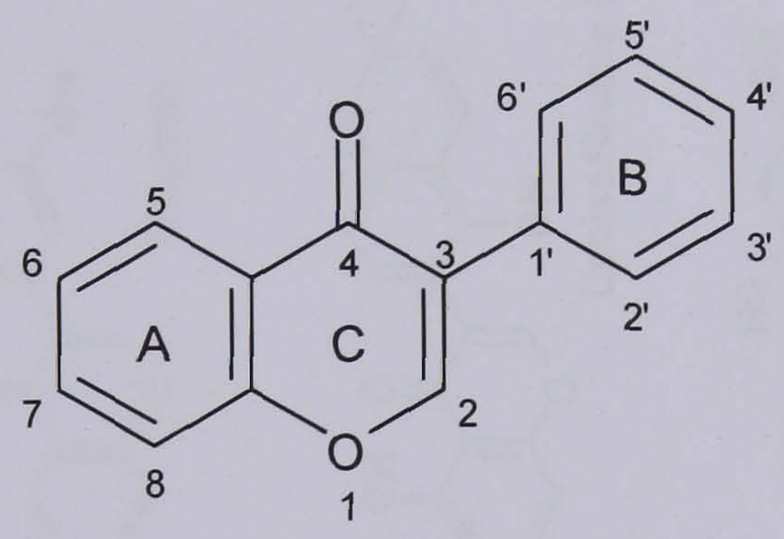


FIGURE 1.6: Isoflavone skeletal structure

The pre-dominate glycoside of isoflavones is the β-glucoside substituted on the C-7 position, whilst their less common glycoside forms include acetyl-β-glucoside and 6'-O-malonyl-β-glucoside. Besides the glycosides moieties, other functional groups attached to isoflavone skeleton are hydroxyl-, acetyl-, glucoronide, sulphate, malonyl, halogens groups, sulphoglucuronide (*84, 91, 92*).

So far, over three hundred and fifty isoflavones aglycones have been reported. Some of the isoflavones are shown in FIGURE 1.7:

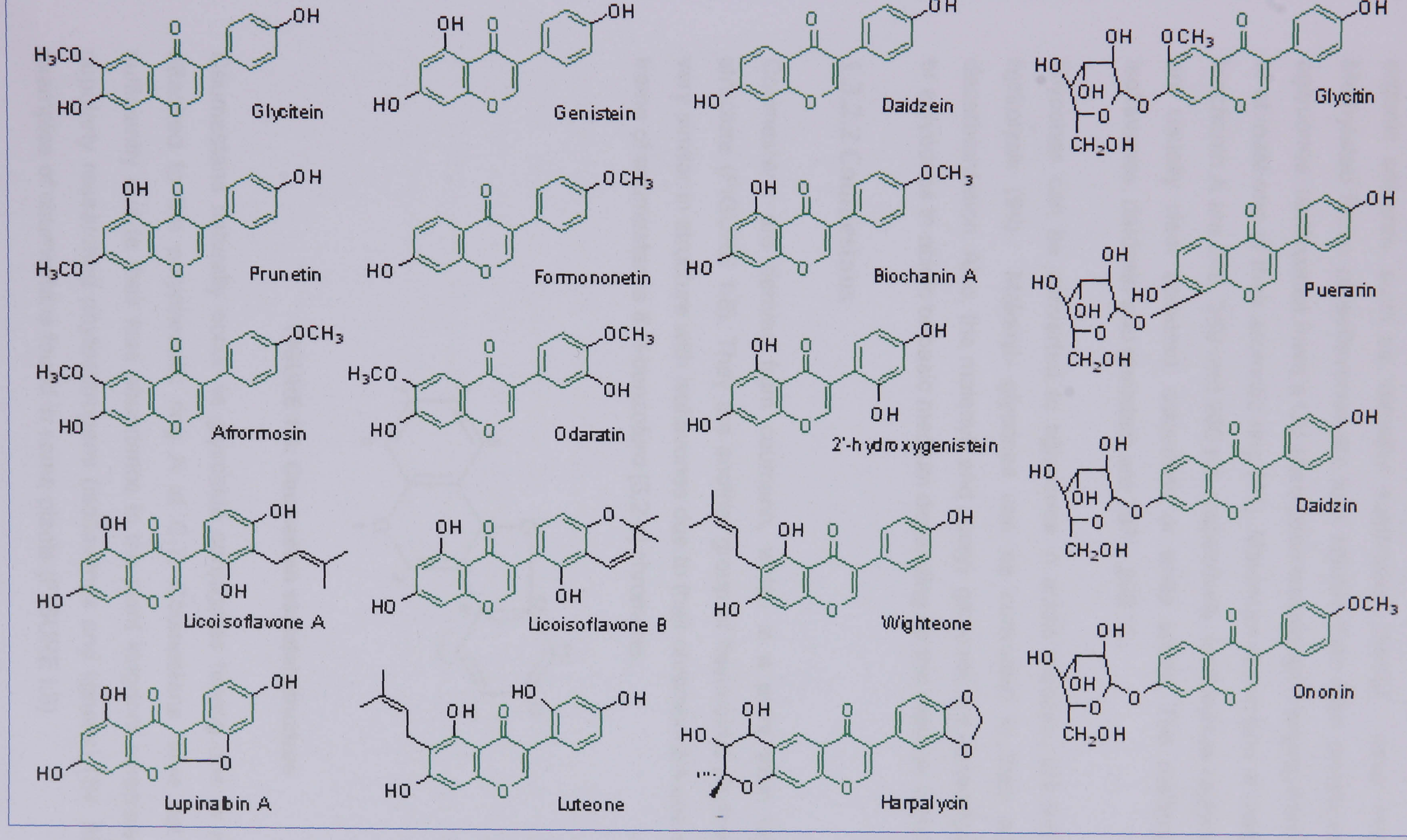


FIGURE 1.7: Structures of some examples of isoflavones isolated from various plants (*18, 46, 84, 91, 92*)

Depending on the attached functional groups or conjugates, isoflavones could either be hydrophobic or hydrophilic. In their nature forms, they exist as glycosides and they are highly hydrophilic. However, their aglycones are highly hydrophobic and only soluble in organic solvents, such as, dimethyl sulphoxide, diethyl - ether methanol, acetonitrile. Methylated forms of isoflavones are less soluble than their corresponding unsubstituted aglycones. Isoflavones have a UV absorption wavelength ranging from 250 to 270 nm due to at least one of their aromatic ring (*91*). Maximum absorption of daidzein, genistein and biochanin A are 249, 259 and 260 nm respectively. Isoflavones (aglycones and glycones) are usually clear coloured, colourless or white solids. The melting points of popular isoflavones, daidzein and genistein are 297 – 298 °C.

Glycones can be converted to aglycones in acidic medium, pH 3-5 (*93*) or enzymatic hydrolysis (*94*). Malonyl- glycones can be converted to their acetyl-derivatives by decarboxylation. Also, the malonyl- and acetyl- glycones are converted to either glycones or aglycones in acidic or basic medium depending on the medium concentration.

1.3.2.2 Coumestans

Coumestans are derived from coumarin, which is a plant toxin with a benzopyrone structure (FIGURE 1.8). They are another group of flavonoidal phytoestrogen, which are very similar in structure with isoflavones due to their chromen-2/4-one skeletal frame. The frame of coumestan is *6 H*-benzofuro [3,2-c] chromene.

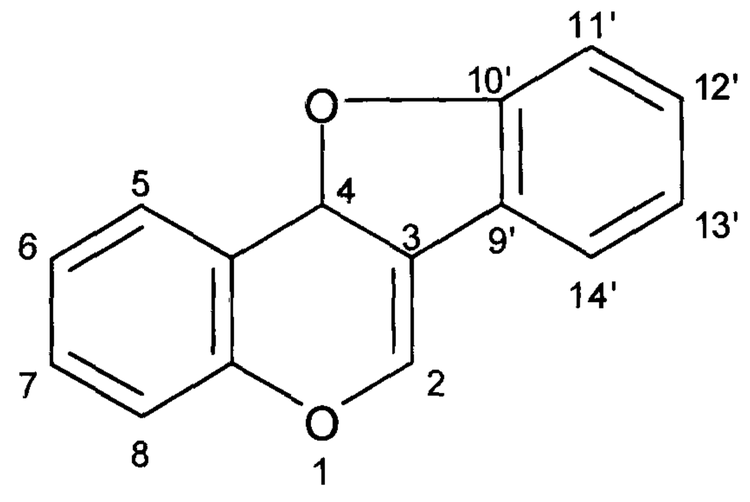


FIGURE 1.8: Coumestan skeletal structure

Coumestans naturally occur as glycoside conjugates having one or more sugar moiety attached to the polyphenolic ring A at C-7. Coumestans have not been researched sufficiently due to their less abundance in the plant kingdom compared to the other two popularly researched phytoestrogens (isoflavones and lignans) (*84, 95*). Here are some examples of coumestans found in some plants (FIGURE I.9).

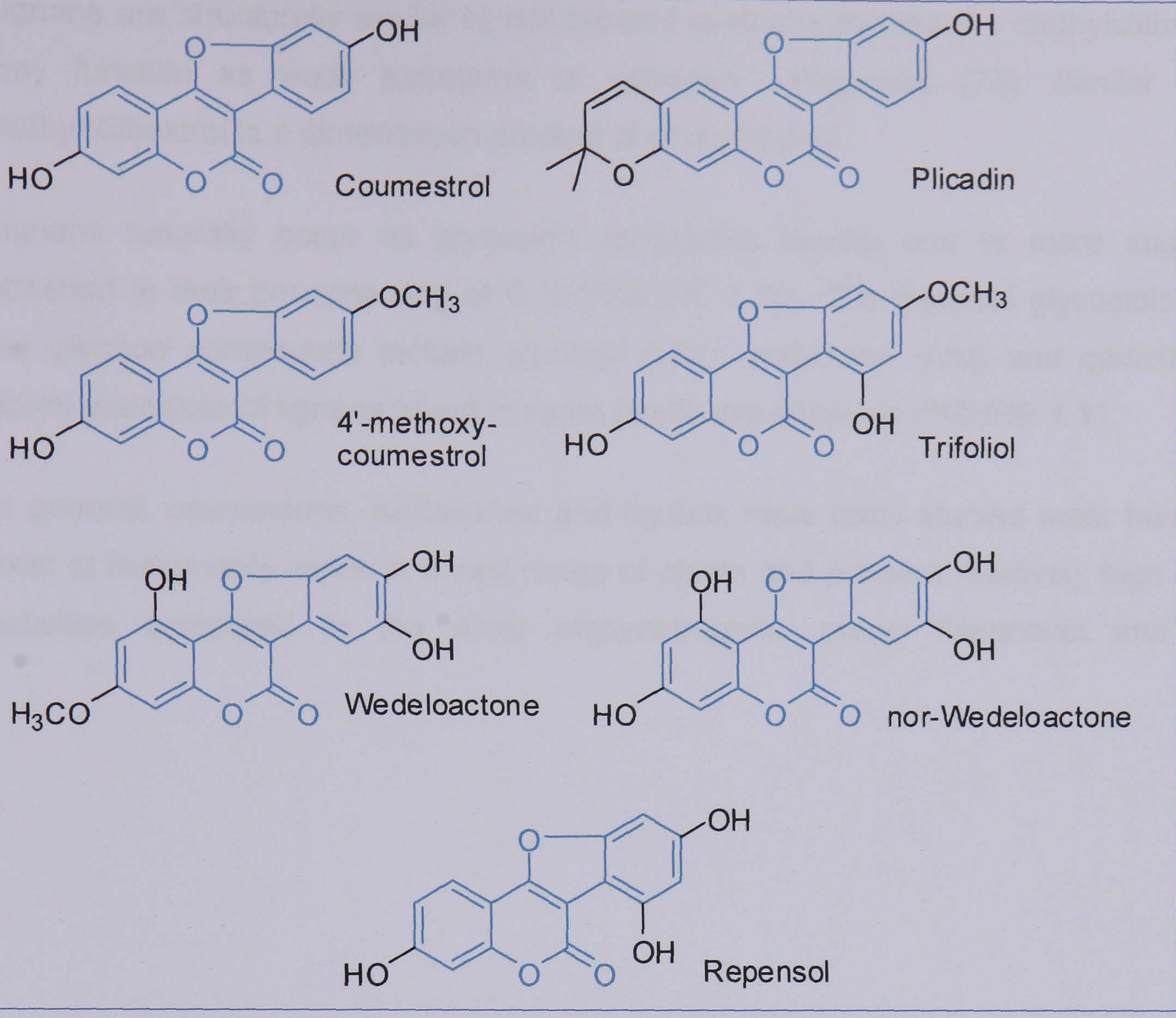
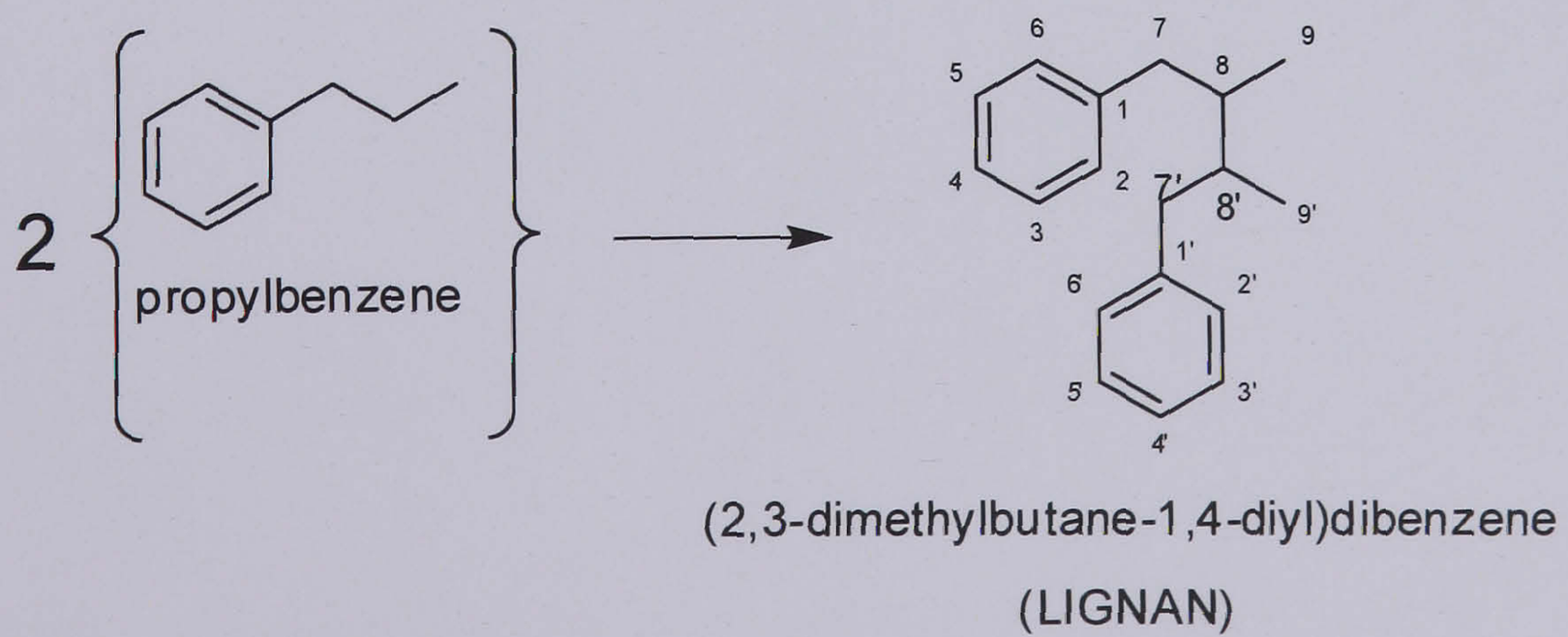


FIGURE 1.9: Structures of some examples of coumestans isolated from plants (*84, 91*)

1.3.2.3 Lignans

Compared to the isoflavones and coumestans, lignans are not produced naturally by the plants. They are biosynthesised by dimerisation of propylbenzene to form (2,3-dimethylbutane- 1,4 – diyl) dibenzene (FIGURE 1.10). Non-flavonoidal lignans are one of the few exceptional phytoestrogens classes possessing a bridge linking the two benzene rings. Common examples of lignans are secoisolariciresinol and matairesinol.



(2,3-dimethylbutane-1,4-diyl)dibenzene

(LIGNAN)

FIGURE 1.10: Dimerisation biosynthesis of lignan

Lignans are structurally similar to natural and synthetic estrogens - diethylstilbestrol; and may function as weak estrogens or estrogen antagonists (*75*). Similar to lignan, diethylstilbestrol is a dimerisation product of propylphenol.

Lignans naturally occur as glycoside conjugates having one or more sugar moiety attached to their benzene ring at C-5 (FIGURE 1.10). The possible glycosidic groups in the glycone compounds include glucosyl (Glc), arabinosyl (Ara) and galactosyl (Gal). Some examples of lignans found in some plants are shown in FIGURE 1.11.

In general, coumestans, isoflavones and lignans have been studied most because they exist at reasonable levels in a vast range of plants and possess relatively high estrogenic activities compared to the other phytoestrogens: prenyl flavonoids and stilbenes

Matairesinol

HydroxyMAT

Nortrachelogenin

a-Conidendrin

Syringaresinol

Todolactol

Seco-isola riciresinol

Seco-isola riciresinol Di-glucoside

Eudesmin

Podophyllotoxin

Lariciresinol

Pinoresinol

Lignan A

Cyclolariciresinol

FIGURE 1.11: Structures of some examples of lignans isolated from plants (*30, 96-99*)

1.3.3 Chemical ecology of phytoestrogens in plants

Plants usually produce defence chemicals against herbivores, competitors, pathogens and abiotic stresses. These defence chemicals are either active (produced after attack) or passive (always present). The active defence chemicals are called 'phytoalexins'. As a result, related plant families and species produce similar and specific types of phytoalexins. Phytoalexins are polyphenolic compounds including phytoestrogens, which are plants secondary metabolites produced from the metabolism of primary metabolites, e.g. proteins, carbohydrates, fatty acids. Specifically, phytoestrogens function in the plants as plants defence compounds against other competing plants, herbivores, infectious micro-organisms (e.g. fungi, viruses, bacteria) and pests (e.g. insects and snails) (*50*) (*100*). In the plants' defence against the herbivores, it is hypothesised that secondary metabolites are synthesised to play a controlling role in the herbivores' reproductive system (*100, 101*). Other functions of phytochemicals including phytoestrogen in plants include: regulate plant hormones, trigger hypersensitive cell death initiation (i.e. apoptosis-like defence) and protect against harmful UV radiations (*102*). Another important function of phytoestrogen is to promote symbiosis with rhizobacteria (*102, 103*), for nodules production in legume crops, which converts atmospheric nitrogen to ammonia in the soil for plant growth (*104*).

Flavonoids are biosynthesised from 4 - coumaroyl - coenzyme A (CoA), which is derived from phenylalanine. With the combination of 4-coumaroyl – CoA to malonyl - CoA, chalcones are produced, following which with the conjugate ring closure of chalcones, isoflavanones are produced. Isoflavanones are then reduced to isoflavones following dehydration (*105, 106*) as shown in FIGURE 1.12. In addition, isoflavones or isoflavanones could be methylated by isoflavones- or isoflavavones - O - methyltransferase (IOMT) to form methoxy-isoflavones or methoxy-isoflavanones.

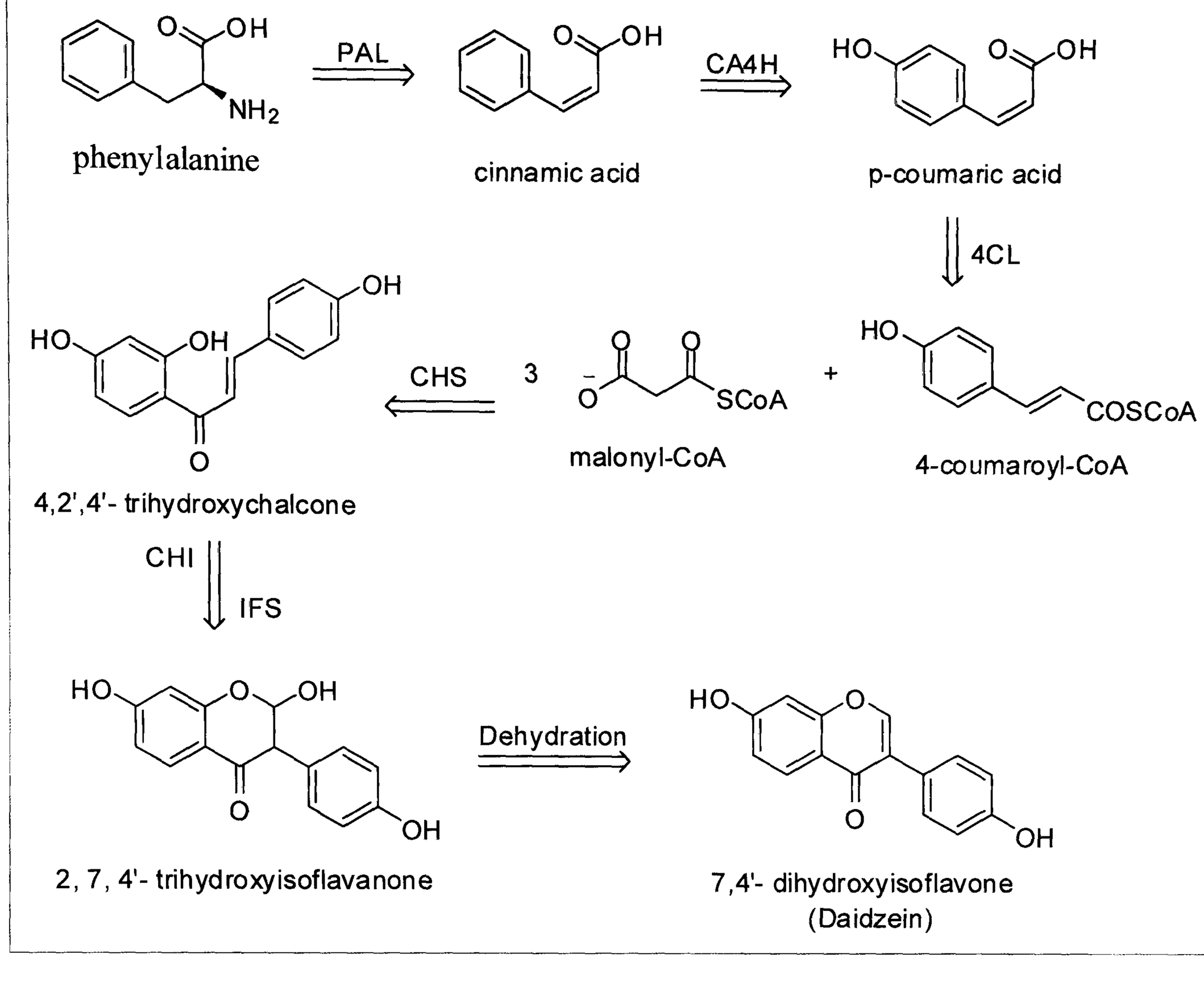


FIGURE 1.12: Biosynthetic production of isoflavones (adopted from (*106*)

PAL- phenylalanine ammonia-lyase; CA4H- cinnamate-4-hydroxylate; 4CL- 4-coumarate CoA ligase; CHS- Chalcone synthase; CHI- Chalcone-flavanone isomerase; IFS- isoflavone synthase.

1.3.4 Occurrence in food

Following the discovery of isoflavones in soy meal by Waltz, analysis and testing of several (> 300) plants and foods for different types of phytoestrogens have been reported (*42, 45, 59, 60, 70, 71, 107-111*). Phytoestrogens are extensively spread in the plant kingdom, thereby present in minimal (< 0.01 µg/ 100g) to high concentrations (> 500 mg/ 100g) in some plants and foods. Phytochemicals have been reported to be present especially in fruits and vegetables (*59, 74, 109, 112*). Most especially, leguminous plants have been reported to have the most amount of phytoestrogen as glycones and aglycones.

Isoflavones are found abundantly in most legumes. The major sources of isoflavones, coumestans and lignans are soy bean (*Glycine max*), alfalfa sprouts (*Medicago sativa*) and flaxseed (*Linum usitatissimum*) respectively. Furthermore, of these three foods, flaxseed has been found to contain the highest amount of phytoestrogen – approximately 4 mg/g, containing mainly lignans. Other sources of lignans are linseeds rye, teas, nuts, whole grains and berries (*97, 112*). Likewise, other sources of coumestans are mung beans and soy sprouts. Soybean and alfalfa sprouts specifically have been reported to contain 1-3 mg/g and 5 μg/g respectively. Also, red clover (*Trifolium pratense)* has been stated to have high isoflavone content; although, red clover plants are not easily consumed by humans in the everyday diet, nonetheless, isoflavones are extracted from these plants and manufactured into medicinal tablets forms.

To date, over three hundred different phytoestrogens have been discovered in various plant samples (foods and herbs). As summarised in TABLE 1.2, foods, plants and herbs from the Far Eastern countries (such as: Japan, China and Korea) and Western countries (such as United Kingdom, Finland and United States of America) have been extensively studied for their phytoestrogen content.

TABLE 1.2: Literature of some foods analysed in different locations.

SAMPLE	Country	Phytoestrogens assessed (μg/100g wet weight)								Ref
		Formono-netin	Biochanin A	Daidzein	Genistein	Glycitein	Coumestrol	Secoisolari-ciresinol	Matai-resinol	
Soy beans	Canada	5.2	-	56621.4	44213.4	2809.4	1.5	79.1	1.8	(*113*)
Flaxseed		34.1	-	58.2	173.2	56.0	46.8	375321.9	153.3	
Chick peas		0.5	-	0.1	1.1	0	0.1	0.6	0.1	
Alfalfa sprout		384.3	-	1.7	7.5	0.6	2.2	18.4	0.1	
Flaxseed	Finland	0	0	0	0	-	0	369900.0	1087.0	(*52*)
Soy flour		30.2	74.4	67369.0	96914.0	-	0	130.4	trace	
French beans	UK	0.0	16.0	124	381.0	20.0	31.0	1009.0	0	(*114*)
Kudzu (root)	USA	7090.0	1400.0	185000.0	12600.0	-	1570	30.7	trace	(*112*)
Cowpea		0	0	30.3	55.7	-	Trace	195.0	nd	
Black eyed beans		5.5	7.7	20.5	11.4	-	7.7	196.0	nd	
Groundnut		6.8	6.5	49.7	82.6	-	0	333.0	trace	
Soybean seeds	Japan	nd	nd	100650.0	138240.0	-	-	-	-	(*48*)
Clover sprout	Hawaii	2280.0	440.0	nd	350.0	-	28060.0	-	-	
Alfalfa sprout		340.0	nd	nd	nd	-	4680.0	-	-	
Pumpkin seeds	UK	3.0	7.0	<1.0	5.0	2.0	<1.0	510.0	11.0	(*70*)

- not measured; nd: not determined because of low concentrations.

Phytoestrogens are generally found in most everyday diet ranging from very low or absent to very high concentrations, for example soybeans have been found to contain the highest amount of isoflavones (2-3 mg/g) in 0.4 - 9.5 mg/g total phytoestrogen. The variation in the phytoestrogen levels have been found to be due to crop location, time of harvest and genetics (*115, 116*). Phytoestrogen are generally found in mainly plants, however until recently, they have been also found to be present in other food sources - animals based foods: for example meat, milk, seafood (*111*). With the acknowledgement of the health benefits of phytoestrogens, food industries have been allowed to claim on the potential benefit of phytoestrogen stating that the consumption of foods containing 6.25 g of soy protein per serving can reduce heart disease risk (Food and Drug Administration, 1999).

1.4 Exposure to phytoestrogens

Suggestions based on epidemiological and migration studies have shown that racial characteristics and other factors including lifestyle, diet, and fat or fibre intake may play a role in the etiology of certain chemoprophylaxis diseases (*84*).

To date, the level of exposure of phytoestrogens to Africans has still not been reported. Whilst in contrast, the levels of isoflavones consumption have been extensively studied within the South East Asians, Europeans and Americans. It was stated that South East Asians (from China, Japan, etc) consume the highest amount of isoflavones as their diet is mainly soy food products- consuming an average of 25 – 100 mg isoflavones per day (*117*). Americans and British consume very little amount of phytoestrogen: no more than 1 – 3 mg/day, mainly lignans (*49, 87*).

Epidemiological studies on the relationship between disease risk and diet have shown that food has a direct impact on health (*118*). Several epidemiological studies on different populations have been conducted to determine the relationship between phytoestrogen intake and the risk of their claimed preventive diseases, e.g. hormonal related cancers, cardiovascular diseases (CVD), tumour growth. The populations studied to date include East Asians (*53, 63, 119-123*), Europeans (*68, 124-126*), Americans (*62, 113, 127-129*) and Australians (*130*).

However, there were a few contradictory reports published on these studies presenting either a negative or no association with some of these diseases – especially with breast and prostate cancers (*119*). As reviewed by Thompson et al (*113*), some of the differences in results could be due to the trueness of the database (*62, 131, 132*), the

differences in food sources, i.e. cooked, raw, processed, etc (*68*), non-generalisation of the phytoestrogen source, i.e. some based the phytoestrogen intake on soy products consumption only (*121, 124*).

1.5 Phytoestrogen metabolism

Intestinal bacterial metabolism of dietary components can result in the production of metabolites, which are more biologically active than their precursor thus influencing host's health [83]. Research on phytoestrogen metabolism and absorption was encouraged in the early 1980s by Axelson and Setchell, who discovered equol in urine and the levels of isoflavones were higher than endogenous estrogens (*83, 133*). Afterwards, it became possible to assess their concentrations in other bio-fluids e.g. in bile, faeces, saliva, semen and breast-milk (*67, 134, 135*).

Phytoestrogens like human estrogen have the potential to act as either estrogenic or anti-estrogenic compounds depending on the amount that is ingested and absorbed by the body (*136*), receptor status, presence or absence of endogenous estrogens, and the type of target organ or cell (*95, 137-140*). However, it has been found that phytoestrogens including their metabolites have much weaker estrogenic activity (10^{-2} to 10^{-5}) than those of the human estrogens- 17β-oestradiol (*41, 136, 141*). Generally, phytoestrogen concentrations circulating in the body are 100 times higher than endogenous estrogens (*90*) and with the consumption of a soy – diet, the isoflavones concentration is said to be about 100 to 1000 times higher than estrogen concentration (*142*). Thus, this leads to the competition between the estrogen and phytoestrogen for the binding ER - β, with the phytoestrogen being favoured over the estrogen (*143*). With this, more endogenous estradiol is excreted due to the struggle of estrogen receptor binding affinity, resulting to lower estrogen potency (*90, 144*).

1.5.1 Phytoestrogen metabolites

1.5.1.1 Intestinal microflora metabolism pathway of phytoestrogen

Intestinal microflora are involved in different processes within the body including metabolism of dietary compounds and are important in maintaining host health (*145, 146*). With the consumption of bio - inactive phytoestrogen glycones in foods, these glycones undergo β-glucosidase hydrolysis in the jejunum. Thus, the sugars moieties are broken down leaving their aglycone forms as active compounds in the intestine. The β-

glucosidases are produced by many species of intestinal bacteria, but the *bacteria* largely responsible for cleaving the glycosidic bond within the small intestine of mammals are the *Lactobacilli, Bacteroides,* and *Bifido* (*147*).

Before absorption occurs, intestinal bacteria may further metabolize the phytoestrogen aglycones to their metabolites. For example, genistein formed from biochanin A or genistin, is further metabolised to p-ethyl phenol which is non-estrogenic or biologically inactive (*148*), whilst daidzein, which is formed from formononetin is reduced to equol and o-desmethylangolensin (o-DMA). The two major metabolites of daidzein have estrogenic activity (*13, 46*), but this process happens to a substantial extent in only 30 - 50 % of people. On consumption of these foods, these aglycones are further metabolised in the distal intestine with the formation of dihydrodaidzein and dihydrogenistein and then finally equol and *o*-DMA (*147*). On the other hand, in lignan metabolism, the gut bacteria in the colon convert secoisolariciresinol to enterodiol, which is readily oxidised to enterolactone, whereas matairesinol is metabolized to enterolactone (FIGURE 1.13).

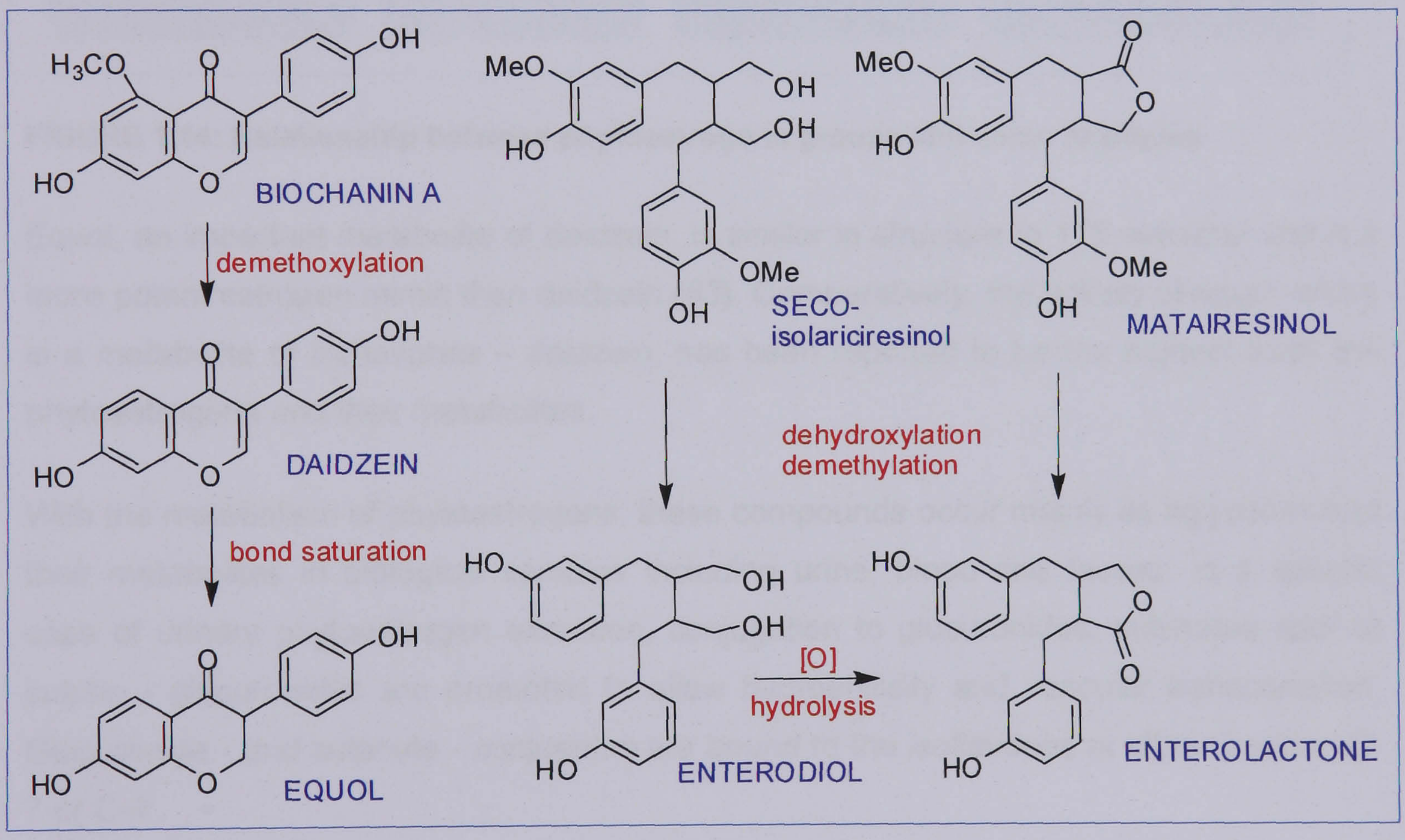


FIGURE 1.13: Gut/ Intestinal metabolism of different phytoestrogens

A summary of the general relationship between the phytoestrogens and their metabolites with some examples is shown in FIGURE 1.14.

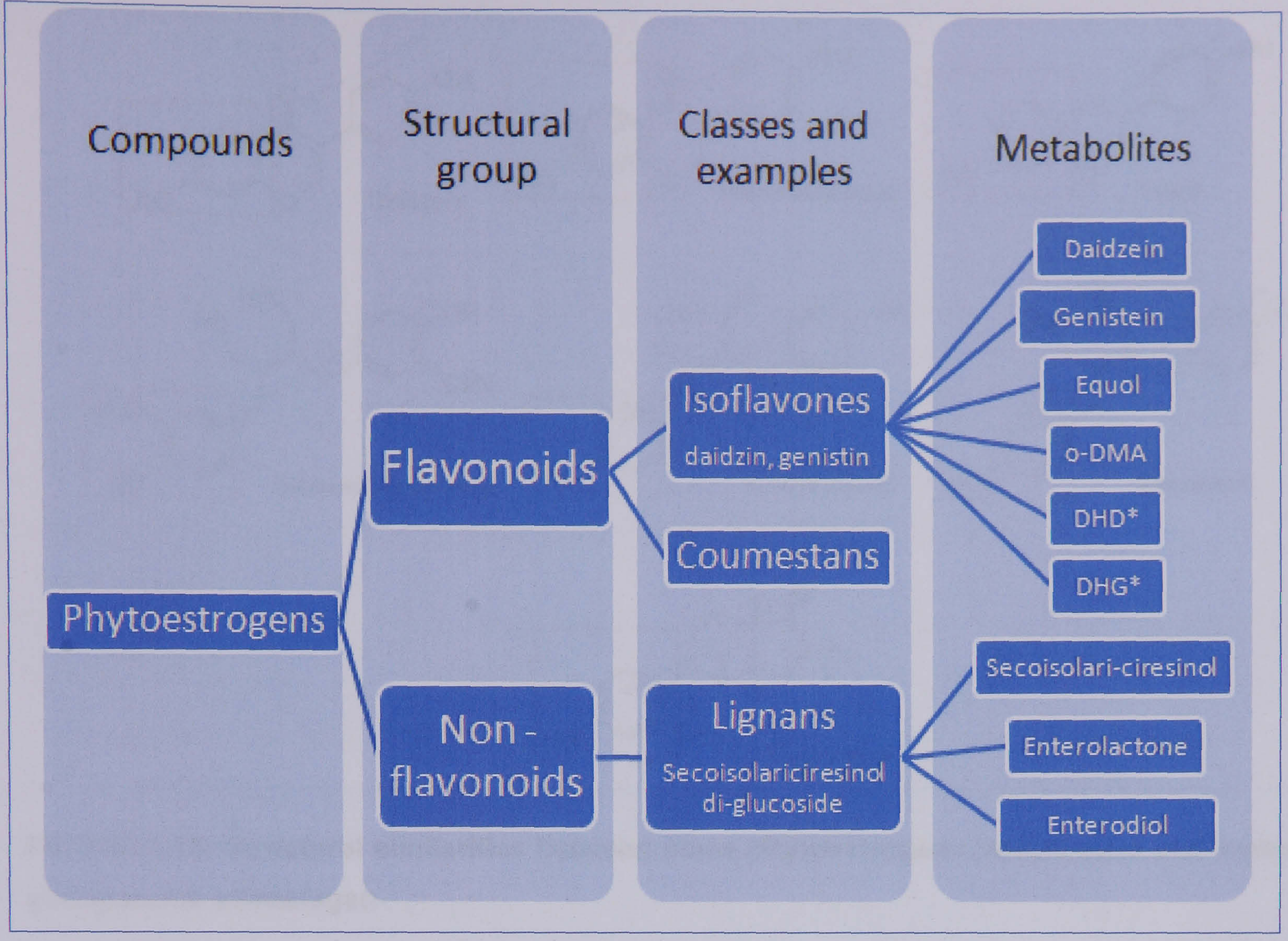


FIGURE 1.14: Relationship between phytoestrogens groups and some examples

Equol, an important metabolite of daidzein, is similar in structure to 17β-estradiol and is a more potent estrogen mimic than daidzein (*83*). Comparatively, the activity of equol, which is a metabolite of isoflavones – daidzein, has been reported to be the highest of all the phytoestrogens and their metabolites.

With the metabolism of phytoestrogens, these compounds occur mainly as aglycones and their metabolites in biological samples including urine, blood and faeces. In a specific case of urinary phytoestrogen excretion, conjugation to glucuronides, sulphates and/ or sulpho - glucuronides are promoted to allow hydrophilicity and vascular transportation. Glucuronide - and sulphate - conjugates are bound to the isoflavones at either position C-7 or C-4'.

Similar to phytoestrogen and estradiol structures, Setchell and Adlercreutz reported that the distance between the two aromatic hydroxyl groups on the metabolites nucleus is almost identical to the distance between the C-3 and C-17 hydroxy groups of estradiol (FIGURE 1.14). In addition, it was reported that it is the C-3 phenolic group (red box), which is required for estrogenic activity (*74*).

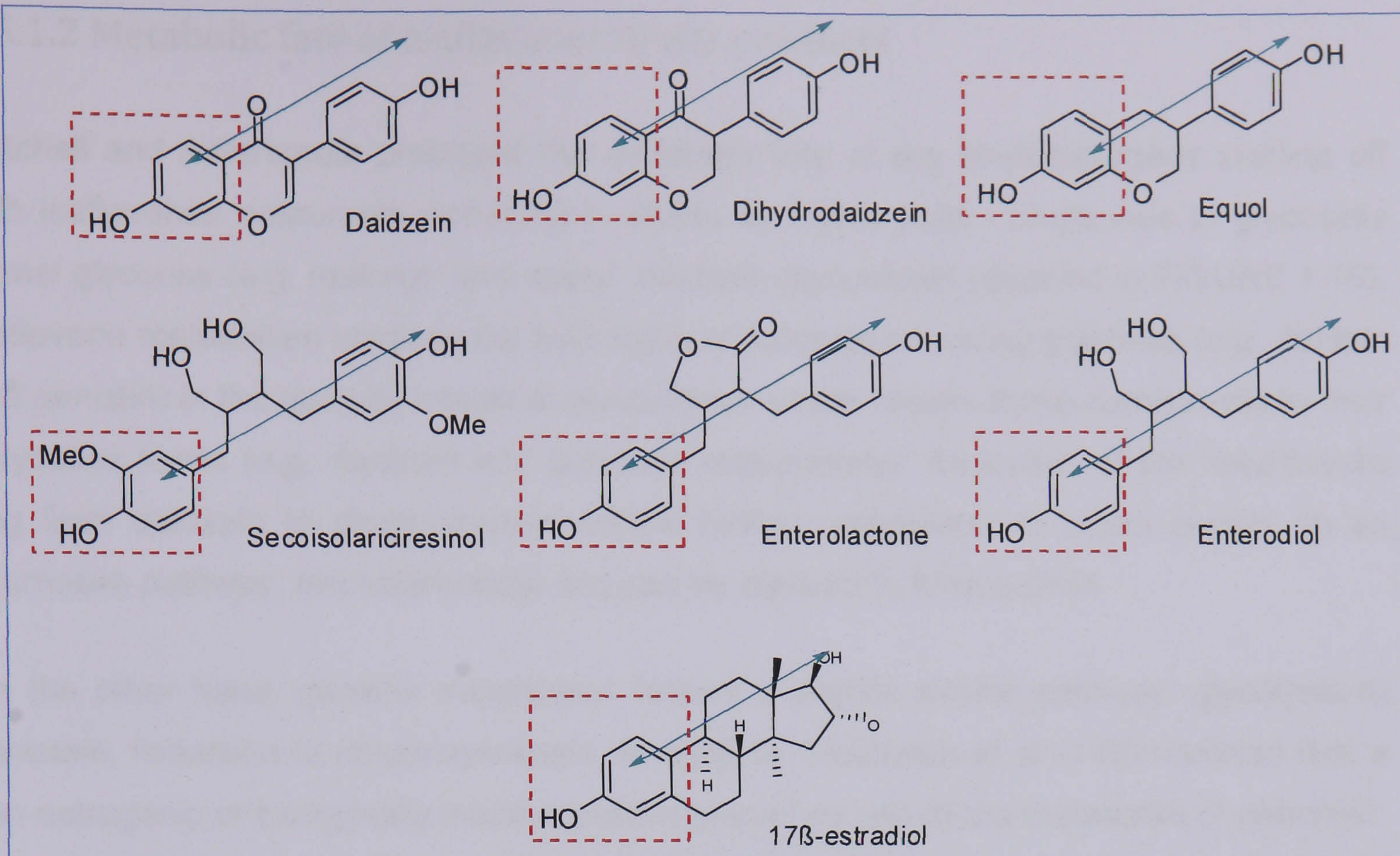


FIGURE 1.15: Structural similarities between some phytoestrogens, metabolites and human endogenous estrosfagen

On the absorption and circulation of aglycones in the body, excess phytoestrogens and their metabolites are filtered and transported from the blood stream to the liver, where β-glucuronide (glucuronic acids) and sulphates are conjugated to these compounds and excreted in urine (*95*). However, the level of sulphate derivatives is less than the glucuronic acid derivatives due to the hepatic glucuronyl transferase capacity. Generally, all of the metabolites are absorbed in the blood and removed as waste products in urine. The liver efficiently transfers isoflavones from the blood into bile thereby returning the isoflavones to the gut. Identical to endogenous steroids, phytoestrogens undergo a process called enterohepatic circulation, whereby these groups of compounds are deconjugated in the intestine and re-absorbed or excreted in the faeces (*95, 135*). 10% to 30% of ingested phytoestrogen is excreted in urine whilst it is significantly less scaled (< 10%) in faecal excretion (*135, 149*).

Generally, the conversion from glycones to aglycones and metabolites has very high inter- and intra - individual variations. These variations are influenced by many factors including age, gender, use of antibiotic, bowel disease, gut motility, gastric pH, bile secretion, stress, diet and intestinal transit time (*150*).

1.5.1.2 Metabolic fate of isoflavones in soy products

Setchell and Adlercreutz proposed the metabolic fate of soy phytoestrogens starting off with isoflavones precursors occurring in plants as highly polar compounds in glycosidic forms/ glycones (e.g. malonyl- and acetyl- daidzein glycosides) (detailed in FIGURE 1.16). Isoflavone metabolism involves the hydrolysis of naturally occurring glycones (e.g. daidzin and genistin) in the body by intestinal glycosidase, which breaks these compounds to their aglycones forms (e.g. daidzein and genistein respectively). Reduction of the heterocyclic ring from daidzein to dihydrodaizein before further metabolism to equol occurs. In an alternative pathway, the heterocyclic ring can be cleaved to form *o-DMA*.

On the other hand, genistin metabolism follows a slightly similar pathway: glycolysis to genistein, reduction to dihydrogenistein. In addition, Heinonen et al (*148*) reported that a non-estrogenic or biologically inactive p-ethyl phenol as one of the metabolite of genistein. With the ingestion of genistein, daidzein and glycitein, 37%, 52% and 47% (respectively) are recovered in urine following metabolism (*151-153*) as unconjugated and conjugated forms (i.e. total concentrations).

Furthermore, it was explained by Axelson et al (*133*) that the extent of phytoestrogen metabolism varied in individuals, which could be as a result of other bio-active components in their diet amongst other factors, e.g. gut microflora bacteria. Following this lead, Cassidy et al. (*154*) investigated and found that as a result of intestinal fermentation on a high carbohydrate milieu diet, there was a noticeable difference of phytoestrogen metabolism. He observed in this *in-vitro* colonic model study that the production of metabolite equol was higher in a high carbohydrate diet.

Also, reasons for inter-individual differences in the ability to harbour the equol – producing and o-DMA - producing bacteria in the humans remain unknown (*145*). Controversial results by Adlercreutz et al (*46*), Rowland et al (*150*) and Lampe et al (*65, 155*) have been reported on the equol production ability based on diet.

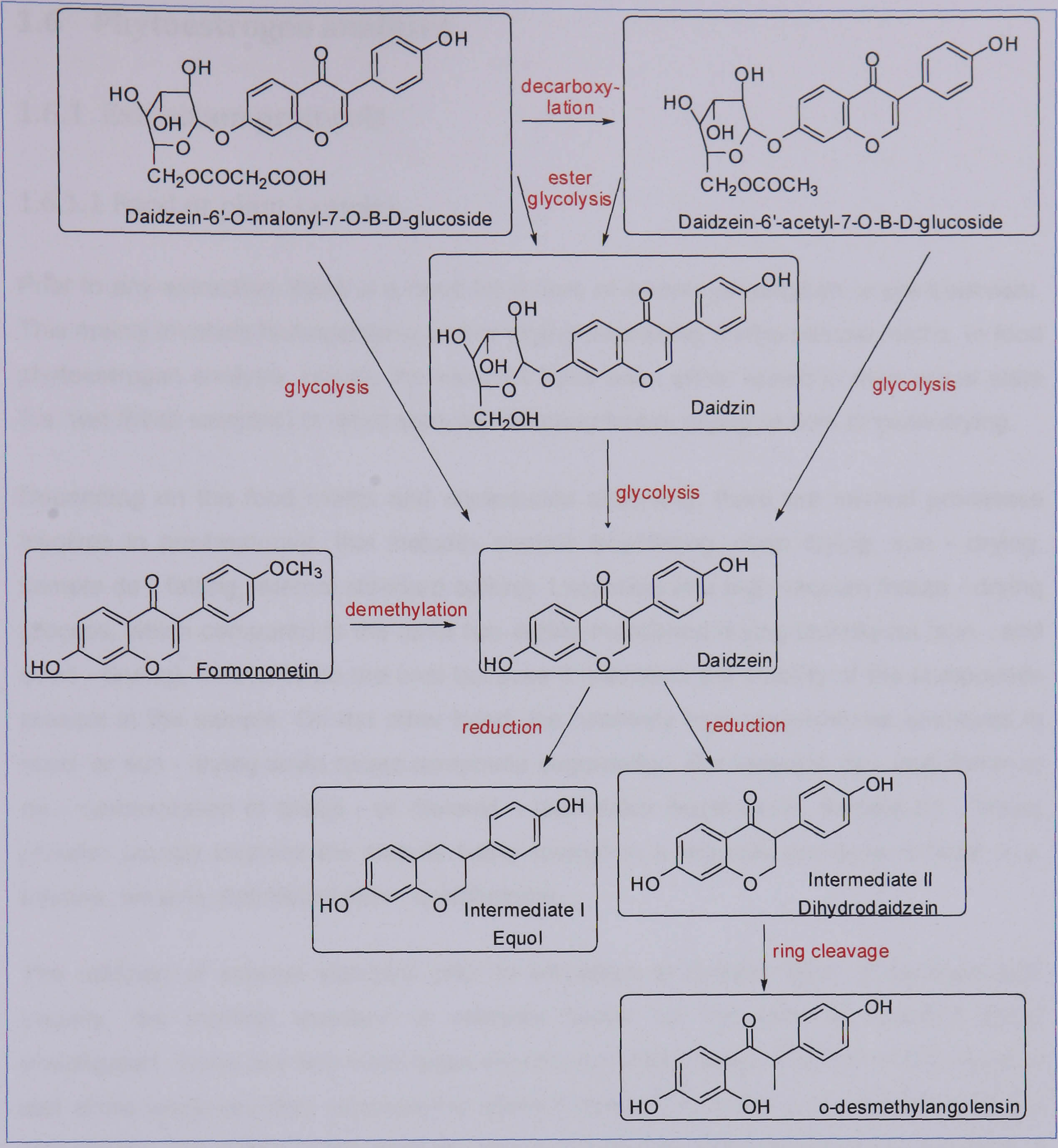


FIGURE 1.16: Schematic metabolic fate of daidzin (adopted from (*37, 38, 133*))

1.6 Phytoestrogen analyses

1.6.1 Extraction protocols

1.6.1.1 Food or plant samples

Prior to any extraction, there is a need for a form of sample preparation or pre-treatment. This mainly involves homogenising and or drying depending on the sample matrix. In food phytoestrogen analysis, usually the samples have been either tested in their actual state (i.e. wet /fresh samples) or dried samples following freeze drying or sun- or oven-drying.

Depending on the food matrix and compounds analysing, there are several processes involves in pre-treatment, this includes sample lyophilising, oven drying, sun - drying, sample de - fatting, internal standard spiking. Lyophilisis is a high vacuum freeze - drying process, which compared to the other two earlier mentioned drying techniques (sun - and oven - drying), seems to be the best because it maintains the stability of the compounds present in the sample. On the other hand, the relatively high temperatures employed in oven- or sun - drying could cause compound degradation. For instance, de - acetylation or de - carboxylation of acetyl - or malonyl – glucosides respectively. Sample de - fatting process usually involves the sample being soaked in a lipophilic non-polar solvent, e.g. hexane, toluene, petroleum ether, cyclohexane.

The addition of internal standard prior to extraction is another type of pre-treatment. Usually, the internal standard is selected based on the type of recovery being investigated. There are two main types of recovery experiments. The first is the recovery test of the whole process, whereby the internal standard is added at the beginning of the experiment. The other is the recovery test at a targeted step, whereby the standard is added at a certain step of extraction process, e.g. after the hydrolysis, with the aim of calculating the aglycones extraction recoveries only.

Usually, the sample pre-treatment process depends on the type of analytical platform to be used, whereby LC-MS needs little or no sample pre-treatment whereas GC-MS need extensive pre-treatment. The most common phytoestrogen extraction method from food is the solid extraction with aqueous organic solvents, such as acetonitrile, methanol, acetone, and ethanol. Occasionally, acidified solvents are used for a two-in-one step: extraction and hydrolysis; however, one of drawbacks is the phytoestrogens' stability in this condition. In general, because the phytoestrogen precursors in foods/ plants exist as

glycosides, this makes them highly soluble in water, and thus aqueous solvents are used for sufficient solubility. On this basis, several methods have been developed varying polarity (water composition), extraction temperatures (room temperature to 80/ 90°C depending on the solvent's boiling point) , times (2 - 24 h) and techniques (stirring, shaking, refluxing, rotary mixing, soxhlet). Typically, either a hot extraction followed by filtration or soxhlet extraction is employed. A number of isolation and extraction of these phytoestrogens have been carried out with varying conditions and techniques (*156-158*). A summarised table (TABLE 1.3) illustrates these extraction conditions.

Griffith et al developed an extraction of phytoestrogens from food using acetonitrile-water-dimethyl sulphoxide (DMSO) without acidification, thus extracting malonyl and acetyl isoflavones from soy proteins (*159*). However, more drastic extraction conditions using acidified or alkaline aqueous organic solvents have been employed (*81, 93, 159, 160*), whereby the phytoestrogens glycosides in their natural forms are hydrolysed and extracted in one step. The drawbacks with this extreme method are the isolation of only the aglycone forms and also possibly, there is some likelihood of degradation of some phytoestrogens. The preferred extraction method used is aqueous (70 – 80 %) methanol or aqueous ethanol over a minimum period of two hours either at hot temperatures (*52, 113, 161*) or low temperatures (*162*).

TABLE 1.3: Summarised extraction protocols from literature

FOOD SAMPLE	EXTRACTION CONDITIONS	ANALYTICAL PLATFORM	COMPOUND ANALYSED	REF
Soybeans	Time: 2 h; Temperature: RT Technique: Stirring Solvent: 58% ACN (aqueous)	HPLC	DAI, GLY, GEN, G-DAI, G-GLY, G-GEN, A-G-DAI, A-G-GLY, A-G-GEN, M-G-DAI, M-G-GLY, M-G-GEN	(*160*)
Animal origin foods e.g. dairy product, meat and fish	Time: - Technique: SPE Solvent: 10% MeOH in 0.1% sodium acetate, pH 5	LC-MS	BIO A, GEN, DAI, FOR, GLY, SEC,MAT, COU, EQU, ENL,END	(*70, 111*)
LEGUMES: e.g. kudzu root, chickpea, split peas, red clover, alfalfa sprout, lentil, cowpea, groundnut, black eyed beans Others: Flaxseed, soy flour, sunflower seeds, candy bar, bread.	Time: 15-24 h (overnight); Temperature: RT Technique: Soaking Solvent: Distilled water; hexane (if fatty food)	ID-GC-MS-SIM	BIO A, GEN, DAI, FOR, GLY, SEC,MAT, COU	(*52, 112*)
Cashew nut, mung beans, sesame seeds, asparagus, broccoli, carrot, celeriac, courgette, strawberry	Time: 2 h + 15 min ; Temperature: 40-70 °C Technique: Sonication + centrifugation Solvent: 70% EtOH in 0.1M sodium acetate, pH 5	HPLC-CED	Lignans only - SECO, ISOL, LARI	(*96*)

TABLE 1.3: Summarised extraction protocols from literature (cont'd)

FOOD SAMPLE	EXTRACTION CONDITIONS	ANALYTICAL PLATFORM	COMPOUND ANALYSED	REF
Cashew nut, mung beans, sesame seeds, asparagus, broccoli, carrot, celeriac, courgette, strawberry	Time: 3 h; Temperature: 40-70 °C Technique: Methanolysis Solvent: Sodium methoxide solution	HPLC	Lignans only- SECO, ISOL, LARI	[134]
Soybean meal and Soy-protein isolate (ProFarm873)	Time: 2 h; Temperature: RT Technique: Mixing (platform shaker)/ Sonication Solvent: 80% ACN-Hydrochloric acid (0.1 N)	HPLC	DAI, GLY, GEN, G-DAI, G-GLY, G-GEN, A-G-DAI, A-G-GLY, A-G-GEN, (M-G-DAI, M-G-GLY, M-G-GEN)	(*93*)
	Time: 2 h; Temperature: RT Technique: Mixing (platform shaker)/ Sonication Solvent: 80% MeOH			
	Time: 2 h; Temperature: RT Technique: Mixing (platform shaker)/ Sonication Solvent: 80% EtOH			

DAI- daidzein; GLY- glycitein, GEN- genistein, SECO- secoisolariciresinol, MAT- matairesinol, ISOL- isolariciresinol, A- acetyl, M- malonyl-, G- glycoside, LARI- laresinol, BIO A- biochanin A, EQU- equol, ENL- enterolactone, COU- coumestrol, FOR- formononetin, END- enterodiol. CED – coloumeric - electrode detector.

1.6.1.2 Biological samples

Regarding biological samples, usually these also undergo some form of pre-treatment, for instance with blood samples where serum or plasma is needed, the samples are centrifuged prior to extraction. Usually, in the extractions of phytoestrogens, there are two main isolation methods used, which are either liquid-liquid extraction (LLE) or solid-phase extraction (SPE). Prior to extraction, the samples are either hydrolysed to their bio - active forms or extracted as their inactive/ free forms (e.g. in urine as glucuronides). As a rule, the sample clean - up of biological samples should involve only a few clean-up steps, therefore less handling prior to analysis. However, Adlercreutz et al adopted the developed food extraction protocol for the extraction of phytoestrogens in blood/ urine (*12, 13, 46, 163, 164*). The developed extraction protocol involved hydrolysis, LLE and ion exchange column chromatography (two columns - DEAE and QAE).

1.6.2 Analytical techniques: chromatography & spectrometry

Chromatography is a well-known technique involving the separation of mixtures of compounds (analytes) based on their physical and chemical interaction between a solid stationary phase and a gaseous or liquid mobile phase. Chromatographic separation is governed by the polarity or boiling point of compounds in the test mixture.

Generally, there are four main types of chromatography namely: liquid chromatography (LC), gas chromatography (GC), thin-layer chromatography (TLC) and paper chromatography (PC). These types have been named and classed based on the nature of either their stationary phase (thin layer and paper chromatography) or mobile phase (liquid and gas chromatography).

TLC and PC are the least popular and oldest of chromatography. These two techniques are very similar in several ways. One main similarity is the capillary action separatory principle used to draw up solvent along a stationary phase of paper strip or thin layer of absorbent material (silica gel on a flat inert sheet) in PC and TLC respectively, whilst the analytes are separated during this process. Nonetheless, with the great potential of TLC technique, it has recently been modified with the automation of this process to high performance TLC (HP-TLC) to allow better resolution, more accurate quantitation and faster throughput analysis of samples.

Liquid chromatography was originated in the late 1890s by Russian botanist -Mikhail Semenovich Tswett, who called it liquid-solid chromatography, to separate mixtures of similar plant pigments (*165*). Following this study, two British biochemists (Nobel Prize winners)- Archer John Porter Martin and Richard Laurence Millington Synge developed liquid-liquid chromatography, which they used to separate some acetyl amino acids in protein. In their published study, the idea of gas-liquid chromatography and high-pressure liquid chromatography were conceived.

A few years later, GC was originally developed by Martin and James for the micro-analysis of fatty acids (*166*). GC is a technique used to separate mixtures of compounds, which are gases, volatile liquids and derivatised liquids. Typically, a liquid stationary phase and a gaseous mobile phase are employed in GC. GC involves a gaseous sample (if not, derivatisation is required) being injected into the head of a column and with the aid of the mobile phase, separation is achieved within the column based on their different boiling points and volatility.

General principles of chromatography

Prior to any chromatographic analysis, there is a need for sample preparation, which could be either minimal or intensive. In the case of GC, there is a need for derivatisation of involatile compounds, which is quite intensive and time-consuming; however, very minimal sample preparation/ pre-treatment is needed for LC analysis.

The principles of GC and LC are very similar; however, the key differences between these two techniques are the mobile phase and stationary phase. In GC, the mobile phase is chemically inert and gaseous- usually either helium or nitrogen is used as the carrier gas, whereas in LC, the mobile phase is a mixture of solvents, usually aqueous, organic or buffer solutions. In both cases the mobile phase selection is based on the chromatographic detector. Also, in most cases, the GC stationary phase is a column of liquid, gum or elastomer rather than that of LC being a column of porous solid with a thin film of liquid, could be water, aiding in equilibration of stationary and mobile phases during partitioning (*167*). Also, the separation of analytes in a mixture employing GC is based on mainly their boiling point differences; whilst this is not important in LC separation.

1.6.2.1 Gas chromatography

The principle of GC involves three main steps: sample injection, separation in column and peak detection.

Samples are introduced into a chromatographic system using a syringe firstly at the injection port either manually or automatically with an auto- sampler. In the injection system, for GC, following a rubber septum, the injector is usually a small heated space fixed to the top of a GC (*167*).

Typically in GC, this headspace is set at a higher temperature than that of the oven, so as to instantly volatise the sample. Prior loading onto the GC column head, there are two possible injection modes - split and splitless. In split mode, only a small proportion of the sample goes into the head of the column (1%) and remaining goes to the atmosphere; whilst in splitless mode, the entire injected sample goes onto the head of the column. Comparing these two injection modes, due to the venting of sample in the split mode, the sensitivity is reduced compared to splitless. On the other hand, due to the precision and speed of the injection in split mode, the accuracy and separation speed are improved compared to splitless mode.

The separation step takes place in the column, which is the stationary phase. Prior to separation of a mixture solution in any column, the stationary phase is equilibrated with a constant flow of mobile phase. Thereafter, separation in the GC column is achieved with the introduction of sample into a continuous flow of mobile phase in system by a syringe at the column head in either split or splitless injection mode.

The method of separation in the column is partitioning, whereby the analytes are intermittently interacting and exchanging with both the mobile phase and stationary phase. GC separations are thus based on the volatility and polarity of the test compounds (analytes). As the analytes in the mixture are attracted differently to the stationary phase, due to boiling point and polarity differentiations; there are consequently different mobilities of individual analytes resulting to different retentions on the column. Here, compounds with high volatility elute from the column earlier than those with low volatility.

The typical examples of GC columns are coiled, narrow and very long quartz columns containing liquid (nujol, didecyl phthalate), gum (carbowax, diethylene glycol succinate -

DEGS) and elastomer (silicone). These GC stationary phases are usually coated very thinly on a capillary column of varying lengths from 5 to 30 metres long.

Following the separation in the column, the separated analytes in the mixture (eluents) are introduced and monitored in a detector. In GC, the most universal detector is flame ionisation detector (FID), where the eluents is mixed with hydrogen and air and continuously burnt in a flame producing ions and electrons. These ions and electrons result in electric currents, which are monitored electronically by chart recorders resulting as the chart outputs - chromatograms.

On the chromatogram, each eluent resulting as an individual compound is represented as a peak. The area under each peak defines the quantitative determination of the compound in chromatograms. The peak is quantitatively proportional to the concentration of compound present (*168*).

1.6.2.2 Liquid chromatography

Liquid chromatography is a technique used to separate mixtures of compounds (analytes) in liquid form, employing a liquid mobile phase in which the analytes flows through to a solid stationary phase. These two phases (stationary and mobile) can be varied to greater extent to facilitate the best separation of compound mixtures compared to the other chromatographic techniques.

The principle of liquid chromatography also involves three main steps: sample injection, separation in column and eluent/ peak detection.

Samples are introduced into a chromatographic system using a syringe firstly at the injection port either manually or automatically with an auto- sampler. In the LC injection system, there is a loop injector where the sample (mixture of compounds) is loaded prior to its introduction to the head of an LC column (*167*). With the aid of an injection port valve in LC, sample is introduced using a switch mechanism: load/ inject into the continuous flow of mobile phase in the system. Thereby, forcing the sample towards the column (stationary phase) where separation takes place.

The second step is the separation of the components of the sample mixture, which takes place in the column (stationary phase). In any column, the stationary phase needs to be equilibrated with a constant flow of mobile phase prior component separation of a mixture

solution. The method of separation in the LC column is the same as in GC, which is partitioning, whereby the analytes are intermittently interacting and exchanging with both the mobile phase and stationary phase. LC separation is thus based on the chemical and physical properties of analytes, including polarity. As the analytes are attracted differently to the stationary phase, due to polarity and hydrophilic interactions differentiations; there are consequently different retentions in the column. Furthermore, in LC, the elution of compounds also depends on the type of LC mode used. The different types of LC modes are ion exchange chromatography, size exclusion chromatography, normal phase and reverse phase LC. However, the most popular is the reversed phase LC, in which the mobile phase is polar and the stationary phase is non-polar, thus, leading to the most polar compounds eluting faster than the less polar compounds.

There are different types of stationary phase packings /materials for liquid chromatographic techniques. The typical examples of LC columns are metal columns packed with materials such as octadecyl silane (ODS/ C18), octa-silane (C8), cyano (CN) and silica gel. However, the most popular LC column being C18 bound to silica.

In HPLC, the LC is equipped with a high pressure pump, which operates with a forward stroke piston movement that pushes solvent through a non-return valve on the column side, and the backward stroke draws even more solvent from the solvent reservoir into the pump cylinder (*169*). This pump is very helpful in creating enough pressure to move the mobile phase (solvent) through the stationary phase (column) towards the detector.

The last step is the detection and output displaying of the separated analytes. Following the separation of compounds in the column and with the help of a high pressure pump, the different compounds in the mixture (eluents) are introduced and monitored in a detector. There are different types of LC detectors, which include UV, fluorescence, electrochemical, mass spectrometry (MS) and nuclear magnetic resonance (NMR). The most popular and universal detection used to date is the UV detector. In UV detection, a radiation source of a deuterium lamp is used to monitor the analytes as they elute from the column at a specified wavelength (UV) or scan of wavelengths over a specific range (diode array - DA). Following these scans, electrical signals are sent to the amplifier, which records the detected voltages into a chart recorder resulting as the chart outputs-chromatograms.

Similarly, on the LC chromatogram, each eluent resulting as an individual compound is represented as a peak. The area under each peak defines the quantitative determination of the compound in chromatograms. The peak is quantitatively proportional to the concentration of compound present (*168*).

1.6.2.3 Hyphenated chromatographic techniques

GC-MS is one of the oldest and most successful hyphenations of separation techniques to a MS. As mentioned earlier, GC is based on a programmed temperature gradient separation with the most volatile components eluting first. In addition, GC offers a stable, precise, sensitive and high-resolution separation system. Although, capillary GC provides an efficient and high-resolution separation method, there are few practical problems in hyphenating the interface of the MS.

In GC-MS, the two main types of ionisation are electron impact (EI) and chemical ionisation (CI). EI has the advantage of better structural elucidation than CI because EI spectrum usually contains weak molecular ion showing extensive fragmentation into diagnostic daughter ions. CI, which is a 'gentle' ionisation, provides mostly molecular ion information, thus helpful in the conformation/detection of the molecular mass of the unknown compounds.

To improve, the potential of GC and GC-MS, just recently, a double GC (GC-GC) separation has been introduced, which has been coupled to MS detectors (ToF and ToF-ToF) (*170*). Thus, with the limitation of selectivity on a single GC separation, combining two different columns led to the improvement of separation by varying the analytes' selectivity on each column.

LC-MS hyphenation has proven to be less straightforward and more technically demanding than GC-MS. Similar to capillary GC; capillary LC also offers benefits for the analysis of complex mixtures, i.e. as an efficient and high-resolution separation method.

Furthermore, the discovery of the tandem LC (LC-LC) coupled to tandem mass spectrometry (MS-MS) has improved the excelled potential of MS. However, MS-MS was originally introduced as a way to attain fragmentation of ions generated in the source with the aid of a collision cell. The precursor ion is collected into this cell and then bombarded with higher energies to dissociate and produce the product ions, which are analysed with

the second MS. Following the fragmentation pattern of unknown compounds, structures of unknown compounds can be clearly elucidated.

1.6.2.4 Spectrometric techniques

General Principle of mass spectrometry

Mass spectrometry is one of the most powerful analytical qualitative and quantitative tools with a sensitivity of 1 µg to less than 0.1 pg. One of the many important advantages of MS is the sample separation diversity based on its hyphenation potential, therefore, leading to its scores of application in various analyses.

The principle of mass spectrometry involves four main steps, which are: sample introduction, ionisation of analyte gas molecules, *m/z* ion separation and ion detection.

In inlet system, the sample is introduced into the instrument by means of various types of inlet modes depending on the platform used, i.e. lone or hyphenated MS. Usually with a lone MS, the sample is introduced using a syringe or direct probe; whilst for the hyphenated techniques, possible inlets include GC, LC and CE. On the introduction by any of these inlets into a vacuum system, with the aid of a nebuliser gas, the sample is volatised into small droplets that travels into the ion source.

Following the volatising of samples, the formed droplets travel to the ion source where they are ionised/ charged. With the various ion sources, different ionisation agents are employed, which include: energetic electrons, ion beams (bombardment), reagent gaseous ions, high electric fields, high potential electrodes, high temperature, laser beams and fission fragments from ^{252}Cf. There are two main types of ion sources for MS, these are: gas phase (volatile) type- and desorption (non-volatile) type- ion sources. In gas phase type-, the ion sources are electron impact (EI), chemical ionisation (CI) and field ionisation (FI). Whilst for desorption type- sources include electro-spray ionisation (ESI), thermospray ionisation (TS), field desorption (FD), plasma desorption (PD), fast atom bombardment (FAB), matrix- assisted laser desorption ionisation (MALDI). Alternatively, these ion sources could be classed into two based on their ionisations: hard and soft ionisation. In soft ionisation, little energy is used for fragmentation, thus resulting in reduced fragmentation, whereas in hard ionisation, there is an excess of energy used for the fragmentation causing an extensive fragmentation. Hence, hard ionisation has the

potential of complete structural elucidation of unknown compounds compared to soft ionisation, where mainly the exact molecular mass/ elemental composition of an unknown compound is possible. Examples of hard ion sources are EI, CI and FAB, and that of soft ion sources are ESI, MALDI and FI.

With the ionisation of the samples, the mixture of charged ion droplets travel through to be separated based on *m/z* in the mass analyser. Examples of analyser are magnetic sector, double- focusing analyser, quadrupole, time-of-flight (TOF) and ion trap. The ion separations are based on kinetic energies of ions, angular aberrations or trajectory energies.

Following the separation in the analyser, the ions are pushed into the detector, where each ion is detected and recorded as mass spectrum based on its *m/z* and intensity. Examples are electron multiplier, photomultiplier and array detector.

Principle of electro-spray ionisation

Sample solution passes through a short length of stainless steel capillary tube, at which a high negative or positive electric potential 1 – 5 kV is applied. This high potential causes the sample to instantaneously vaporise or nebulise into a jet or spray of very small droplets of solution in solvent vapour on exiting the tube (FIGURE 1.18). To assist evaporation of droplets and breaking up of unwanted cluster ions, a drying gas (nitrogen) flows along and past the end of the capillary in the evaporation region. While the droplets move from the tube exit into the evaporation chamber, which can be heated slightly to avoid condensation, solvent evaporates rapidly from the surfaces and the droplets get smaller. As the droplets shrinks, there is an excessive positive or negative electrical charge on their surface resulting in electrical surface charge density increase, until the natural repulsion "coulomb explosion" between like charges causes the release of ions and neutral molecules. The end of the capillary tube is aimed at a target, which is a small hole at the opposite end of the evaporation region. After vaporising from the surface of a droplet, solvent molecules of low molecular mass conveniently diffuse away from the line of sight trajectory to the inlet target.

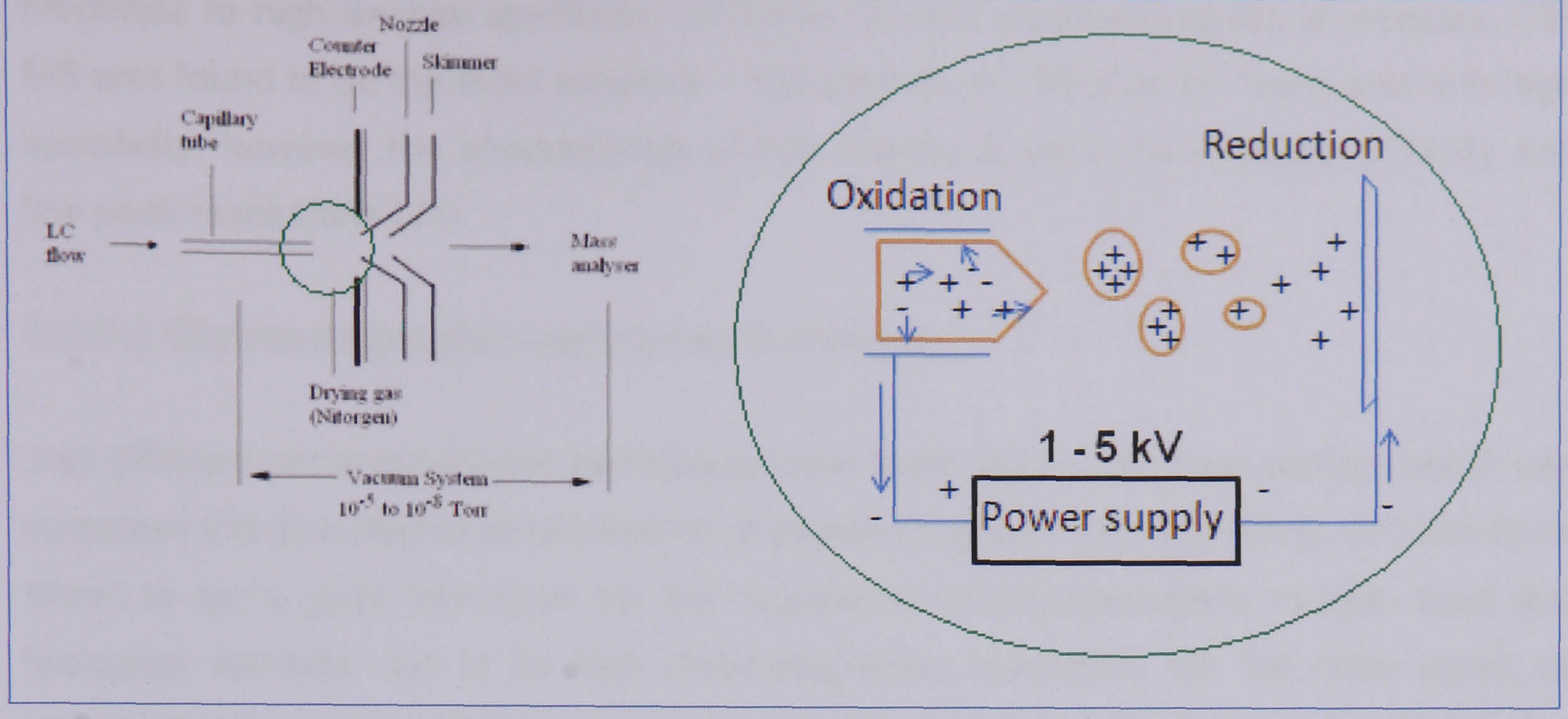


FIGURE 1.17: Schematic diagram of an electrospray.

1.6.3 Analysis of phytoestrogens

The novel discoveries of soy - phytoestrogen by Waltz (*18*) and metabolite - equol by Marrian (*19*) have led to numerous studies on phytoestrogens analyses. Some of the reported studies include food and biological sample analyses (*22, 24, 27, 33, 34, 38, 42, 43, 107, 162, 171-179*). However, especially, over the last decade, according to the popular search engine - web of science, there has been an increment in the number of annually published articles from 187 in 1998 to at least 330 from 2003 onwards (approximately 4152 in total to date) with 2006 peaking the chart with 453 publications. In these journals, different studies on phytoestrogens have been reported in different fields – including food analysis (*180*), biological sample analysis (*152, 181-183*), biological activities – *in vivo* and *in vitro* studies (*183-185*), estrogenic potency (*186, 187*), animal studies (*188-190*), human studies (*191*) and epidemiological studies (*126*).

A variety of techniques have been used in phytoestrogen analyses, which depend on what samples are being analysed – sample matrix, analytes' forms. Some of these techniques are: GC-MS (*46, 52, 163, 192-195*), HPLC-UV/DAD (*43, 91, 179, 196*), HPLC-fluorescence (*197*), HPLC-ECD (*198*), HPLC-MS (*196, 199, 200*), capillary electrophoresis (CE)- UV/DAD (*201*), CE-fluorescence (*202*), CE-ECD (*203*) (*204*), CE-MS (*205*) and RIA (*50, 190*).

Generally, the CE techniques coupled to any of the above mentioned detectors (ED, UV/DAD, MS, fluorescence) gave good sensitivity ranging from 0.1 to 50 fmol with limited/

moderate to high analyte specificity. Of these CE and chromatographic techniques, CE-MS was found to be the most sensitive - 100 attomol (0.1 fmol or 10^{-13} mol) and with high specificity, however the shortcomings of this technique were the interface difficulty and low peak resolutions (*91*).

1.6.3.1 Chromatographic separation techniques

Two different chromatographic techniques have been used in the past and are still in use to assess the quantitative determination of phytoestrogens – GC and HPLC. GC has been found to be a good technique for the separation of phytoestrogens in both food and biological samples due to its high chromatographic resolution. On the other hand, as phytoestrogens are involatile compounds, having at least one hydroxyl group; there is a need for the lengthy labour intensive sample preparations (derivatisation) prior analysis, which is a major drawback. Derivatisation to their trimethylsilyl (TMS) derivatives could be quite tedious, however the specificity of analytes on this platform compensates for this problem. This involves sample drying, addition of derivatising reagent, incubating at high temperature (usually 60 °C) for a minimum of 2 h, and lastly an optional solvent drying for achieving even higher response/ sensitivity. GC for quantitation purposes has been coupled to several detector platforms, especially MS. The great sensitivity and specificity of MS together with the high resolution of GC make this platform a mainstay of phytoestrogen analysis. This technique has even been improved in sensitivity (less matrix interference) with the possibility of monitoring chosen *m/z* molecular precursor or product ions of analytes over the chromatographic run. This method is called selected ion monitoring mode (SIM). Hence, these advantages have made GC-MS to be applicable in both food analyses and biological samples phytoestrogen analyses.

Adlercreutz and his research group developed the use of isotopic dilution GC-MS (ID-GC-MS) using stable isotopes of ^{2}H or ^{13}C labelled phytoestrogen internal standards in both sample analyses (*46, 135, 164*). The advantage of the isotopic labelled internal standard is the high precision of analyses; however these isotopic labelled standards are expensive.

In LC, different methods using various stationary and mobile phases have been developed for the separation and isolation of phytoestrogens in foods and biological samples. Some of these developed methods are being continuously modified with the

promising potential of HPLC coupled to a highly sensitive detector. The first LC separation of phytoestrogen was performed using C18 column with a UV detector (*43*).

Comparing the results from the either GC or LC coupled to different detectors (UV/DAD, fluorescence, ECD, and MS) reported for phytoestrogen analysis, LC-ECD was found to be the most sensitive - 20 fmol with good specificity but incapable of novel compound detection (*91*). Whereas, GC-MS was found to have the potential of novel compound detection as well as having high resolution, high specificity and high sensitivity- 50 fmol; although, intensive sample preparations (clean-up and derivatisation) are required and the chemistry complexity of derivatised compound are the drawbacks faced with this technique.

Also, comparing the GC and LC techniques; as there are many merits with the GC over LC platform analysis, there are also a couple of drawbacks with the GC compared with the LC. For instance, there are more potential of method development and modification using LC than GC. This is based on that fact that both stationary phase and mobile phases can more easily be changed with vast variations of columns and mobile phases compared to GC where the mobile phase could only be either nitrogen or helium with only a few selections of GC columns available. Besides these differences, there is an advantage of LC over GC; in which there is little or no need labour intensive sample preparation protocols (derivatisation) required prior LC. Here, the sample preparation involved in LC is filtering prior injection.

Generally in phytoestrogen analyses, as reviewed by Merken et al, the most popular LC method is the reversed phase which works on both non-polar and polar analytes separated on a non-polar stationary phase (e.g. C18) with a polar (aqueous solvent/ buffer) mobile phase (*61*).

The detection sensitivity with the universal UV detector was improved using fluorescence (*178, 179, 206*) or electrochemical detector (*174, 176, 177*). According to Setchell et al, ECD increased the sensitivity by a factor of two to six compared to DAD (*177*). Furthermore, the coupling of LC and MS has led to an increment in the sensitivity response and specificity of phytoestrogens as reviewed by Wang (*91*). The typical MS detectors used for these studies, including quadrupole (most popular) (*177*) and ion-trap (*159*).

1.6.3.2 Non - chromatographic separation techniques

The other non-chromatographic techniques used for the separation and analyses of phytoestrogens include capillary electrophoresis, radio- immunoassays and time-resolved fluoro-immunoassays.

Capillary electrophoresis offers advantages of rapid, high-resolution separation, very small sample volume (nL) (*91*). On the basis of most phytoestrogens being weak acids, capillary zone electrophoresis (CZE) has been applicable in phytoestrogens studies. As CE is very similar to LC, most of the LC detectors are compatible with CE. These include: UV (*201*), DAD (*54*), ECD (*203, 207*). However, comparing these to techniques on UV, Mellenthin et al reported the sensitivity of CE to be lower than LC due to the sample injection volume (nL) used; in addition, the separation resolution on CE is less than that of LC (*208*). Additionally, Rostagno, mentioned the poor quantitative reproducibility of CE in isoflavone analyses, as a result of inconsistent flow rates and injection amounts (*94*).

Spectroscopic methods have also been used for the estimated qualitative analysis of phytoestrogens in samples. These include: UV, infra-red (IR), near IR (NIR) and mid-IR (MIR) spectroscopy (*209*). As these techniques are rough, there is a need for sample clean - up, e.g. SPE prior analysis to avoid time-consuming data interpretation.

Wang et al was the first to use MALDI coupled to ToF-MS for the identification of isoflavones in soy samples (*210*). This was facilitated using 2',4',6'-trihydroxy-acetophenone and 2,5 - dihydroxybenzoic acid as the MALDI matrices for the isoflavones.

The application of immunoassays in the phytoestrogen analysis has the advantage of low running cost, high sample throughput, small sample volume and high sensitivity. However, the drawback is the imperfect specificity of the anti - bodies. The three major types used are time resolved fluoro-immunoassay (TR-FIA) (*190, 211*), radio-immunoassay (RIA) (*212, 213*), and enzyme-linked immunosorbent assay (ELISA) (*214*). These assays have particularly been used in biological sample analysis- urine, plasma and tissue. Of these three types, TR-FIA is the most sensitive technique (10 to 100X > RIA) and ELISA being the least sensitive (100X < RIA).

1.7 Phytoestrogen metabonomics

1.7.1 Introduction of metabonomics

Metabonomics originally known as 'metabolic profiling' was started in the late1960s with the chromatographic detection of some metabolites in human bio-fluids, following which in 1971, Pauling et al reported biological variability in urine samples analysed by GLC (*215*). However, one of the first reports published on metabonomics (GC used) was by Gates (*216*). Metabonomics is a new technique pioneered by Nicholson, whose research group defined metabonomics as "the quantitative measurement of time-related multi-parametric metabolic responses of an intact living system to patho-physiological stimuli or genetic modification" (*217*). Also, as stated by Pasikanti et al, metabonomics is used to identify subtle changes in metabolic profiles between biological systems of different physiological or pathological states (*217*). It measures multiple endogenous metabolites of complex system at various levels of bio-molecular organisation (cells, tissues or bio-fluids) resulting to the predictive modelling for different factor group and/ or disease classification with the use of pattern recognition chemometric statistical (multivariate) analysis. In a less complicated way, metabonomics seeks to measure metabolites produced by an organism and relate changes in these to the state of disease or intoxication (*218*).

Metabonomic studies are classified mainly into two major categories namely: class specific and dynamic studies. The class specific studies involve diagnosis or toxicological classification, whereby groups of control 'placebo' and unknown are identified; whilst dynamic studies involve temporal progression of a treatment, whereby different types of variations in the observations or objects are evaluated (*219*).

Metabonomics help in the discovery and identification of potential biomarkers influencing drug/ analyte metabolism by monitoring the metabolic profiles (*220*). These biomarkers could be the drug metabolites or small endogenous metabolites showing either the presence or metabolism of the drug, toxin or bio - active compound response. Hence, metabonomics shows a better understanding of the drug's effect in an *in-vivo* stimulated scenario. Metabonomics is interpreted by a combination of metabolism monitoring, spectroscopy or spectrometry data collection and multivariate statistical analysis methods (pattern recognition).

Levels of metabolites in bio-fluids can be influenced by environmental and lifestyle effects including diet, which greatly influences many diseases and modulate the complex internal community of gut micro-organism 'microbiome' (*221-223*). Evaluation of human bio-fluids samples are complicated by a high degree of normal physiological variations caused by genetics and lifestyle differences. The human bio-fluids studied in metabonomics include: urine, blood, seminal fluid, saliva, cerebrospinal fluid, amniotic fluid, blister fluid, cyst fluid, synovial fluid, lung aspirate, dialysis fluid and digestive fluid (*224, 225*). Of all of these fluids, urine and blood are the commonly used in metabonomics studies due to their non-invasiveness in collection; nonetheless, urine is more readily available compared to blood collection.

Pattern-recognition techniques application to human nutrition

With the consumption of food in human nurtition, many metabolites (both endogenous and diet induced metabolites) are formed, produced, circulating and possibly excreted. In the process of identificating and analysing these circulatory or excreted known and unknown metabolites in biological samples, different statistical pattern recognition techniques are carried out on the collected (transformed) spectroscopic data. The collection and analysis of these data are by statistical analysis software, such as AMIX® and SIMCA-P®. These softwares are designed to import raw spectral data, manipulate and transform to fit the necessary multivariate analysis, thus producing summarised overview analyses of multi-dimensional observations and variables.

In metabonomics, as a result of the complexity of collected and analysed spectral data, chemometrics multivariate statistical analysis and projection methods are applied to achieve latent information from these data. Multivariate statistical analyses visually describe and simplify the effects of diet on various metabolites, since it is capable of simultaneous analysis of metabolites in a single NMR or MS run. These multivariate methods include pattern recognition by principal component analysis (PCA), and classification analysis using projections to latent structures model by soft independent modelling of class analogy (SIMCA), partial least square- discrimination analysis (PLS-DA) and orthogonal-partial least square - discrimination analysis (OPLS-DA). SIMCA and PLS-DA are used for classifying, predicting or projecting toxic or therapeutic responses/ effects; although, PLS-DA is a more powerful discriminatory technique.

^{1}H-NMR spectra of biological samples are usually expressed as complex multidimensional data, which are then dimensionally reduced by binning to single variable sets. These independent binned variables are correlated to produce principal components. Following the analysis of the variances in PCA, classification methods are carried out on the data set. SIMCA is a type of supervised classification, where different PCA models are created for each known class of observations, with the hope of identifying unknown class or predicting new observations into the former models (*226*).

Principal Component Analysis (PCA)

Firstly, in pattern recognition techniques, data collected are evaluated by PCA (score plot, scatterplot, loading plot, distance - to - model - X (DModX) plot and correlation loading plot analyses). PCA is the simplest form of the true eigenvector-based multivariate analysis used for identifying patterns in data groups based on their differences. As mathematically defined by Joliffee (*226*), PCA is an orthogonal linear transformation, which transforms the data to a new coordinate system such that the greatest variance by projection of data comes to lie on the first coordinate (first principal component), the second greatest variance on the second coordinate, etc. Furthermore, PCA is useful in predictive modelling of data by calculating and progressing independent data axes under the Gaussian assumption.

Partial least square- discrimination analysis (PLS-DA)

PLS method is a classification based on the quantitative relationship between two data sets - X (chromatographic / spectroscopic spectral data) and Y (either quantitative values-concentration or qualitative values – class, gender, age group and treatment: control or disease) (*226*). Further to this, orthogonal PLS (O-PLS) method is used to identify the variables (X- factors) influencing sample classifications by splitting X data into two sets: one, which is linearly related to Y and the other, is unrelated (orthogonal) to Y (*226*).

Pattern - recognition techniques have been used to identify potential biomarkers in several studies based on the biological diversification of metabolites (*227-231*).

Analytical techniques used in metabonomics

Metabonomics studies have been principally carried out using NMR and MS. One of the first works done on endogenous metabolites using NMR was in early 1980 by Nicholson (*232*) and Bales (*233, 234*); whilst the first using MS was just recently developed as MS is

a more complex 2D - analytical platform for metabonomics by Fiehn (*235*). To date, many improvements on these two analytical technologies have resulted to their escalating uses in metabonomic analysis on complicated biological systems (cells, tissues and bio-fluids) was reported by various researchers (*236-241*).

NMR methods have been used extensively in metabonomics studies compared to MS, which only recently has been added as a complementary technique due to its sensitivity and molecule identification. The two determining variables responsible for separation of groups in pattern recognition techniques (PCA, PLS-DA, SIMCA) by MS are mass and retention times. Due to the potential of easier coupling/ hyphenation of MS to chromatographic techniques, several studies have been carried out to monitor the in-vivo studies of drugs/ toxins/ bio-functional compounds using GC-MS (*242-248*), GC-GC-MS (*249-251*), LC-MS (*247, 248, 252, 253*) and UPLC (*254, 255*).

Proton NMR is a technique that provides an overview of all protons contained in the compounds in a sample solution represented as spectral lines. Due to this potential, multi-component analysis of a vast number of metabolites can easily be carried out in a short period of time (5- 20 minutes depending on number of scans) on a wide calibration range. The application of some database libraries, e.g. Chenomx and Bruker-Biospin, has even made bulky and wide range quantitation easier. Typical proton NMR spectra of biological samples contain thousands of spectral lines that arise from hundreds of the most abundant small molecules in bio-fluids (*256*). In these spectra, 20 – 75 metabolites and over 200 metabolites have been identified to be commonly present in tissues and urine samples respectively (*257*). ^{1}H-NMR spectroscopic approaches generally provide high level of reproducibility and detail in each fingerprint.

Application of Chemonx NMR Suite in metabonomics studies

Due to complexity of spectral lines in ^{1}H-NMR based metabonomics studies; there is a need for deconvolution of NMR spectra. With the aid of this customizable library, search, identification and quantification of specific metabolites and endogenous biomarkers have been made possible and easier as shown in FIGURE 1.19.

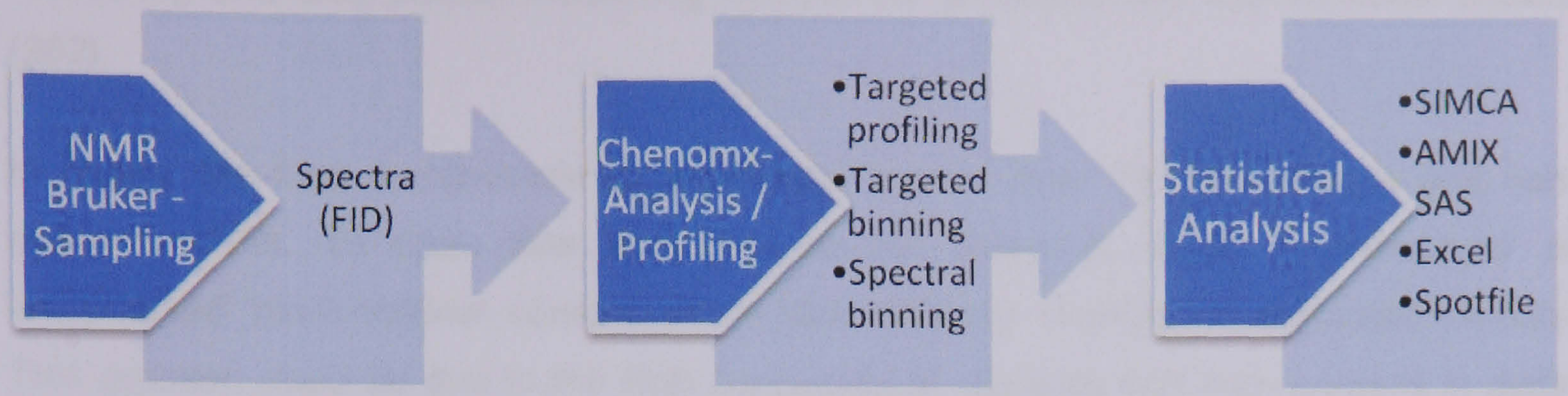


FIGURE 1.18: NMR/ Chenomx - based metabonomics flowchart

One of the many advantages of Chenomx software is its target binning and profiling potential. The principle of target binning and profiling is based on comprising of known compounds of interest, which have been defined into a database library and then targeted for analysis. This is a novel application in metabonomics studies for mixture analysis, whereby specific metabolites in biological samples are identified and quantified based on a database library with direct data comparison under potentially varying conditions, i.e. water suppression, baseline, noise, relaxation times, and scaling factors; followed by the validation of these profiled metabolites by PCA-based pattern recognition.

In target binning, three main factors are acknowledged before following the target binning algorithm, these are: experimental spectra, compound library and the pH range. All of these factors, if not carefully examined, would lead to the variations in the NMR detected peak position and line width (*257*). Here, multi-dimensionalities of data received are reduced to quantified metabolites, which are used as input variables in multivariate analysis, e.g. PCA and PLS-DA (*257*).

Some of these metabolites measured simultaneously using NMR with the aid of Chenomx have been studied individually or in groups employing other analytical platforms, these include: the ketone and amino groups (*258*).

1.7.2 Metabonomics in human nutrition

Metabonomics is used as a non-invasive type of *in-vivo* study. Besides its much application in toxicology and pharmacology, metabonomics has been newly introduced into human nutrition (*259-261*). The many applications of metabonomics in human nutrition include quality/ authenticity detection, food component analysis, consumption

monitoring and biochemical monitoring nutrition by controlled diet interventional studies (*262*).

In reality, the food metabolome is comprised of more than 15,000 nutrients and non-nutrients (*259*). To date, little is known on the extent to which 'hypothesised' or 'ascertained' bio-functional compounds in diet influence changes in metabolic profiles. This problem could be due to the high complexity of nutrients and non-nutrients in diets, which may cause analyte-analyte interaction. In addition, drug is amongst other important factor that influences these metabolic profile changes causing drug-analyte interaction. Accordingly, there is a need for a controlled interventional study in investigating metabolic effects of diets. So far, several diet induced (interventional) studies have been followed in line of metabolic effects investigations, even in non-human subjects. Some examples of these studies are whole grain and refined wheat flour on rats (*263*), hyperlipidic diets on rats (*264*), fresh leaf and green tea on rats (*265*), black tea on rats (*266*), soy products on humans (*230*), green tea on humans (*267*), etc. These studies have been made possible with the sophisticated and highly developed potential of analytical platforms especially NMR over MS, where all metabolites can be simultaneously analysed. Before the possibility of simultaneous analysis of a wide range of metabolites, previous studies on bio-fluids were restricted due the un-stretched potentials of analytical platforms used. For these reasons, metabonomics studies carried to date could be either targeted metabolic profiling (*268, 269*) or total metabolic profiling (*270*).

Nonetheless, due to the lack of sensitivity of NMR, immunoassays and MS have been adopted for metabolic profiling/ metabonomics. Studies carried out monitoring oxidative stress and other effects are shown in TABLE I.4.

TABLE 1. 4: Metaboniomics studies reported in literature

DIET	BIOCHEMICAL EFFECT	METHOD	REF
Soy	Osmolytic fluctuation Energy metabolism	NMR	(*231*)
High pressurized Vegetable soup - Gazpacho	Oxidative - stress Inflammatory	IMA	(*271*)
Red grape juice	Oxidative - stress Inflammatory Lipidemic	IMA	(*272*)
Fruits & vegetables	Oxidative - stress	IMA	(*273*)
Extra virgin olive oil	Oxidative - stress Inflammatory	IMA	(*274*)
Whole grain*	Oxidative - stress	NMR	(*263*)
Macadamia nuts	Oxidative - stress Inflammatory	NMR	(*275*)

All human subjects except * where rat subjects were used.

In addition, to the advancement of metabonomic studies, challenges faced by these studies have been addressed and corrected. Further to this, based on the convolution of the collected multi-parametric spectroscopic data, the two main problems are the complexity of data and identification of all the metabolites present in the bio-fluid. Comparing the complexity of data collected on the two metabonomic compatible analytical platforms, MS has two major disadvantages over NMR (*276*). One is that MS has a higher sensitivity; roughly thousands times greater, thus more complex data are collected and consequently harder compound identification. And secondly, MS has a lower selectivity, whereby only certain ionisable compounds can be observed, whilst in ^{1}H-NMR, all metabolites contain at least one hydrogen atom, which are observed in the spectrum. However, these spectra complexity problems have been partially solved by compiling known metabolites in metabolome libraries, where common metabolites in bio-fluids are identified and analysed. To date, some libraries containing information, including spectra on 100 – 300 metabolites have been developed; these include Chenomx, Bruker, PubChem and Biocrates. Ideally, the problem faced with complexity of spectral data could be solved with the development of an extensive metabolome library with the complete identification of all of the metabolites using lone or coupled analytical platforms.

Considering all of these issues, there is still a need to build a small consensus metabolome library containing a list of commonly present metabolites in bio-fluids. These metabolites should be present as a result of priority metabolism and pathways to carbohydrate, fat, protein, energy, mineral, trace elements and mineral metabolism (*259*). Other pathways worth considering include: reproductive, satiety, tissue-specific, signalling and cell regulatory (*259*).

1.7.3 Biochemical effects of phytoestrogen on other metabolites

The effect of the phytoestrogen on estrogen following a soy intervention was first reported by Xu et al (*277*). It was established that with increased isoflavone consumption, the urinary excretion of the total estrogens and their hypothesised metabolites were reduced, and thus, suggesting phytoestrogens anti - cancer protective properties.

In metabonomics, with the improvement and capabilities of analytical platforms like MS or NMR, very little work has been conducted on the effects of phytoestrogen on small molecular metabolites. It was only in 2003, that Solanky's research group were the first to investigate the biochemical effect in metabolic modification on the human plasma profile following soy interventional study (*231*). Following this, another similar study on urine samples was reported by the same group (*230*). It was concluded by this research group that isoflavones have effects on the metabolism associated with osmolytic fluctuations and energy (from lipoprotein, amino-acids and carbohydrate profiles) metabolism.

1.7.4 Metabonomics in physiological monitoring of human nutritional diet

Methods of monitoring the effects of a limited number of specific metabolites brought about the development of metabonomics in NMR and MS, where simultaneous analysis of several essential metabolites can be studied. These led to the exploration of biochemical effects of certain compounds in different nutritional foods. However, it has been reported that subtle changes have been noticed in some key metabolites following a diet intervention. These include: allylmercapturic acid in garlic (*278*), polyphenols in chamomile tea (*279*), salicyluric and salicylic acid in fruits and vegetables (*280*), quercitin in fried onion (*281*), isoflavones in soy product (*230, 231*) and caffeine in coffee (*282*). Generally, the effects on the human gut micro - biota and the term effects following

interventions can be deduced depending on the design of experiment (e.g. time length of study, sampling type).

1.7.5 Endogenous metabolites

Endogenous metabolites are products of metabolic pathways. The largely studied metabolisms and their pathways include carbohydrate, energy, fat and proteins. Some of the metabolites produced by energy metabolism are the ketone groups. Ketones are energy end products of excessive fatty acid breakdown rather than carbohydrate breakdown. The compounds in this group include acetone, acetoacetate, acetoacetic acid, 3-hydroxybutyrate. These are only present in the urine if their level in blood goes above a certain level (0.6 mM) after being filtered by the kidney. This condition is likely to occur when carbohydrates are in short supply, cannot be metabolised or not free for metabolism. Smith et al. (*283*) found that the acetone levels were related to the ovulation cycle using women of different groups, i.e. pre- and post - menopausal women volunteers.

Ketone classed metabolites, which include acetone and acetoacetate were previously quantified using spot - test strip readers by reflectrometry (*258*) and acetone using colorimetric method (*258*). However, due to the inadequacy and inability to perform multitude metabolites' (different classes) analyses simultaneously using these techniques, there is the need to investigate better techniques. These techniques should be more sensitive, less sample preparation required, shorter time of analysis and the potential to analyse different classes of metabolites simultaneously.

The metabolites produced by protein metabolism are amino acids and their derivatives. These are compounds containing both amine and carboxyl functional groups (carboxylic acids). Generally, amino acids are very important in life (following water) as they form the building blocks for proteins in the body and synthesis of nucleic acids. There are two types of amino acids: essential amino acids and non - essential amino acids. Essential amino acids are those supplied in the diet and cannot be synthesised by the body. The eight examples are phenylalanine, valine, leucine, lycine, tryptophan, threonine, isoleucine and methionine. Conversely, non - essential amino acid are those amino acids that are synthesised by mammals via the essential amino acids or the break down of proteins: thus, they are not necessary in the human diet. Examples include alanine, glycine, glutamine, glutamic acid, praline and tyrosine. Halvor et al was the first to report the analysis of two non-essential amino acids: alanine and glycine in urine following the ingestion of these amino acids by human subjects. Alanine is an amino acid acting as one

of the aliphatic side chain of a protein structure and it is important for the metabolism of glucose and tryptophan (aromatic side chain). It is synthesised from pyruvic acid, which is an intermediate breakdown product of carbohydrate. Excessive high levels of alanine have been suggested to be linked to glandular fever and chronic fatigue syndrome. Sources of alanine include: meat, poultry, eggs, fish and dairy products. Other aliphatic side chains of the protein include: glycine, isoleucine, leucine, praline and valine.

Pyruvate is an endogenous metabolite of glycolysis, which is an aerobic metabolism of glucose. A glucose molecule breaks down in the cytoplasm into two molecules of pyruvate, where it goes through energy metabolic pathways (Kreb's cycle). The Kreb's cycle is an energy metabolism cycle resulting to the formation of vital ATP or ADP. Besides being produced by the body on the consumption of sugar products, pyruvates are present in some food sources. Following the breakdown of glucose into pyruvate molecules, carbon dioxide molecules are removed in the mitochondria as waste products leaving behind acetaldehyde. Continuing the process in the Kreb's cycle, other organic compounds (NADH, water, acetyl CoA, oxaloactate, Co-A) are added, some parts removed to form citrates.

Citrate is an intermediate substrate of the Kreb's cycle, thus it is important in glycolysis. It is important for the synthesis of fatty acids synthesis because it carries acetyl groups across the inner mitochondrial membrane (*284*). Citrates are also involved in carbohydrate catabolism, which is the breakdown of sugar - glycolysis (energy metabolism).

Glycine was first isolated from gelatine and it is synthesised from serine and threonine in the liver. The roles of glycine include acting as an inhibitory neurotransmitter in the central nervous system, improving glycogen storage by freeing up glucose for energy needs, retards muscle degeneration by supplying extra creatine in the body and increasing the conformation range of folding polypeptide chains (*284*). Also, it is useful for the synthesis of haemoglobin, glutathione, DNA and RNA, as well as serving as a basic nitrogen source for the manufacture of many other amino acids. No clear toxicity has emerged from glycine studies. Foods found to contain glycine are the high protein foods - fish, meat, poultry, dairy products and beans. Glycine can be methylated to dimethylglycine (DMG), which is part of the one-carbon pathway that allows for the donation and acceptance of methyl groups. The one - carbon pathway is extremely important for the synthesis of

steroids such as the androgenic and estrogenic hormones as well as cortisone-like hormones

Creatine is a metabolite of glycine, which is an amino acid metabolism metabolic intermediate. It is an energy by-product found in the liver, where it is transported into muscles and some of which are further metabolised to creatine phosphate. Although, creatinine is a breakdown product of creatine phosphate, the body actively secretes only a small ratio. As both creatine and creatinine are found in muscles, they are dependent on skeletal muscle mass; hence, women have lower levels than men. Also, creatinine is used to monitor the normality of the kidney functioning. The level of creatinine is usually quite fixed with a normal diet; however it has been found that this metabolite increased 10 – 30 % higher in high - meat consumers. Additionally, it was reported that with age, as the muscle mass reduces and kidney functioning reduces; only slight changes have been noticed in these metabolites levels.

Protein waste products, creatinine and urea levels/ ratios are used for the test of the effectiveness of kidney waste filtering from the blood, i.e. renal functioning. This measurement reflects glomerular filtration rate (GFR), which is clinically important (*285*). The level of measured metabolites is inversely proportional to the effectiveness of the kidney (*286*). Urea is produced from ammonia and readily excreted through the kidney to urine.

Hippurate is an endogenous conjugate that belong to a group of uremic toxin (*287*). It is an ester or salt of hippuric acid; benzene with an amino acid – glycine attached to it. Hippurate is synthesised in the kidney and liver from benzoate and glycine (*288*). It takes part in renal ammoniagenesis by modulating γ-glutamyl transpeptidase (γ-GT) activity (*288*). In addition, it was reported that hippurate controls metabolic acidosis (*287*).

Methylamine (MA), dimethylamine (DMA), trimethylamine (TMA) and trimethylamine-N-oxide (TMAO) are ammonia derivatives with a characteristic fishy/ ammonia-like odour. MA is the simplest form of these derivatives. TMA is a nitrogenous-based decomposition product of plants and animals. It has been associated with taking an excess of choline and carnitine. TMAO is an oxidation product of TMA.

1.8 Hypothesis

Comparing the incidences of certain chronic diseases in the African population with the Western population, for instance, there are about 3 fold less affected African by breast cancer (incidence) compared to Europeans (developed countries) (*289*). In addition, according to the American cancer society (ACS) cancer atlas, the incidence of breast cancer is significantly lower in Japan, Thailand, Nigeria, and India compared to developed Western countries - Denmark, New Zealand, U.K and the United States. Supporting these incidence differences in part may be justified by genetics or cultural and geographical differences leading to the environmental factors including dietary habits. Diet is influenced by culture and environment amongst other factors; racial differences in relation to certain diseases are still inconclusive. To date, there is no knowledge on the background of the phytoestrogens in Africans.

Following the pre-mentioned observations and ideology, this led to a set of questions raised. The first question being: what are the general concentrations of phytoestrogens in African foods? Based on the findings reported by researchers (*115, 116, 290*) showing that there are variations in phytoestrogens concentrations due to location differences, there is a need to analyse specific food samples consumed in their respective continents rather than generalising phytoestrogen contents based on food similarities. Thus, there is a need to establish the phytoestrogen content in some African – grown staple foods, which have not been reported before.

The second set of questions include: do Africans consume phytoestrogens and at what exposure concentration level? To what extent do Africans metabolise phytoestrogens? Are there differences in metabolism in African based on any influencing factor? As earlier reported in the East Asian countries, do phytoestrogens have the same effects on other populations? Further to this on soy - metabolism, Rowland et al (*150*) reported phytoestrogen metabolism to be individual - dependent based on their habitual diet and microflora state. With these findings and other researches (*155, 291, 292*), only selected populations have been considered, e.g. East Asians, Europeans, Americans. Thus, this led to the need to understand soy – phytoestrogen metabolism in Africans following soy - intervention 'soy – challenge' and investigating their possible metabolism - influencing factors. The three influencing factors considered were geographical location, gender and age differences.

Other questions considered in this research include: what are the effects of phytoestrogens on other metabolites following soy - intervention, particularly for a population with a diverse diet? Do phytoestrogen metabolites have any relationship with endogenous metabolites? What are the factors influencing the differential metabolism of phytoestrogens? What group of people would benefit most with an interventional diet? To date, only a few studies have been conducted and reported on the ^{1}H-NMR based metabonomics being a fairly new technique in relation to soy - phytoestrogens (*230, 231*). So far studies reported have been selectively and limited in subject sampling and result analysing. Hence, there is a need to further explore metabonomics using Africans as the case study. Additionally, to investigate the relationship between soy - metabolites endogenous metabolites, both a robust analytical technique for multitude analyses of metabolites simultaneously and statistical softwares for quantitative and statistical analyses are required to be employed

1.9 Aims and Objectives

The phytoestrogen family in the long history of natural medicinal products has already proven to be a remarkable inspiration and challenge in the world of therapeutic science. This current research is based on three major frameworks- food analysis, soy - phytoestrogen interventional study and ^{1}H-NMR metabonomics. The aims and objectives of the work described in the thesis are to contribute to the science research programme as follows:

In the food analysis, the principal aim was to analyse and quantify phytoestrogen concentrations in some African staple foods. The objectives were:

1. Develop reliable and efficient extraction and isolation of phytoestrogen from six staple African grown foods - beans, cassava, plantain, pumpkin leaves, rice and yam.
2. Develop, validate and implement novel chromatographic methods to quantitate the phytoestrogens (aglycones) – biochanin A, coumestrol, daidzein, formononetin, genistein, matairesinol and secoisolariciresinol present in African foods.
3. Investigate novel phytoestrogenic compound(s) present in these foods.

In the soy interventional study, following the recruitment of fifty healthy African subjects (25 UK and 25 Nigeria), the aim was to investigate what factors influence the different metabolism rates of phytoestrogen following this soy-interventional study. The objectives include:

1. To compare, develop and optimise extraction and isolation of phytoestrogens in the collected urine samples.
2. To develop, validate and implement novel chromatographic methods to quantitate the phytoestrogens (aglycones) and their metabolites – coumestrol, daidzein, o-desmethylangolensin, equol, enterodiol, enterolactone, genistein, matairesinol and secoisolariciresinol present in urine, using either formononetin or biochanin A as internal standard.
3. To quantitatively determine the average phytoestrogen consumption based on random spot urinary excretion level of phytoestrogen and their metabolites on these UK and Nigeria subjects.
4. To assess the extent to which Africans metabolise soy - phytoestrogen

In the final study, ^{1}H-NMR metabonomics study, the aim was to investigate what factors influence the differences in individual metabolic profile of phytoestrogen and their metabolites following this soy-interventional study using MVA and the objectives include:

1. To assess phytoestrogens influences on other metabolites: *in-vivo* study
2. To quantitatively analyse the biochemical effects of soy on other key metabolites observed in human urine.
3. To understand the *in-vivo* relationship of phytoestrogens on other metabolites using ^{1}H-NMR, with the aid of statistical chemometrics softwares – Chenomx Suite, SIMCA-P, AMIX.

Chapter II

Phytoestrogens in some African staple foods

2.1 Introduction

Analyses of phytoestrogen in foods are deemed important as the 'groundwork' to determine their intake in populations leading to the investigation of their potential health effects- antagonistic or agonistic effects. Several extraction methods of phytoestrogens in foods have been developed (*64, 70, 92, 173, 203, 207, 293-301*). Quantitations of phytoestrogens in foods have been extensively studied using various analytical techniques. Chromatographic techniques (GC and LC) have been the popular separatory techniques used, just as UV/ DAD and MS are the popular detection techniques (*48, 157, 158, 173, 192, 299, 301-304*).

Plant-based foods especially fruits and vegetables are the known key sources of phytoestrogen (*74*), however they are also present at diminutive levels in animal-based food as well (*111*). So far, the levels of phytoestrogens including isoflavones, lignans and coumestans have been studied extensively in some cases, within the commonly consumed foods in East Asia (*193, 305*), Europe (*132, 306, 307*), America (*53, 123, 180*) and South Asia (*308*). It was from these studies that the foods found to contain the highest levels of isoflavones, coumestans and lignans are most soy products, alfalfa sprouts and flaxseed respectively.

In this chapter, an efficient and reliable extraction method was developed with the modification of the '*gold-standard*' method developed by Adlercreutz et al (*309*). Following this, fairly novel and rapid simultaneous analyses of isoflavones, coumestan and lignans (FIGURE 2.1) were developed using GC, HPLC and UPLC coupled to MS. GC-FID was also considered for this analysis.

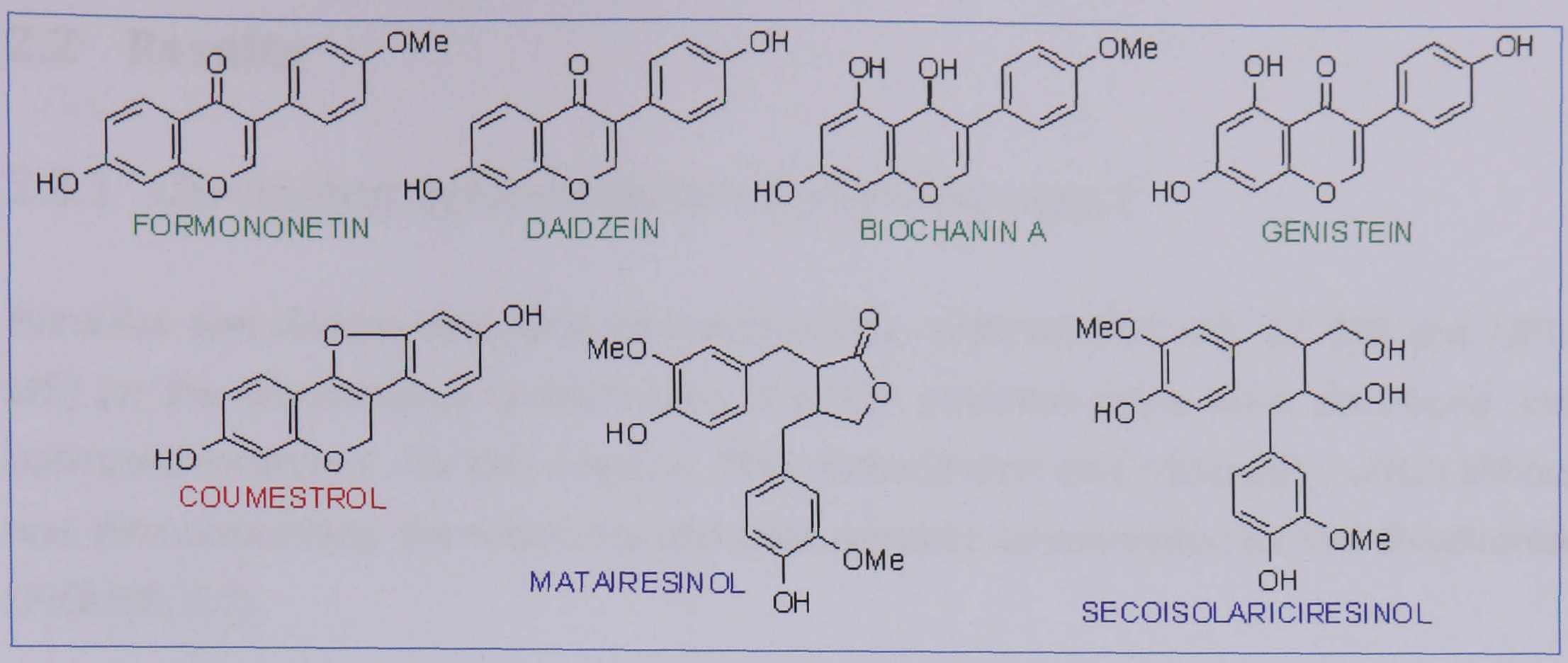


FIGURE 2.1: Structures of the seven phytoestrogens analysed.
Colour labelled: green- isoflavones, red- coumestan and blue- lignans.

Both GC and LC were employed to investigate their sensitivity in simultaneous quantitation of seven phytoestrogens, namely, formononetin, biochanin A, genistein, daidzein, coumestrol, secoisolariciresinol and matairesinol. GC-MS, LC-MS and UPLC-MS in the selected ion monitoring (SIM) mode proved to be highly sensitive, precise and reliable for the simultaneous analysis of all the seven compounds in the food analyses.

Although the information on phytoestrogen in foods studies has increased over the years, there is no entry on the existing database available on their composition/ levels in African foods. Consequently, this led to the analysis of six foods grown in Africa, which are commonly consumed African foods, namely: plantain, cassava, pumpkin leaves, puna "sweet" yam, rice and brown beans. These are chosen as they best represent the major food groups consumed by Africans, namely roots and tubers, legumes, cereals, vegetables and fruits. Roots and tubers, cereals and legume form the major source of nutrients in Nigeria (*310, 311*).

Even though GC-FID has low selectivity, simultaneous separation was achieved. Furthermore, incorporating this technique with the coupling to a mass spectrometer operated in the SIM mode achieved the sensitive separation of all ions based on the monitoring of the selected ions during the run (in MS) and partition equilibrium (in GC or LC). GC-MS-SIM, LC-MS-SIM and UPLC-MS-SIM methods were developed and applied to quantitate the seven phytoestrogens in the six African foods. Currently, no published documentation has shown the simultaneous quantification of these compounds.

2.2 Results

2.2.1 Chromatographic analyses of phytoestrogens

Sensitive and reliable analytical chromatographic methods (GC-MS, LC-MS and UPLC-MS) for the simultaneous quantification of seven phytoestrogens were developed under optimized conditions. For GC analysis, TMS derivatisation was necessary, which although was time-consuming, the sensitivity and peak capacity compensated for this disadvantage (FIGURE 2.2).

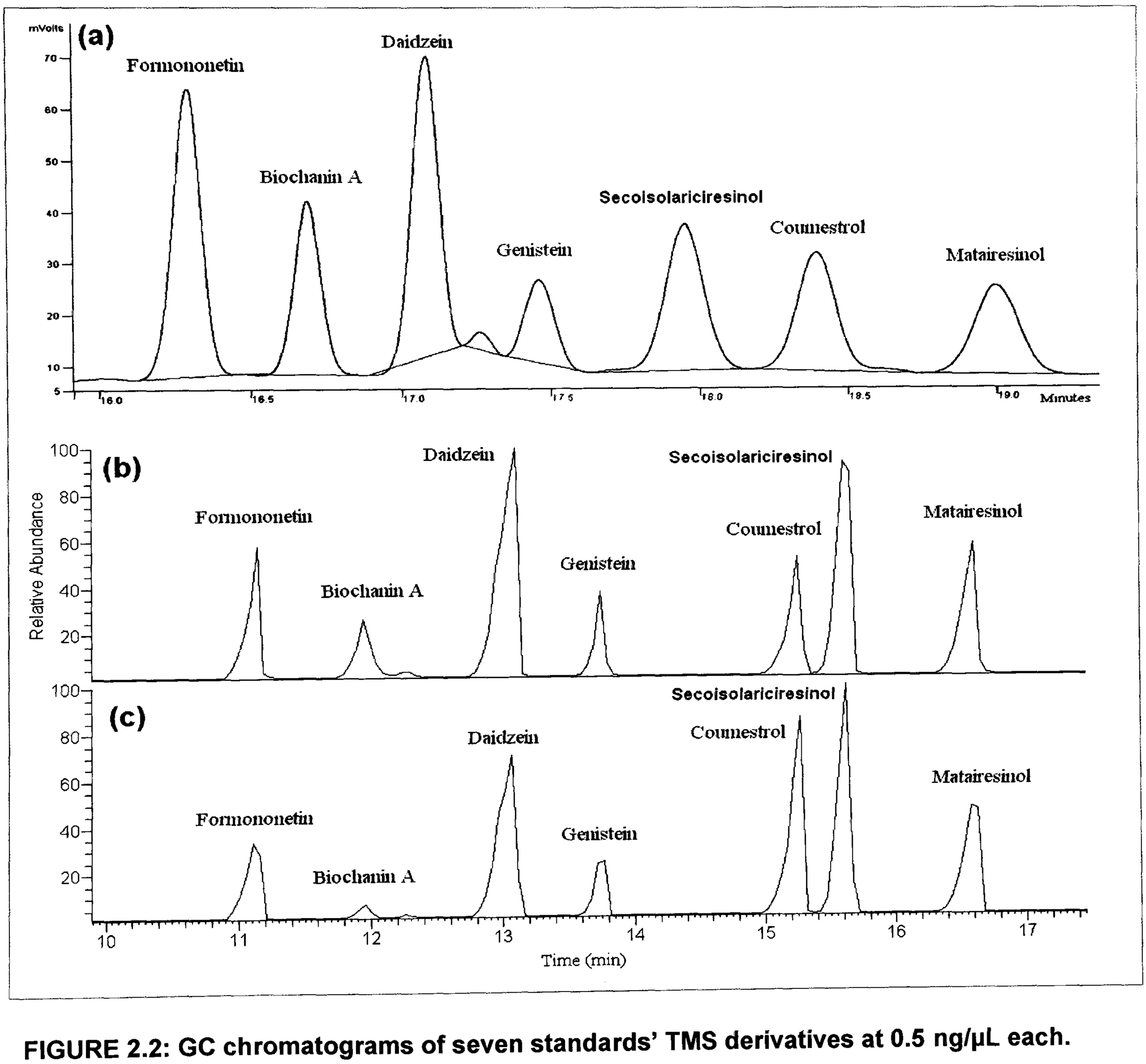


FIGURE 2.2: GC chromatograms of seven standards' TMS derivatives at 0.5 ng/μL each. (a) GC-FID; (b) GC-MS (TIC); (c) GC-MS (SIM).

In GC-MS, the SIM mode was done simultaneously for all the standards due to their distinct peak separation, whereas in LC-MS, more overlapping peaks were observed in the TIC mode, thus SIM was done singly for each standard (FIGURE 2.3).

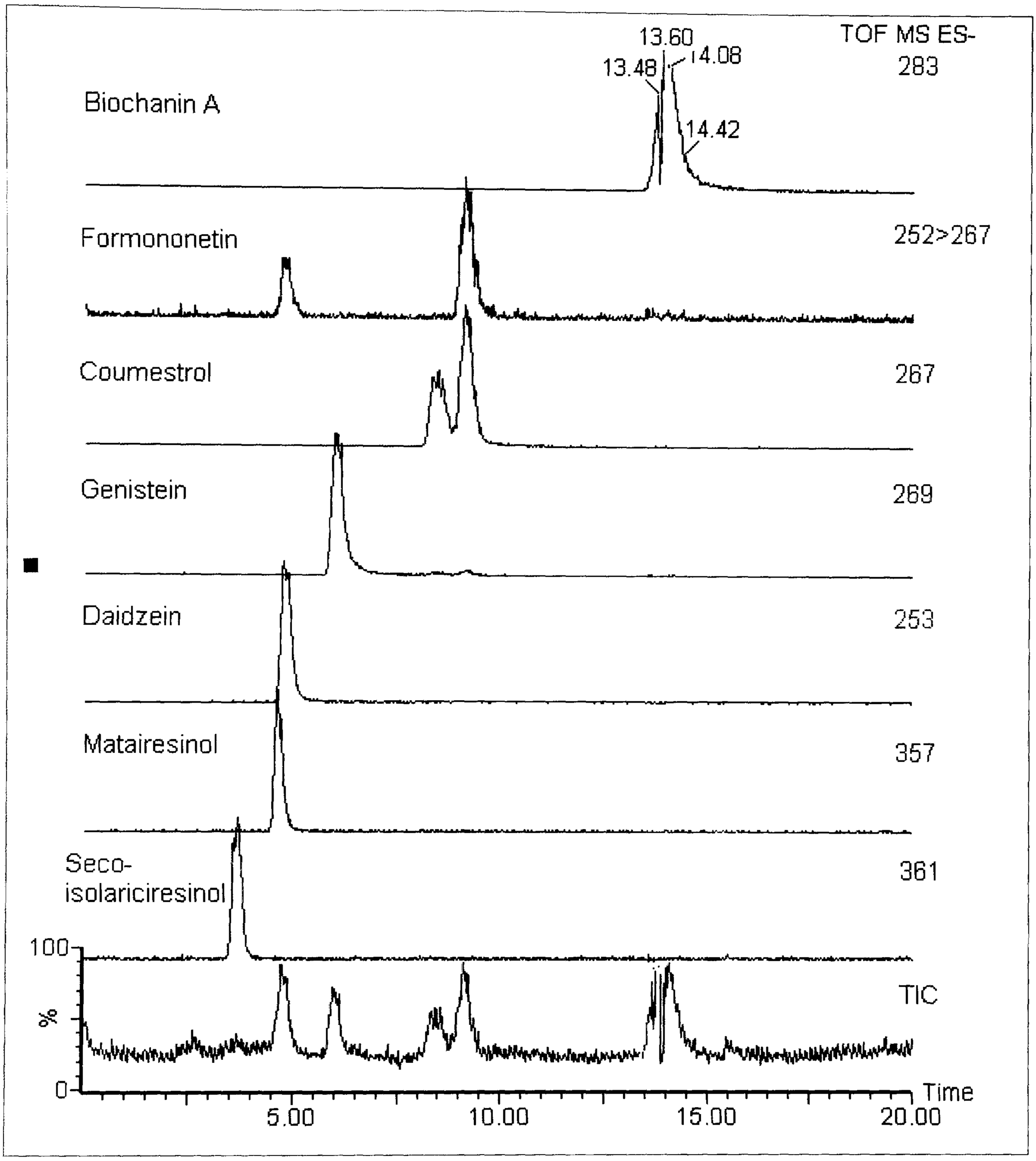


FIGURE 2.3: LC-MS (SIM and TIC) chromatograms of seven standards at 1 ng/µL each.

The elution order in LC-MS was secoisolariciresinol, matairesinol, daidzein, genistein, coumestrol, formononetin and biochanin A. The UPLC-MS elution order was slightly different compared to LC-MS: secoisolariciresinol, daidzein, matairesinol, genistein, coumestrol, formononetin and biochanin A, as shown in FIGURE 2.4

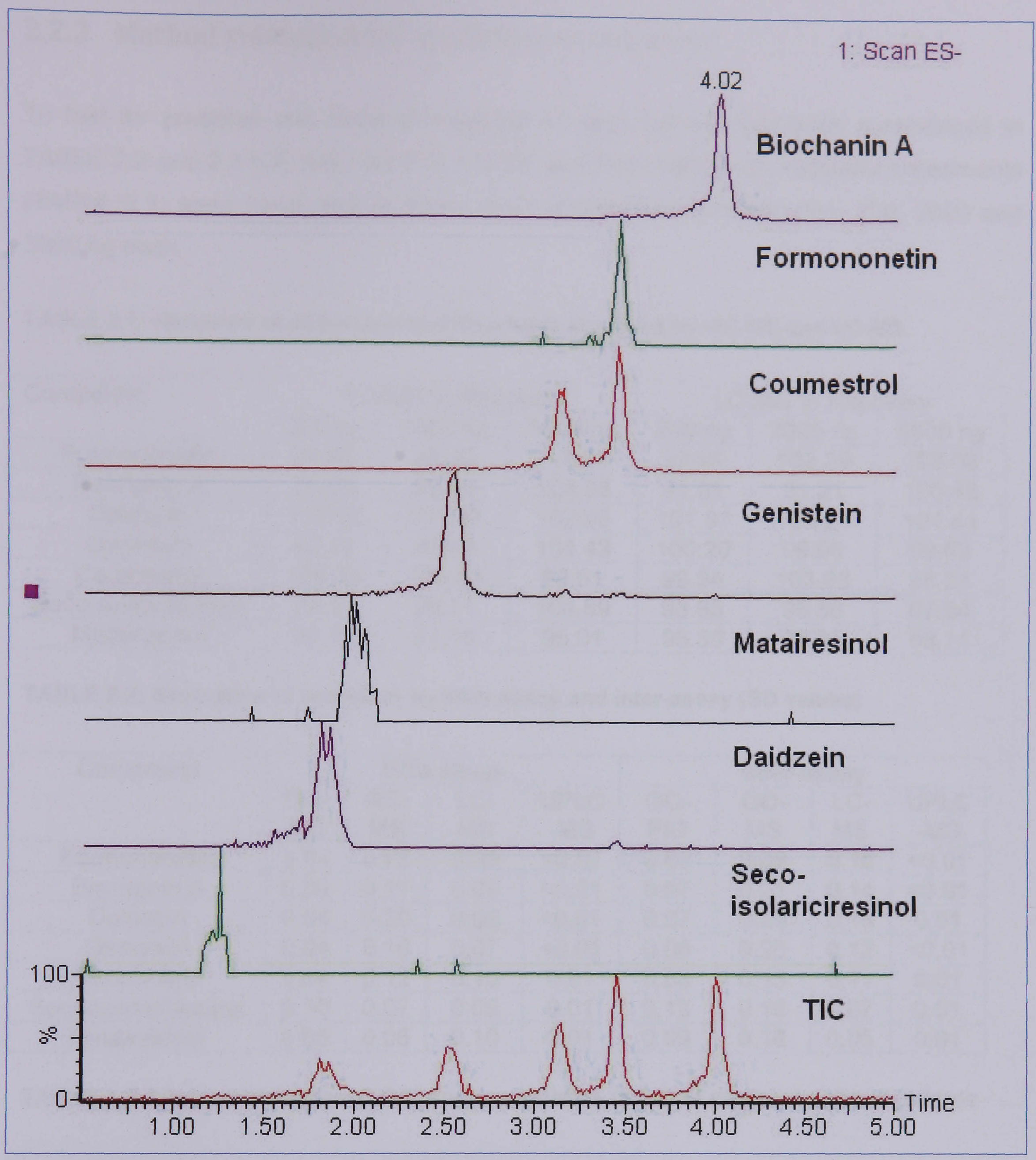


FIGURE 2.4: UPLC-MS (SIM and TIC) chromatograms at 1 ng/µL each

Following the development of optimum separation methods, method validation was carried out on each chromatographic method. The typical validation characteristics considered include: accuracy (trueness/ recovery), precision (repeatability/ intra-assay precision), specificity/ selectivity, detection limits, quantitation limits, linearity and range.

2.2.2 Method validation for quantitative analyses

To test for precision and limits of analyses in each method has been summarised in TABLE 2.2 and 2.3 (GC-MS, GC-FID, LC-MS and UPLC-MS only), recovery experiments (TABLE 2.1) were conducted at three levels of standard concentration- 200, 2000 and 5000 ng each.

TABLE 2.1: Recovery of phytoestrogen standards analysed by GC-MS and LC-MS.

Compound	GC-MS % Recovery			LC-MS % Recovery		
	200 ng	2000 ng	5000 ng	200 ng	2000 ng	5000 ng
Formononetin	94.86	98.32	110.06	99.81	102.29	103.08
Biochanin A	75.05	85.69	104.86	91.01	91.21	100.48
Daidzein	112.79	101.47	102.96	101.37	103.24	101.44
Genistein	83.42	97.98	104.43	100.20	99.98	99.62
Coumestrol	108.34	100.03	99.51	99.94	103.63	98.21
Secoisolariciresinol	79.15	98.71	104.89	93.65	95.88	97.54
Matairesinol	98.15	97.15	95.01	95.50	97.05	98.11

TABLE 2.2: Evaluation of precision by intra-assay and inter-assay (SD values)

Compound	Intra-assay				Inter-assay			
	GC-FID	GC-MS	LC-MS	UPLC-MS	GC-FID	GC-MS	LC-MS	UPLC-MS
Formononetin	0.04	0.75	0.09	<0.01	0.08	0.68	0.16	<0.01
Biochanin A	0.02	0.11	0.09	<0.01	0.07	0.21	0.14	<0.01
Daidzein	0.04	0.20	0.06	<0.01	0.07	0.29	0.13	0.01
Genistein	0.04	0.16	0.07	<0.01	0.08	0.20	0.12	<0.01
Coumestrol	0.04	0.12	0.10	0.01	0.08	0.15	0.11	0.01
Secoisolariciresinol	0.10	0.07	0.09	0.01	0.13	0.16	0.07	0.01
Matairesinol	0.05	0.06	0.10	0.01	0.09	0.10	0.05	0.01

TABLE 2.3: Limits of detection (LOD) and quantitation (LOQ) of phytoestrogen standards

Compound	LOD (pg/µL)				LOQ (pg/µL)			
	GC-FID	GC-MS	LC-MS	UPLC-MS	GC-FID	GC-MS	LC-MS	UPLC-MS
Formononetin	200	120	13	16	667	400	43	53
Biochanin A	500	40	25	16	1665	133	83	53
Daidzein	200	40	33	21	667	133	109	65
Genistein	500	100	15	20	1665	333	50	63
Coumestrol	400	40	35	16	1333	133	116	53
Secoisolariciresinol	400	40	200	150	1333	133	667	500
Matairesinol	400	160	120	3	1333	533	400	10

2.2.3 Quantitation of phytoestrogen contents in food samples

Following method development, validation (TABLEs 2.2 & 2.4) and calibration, quantitation of the six African foods were performed using both GC and LC methods. The results of these are shown in TABLE 2.4. Generally, it was observed that the GC-MS results showed to be slightly higher than the LC-MS data in most cases, thus giving slightly different concentration levels in the foods analysed. This problem could be due to either human errors in manual injection or split mode used. In this study, splitless injection mode was used, thus there was a possibility of over - estimation of analyses.

Dehydrated soy powder was also assessed for the phytoestrogen concentration as a reference to test method reliability in comparison to other soy powder analysed by other researchers. All foods analyzed contained at least two out of seven phytoestrogens analysed. Daidzein and genistein were found to be present in at least 4 of the 6 foods. Dehydrated soy powder was found to contain the highest level of phytoestrogen, i.e. ≥ 2.27 mg/g. No lignan was detected in soy powder. Following this, pumpkin leaves was found to contain the highest phytoestrogen level with ~10 µg/g. Biochanin and matairesinol were found to be absent in most of the foods analysed; whilst the main phytoestrogens detected in most of the samples are genistein, daidzein, coumestrol and formononetin (accordingly). On analysing the results in TABLE 2.4 by comparing the two chromatographic techniques used, i.e. GC and LC, there were slightly high variances in some of the analysed phytoestrogen levels. Although, the LC was observed to have (slightly) higher concentrations in some cases, these could be summed to be mainly due to derivatisation differences in each compound. Considering the LC-MS and UPLC-MS data, these two techniques showed to give closely similar values compared to GC-MS. In this particular study, LC could be considered a better chromatographic technique for the analysis of phytoestrogen based on the results obtained.

2.2.4 Identification of novel phytoestrogenic compounds in an African staple food

Interestingly, on analysing pumpkin leaves *'ugu'*, striking peaks besides the standard peaks were observed between 15 – 24 minutes in the GC analysis. This led to the identification of three novel phytoestrogenic compounds shown in FIGURE 2.5.

TABLE 2.4: Phytoestrogen levels in African foods and dehydrated soy as reference

FOODS	Analytical Method	Concentration µg/g							
		Formononetin	Biochanin A	Daidzein	Genistein	Coumestrol	Seco*	Matairesinol	Total
Brown beans	GCMS			0.18 ± 0.02		2.64 ± 0.10	0.22 ± 0.05		3.04
	LCMS			0.64 ± 0.09	0.65 ± 0.08	0.27 ± 0.42	1.32 ± 0.14		2.88
	UPLC-MS	0.11 ± 0.01	0.10 ± 0.01	0.74 ± 0.08	0.53 ± 0.06	0.09 ± 0.01	2.38 ± 0.10		3.95
Cassava 'garri'	GCMS	0.45 ± 0.02	0.07 ± 0.02	0.05 ± 0.01			0.26 ± 0.02		0.83
	LCMS	0.21 ± 0.03		0.14 ± 0.03	0.20 ± 0.03		0.98 ± 0.08	0.14 ± 0.01	1.67
	UPLC-MS				0.18 ± 0.02		1.08 ± 0.05	0.17 ± 0.01	1.43
Plantain	GCMS	0.34 ± 0.07		0.16 ± 0.01	0.30 ± 0.02	0.25 ± 0.01	0.10 ± 0.01		1.15
	LCMS	0.16 ± 0.02	0.24 ± 0.03		0.28 ± 0.05	0.22 ± 0.02			0.90
	UPLC-MS	0.17 ± 0.02	0.21 ± 0.02		0.27 ± 0.04	0.21 ± 0.02			0.86
Pumpkin leaves 'ugu'	GCMS	4.96 ± 0.17		3.31 ± 0.18	2.80 ± 0.08	1.61 ± 0.06			12.68
	LCMS	2.38 ± 0.11		1.81 ± 0.06	2.65 ± 0.10	2.98 ± 0.08			9.82
	UPLC-MS	0.47 ± 0.08	0.82 ± 0.10	1.86 ± 0.05	2.08 ± 0.06	0.43 ± 0.02			5.66
Rice	GCMS					0.73 ± 0.03	0.44 ± 0.06		1.17
	LCMS				0.21 ± 0.01	0.28 ± 0.01		0.10 ± 0.01	0.59
	UPLC-MS	0.06 ± 0.02	0.06 ± 0.01		0.06 ± 0.01	0.06 ± 0.01		0.34 ± 0.06	0.58
Yam 'puna'	GCMS			0.57 ± 0.12	0.66 ± 0.18		0.29 ± 0.08	0.26 ± 0.07	1.78
	LCMS	0.17 ± 0.01			0.25 ± 0.07		3.37 ± 0.21	0.36 ± 0.08	3.98
	UPLC-MS	0.07 ± 0.01			0.37 ± 0.08		4.12 ± 0.06	0.34 ± 0.05	4.90
Soy Powder	GCMS	28.40 ± 0.53		425.0 ± 3.80	1818.0 ± 9.60				2271.4
	LCMS	27.89 ± 0.25		545.7 ± 4.12	1916.2 ± 10.2				2489.8

SECO* secoisolariciresinol. Concentration expressed in µg/g (dry weight) ± standard deviation (S.D).

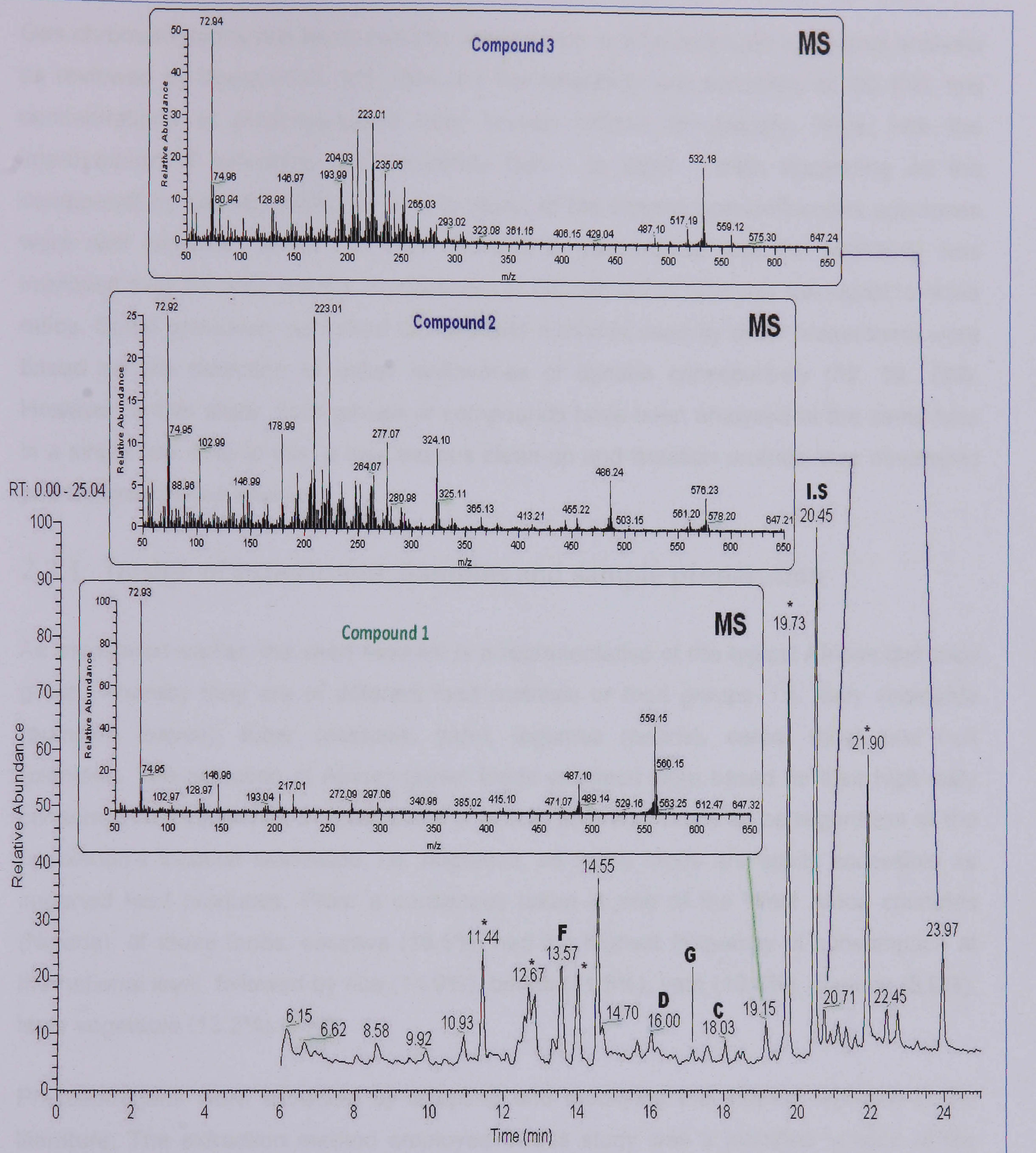


FIGURE 2.5: GC-MS chromatogram and mass spectra of three new phytoestrogenic compounds in pumpkin leaves 'ugu' (inset).

F: Formononetin; D: Daidzein; G: Genistein; C: Coumestrol; I.S: Internal standard; *: Unknown.

2.3 Discussion

Gas chromatography has been used for many years in phytoestrogen aglycones analysis as reviewed by Wang et al (*91*). With the low selectivity and sensitivity of GC-FID, low concentrations of phytoestrogens have proven difficult to quantify. Thus, with the improvement of selectivity and sensitivity (two - to eight – folds depending on the compound) by GC-MS (SIM), co-eluting peaks of the lignans and isoflavones aglycones were well resolved. In the full scan GC-MS, it was noticed that the sensitivity was improved over GC-FID, but the interferences in GC-MS led to relatively low signal to noise ratios. Some previously published GC-MS-SIM methods used by other researchers were based on the detection of either isoflavones or lignans consecutively (*52, 59, 109*). However, in this study, both groups of compounds have been analysed at the same time in a single run. Due to this, a less tedious clean-up and isolation protocol was developed prior to quantitative analysis.

2.3.1 Design of experiment: Samples and sample preparation

As mentioned earlier, the short food-list is a representative of the typical African diet food group, whereby they are of different food matrixes or food groups- i.e. leafy vegetable (pumpkin leaves), tuber (cassava, yam), legumes (beans), cereal (rice) and fruit (plantain). The selection of African-grown foods analysed were based on their high daily consumption frequencies in most parts of Africa. This was found to be regardless of the population's location worldwide, i.e. migration; as these foods are easily accessible as imported food produces. From a consensus taken in one of the West Africa countries (Nigeria), of these foods, cassava (16.5%) had the highest frequency of consumption at the national level, followed by rice (14.9%), beans (11.8%), yam (10.4%), plantain (5.9%), leafy vegetable (13.2%) (*310*).

Phytoestrogens were extracted by adopting and modifying the method reported in the literature. The extraction method employed in this study was a modified version of the *'gold standard'* method developed by Adlercreutz et al (*309*). The developed method by Adlercreutz et al was modified by an increase in the sample amount, solvent volume, choice of extraction solvent, extraction time and temperature. The increased sample amount and volume was supported by Song et al (*312*), who reported a more efficient extraction of isoflavones by adjusting the ratio of extraction solvent to sample amount from > 6 mL/g to 95 mL/g. Achouri et al (*93*) compared extraction solvents and found that 80 %

ethanol gave good extraction yields of isoflavones. Ethanol was reported to be highly effective, lower cost, lower toxicity and environmental compatibility (*94*). In addition, the extraction time was reduced from overnight to 2 h at a high temperature following the report published by Griffith and Collison (*159*) and Rostagno et al (*156*) showing that > 80 % of the isoflavones in samples are extracted within the first five to ten minutes of extraction at a high temperature.

Adlercreutz' general post extraction steps (enzymatic hydrolysis, acid hydrolysis, LLE and column chromatography) were followed to the point of column chromatographic separation on DEAE, which would allow the analysis of isoflavones, coumestans and lignans. The subsequent QAE step, which is generally used for the separation of lignans and isoflavones fractions, was omitted. This was because the lignans and isoflavones were being analysed simultaneously with both GC and LC.

The developed method described in this study was justified to be sufficient for the recovery and assessment of isoflavones, coumestans and lignans based on the calculated recovery values summarised in TABLE 2.1, which averaged 98 % depending on the concentration and analytical technique used to quantitate. Whilst the averaged recovery reported by Mazur et al (*52*), who followed Adlercreutz' method with isotope labelled standards, was 98%.

2.3.2 Chromatographic separation

In this study, simultaneous separations of seven phytoestrogen TMS ether derivatives were developed under optimized GC conditions (FIGURES 2.2 and 2.3). In the past, problems faced with the GC analysis of both lignans and isoflavones simultaneously was due to the column selectivity, similarity in polarity and volatility of their TMS ethers derivatives (*52*). The column selectivity problem was overcome using > 95 % dimethylpolysiloxane: DB-1 and CP-Sil5 CB capillary columns, which are non-polar, 100% dimethylpolysiloxane column with high temperature limits and low bleed capabilities. These columns assisted with the perfect GC separation of these seven compounds compared with other columns made of less than 95% dimethylpolysiloxane. However, there was a difference in the elution order of the seven phytoestrogens between the GC-FID and GC-MS due to the different selectivity by the column manufacturers.

Most of the previous studies (*52, 112*) reported a steep gradient followed by an isocratic temperature program at 280°C, however, the volatility temperatures of phytoestrogen TMS - ether derivatives, ranged from ~250 to 300°C. In addition, lignans' TMS - ether derivatives elute from the GC column at temperatures >290°C, thus this showed why simultaneous analysis was not feasible for other researchers. Also, due to the high temperature limit (300°C) of the DB-1 column with low-column bleed, an efficient gradient temperature program from 150 to 300°C was developed to achieve a good GC separation. GC-FID has never been considered for the analysis of phytoestrogen by any research group. However, in this study, it was noticed that GC-FID was as capable as GC-MS for phytoestrogen analysis based on the chosen column and detector capabilities. It was found to be quite reliable and sensitive based on the effectiveness of the sample preparations and purifications before analysis. FID is well known for its universal response on a wide range of volatile organic compounds, its good linearity and precision. Although, the detection limit of the FID (200 pg) compared with the MS detector (~100 pg) (TABLE 2.1) showed to be approximately two to ten folds higher, it is noteworthy employing the basic GC - FID in the analysis of phytoestrogen considering its price is five folds less than that of GC - MS. Generally, it was found that using the SIM mode in either GC or LC helped with the pico-quantification of these compounds compared to using a total ion count (TIC) mode whereby close peaks could suppress the scanning of peaks of interest especially if these are in nano-quantities.

UPLC was found to be the best analytical technique for phytoestrogen analyses based on pre-treatment simplicity and short analysis time (<5 minutes chromatographic run time), in comparison to both GC analyses, where derivatisation is crucial. In both UPLC-MS and LC-MS analyses, acid modifier trifluoro-acetic-acid was used to improve ionisation of analytes, however as previously reported (*313*), it has no significant effect on retention time.

2.3.3 Detection methods and conditions

Generally, with the SIM mode, each selected ion was chosen based on their corresponding parent ion $[M]^{+\cdot}$. In GC-MS-SIM, these masses were chosen to represent the TMS - ether derivatives of each standard and internal standard. Thus, 72.9, 340.0, 398.0, 411.9, 413.0, 414.0, 471.0, 502.0, 560.0 and 460.0 ions represent TMS group,

biochanin A, formononetin, coumestrol, daidzein, genistein (2-TMS), genistein (3-TMS), secoisolariciresinol, matairesinol and 5α-cholestan-3β-ol respectively.

5α-cholestan-3β-ol was used as the internal standard in GC because of the similarity in structure with phytoestrogens and it was found not to interfere with their detection. Also, due to its volatility, there was no need for prior derivatisation; therefore it can be added prior to GC-MS analysis as an internal standard.

Whilst in (UP-) LC-MS-SIM, the selected ions were 361, 357, 253, 269, 267, 267 / 252, 283 representing secoisolariciresinol, matairesinol, daidzein, genistein, coumestrol, formononetin and biochanin A respectively. Two ions: 267 and 252 were selected for formononetin to differentiate between coumestrol and formononetin, which have the same molecular mass of 268. This peak differentiation was made easy with the ionisation ease of de - methylation of formononetin $(M-H)^-$ from *m/z* 267 to 252.

2.3.4 Method validation

It was observed that intra-assay (accuracy/ repeatability) and inter-assay (precision/ reproducibility) of retention times were good < 0.75, as shown in TABLE 2.2. Both intra- and inter-assay analyses were measured as coefficient of variance (also relative standard deviation - RSD). The inter-assay analysis representing the precision of the developed method with relation to the standards was calculated within a range of 0.02 to 0.10 % for GC-FID, 0.06 to 0.75 % for GC-MS-SIM and 0.06 to 0.10 % for LC-MS-SIM. The intra-assay ranged from 0.07 to 0.13 % on the GC-FID, 0.1 to 0.68 % on the GC-MS-SIM and 0.05 to 0.16 % on the LC-MS-SIM. Thus, in both techniques, the precision were found to be excellent with values < 0.75% RSD, however, GC-FID had better precision in comparison with GC-SIM-MS and LC-MS.

With the calibration of each standard over a range of 100 – 100000 pg/uL, good linearity was achieved based on the calculated correlation coefficient of r^2>0.99 of each standard. The limits of detection (LOD), was calculated as 3 times the signal to noise ratio using a blank, and the limits of quantification (LOQ), which is 3.3 times the LOD (TABLE 2.3). On the overall, LC-MS showed to be more sensitive in this phytoestrogen analysis study.

Recovery of the standards was carried out at concentration levels: 200 ng, 2000ng, and 5,000 ng. The recovery tests of the seven phytoestrogens were found to be between 75 –

113%. On the average, formononetin, daidzein and coumestrol matairesinol had approximately 100% recoveries, whilst biochanin A, genistein, secoisolariciresinol and matairesinol had relatively lower recoveries. On the overall, the recovery analyses using LC-MS showed the highest %recoveries compared to GC-FID and GC-MS.

2.3.5 Quantitation of African food analysis

For the quantitation, GC-MS, LC-MS and UPLC-MS in the selective ion monitoring modes were used. For accuracy, 5α-cholestan-3β-ol and flavone were used as the internal standard in GC and LC analyses respectively. Dehydrated soy powder was used as the reference food sample due to various quantitative works carried out showing that it contains approximately 1-2 mg/g genistein, 0.3 – 0.8 mg/g daidzein; in this study, it was found to contain 0.4 – 0.5 mg/g daidzein and 1.9 – 2.0 mg/g genistein. Corresponding to this, variations in phytoestrogens levels in soybeans has been known to vary more than 4-fold depending on environmental factors such as location, variety, harvesting and processing (*50, 70, 314*). In addition to the dominant daidzein and genistein peaks of soy extract, it was observed in both GCMS and LCMS that soy contained another relatively high peak that was MS-confirmed to be glycitein. Glycitein has been identified by other groups to be present in soy - produces (*93, 113, 156, 303*)

Currently, there is no article in the public domain that has reported the simultaneous quantitation of all of these compounds by GC as well as the phytoestrogen levels in any of these African foods. A similar study by Thompson L.U et al (*113*) analysing some Canadian foods were reported using simultaneous GC-MS analysis of isoflavones and lignans. However in this report, biochanin A was undetectable using this GC-MS analytical method.

The measured food levels of total daidzein, genistein, biochanin A, formononetin, coumestrol, secoisolariciresinol and matairesinol are listed in TABLE 2.4. Generally, from the results obtained from the analysis of the six foods, it can be seen that they all contain at least two out of the seven phytoestrogens assessed. All except pumpkin leaves contained at least one of the two lignans assessed. Pumpkin leaves '*ugu*' was found to contain the highest total amount of the seven phytoestrogen. The foods containing the least total amount of phytoestrogen assessed were cassava '*garri*' and rice.

Although, the comparison of the results with literature was not possible, with this in mind, the levels of phytoestrogens in the African foods were compared to previously published data of similar foods (*59, 109*). Bearing in mind that similar foods from different locations may be significantly different due to variety, cultivation, harvesting, processing, extraction protocol, analytical technique used (*50, 70, 71, 92, 171, 314, 315*). As reported by Mazur et al (*109*), the concentration of cowpea, closely similar to African black-eyed bean was reported to contain 30.3 μg/100g daidzein, 55.7 μg/100g genistein, 195 μg/100g secoisolariciresinol, and trace amount of coumestrol. Comparatively, the level of daidzein, genistein, secoisolariciresinol and coumestrol were calculated in this study to be 69, 59, 188 and 9 μg/100g respectively. Additionally, in this current study, formononetin and biochanin A were found to be 11 and 10 μg/100g respectively. Mazur et al reported in another study (*112*) black – eyed bean to contain 5.5, 7.7, 20.5, 11.4, 7.7, 196.0 μg/100g formononetin, biochanin A, daidzein, genistein, coumestrol and secoisolariciresinol respectively.

Another food considered for comparison is plantain. Liggins et al (*316*) reported no detection of phytoestrogen in this food, however, in this thesis report, plantain was found to contain less than 30 μg/100g each of formononetin, biochanin A, genistein, coumestrol and secoisolariciresinol. A closely related food to plantain: banana, was reported by Thompson (*113*) to contain < 0.3 μg/100g (wet weight) each of formononetin, daidzein, genistein, matairesinol and 0.6 μg/100g (wet weight) secoisolariciresinol. The wet - weight to dry - weight conversion in plantain, from this study was calculated to be 2.5. Thus, after conversion, the level of phytoestrogens reported by Thompson in banana was much lower than that of plantain reported in this study.

Although, the incidence of some of the chronic diseases leading to cancer is reduced in the African populations, there has been no correlation of this to phytoestrogens in their diet and their consumption. Thus, the phytoestrogens content in six commonly consumed African foods were analysed. According to Jason Liggins et al. (*316*), foods containing lower amounts of the isoflavones may have more subtle effects upon long-term health, especially if different sources of compounds are eaten together, generating a cumulative dose. Even though, these analysed foods have been found to contain little amounts of phytoestrogens, however, on the basis of their daily consumption, their consumption level is relatively high.

For instance, a typical local lunch diet in Africa (Nigeria) is made of 100 g (wet weight) each of pumpkin leaves '*ugu*' and cassava '*garri*' consumption respectively per day, the total of the phytoestrogen consumption would be equivalent to ≥ 82 µg for pumpkin leaves '*ugu*': ~20 µg formononetin, ~15 µg daidzein, ~22 µg genistein and ~25 µg coumestrol; and for cassava 'garri'- total of ≥ 356 µg: 210 µg formononetin, 14 µg daidzein, 20 µg genistein, 98 µg secoisolariciresinol and 14 µg matairesinol. Therefore, with the consumption of one of the three meals of the day, ≥ 438 µg phytoestrogens are consumed. Although cassava has a lower concentration of the seven phytoestrogen (in dry weight) compared to pumpkin leaves, based on the calculation of the amount (wet weight) eaten per day- cassava was found to have a higher daily consumption.

Following this case study, a typical Nigerian dinner diet is made of 100 g rice, 50 g beans, 200 g plantain and tomato sauce. The rice would contribute to at least 59 µg phytoestrogen: 21 µg genistein, 28 µg coumestrol and 10 µg matairesinol. The beans would contribute to at least 264 µg phytoestrogens: ~32 µg each of daidzein and genistein, 134 µg coumestrol and 66 µg secoisolariciresinol. The plantain would contribute to at least 72 µg phytoestrogens: 12.8 µg formononetin, 19.2 µg biochanin A, 22.4 µg genistein and 17.6 µg coumestrol. Therefore, the phytoestrogen contribution in this meal is at least 395 µg.

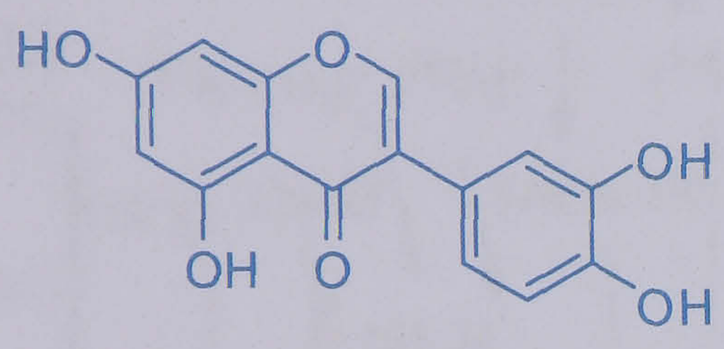
It is noteworthy that these calculated values have been estimated based on the assumption that there are no losses in the food processing, e.g. washing, steaming and frying. In the report published by Kuhnle et al (*114*), cooking showed to have no consistent effect on the phytoestrogen content, as in most of the foods analysed, some phytoestrogens levels slightly increased on cooking whilst others slightly decreased. Also, reported by Grace et al (*317*), isoflavones are present as emulsifiers and preservatives in bread and bakery products, thus, there is the possibility of isoflavone consumption found in bread, consumed as part of the breakfast in Africa.

With these estimates' calculations, the average consumption of phytoestrogen in Nigeria is ≥ 700 µg/day on consuming at least two meals a day. Whilst in comparison, the average isoflavone consumption by women in other countries: UK (Norfolk) was 437 µg/day (*317*), Americans was 630 µg/day (*62*), Singapore was 4.7 mg/day (*318*), Chinese was 333.4 mg/day (*319*) and Japanese was 46.5 mg/day (*320*). Thus, the consumption of isoflavone is higher in Africans, specifically Nigerians compared to Western populations. The majority of the isoflavones have been mainly derived from starchy staple food – pounded yam

(estimated value not included), cassava '*garri*' as detailed. Although, their isoflavone content detected are low, the fact that these foods are consumed in dried powdery forms at large quantities, thus, the overall phytoestrogen mount up to be significant.

2.3.6 Identification of new phytoestrogenic compounds in pumpkin leaves '*ugu*'

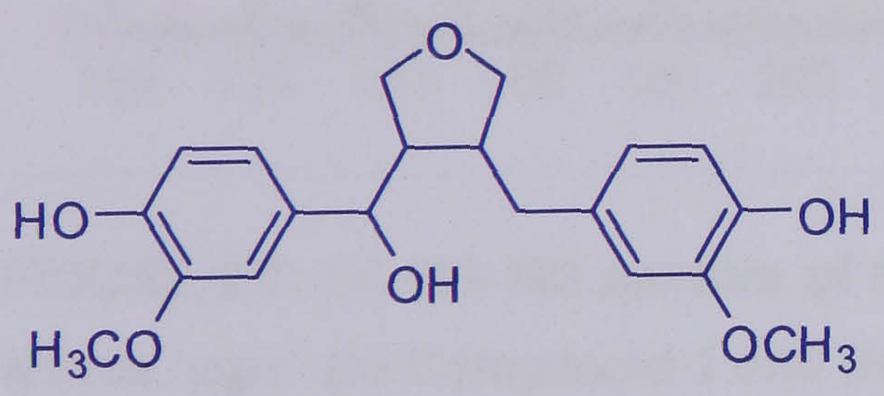
On GC analysis, three new phytoestrogenic compounds were observed to be present in pumpkin leaves 'ugu'. Based on the MS fragmentation patterns in mass spectra, it was deduced that these compounds were one flavonol and two lignans. These compounds (FIGURE 2.6) were interpreted to be 3, 5, 7- trihydroxy-2-(4-hydroxyphenyl)-4*H*-chromen-4-one (compound 1), 4-(hydroxyl(4-(4-hydroxy-3-methoxybenzyl) tetrahydrofuran-3-yl)-2-methoxyphenol (compound 2) and 3-(4-hydroxy-3-methoxybenzyl)-4-((4-hydroxy-3-methoxyphenyl)(methoxy)methyl)dihydrofuran-2(3*H*)-one (compound 3).



(1)

3-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4*H*-chromen-4-one

Chemical Formula: $C_{15}H_{10}O_6$
Exact Mass: 286.05



(2)

4-(hydroxy(4-(4-hydroxy-3-methoxybenzyl)tetrahydrofuran-3-yl)methyl)-2-methoxyphenol

Chemical Formula: $C_{20}H_{24}O_6$
Exact Mass: 360.16

(3)

3-(4-hydroxy-3-methoxybenzyl)-4-((4-hydroxy-3-methoxyphenyl)(methoxy)methyl)dihydrofuran-2(3*H*)-one

Chemical Formula: $C_{21}H_{24}O_7$
Exact Mass: 388.15

FIGURE 2.6: Structures of novel phytoestrogenic compounds in pumpkin leaves 'ugu'

With pumpkin leaves '*ugu*' GC-MS full scan mode analysis (shown in FIGURE 2.4), analysing the mass spectrum obtained at retention time 19.08 min, the observed base peak ion was 559 *m/z* (parent ion 574 *m/z*) as a result of a loss of a methyl radical (15

amu), thus implying the better stability of $[M\text{-}CH_3]^{+\bullet}$ ion compared to the $M^{+\bullet}$ ion. Loss of trimethylsilane $[M\text{-}TMS]^{+\bullet}$ ion resulted to give the 486 amu, followed by consequent loss of three TMS ions showed to be tetrahydroxyl- substituted. The fragmentation pattern was found to be quite similar to genistein's, i.e. main difference was with the additional TMS substitution, thus led to the conclusion of a tetrahydroxyflavonol. With this deduction, LC-MS-MS was used to confirm the proposed compound (FIGURE 2.7).

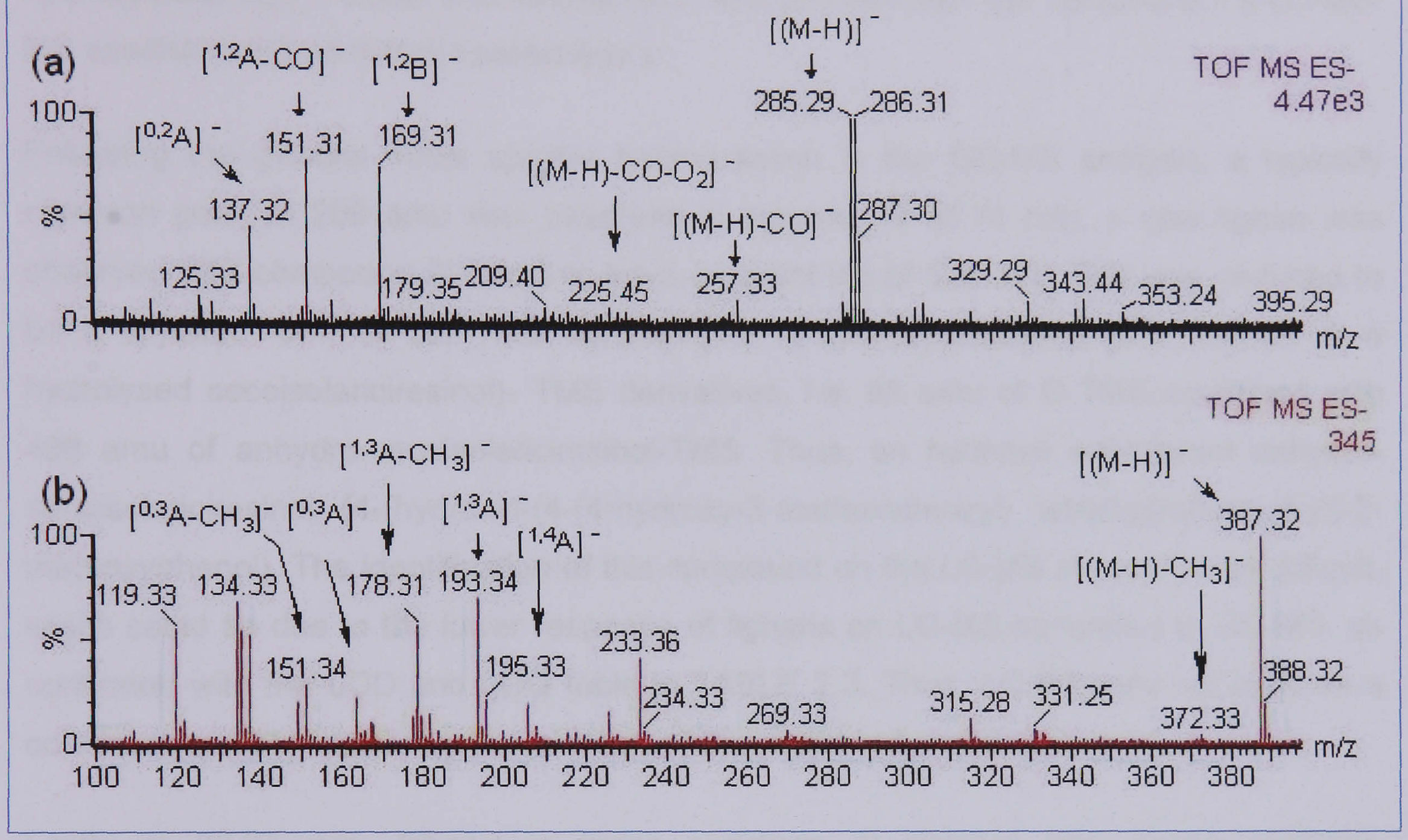


FIGURE 2.7: LC-MS-MS spectra of two new phytoestrogenic compounds found in pumpkin leaves 'ugu'. (a) Compound 1 and (b) compound 3.

From the LC-MS-MS spectra, the proposed detailed pathway of fragmentation was illustrated in FIGURE 2.8 & 2.9 showing the schematic fragmentation pathway of each of the $[M\text{-}H]^-$ ions.

With compound 1: 3, 5, 7- trihydroxy-2-(4-hydroxyphenyl)-4*H*-chromen-4-one, (tetra-hydroxyflavonol), it was observed that there was a loss of carbonyl group from the parent ion indicating $[M\text{-}H\text{-}CO]^-$ ion at 257. The *m/z* 257 underwent a consecutive loss of two oxygen atoms with the re-arrangement of electrons and ions, to generate an ion at *m/z* 225 (at low relative abundance). Interestingly, there were different occurrences of retro-Diels Alder fragmentation reaction (retrocyclisation) with the main reaction being the formation of *m/z* 137 (ring-^{1,3}A) and 151 (ring-^{0,3}A) product ion, especially due to their

relatively high abundance as seen in FIGURE 2.7. On the other hand, following the cleavage ring at ring-^{1,2}A with *m/z* 179 which with the loss of CO gave *m/z* 151 ion. Similar ring-B bond cleavage at different carbon positions occurred, which resulted to the production of *m/z* 169 ion in ring-^{0,4}B- 2[O]. Also, *m/z* 169 ion was achieved following ring-^{1,2}B cleavage in an rDA reaction.

Additionally, the proposed compound was compared with two different suspected tetra-hydroxyflavonoids – lutelin and kaempherol, and of these two, the compound 1's LC-MS-MS spectra matched that of kaempherol's.

Following the general mass spectra fragmentation in the GC-MS analysis, a typically common peak at 209 amu was observed in lignans. At 20.71 min, a new lignan was observed (MS-compound 2) found to have a parent ion of 576 *m/z*. This was deduced to be a silylated form of the TMS derivative of anhydrosecoisolariciresinol (which is a hydrolysed secoisolariciresinol)- TMS derivatives, i.e. 88 amu of O-TMS combined with 488 amu of anhydro-secoisolariciresinol-TMS. Thus, an hydroxyl substituted anhydro-secoisolariciresinol (4-(hydroxyl-(4-(4-hydroxy-3-methoxybenzyl) tetrahydrofuran-3-yl)-2-methoxyphenol). The identification of this compound on the LC-MS showed to be difficult, which could be due to the lower response of lignans on LC-MS compared to GC-MS, as confirmed with the LOD and LOQ table in TABLE 2.3. Thus, LC-MS was not used as a confirmatory test for structural elucidation of this proposed compound.

Lastly at 23.97 min, (shown in mass spectrum 3) another new lignan methoxy-matairesinol was observed. Based on the ions *m/z* 209, 223 ions representing the silylated methoxy-phenyl and its dimethoxy-phenyl ion respectively. The ions at *m/z* 501 and parent ion at *m/z* 532 represent the matairesinol and its methoxyl - derivative: 3-(4-hydroxy-3-methoxybenzyl)-4-((4-hydroxy-3-methoxyphenyl)(methoxy)methyl) dihydrofuran -(3*H*)-one respectively. This compound, compound 3, which is a methoxyl derivative of matairesinol was further proven on LCMS with a $[M-H]^-$ ion peak at *m/z* 388. On immediate assessment of the MS/MS spectra, following Rongxia Liu et al (*321*), the loss of CH_3 are peculiar to methoxylated flavonoids in the negative ESI-MS/MS mode, thus from the result obtained (compound 3) with multiple CH_3 loss, it can be inferred that the compound is a multi-methoxylated flavonoid. This *m/z* 388 ion was observed to undergo a retro-synthetic cleavage- retro-Diels-Alder fragmentation; the proposed pathway for this reaction and other fragmentation of this ion is shown in FIGURES 2.8 - Scheme II.

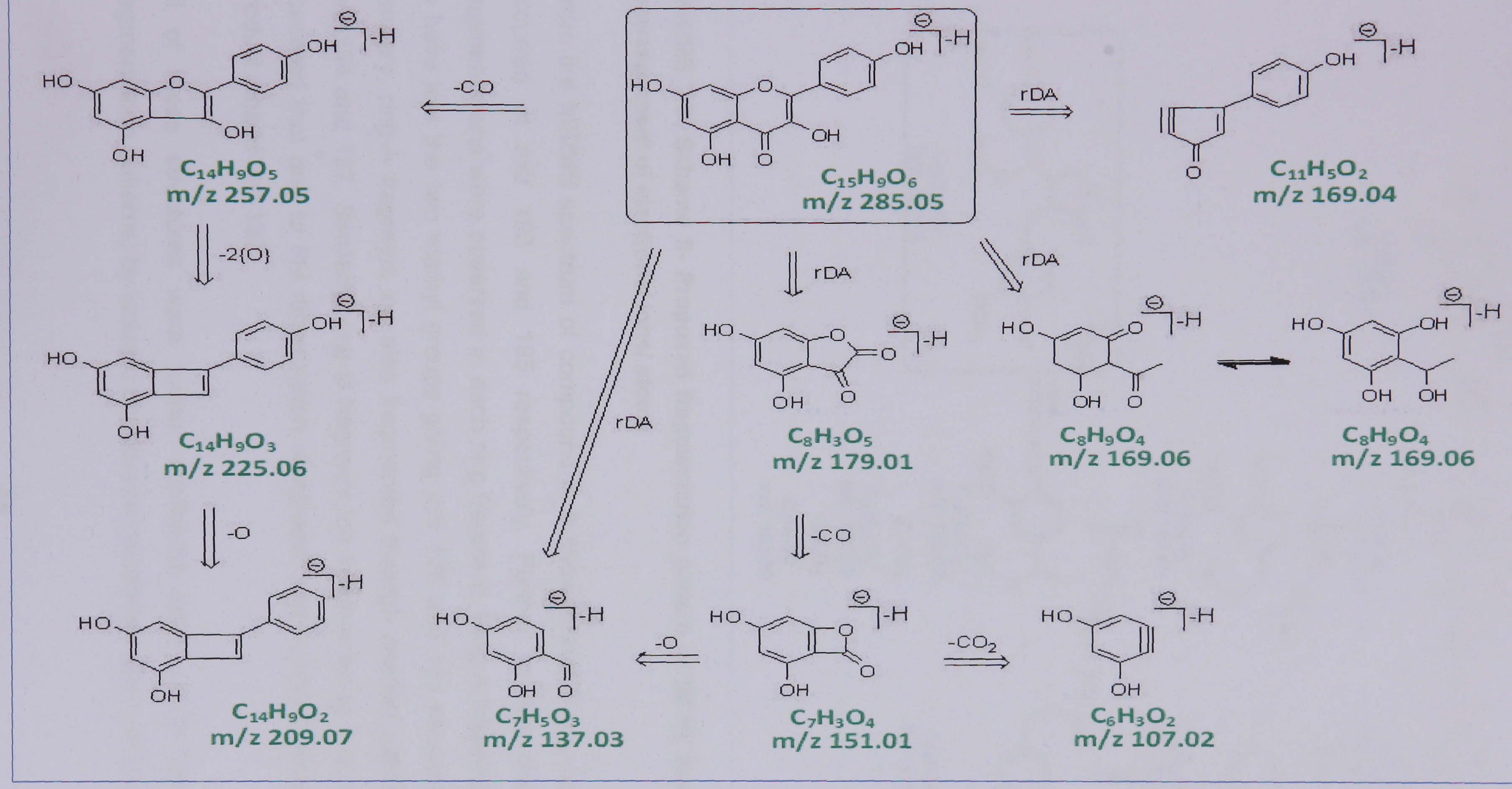


FIGURE 2.8: Proposed Scheme I: Proposed fragmentation pattern of $[M-H]^-$ ion of compound 1. rDA: retro-Diels Alder reaction (retrocyclisation)

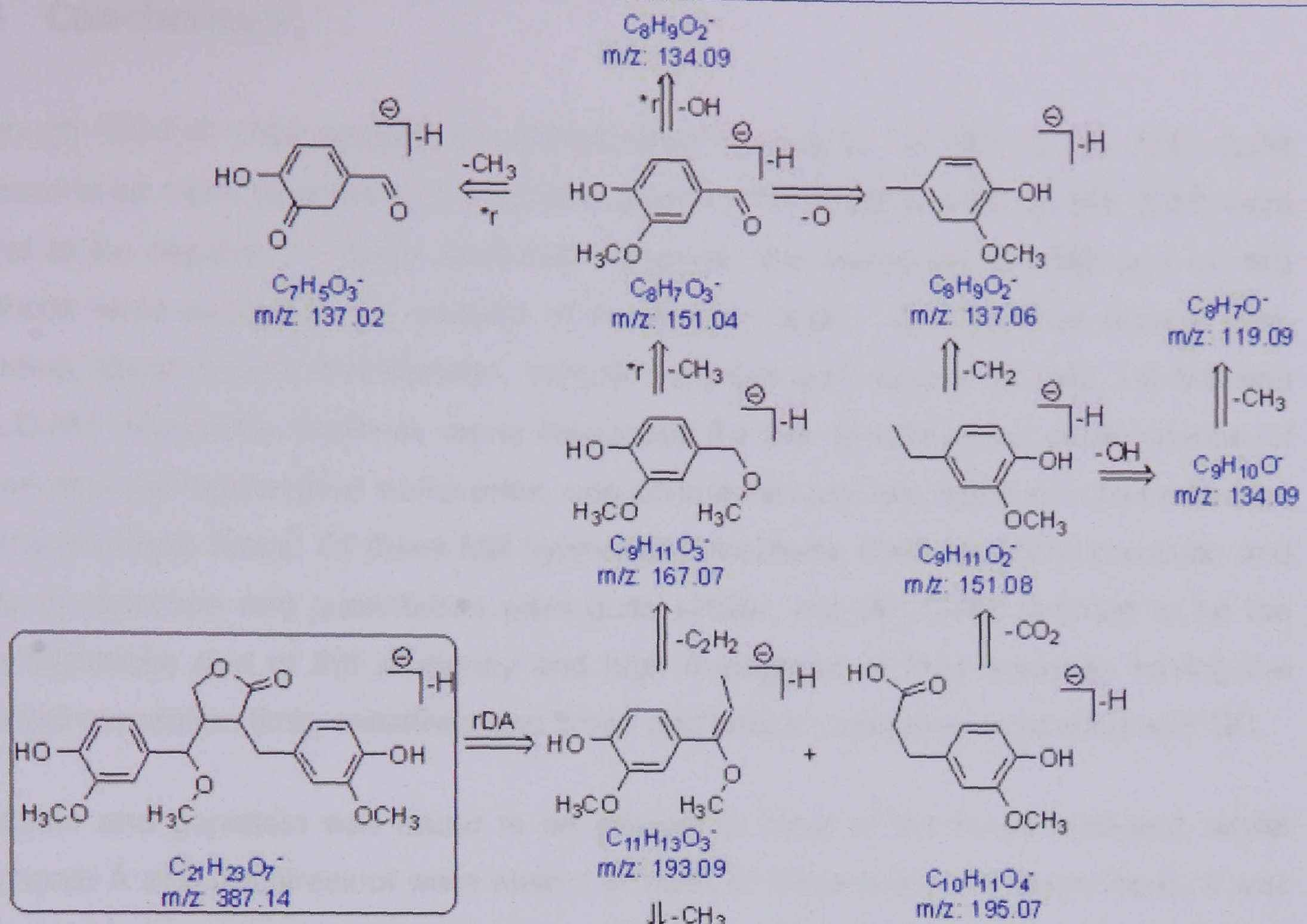


FIGURE 2.9: Scheme II- Proposed fragmentation pattern of $[M-H]^-$ ion of compound 3. *r rearrangement of electrons/ ions/ atoms

From the MS/MS spectrum of compound 3, it showed that A-ring and B-ring fragments occurred at *m/z* 193 and 195 respectively. Further to this cleavage, subsequent fragmentations were observed in each ring fragment. Ring-A fragment ion was observed to have lost the two methyl groups giving *m/z* 178 and 163 sequentially. Whilst on the contrary, ring-A fragment ion also fragmented through another pathway to product ions *m/z* 134 and 137. Similarly, ring-B fragment ion fragmented to *m/z* 134 and 119. It was observed that due to the dimerisation of lignans, ring-A and B showed to have similar product ions of *m/z* 137.

All of these structures were further confirmed (after visual interpretations of the fragmentation patterns) by isotopic abundance calculations/estimations.

2.4 Conclusions

Although GC-FID was capable of phytoestrogen analysis, GC-MS in the SIM mode showed to be more sensitive and accurate. Also, LC-MS (SIM) and UPLC-MS (SIM) were found to be capable for these analyses. However, the developed GC-MS and LC-MS methods were applied in the analysis of six African foods - rice, pumpkin leaves, yam, cassava, brown beans and plantain. Simple, sensitive and reliable GC-MS, LC-MS and UPLC-MS separation methods were developed for the simultaneous determination of seven phytoestrogens (four isoflavones, one coumestan and two lignans) in this selection of African staple foods. Of these MS hyphenated methods, their accuracy, precision and limits of detection and quantitation were quite similar, but UPLC-MS seemed to be the best technique due to the efficiency and high throughput of food analyses having the shortest separation time, sensitivity and time - and labour - intensive comparing with GC.

Daidzein and genistein was found to be present in most of the foods analysed, whilst biochanin A and matairesinol were absent in most of these foods. Of these foods, it was found that pumpkin leaves 'ugu' showed to have the highest total concentration of measured phytoestrogens in this study. Following this, African beans were found to contain 3 – 5 µg/g cumulative phytoestrogen, whilst yam contained 2 – 4 µg/g cumulative phytoestrogen.

In addition, three estrogenic compounds (not included in the test compounds) were found to be present at fairly high amount in pumpkin leaves '*ugu*' excluding the seven assessed phytoestrogens. These include: two lignan compounds and one isoflavone-like compound (specifically isoflavonol).

Based on the results obtained from this study, it is encouraging to analyse the phytoestrogen content in various African foods, as presently there are no data available in the public domain on their phytoestrogen content. Following this, a well - validated food-frequency questionnaire would be beneficial for the assessment of phytoestrogen consumption within the African populations in UK, Africa and Caribbean.

2.5 Experimental

2.5.1 Chemicals, solvents & materials

2.5.1.1 Chemicals

Seven phytoestrogen aglycone standards namely formononetin, biochanin A, genistein, daidzein, coumestrol, secoisolariresinol, matairesinol were procured from Sigma Fluka, UK (Gillingham, Dorset, UK). β-Glucosidase from almonds, *N,O*-bis- (trimethylsilyl)-trifluoroacetic amide (BSTFA) with 1% trimethylchlorosilane and diethylaminoethyl (DEAE) resin were procured from Sigma Fluka, UK (Gillingham, Dorset, UK). Sodium hydroxide pellets, acetic acid glacial 99+% and hydrochloric acid 1.16 S.G. 31 – 33 % were purchased from Fisher Scientific Limited, UK (Loughborough, Leicestershire, UK).

2.5.1.2 Solvents

HPLC-grade solvents: methanol, acetonitrile, water and diethyl-ether were purchased from Fisher Scientific Limited, UK (Loughborough, Leicestershire, UK). Distilled water was purified with Millipore Simplicity 185 (Millipore Corp., USA) to 18.2 $m\Omega.cm^{-1}$ polishing units.

2.5.1.3 Materials: food samples

Six foods (grown and imported from Africa) were purchased fresh from Peckham Street African Market, South East London, UK. These foods were pumpkin leaves '*ugu*' (*Telfaria Occidentals*) purchased as leafy vegetables; cassava (*Manihot esculenta*) '*garri*' as grains, yam (*Dioscorea rotundata*) as a root-tuber, rice (*Oryza sativa*) as grains, plantain (*Musa paradisiacal)* as fruit and brown beans/cowpeas (*Vigna unguiculata*) as grains. Leafy vegetables were chopped and subjected to freeze-drying, whilst the fruit and root tuber were peeled, chopped and freeze-dried. Food samples were stored at 4 – 8°C prior to analysis.

2.5.2 Instrumentation

2.5.2.1 General

Eppendorf Centrifuge 5804R (Germany) was used for sample preparation in liquid- liquid extraction (LLE). *HERA* cell150 incubator (U.K.) was used to incubate the food extracts for enzymatic hydrolysis overnight at 37ºC. Buchii Rotary Evaporator R-210 (Switzerland) equipped with a vacuum pump V-700, heating bath B-491 was used to evaporate extraction solvents and concentrate the soxhlet lipid-free extracts. Millipore Simplicity 185 (U.K.) was used for the purification of water for sample preparation and the HPLC/ LCMS analyses.

Buchii B-811 Extractor system (Buchii, Switzerland) was used for the dried food sample extraction by soxhlet extraction. The set program developed for this study involved 3 steps: extracting, rinsing and drying accordingly. In extraction (step 1), the upper and lower heating were set at 4 and 16 respectively for 2 h, whilst the rinsing (step 2) and drying (step 3) were set at 2 for 5 min and 4 for 30 min respectively.

2.5.2.2 GC-FID

The GC separation was performed on a Varian 3900 connected to flame ionisation detector (FID) was employed with WCOT Fused Silica Stationary phase capillary column: CP-Sil5 CB, 10 m x 0.53 mm, 2 µm film thickness (Chrompack France, Les Ullis, France). The carrier inert gas (nitrogen) and the support gases - air and hydrogen were maintained at 1, 30 and 300 mL/min respectively. The detector and injection temperatures were kept at 250°C and 200°C respectively. The standards and samples were analysed using the programmed temperature of oven as follows: first at 150°C held for 2 min and then increased to 300°C at a ramp of 10 °C/min, and finally held at 300°C for 3 min.

2.5.2.3 GC-MS

Trace DSQ GC 2000 series gas chromatograph (ThermoQuest, San Jose, CA) interfaced with a TriPlus Autosampler (ThermoElectron, USA) was coupled to a quadrupole mass spectrometer operated in electron impact (EI) ionisation mode at 70 eV electron energy. The capillary column used was DB-1 column with dimensions 30 m x 0.25 mm, 0.25 µm film thickness (Jones Chromatography, Mid Glamorgan, UK). The temperatures of the

source and transfer line were set at 200°C and 250°C respectively. The column oven temperature method used started at 200°C, which was held for 2 min, then increased at 5 °C/min to 300°C and held for 3 min at 300°C. Two scans were run through the experiments, i.e. full scan and selected ion-monitoring scan. The full scan was run from *m/z* 50 – 700. At the same time as the SIM scan mode was operated at 72.9, 132 and 340 (formononetin); 135, 356 and 413 (biochanin A); 73, 184 and 398 (daidzein); 73, 414 and 471 (genistein); 191 and 412 (coumestrol); 209, 560 and 650 (secoisolariciresinol); 179, 209 and 502 (matairesinol) and lastly, 445 and 460 (internal standard: 5α-cholestan-3β-ol). Data acquisition and processing were carried out on Xcalibur 1.1 software (Thermo Finnigan). Four 2 µL injections of these derivatised samples were performed for statistically analysis (n = 4).

2.5.2.4 LC-MS

LC operating conditions: The automated LC system (*LC packing, UK*) comprised of a Famos autosampler connected to Ultimate™ plus HPLC. Chromatographic separation was achieved on Phenomenex column- Luna C8(2)- 3 µm, 150 x 2.0 mm on a gradient elution program of mobile phase composed of 35% acetonitrile in water [0.1% formic acid] at a flow rate of 0.15 mL/min. The HPLC experiments were run isocratically over 20 minutes.

MS operating conditions: The structural identification of compounds was achieved by Waters Micromass Quadrupole Time of Flight Micro™ analyzer operated in the negative ion mode electrospray ionisation under the following conditions: Capillary voltage of 3500 V; sample cone voltage and extraction cone maintained at 30 V and 3 V respectively; source temperature was set at 150ºC; both ion and collision energies at 3 V. For LC-MS-MS, i.e. fragmentation, the sample cone voltage and collision energy were increased to 60 V and 50 V respectively.

The total ion count scan was run from *m/z* 100 – 600. Following this: post data analyses, the SIM scan mode was operated at m/z: 267.27 (formononetin); 283.27 (biochanin A); 253.24 (daidzein); 269.24 (genistein); 267.23 (coumestrol); 361.42 (secoisolariciresinol); 357.39 (matairesinol) and lastly 237.24 (internal standard: 3 hydroxyflavone). Data acquisition was done using MassLynx 4.0 Global Mass-Informatics software. Four 5 µL injections of each sample were performed for statistically analysis (n = 4).

2.5.3 Mass spectrometry analysis

2.5.3.1 Standard calibration

Solution of standard compounds: A 100 ng/µL stock methanolic solution of the standards was made. For calibration, after derivatisation, 1:1 serially dilutions were made to give concentration levels/points 40, 20, 10, 5, 2.5, 1.25, 0.625, 0.313, 0.156, 0.078, 0.039 ng /µL. The calibration curve ranges were 2.5 to 40 ng/µL and 0.039 to 40 ng/µL for GC-FID and GC-MS-SIM respectively. Whilst for LC-MS and UPLC-MS, the calibration range was 0.010 to 20 ng/µL.

For the calibration plots, the ratios of the either the analytes' intergrated peak areas (GC-MS and UPLC-MS) or the ion count intensity 'peak height' (LC-MS) to the internal standard's were plotted against their concentrations.

2.5.3.2 Sample analysis

Following the calibrations, the quantitations of analytes in the samples were acheived by calculating the ratios of the peak areas (GC-MS and UPLC-MS) or ion count intensity (LC-MS) to that of the internal standard's and comparing with the non-weighted calibration curves.

2.5.4 Sample preparation

2.5.4.1 Extraction of phytoestrogens: Liquid (soxhlet) extraction

Approximately 1 g of powder dried food sample was accurately weighed and extracted in 100 mL of 80% ethanol using a Buchii extractor system with the program described above. The resulting extract was defatted with 50 mL of hexane twice, thereafter; the resulting extract was evaporated using a rotary evaporator. The dried extract was re-dissolved in 5 mL of 100 mM sodium acetate buffer (pH 5).

2.5.4.2 Enzymatic hydrolysis

1 mL assay of each extract solution were hydrolysed with 10 U of β-glucosidase from almonds incubated overnight at 37°C (minimum of 15 h). The hydrolyzed aglycones were extracted twice using cold 5 mL diethyl ether by partitioning or LLE.

2.5.4.3 Acid hydrolysis

Thereafter, the aqueous phase concentration was adjusted to 2 N hydrochloric acid, vortexed and incubated in a tightly sealed vial at 100 °C for 2.5 h. These resulting acid hydrolysates were neutralised with 1.3 - 1.4 mL of 2 M sodium hydroxide to pH between 3 and 5. The resulting aglycones were extracted by partitioning with diethyl ether as carried out in the enzymatic hydrolysis. The enzyme and acid hydrolysate- aglycone ether phases collected were combined, completely dried under nitrogen and re - dissolved in 0.5 mL methanol.

2.5.4.4 Purification/ isolation of phytoestrogen aglycones

Ion exchange column chromatography was carried out using DEAE-OH on the resulting methanolic assay. The methanolic solutions were applied onto Pasteur pipettes columns packed with 0.5 X 1.5 cm DEAE-OH resins. Neutral steroids were eluted with ten times the bed volume (~5 mL) prior the elution of phytoestrogen aglycones with 5 mL of 100 mM acetic acid in methanol. This fraction was collected, dried down and redissolved in 0.5 mL methanol prior analysis.

2.5.4.5 Derivatisation of samples prior GC analyses

Aliquots of 100 µL from each isolated extract were taken out into different vials, dried down completely and 100 µL BSTFA was added for derivatisation. For the completion of derivatisation, these solutions were either incubated overnight at ~30 °C or incubated for 3 h at 50 °C. Thereafter, the sample in BSTFA solvent was analysed on GC-MS and GC-FID.

Chapter III

Absorption and metabolism of soy - phytoestrogens

3.1 Introduction

Phytochemicals, which include polyphenolic compounds, have attracted increasing interest for their hypothesised and confirmatory claims on health benefits reviewed by different research groups (*185, 322-324*). Following ingestion, (small-) intestinal bacteria metabolises dietary components to produce metabolites, which are more bio- active than their precursor thus influencing host's health (*324*). Generally, it has been found that phytochemicals, especially the flavonoids have been regarded as one of the important natural product compounds found in diet-based supplies. This in addition to their potential health benefits made them one of the well studied and intriguing compounds. With phytoestrogens consumption, the glycones conversions to aglycones were found to be highly influenced by certain factors: age, gender, stress and diet. (*150*). In addition, following soy - phytoestrogen metabolism, suggestions have been made based on observation and interventional studies that the ability to produce metabolites - equol and o-DMA may be associated with reduced risk of certain cancers including breast and prostate cancers (*145*).

Based on the extensive studies of phytoestrogens and their biomarkers in Asians (*12, 13, 122*), Americans (*62*), Europeans (*187, 325*) and Australians (*130*), no publicly - known study has been carried out on either indigenous Africans or British - Africans in regards to phytoestrogen studies. Further to this, two main questions to be considered in this study are: firstly, what are the typical bioavailability levels of phytoestrogens and their metabolites in Africans (baseline urine analysis). Secondly, although, it has been generalised that less than 40% of the human population depending on the ethnical race, can metabolise daidzein to equol, it is unknown whether African can metabolise soy phytoestrogens- daidzin and genistin, and to what extent. This second question was addressed by following a soy 'challenge'/ soy intervention. Soy-milk intervention was designed to provide enough circulatory phytoestrogens in the subjects so as to detect phytoestrogens and their metabolites in the bio-fluid, i.e. monitor the soy metabolism in bio-fluid (urine) analysis. Numerous studies have shown that bio-fluids are good '*in-vivo*'

representative media for the investigation of potential biomarkers of bio - active compounds of interest.

Although dietary habits influence gut metabolism, on the other hand with regards to daidzein and/ or genistein metabolism, according to Hedlund's research group (*325*), the dietary habit responsible for production of the highly bio-active equol is unknown. This was further confirmed with conflicting results following low-fat (*150*), high fibre (wheat-bran) diet (*155*) and long soy consumption period (1-month intervention) (*65, 291*). Thus, one of the reasons for the differences in equol production could be due to the differences in the strains of gut bacteria, thus inter-subject variability. On this basis, asides habitual diet, other extrinsic factors could be responsible for equol production- age, gender, location. Thus, in this study, these factors were considered for the differences in soy metabolism.

With the aim of detecting and monitoring phytoestrogens and their metabolites in urine, a fast, efficient, specific and selective extraction protocol using solid phase extraction (SPE) was developed. This developed protocol could also be applicable to blood serum. Following extraction, liquid chromatography-mass spectrometry (LC-MS) was employed to analyse pico - to nano - levels of metabolites in the urine. In this study, four phytoestrogens and four phytoestrogen metabolites were analysed as shown in FIGURE 3.1.

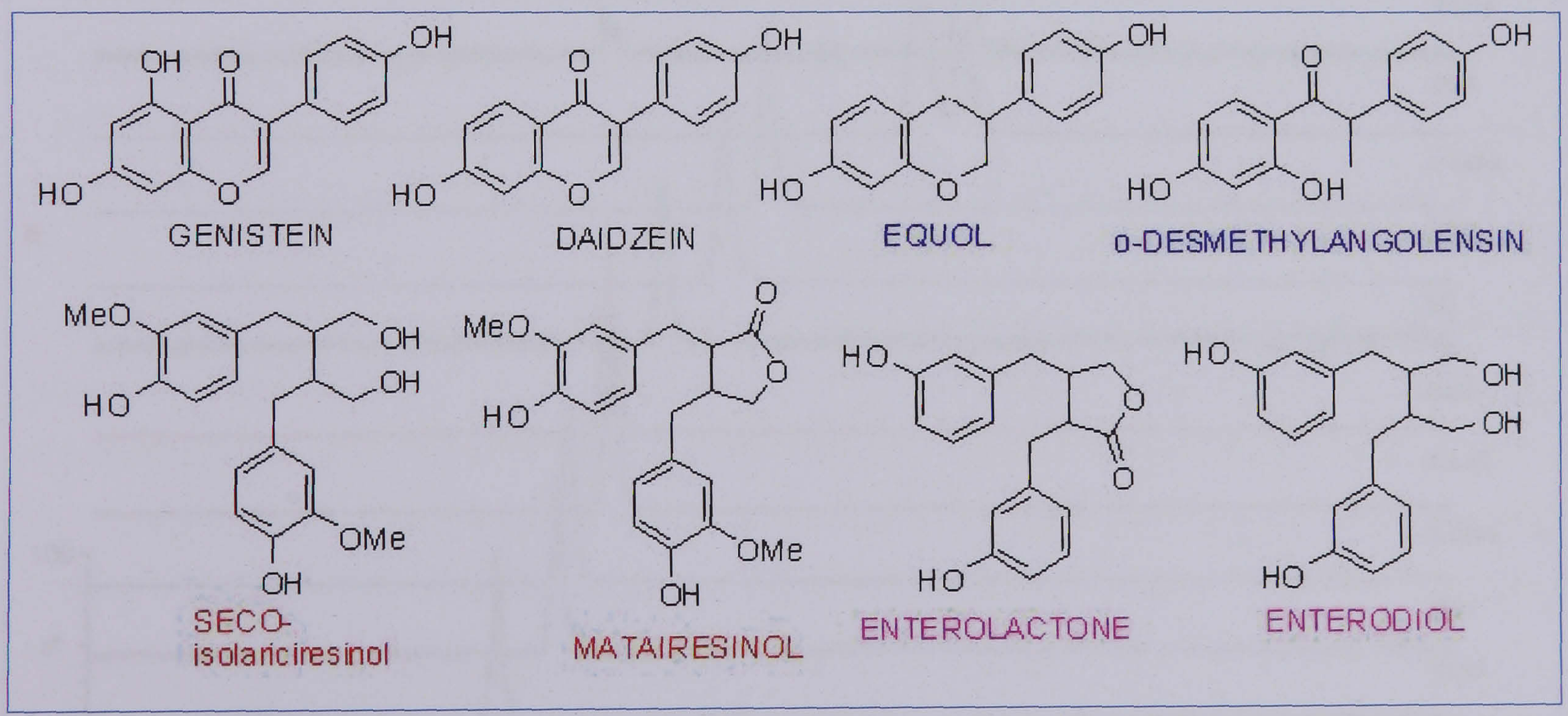


FIGURE 3.1: Phytoestrogens and metabolites analysed in this study

3.2 Results

3.2.1 Chromatographic analyses of phytoestrogens and their metabolites

In this study, two relatively fast, selective and sensitive chromatographic methods using a mobile phase of 35% methanol in water (Method A - UK) and 35% acetonitrile in water containing 0.1 % formic acid (Method B - Nigeria) were developed following several attempts. Although, method A was found to be effective in the analysis of the UK cohort samples, it was observed that there were mounting residual interferences on the LC-MS instrument due to methanol solvent contamination. Thus, a new method using a different organic solvent- acetonitrile (Method B) had to be developed. The optimum separations were achieved in less than 20 minutes at 0.15 mL/min flow-rate at 25°C. With the advantage of selective-ion monitoring mode whereby peaks were separated based on both *m/z* ratio of parent ion (MS) and elution time off column (LC), this showed to produce distinct resolutions of all analytes' peaks (FIGURE 3.2).

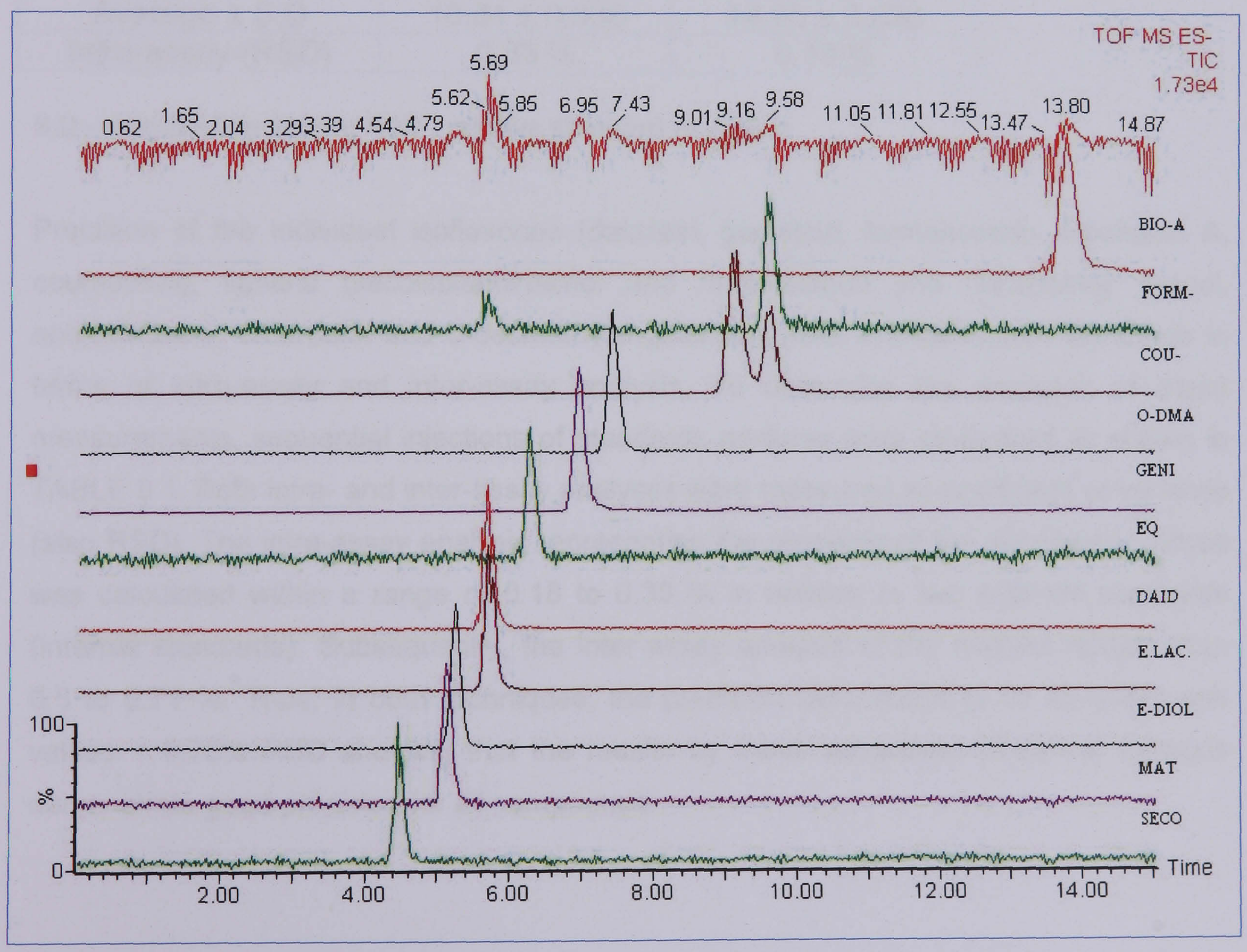


FIGURE 3.2: LC-MS-SIM chromatogram of mixture of 11 standards at 100 pg/µL

Each of the two methods was used for the two cohorts' analyses, i.e. method A was used for the UK cohort whilst method B was used for the NIGERIAN cohort.

3.2.2 Method validation for quantitative analyses

Biochanin A (Method A) or formononetin (Method B) was used as internal standard to evaluate the precision of the method for validation. The retention times of each standard compound was calculated and shown to be within 0.5% variances (TABLE 3.1).

TABLE 3.1: Precision calculation using internal standards.

Injection	Retention time (min)	
	Biochanin A	Formononetin
1	10.81	14.68
2	10.85	14.61
3	10.81	14.66
4	10.75	14.64
5	10.82	14.65
Average ± S.D	10.81 ± 0.036	14.65 ± 0.026
Intra-assay (RSD)	0.33 %	0.18 %

S.D - standard deviation; RSD - relative standard deviation

Precision of the individual isoflavones (daidzein, genistein, formononetin, biochanin A, coumestrol), lignans (secoisolariciresinol and matairesinol) and metabolites (equol, enterolactone, enterodiol and o-desmethylangolensin) were evaluated with standards in terms of intra-assay and inter-assay analysis. To determine the precision of these measurements, sequential injections of standards mixtures were performed as shown in TABLE 3.1. Both intra- and inter-assay analyses were measured as coefficient of variance (also RSD). The intra-assay analysis representing the precision of the developed method was calculated within a range of 0.18 to 0.33 % in relation to two different standards (internal standards). Subsequently, the inter-assay analysis of the method ranged from 0.5 to 0.71 %. Thus, in both techniques, the precision were found to be excellent with values < 0.75% RSD showing that the results by these automated analytical methods demonstrate good precision for all compounds.

3.2.3 Quantitation of phytoestrogens- Calibration plots

The standard stock solution containing at least eight standards at a concentration of 100 ng/μL was serially diluted with the aim of calibration, resulting to six calibration solutions at 1000, 800, 600, 400, 200 and 50 pg/μL. These solutions were injected into the LCMS employing either developed method A or B mentioned earlier. Based on the calibration plots of each standard, both methods showed good linearity as expressed by the individual r^2 coefficient values being closer to 1, ($r^2 > 0.99$).

In summary of the calibration plots, TABLE 3.2 a & b display detailed summarised linearity regression of concentrations versus chromatographic response in terms of ion count per second (cps) and their r^2 coefficient.

TABLE 3.2: a. Test for Method Linearity (Method A)

Compound	Slope	R^2 value	LOD* (pg/μL)	LOQ* (pg/μL)
Secoisolariciresinol	y = 0.347x + 19.648	0.9970	20	60
Matairesinol	y = 0.2224x + 31.85	0.9924	30	90
Enterodiol	y = 1.6663x	0.9927	15	45
Daidzein	y = 8.149x + 272.54	0.9941	10	30
Enterolactone	y = 2.6411x	0.9977	10	30
Equol	y = 0.3519x + 20.415	0.9913	50	150
Genistein	y = 6.5963x	0.9953	10	30
o-desmethylangolensin	y = 2.1661x + 75.766	0.9928	10	30

LOD*: limit of detection and LOQ*: limit of quantitation

TABLE 3.2: b. Test for Method Linearity (Method B)

Compound	Slope	R^2 value	LOD (pg/μL)	LOQ (pg/μL)
Secoisolariciresinol	y = 0.1568x + 17.637	0.9902	15	45
Matairesinol	y = 0.0933x + 11.564	0.9924	30	90
Enterodiol	y = 0.3497x + 29.111	0.9918	15	45
Daidzein	y = 1.3025x + 54.418	0.9948	10	30
Enterolactone	y = 0.2456x + 23.114	0.9918	10	30
Equol	y = 0.0157x + 3.5	0.9965	55	165
Genistein	y = 0.2614x + 30.779	0.9901	10	30
o-desmethylangolensin	y = 0.1565x + 22.1	0.9907	15	45

LOD*: limit of detection and LOQ*: limit of quantitation

3.2.4 Recovery

A recovery experiment was performed using the proposed extraction and analytical protocols to validate both the efficiency of the extraction method and the precision/ accuracy of chromatographic and spectrometric methods for the quantitation of phytoestrogens and their metabolites in bio-fluids. Using a 'blank sample', in this case-water, known amounts of stock solution were spiked into the blank, in the same way as the samples, extracted and analysed with the developed methods. The selected concentrations were 50, 2000 and 5000 ng each made up in 1 mL solution. These recovery experiments were carried out in duplicates for each of the three levels of standard spiked in the blank. Similarly, a recovery experiment was performed using liquid-liquid extraction at 5000 ng.

TABLE 3.3: Recovery of phytoestrogens and their metabolites in spiked 'blank'.

Compound	SPE (%)			LLE (%)
	50 ng	2000 ng	5000 ng	5000 ng
Secoisolariciresinol	94.94 ± 0.91	99.68 ± 0.26	96.73 ± 2.57	98.53 ± 3.45
Matairesinol	100.24 ± 1.09	100.02 ± 0.52	101.41 ± 2.32	99.61 ± 2.82
Enterodiol	94.69 ± 0.71	99.34 ± 0.89	97.98 ± 0.94	94.53 ± 3.64
Daidzein	104.95 ± 2.26	102.54 ± 0.77	108.83 ± 2.15	100.38 ± 2.19
Enterolactone	101.17 ± 2.69	98.83 ± 0.81	105.44 ± 1.79	97.42 ± 3.27
Equol	97.82 ± 3.56	99.15 ± 0.63	99.21 ± 1.38	92.76 ± 1.75
Genistein	98.35 ± 0.56	100.72 ± 0.95	99.90 ± 1.37	99.95 ± 0.93
o-DMA	96.75 ± 2.27	98.24 ± 0.74	97.56 ± 2.14	95.27 ± 2.48

The recovery tests of the phytoestrogens and their metabolites in this study were found to be between 96 – 100 % as shown in TABLE 3.3. On the average, all the standards gave approximately 100 % recoveries using SPE; whilst for LLE, the recovery obtained was between 92 – 98 % for the metabolites. Thus, as the recovery obtained using SPE was higher than that of LLE; SPE was adopted as the final clean-up or extraction of phytoestrogen from the urine in this study.

3.2.5 General changes in phytoestrogen metabolic profile

A typical (overlaid) LC-MS-TIC and LC-MS-SIM chromatogram of an African urine sample collected during soy consumption as shown in FIGURE 3.4. Observing the LC result, the hydrophilic and high polar compounds and the solvent gave the most abundant responses resulting as the dominant peak observed in the TIC; however on implementing the SIM mode, the metabolites were easily visible as sharp peaks.

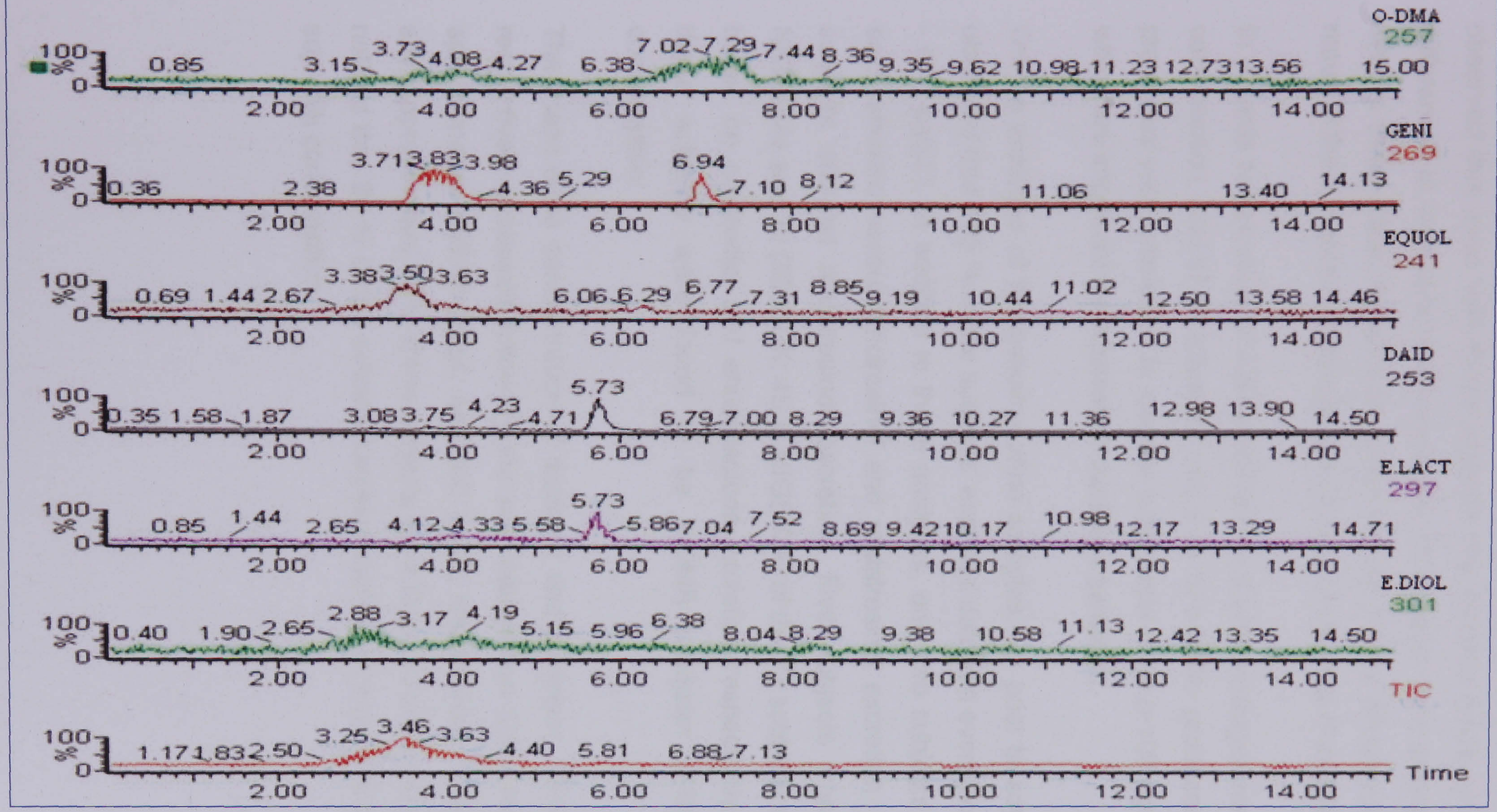


FIGURE 3.3: Typical LC-MS (TIC and SIM) overlaid chromatogram on a subject's urine.

Collected on intervention Day 3 after soy-milk consumption (during/ Phase II).

From the entire subjects' results obtained representing the baseline (day 1) of urine samples, it showed that 39, 23, 18, 10, 22 and 9 out of 42 subjects excreted daidzein, genistein, equol, o-DMA, enterolactone and enterodiol respectively. Furthermore, it was observed that even with those subjects that excrete these compounds, there were vast differences in their excretion levels and as a result, their standard deviation values were two- to three- folds higher than mean values. These differences could be reported as a result of the variation in habitual diets of each individual (FIGURE 3.4).

In regards to excretion of substantial levels of phytoestrogen metabolites prior to soy-milk consumption, (out of 42 subjects) only 5 of 10 o-DMA- producers and 17 out of 18 equol-producers were observed to excrete substantial level (>100 pg/μL) of o-DMA and equol, which are important metabolites of daidzein/ genistein.

On the analyses of the baseline urine samples, i.e. prior to soymilk consumption, it was observed that >95 % of the subjects' excreted daidzein even though at very low levels (10 - 675 pg/μL). In addition to these analyses, only two subjects (both in Nigerian cohort) each showed secoisolariciresinol and matairesinol excretion, whilst one of twenty UK subjects showed matairesinol excretion. Five subjects from each cohort (UK and NIGERIA) and 22 (55% UK: 45 % NIGERIA) of the 42 subjects showed lignan metabolism based on enterodiol and enterolactone excretion respectively. Further to this, 30% of these subjects were found to be o-DMA producer (>100 pg/μL) following soy consumption.

The mean urine concentration of daidzein and genistein: 73.91 pg/μL and 53.30 pg/μL respectively increased following the soy intervention (20 – 40 mg/ serving or day) to approximately 435.66 pg/μL and 406 pg/μL respectively. Following the consumption of soy, approximately 50% of the subjects were found to be equol producers, however it was noticed that 20 % of the cohorts excreted fairly high level of equol (>1000 pg/μL) prior to soy-milk consumption.

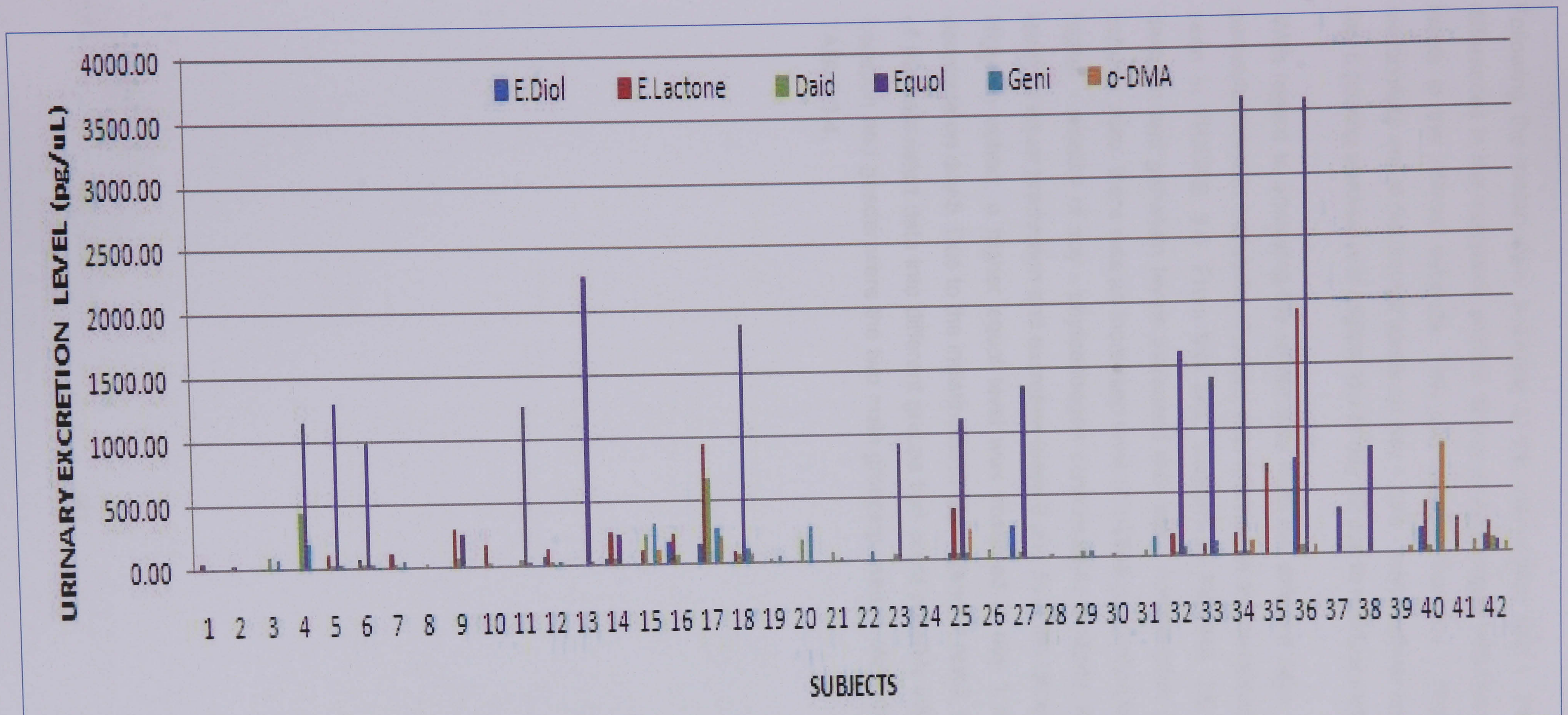
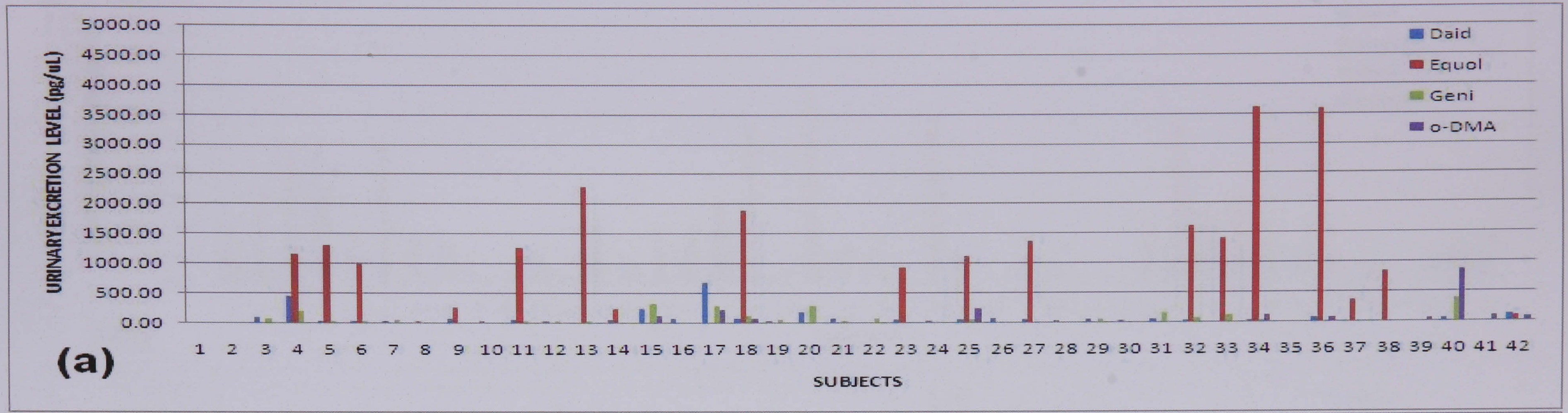
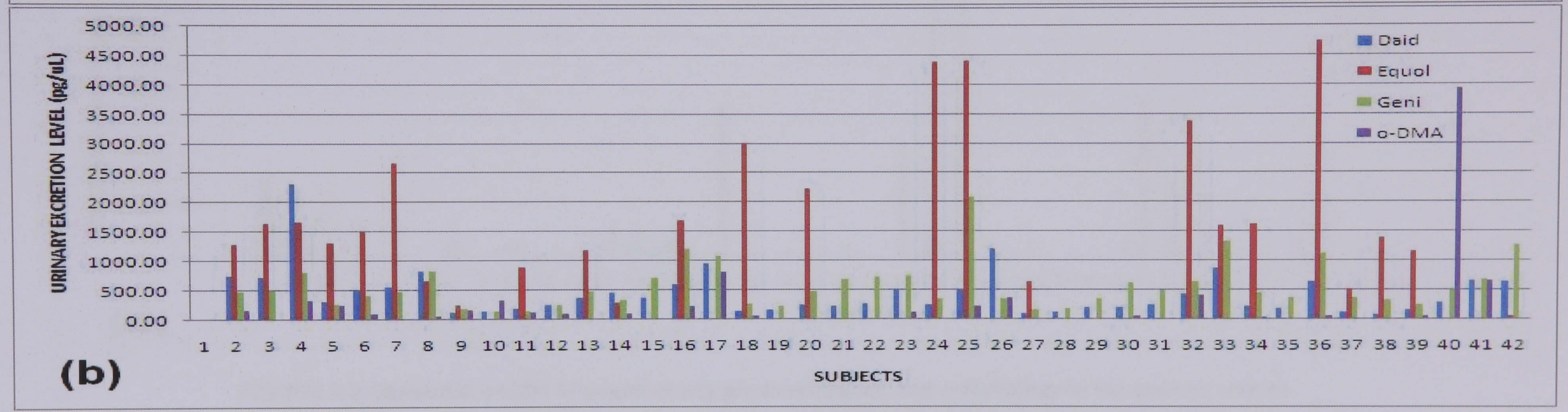


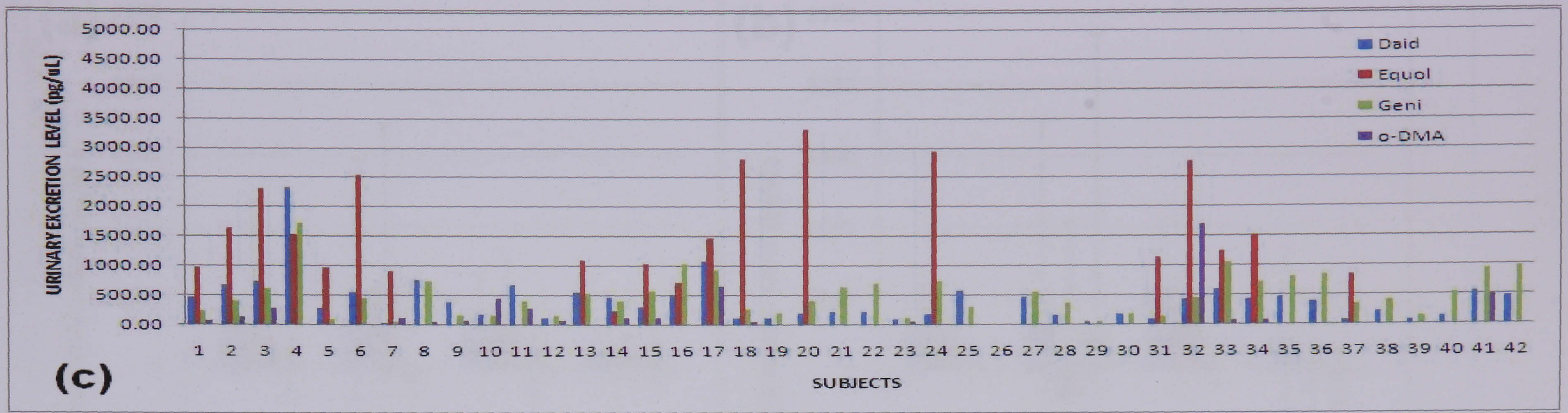
FIGURE 3.4: Representative plot of six metabolites excretion levels in African subjects.

Following the metabolites' analyses of the intervention day 1, FIGURE 3.5 shows the differences in the metabolic profiles of soy phytoestrogen and their metabolites excretion levels in the African subjects. This plot showed that soy - phytoestrogens and their metabolites are at the lowest levels on day 1 (pre - soy), highest on day 3, slightly high on day 5 (during phases) and slightly low at day 12 due to residual metabolites excretion.

With regard to analysing the whole data from the different days, it was noticed that no particular trend in metabolic changes was followed in the metabolism of phytoestrogen as seen in FIGURE 3.6. From this plot, similarly to FIGURE 3.5, it was observed that daidzein and genistein levels increased with soy - consumption (days 3 and 5) in both cohorts. Also, there was an increased level of o-DMA production in both cohorts following higher - amount of soy - phytoestrogen consumption. Similarly, there was an increased level on equol production and excretion (days 3 and 5) in the UK cohort; however with the Nigerian cohort, a higher equol level was observed in day 1 only, with an averaged decrease on day 5. Due to the indistinctness of the general result, this led to the assigning of the calculated data into different groups that could possibly influence soy metabolism. Location and gender were the two main grouping used in clarifying the data as shown in TABLE 3.4.


URINARY EXCRETION LEVEL (pg/uL)
5000.00
4500.00
4000.00
3500.00
3000.00
2500.00
2000.00
1500.00
1000.00
500.00
0.00
Daid
Equol
Geni
o-DMA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42
SUBJECTS
(a)


URINARY EXCRETION LEVEL (pg/uL)
5000.00
4500.00
4000.00
3500.00
3000.00
2500.00
2000.00
1500.00
1000.00
500.00
0.00
Daid
Equol
Geni
o-DMA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42
SUBJECTS
(b)



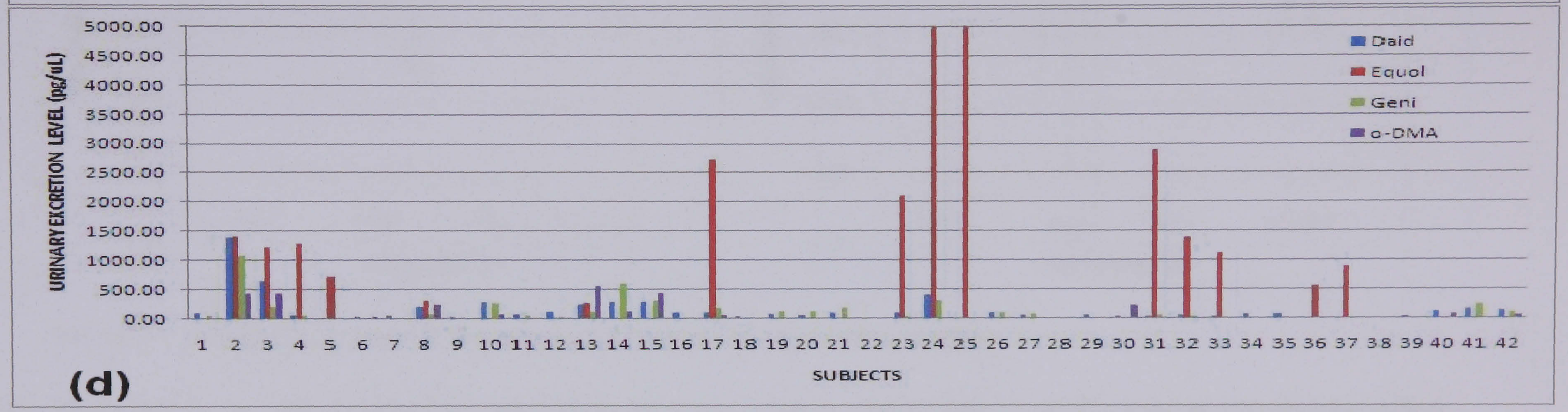


FIGURE 3.5: Metabolic profile changes of soy phytoestrogens and metabolites in the general cohort.

(a) Day 1 (pre-soy); (b) Day 3; (c) Day 5; (d) Day 12 (post-soy). Daid- daidzein, Geni- genistein; o-DMA- o-desmethylangolensin.

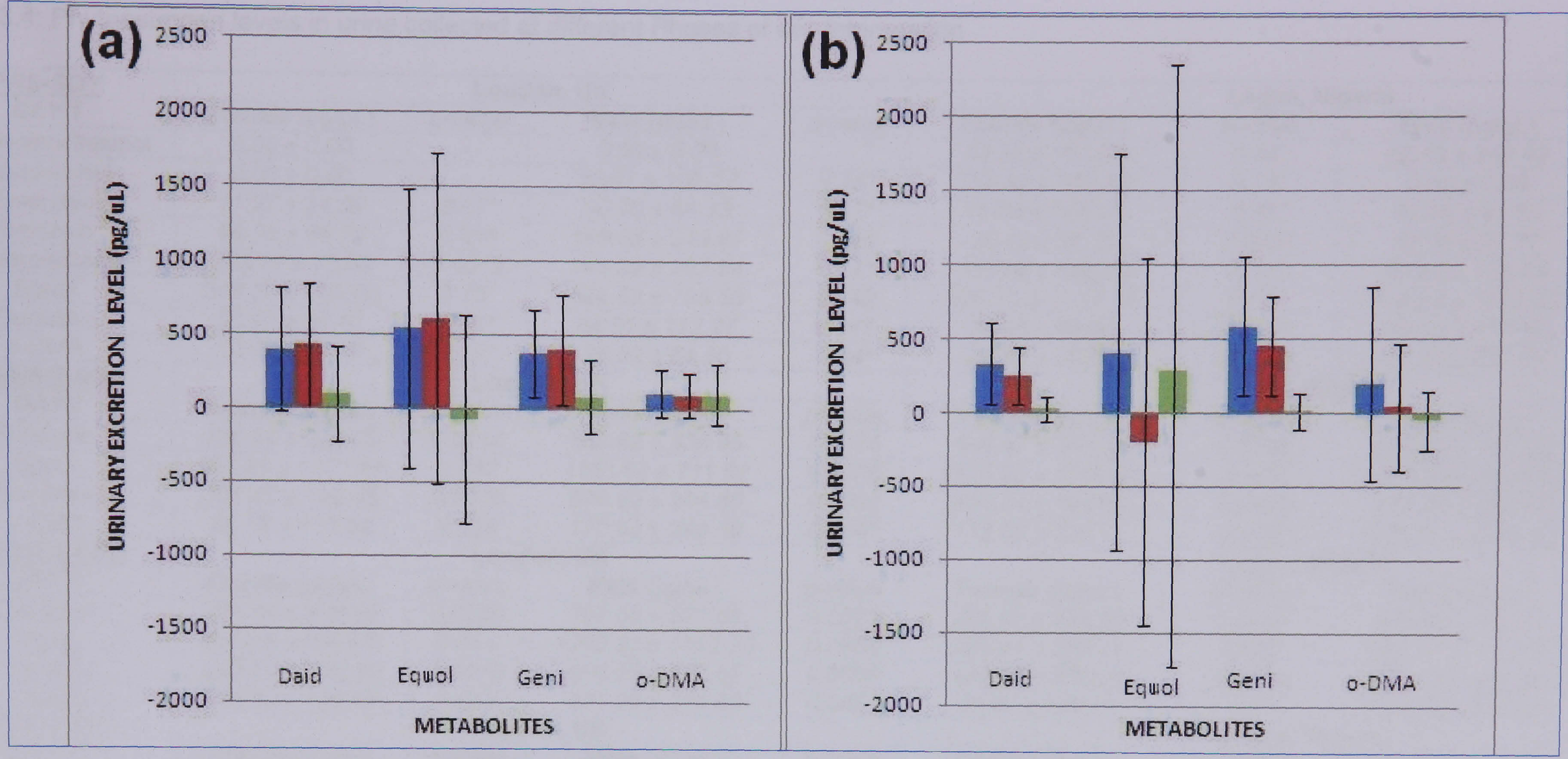


FIGURE 3.6: Comparison plots of changes in urinary excretion levels of phytoestrogen

(± standard deviation) from Day 3, 5 and 12 using Day 1 as baseline reference. • Day 3; •Day 5; •Day 12. (a) UK cohort and (b) Nigerian cohort. Daid- daidzein, Geni- genistein; o-DMA- o-desmethylangolensin.

TABLE 3.4: Phytoestrogen levels in urine collected at different Phases of soy intervention

PRE-SOY DAY 1	**London, UK**				**Lagos, Nigeria**			
	Female (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*	**Female** (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*
Secoisolariciresinol	*0.00 ± 0.00*	-	*0.00 ± 0.00*	-	*50.36 ± 167.03*	*0.34**	*72.49 ± 240.42*	*0.34**
Matairesinol	*0.00 ± 0.00*	-	*33.91 ± 107.22*	*0.34**	*157.97 ± 356.63*	*0.17**	*0.00 ± 0.00*	-
Enterodiol	*11.27 ± 24.09*	*0.17**	*30.48 ± 64.33*	*0.17**	*72.98 ± 228.26*	*0.31**	*51.18 ± 91.80*	*0.094**
Daidzein	*56.08 ± 68.76*	*0.030*	*166.85 ± 219.87*	*0.040*	*33.49 ± 31.72*	*0.0057*	*46.06 ± 32.77*	*0.0009*
Enterolactone	*103.27 ± 75.58*	*0.0019*	*151.99 ± 297.94*	*0.14**	*310.49 ± 586.30*	*0.11**	*87.69 ± 135.89*	*0.058**
Equol	*343.57 ± 680.90*	*0.15**	*594.83 ± 795.55*	*0.042*	*587.72 ± 1117.14*	*0.11**	*768.64 ± 1123.67*	*0.047*
Genistein	*53.03 ± 97.01*	*0.12**	*84.51 ± 112.27*	*0.041*	*28.45 ± 48.50*	*0.080**	*50.02 ± 111.50*	*0.17**
o-DMA	*18.82 ± 41.47*	*0.19**	*20.52 ± 64.90*	*0.34**	*38.69 ± 70.21*	*0.098**	*96.33 ± 261.24*	*0.25**
DURING-SOY DAY3	**London, UK**				**Lagos, Nigeria**			
	Female (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*	**Female** (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*
Daidzein	*320.64 ± 223.05*	*0.0014*	*689.82 ± 636.95*	*0.0076*	*449.97 ± 364.29*	*0.0022*	*294.87 ± 170.26*	*0.0002*
Equol	*853.61 ± 1171.51*	*0.047*	*1163.82 ± 711.57*	*0.0006*	*1127.20 ± 1790.51*	*0.063**	*1043.56 ± 1543.33*	*0.049*
Genistein	*301.49 ± 195.48*	*0.0009*	*599.22 ± 264.86*	*0.0006*	*680.34 ± 594.64*	*0.0035*	*577.65 ± 292.46*	*0.0001*
o-DMA	*93.19 ± 112.54*	*0.028*	*170.03 ± 245.38*	*0.056**	*118.55 ± 210.71*	*0.092**	*408.24 ± 1174.47*	*0.28**
DURING-SOY DAY 5	**London, UK**				**Lagos, Nigeria**			
	Female (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*	**Female** (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*
Daidzein	*284.46 ± 208.55*	*0.0020*	*787.56 ± 591.59*	*0.0023*	*286.91 ± 235.65*	*0.0024*	*286.02 ± 141.72*	*0.0001*
Equol	*855.28 ± 890.06*	*0.014*	*1295.82 ± 1162.25*	*0.0065*	*292.41 ± 509.17*	*0.086**	*657.21 ± 1178.65*	*0.094**
Genistein	*245.77 ± 166.46*	*0.0012*	*696.73 ± 447.52*	*0.0008*	*444.72 ± 385.06*	*0.0033*	*538.44 ± 252.90*	*0.000*
o-DMA	*108.32 ± 126.49*	*0.024*	*131.09 ± 214.93*	*0.086**	*51.67 ± 148.29*	*0.27**	*163.35 ± 508.69*	*0.31**
POST-SOY DAY 12	**London, UK**				**Lagos, Nigeria**			
	Female (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*	**Female** (pg/μL)	*p-value*	**Male** (pg/μL)	*p-value*
Daidzein	*97.17 ± 93.51*	*0.0094*	*315.13 ± 418.47*	*0.041*	*75.45 ± 113.72*	*0.052**	*57.25 ± 44.24*	*0.0016*
Equol	*427.19 ± 904.37*	*0.17**	*364.80 ± 552.13*	*0.066**	*1296.08 ± 2442.40*	*0.11**	*637.91 ± 1020.94*	*0.065**
Genistein	*105.25 ± 179.02*	*0.096**	*208.28 ± 324.33*	*0.073**	*40.03 ± 91.00*	*0.18**	*63.31 ± 80.69*	*0.026*
o-DMA	*78.47 ± 175.97*	*0.19**	*164.91 ± 204.79*	*0.031*	*30.94 ± 67.73*	*0.16**	*0.00 ± 0.00*	-

Identical values, i.e. no p-value calculated. * p values > 0.05.

Following the grouping of datasets, on examining the baseline (day 1) results in TABLE 3.4, it showed that none of the hypothesised factors (location and/ or gender) showed to influence soy- phytoestrogen and their metabolites excretion except in a few odd cases. These few cases were noticed in daidzein all four groupings, equol Nigerian and UK male groupings and genistein UK male only; all with p-values < 0.05. Similarly, this non-grouping ($p > 0.05$) was observed with secoisolariciresinol, matairesinol, enterodiol and enterolactone excretion except with UK females ($p = 0.0019$).

With soy metabolism in day 3 & 5, daidzein excretion showed to be well grouped with all p- values < 0.05. Genistein excretion was found to be significantly different in each grouping except in UK male subjects. Equol production/ excretion showed to be slightly location-influenced, i.e. UK cohorts' (both male and female) showed $p < 0.05$, implying no significant difference within each group, whilst the Nigerian cohort showed to be significantly different with $p > 0.05$. Comparatively, o-DMA production and excretion were significantly different with $p > 0.05$ in all grouping except UK females. On the overall, the profile of day 3 and 5 in TABLE 3.5 showed to be comparable.

Focussing on the metabolic effect changes of soy-milk consumption (day 3 & 5), daidzein and genistein showed improved p-values (i.e. in all groupings $p < 0.01$) compared to the day 1 (pre-soy) sample analyses. Thus, similarity in these soy phytoestrogens excretions implies the similarities of soy metabolism within the selected African subjects.

Finally, considering the post-effect of soy consumption on day 12, the results were found to be similar to the pre-soy intervention on day 1, only with slight higher level than the baseline levels, particularly with daidzein and genistein. As well, equol excretion showed to be significantly different in the four groups (locations & genders) with $p > 0.05$. Similarly, o-DMA production was significantly different in each of the two cohorts' female subjects, whilst the UK male subjects were found to be insignificantly different with $p = 0.031$. Contrary to this, Nigerian male subjects were not o-DMA producers following their regular habitual diets.

Although, the grouping can explain soymilk metabolism taking into account the p-values of some of the phytoestrogens and their metabolites with $p < 0.05$, a more detailed analysis is necessary to show and interpret soy metabolism.

3.2.6 Factors influencing soy metabolism

Focussing on soy metabolism with the aim of understanding metabolic changes, soy phytoestrogens and their metabolites were only targeted and discussed in this section. Considering daidzein metabolism and excretion, upon analysing the general data, there were no observed similarities in the urinary excretion levels. Thus, this led to the grouping into the two cohorts based on location, which was then further classified into gender groups. With this taxonomy as shown in FIGURE 3.7, there were noticeable similarities amongst the subjects in some of the derived groups.

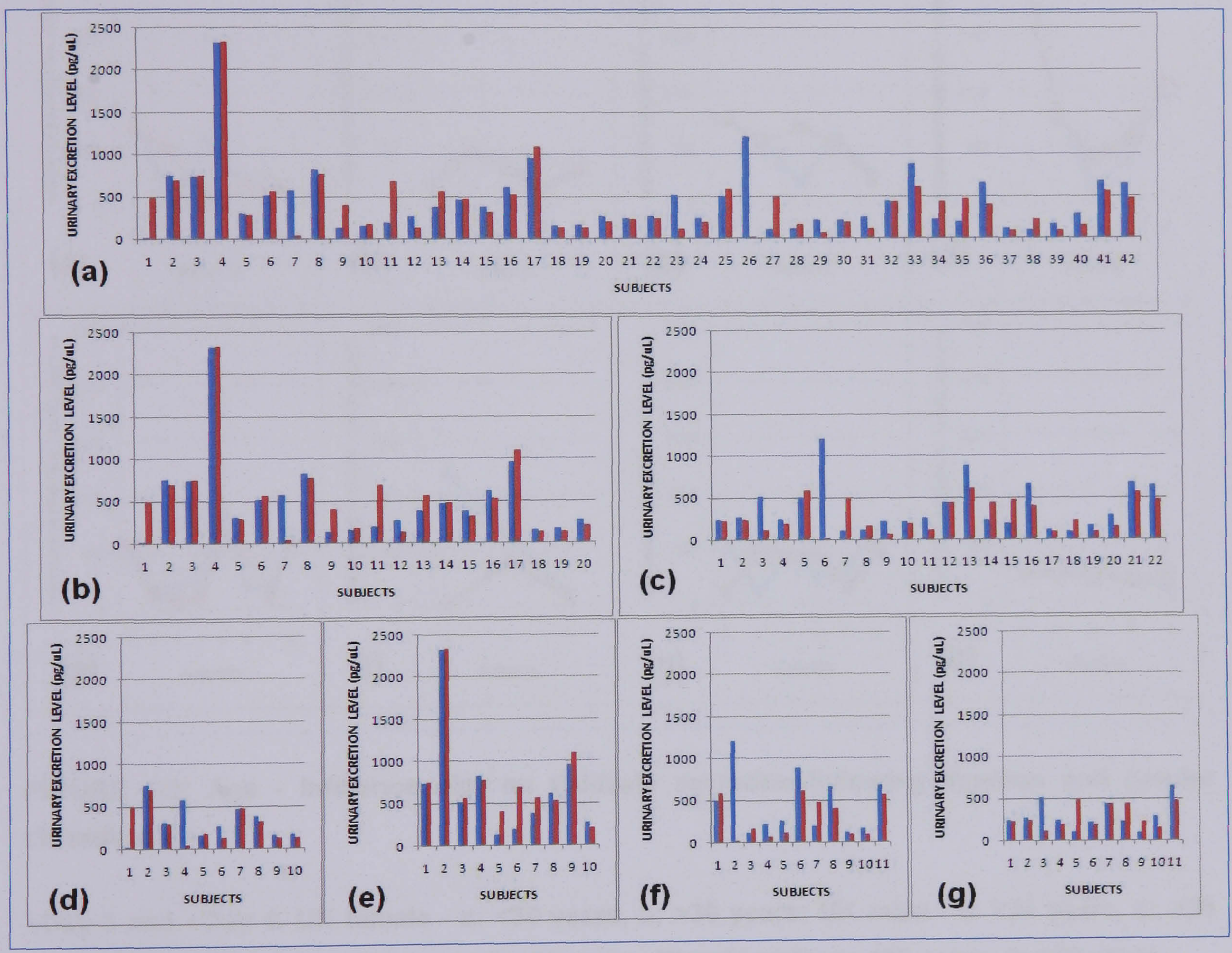


FIGURE 3.7: Taxonomy of hypothesised factors influencing daidzein excretion during soy consumption.

•Day 3 and •Day 5: a: general cohort; location categorisation- b: UK, c: Nigeria; gender categorisation- d: UK female, e: UK male, f: Nigeria female and g: Nigeria male.

On visual interpretation of this taxonomy of daidzein metabolism, it can be seen that the UK male had the highest urinary excretion level following soymilk consumption. Also, it can be reported that the male subjects in the Nigeria cohort (FIGURE 3.7g) showed to have similar biochemical effects on daidzein, i.e. urinary excretion level between 50 – 650 pg/µL, whilst the other three had high variances (slightly in UK female: FIGURE 3.7d) within each groups. Due to this high variation between groups, age was considered a possible factor, which could influence soy metabolism in terms of daidzein excretion.

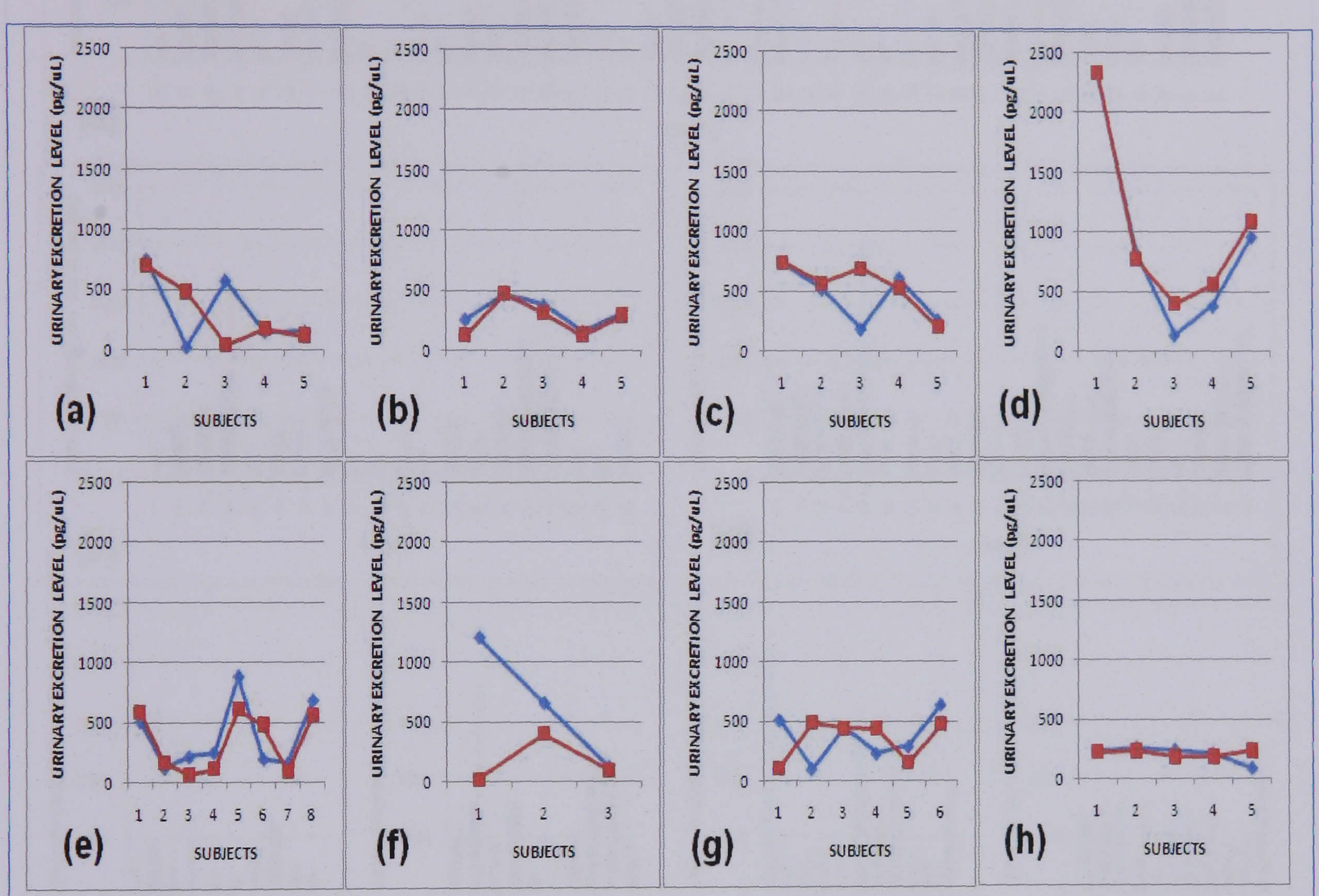


FIGURE 3.8: Age - influence plot on daidzein excretion following location and gender classification.

•Day 3 and •Day 5. UK female - a: <30 years, b: >30 years; UK male - c: <30 years, d: >30 years; Nigeria female- e: <30 years, f: >30 years; Nigeria male- g: <30 years, h: >30 years.

FIGURE 3.8 showed perfect groupings in Nigeria male <30 & >30 years, reasonable grouping in UK female >30 years, weak grouping in UK female <30 years, UK male <30 years, Nigeria female <30 years, and poor groupings in UK male >30 years and Nigeria female >30 years.

Following the analyses on daidzein excretion to monitor metabolism, genistein excretion was also considered in a likewise investigation, i.e. location followed by gender taxonomy.

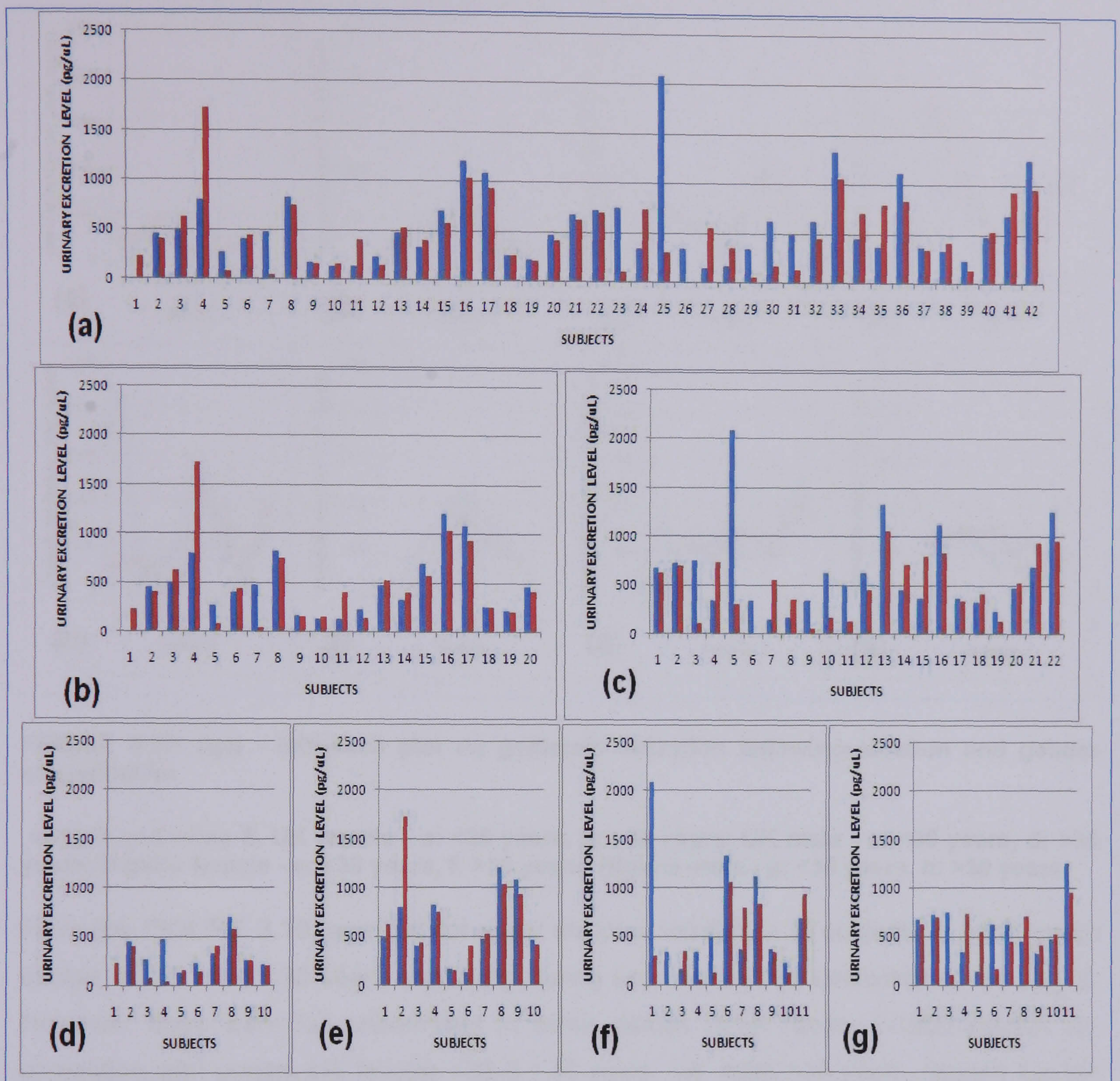


FIGURE 3.9: Taxonomy of hypothesised factors influencing genistein excretion (during soy consumption)

•Day 3 and •Day 5: a. general cohort; Location categorisation- b. UK, c. Nigeria; Gender categorisation- d. UK female, e. UK male, f. Nigeria female and g. Nigeria male.

As seen in FIGURE 3.9, there seemed to be fair similarity in excretion of genistein within each group, however the group with the closest urinary excretion levels was male NIG group. Nonetheless, there is a need for further classification of subjects with the aim of

attaining likeness in excretion levels of genistein. As carried out earlier, age factor is considered as one of the determinant in soy metabolism.

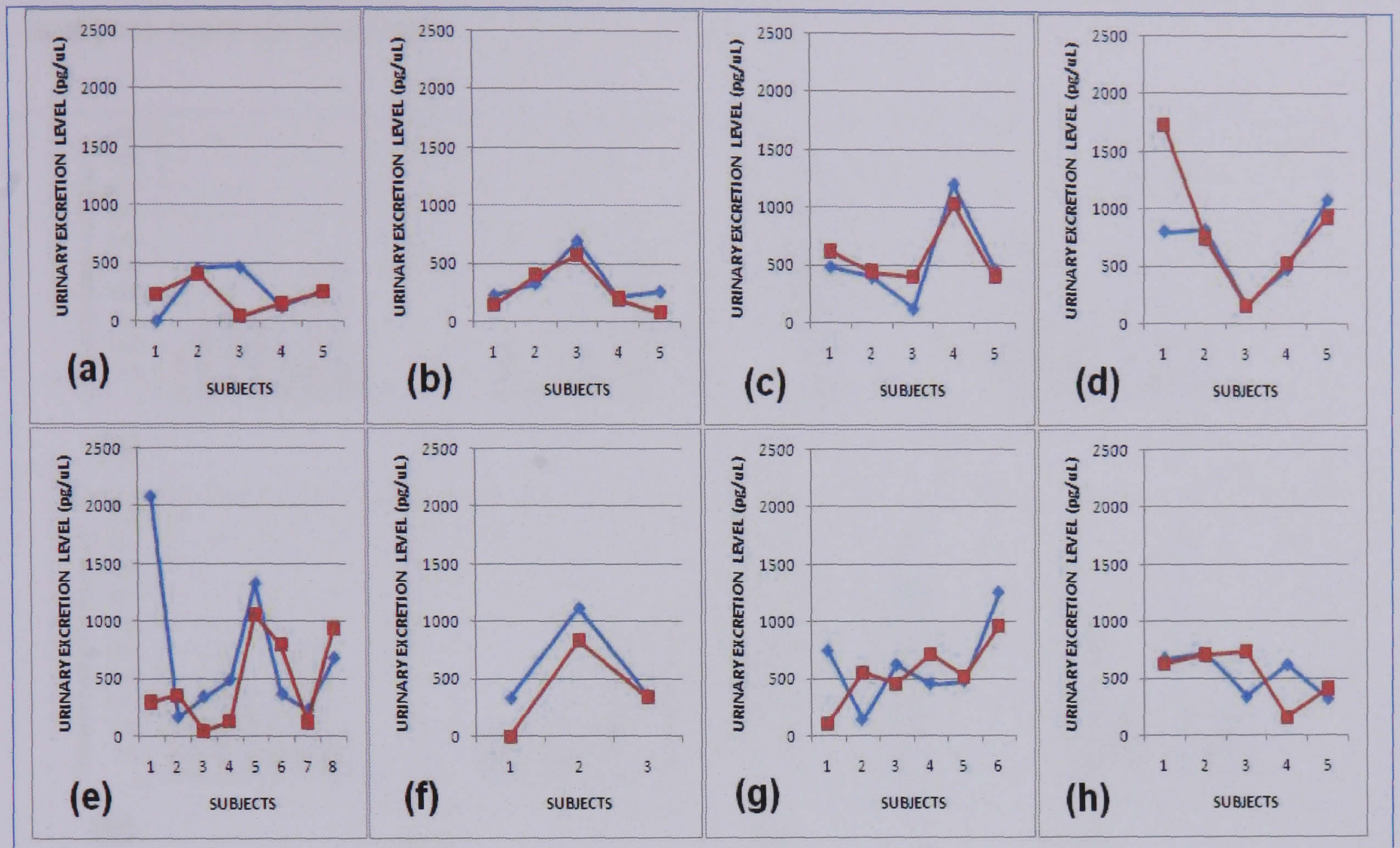


FIGURE 3.10: Age - influence plot on genistein excretion following location and gender classification

•Day 3 and •Day 5. UK female - a: <30 years, b: >30 years; UK male - c: <30 years, d: >30 years; Nigeria female - e: <30 years, f: >30 years; Nigeria male - g: <30 years, h: >30 years.

From the FIGURE 3.10, age did not show obvious correlation of subjects in each group except in UK male <30 years especially during soy intervention consumption on Day 3. However, there were fair uniformities in some groups (level range: 0-1000 pg/µL), i.e. correlation with groups UK female <30 & >30 years, UK male >30 years, Nigeria female <30 years and Nigeria male > 30 years.

Pertaining to daidzein and genistein metabolites i.e. equol and o-DMA production, similar analyses were carried out.

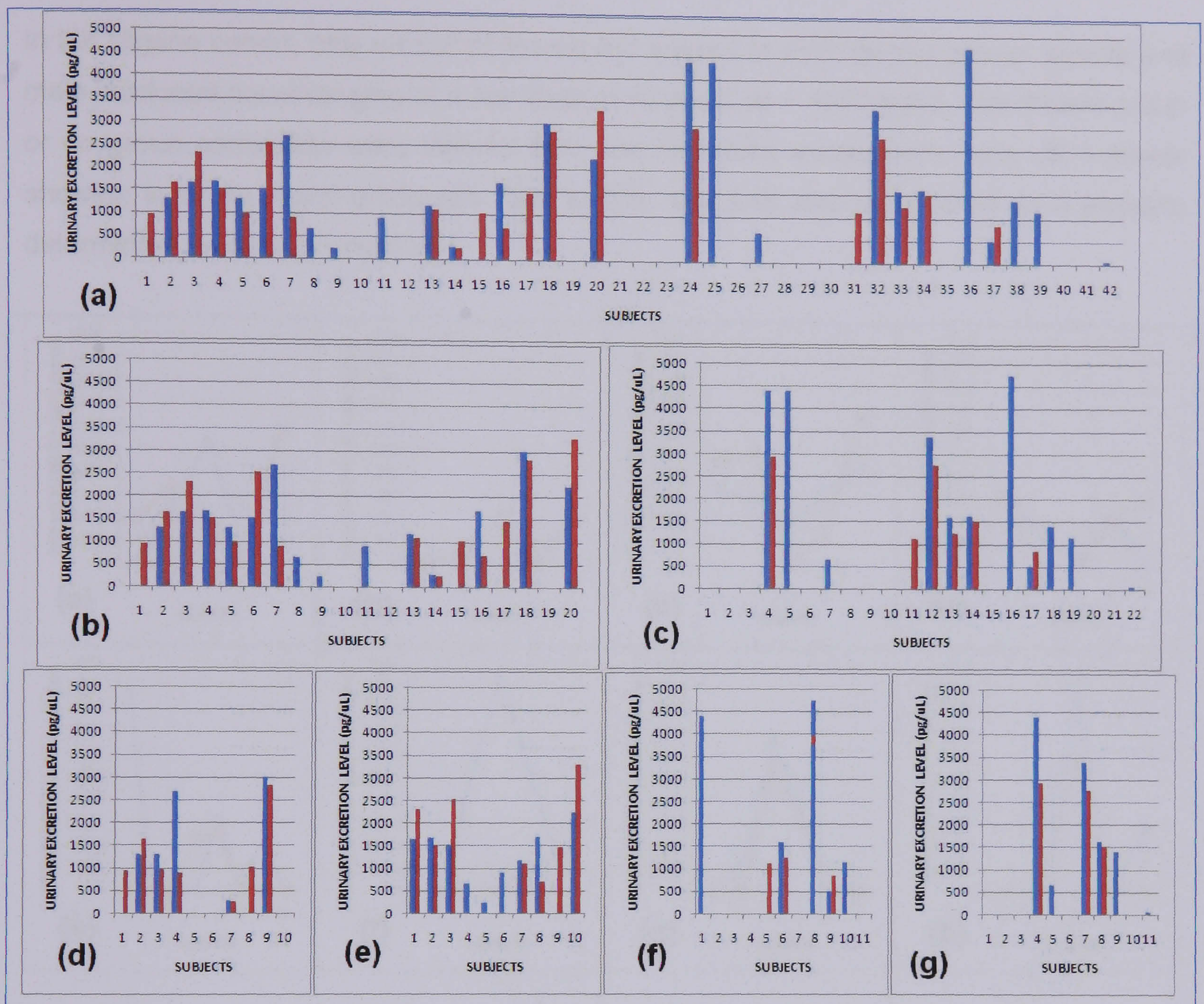


FIGURE 3.11: Taxonomy of hypothesised factors influencing equol production and excretion during soy-milk consumption (Phases II & III).

•Day 3 and •Day 5: a. general cohort; Location categorisation- b: UK, c: Nigeria; Gender categorisation- d. UK female, e. UK male, f. Nigeria female and g. Nigeria male.

Generally, it was found that equol production was not consistent during the soy consumption as some subjects produced in only one out of the two samples assessed during soymilk intervention. As seen in FIGURE 3.11, the UK cohort showed more subject producing equol compared to the Nigeria cohort subjects, however, those equol producers in the Nigeria cohort had higher equol excretion close to 5000 pg/μL (max) than the UK

subjects (~3500 pg/μL at max). Subsequently, it was observed that at either or both during phases (days 3 & 5), all the male UK subjects produced equol ranging from 100 to ~3500 pg/μL, whilst 7 out of 10 female UK subjects produced equol ranging from 100 to ~3000 pg/μL.

In the Nigeria cohort, only six out of the eleven subject in each gender group- female and male produced equol ranging at a low level of 40 pg/μL to ~ 4500 pg/μL. Insufficient group or no group correlation were derived from the taxonomy except with male UK subjects showing to all be equol-producers. Due to this, age was also considered as a possible determinant in equol production.

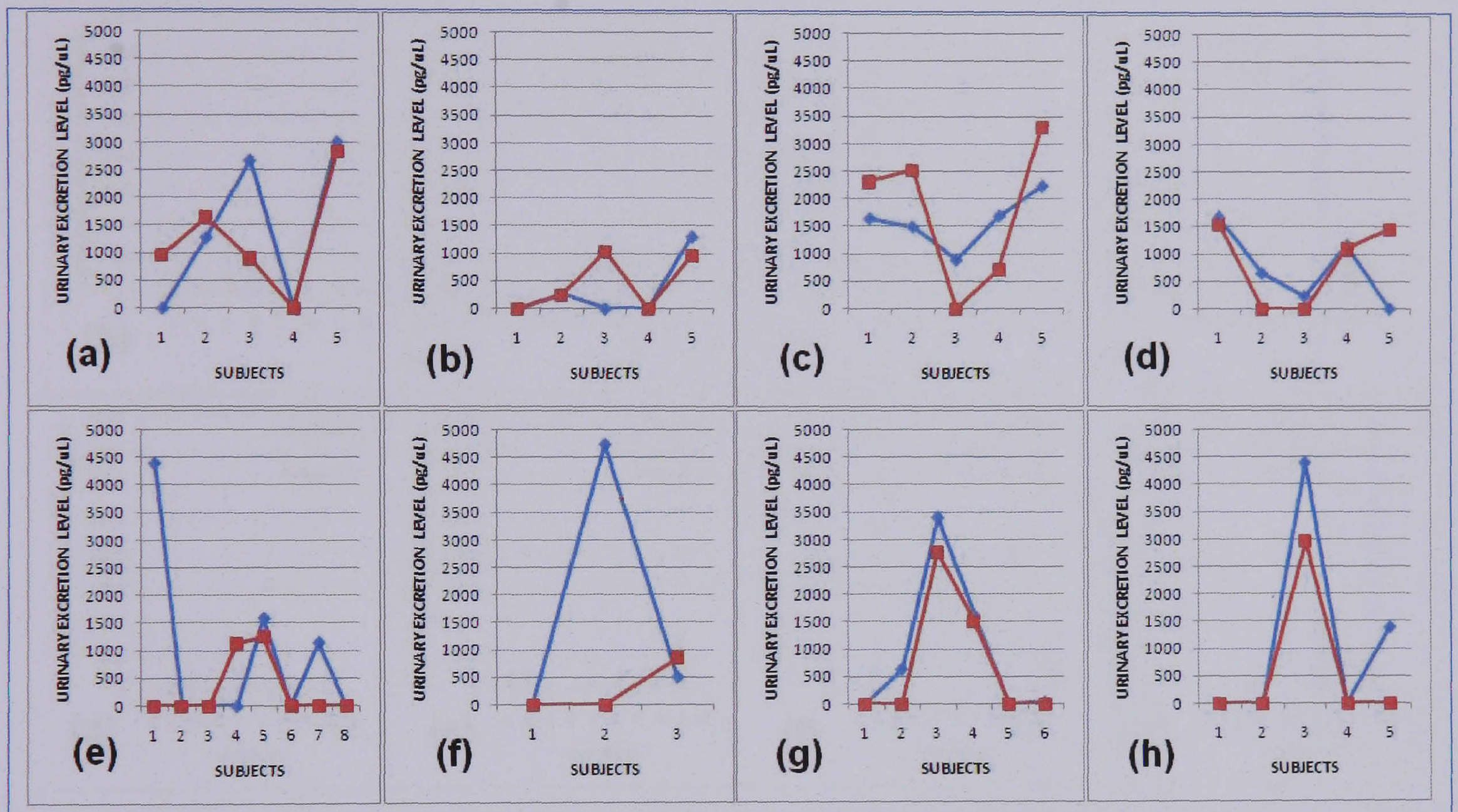


FIGURE 3.12: Age - influence plot on equol production and excretion following location and gender classification.

•Day 3 and •Day 5. UK female - a: <30 years, b: >30 years; UK male - c: <30 years, d: >30 years; Nigeria female - e: <30 years, f: >30 years; Nigeria male - g: <30 years, h: >30 years.

FIGURE 3.12 above shows no valid age influence or correlation between the equol production and excretion as the levels of urinary equol excretion seemed to vary within each of the eight groups. The only convincing correlation found was with the male UK <30 years.

Similarly, considering o-desmethylangolensin production and excretion following soy-milk consumption during intervention, parallel analyses were carried out on the calculated o-DMA levels.

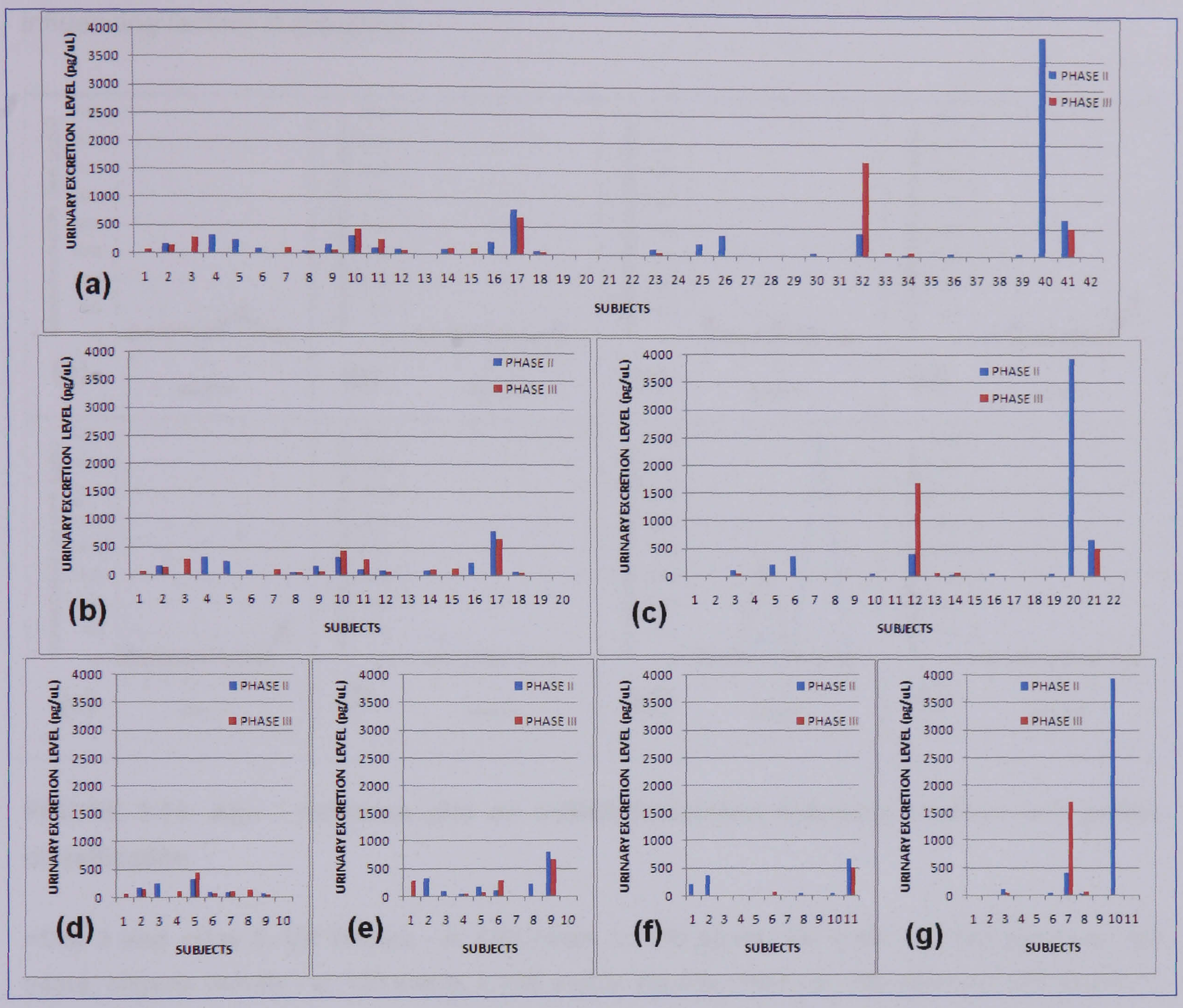


FIGURE 3.13: Taxonomy of hypothesised factors influencing o-DMA production and excretion (during soy consumption).

•Day 3 and •Day 5. a: general cohort; Location categorisation- b: UK, c: Nigeria; Gender categorisation- d: UK female, e: UK male, f: Nigeria female and g: Nigeria male.

Visual analysis of the charts in FIGURE 3.13 (g) showed that two subjects (both male NIG subjects) produced and excreted high levels of o-DMA, which was approximately 2000 – 4000 pg/µL compared to the other subjects excreting roughly 0 – 1000 pg/µL. The UK cohort was found to have a better correlation within the group, i.e. having similar o-DMA excretion levels. The female UK groups appeared to be significantly similar having urinary excretion levels <500 pg/µL. Out of 11 gender-grouped NIG subjects each, 6 and 5 female and male subjects respectively were o-DMA producers. On the other hand, out of 10

gender - grouped UK subjects each, 9 and 8 female and male subjects respectively were o-DMA producers.

Further to this taxonomy on o-DMA producers, age was considered as one of the influencing factors in grouping.

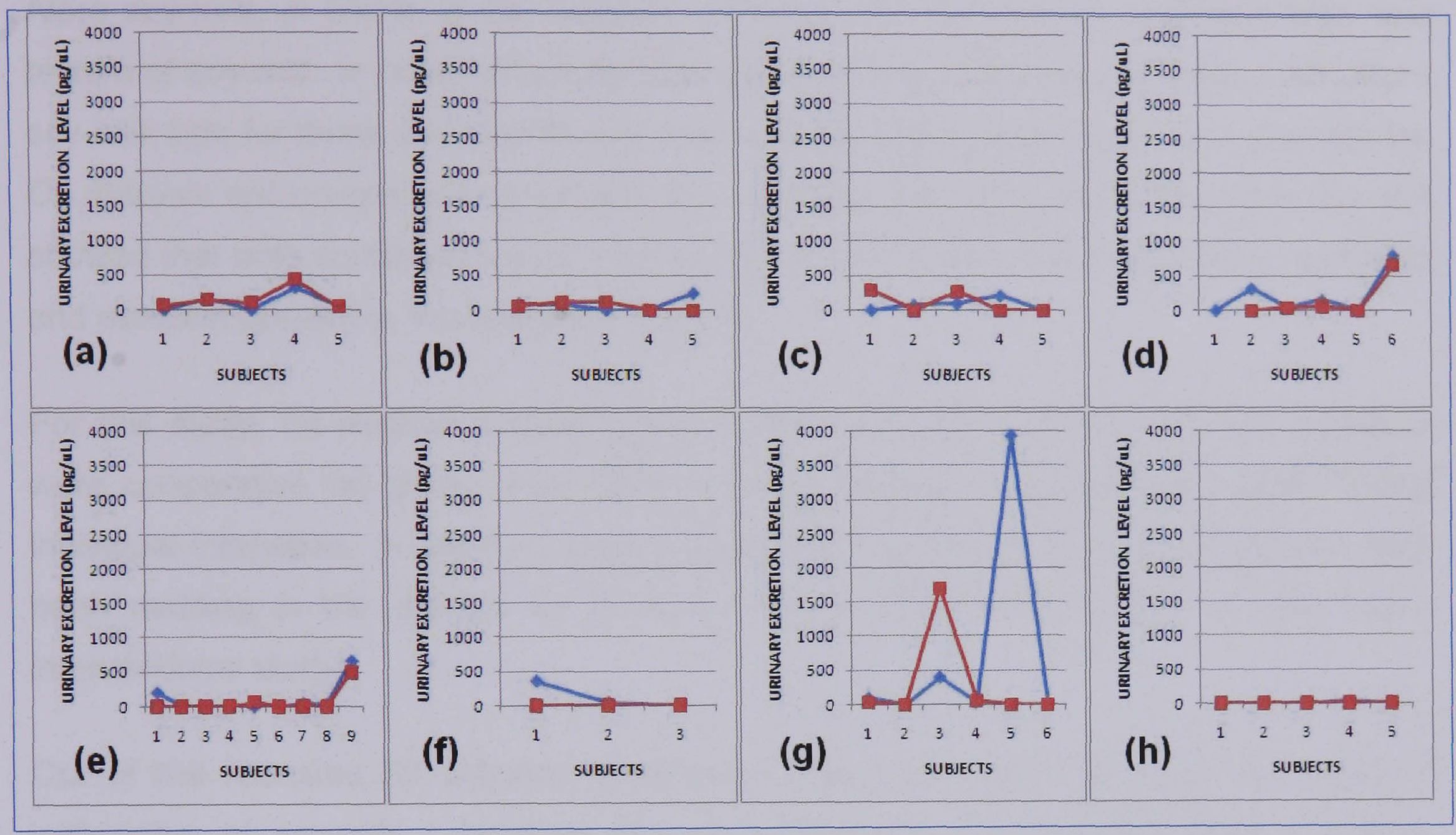


FIGURE 3.14: Age - influence plot on o-DMA production following location and gender classification.

•Day 3 and •Day 5. UK female - a: <30 years, b: >30 years; UK male - c: <30 years, d: >30 years; Nigeria female - e: <30 years, f: >30 years; Nigeria male - g: <30 years, h: >30 years.

Confirmatory to FIGURE 3.14, there are well-defined correlations observed within seven of the eight groups, i.e. except male NIG <30 years. Although, distinct patterns of good correlation were observed as mentioned earlier, with the exclusion of one subject, especially in UK male >30 years and NIG female <30 years, there would have been an even better correlation of subjects in these two groups.

3.3 Discussion

3.3.1 Recruitment of subjects, sample collection

Recruiting of subjects was a big challenge especially to overcome their attitudes towards Alpro soy-milk, in terms of the negative attitudes and perceptions regarding taste and texture of soy-milk. In order to reduce this response, this study was carried out with Alpro soy-milk light for those participants who preferred the lighter soya taste rather than full fat. On analysis and comparison by LC-MS, the results on both light and full fat Alpro soy-milk showed that both contained similar high concentrations of the main isoflavones: genistein and daidzein present in abundance in soy-milk.

For this study, no medical screening was carried out prior analysis; however, subjects were cooperative on giving information on their medical, personal and dietary during individual interviews. In addition, each subject was selected on the basis that they have been residing in the country for at least 3 years prior to participation in this soy - interventional study.

Out of the recruited 50 subjects, 8 subjects pulled out of the study on the basis of indigestion of soy-milk / overload of milk resulting to uncomfortable bloated feeling, cramping and gassing; thus only 42 completed the whole study.

Although, few studies have been reported on soy intervention or 'challenge', so far most of these studies have employed diet intervention or administration of tablets supplements (*66, 67, 326-328*) have employed subject size of 8 subjects (Asian and British women; aged 30 - 55 years) only. The study design involved consumption of soy nutrition food (acetyl - isoflavones) for 7 days, with the collection of urine sample before (as baseline), first three days of 4 consecutive weeks. However, an observed criticism was there was no need and aim for this long sample collection period, as the metabolite's effect over this 4-week period was not mentioned. Another study led by Kano et al (*329*) recruited 12 subjects (9 men and 3 women; Asian; 25 – 51 y) following a 7 - day intervention with the consumption of 100 mL of soy-milk on day 2 and the collection of urine and blood at specific times, e.g. urine: 0, 2, 4, 6, 8, 12, 24, 36 and 48 hours after consumption. Although, this study appeared quite thorough, i.e. over a period of 2 – days with the collection of two bio –fluids, urine and blood, at selective hourly monitoring, however, the effects of long - term soy consumption and effects of possible influencing factors were not

considered as the whole results seemed to have been generalised. Lastly, another similar study was reported by Hedlund (*325*) using 45 subjects (Americans; males only; 19 - 65 y) following a 7 - day soy intervention with the consumption of 330 mL soy-milk daily for 1 - week. However, compared to this present study with urine samples, the biological samples collected by Hedlund were blood (plasma and serum) and prostatic fluid.

3.3.2 Sample clean-up comparison between LLE and SPE

Solid phase extraction was able to extract all the compounds of interest, which include isoflavones (formononetin, biochanin A, daidzein, and genistein), lignans (secoisolariciresinol and matairesinol) and metabolites (enterodiol, equol, enterolactone, and o-desmethylangolensin), whilst, LLE was able to recover isoflavones, lignans and their metabolites at lower recovery levels (as shown in TABLE 3.4). The reason for a better recovery of all the analytes using SPE may be due to equol, and o-DMA having better binding affinities for the column packing than for organic solvent, whilst in LLE, both of these compounds seemed to have had a higher affinity for the aqueous layer than the organic supernatant layer, and so stayed in the aqueous layer. Noteworthy, The recovery achieved in this study was better than that of some reported studies ranging from ~ 90 % to 100% (*330*). In addition, LLE recovery using MTBE on o-DMA showed to be low at ~ 93% (*330*), which is lower than the recovery achieved in this study (~97 % SPE; ~95 % LLE).

3.3.3 Chromatographic separation – method development

Several methods were tried for simultaneous separation of the phytoestrogen and their metabolites standards. These methods included both isocratic and gradient elution modes using methanol, acetonitrile, mixture of acetonitrile and methanol, water, aqueous methanol and aqueous acetonitrile. Eventually, an isocratic chromatographic method using 35% methanol (Method A) or acetonitrile (Method B) in water at 0.15 mL/min on Phenomenex column, C8(2) was developed for a simultaneous LC-MS separation of all standards. Using method A, the elution order of these standards where found to be secoisolariciresinol (4.26 min), matairesinol (4.79 min), enterodiol (5.62 min), daidzein (5.69 min), enterolactone (5.85 min), equol (6.65 min), genistein (6.95 min), o-desmethylangolensin (7.43 min), coumestrol (9.15 min), formononetin (9.58 min) and biochanin A (13.75 min). Although, co-elution of peaks were observed, the use of SIM

mode overcame this problem as the co-eluting compounds have different molecular masses 'M', whereby their $[M-H]^-$ ion were been monitored in the run.

3.3.4 Spectrometric separation and method development

Spectrometric method used for phytoestrogen and their metabolites analyses include EI (using GC-MS (*66*)), ESI (*331, 332*) and APCI (*330*). Of these, ESI is the most popular technique used. In this study, the two main ESI modes, which are positive $[M+H]^+$ and negative $[M-H]^-$ ions were considered for their analyses. However, it was observed that the ESI in negative mode was better than that of the positive mode mainly due to its improved sensitivity and signal responses. The possible reason for this improved sensitivity in negative mode was due to these polyphenolic compounds possessing hydroxyl groups (R-OH) that are likely to loss a proton to give an alkoxide ($R-O^-$). Generally, negative ESI is the more popular ESI technique reported by most LC-MS phytoestrogen studies (*70, 333, 334*).

The parameters of the mass spectrometer were tuned to ensure the best sensitivity and resolutions of mass peaks were obtained. The governing parameters that had the most effects were the capillary, MCP, steering, tube lens, flight-tube and reflectron voltages.

3.3.5 Method validation: limits of detection and quantitation

With a calibration linearity range, 5 to 1,000 pg/µL, good linearity was achieved with the r^2 coefficients for all standards in LC/MS-SIM found to be greater than 0.99. The limits of detection (LOD) was calculated as 3 times the signal to noise ratio using a blank, and the limits of quantification (LOQ), which is 5 times the signal to noise ratio, both were found to be range from 10 to 55 pg/µL and 30 to 165 pg/µL (0.2 to 1.09 pmol and 0.60 to 3.26 pmol) respectively in all standards. Equol had the highest LOD and LOQ values as a result of its less susceptibility to negative ionisation by ES. Comparative to other studies, the sensitivity of this technique towers that of some previous literature values of 0.31 to 37 pmol LOD (*330*).

3.3.6 Intervention diet sample analysis

The level of daidzein and genistein found to be present in soymilk after hydrolysing was 48 – 53 ng/µL daidzein and 53 to 75 ng/µL genistein. Thus, it is deduced that the ingestion

of 200 mL soy-milk would account for roughly 9.6 to 10.6 mg daidzein and 10.6 to 15.0 mg genistein consumption. For method validation, the values obtained were compared with the range of values reported by other researchers. The concentrations of the phytoestrogens present in the analysed samples – soy-milk tallied with the Alpro® expected values of 0.040 – 0.060 mg/mL each of daidzein and genistein.

3.3.7 LC-MS metabolic profiling

Phytoestrogen levels in the urine samples were calculated based on the ion intensity responses (counts per second, cps) obtained from the integration of the selected ion using SIM mode. The obtained peak intensity is compared with the calibration curves values to work out the concentrations of the compounds in the test solution. The concentration in the test solution was re-calculated to obtain the final concentration in the crude urine samples.

After extraction and analysis, the samples were analysed and the peak areas recorded for the analytes compared to "recovery standards". These recovery standards were made up in HPLC mobile phase to produce a final concentration the same as that produced through the extraction of biological standards, assuming 100% recovery. The concentrations of metabolites in each urine samples were calculated with reference to the internal standard using the analyte peak areas averaged for triplicate injections for each cleaned-up sample. With the use of SIM in this study, the sensitivity was improved whereby the urine sample concentration, i.e. sample volume reduction was not necessary.

Also, to ensure possible errors are corrected, a known concentration of internal standards – biochanin A were added to each sample prior LC-MS analysis. The correction factor was calculated based on the ratio of peak responses of the sample to internal standard.

Some of previous studies published on metabolites concentrations showed to be corrected with creatinine, which is excreted at a fixed rate by glomerular filtration (*317*) thus, correcting for possible error in dilution. However, in this study, this measurement was not performed.

3.3.7.1 DAY 1: 1st sampling- baseline 'control' pre- intervention (PHASE I)

The aim of taking 'control' spot urine before the consumption of soy-milk was to know the amount of phytoestrogens consumed on a regular basis based on their metabolites

concentrations in bio-fluids. It has been reported that the amount of phytoestrogen detected in urine is about 10 – 30 % of the total amount of phytoestrogen consumed (*47, 152, 153*). Thus, the consumption of phytoestrogens can be known and generalised within communities and populations based on their biological sample analysis prior any diet challenge or interventional study. However, this method is 'blind' consumption estimation as it was found from this urine results that 100% recoveries are not achieved in all individuals on each compound due to inter-subject variability in gut metabolism.

In this study, the general profile of the African folks' baseline urine showed that there is little or no consumption of phytoestrogens in the UK based population as shown in FIGURE 3.4 and TABLE 3.4 with low levels of phytoestrogen and their metabolites excretion.

Roughly 20%, 50%, 90%, 40% 50% and 25% of this whole African cohort were found to produce enterodiol, enterolactone, daidzein, equol, genistein and o-DMA respectively on eating their regular diet prior the 'soy challenge' study. Daidzein was found to be present in almost of the subjects but in very small quantities ranging from 12 to 674 pg/µL with an average of 73 pg/ µL in UK cohort and 40 pg/ µL in Nigerian cohort. It can therefore be said that there is at least a low consumption of phytoestrogens within the African folks with the Nigerian cohort showing to consume less phytoestrogen than in London, UK. Interestingly, the presence of equol in some subjects could be due to both the metabolism of phytoestrogens ingested or consumption of cow milk containing casein, which is found to contain a lot of equol (*38, 111*). Comparing the urinary levels of phytoestrogens in spot urine (base-line) of women gave:- 88.1: 44.8 pg/ µL (daidzein), 38.9: 40.7 pg/ µL (genistein), 0.96: 28.6 pg/ µL (o-DMA), 0.26: 465.6 pg/ µL (equol), 34.4: 42.1 pg/ µL (enterodiol) and 462.0: 206:9 pg/ µL (enterolactone) Norfolk study (*317*): present study. These results showed that although daidzein have lower urinary excretion in Africans, the excretion of metabolites (o-DMA and equol) were distinctly higher than the Norfolk female subjects. Thus, there is the possibility of higher ingestion of isoflavones in Africans due to the completion of isoflavone metabolism to equol and o-DMA. On the other hand, the enterolactone results showed Norfolk, UK women consume higher amounts of lignans than Africans.

o-DMA was found at a fairly high concentration in only one subject (NIG cohort) with fair amount of daidzein and genistein but no equol, thus it may be deduced that this subject

had consumed high amounts of phytoestrogens prior the sampling, thus, resulting in the metabolism of isoflavones to o-DMA and not equol.

Merely examining the LCMS profile based on FIGURE 3.5, there are no specific groupings/ clustering of the populations in terms of phytoestrogen levels. This could be due to the fact that all of the subjects used have different/ varying lifestyle and dietary habits. For instance, for the mature students / workers >30 y used in this study, they have probably had more African based food for breakfast and dinner but probably similar lunches with the rest of the other subjects due to the few lunch choices available around. Whilst for the young subjects/volunteers, from the recorded diet dairy and food frequency questionnaire, it showed that they consume very little African-based food.

It was noticed that 50 % of the population used showed to have consumed lignans based on the detection of enterolactone in their urine. However, only 40% of these enterolactone producers were also enterodiol producers.

3.3.7.2 DAY 3: 2nd sampling- During interventional diet (PHASE II)

Soy – phytoestrogens: daidzein and genistein were found to be present in Alpro soy-milk as mainly glycones (>95%), and it was observed that with the consumption of Alpro soy-milk on Day 3 & 5 (during intervention- Phase II & III), free-forms daidzein and genistein (aglycones) were detected in urine. Knowing this, enzymatic hydrolysis was performed on the urine samples in order to complete glycolysis of ingested phytoestrogen glycones with the aim of measuring the accurate amount of 'glycolysed'-daidzein/ genistein only. Phytoestrogens and their metabolites levels were observed to have increased upon the consumption of Alpro soymilk as shown in FIGURE 3.5. The phytoestrogens- daidzein and genistein were found to have increased in all subjects by 0 – 2000 % and 0 – 30000 % respectively. With the quantitation of the soy - phytoestrogen in Alpro soy-milk and urine performed, daidzein was found to be 0.1 – 2.3 ng/µL, 48 – 53 ng/µL in urine and milk respectively. Whilst for genistein, it was found to be 0.10 – 1.6 ng/µL, 53 – 75 ng/µL in urine and milk respectively. Thus, assuming each subject excretes 1.4 L urine daily, the levels of daidzein and genistein recovered in urine was following gut metabolism was calculated to be 0.04 – 0.36% and 0.03 – 0.26 % respectively, i.e. <0.5% phytoestrogen ingested undergo glycolysis - metabolism only.

Numerous previous studies (*13, 46, 130, 148, 150*) have mentioned that 30 - 40 % population can produce equol, which tallies with the finding of this study. Approximately 40% of the whole cohort analysed were found to be significant equol producers producing over 1000 pg/ µL excretion level. Conversely, studies show that 70 % of population are potential o-DMA producers (*135, 292*), which is slightly higher than obtained in this study, as only 55% were o-DMA producers. Intrinsically, ten out 25 equol producer were not o-DMA producers and similarly, eight out of 23 o-DMA producers were not equol producers. Thus, approximately 35 % (15 out of 42) of the representative sampling are good/ high responders to soy phytoestrogen based on their complete metabolism to o-DMA and equol and not either metabolite as reported by other research groups (*145, 292*).

As seen in TABLE 3.5, there were better p- values, i.e. $p < 0.05$ observed in the Phase II (Day 3) samples on the overall due to the generalised effects of soy-milk on most of the subjects. In this phase, following location and gender grouping, daidzein and genistein showed to show no significant difference with p-values < 0.01.

3.3.7.3 DAY 5: 3rd sampling- During interventional diet (PHASE III)

A second during sampling was carried out to ensure the monitoring of phytoestrogens in the body whereby these two phases' samples showed similar patterns. It was noticed that in most of the subjects each followed a regular consumption and sampling timetable based on the slight difference in concentrations, as shown in FIGURE 3.6, however slight variations were observed. The main difference noticed in both second and third samples was that the second sample concentrations showed to be slight higher in most subjects than those of the third samples. The difference could have been due to either the slight disparity in sample collection times or metabolic rate of each individual.

3.3.7.4 DAY 12: 4th sampling - Post diet intervention (PHASE IV)

A forth sample was carried out to observe the metabolic effect of soy after a week of soy-milk consumption. From FIGURE 3.6, it showed that there were still traces of soy metabolism even after a week with higher levels in the post- soy (Phase IV) samples collected on Day 12. Daidzein and genistein averaged urinary excretion levels increased from 75 to 130 and 50 to 100 pg/µL respectively (day 1 to 12/ Phase I to IV). Although, it was reported that the urinary recovery of daidzein and genistein is completed with 24 - 36 h (*335*), in this study, the urinary recovery showed to be longer than 168 h due to the

relatively increased concentrations in comparison to the baseline (Day 1). Further to this, the diet of the subjects were unchanged over this period, thus there is no possibility of more phytoestrogen been introduced into the circulatory system.

3.3.7.5 Factors influencing soy metabolism

As previously reported by researchers (*150, 152, 155*), several factors influence soy metabolism, these include: diet, stress, gut motility, antibiotics; however, three main factors were not considered or mentioned in these studies. The three factors are geographical location, gender and age. In this study, it was found that location and age were major discriminatory factors in soy metabolism. For instance, equol and o-DMA productions were location and gender influenced especially in the UK cohort; whilst age was not a conclusive influencing factor in equol production.

3.4 Conclusions

A simple, selective and specific extraction method using solid-phase extraction (SPE) was developed for the isolation of phytoestrogens and their metabolites applicable to two different bio-fluids: urine and blood serum. The recoveries achieved based on the SPE method ranged from 96 – 100 %. In this study, 42 recruited African subjects (UK- and Nigeria- based; female and male; > 20 y) were successfully assessed for the detection and monitoring of phytoestrogens and their metabolites following a soy-intervention. Alpro soy-milk was the interventional diet, which was analysed to contain approximately 50 ng/µL daidzein and 65 ng/µL genistein.

With urine sample analysis, sensitive, reliable and validated LC-MS methods were developed for the analysis of phytoestrogens and their metabolites. High variances of phytoestrogens and their metabolite urinary excretion levels were observed due to individualism of each subject. It was also found that some of the subjects excreted high levels of equol (> 1000 ng/µL); this was found to be due to the consumption of cow milk. These findings showed and acknowledged that cow-milk made with casein contained a high level of equol (*111, 334*). In addition, results showed that 50 % of the recruited African subjects consume lignan-rich food based on their detection and their metabolites'.

The consumption of one- 200-mL of Alpro soy-milk daily over a period of five days was calculated to contain an equivalent of ~21 mg (total) phytoestrogen. Results showed that

<0.5% of phytoestrogen consumed undergoes glycolysis – metabolism only. Further to this, it was reported that only 10 – 30 % of the total amount of phytoestrogen is excreted in urine, thus 9.5 - 29.5 % of phytoestrogen ingested undergo either further metabolism to o-DMA/ equol and/ or enterohepatic circulation of glucuronide conjugates.

Following the analyses of the samples during-soy consumption (days 3 & 5), results showed that there were correlations of selective factors with the excretion/ production of phytoestrogens (daidzein and genistein) and their metabolites (equol and o-DMA). These selective factors include location and gender.

Also, this study showed that 35 % of the African subjects were equol and o-DMA (combined) producers, i.e. good responders to soy - phytoestrogen. Specifically, 40 % were equol- producers and 55 % were o-DMA producers (both mainly UK- African subjects).

Based on the results of the Day 12 (Phase IV/ post-soy) samples, with the abstinence of phytoestrogen-rich foods, results showed that either the half life of these phytoestrogens are over one week or phytoestrogens are consumed in their regular daily diet.

Location and gender were found to be influencing factors of soy metabolism in the production of equol and o-DMA, in addition, age showed to influence o-DMA production.

With the potential of the well - developed extraction and analysis methods, other metabolites of phytoestrogens could be assessed to understand the proportion distribution of different metabolic reactions in the urinary excretion of phytoestrogens. Also, being a preliminary study of phytoestrogen study in Africans, an epidemiological study of a larger subject size on this population needs to be examined. Complementing this epidemiological study, a well structured and validated food-frequency questionnaire needs to be targeted, as there are currently no available data in the public domain in recent years.

3.5 Experimental

3.5.1 Chemicals, solvents & materials

3.5.1.1 Chemicals/ standards

Formononetin, biochanin A, genistein, daidzein, coumestrol, secoisolariresinol, matairesinol, equol were procured from Sigma Fluka, UK (Gillingham, Dorset, UK). The other metabolites standards including o-desmethylangolensin (o-DMA), enterolactone and enterodiol were purchased from Plantech, UK (Reading, Oxfordshire, UK). The enzyme-β-glucosidase from almonds was purchased from Sigma Fluka, UK (Gillingham, Dorset, UK).

3.5.1.2 Solvents

HPLC-grade methanol, acetonitrile and diethyl ether were purchased from Fisher Scientific Limited, UK (Loughborough, Leicestershire, UK). Distilled water was purified with Millipore Simplicity 185 (Millipore Corp., USA) to 18.2 $m\Omega.cm^{-1}$ polishing units.

3.5.1.3 Materials

Alpro® is a popular brand of soy-milk, which was found to be one of a very few imported soy-milk distributed in Africa. The isoflavones contents ranged from 40 – 45 % daidzein and 50 – 55 % genistein. The difference/ range is due to the flavours. With a serving of one cupful of Alpro® soy-milk, the total isoflavones range is 20 to 25 mg. It can be rationalized that based on the amount of phytoestrogens consumed in one serving over a period of five days, there should be enough metabolites to be produced, absorbed and accumulated in the body. The purpose for the soy-milk consumption was to provide sufficient levels of phytoestrogens to enable metabolic detection and profiling.

3.5.2 Instrumentation

Polymeric Strata-X™ reversed phase ODS 30 mg/ 1 mL sorbents cartridges were purchased from Phenomenex (U.K.). Sterile pipette tips (0.02 mL, 0.2 mL, 0.5 mL and 1 mL and 5 mL) and centrifuge tubes were obtained from Fisher Scientific (U.K.).

Eppendorf Centrifuge 5804R (Germany) was used for sample preparation in liquid- liquid extraction (LLE). *HERA* cell 150 incubator (U.K.) was used to incubate the bio-fluids overnight at 37ºC.

Waters® Micromass Quadrupole Time of Flight Micro™ was linked to a LC packing HPLC with the following operating conditions.

LC operating conditions: The automated LC system (LC packing, UK) comprised of a Famos autosampler connected to Ultimate™ plus HPLC. Chromatographic separation was achieved on Phenomenex column- Luna C8(2)- 3 µm, 150 x 2.0 mm on a gradient elution program of mobile phase composed of 35 % methanol (Method A) or acetonitrile (Method B) in water (0.1 % formic acid) at a flow rate of 0.15 mL/min. Either LC experiments were run isocratically under 20 minutes.

MS operating conditions: The structural identification of compounds was achieved by Waters Micromass Quadrupole Time of Flight Micro™ analyzer operated in the negative ion mode electrospray ionisation under the following conditions: Capillary voltage of 3500 V; sample cone voltage and extraction cone maintained at 30 V and 3 V respectively; source temperature was set at 150ºC; both ion and collision energies at 3 V. For LC-MS-MS, i.e. fragmentation, the sample cone voltage and collision energy were increased to 60 V and 50 V respectively. Data acquisition was done using MassLynx 4.0 Global Mass-Informatics software.

3.5.3 Design of experiment

3.5.3.1 Ethical approval

Ethical approval for this research was reviewed and approved by the London Metropolitan University ethics committee and all procedures complied with National Health and Medical Research Council standards.

As compliant to the ethics, all subjects were interviewed and signed informed consents were obtained from all participants before commencement of the study and all participant samples and information were confidential and remained completely anonymous.

Apart from the ethics discussed with each of the subjects, a brief talk was given about the study, filling of forms and also personal information were requested: these include their contact details – email and telephone numbers, age range: 20 – 25 y, 25 – 30 y, 30 – 35 y and 35 – 40 y: generalised as <30 and > 30 y, dietary group - vegetarian (no animal meat products), vegan (no dairy or animal meat), omnivores (consuming animal and plant products), or macrobiotic (no dairy, meat or processed food including flour) and body mass index (BMI) measurement whereby their weight (kg) and height (m) were taken for the NIG cohort. All of these subjects' personal information was stored also stored in a password-protected document.

3.5.3.2 Subjects recruitment

Africans were targeted to facilitate recruitment of the individuals for this diet interventional study. Two cohorts were designed for this interventional study- UK and NIG cohorts.

London, UK: Twenty-five healthy volunteers aged between 20 and 40 y were selected for a 2-week diet interventional study. Twenty-two of these volunteers were of African origin, two were South Asian and one was East Asian. Of the total twenty-five, thirteen were male volunteers and the remaining twelve were female volunteers. 80% of these volunteers were London Metropolitan University North campus students/staff and the remaining were workers around London City. There were no criteria for the selection of the volunteers based on their dietary habit, these include vegetarian (no animal meat products), vegan (no dairy or animal meat), omnivores (consuming animal and plant products), and macrobiotic (no dairy, meat or processed food including flour) diet groups were used in this study.

Lagos, Nigeria: Twenty-five healthy volunteers were selected for a 2-week diet interventional study. Twenty-three were aged between 20 and 40 y and the remaining two were > 65 y (retired citizens). All of these volunteers were of African origin. Of the total twenty-five, thirteen were male volunteers and the remaining twelve were female volunteers. 90% of these volunteers were University of Lagos campus students and the remaining were workers/retirees around the city of Lagos. All the subjects were found to be omnivores (consuming animal and plant products).

3.5.3.3. Study preparations

Intervention preparations and protocol

A cupful of *Alpro*® soy-milk equivalent to 150 - 200 mL was consumed each day for five days (Monday to Friday). During the period, subjects were allowed to continue their regular diets. Due to the possibility of variation in each person's diet, a week food dairy was kept during this diet intervention week. However, the subjects were advised to avoid soy products during the following week before the collection of the final urine sample.

Each subjects pack contained an information sheet, consent form, candidate information, food frequency questionnaire based on African staple foods and soy-milk for the primary purpose of estimating their weekly consumption, 5-day food dairy to establish whether there were any significant differences that could affect the study outcome, four plastic cups, four sterile urine sample collection tubes, strips of para-film and one small sample bag for containing each of the filled tubes to avoid spillage (risk assessment) – see appendix. All tubes were individually labelled with random figure numbers before distribution to the participants for the purposes of anonymity.

London, UK: This study required the participants to make four visits to the Research laboratory at the New Science Centre, London Metropolitan University – Monday, Wednesday, Friday and the following Friday. At the first visit, subjects were invited to the interview/seminar room, briefed about the study, interviewed and requested to provide spot urine for the baseline measurement of phytoestrogens. On this day, subjects were given above listed items (in subject pack) and sent home with a continuous 5-day supply of Alpro soy-milk. Afterwards on the remaining three days, urine samples are collected and stored.

Lagos, Nigeria: This study required the participants to make six visits to the Department of Chemistry, Faculty of Science, University of Lagos. On the first day, subjects were briefed about the study, interviewed, requested to fill forms, provide spot urines and consume a cupful Alpro® soy-milk (~200 mL). In the five drinking visits, evening consumption was followed, most especially on the days prior to early morning urine collection – Tuesday and Thursday, all subjects were invited for a late evening drinking session due to the fact that the time span of ingested phytoestrogens vary from 4 to 13 h. Also, the subjects were asked to fill their food dairy form as explicitly detailed as possible.

3.5.3.4. Sample collection

Prior to starting the study, a control "blank" spot urine sample was taken on the first day of interventional study (before the consumption of soy-milk) for analysis as the baseline urine. During this diet study, at the middle of the week (Wednesday) and last day (Friday) of the soy-milk consumption early morning urine samples were collected. A week after the diet study (Friday), another spot urine sample was taken for analysis. All the urine samples collected were stored by freezing at -20 °C prior batch processing: extraction and analysis. In the case of the Nigerian study, the frozen samples were shipped to UK in cold conditions.

3.5.4 Methodolgy

3.5.4.1 Extraction of phytoestrogens in Alpro® soymilk

0.5 mg β-glucosidase from almonds (~4.7 U) was accurately weighed into a 1.5 mL eppendorf tube and 0.5 mL soy-milk was added. This solution was vortex for 2 minutes incubated at 37 °C overnight. Following this, to extract the aglycones, 0.5 mL diethyl ether was added to the hydrolysed solution and centrifuged for 2 minutes at 2000 g. The ether layer (supernatant) was collected and the step was repeated twice with the (three) ether fractions combined and completely dried under nitrogen gas prior storage at -20 °C for further analysis. This entire process was repeated at least thrice for each random selection of Alpro soy-milk used for this study.

3.5.4.2 Extraction of conjugated phytoestrogens in urine samples

Frozen human urine samples were thawed at room temperature. 0.5 mL urine was combined with 0.1 mL of a β-glucosidase solution (32 U/mL) in 0.1 M sodium acetate buffer (pH 5), vortexed for 2 minutes and incubated at 37 °C overnight. The resulting enzyme-treated urine was subjected to SPE.

Solid phase Extraction (SPE) – Enzyme-treated urine (hydrolysed phytoestrogens): Purification of enzyme-treated urine was carried out by SPE on strata-X™ reversed phase column cartridges. Before SPE was carried out, the cartridges were conditioned with 600 µL of 10 % methanol and 600 µL of water. Following this, 600 µL of the hydrolysed urine sample which was incubated overnight at 37 °C with β-glycosidase was then loaded onto

the cartridge and was then washed with 600 µL of 10 % methanol. The cartridge was left to dry for one minute before the elution of aglycones in 650 µL acetonitrile / methanol (1:1), which was collected. The collected eluent was dried under nitrogen and then reconstituted in 0.5 mL methanol.

3.5.4.3 Recovery- extraction of unconjugated phytoestrogens

A recovery experiment was also carried out using 50, 2000 and 5000 ng each of the standards: isoflavones (daidzein and genistein) and lignans (secoisolariciresinol and matairesinol) and the metabolites (enterodiol, equol, enterolactone and O-DMA). The standard solution was subjected to SPE (all three spiked solutions) and LLE (5000 ng only) as described below. The recovery was carried out in batches alongside the actual urine sample clean - up.

Liquid - liquid extraction (LLE) – Spiked Blank (conjugated phytoestrogens): To the standard-spiked blank (recovery experiment), 0.5 mL of diethyl ether was added, centrifuged for 2 minutes at 2000 g and the supernatant ether layer collected. This was repeated twice and the collected fractions combined and completely dried off using nitrogen gas. Subsequently, a blank was also carried out using Millipore purified water.

3.5.4.4 Method development using LC-MS

Prior to sample analyses, a mixture of the twelve standards (biochanin A, formononetin, daidzein, genistein, coumsterol, secoisolariciresinol, matairesinol, enterodiol, equol, enterolactone, p-ethylphenol and o-DMA) in methanol was used for the development of a good separation on the Waters Micromass qTOF LCMS). Following the development of good LCMS chromatographic separation methods after several methods were tried. The final LC methods (Methods A & B) was achieved with the greatest sensitivity, specificity and a good and distinct separation of all the standard compounds, i.e. resolution.

The LC-MS system was validated for the accuracy and reliability of generated data. Calibration curves for the standards and metabolites were constructed relating the concentration of each of the analysed samples which were determined and compared with the analysed samples in the calibration.

3.5.4.5. Sample analysis on LC-MS-SIM

Phytoestrogens in the urine were separated chromatographically using a Phenomenex, Prodigy 150 x 4.6 mm internal diameter C8(2) reversed phase column which was pre equilibrated using a gradient elution system made up of methanol or acetonitrile in water before being analysed by LC-MS-SIM. Also, in order to increase the sensitivity of detection LC-MS-SIM was used for the calibration of the isoflavones and their metabolites. Only ions of selected mass-charge ratio were monitored in this scan mode. SIM for ions $[M-H]^-$ of (*m/z*: 283) biochanin A, (*m/z*: 252) formononetin, (*m/z*: 267) coumestrol, (*m/z*: 257) o-DMA, (*m/z*: 269) genistein, (*m/z*: 241) equol, (*m/z*: 253) daidzein, (*m/z*: 297) enterolactone, (*m/z*: 301) enterodiol, (*m/z*: 357) matairesinol and (*m/z*: 361) secoisolariciresinol were used in this study.

Chapter IV

^{1}H-NMR-based metabonomics of soy - phytoestrogens.

4.1 Introduction

Metabonomics have been generally applicable in toxicology and in drug discovery; however, its potential in nutrition (both in human and animal) has just only been explored in recent years. Various researches involving induced diet metabonomics are now studied with the aim of understanding the biochemical effects of major components in these diets, i.e. identifying their bio-functionality by simulating *in-vivo* set-ups. Metabonomics can be professed to be better than conventional *in-vivo* studies because it is subject-specific (not-generalised) and thus gives better insights into the integrated function of individual complex bio-system at a system level (*336*). Metabonomics have be used to help clarify Inconsistencies of the biological effects of certain compounds in diets have been based on *in-vivo* and *in-vitro* studies. So far, little progress has been achieved in its application to the diet and nutrition, especially thorough understanding of metabolic effects of vital bioactive components in the diets. With the introduction of an external stimulus (e.g. diet) monitoring metabolic processes based on simultaneous analysis of hundreds of low molecular weight metabolites has been facilitated with the aid of effective and non-destructive high resolution ^{1}H-NMR on biological samples and bio-fluids (*337, 338*). In fact, the diets studied to date in metabonomics, *as earlier mentioned in Chapter I*, include: fruits and vegetables, chamomile tea, garlic, soy products, grape juice, olive oil, nuts and whole grains.

Soy-milk prepared from soybeans is known to contain mainly polyphenolic isoflavones: daidzein, genistein, glycitein, daidzin, genistin and glycitin. This group of compounds has been found to possess significantly high estrogenic activities compared to estradiol, a human estrogen, due to their similarity in structures (*1*). In the last five years numerous studies on the health benefits of phytoestrogens in combating hormone dependent diseases (*1-4*), bone health diseases such as osteoporosis (*5*), cardiovascular diseases (*4, 5*), and hypercholesterolemia (*4, 5*) have been reported. Most of these studies and results have been based on epidemiological and *in vitro* simulative studies, for instance, using cell lines. In relation to the growing interest on dietary phytoestrogens, with

isoflavones being the important chemical component found abundantly in most soy products, little or no work has been reported on their biochemical effects based on *'stimulated in-vivo'* study, i.e. metabonomics. The only related studies were reported by Solanky's research group (*230, 231*) on distinguishing between biochemical effects following a textured vegetable protein or miso intervention. The extent to which soy phytoestrogens affect the human system has not been fully investigated. Additionally, it was noted that this study was not explicatory because only six subjects are involved and only a few particular endogenous metabolites (< 10) were investigated, yet the human bio-fluid has been known to constitute over 300 of these metabolites.

Based on the potential of assessing these numerous endogenous metabolites in urine and with the little information explored by Solanky et al (*230, 231*), this instigated the investigation of soy phytoestrogens in ^{1}H-NMR based metabonomics using larger subjects' cohort and number of metabolites. Coherently, the whole metabolic effects of phytoestrogens were deemed complete by assessing at least three key Phases of intervention: pre-, during and post consumption. Thus, this led to the qualitative and quantitative study of >20 targeted endogenous metabolites found to be common in the urine samples as well as the qualitative study of the entire ^{1}H-NMR spectrum region for pattern recognition.

Phytoestrogens and their metabolites have not been investigated in ^{1}H-NMR mainly due to the poor solubility in aqueous solution and poor ^{1}H-NMR sensitivity. With the aid of computer-aided softwares- SIMCA-P+® and Chenomx®, statistical pattern recognition techniques (qualitative analysis) and quantitation of 22 endogenous metabolites (shown in FIGURE 4.1) respectively were achieved.

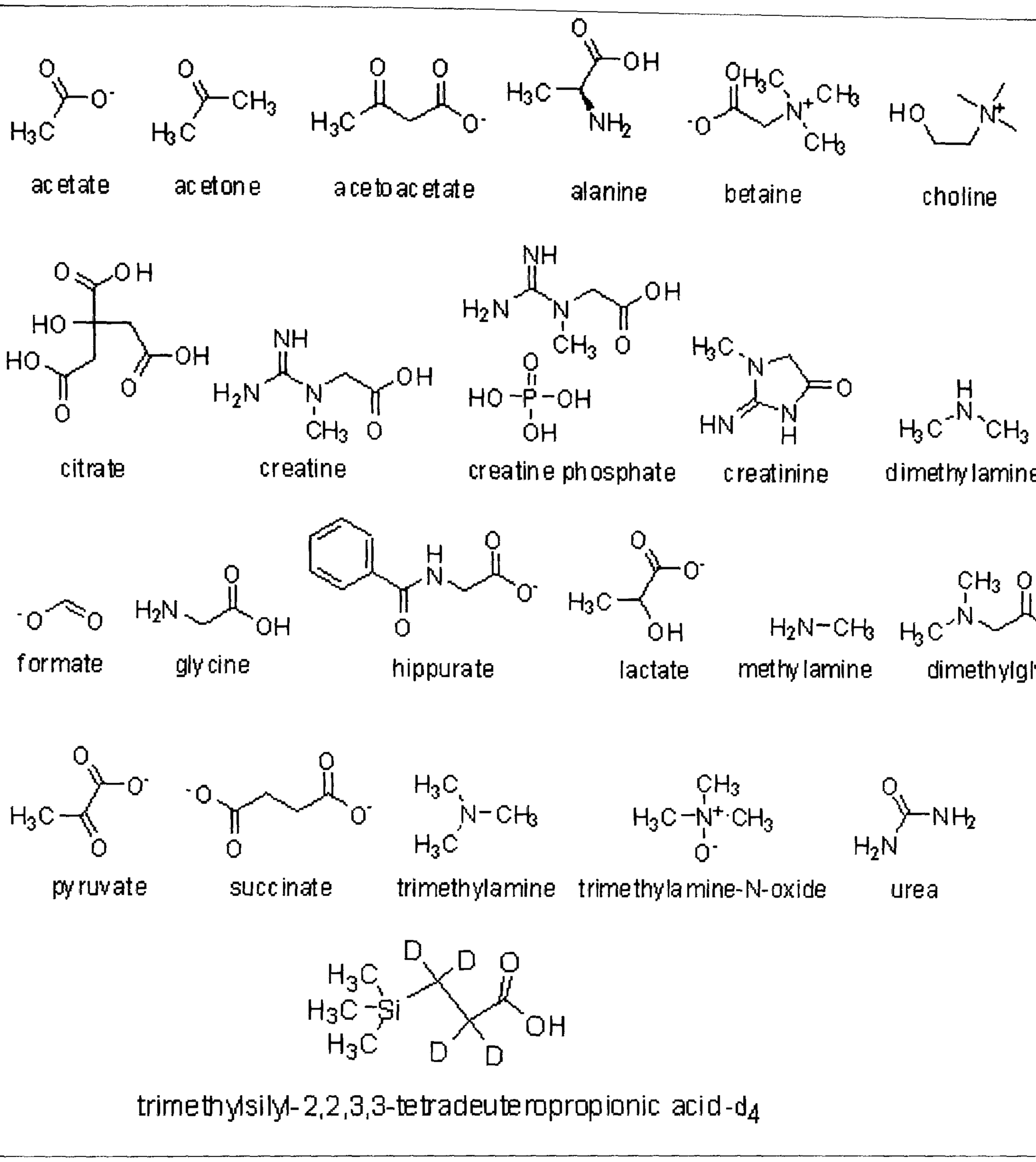


FIGURE 4.1: Structures of the endogenous metabolites analysed in this study and reference compound (TSP-d_4)

More metabonomics investigations are needed for the understanding of consumption, metabolism and biochemical effects of soy - phytoestrogens. However, no report has been published on the study of the least studied human population: African - population. In particular, the consumption and biochemical effects of soy -products within the Africans is unknown and worthwhile investigating. Furthermore, based on the consequential analysis and results found in chapter III, the use of these subject groups claimed to be exploratory as a pilot test for metabonomics study. With this aim, quantitation and multivariate analysis (MVA) were performed. Two types of SIMCA-P+® aided-MVA analyses were conducted in this study: unsupervised and supervised MVA. The only

unsupervised MVA tool is Principal Component Analysis (PCA); whilst the supervised MVA tools include: Partial Least Squares- Discriminant Analysis (PLS-DA) and Orthogonal- Partial Least Squares- Discriminant Analysis (O-PLS-DA). Following this, three main factors were investigated for pattern recognition based on supervised multivariate analysis (MVA)- O-PLS-DA. These include: location, gender and age.

4.2 Results

4.2.1 General changes in spectra following intervention

A typical ^{1}H-NMR spectrum of the African urine sample characterised with all 22 analysed endogenous metabolites (excluding urea) is shown in FIGURE 4.2 a-c. Generally, acetone (2) and creatinine (10) gave the most abundant peaks.

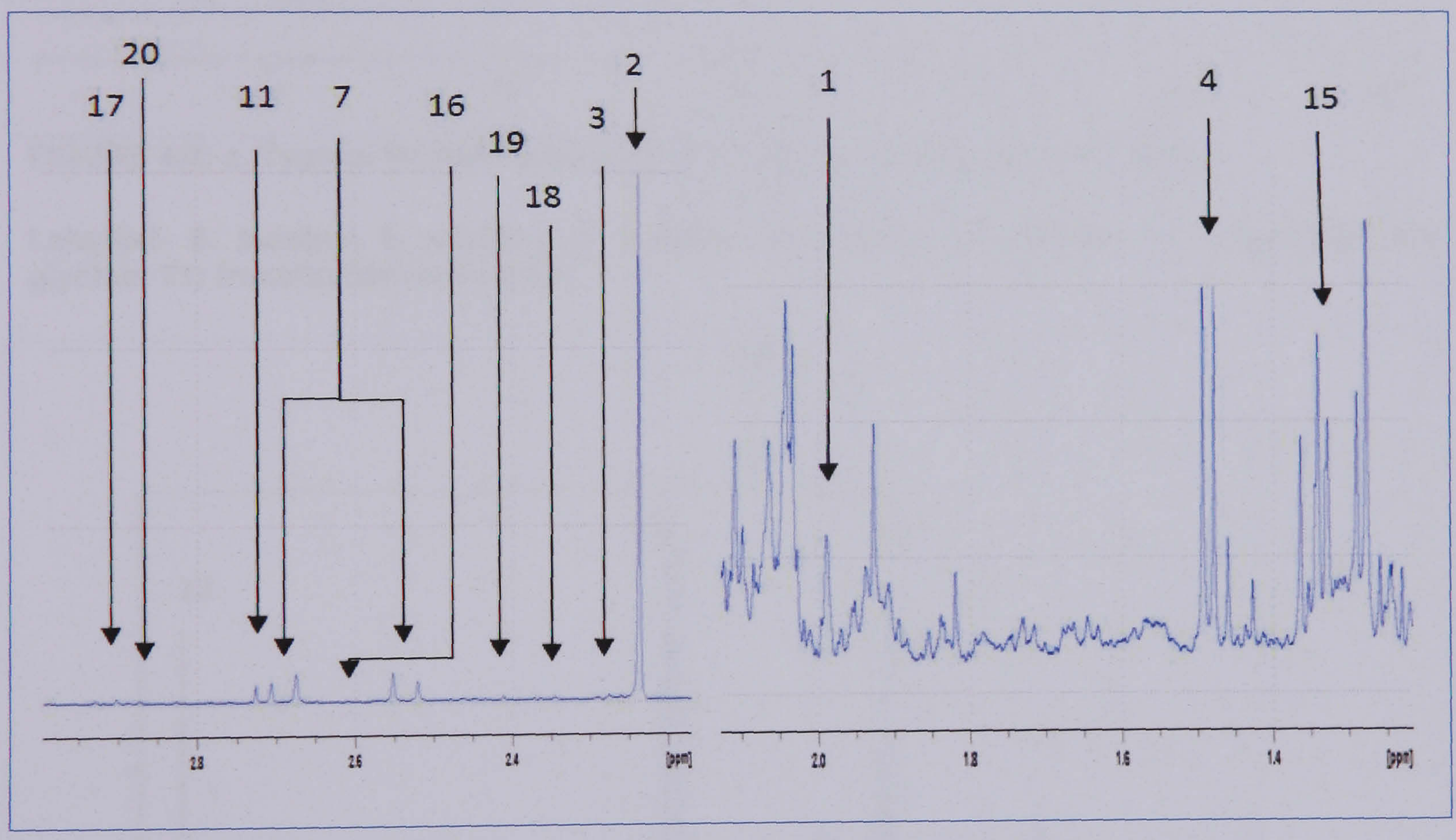


FIGURE 4.2: a. Typical 1H-NMR spectrum (δ: 1.2- 3) region of a subject's urine.

Labelled- 1: acetate; 2: acetone; 3: aceto-acetate; 4: alanine; 7: citrate, 11: dimethylamine, 15: lactate; 16: methylamine; 17: dimethylglycine; 18: pyruvate; 19: succinate; 20: trimethylamine

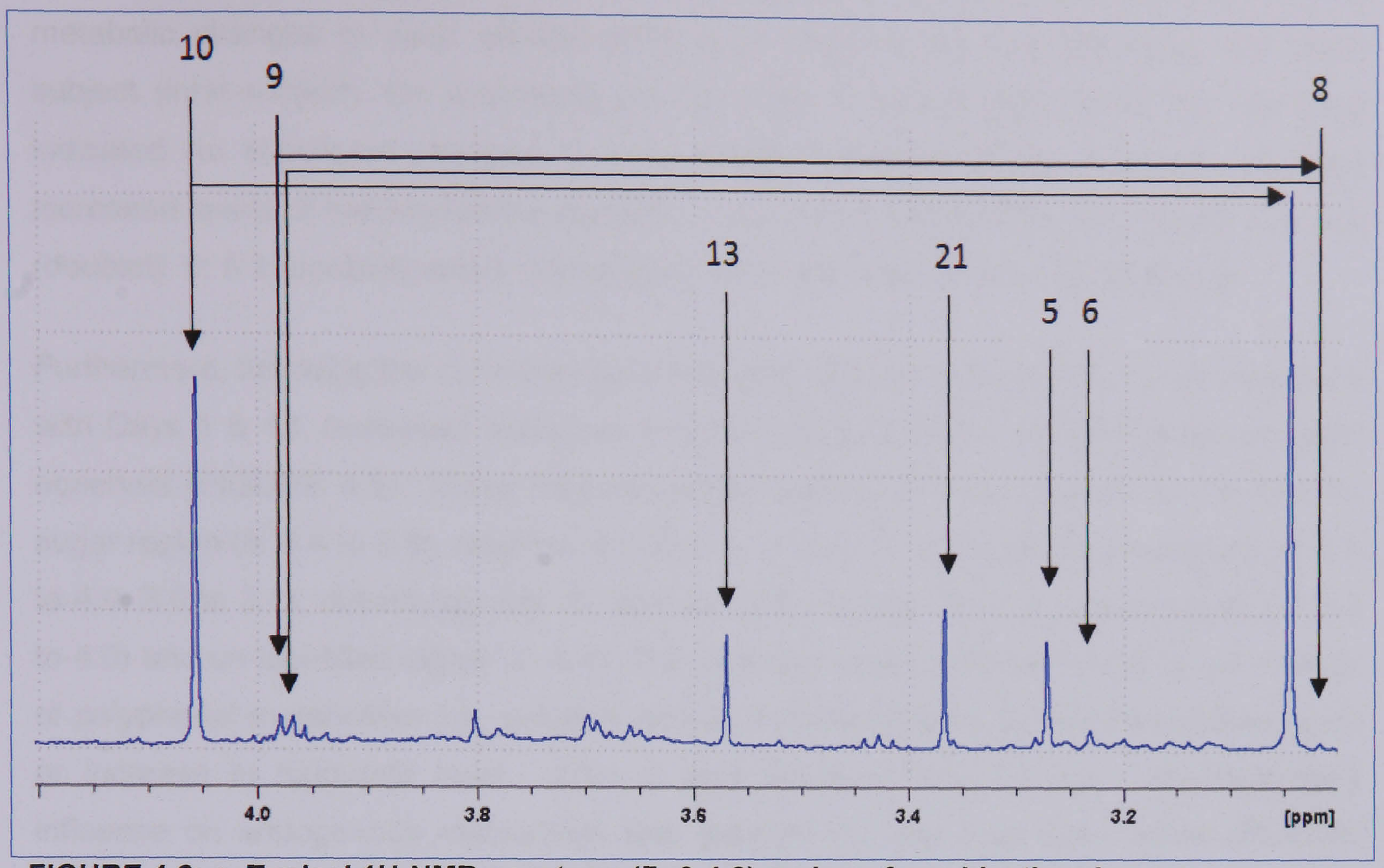


FIGURE 4.2: a. Typical 1H-NMR spectrum (δ: 3-4.2) region of a subject's urine.

Labelled- 5: betaine; 6: choline; 8: creatine; 9: creatine phosphate; 10: creatinine; 13: glycine; 21: trimethylamine-N-oxide

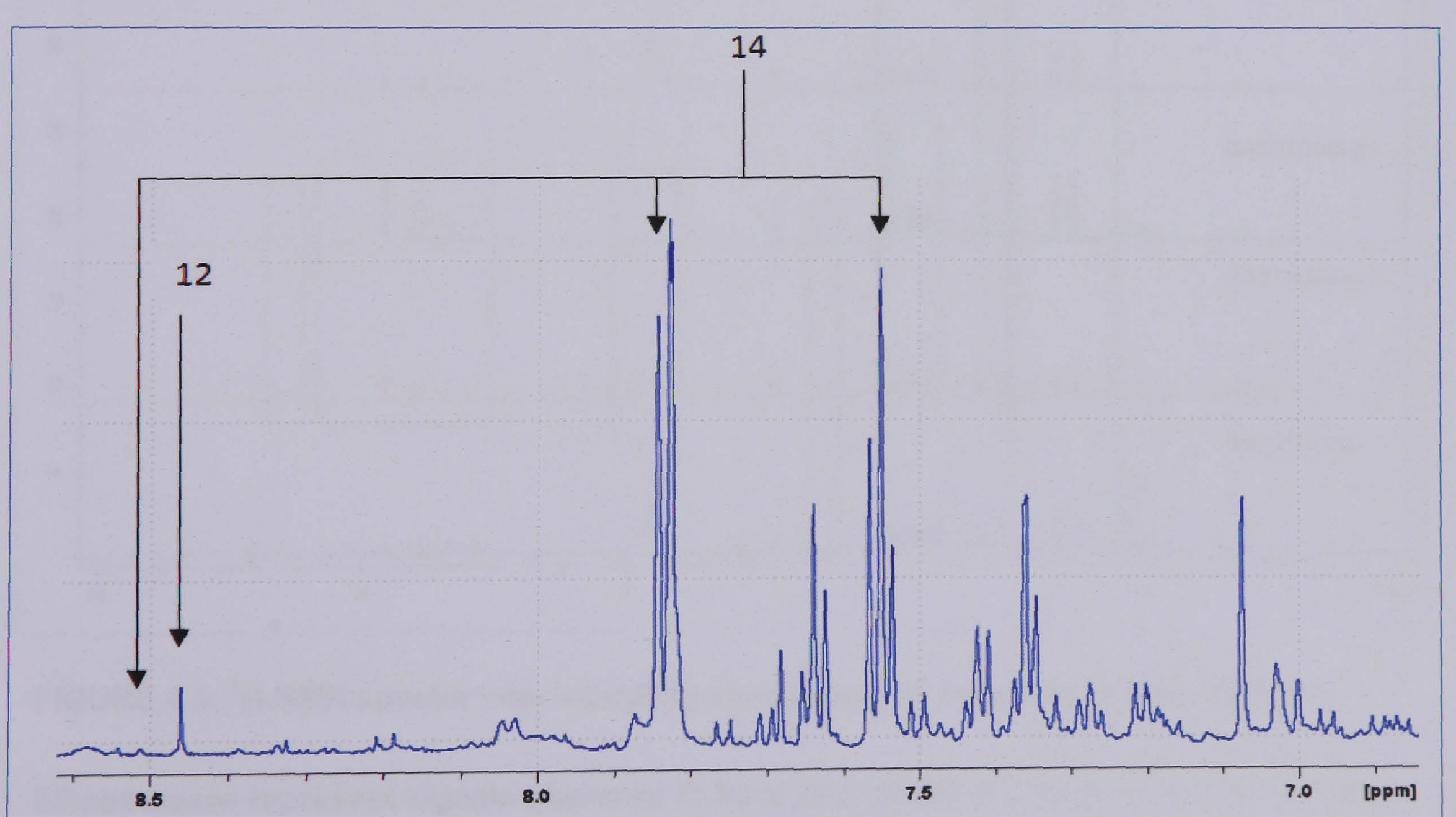


FIGURE 4.2: b. Typical 1H-NMR spectrum (δ: 6.8 - 8.7) region of an African subject's urine.

Labelled- 12: formate; 14: hippurate.

From visual assessment of all the ^{1}H-NMR spectra, it was shown that there were metabolic changes in each sample (phase) of intervention (inter-sampling) and each subject (inter-subject). On examining pre-soy (Day 1) and post-soy (Day 12), samples indicated no significant changes in their overall metabolic profile, however, (slightly) increased levels of trimethylamine signal (δ: ~2.2) and three un-identified signals at δ: 4.4 (doublet), δ: 5.3 (doublet) and δ: 8.9 (singlet) were observed on Day 12 (post-soy).

Furthermore, following the consumption of soy-milk (shown on Days 3 & 5) in comparison with Days 1 & 12, increased variances in certain regions of the ^{1}H-NMR spectrum were observed (FIGURE 4.5). These regions include: certain aromatic region (δ: 7.4 to 7.8), sugar region (δ: 3.4 to 3.8), creatine (δ: 4.0 to 4.1; 3.0 to 3.1), creatinine phosphate (δ: 3.9 to 4.0; 3.0 to 3.1), dimethylglycine (δ: 3.25 to 3.35), formate (δ: 8.4), sugar region (δ: 3.5 to 4.0) and un-identified signal (δ: 4.4). The aromatic region changes could be as a result of polyphenol metabolism (i.e. possibly traces of phytoestrogen & their metabolites) and/ or increase in hippurate level. Also, it was observed that the soy - phytoestrogen influence on endogenous metabolites was greatest on Day 3 as seen below (FIGURE 4.3).

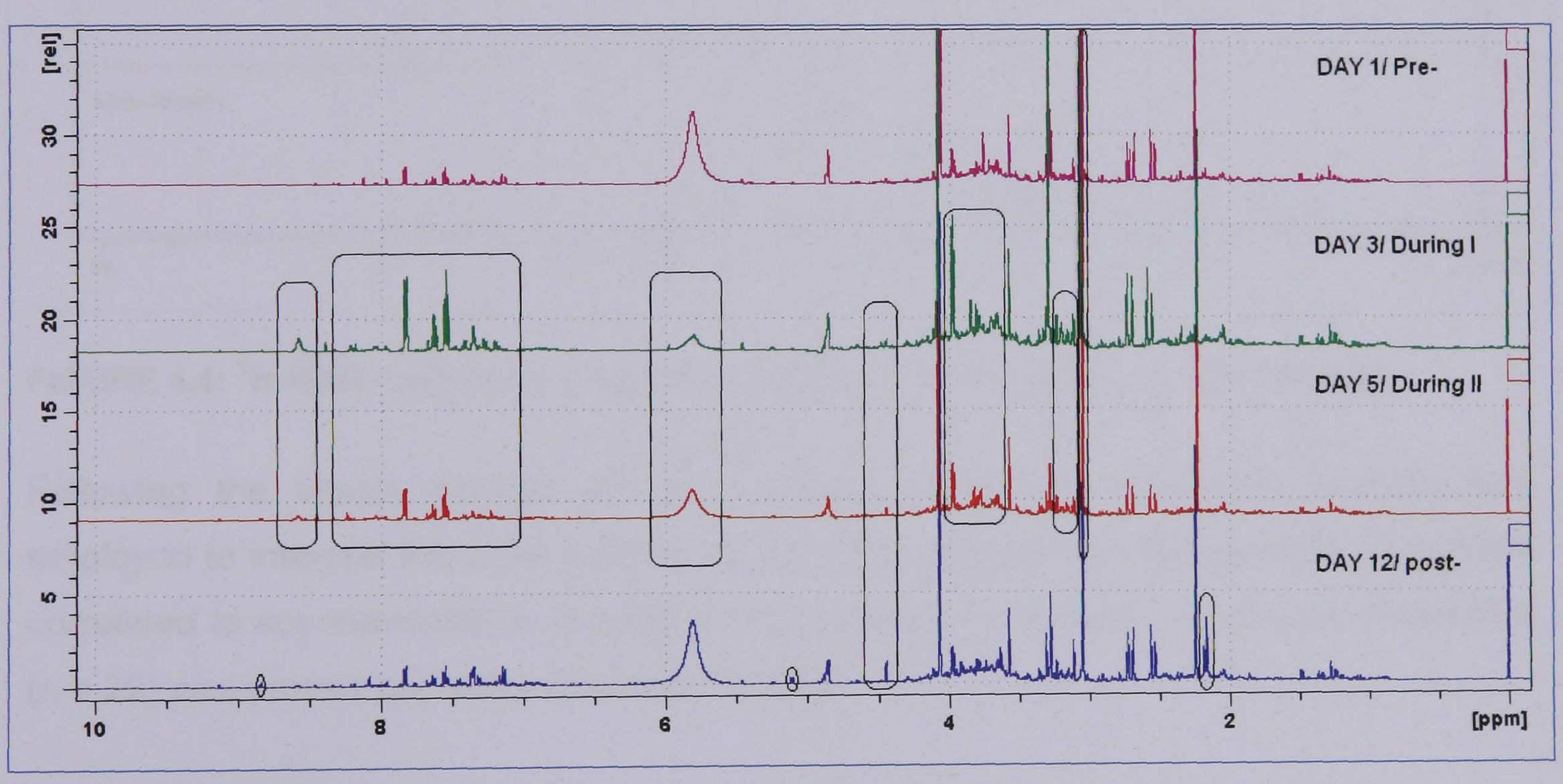


FIGURE 4.3: ^{1}H-NMR spectra inter-sampling comparison of metabolic profile changes.

Ellipse-boxes represent signals observed to have changed (↑ or ↓) during the four Phases.

On the other hand with respect to the different subjects' spectra, there were noticeable metabolic profile changes in inter-subjects as represented in FIGURE 4.4. Comparing the

^{1}H-NMR spectra of subjects based on their different locations and genders showed noticeable changes in certain regions- aromatic (δ: 6.8 to 8.5), sugar (δ: 3 to 4) and selected aliphatic (δ: 2.4 to 3.0) regions. Conversely, insignificant or no changes were observed in (δ: 0.8 to 1.6), and the urea peak had no specific pattern in either location or gender groups.

Generally, the female subjects showed to have more metabolites than the male subjects, and the Nigeria subjects showed the presence of more metabolites than the UK subjects. Thus, these could be as a result of differences in the demographical regular African diets metabolism based on their microflora system and consequently gender.

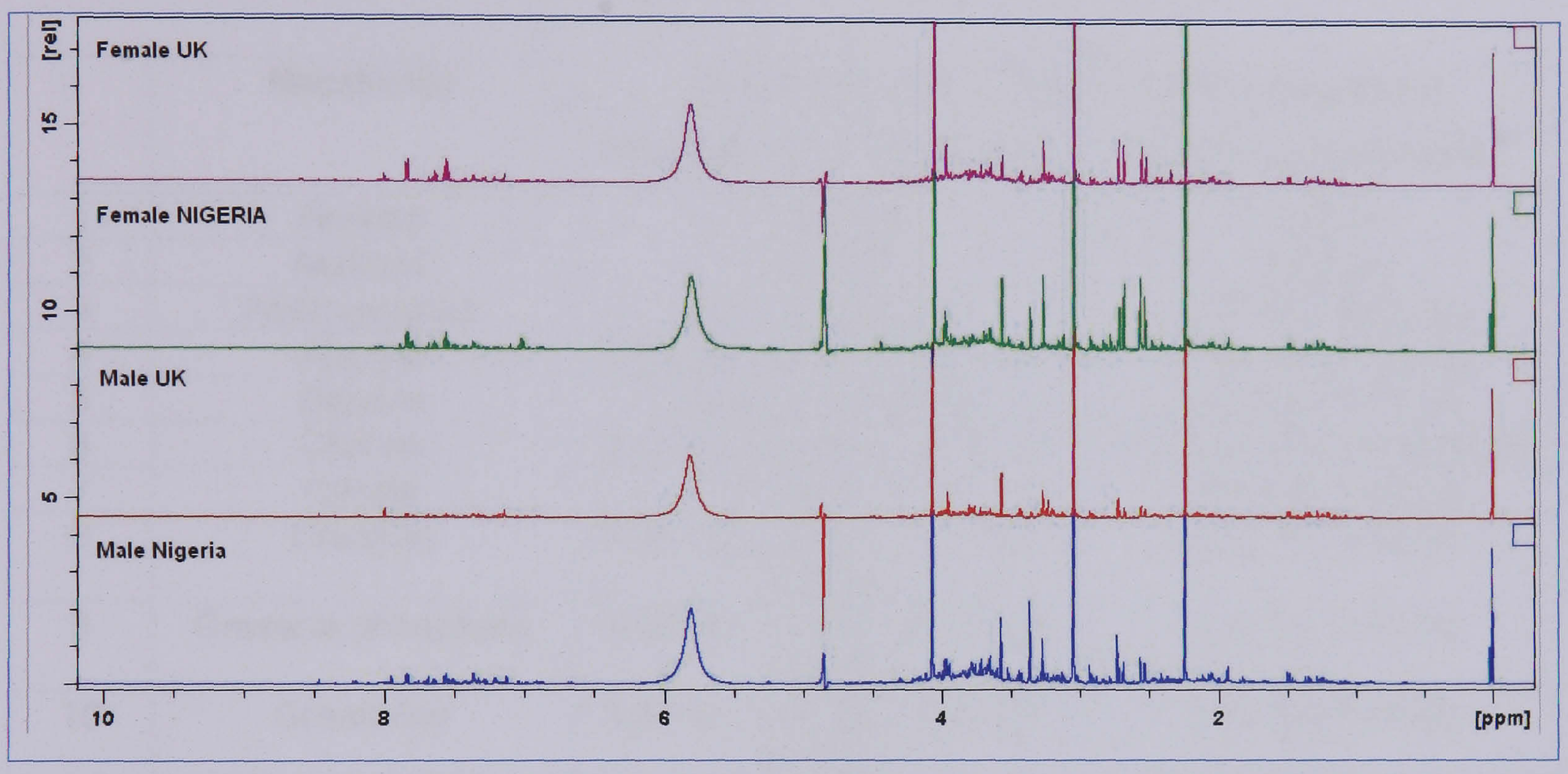


FIGURE 4.4: ^{1}H-NMR spectra inter-subject comparison of metabolic profile changes.

Following the visual analyses of the collected ^{1}H-NMR, multivariate analysis was employed to interpret the extensively bulky spectral data and identify potential biomarkers correlated to soy intervention. In addition, quantitation of all earlier mentioned metabolite (n = 22) was carried out using Chenomx® Profiler.

4.2.2 Quantitation of endogenous metabolites

With the aid of ChemDraw® Ultra 10.0 and incorporated metabolites library in Chenomx®, TABLE 4.1 gives the chemical shifts and splitting pattern of those metabolites used for quantitation in this study. Comparing the chemical shifts predicted by ChemDraw® (FIGURE 4.5) with the actual by Chenomx®, it was noticed that there were significant

chemical shift differences due to the ChemDraw® calculated data, which did not include pH/ ionic strength conditions in the actual measurements. In addition, as noted in TABLE 4.1, some of the chemical shifts were not used for quantitation due to the prediction falsity or insignificance of the peak (i.e. signal intensity).

As shown with the pattern, most of these metabolites have simple splitting patterns- singlet and / or doublet, except alanine and lactate (quartet), choline (triplet) and hippurate (multiplet). The simplicity of these multiplicities of metabolite signals made it easy for manual identification and quantitation, thus reducing possibilities of human error.

TABLE 4.1: Annotation of 1H-NMR signal chemical shift and multiplicity

	Metabolite	Chemical shift (δ) / ppm (splitting pattern)	
		Predicted by ChemDraw®	Actual by Chenomx ®
1	Acetate	2.20 (s)	1.95 (s)
2	Acetone	2.09 (s)	2.22 (s)
3	Aceto-acetate	2.25 (s), 3.55 (s)	2.27 (s), 3.43 (s)
4	Alanine	1.23 (d), 3.67 (q), 8.81 (b)*	1.45 (d), 3. 70 (q)
5	Betaine	3.30 (s), 4.35 (s)	3.25 (s), 3.87 (s)
6	Choline	3.30 (s), 3.43 (t)*, 3.97 (t)*	3.19 (s), 3.50 (t), 4.05 (t)
7	Citrate	2.61 (s)	2.54 (d), 2.68 (d)
8	Creatine	3.04 (s), 3.96 (s), 6.63 (b)*, 7.84(b)*	3.02 (s), 3.92 (s)
9	Creatine phosphate	3.04 (s), 3.96 (s), 6.63 (b)*, 7.84 (b)*	3.02 (s), 3.93 (s)
10	Creatinine	3.04 (s), 3.47 (s), 7.84 (b)*, 8.00 (b)*	3.06 (s), 4.08 (s)
11	Dimethylamine	2.00 (b)*, 3.26 (s)	2.71 (s)
12	Formate	9.60 (s)	8.44 (s)
13	Glycine	4.24 (s), 8.70 (b)*	3.55 (s)
14	Hippurate	4.31 (s), 7.63 (m), 7.70 (d), 8.03 (b)	3.95 (s), 7.55 (m), 7.62 (m), 7.83 (d), 8.51 (b)
15	Lactate	1.28 (d), 4.14 (b), 4.26 (qd)	1.31 (d), 4.10 (b)
16	Methylamine	2.00 (b)*, 2.47 (s)	2.59 (s)
17	Dimethylglycine	2.75 (s), 3.30 (s)	2.91 (s), 3.71 (s)
18	Pyruvate	2.17 (s)	2.36 (s)
19	Succinate	2.73 (s)	2.41 (s)
20	Trimethylamine	2.26 (s)	2.87 (s)
21	Trimethylamine-N-oxide	2.90 (s)	3.31 (s)
22	Urea	6.00 (b)	5.76 (b)

Splitting Pattern- s: singlet; d: doublet; t: triplet; q: quartet, m: multiplet, b: broad. * ignored chemical shifts predicted by ChemDraw®, i.e. not used in quantitation

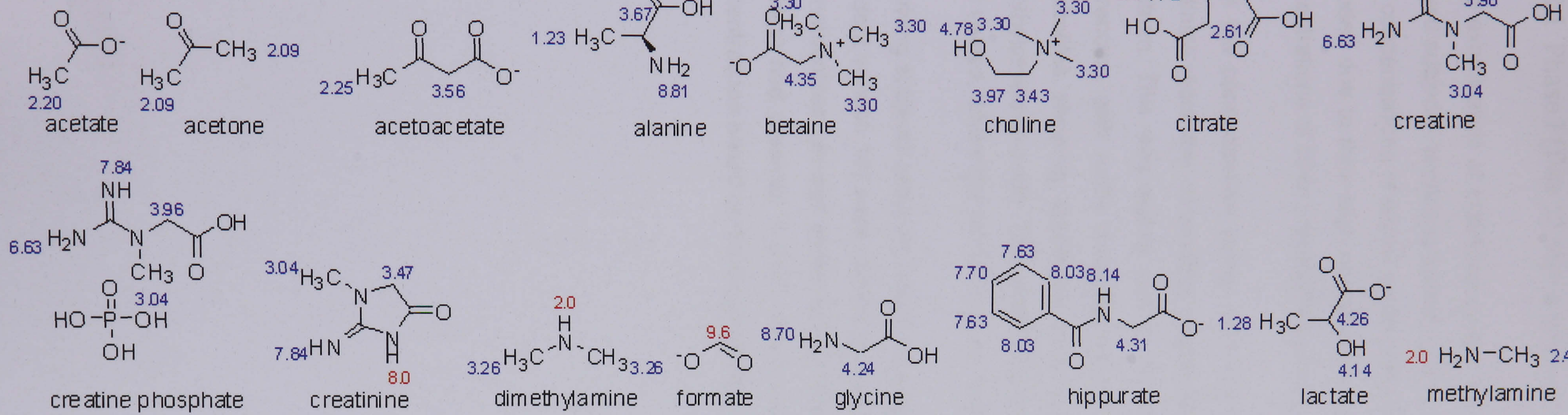


FIGURE 4.5: Annotations of chemical shift (predicted) in endogenous metabolites

4.2.2.1 Phase I (Day 1) soy intervention: Pre-Soy/ Control

Close examination of individual control samples collected before the start of the intervention showed subjects' variances based on the identified and quantified metabolites. Following the high concentrations of some metabolites (based on normality test), creatinine and urea were excluded due to their high concentration and variance, thus the variance in smaller valued concentrations of other metabolites were pronounced (FIGURE 4.6).

From the concentration values and the plot in FIGURE 4.6, it was found that urea and creatinine were the dominating metabolites showing high variances (S.D) in the ^{1}H-NMR spectrum. This was mainly due to their high excretion levels as well as the inter-subject differences- with some subjects excreting high levels and others at low levels. Other metabolites showing significant variances were acetone, citrate, glycine, hippurate and trimethylamine-N-oxide. This was further confirmed with the average and standard deviation (SD) values, where the SD ≥ 0.9 of average (shown in TABLE 4.2).

Following subjects grouping based on location and gender, one-sample p-tests using Minitab (Student version 10) were carried out to test significant differences between concentrations. All metabolites (except acetone, hippurate and TMAO in UK females; acetone in Nigeria females) had p-values < 0.01, which implied no significant differences in the metabolites concentrations based on these groupings- location and gender.

PAGE

MISSING

IN

ORIGINAL

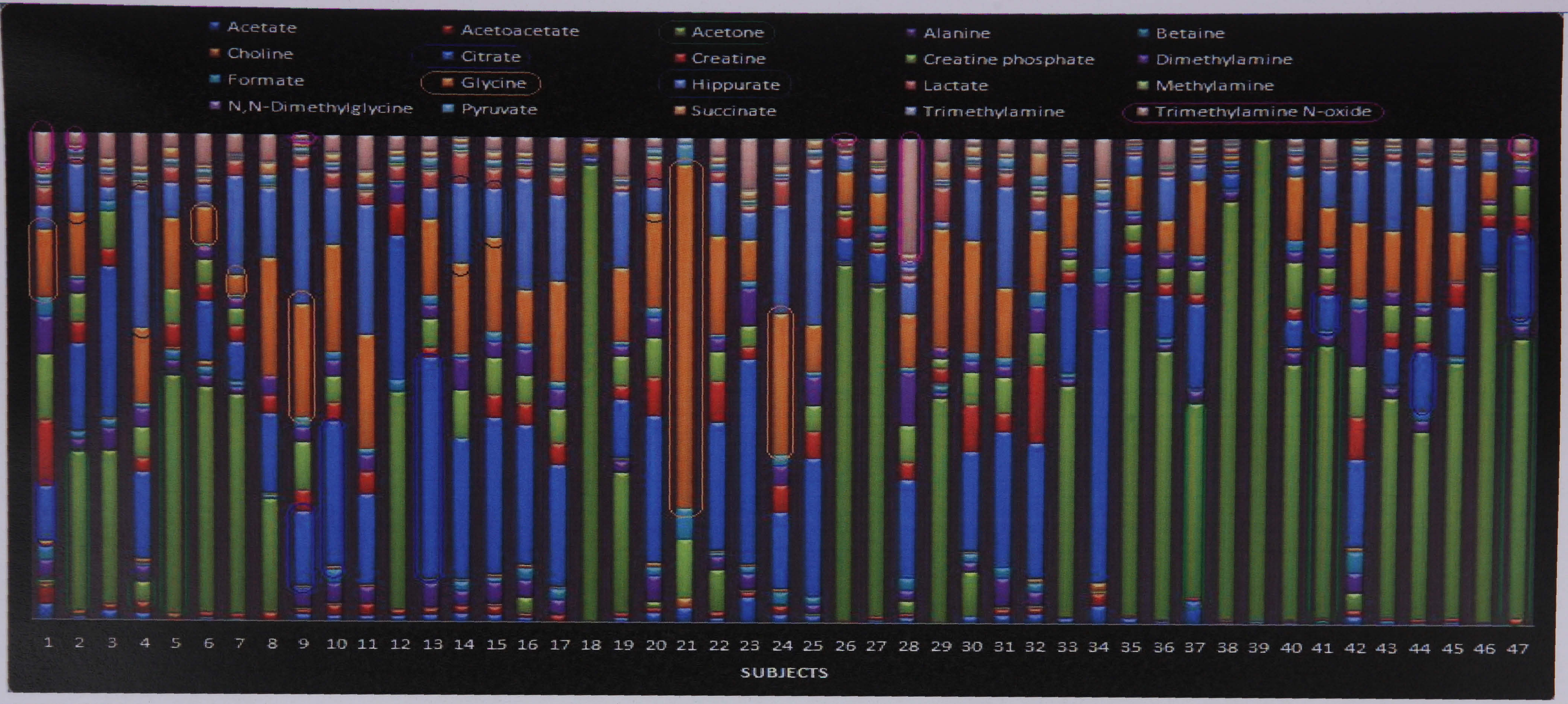


FIGURE 4.6: Bar-Chart of all subjects' metabolites concentrations (normalised) excluding urea and creatinine.
Ellipse showing the variances in the five selected metabolites - acetone, citrate, glycine, hippurate and TMAO.

TABLE 4.2: Concentration (mM) levels of metabolites in location- and gender- grouped control samples (Phase I intervention)

Metabolites	UK				Nigeria			
	Female	*p-value*	Male	*p-value*	Female	p-value	Male	p-value
Acetate	0.082 ± 0.057	*0.0015*	0.069 ± 0.036	*0.0000*	0.141 ± 0.046	0.0007	0.124 ± 0.094	0.0008
Aceto-acetate	0.051 ± 0.032	*0.0031*	0.079 ± 0.039	*0.0000*	0.141 ± 0.059	0.0003	0.085 ± 0.055	0.0001
Acetone	1.695 ± 3.138	*0.1400**	2.687 ± 2.680	*0.0052*	4.716 ± 4.447	0.031*	4.067 ± 3.273	0.0003
Alanine	0.207 ± 0.165	*0.0033*	0.193 ± 0.079	*0.0000*	0.233 ± 0.147	0.0058	0.211 ± 0.123	0.0002
Betaine	0.086 ± 0.064	*0.0035*	0.064 ± 0.040	*0.0002*	0.154 ± 0.092	0.0021	0.129 ± 0.070	0.0000
Choline	0.036 ± 0.026	*0.0028*	0.039 ± 0.017	*0.0000*	0.055 ± 0.024	0.0001	0.034 ± 0.018	0.0000
Citrate	1.372 ± 0.827	*0.0005*	0.970 ± 0.720	*0.0007*	1.909 ± 0.901	0.0005	1.624 ± 1.134	0.0001
Creatine	0.379 ± 0.292	*0.0026*	0.241 ± 0.155	*0.0002*	0.472 ± 0.181	0.0005	0.335 ± 0.235	0.0001
Creatine phosphate	0.339 ± 0.234	*0.0024*	0.427 ± 0.214	*0.0000*	0.566 ± 0.314	0.0014	0.340 ± 0.217	0.0002
Creatinine	6.346 ± 3.932	*0.0013*	6.917 ± 4.271	*0.0002*	14.28 ± 6.91	0.0001	9.592 ± 5.814	0.0000
Dimethylamine	0.188 ± 0.123	*0.0009*	0.195 ± 0.103	*0.0000*	0.402 ± 0.163	0.0002	0.314 ± 0.146	0.0000
Formate	0.103 ± 0.061	*0.0010*	0.081 ± 0.031	*0.0000*	0.144 ± 0.086	0.0021	0.104 ± 0.055	0.0000
Glycine	1.010 ± 0.761	*0.0023*	0.924 ± 0.581	*0.0002*	1.665 ± 0.862	0.0009	1.069 ± 0.726	0.0002
Hippurate	0.795 ± 0.825	*0.0140**	0.895 ± 0.779	*0.0022*	1.293 ± 0.921	0.0054	0.705 ± 0.664	0.0024
Lactate	0.166 ± 0.142	*0.0050*	0.131 ± 0.052	*0.0000*	0.299 ± 0.148	0.0018	0.162 ± 0.115	0.0003
Methylamine	0.038 ± 0.024	*0.0024*	0.023 ± 0.011	*0.0000*	0.047 ± 0.018	0.0004	0.029 ± 0.021	0.0002
N.N- Dimethylglycine	0.034 ± 0.018	*0.0005*	0.027 ± 0.007	*0.0000*	0.047 ± 0.026	0.0013	0.052 ± 0.023	0.0000
Pyruvate	0.071 ± 0.042	*0.0010*	0.068 ± 0.033	*0.0000*	0.105 ± 0.0516	0.0001	0.082 ± 0.050	0.0000
Succinate	0.099 ± 0.089	*0.0100*	0.049 ± 0.031	*0.0002*	0.145 ± 0.078	0.0006	0.112 ± 0.065	0.0000
Trimethylamine	0.020 ± 0.013	*0.0009*	0.021 ± 0.010	*0.0000*	0.041 ± 0.012	0.0000	0.026 ± 0.013	0.0000
Trimethylamine-N-oxide	0.186 ± 0.208	*0.0200**	0.251 ± 0.185	*0.0007*	0.411 ± 0.143	0.0003	0.395 ± 0.195	0.0000
Urea	84.86 ± 42.09	*0.0003*	74.64 ± 23.73	*0.0000*	72.83 ± 27.95	0.0005	70.47 ± 27.60	0.0000

*p-value >0.01

4.2.2.2 Comparative study of metabolic profile changes in intervention

In general, acetate, acetone, choline, citrate, formate, hippurate, pyruvate, succinate, TMA and TMAO showed to have great variances based on SD ≥ average. However, in the Phase II of the intervention (during), it was observed that the variances in terms of SD showed to be reduced in all measured metabolites except acetate, choline, succinate, TMA and urea. This could be due to the converging of groups (group clustering) resulting from subjects having slightly similar effects/ response to the soy-milk consumption.

Considering the general cohort, upon the ingestion of soy and comparing the Day 1 and Day 3 of the soy intervention, results showed either increased or decreased changes in metabolic profiles. Clearly shown in FIGURE 4.7, choline, creatine-phosphate and trimethylamine increased through the soy-intervention study (Days 1-12). However, unusual trends of un-parallelism (non-uniformity) in the during-soy phases (Days 3 & 5) were observed with creatine, creatinine, dimethyl-glycine, methylamine and succinate. Following this generalised biochemical changes from Phase I-IV, a more selective analysis was performed on the dataset by grouping the subjects (location and gender) and excluding outliers to identify potential biomarkers responsible during and after soy-metabolism. Wholly, creatine, creatinine, MA, TMAO were found to decrease after soy consumption, Days 3 - 12 (different phases II-IV) in all subject groups (except TMAO in Day 3/Phase II of the UK female group only).

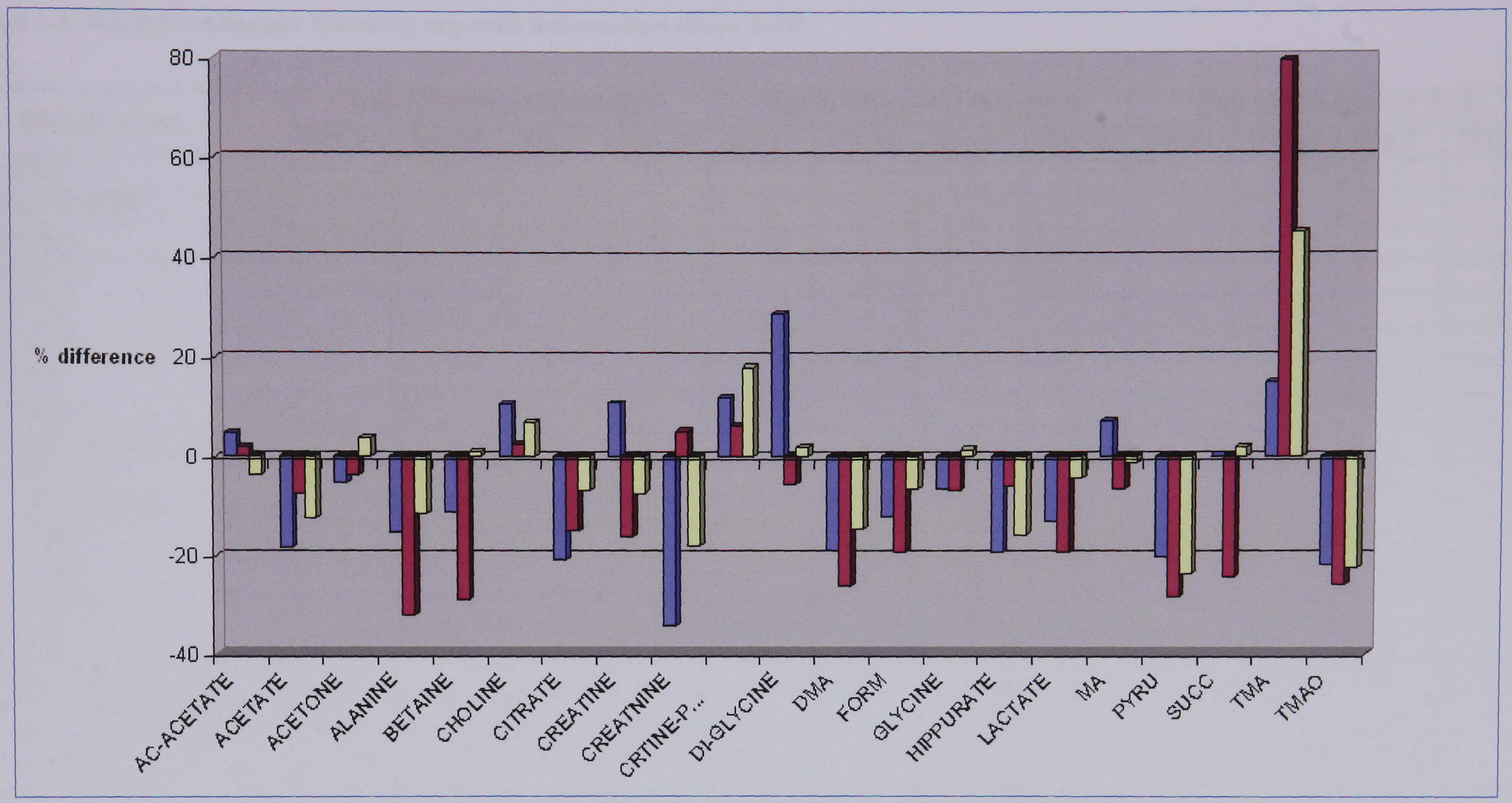


FIGURE 4.7: Metabolic effects of soy - phytoestrogen on endogenous metabolites over two-weeks (expressed as % difference) excluding urea.

●: Difference between Days 1-3 (Phase I-II); ●: Difference between Days 1-5 (Phase I-III); ●: Difference between Days 1-12 (Phases I -IV).

TABLE 4.3: Metabolic changes following soy-milk intervention (Days 3-12)

Metabolites	Day 3/ Phase II/ During I				Day 5/ Phase III/ During II				Day 12/ Phase IV/ Post			
	UK F	UK M	Nig F	Nig M	UK F	UK M	Nig F	Nig M	UK F	UK M	Nig F	Nig M
Acetate	↓	↑	↑	↑	↓	↓	↑	↓	↓	↓	↑	↑
Aceto-acetate	↑	↑	↓	↓	↑	↑	↓	↓	↑	↓	↓	↑
Acetone	↓	↓	↑	↓	↑	↓	↑	↑	↑	↑	↑	↑
Alanine	↑	↑	↑	↑	↑	↓	↑	↓	↓	↑	↑	↑
Betaine	↓	↑	↓	↓	↓	↓	↓	↓	↓	↑	↓	↓
Choline	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑
Citrate	↑	↑	↓	↓	↓	↓	↓	↓	↓	↑	↓	↑
Creatine	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
Creatine phosphate	↑	↑	↓	↓	↑	↑	↑	↑	↑	↓	↓	↑
Creatinine	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
Dimethylamine	↓	↓	↓	↓	↓	↑	↓	↓	↓	↑	↓	↓
Formate	↓	↑	↑	↓	↓	↓	↓	↓	↑	↑	↓	↑
Glycine	↑	↓	↓	↑	↑	↓	↑	↓	↑	↑	↑	↑
Hippurate	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓
Lactate	↑	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑
Methylamine	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
N,N- Dimethylglycine	↓	↑	↑	↓	↓	↓	↑	↓	↓	↓	↓	↑
Pyruvate	↓	↑	↓	↓	↓	↓	↓	↓	↓	↑	↓	↓
Succinate	↓	↓	↑	↓	↓	↓	↑	↓	↓	↓	↑	↓
Trimethylamine	↓	↑	↑	↑	↓	↓	↓	↑	↓	↑	↑	↑
Trimethylamine-N-oxide	↓	↓	↓	↓	↑	↓	↓	↓	↓	↓	↓	↓
Urea	↓	↑	↑	↑	↓	↑	↑	↑	↓	↑	↑	↓

↑ Increase in quantitated level (averaged) ↓ Decrease in quantitated level (averaged)

According to TABLE 4.3, phases- during II & III on Days 3 & 5 showed that soy metabolism had a common biochemical effect, i.e. decrease in creatine and creatinine concentrations. Other metabolites- dimethylamine (Day 3/ Phase II only), formate (Phase III), hippurate (Day 3/ Phase II), lactate (Day 5/ Phase III) and methylamine (Day 5/ Phases III) also showed concentration decrease. However, the during-phases II & III (Days 3 & 5) showed to be significantly influenced by location: specifically, aceto-acetate, citrate (Day 3/ Phase II only), creatine- phosphate (Day 3/ Phase II only) and hippurate (Day 5/ Phase III only). Furthermore, on Days 3 & 5 (Phases II & III) choline showed to be gender- influenced, as glycine on Day 5/ Phase III.

On Day 12 (Phase IV/ post-), it was observed that acetate, formate and citrate were the only significant metabolites influenced by location-differences; and citrate to be gender-influenced. In addition, creatine, creatinine, hippurate, lactate, methylamine had all decreased in the post-soy analysis, whilst glycine had a general increased level.

4.2.3 Multivariate statistical analysis (MVA) of the whole cohort (UK and Nigeria)

4.2.3.1 Whole spectrum binning (standard binning techniques)

To achieve differences among the four Phases of intervention, the ^{1}H-NMR spectral data from all Phase intervention samples were selected to be digitised for MVA using correlation method of PCA-X, PLS-DA and O-PLS-DA.

I. Unsupervised MVA Principal Component Analysis

Close examination of the individual (subject) metabolic changes through the intervention showed a distinct differentiation between all the four Phases. With the exclusion of TSP ($\delta \leq 0.05$) bucket, the model of the PCA-X of combined four different groups based on the Phases I-IV showed seven extreme outliers, thus these were also excluded. The PCA-X plots of the urine ^{1}H-NMR spectra as shown in FIGURE 4.8 displayed overlaid clustering of all four groups with no specific pattern, thus no discrimination of classes was observed. This was as a result of both inter-sampling and inter-subject variations. Observing the model overview plots FIGURE 4.8 (a) & (d), it showed that a better model fitness and predictivity (from 0.6 to 0.9 and 0.3 to 0.6: R^2X and Q^2X cumulative values respectively) was achieved with the exclusion of outliers. Comparing the loading plots in FIGURE 4.8

(c) & (f), with the exclusion of outliers it showed that certain region of the aromatic region (δ: 7.2 to 8.2) was found to be mainly responsible for the outliers. Generally, from this loading plot in FIGURE 4.8(c), it was observed that the positively influencing region (yellow ellipses) of the spectra causing the outliers variances were δ: 0.05 - 1.0, 5.7 - 6.0 and 6.0 - 9.0. With the outlier exclusion, the loading plot in FIGURE 4.8 (f) showed regions δ: 0.05 – 0.5, 4.5 - 5.0, 5.5 - 6.5 and 8.0-8.5 showed regions influencing slight variance in scores (data).

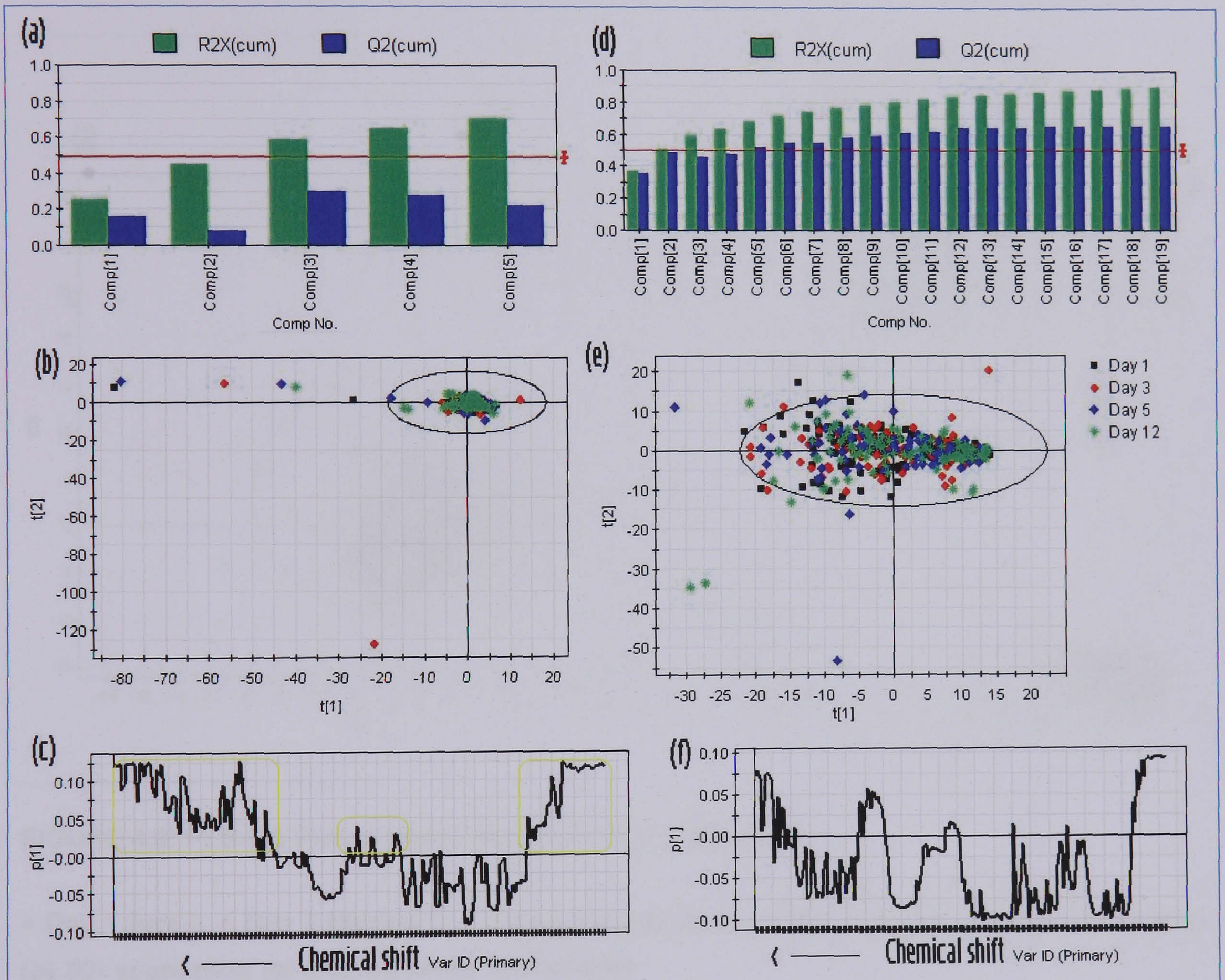


FIGURE 4.8: PCA-X (PCA-Class) plots of Phase I-4.

■Day 1 (pre-), ◆Day 3 (during I), ◆Day 5 (during II) and ✱Day 12 (post-) intervention sample data.

Whole dataset: (a) corresponding model overview plot; (b) score plot; (c) loading plot; Whole dataset excluding outliers: (d) corresponding model overview plot; (e) score plot; (f) loading plot. * indicates R^2 & Q^2 >0.5, which is the critical expected for good model fit and predictivity respectively.

II. Supervised MVA: PLS-DA

Due to the insignificant clustering and the un-clarity of the clustering pattern of all groups, supervised MVA was employed for an enhanced discriminant analysis. The basic supervised MVA: PLS-DA, where addition variable (y-variable) which is related to/ dependent on the x-variables was used. In this PLS-DA study, classes/ phase numbers represented qualitative y-variables and these y-variables showed to cause the significant separation of the groups (loading score plot). This has been summarised in FIGURE 4.9.

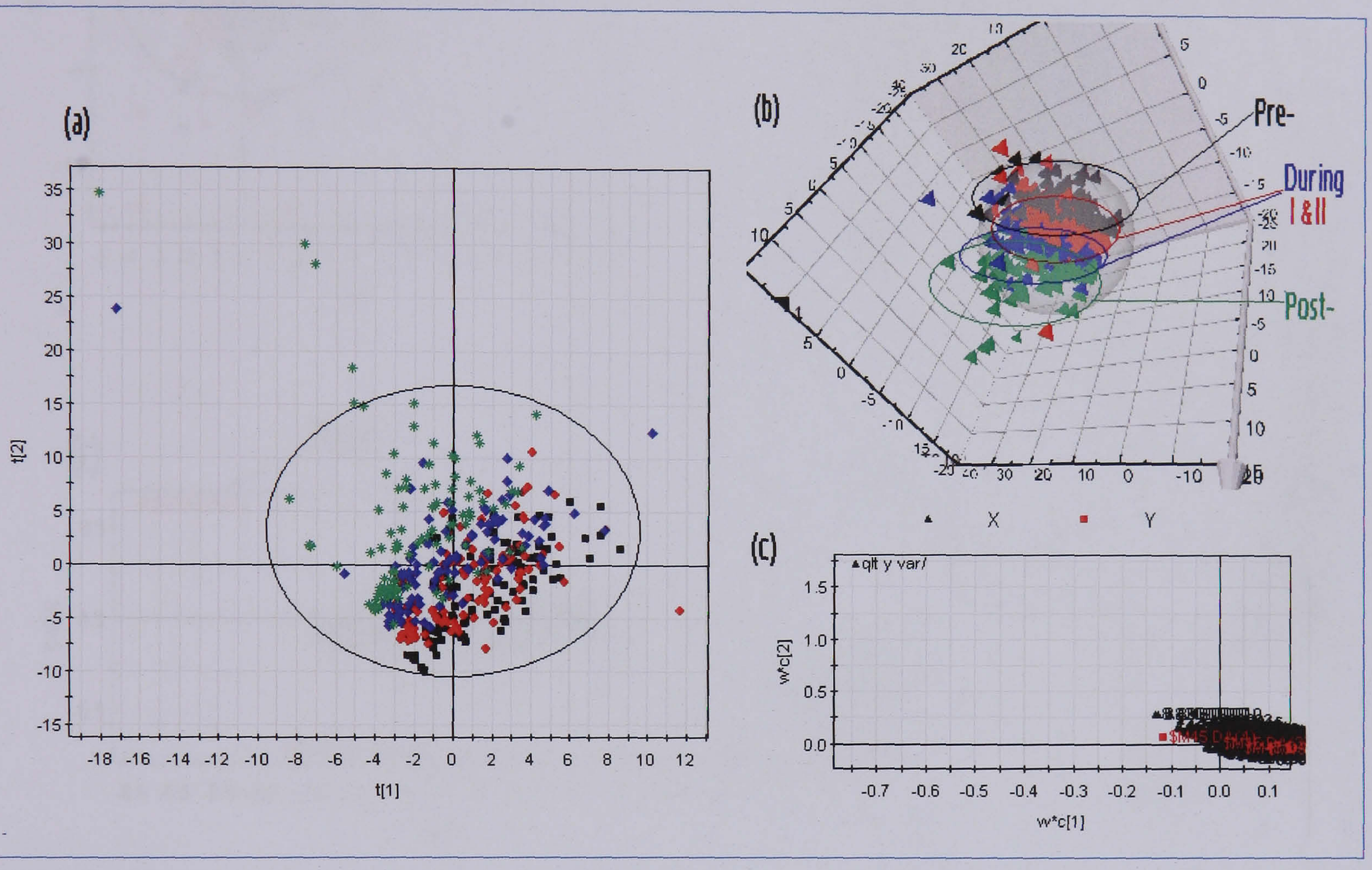


FIGURE 4.9: PLS-DA Plot of whole data excluding outliers.

■ Day 1 (pre-), ◆ Day 3 (during I), ◆ Day 5 (during II) and ✱ Day 12 (post-). (a) 2D score plot; (b) 3D- score plot; (c) loading score scatter-plot

Comparative to PCA, in PLS-DA, the clustering of the four Phase intervention (Days 1, 3, 5 & 12) groups were better separated, however, partial overlaps of data were also observed as seen in FIGURE 4.9b. Based on the loading plot, the positions of the clusters in the 2D score-plot were represented as new computer-generated y-variables (identified as the class i.ds). In addition, these new y-variables were observed to be close together leaving out the initially assigned class y-variable as the main discriminatory variable in the dataset.

With the PLS-DA completed, O-PLS-DA was performed to further improve the separation and to investigate the metabolic effects following the soy intervention with the aim of having clear separations of the cluster groups in the whole cohort dataset.

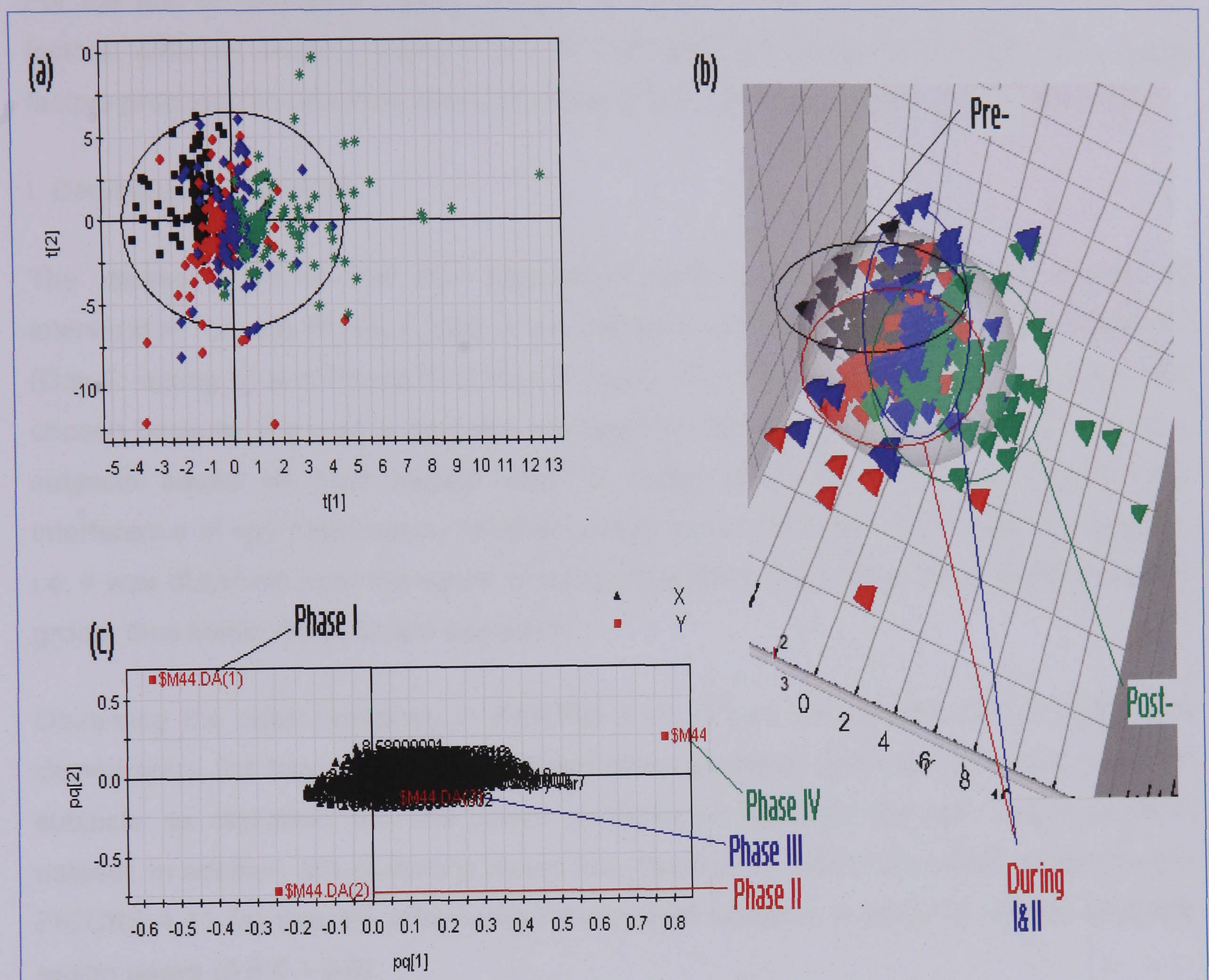


FIGURE 4.10: O-PLS-DA Plots of whole data excluding outliers.

■, ◆, ◆ and ✱ represent Phase I, II, III and IV respectively. (a) 2D- score plot; (b) 3D- score plot; (c) 2D- loading scatter plot.

Looking at the O-PLS-DA plot, it seems there was not much difference in these two plots (score plots); however, the loading plot of O-PLS-DA showed classes/ cluster groups to be better separated. Thus, the computer-generated y-variables showed to be more discriminating than the proposed y-variables.

4.2.3.2 Proposed factors influencing metabolic changes following soy intervention

For the test of metabolic profile changes by MVA caused by hypothesised influencing factors, different analysis using PCA, PLS-DA and O-PLS-DA were performed. These factors proposed to influence inter-variations in MVA include: location, gender and age.

I. Demography influences

The dataset used to test this hypothesis contained samples from all Phases of intervention, i.e. the Phase I (Day 1: control/pre-), Phase II (Day 3: during I), Phase III (Day 5: during II) and Phase IV (Day 12: post-). The Phase I (Day1) samples were pilot chosen because the control samples represent the general metabolic profile of individual subjects based on their regular diets & habits within the community without the interference of soy metabolism. Another reason for this selection is for dataset simplicity, i.e. it was observed from the earlier analysis, that there were slight differences within this group, thus fewer variables are expected.

Observing the data clustering in FIGURE 4.11, it can be reported that in relation to demography, the Nigerian cohort produced better sampling type than the counterpart- UK subjects: as explained with the tighter clustering in excluded aromatic region variable-dataset. In addition, the clustering shows from the loading plot of the model represented in FIGURE 4.11 (a) that the influencing inter-subject variation x-variables are the aromatic region peaks (δ 6.0 – 9.0).

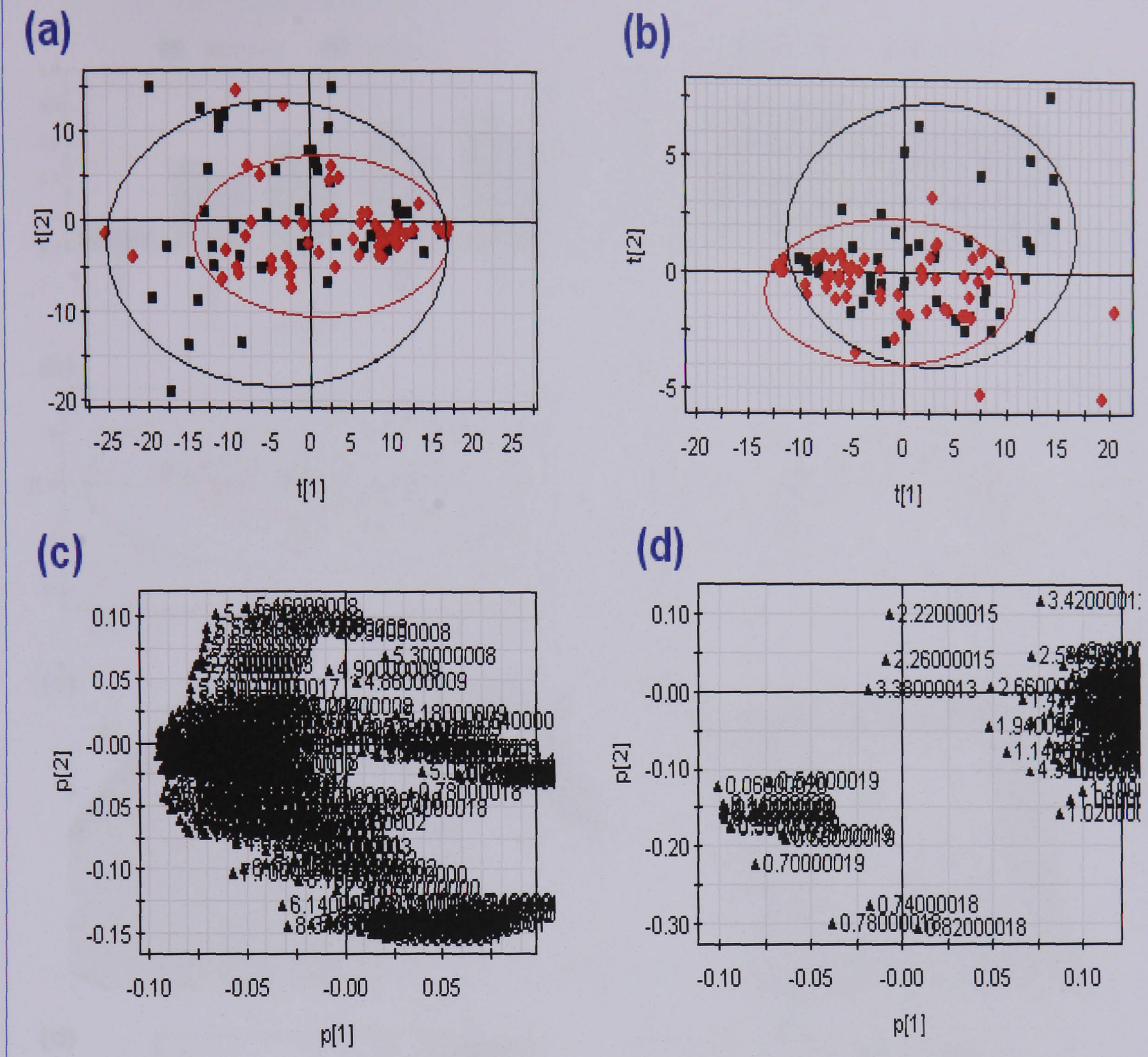


FIGURE 4.11: PCA-X comparison of location classification based on 2D-scoreplots.

◆ Nigeria, ■ UK: (a) whole region (δ: 0-9), (b) aliphatic region (δ: 0-4.5), (c) loading score plot of whole region (δ: 0-9) and (d) loading score plot of aliphatic region (δ: 0-4.5).

Although, unsupervised PCA-X shows a slight differentiation between the two groups, supervised MVA was necessary to confirm the discrimination of these two groups. With this, PLS-DA and O-PLS-DA were performed on the latter model set- aliphatic region (δ: 0 – 4.5). From FIGURE 4.12, it can be confirmed that location was a discriminatory factor, especially based on the O-PLS-DA, where two obvious group clusters were observed in both 2D- and 3D- scoreplots. In addition, the model overview plot showed both PLS techniques having good R^2 and Q^2 values, i.e. > 0.5, thus showing good fitness and predictivity of dataset.

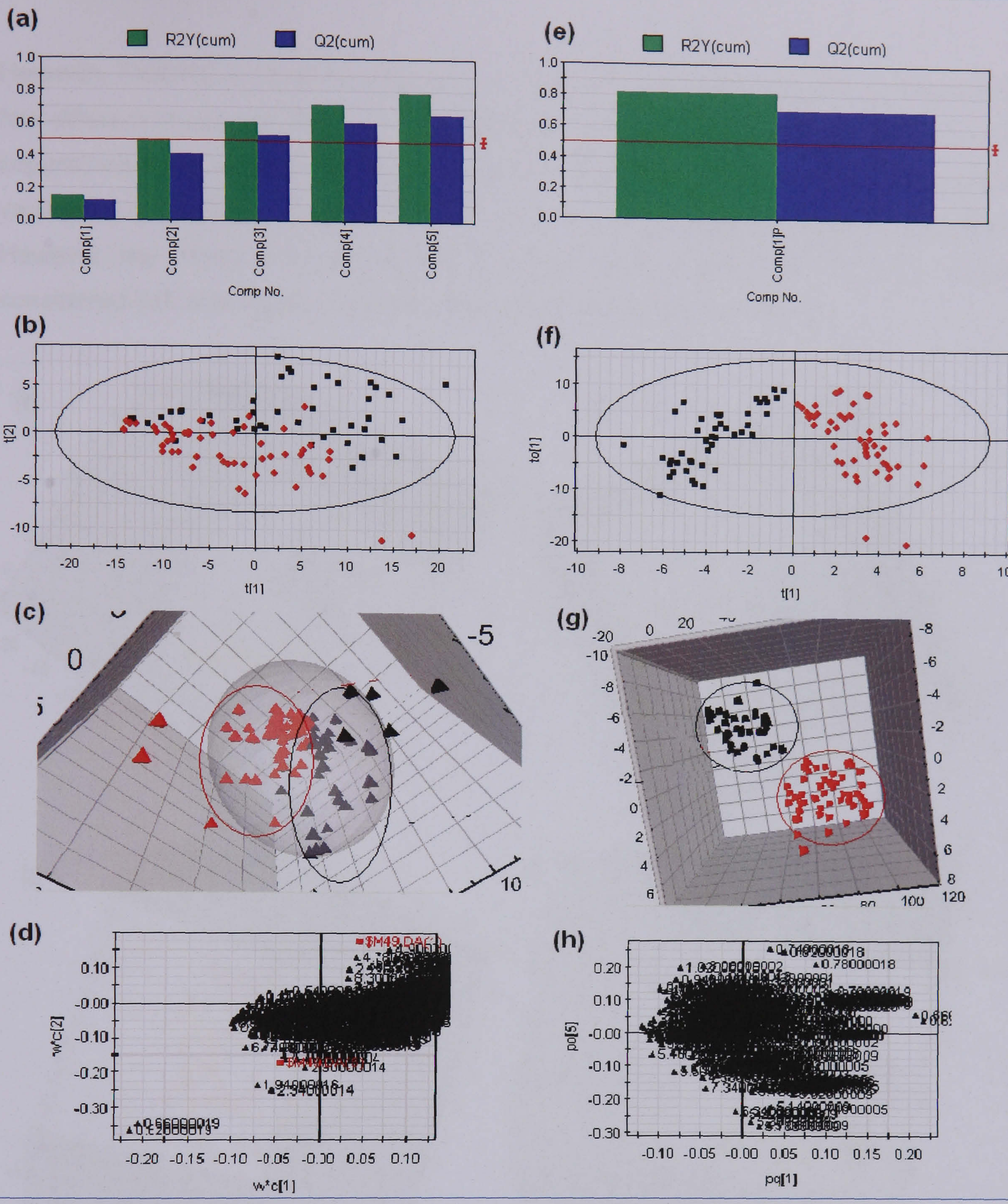


FIGURE 4.12: Supervised MVA (PLS) plots of Phase I location difference dataset (δ: 4.6 – 5.3 excluded).

◆ Nigeria, ■ UK: (a) corresponding PLS-DA model overview plot; (b) 2D-PLS-DA score plot; (c) 3D-PLSDA score plot; (d) corresponding PLS-DA model loading plot; (e) corresponding O-PLS-DA model overview plot; (f) 2D- O-PLS-DA score plot; (g) 3D-O-PLS-DA score plot; (h) corresponding O-PLS-DA model loading plot. * indicates R^2 & Q^2 >0.5, which is the critical expected for good model fit and predictivity respectively.

Herewith, FIGURE 4.13 below shows the O-PLS-DA score plots on differentiations of the four different Phase I-IV (Days 1 - 12) intervention samples, whereby relatively clear group separations were achieved as seen in PC1 and PC2 score plot based on demographic variation, i.e. location – Nigeria and UK as observed in all four phases independently. However, the Phase IV (post-) showed a partial overlap- grouping of the two locations considered (UK and Nigeria cohorts) compared to the other three phases.

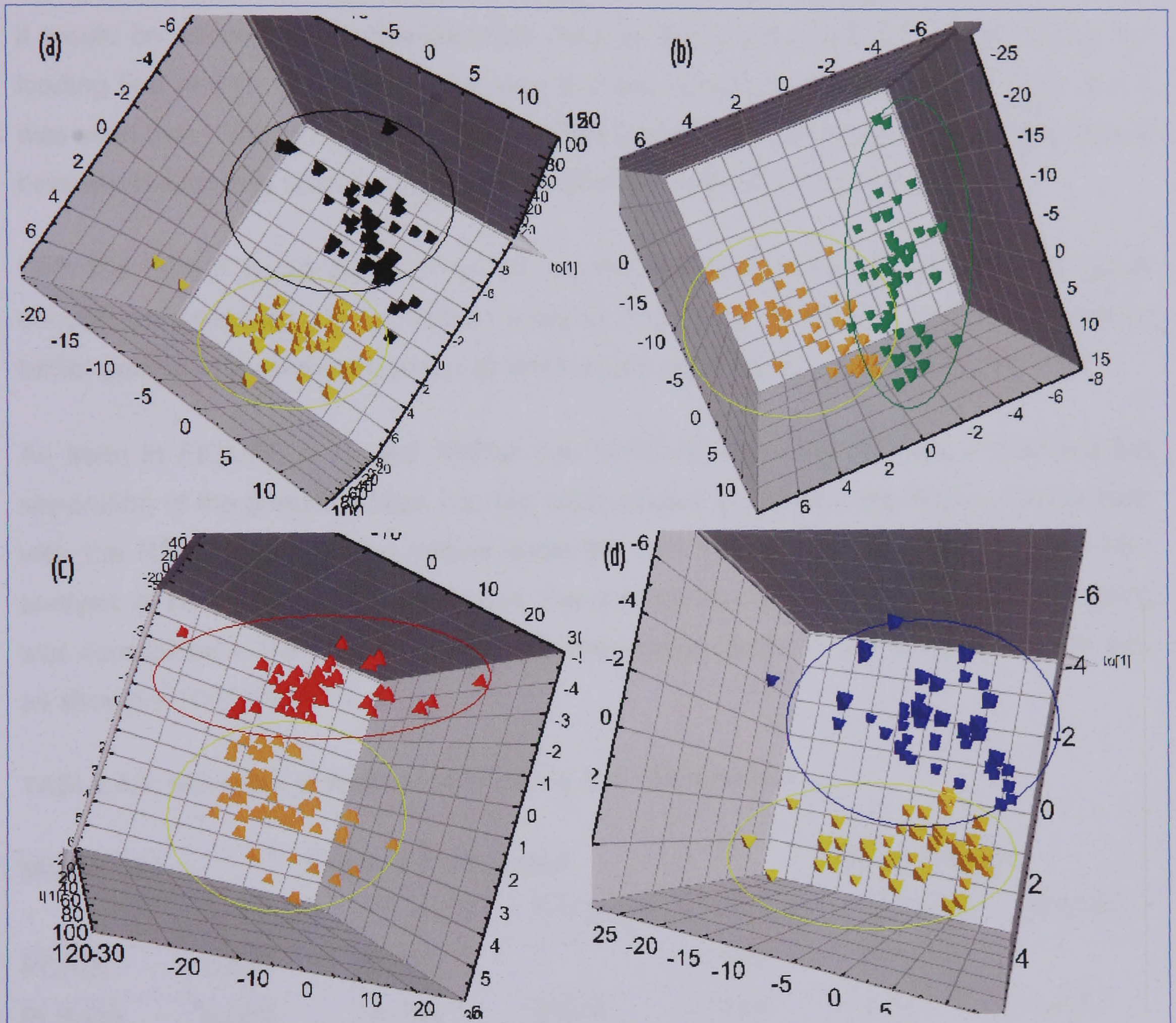


FIGURE 4.13: Supervised MVA (O-PLS-DA) 3D- score plot of Phase I-IV of soy intervention comparison based on location.

(a) Phase I (pre-): Day 1 ■UK ▲Nigeria; (b) Phase IV (post-) : Day 12 ✳UK ▲Nigeria; (c) Phase II (during I): Day 3 ●UK ▲Nigeria; (d) Phase III (during II) Day 5 ◆ UK ▲Nigeria.

II. Gender-influence in MVA

In order to assess the effect of gender on the metabolic response, supervised and unsupervised MVA models were generated using the Phase I and/ or combined Phase II & III intervention. ^{1}H-NMR spectra were used as x- variables and the genders- male and female groups (Groups 1 and 2 respectively) as the dummy qualitative y variables. Unsupervised PCA-X model was constructed as the pilot test, however no clear clustering pattern was observed, even with high R^2 & Q^2 cumulative values (>70%) indicating good fitness and prediction of PCA model. Although, these values were high, the plot still shows it would be difficult to discriminate new data as each group did not cluster. Using the loading line plot t1, t2, t3, etc to interpret the association between score t1, t2, t3, etc, it was seen that Phase I (Day 1) and Phase II/III (Days 3/5) models have similar differences between the loading blocks, i.e. similar correlation between x-variables.

With the PCA-X score plots, it can be reported that the group separation is not good enough for prediction. Thus, further analysis tool using PLS-DA & O-PLS-DA showed better gender- discrimination plots (all MVA model plots summarised in FIGURE 4.14).

As seen in FIGURE 4.14 and TABLE 4.4, O-PLS-DA seemed to have maximised the separation of the gender groups into two clear subset/ group clusters. This in combination with the R^2 & Q^2 cumulative values show the two models are well fitted for predictive analysis of new data set. Following this, the combination of location and gender grouping was considered in the Phases I - IV of soy intervention (Days 1 - 12) using O-PLS-DA only as shown in FIGURE 4.15.

TABLE 4.4: Summary of model 4.14 R^2 (cum) & Q^2 (cum) values

Model 4.14	Phase II & III dataset			Phase I dataset		
	R^2X (cum)	Q^2 X(cum)	R^2Y (cum)	R^2X (cum)	Q^2 X(cum)	R^2Y (cum)
PCA-X	0.899	0.709	-	0.846	0.709	-
PLS-DA	0.665	0.470	0.674	0.594	0.281	0.472
O-PLS-DA	0.715	0.592	0.736	0.754	0.521	0.785

- Not applicable

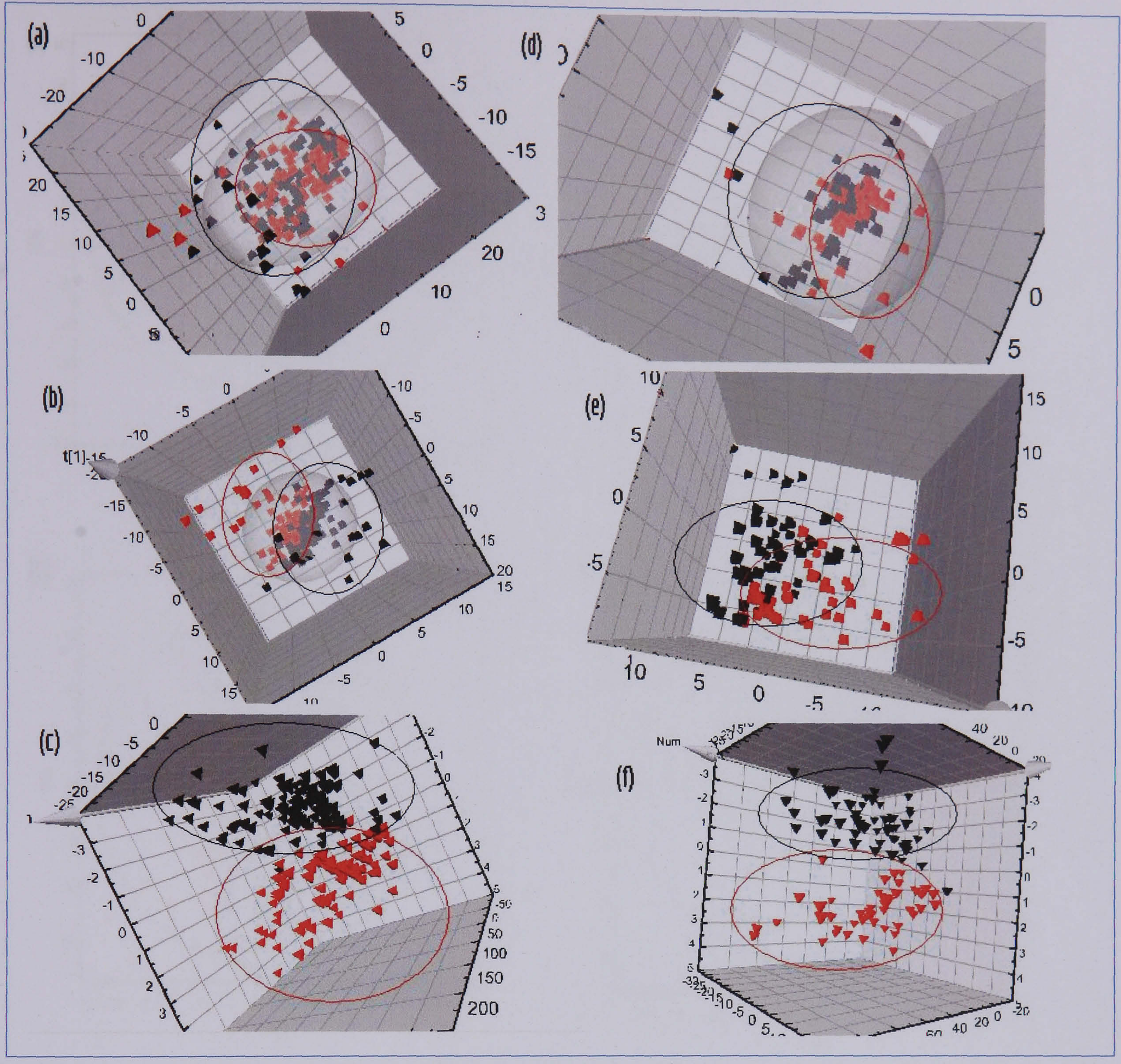


FIGURE 4.14: Comparison of gender using three MVA -3D score plots (δ: 0-9).

• Female ▪ Male. Phase II & III combined dataset: (a) PCA-X score plot; (b) PLS-DA score plot; (c) O-PLS-DA score plot. Phase I only dataset: (d) PCA-X score plot; (e) PLS-DA score plot; (f) O-PLS-DA score plot.

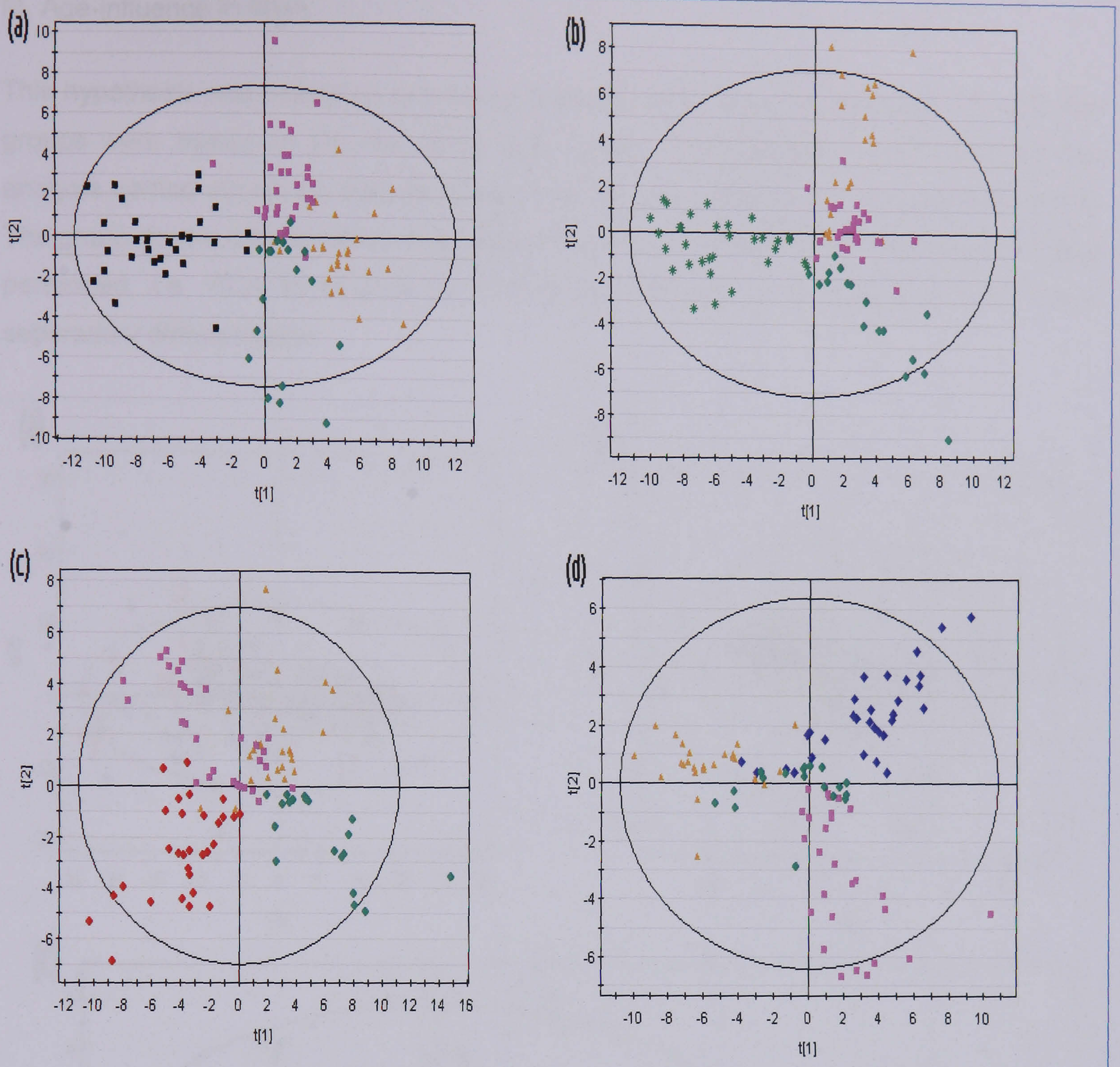


FIGURE 4.15: Supervised MVA (O-PLS-DA) 2D- score plot of groupings based on location & gender.

■, ✱, ● and ◆ represent UK male cohorts in Phases I, IV, III and II respectively and ◆ UK female cohorts; Nigeria: ▲ male and ■ female cohorts respectively. (a) Day 1 (Phase I / pre-); (b) Day 3 (Phase IV/ post-); (c) Day 5 (Phase II/ during I); (d) Day 12 (Phase III/ during II)

From FIGURE 4.15, the O-PLS-DA showed the separation of groups based on location in PC 1 and gender in PC 2. Clear group separation was achieved in Day 1 (Phase I/ pre-) & Day 3 (Phase II/ during I), and thereafter, merging and overlapping of some groups were observed in Day 5 (Phase III/ during II) & Day 12 (Phase IV/ post), especially the female subjects (6 & 7) showed slight convergence.

III. Age-influence in MVA

This hypothesis was tested on the during phases- Phase II and III datasets. The selected groups were based on the five listed age ranges. Unsupervised PCA-X was the first analysis carried out on the dataset followed by the unsupervised PLS-DA and O-PLS-DA. The group cluster differentiation was explained as more selective data discrimination were performed, i.e. PCA to PLS-DA to O-PLS-DA. FIGURE 4.16 illustrates these group separation/ differentiation.

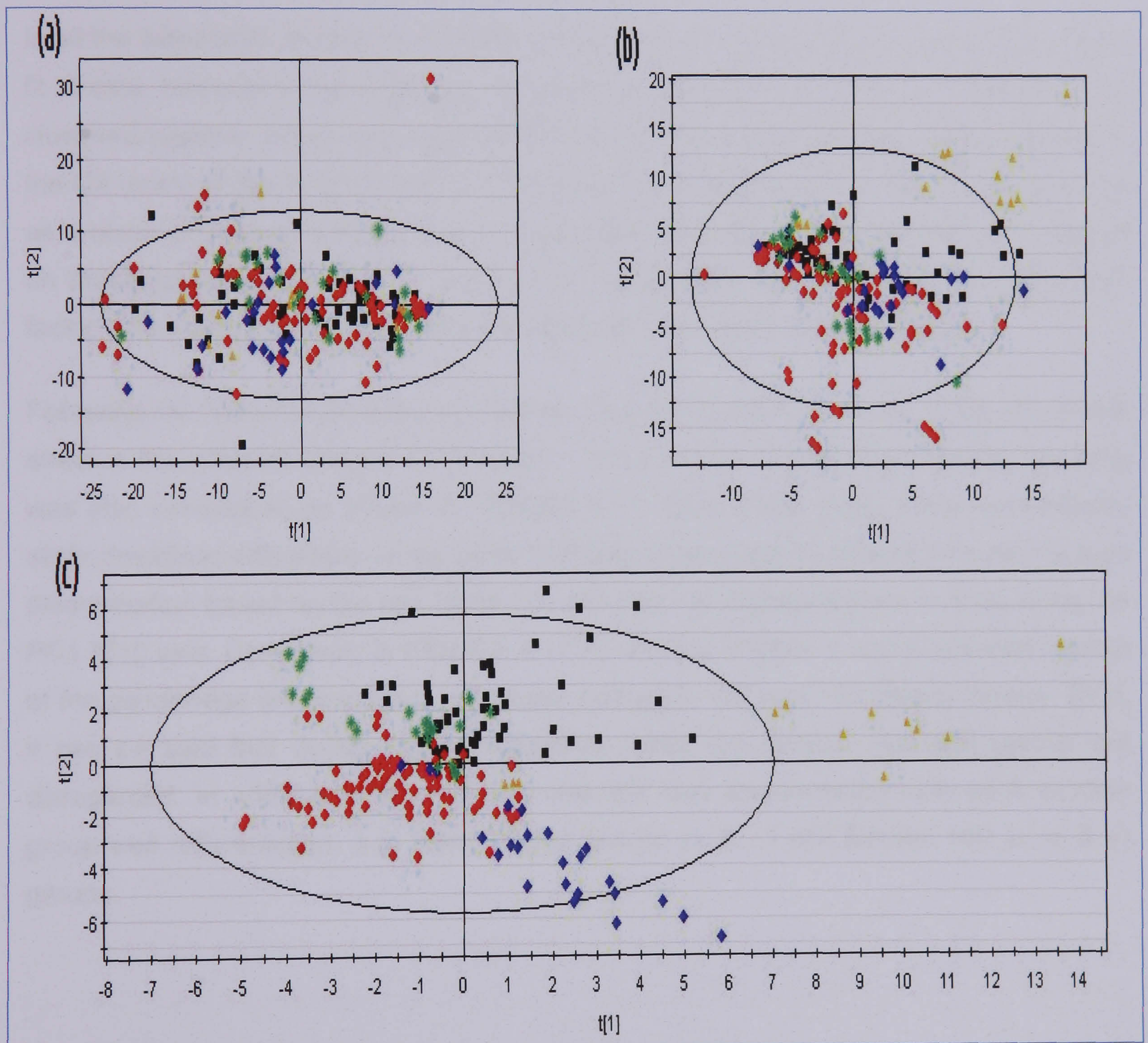


FIGURE 4.16: Comparison of three MVA in 2D- score plots of age y-variable dataset (δ: 0-9)

Combined Days 3 & 5 (Phase II & III/ during) dataset. ■ 20-25 yrs; ◆ 25-30 yrs; ◆ 30-35 yrs; ✱ 40-45; and ▲ >60 yrs. (a) PCA-X score plot; (b) PLS-DA score plot; (c) O-PLS-DA score plot.

TABLE 4.5: Summary of model 4.15 R^2 (cum) & Q^2 (cum) values

Model 4.15	Phase II & III dataset		
	R^2X (cum)	Q^2 X(cum)	R^2Y (cum)
PCA-X	0.900	0.708	-
PLS-DA	0.846	0.394	0.683
O-PLS-DA	0.747	0.414	0.551

- Not applicable

In all the listed MVA models of FIGURE 4.16, the R^2 (cum) were all >0.7, implying a good fit of data, however visual inspection of the score plots (both 2D- and 3D-) showed to be clustered together around the origin rather than separate into different groups. Although, the Q2 (cum) of the PLS-DA and O-PLS-DA were reported to be low (<0.5), this could be as a result of noise or there is a poor relationship between the y-variables. Thus, based on these conflicting inferences, age can not conclusively be classified as a dependent factor of soy metabolism measuring the metabolic changes in urine samples.

Following the classification based on during – soy (Days 3 & 5) (FIGURE 4.16), the overall effect in the different phases I-IV (Days 1 - 12) based on location-gender-age grouping was also considered as shown in FIGURE 4.17. From these plots, there were notably slight metabolic differences in the Days 1 (Phase I/ pre-) and 12 (Phase IV/ post-) groups differentiation based on the age factor i.e. <30 and >30 Nigeria female or male along the PC1 (t[1]) axis. Conversely in FIGURE 4.17 (c) and (d), Phases II and III showed overlap of the gender-age group clusters, especially noticed in <30 and >30 Nigeria female. Thus, it can be said that during soy consumption, these two factors- age and gender are disregarded. In addition to this, location was still very discriminatory with each location group well differentiated, e.g. male: >30 yr groups (● &) and female: >30 yr (● & ●) groups.

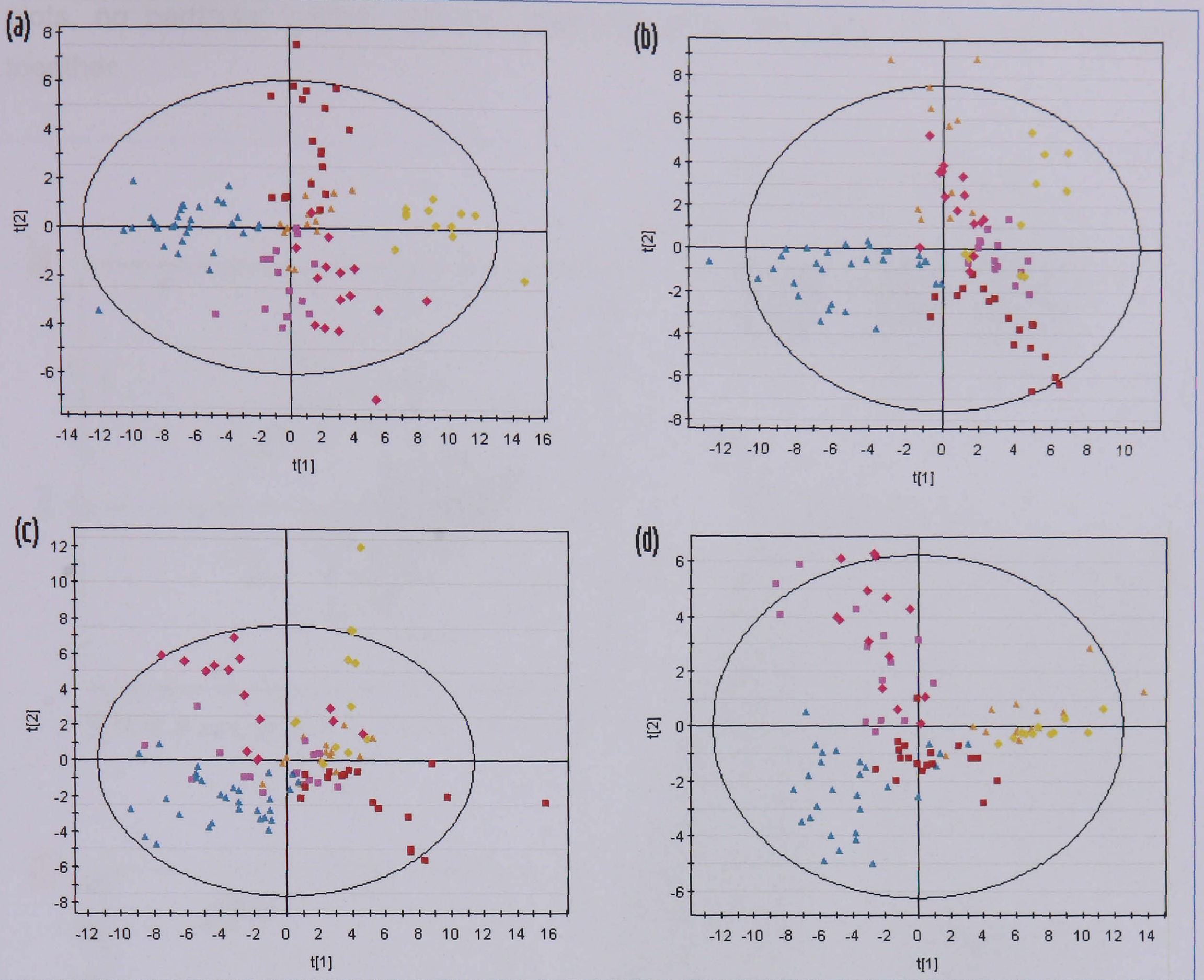


FIGURE 4.17: Supervised MVA (O-PLS-DA) 2D- score plot of Phase I-IV of soy-intervention based on location & gender differences.

<30 yr Nigeria: ◆ male and ■ female; >30 yr UK: ▲ male and ■ female; >30 yr Nigeria: ▲ male and ◆ female. (a) Day 1 (Phase I/ pre-); (b) Day 12 (Phase IV/ post-); (c) Day 3 (Phase II/ during I); (d) Day 5 (Phase III/ during II).

4.2.3.3 Target binning

Identified and quantified data of twenty-four metabolites from the Phase I-IV intervention samples using Chenomx® Suite Profiler generated 39- varied sized-buckets which were exported into SIMCA-P+® for MVA. By means of unsupervised MVA tool- PCA-X was performed on the dataset as shown in FIGURE 4.18. From the analysis of these plots, it can be reported that the model overview plot showed the R^2 & Q^2 (cum) were >0.5 implying that the model had a good data fit and predictivity. Although, based on the score

plots, no particular pattern can be observed, being that most of the data clustered together.

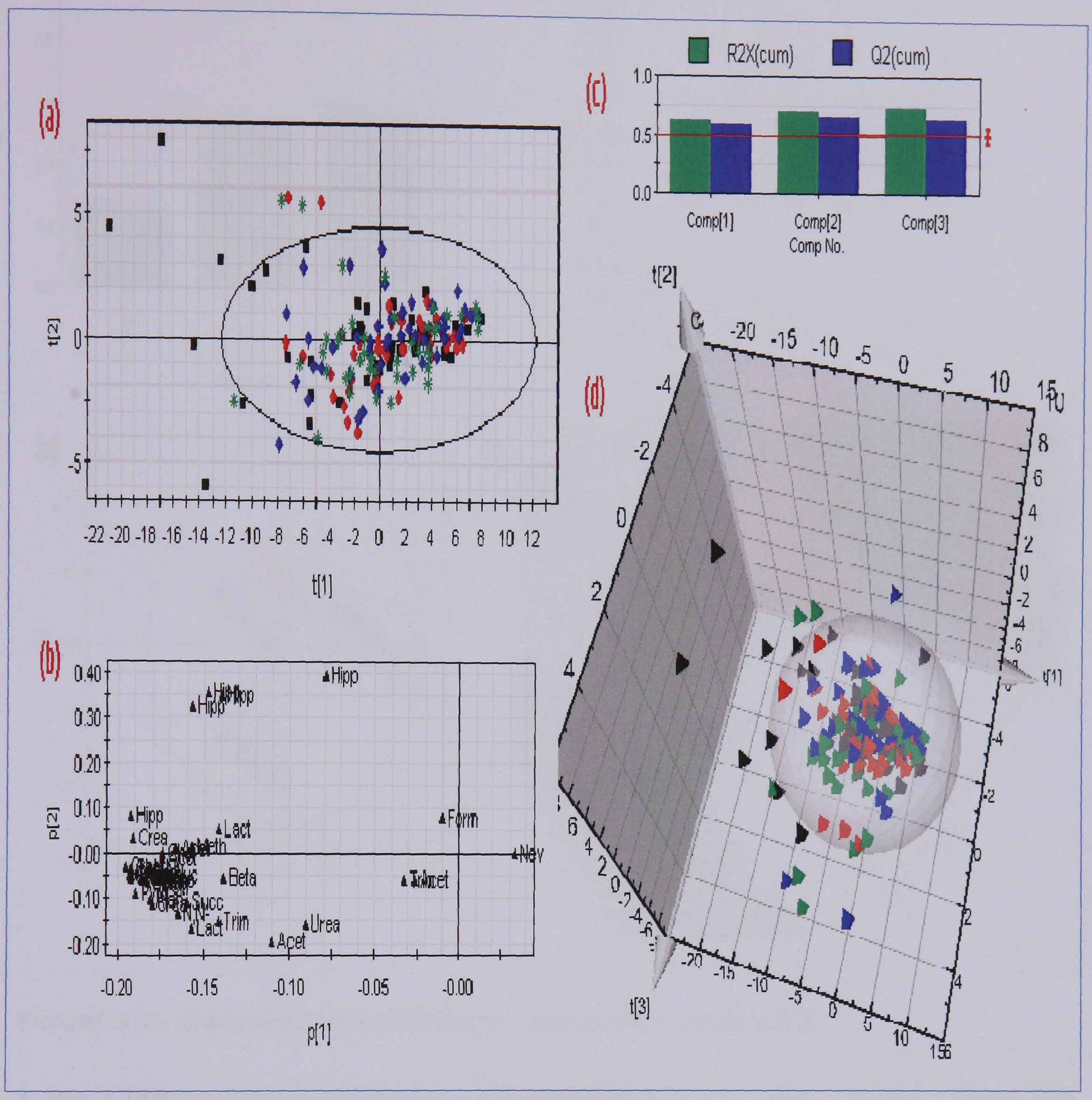


FIGURE 4.18: Unsupervised MVA (PCA-X) of Chenomx® profiled dataset.

■ Day 1 (pre-), ● Day 3 (during I), ◆ Day 5 (during II) and * Day 12 (post-): Phase I-IV respectively. (a) 2D- PCA-X score plot; (b) 2D loading score plot; (c) model overview plot; (d) 3D- score plot. * indicates R^2 & Q^2 >0.5, which is the critical expected for good model fit and predictivity respectively.

Due to the insufficient discrimination of these groups, further MVA tools were performed, i.e. PLS-DA and O-PLS-DA.

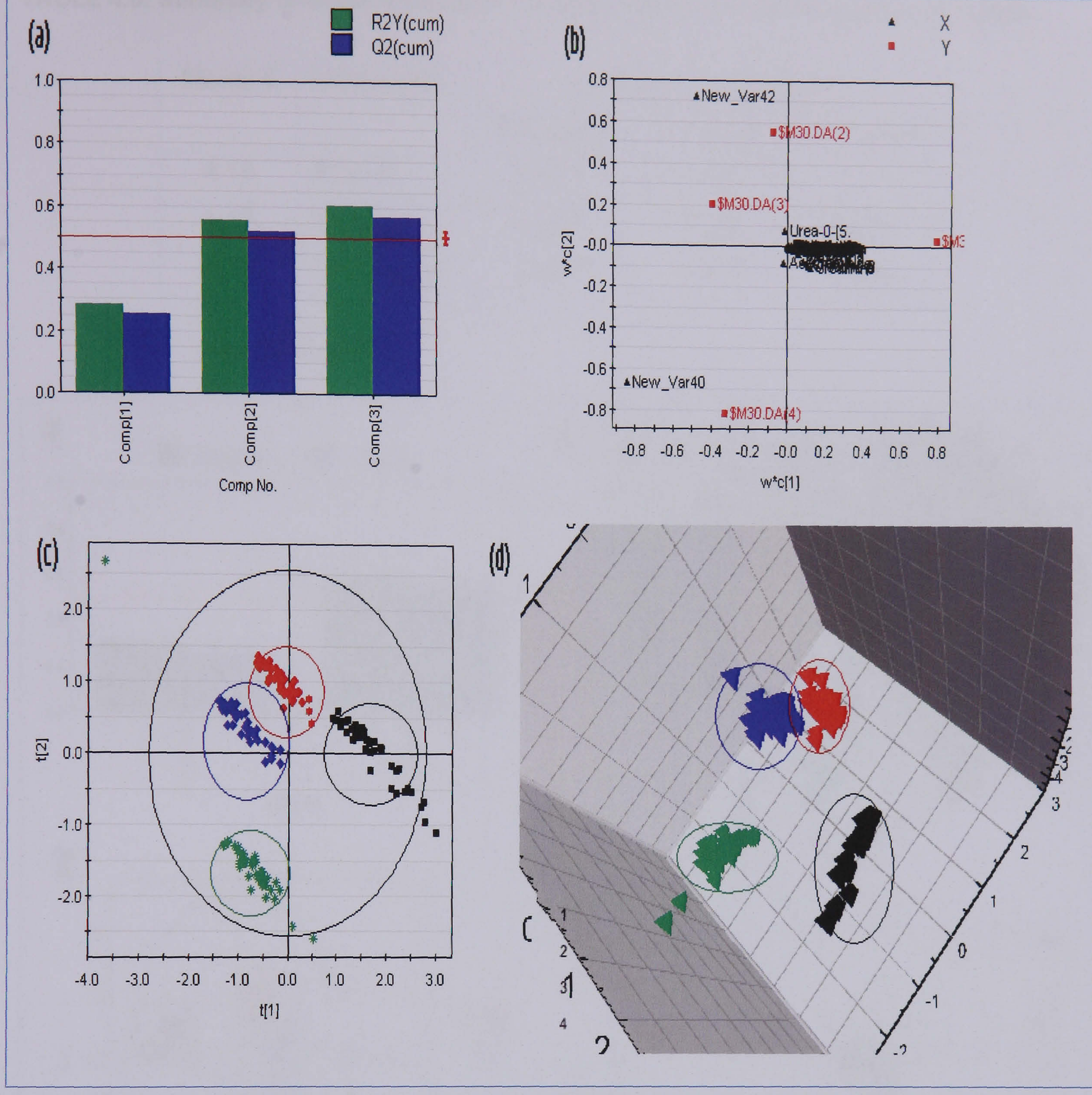


FIGURE 4.19: Supervised MVA (PLS-DA) of Chenomx® profiled dataset.

■ Day 1 (pre-), ◆ Day 3 (during I), ◆ Day 5 (during II) and * Day 12 (post-): Phase I-IV respectively. (a) Model overview plot; (b) loading score plot; (c) 2D- PLS-DA score plot; (d) 3D- PLS-DA score plot. * indicates R^2 & Q^2 >0.5, which is the critical expected for good model fit and predictivity respectively.

In support of the score plots in PLS-DA and O-PLS-DA (FIGURES 4.19 & 4.20 respectively), it can be further confirmed that target binning multivariate statistical analysis is a better binning technique to standard whole region (simple rectangular) binning. Additionally, with the R^2 & Q^2 cumulative values (tabulated in TABLE 4.6) in all models having both values > 0.5 each, further provides the advancement of target binning.

TABLE 4.6: Summary of model Chenomx®- profiled dataset: R2 (cum) & Q2 (cum) values

Model #	MVSA tool	Phase II & III dataset		
		R^2X (cum)	Q^2 X(cum)	R^2Y (cum)
4.16	PCA-X	0.713	0.667	-
4.17	PLS-DA	0.587	0.606	0.571
4.18	O-PLS-DA	0.719	0.550	0.596

-Not applicable

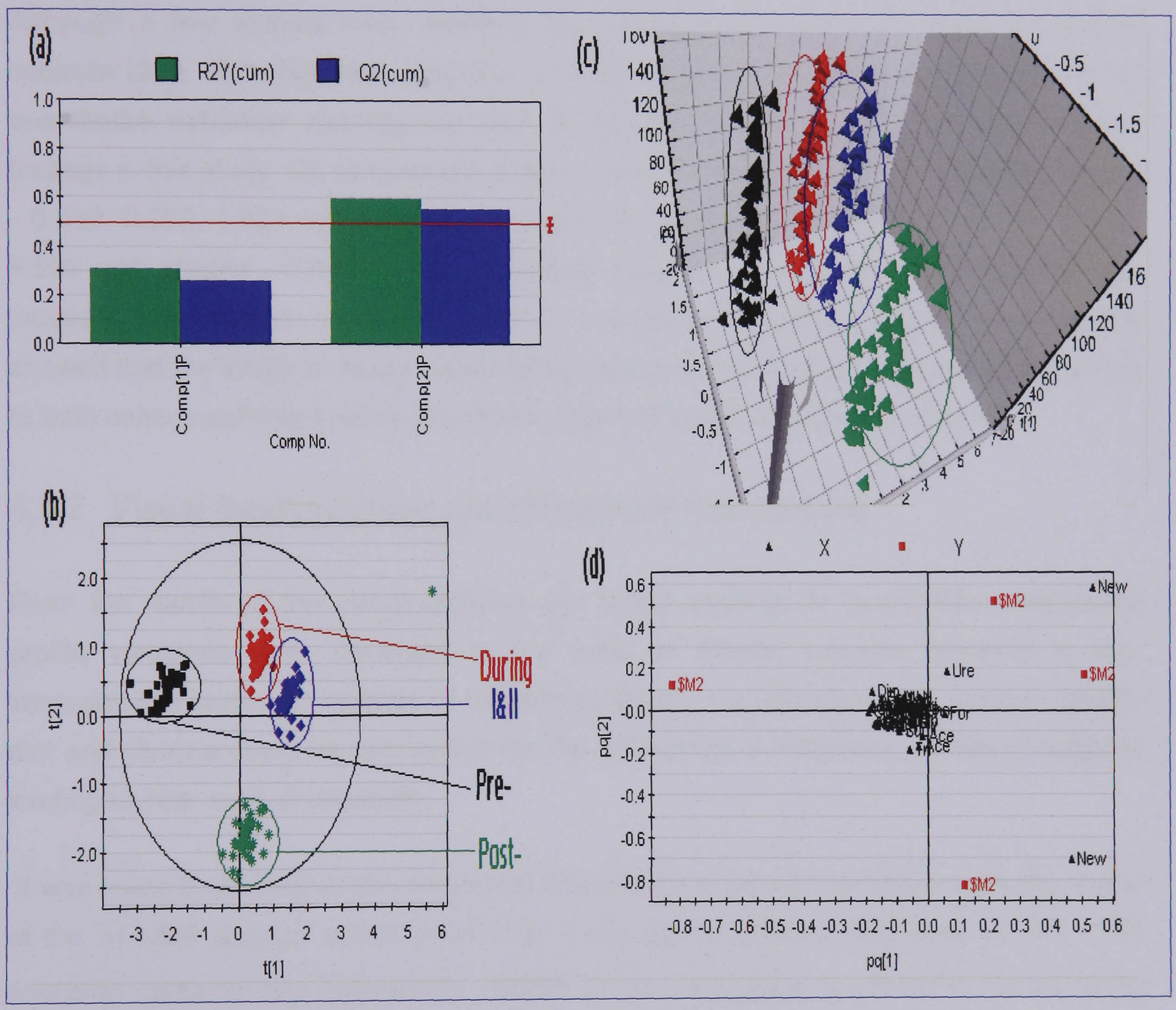


FIGURE 4.20: Supervised MVA (O-PLS-DA) of Chenomx® profiled dataset.

■ Day 1 (pre-), ● Day 3 (during I), ◆ Day 5 (during II) and ✱ Day 12 (post-): Phase I-IV respectively. (a) Model overview plot; (b) 2D- O-PLS-DA score plot; (c) 3D- O-PLS-DA score plot. * indicates R^2 & Q^2 >0.5, which is the critical expected for good model fit and predictivity respectively.

4.3 Discussion

Metabonomics was applied in this study to observe the biochemical effect of soy-metabolism with the consumption of soymilk. The metabolic effect during (Phase II & III) and after-intervention (Phase IV) using ^{1}H-NMR were investigated in 47 African subjects selected from two different locations- UK and Nigeria.

4.3.1 Quantitation of endogenous metabolites

Although a few studies have reported the endogenous metabolites levels in human subjects (*230, 257*), however, only one of these studies used Chenomx software for the quantitative validation, whereby the level of metabolites reported were comparable to the findings in this study. On comparison of these results, the urinary excreted levels of citrate - 0.830: 0.970: 1.624 mM; creatine: 0.188: 0.241: 0.335 mM; creatinine - 6.830: 6.917: 9.592 mM; glycine - 0.809: 0.924: 1.069 mM; hippurate - 1.012: 0.895: 0.705 mM; and lactate - 0.097: 0.131: 0.162 mM (Weljie's: UK cohort: NIGERIA cohort). These results showed that the Weljie's results were closely similar to the UK cohort. These could be due to both cohorts sampling being collected in Western countries: Canada and UK.

4.3.2 Visual Interpretation and Chenomx® quantitation

From the results of the visual analysis and target profiling for quantitation, metabolite profile variations were observed in the urine of African subjects following a soy-interventional study. Complexity of influencing factors e.g. genetics, age, gender, health, diet and lifestyle could be responsible for the differences in metabolic profiles of subjects leading to inter-subject variability.

It was found that visual analysis showed an increase of signal intensity in aromatic region of the ^{1}H-NMR spectra, which could have been due to elevated hippurate (δ: 7.5 - 7.9) excretion following soy intervention. Based on the calculated concentration levels using Chenomx®, it was found that the hippurate decreased following intervention, thus implying that the observed changes in the aromatic region was due to the excretion of polyphenolic/ aromatic metabolites. On the whole, creatine, creatinine and TMAO (excluding UK female on Day 5/ during II) showed general decrease in excretion following soy intervention on Days 3 - 12, i.e. Phases II - IV. On the other hand, with respect to the post-effect of soy consumption (Day 12/ Phase IV), lactate, methylamine and hippurate

decreased in excretion; whilst this resulted in increased acetone and glycine excretion. Complimentarily, acetate excretion was found to be influenced by location as citrate and choline were found to be influenced by gender. With respect to Day 3 (Phase II/ during I), increased urinary alanine excretion and decreased dimethylamine, hippurate, methylamine and TMAO excretion were observed. With the classification of subjects into gender and location: aceto-acetate and creatine-phosphate were found to be location dependent; as choline was gender dependent. Furthermore, on Day 5 (Phase III/ during II), increased urinary creatine-phosphate and decreased urinary betaine, formate, lactate, methylamine and pyruvate were observed in the general observation of subjects. Specifically in Phase III on Day 5, aceto-acetate and hippurate were location dependent; as alanine, choline and glycine were gender dependent.

4.3.3 Biochemistry of metabolites- overall biochemical effect of phytoestrogen

First-and-foremost, it is important to take into account that microbial production has been reported to contribute to metabolites concentration elevation. As well, some of these metabolites are present in the mammalian diet, thus upon ingestion, the levels increase. Besides this, degradation of mammalian choline converts it to creatinine via dimethylglycine. Thus, the elevation of choline complementarily influenced the elevation of creatinine and dimethylglycine. In addition to this as reported by Wiback et al (*339*), the elevation of lactate and pyruvate indicates increased glycolysis due to the oxidation of glucose to lactate and pyruvate. Thus, this study showed that the reduction of these metabolites implies reduced glycolysis after Day 3 through to the end of soy-intervention (Days 3 - 12/ Phases 2 - 4).

Increased levels of aceto-acetate, choline, creatine, creatinine phosphate, dimethyl-glycine and TMA on Days 1 & 3 (Phases I & II) were observed as shown in FIGURE 4.7. Following this, comparing the during-Phases II & III (Days 3 - 5) showed that there were parallel changes in some metabolites, except in creatine, creatinine, dimethylglycine and methylamine. Thus, an increase in choline would result to an increase in methylamine, as found on Day 5 (Phase III/ during II). On the final day (Day 12/ Phase IV/ post-soy), negligible elevations of concentration levels were noted in acetone, choline, creatine - phosphate, dimethylglycine, glycine, and succinate; as well as a significant elevation in TMA.

Considering the plot in FIGURE 4.7 in every respect, metabolites including creatine, creatinine, dimethylglycine, methylamine and succinate were peculiar based on their irregular metabolic change pattern. The problem being the un-parallel response in Days 3 & 5 (Phases II & III- during) observed. On Day 3, there was an increase in creatine, dimethylglycine, methylamine and succinate; and decrease in creatinine. Whilst on Day 5, these were vice-versa. As creatinine is a waste product of muscle metabolism, thus, its metabolic pattern could be explained by its production dependence on muscles, for instance, muscles under stress produce more creatinine. Consequently, creatinine production and excretion are highly individual-dependent.

The major amino-acids: glycine and alanine are known to be responsible for collagen synthesis, and in relation to this study, both metabolites were found to have decreased. Thus, this finding implies the reduction in the synthesis of collagen.

4.3.4 Multivariate analysis interpretation

With the visual analysis deduction of largely-influencing metabolites causing inter-subject variation being in the aromatic region, MVA tools- PCA-X was employed to further confirmed this as shown in FIGURE 4.11. Unsupervised MVA (PCA-X) plots of Phases I-IV showed no distinct separation of the different phase-groups (Days 1 - 12). This result is in accordance with previous studies reporting the non - uniformity of biochemical effects of soy (supplements) (*340*). This was further confirmed using individual subject PCA-X analyses, whereby there were noticeable differences in the four phases for each individual. These subtle changes were found to be inconsistent in subjects, which could be as a result of soy - response difference in each subject. However, due to these subtle metabolic changes and sample size, the overall PCA-X plot showed to have no clear distinction between the phases. With this drawback, supervised MVA was employed as the preferred tool for clarity of this metabonomic study.

4.3.4.1 Influencing factors in metabolic biochemical effects of soy-metabolism

As a result of inter-subject variations as seen in unsupervised MVA (PCA-X), more selective MVA (supervised) and classification had to be performed using influencing factors variables- location, gender and age. With all the MVA tools used for location-dependent classification, it can be reported that the NIGERIAN cohort had less metabolic variation throughout Phases I – IV due to better group clustering even with the PCA-X plot

in FIGURE 4.11. With the aid of loading plot, aromatic region was less-influencing within this group showing less variation in polyphenolic consumption, i.e. very similar dietary habit. Interpreting the O-PLS-DA score-plot of the location difference in post-soy (Phase IV), an overlap of groups was observed possibly due to similarity in metabolic response to soy metabolism.

Also, the observed convergence of female groups (group-overlap) in UK and NIGERIA could be due to the parallel responsiveness of UK and NIGERIAN women to soy compared to their counterparts- men. Additionally, it was found that there was no age disparity during soy consumption- Days 3 & 5 (Phase II and III) but otherwise in the pre- and post- consumption (Phase I and IV) respectively.

On the whole, based on the different clustering and separation of groups, it can be noted that the investigated influencing factors- location and gender play vital roles in soy metabolism. This was similarly reported in epidemiological studies by [intro ref] with the disparity in soy-responders following location differences and migration effects.

4.3.4.2 Target profiling & binning

Generally, target profiling is an efficient and valuable method for identifying and quantifying metabolites with the aid of a defined metabolome library. Following target profiling for the quantitative analysis, target binning was employed for chemometrics. This is proposed to be a better binning technique than whole spectrum binning because it is more selective, specific and precise, where other un-profiled artefacts in the spectrum regions are excluded prior to chemometrics. In addition, there is a compensation of possible sample changes due to water suppression, baseline, pH/ ionic strength and binning region/ dimension.

On the other hand, one of the few disadvantages is that only the profiled metabolites are correlated as observed in the loading plot. In point of fact, there is no potential of discovering unlisted potential biomarker(s) causing the differences in metabolic effects. Other disadvantages include: prior knowledge of targeted metabolites of interest, availability or build-up of metabolites library and slightly time-consuming to profile/ quantify each metabolite.

In this study, unsupervised PCA was carried out as the preliminary discrimination test on the x- variable buckets of dataset. Samples were inputted blindly with the aim of achieving

discriminatory cluster groups based on key factors including environmental factors. Possible descriptive factors used for discrimination include location, gender, age, smoking status, diet and weight (or BMI). Following the basic test by PCA, whereby group clusters are not sufficiently defined and classed, another MVA used was PLS-DA, a supervised classification technique. In PLS-DA, qualitative y- variables used were the classes, i.e. either hypothesised factors classes or interventional Phase classes. This classification is theoretically expected to be a better discriminatory analysis. Following the calculation of PLS-DA, orthogonal-PLS-DA (O-PLS-DA) was calculated on the data set, for the cogent interpretation of hypothesised discriminatory factors.

In addition to the examination of the models' score plot and loading plot, the model overview plot is important in showing how well the model fits and predicts data in terms of R^2 and Q^2 for each principal component. From the results, it showed that target binning was a better binning technique compared to standard whole spectrum binning.

4.4 Conclusions

With the aim of quantitation of endogenous metabolites that are found to be vital in major metabolism- energy, fat and protein, twenty-two metabolites in African subjects' urine were simultaneously assessed with the aid of Chenomx Suite. These metabolites' levels at the baseline (before soy - intervention) were found to be comparable to the only other reported similar study ranging from 5 μM to 15 mM. Of these metabolites: acetone, citrate, creatinine, glycine, hippurate and trimethylamine-N-oxide and urea showed significant variances in the whole/ general cohort. Prior to further analysis, data taxonomy was performed on the dataset based on proposed factors influencing soy metabolism. Of the proposed factors, two main factors showed to influence differences in individual metabolism following soy-intervention. These two factors were location (Nigeria and UK cohorts) and gender (male and female groups).

Following this on the basis of these quantified amounts, it was observed that with soy consumption, there was a reduction in glycolysis, as a result of decrease in lactate and pyruvate. Also, there was a reduction in collagen as a result of reduced levels of glycine and alanine following soy consumption (Days 3 & 5 - Phases II & III/ during).

In the other investigation of generalised metabolites analysis using multivariate analysis, it showed that target binning was a better clustering or discriminatory analysis technique

compared to the standard whole spectrum analysis. Further to this, of all the MVA tools, orthogonal-partial least square- discriminatory analysis (O-PLS-DA) showed to be the best expressive technique to differentiate the proposed theories of influencing factors. Repeatedly, location and gender seemed to be the main influencing factors as shown in MVA.

The observations made in this study with regards to influencing factors suggest that the use of other factors have potential to be useful for more understanding of soy metabolism by *in-vivo* relationship of phytoestrogen metabolites on endogenous metabolites.

4.5 Experimental

4.5.1 Chemicals and materials

4.5.1.1 Chemicals

Monobasic sodium phosphate, dibasic sodium phosphate, trimethylsilyl-2,2,3,3-tetradeuteropropionic acid (TSP) and deuterated water (D_2O) were purchased from Fisher Scientific, UK. Distilled water was purified with Millipore Simplicity 185 (Millipore Corp., USA) at 18.2 $m\Omega.cm^{-1}$ polishing units.

4.5.1.2 Material: descriptive information of interventional diet

Alpro soya-milk was procured from local supermarkets for both cohorts. Both subject groups (UK and Nigeria) consumed 200 mL Alpro soya-milk at the each day for five consecutive days. The serving of cholesterol and lactose-free Alpro-soya milk in compliant with guideline daily allowance (GDA) contained approximately 8 mg daidzein, 10 mg genistein, 84 kcal (354 KJ) energy, 6.6 g protein, 5.6 g sugar carbohydrate, 0.6 g saturated fat, 0.8 g mono unsaturated fat, 2.4 g poly unsaturated fat (2.18 g omega 6; 0.28 g omega 3), 1.2 g fibre, 0.3 g salt (0.1g sodium), 240 mg calcium, 0.48 mg vitamin B2 and 0.3 µg vitamin B12 [manufacturer's specifications].

4.5.2 Design of experiment: subjects and sample collection

The design of experiment (DOE) of this metabonomic study following soy intervention was based on a dynamic-sampling type study. Its importance is to maximise the information

content, thereby increase the chances of capturing all possible variation of responses and avoid biased evaluation of individual, intra- and inter- observations (*219*).

Ethical approval for this research study was reviewed and approved by the London Metropolitan University human research ethics committee and all procedures complied with National Health and Medical Research Council standards. Two cohorts of twenty five African subjects (UK and NIGERIA; males and females) were recruited to participate in the soy-interventional study. To control this study, it was ensured that these two African cohorts (UK and NIGERIA) were all residents in their two different locations for at least two years prior recruitment. As compliant to the ethics, all subjects were interviewed and signed informed consents were obtained from all participants before commencement of the study and all participant samples and information were confidential and remained completely anonymous. All factors that could cause variations in results e.g. age, BMI, lifestyle habits (tobacco, drinking, social, etc.) were noted; however no inclusion/exclusion criteria were followed prior sampling analysis. In addition, four control subjects were enrolled: two Asians (UK) and two >65 yrs old Africans (NIGERIA). A summarised table illustrating this distribution is shown in Table 4.7. Of these recruited subjects, five subjects did not complete the study due to personal reasons.

On the induction day and a week after intervetion, both subject groups (UK and Nigeria) supplied spot urine samples collected as the pre- and post- (Day 1-Phase I and Day 12-Phase IV) samples respectively. During the consumption of Alpro soya-milk at the guideline daily allowance (GDA) of 200 mL each day for five consecutive days, two different early morning urine samples were collected (Days 3 & 5/ during phases II & III). All of these samples were stored without preservatives at -20°C or -80°C.

TABLE 4.7: Characteristics of study groups

	UK group	Nigeria group
No of subjects	25	27
Demographic selection		
African	23	27
Non-African (Asian)- CONTROL	2	-
Gender (Female/male)	12/13	12/15
Age Ranges (years)		
20-25	12	13
25-30	6	10
30-35	5	-
35-50	2	2
>60 - CONTROL	-	2
BMI (kg/m^2) *		22.02 ± 3.788
Range		(17.53 – 32.89)
Smoking Status (non-smoker/ smoker)		17/10
Diet Status (non-vegetarian/ vegetarian)	21/4	27/0

BMI (kg/m^2)* expressed as mean ± standard deviation

4.5.3 Sample preparation

0.1 M buffer solution (pH 7.4) was prepared in 30:70 (D_2O:water) mixture containing 1.44 mM of reference trimethylsilyl-2,2,3,3-tetradeuteropropionic acid (TSP).

To thawed-out samples, 600-µL aliquot of urine sample was combined with 300-µL phosphate buffer solution in 0.5-mL polypropylene Microfuge tube (Fisher Scientific, UK). Following this, the buffered urine solution was left to stand at room temperature for 10 minutes before being subjected to centrifugation. Centrifugation was carried out at 13,000 g for 10 min to homogenize solution using Eppendorf Centrifuge 5804R (Germany) operating at 4ºC. 850-µL aliquot of the supernatant was transferred into a 5-mm NMR tube for ^{1}H-NMR analysis. This sample preparation procedure was carried out in triplicate on each urine sample.

4.5.4 Analysis: ^{1}H-NMR spectroscopy

All ^{1}H-NMR spectroscopic analyses were measured in buffer-deuterated water (D_2O) mixture employing a Bruker® Avance AV-500 NMR spectrometer. Prepared urine samples in 5mm-NMR tubes were analysed in this spectrometer operating at 500.13 MHz and a temperature of 303 K. The acquisition type used for the collection of the ^{1}H-NMR spectrum was by standard pulse sequence for water peak suppression obtained using 1D-NOESY pulse sequence with water pre-saturation (*341*)]. The sequence used [RD-90°-t_1-90°-t_m-90°-FID], where recycle delay (RD) is 2-s, fixed interval (t_1) is 10-µs and tm is 100-ms.

All spectra were acquired at 14.0019 ppm (7002.801 Hz) sweep width, 2.34-sec acquisition time and 128 transients. FID spectra were zero filled to 32 k data points, exponential 0.3 Hz line broadened, Fourier transformed and manually fine-tuned (phase and baseline correction) using the pre-installed computer-aided software for data acquisition and analysis was Bruker Topspin® software- version 2.0 (Bruker Biospin®, Rheinstetten, Germany).

4.5.5 Identification and quantitation of endogenous metabolites

Acquired ^{1}H-NMR spectra were exported into the Chenomx® NMR suite 5.1 for general identification and quantitation of present metabolites in the urine. This development is been supported by the complementary database containing almost 300 metabolite compounds (standards). Exported spectra are further fine-tuned to enable good mathematical modelling of the metabolites in the samples with pure standards compounds spectra in the library.

With regards to metabolite quantitation, each Lorentzian peak shape models of reference compounds in the library were individually superimposed upon supposed representative peaks of the sample NMR spectrum. This deduction was based on both the exact chemical shift and the splitting pattern as shown in FIGURE 4.21. For example,

FIGURE 4.21: Example of superimposition of Lorentzian peak shape models of reference on raw ^{1}H-NMR data.

With the aid of the TSP reference, quantification of metabolites was possible based on the known concentration of TSP. Each pure compound (standard) in the database had been fit to record peak centres and homonuclear J-coupling constant at pH 7. Quantitation in the Profiler module was established by individual chemical shift and splitting patterns (multiplicities). Basically, identification and quantitation was carried out by overlapping Lorentzian peak shape model of each standard against the exported tuned spectra and visually ensuring these peaks are at the same level on y- (concentration) axis.

Using Chenomx® NMR suites 5.1 Profiler, metabolites listed in the compiled library of 22 endogenous metabolites have two concentrations each: actual and potential. The potential concentrations are calculated and generated by this software, whilst the actual concentrations are the concentrations calculated by the examiner. Usually these two concentrations are expected to be close in value (also shown) especially for simple patterned signals (FIGURE 4.22).

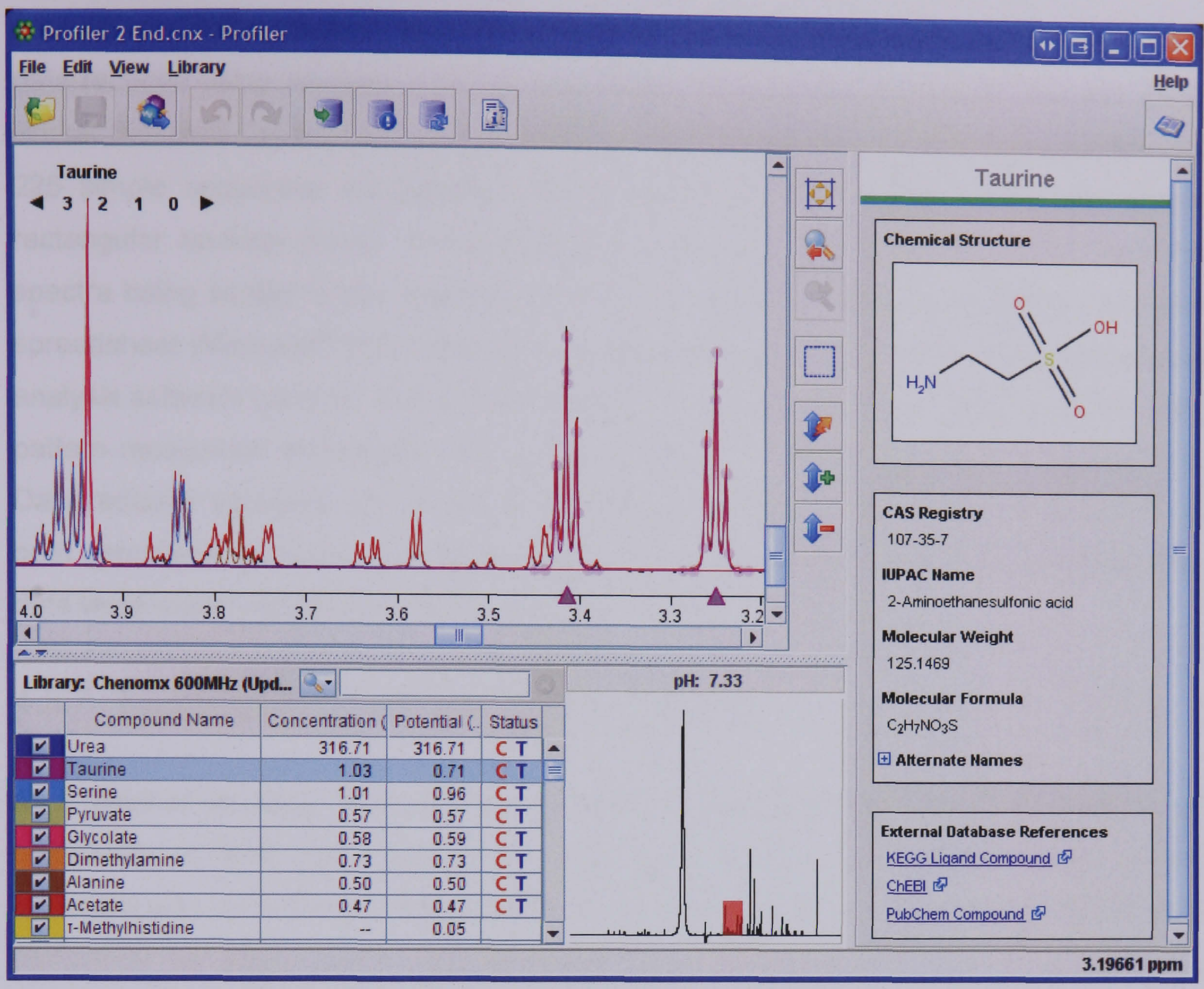


FIGURE 4.22: Typical Chenomx® Profile page

The selection of the targeted metabolites were established on their vast representation of different metabolic biochemical reactions in the body, e.g. in blood, liver and muscles. Further to this, the selection of these metabolites was also based on their simple ^{1}H-NMR chemical shifts and multiplicity, thus making quantitation easier with minimal misinterpretation errors. In this analysis, an exact concentration of TSP was used as a reference in all samples with the aim of matching the NMR spectroscopic conditions and outputs. The other function of using TSP was for spectrum signal alignment, due to its proton zero- (approximately) chemical shift.

4.5.6 Data reduction (binning) prior multivariate statistical analysis

In this study, two main types of binning were considered: target binning and whole spectral binning. Target binning is a new advanced technique where only identified and quantified (in this case by Chenomx® Profiler), signals/ peaks are data-reduced into customised varied width sized bins prior multivariate statistical analysis.

Alternatively in the standard spectral binning, acquired ^{1}H-NMR spectra were sequentially data-reduced using Analysis of Mixtures software package (Bruker AMIX®, Version 3.6.8, Bruker Biospin® GmbH, Germany) across the ^{1}H-NMR regions δ 0 - 9.0. This resulted into 225 simple sequential rectangular integral segments of 0.04 ppm width-sized simple rectangular buckets (bins). The integration mode used was sum of intensity with the spectra being scaled to the reference (TSP). The resulting data was exported into Excel spreadsheet (Microsoft® 2007) due to its compatibility with SIMCA –P+® multivariate data analysis software (version 12.0.0.0, Umetrics, Umeå, Sweden). Multivariate analysis using pattern recognition techniques were performed using standard SIMCA-P+® applications. Data-reduced variables are imported from Excel (Microsoft® 2007) into SIMCA-P+®. In both generated bin types, i.e. target profiled and whole spectral computer-generated bin, data were mean centred and Pareto scaled for the MVA.

4.5.7 Multivariate statistical analysis (MVA)

As a result of the large observation size, subtle metabolic changes due to physiology, and complexity of NMR data, there is the need for a sensitive and selective form of data dimensionality reduction for data analysis, thus leading to MVA. Two classes of MVA were performed on both binned data sets, these are unsupervised and supervised. The unsupervised statistical analysis was by Principal Component Analysis (PCA), whilst the supervised statistical analysis was by Partial Least Square –Discriminatory Analysis (PLS-DA) and orthogonal Partial Least Square – Discriminatory Analysis (O-PLS-DA).

PCA is a pattern recognition technique, which generates a model showing its graphical projection, and summarises displays (plots) of the correlation of observations and their variables. In unsupervised MVA (PCA), the data buckets are calculated as x-variables with no additional information as y-variables, resulting to a 'blind' classification based on their correlation of observations (shown in score-plot) or covariance of x-variables only (shown in loading plot). On the other hand, PLS-DA further classifies based on two variables: y- and x- variables, where the y- variables could be either qualitative 'dummy' or quantitative 'time-related, dose-related'. The 'dummy' y-variables created for these analyses where based on the different classes to differentiate, e.g. four classes of Phases I-IV, two classes of locations, two classes of genders, two classes of age groups or combinations- four classes of location and gender. Further-to-this, O-PLS-DA is more advanced than PLS-DA, where an addition set of y-variables, orthogonal (unrelated) to the other variables, is computer generated.

The score-plot is a summarised correlation plot of spectrum 'observation', whilst, it's complementary plot: loading plot shows the summarised relationship between the variables, whereby using x-variables only in PCA and x-, y- variables in PLS. Usually, the patterns are similar in both score and loading score-plots. Each observations or variable is represented in the score-plot or loading score-plot respectively as a point, whereby the closeness shows similarities. Additionally, the loading plot showed the potential ^{1}H-NMR regions responsible for clustering and discrimination, thus, the detection of potential biomarkers.

MVA applications were employed to explore the clustering and discriminatory behaviours of the observations based on different influencing factors as mentioned earlier, location, gender and age. Generally, all samples or specific samples were included in the various analyses for model generation, whereby outliers were excluded based on Hotelling's test (MVA Student t-test). No variable or spectral region was excluded except the TSP region (δ: -0.05 to 0.05).

Chapter V

Correlation-ship of metabolites

5.1 Introduction

The observations and complex results of metabolites levels both soy - phytoestrogen metabolites and endogenous metabolites reported in Chapters III and IV prompted an investigation into the relationship between these compounds. The study was a combination analysis of class specific and dynamic analyses. The soy metabolite analysis was based on a class specific analysis whilst the endogenous metabolites' was a dynamic sampling analysis. Since most of the published reports are only partially or not understood in regards to a linkage between these two groups of metabolites, thorough-correlations of these metabolites groups is needed to be achieved to acquire a better insight on soy metabolism.

The aim of this chapter is to attempt exploring the urinary excretion dependency/ interactions of soy-metabolites and endogenous metabolites. To date, no report has observed the correlation of soy-metabolites and/ or endogenous metabolites following long-term or short-term soy - phytoestrogen intervention.

Before going into details of the correlation-ship, it is worth-while to recall the analysed soy metabolites and the endogenous metabolites. The soy-metabolites considered for this relationship are daidzein, genistein, equol and o-DMA. Whilst, the twenty-two endogenous metabolites include: acetate, aceto-acetate, acetone, alanine, betaine, choline, citrate, creatine, creatine-phosphate, creatinine, dimethylamine, formate, glycine, hippurate, lactate, methylamine, dimethylglycine, pyruvate, succinate, trimethylamine, trimethylamine-N-oxide and urea.

5.2 Results & Discussion

5.2.1 Correlation-ship in soy - phytoestrogen metabolites

On the basis of investigating correlation effects, the selected results applied for these analyses were the urinary excretion levels of samples collected during the soy consumption (Days 3 & 5) only. Pearson correlation was used on the quantitative results of urinary excretion levels of four soy - phytoestrogens (daidzein, genistein, equol and o-desmethylangolensin) and twenty-two endogenous metabolites, there were noticeable correlations within either group of compounds.

Prior to correlation calculations, urinary excretion levels of metabolite changes were examined rather than individual levels. Considering the correlation of soy - phytoestrogens and their metabolites as shown in TABLE 5.1, there were only two significant positive linear correlations between genistein and daidzein ($r = 0.460$, $p = 0.007$) and o-DMA and daidzein ($r = 0.421$, $p = 0.015$), whilst the other metabolites showed no significant positive correlation ($r < \sim 0.300$, $p > 0.080$).

TABLE 5.1: Pearson correlation of Δ soy - phytoestrogens and their metabolites (General cohort): Day 3

	Daidzein	**Equol**	**Genistein**
Equol	0.085; 0.639		
Genistein	0.460; 0.007	0.299; 0.091	
o-DMA	0.421; 0.015	0.014; 0.937	0.301; 0.088

Pearson Correlation; p-value

With the result obtained as seen above shows that generalised correlation did not show much distinct linear correlation. Thus, the classification adopted in previous *Chapter 3* and *4* were implemented resulting to TABLE 5.2.

TABLE 5.2: Pearson correlation of Δ soy - phytoestrogens and their metabolites (Location Classified)

	UK cohort Pearson Correlation r; p-value			NIGERIA cohort Pearson Correlation r; p-value		
	Daidzein	**Equol**	**Genistein**	**Daidzein**	**Equol**	**Genistein**
Equol	0.290; 0.295			-0.028; 0.913		
Genistein	0.631; 0.012	0.319; 0.247		0.426; 0.078	0.293; 0.238	
o-DMA	0.450; 0.092	-0.042; 0.883	0.526; 0.044	0.481; 0.043	0.046; 0.858	0.258; 0.301

TABLE 5.2 showed that the genistein and daidzein correlation (positive) observed in the general cohort was mainly influenced by the UK cohort (r = 0.631; p-value = 0.012), whilst the other correlation of o-DMA and daidzein was observed with the NIGERIA cohort. Additionally, o-DMA and genistein showed a significant positive correlation (r = 0.526, p = 0.044) in the UK cohort only. Following this, the other metabolites in either location cohort showed no significant correlation (p > 0.05). In support of the statistical values, the matrix-plot displayed in FIGURE 5.1 shows the linearity or non-linearity pattern of results.

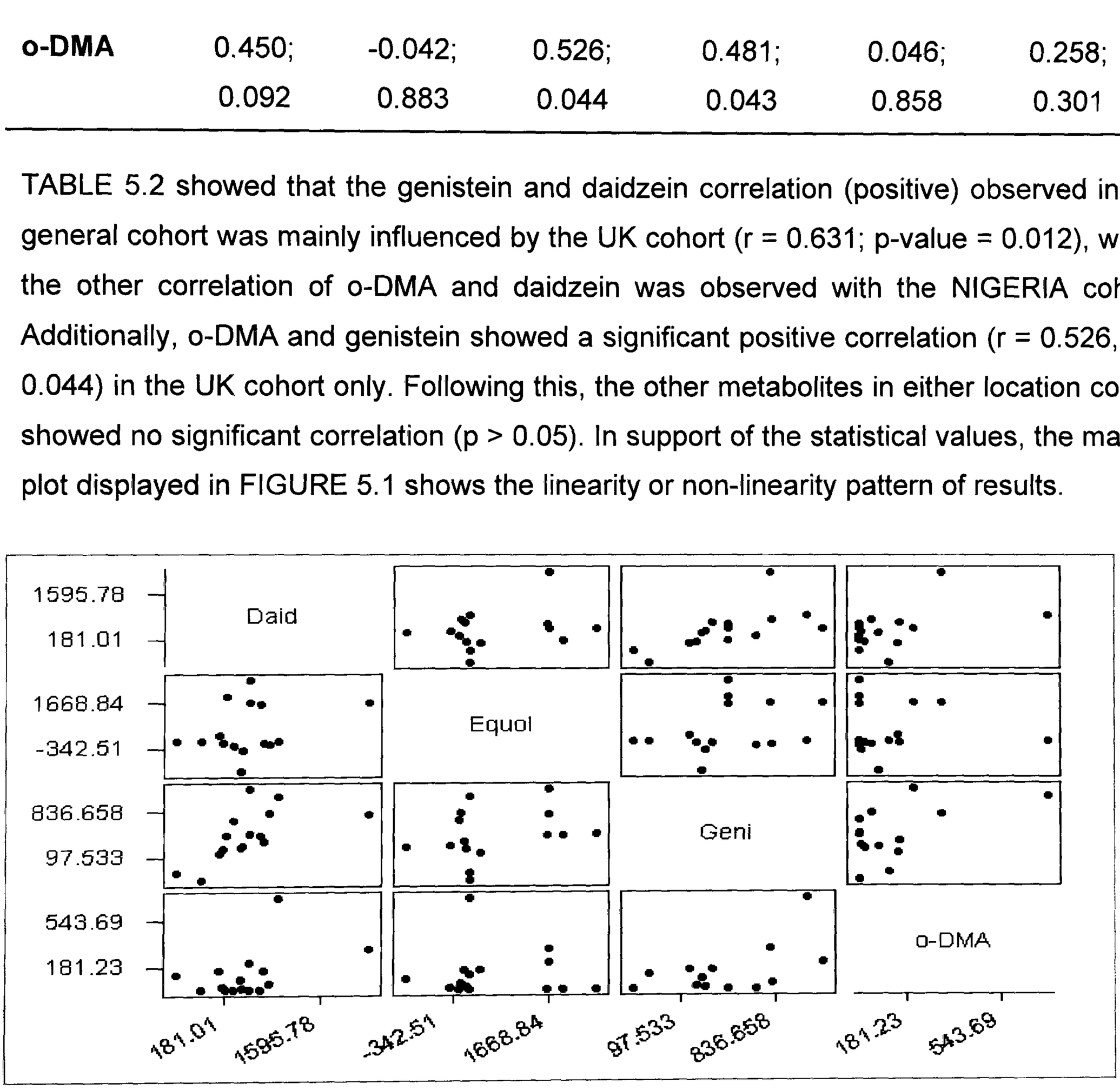


FIGURE 5.1: Matrix-Plot of soy - phytoestrogens and their metabolites within UK cohort

Specifically, there was a distinct linear pattern with one/ two outliers observed in genistein-daidzein plot, whilst the others showed no correlation pattern.

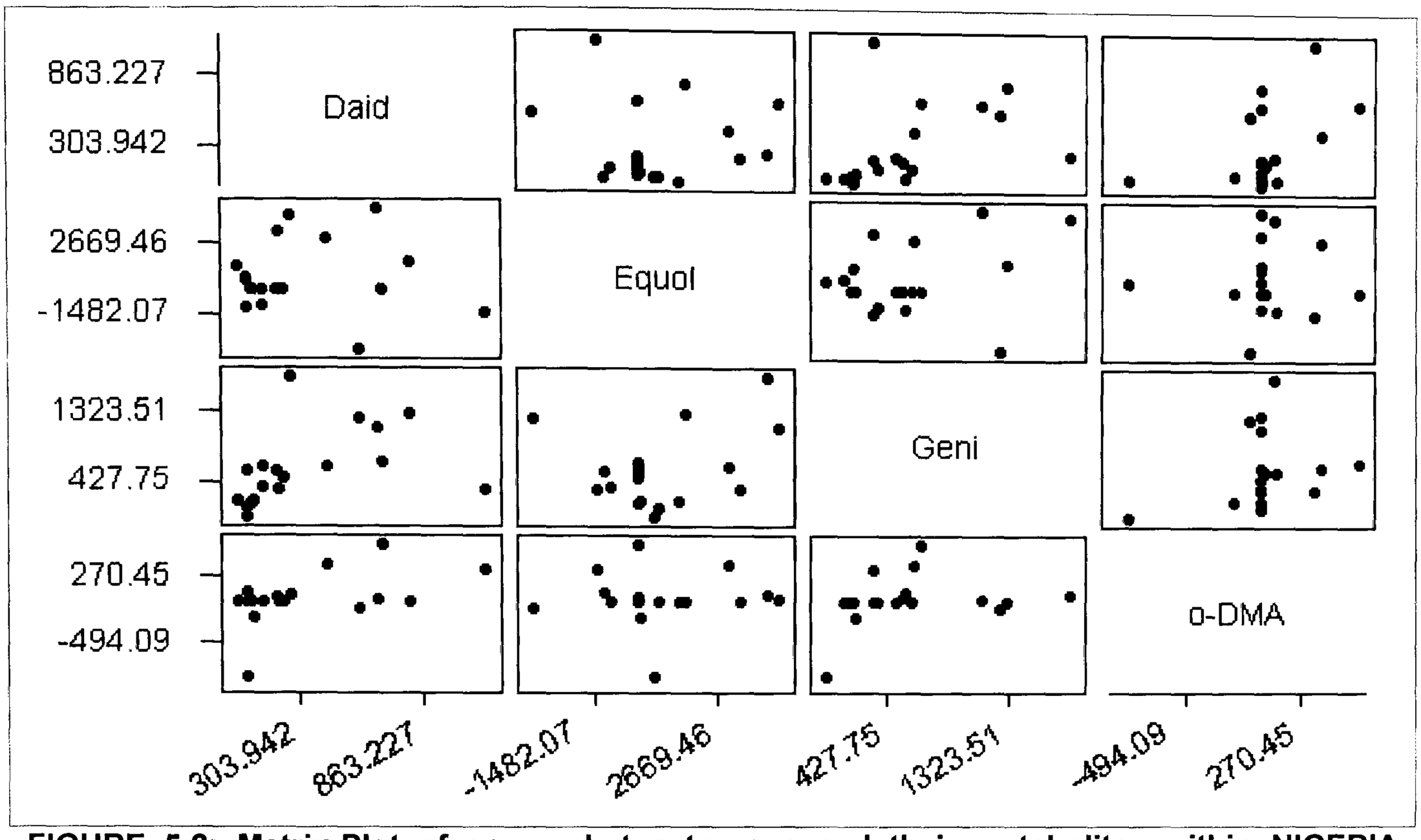


FIGURE 5.2: Matrix-Plot of soy - phytoestrogens and their metabolites within NIGERIA cohort

Comparing the two location-based cohorts (FIGURES 5.1 & 5.2), the UK cohorts showed to have a better significant linear correlation than the Nigeria cohort. Following this location classification of the whole cohort, in the UK cohort, the only significant linear correlation was observed in the female group, which showed a correlation of daidzein and genistein with $r = 0.743$ and $p = 0.035$ compared to that of the male with $r = 0.417$, $p = 0.304$.

Whilst in the Nigerian cohort, the male subjects showed a positive strong linear correlations was obtained with genistein- daidzein with $r = 0.827$ and $p = 0.011$. In addition to this (NIGERIA - male), a good negative correlation was achieved with genistein and equol ($r = -0.676$), however this was found to be statistically insignificant ($p = 0.066$). On the contrary, the Nigerian female group showed strong positive significant linear correlation with genistein and equol only ($r = 0.796$, $p = 0.003$).

Similar to the correlation analyses performed on the metabolite urinary level Δ on Day 3 (in comparison to Day 1), Day 12 (post-soy) changes were analysed for correlations. The general cohort was considered for preliminary correlation test as shown in TABLE 5.3. This showed that there were significant correlations in three pairs of metabolites: genistein and daidzein, o-DMA and daidzein and o-DMA and genistein ($r > 0.4$, $p < 0.02$).

TABLE 5.3: Pearson correlation of Δ soy - phytoestrogens and their metabolites (General cohort): Day 12 (post-soy)

	Pearson Correlation r; p-value		
	Daidzein	**Equol**	**Genistein**
Equol	0.196; 0.266		
Genistein	0.876; 0.000	0.202; 0.253	
o-DMA	0.432; 0.011	-0.033; 0.854	0.518; 0.002

Again, the correlation between the soy - phytoestrogens and their metabolites were investigated based on classification followed in Chapters III and IV. Firstly, as location was considered to be one of the criteria for the metabolism, this was tested for correlation-ship of phytoestrogens and their metabolites.

TABLE 5.4: Pearson correlation of Δ soy - phytoestrogens and their metabolites (Location Classified): Day 12 (post-soy)

	UK cohort Pearson Correlation r; p-value			**NIGERIA cohort** Pearson Correlation r; p-value		
	Daidzein	**Equol**	**Genistein**	**Daidzein**	**Equol**	**Genistein**
Equol	0.748; 0.001			0.322; 0.178		
Genistein	0.943; 0.000	0.760; 0.001		0.551; 0.014	0.224; 0.357	
o-DMA	0.647; 0.007	0.287; 0.281	0.560; 0.024	-0.095; 0.698	-0.056; 0.820	0.445; 0.056

The results obtained showed that the UK cohort showed better overall correlation-ship than NIGERIA cohort. In other words, there were five out of a possible six correlation pairs of metabolites in the UK cohort ($r > 0.55$; $p < 0.03$) compared to only one significant correlation in the NIGERIA cohort ($r = 0.55$; $p = 0.014$) with daidzein and genistein. The correlations in the UK cohort include: equol and daidzein, genistein and daidzein, genistein and equol, o-DMA and daidzein and o-DMA and genistein.

Although, the statistical values showed there are correlation-ships with five paired metabolites, however the positive slopes indicating the linear associations between some of these paired metabolites, specifically genistein and equol & o-DMA and daidzein showed not to be normally distributed as shown in FIGURE 5.3. Thus correlation coefficients statistically generated may not be relevant. Whilst equol and daidzein showed to be a fairly distributed dataset, however there were obvious outliers present which possibly influenced the correlation.

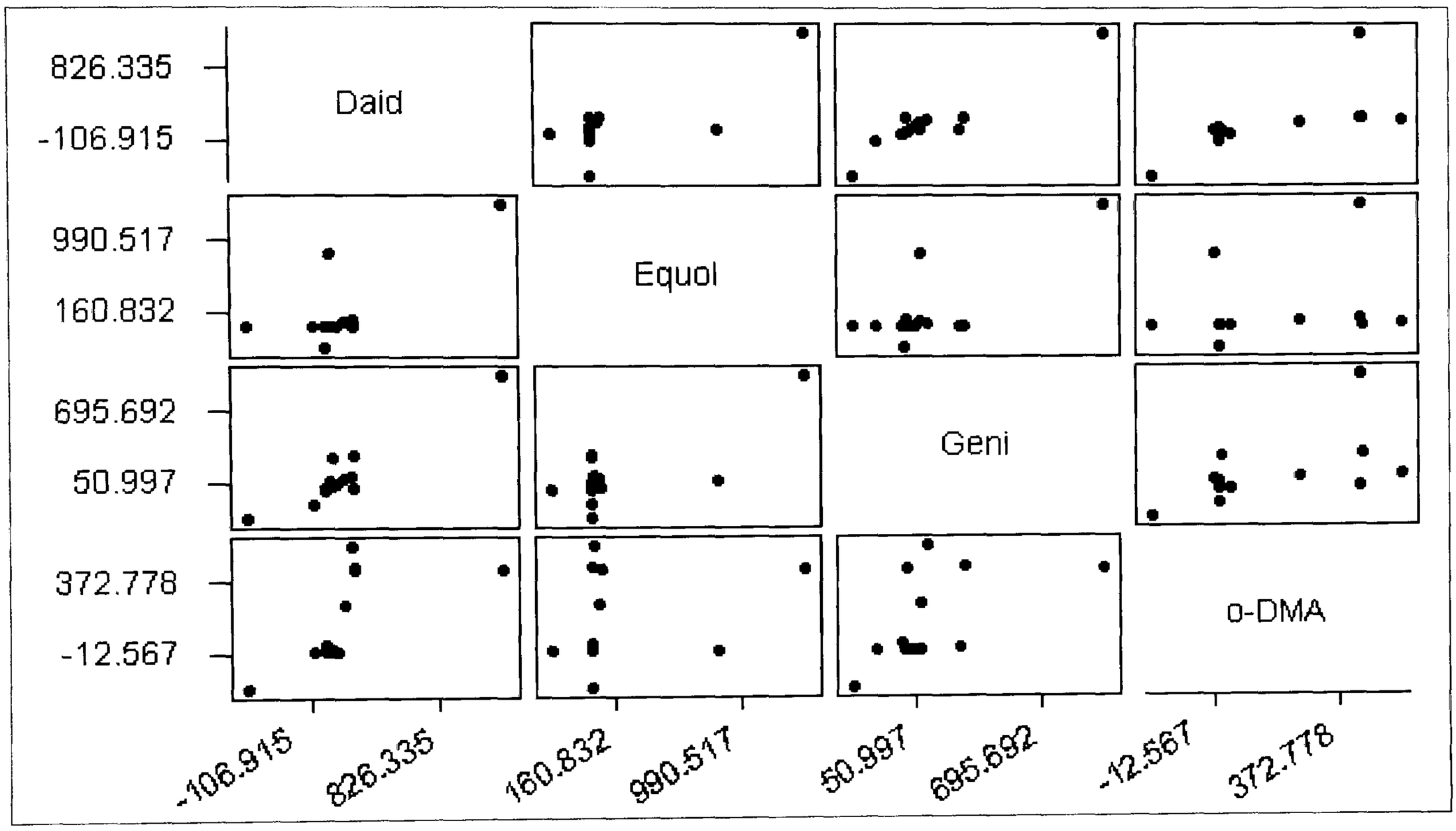


FIGURE 5.3: Matrix-Plot of soy - phytoestrogens and their metabolites within UK cohort- Day 12 (post-soy)

Further to this location - classification on the Day 12 (post-soy) samples, gender - classification was also considered for correlation analyses.

With UK - female group, the only significant correlation found was with o-DMA and daidzein with $r = 0.841$ and $p = 0.009$, which on further examining the matrix-plot showed the variables were not linear but the r value was greatly influenced by an outlier (FIGURE 5.4).

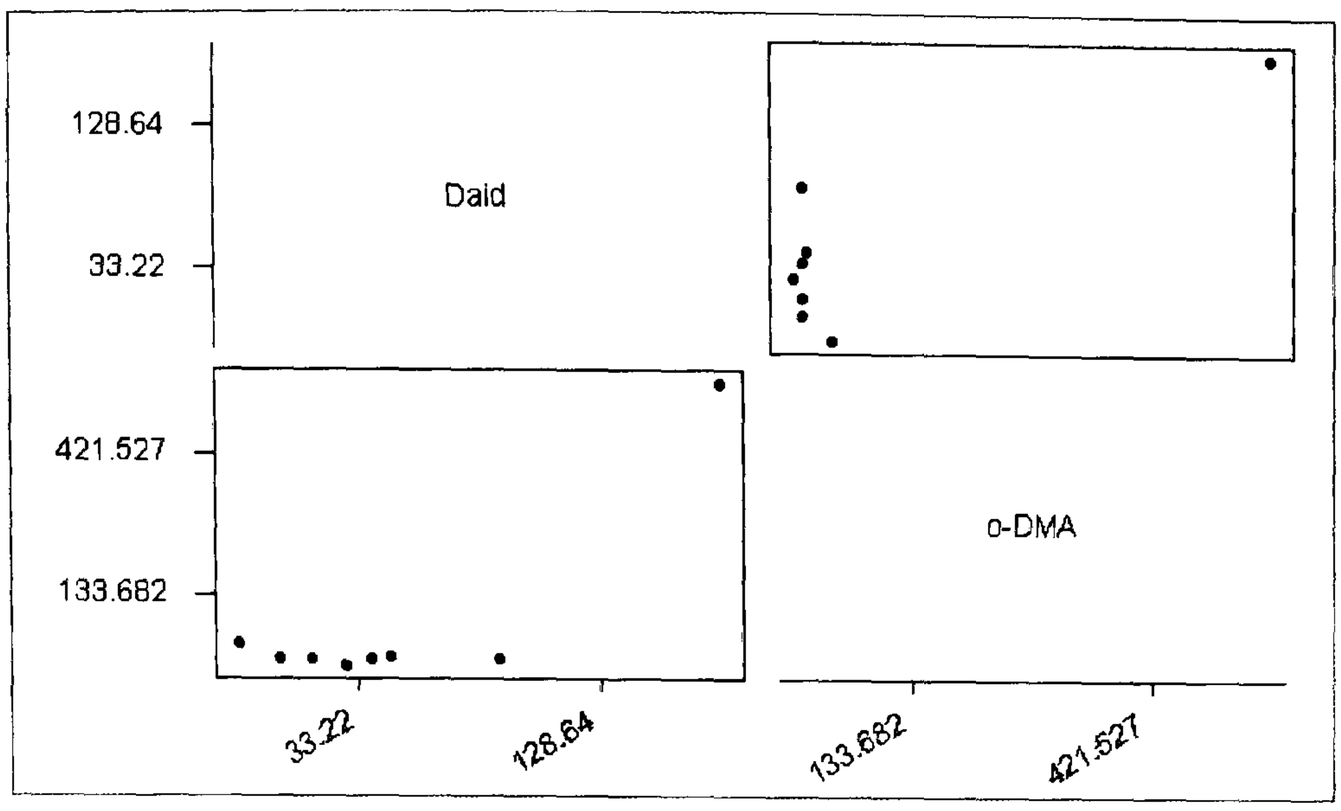


FIGURE 5.4: Matrix-Plot of o-DMA and daidzein within UK-female group- Day 12 (post-soy)

Similarly in FIGURE 5.5 & TABLE 5.5, considering the matrix plot and statistical values, UK-male group showed to have good correlation with daidzein-genistein, o-DMA-genistein and daidzein-o-DMA, whereas, equol correlation to other metabolites seemed to be non-linear relationships.

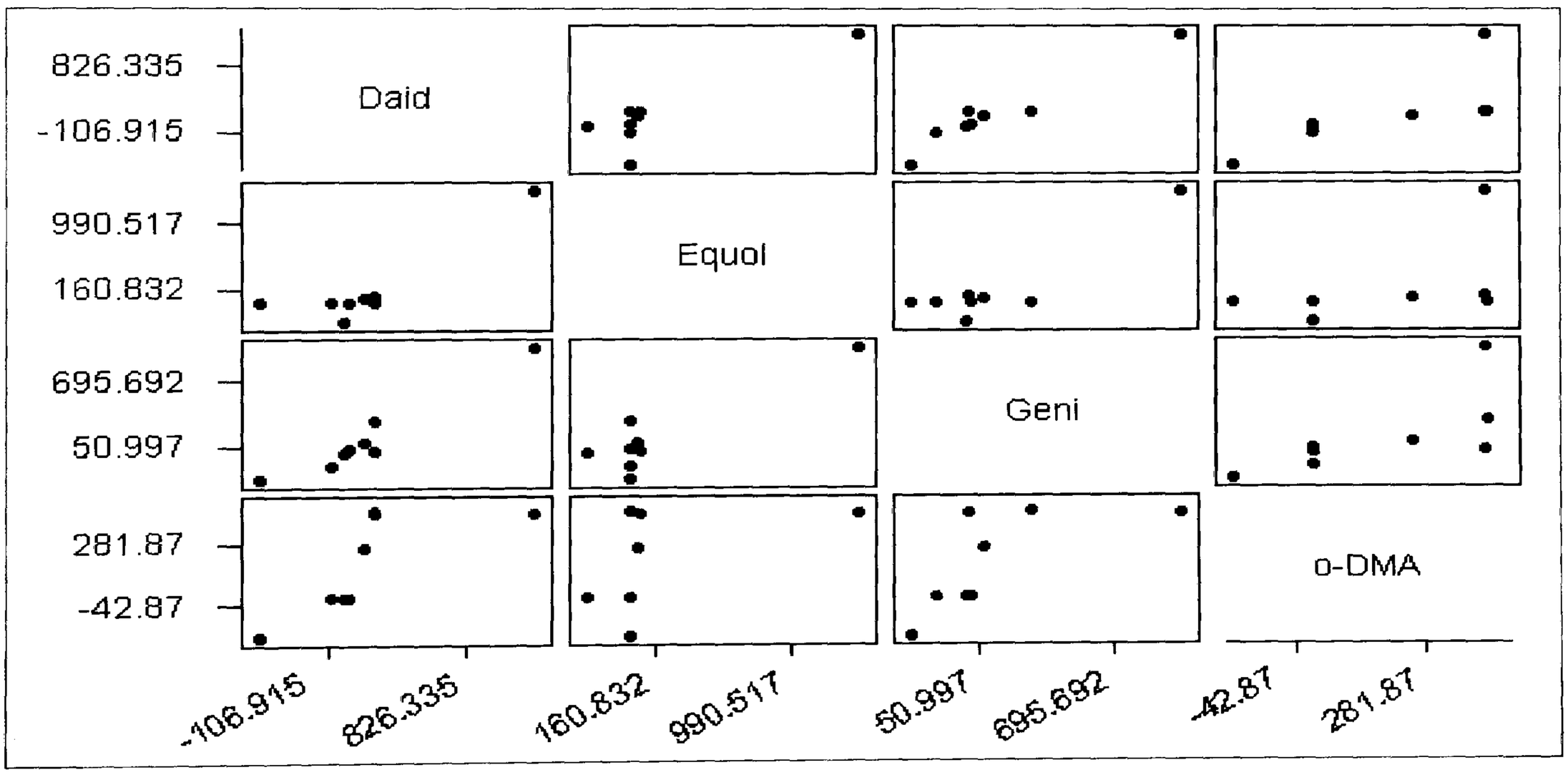


FIGURE 5.5: Matrix-Plot of soy - phytoestrogens and their metabolites within UK-female group- Day 12 (post-soy).

TABLE 5.5: Pearson correlation of Δ soy - phytoestrogens and their metabolites (UK-Male group): Day 12 (post-soy)

	Daidzein	Equol	Genistein
Equol	0.875; 0.004		
Genistein	0.965; 0.000	0.895; 0.003	
o-DMA	0.740; 0.036	0.479; 0.229	0.689; 0.059

Pearson Correlations; p-value

With the NIGERIA-female group, the only significant correlation was daidzein and genistein (r = 0.685, p = 0.014) and with the matrix-plot (FIGURE 5.6), there was a true linear relationship between both variables.

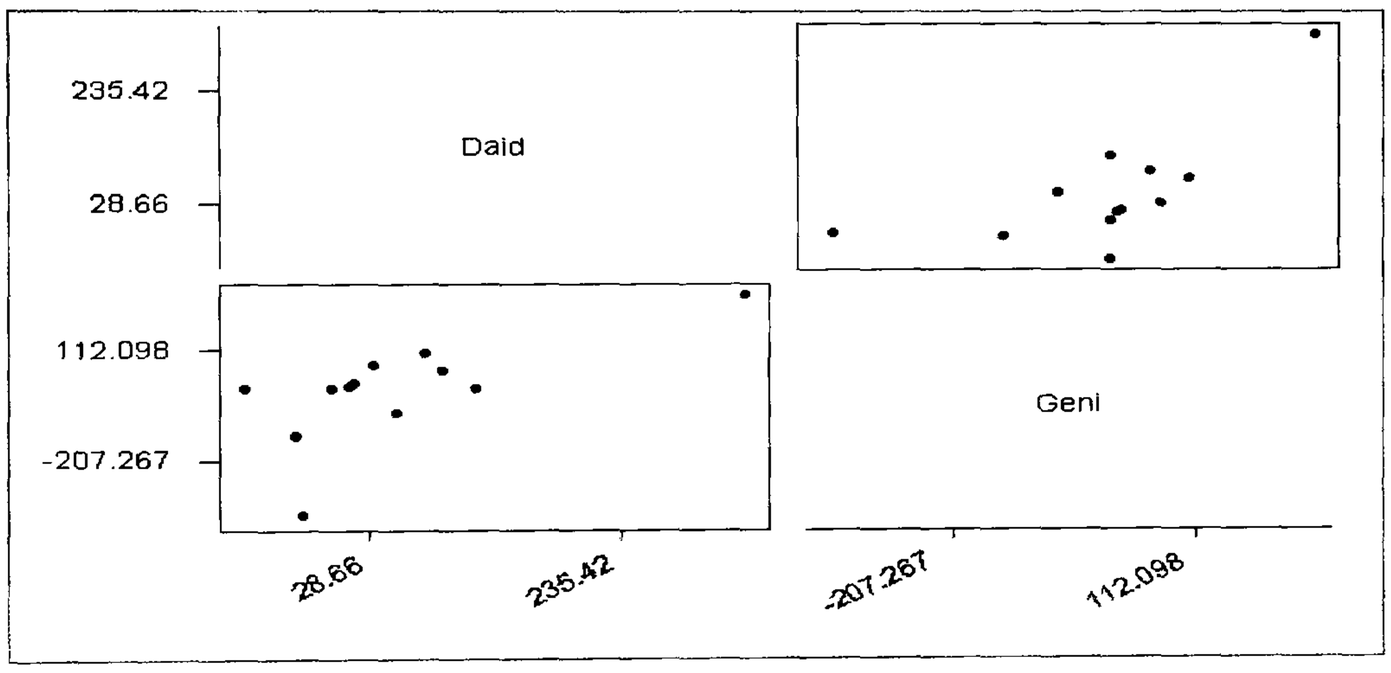


FIGURE 5.6: Matrix-Plot of o-DMA-daidzein within NIGERIA-female group- Day 12 (post-soy)

With NIGERIA-male group, the only significant correlation was found to be negatively influenced, i.e. r = -0.908 (p = 0.005) in o-DMA-genistein. However, its matrix-plot showed correlation coefficient (relationship slope) to be biased by an outlier. Of all the published reports available, most of the correlation-ships reported are between urinary excretion and serum levels (*317, 320*), however in this study, inter-relationships were aimed to analyse soy - metabolite correlation-ship.

5.2.2 Pearson correlation in endogenous metabolites

Following the test of Pearson correlation with soy - phytoestrogens and their metabolites, endogenous metabolites were also considered. In this chapter, correlation in twenty-two endogenous metabolites (assessed in *Chapter 4*) were analysed and reported following a pilot test. A pilot test was first performed for screening purposes and this showed there was no significant relationship between acetone, hippurate and the other metabolites.

On performing Pearson correlation test on the endogenous metabolites, it showed that there were significant linear correlations in different metabolites with $r > 0.05$ and $p < 0.05$.

In support of the statistical values, TABLE 5.6 shows the overall statistical Pearson correlation amongst 22 endogenous metabolites in the general cohort on Day 3 (Phase II/ during-soy).

T

	Acetate	Acetoace	Acetone	Alanine	Betaine	Choline	Citrate	Creatine	Creatine Phosp	Creatini	Dimethyl Amine	Formate	Glycine	Hippurat	Lactate	Methylam	N,N-Dime Glycine	Pyruvate	Succinat	Trimethy Amine	Trimethy N-oxide
Acetoace	-0.172 0.339																				
Acetone	0.292 0.100	-0.100 0.580																			
Alanine	**0.430** **0.012**	**0.376** **0.031**	0.058 0.748																		
Betaine	0.127 0.481	**0.586** **0.000**	-0.045 0.805	**0.707** **0.000**																	
Choline	0.176 0.327	0.154 0.391	-0.167 0.353	**0.502** **0.003**	**0.365** **0.037**																
Citrate	0.249 0.163	0.288 0.104	0.180 0.317	**0.585** **0.000**	**0.638** **0.000**	**0.416** **0.016**															
Creatine	0.114 0.528	**0.428** **0.013**	0.180 0.317	**0.535** **0.001**	**0.527** **0.002**	0.176 0.327	**0.390** **0.025**														
Creatine Phosp	**0.409** **0.018**	**0.349** **0.046**	-0.012 0.948	**0.838** **0.000**	**0.682** **0.000**	**0.412** **0.017**	**0.497** **0.003**	**0.738** **0.000**													
Creatini	**0.363** **0.038**	**0.459** **0.007**	0.163 0.366	**0.881** **0.000**	**0.678** **0.000**	**0.514** **0.002**	**0.667** **0.000**	**0.701** **0.000**	**0.824** **0.000**												
Dimethyl Amine	**0.404** **0.020**	**0.402** **0.020**	0.151 0.401	**0.795** **0.000**	**0.664** **0.000**	**0.528** **0.002**	**0.782** **0.000**	**0.498** **0.003**	**0.664** **0.000**	**0.908** **0.000**											
Formate	**0.684** **0.000**	0.115 0.524	0.238 0.182	**0.489** **0.004**	0.181 0.314	0.199 0.266	0.178 0.320	0.047 0.794	0.309 0.081	**0.406** **0.019**	**0.433** **0.012**										
Glycine	**0.394** **0.023**	**0.429** **0.013**	0.063 0.726	**0.937** **0.000**	**0.702** **0.000**	**0.464** **0.007**	**0.588** **0.000**	**0.557** **0.001**	**0.838** **0.000**	**0.827** **0.000**	**0.723** **0.000**	**0.400** **0.021**									
Hippurat	0.036 0.843	**0.383** **0.028**	0.084 0.641	**0.652** **0.000**	**0.397** **0.022**	**0.390** **0.025**	0.262 0.141	0.298 0.092	**0.414** **0.016**	**0.562** **0.001**	**0.411** **0.018**	0.180 0.315	**0.547** **0.001**								
Lactate	**0.763** **0.000**	0.278 0.118	0.221 0.216	**0.642** **0.000**	**0.430** **0.013**	0.221 0.217	0.290 0.101	**0.471** **0.006**	**0.673** **0.000**	**0.604** **0.000**	**0.511** **0.002**	**0.657** **0.000**	**0.643** **0.000**	0.249 0.162							
Methylam	**0.360** **0.040**	**0.472** **0.005**	0.095 0.599	**0.808** **0.000**	**0.689** **0.000**	**0.525** **0.002**	**0.522** **0.002**	**0.643** **0.000**	**0.782** **0.000**	**0.804** **0.000**	**0.705** **0.000**	**0.437** **0.011**	**0.749** **0.000**	**0.532** **0.001**	**0.677** **0.000**						
N,N-Dime Glycine	**0.373** **0.033**	0.342 0.051	0.283 0.110	**0.702** **0.000**	**0.547** **0.001**	**0.424** **0.014**	**0.590** **0.000**	**0.628** **0.000**	**0.593** **0.000**	**0.783** **0.000**	**0.755** **0.000**	**0.402** **0.020**	**0.623** **0.000**	**0.484** **0.004**	**0.498** **0.003**	**0.633** **0.000**					
Pyruvate	**0.539** **0.001**	**0.476** **0.005**	0.134 0.456	**0.858** **0.000**	**0.653** **0.000**	**0.500** **0.003**	**0.628** **0.000**	**0.606** **0.000**	**0.848** **0.000**	**0.895** **0.000**	**0.812** **0.000**	**0.527** **0.002**	**0.849** **0.000**	**0.531** **0.001**	**0.709** **0.000**	**0.748** **0.000**	**0.738** **0.000**				
Succinat	**0.600** **0.000**	0.301 0.089	0.113 0.532	**0.634** **0.000**	**0.505** **0.003**	**0.426** **0.013**	**0.697** **0.000**	0.325 0.065	**0.513** **0.002**	**0.617** **0.000**	**0.725** **0.000**	**0.440** **0.010**	**0.656** **0.000**	0.291 0.100	**0.616** **0.000**	**0.665** **0.000**	**0.596** **0.000**	**0.691** **0.000**			
Trimethy Amine	**0.521** **0.002**	0.185 0.303	0.214 0.233	**0.724** **0.000**	**0.503** **0.003**	0.217 0.224	**0.557** **0.001**	**0.477** **0.005**	**0.625** **0.000**	**0.642** **0.000**	**0.646** **0.000**	0.302 0.087	**0.721** **0.000**	0.323 0.067	**0.535** **0.001**	**0.558** **0.001**	**0.735** **0.000**	**0.636** **0.000**	**0.634** **0.000**		
Trimethy N-oxide	0.319 0.070	0.272 0.126	0.238 0.183	**0.482** **0.005**	**0.542** **0.001**	**0.547** **0.001**	**0.497** **0.003**	**0.362** **0.038**	**0.415** **0.016**	**0.579** **0.000**	**0.677** **0.000**	**0.481** **0.005**	**0.461** **0.007**	0.191 0.287	**0.377** **0.031**	**0.618** **0.000**	**0.652** **0.000**	**0.511** **0.002**	**0.529** **0.002**	**0.433** **0.012**	
Urea	**0.437** **0.011**	**0.391** **0.024**	-0.105 0.561	**0.723** **0.000**	**0.563** **0.001**	**0.409** **0.018**	**0.618** **0.000**	**0.415** **0.016**	**0.624** **0.000**	**0.722** **0.000**	**0.732** **0.000**	**0.445** **0.009**	**0.727** **0.000**	**0.404** **0.020**	**0.579** **0.000**	**0.566** **0.001**	**0.624** **0.000**	**0.773** **0.000**	**0.687** **0.000**	**0.633** **0.000**	**0.466** **0.006**

Acetate. Acetoacetate. Acetone. Alanine. Betaine. Choline. Citrate. Creatine. Creatine-phosphate. Creatinine. Dimethylamine. Formate. Glycine. Hippurate. Lactate. Methylamine. N,N-dimethylglycine. Pyruvate. Succinate. Trimethylamine. Trimethylamine-N-oxide. Urea. *Boxed- p<0.05

Considering TABLE 5.6, it was seen that most of the endogenous metabolites urinary excretion showed to be correlated to each other, precisely 157 out of 210 correlated pairs of metabolites. In addition, the significantly correlated metabolites showed to have r-values greater than ~0.35, which implies a reasonably fair to high correlation between metabolites pairs.

Of these metabolites: alanine, creatine-phosphate, creatinine, dimethylamine, glycine, methylamine, pyruvate and urea showed to be significantly correlated to the other metabolites assessed except acetone, followed by betaine and N,N-dimethyl glycine . Acetone was the least correlative because it showed no significant correlation with any other metabolites analysed.

Although, the statisical values of the p-value of each correlation showed to be less than 0.05 as test of significance, there was still a need to visually analyse the matrix plot for un-influenced linear correlation slope. On this basis, FIGURE 5.7, in support of the statistical values that alanine, creatine-phosphate, creatinine, dimethylamine, glycine, methylamine, pyruvate and urea appeared to be linearly correlated to most metabolites ($r > 0.35$), as there were individual linear patterns and slopes observed in the matrix plot. In addition, some of the other pairs showed to have slightly obvious linear patterns as expected from the statistical data obtained in TABLE 5.6.

.

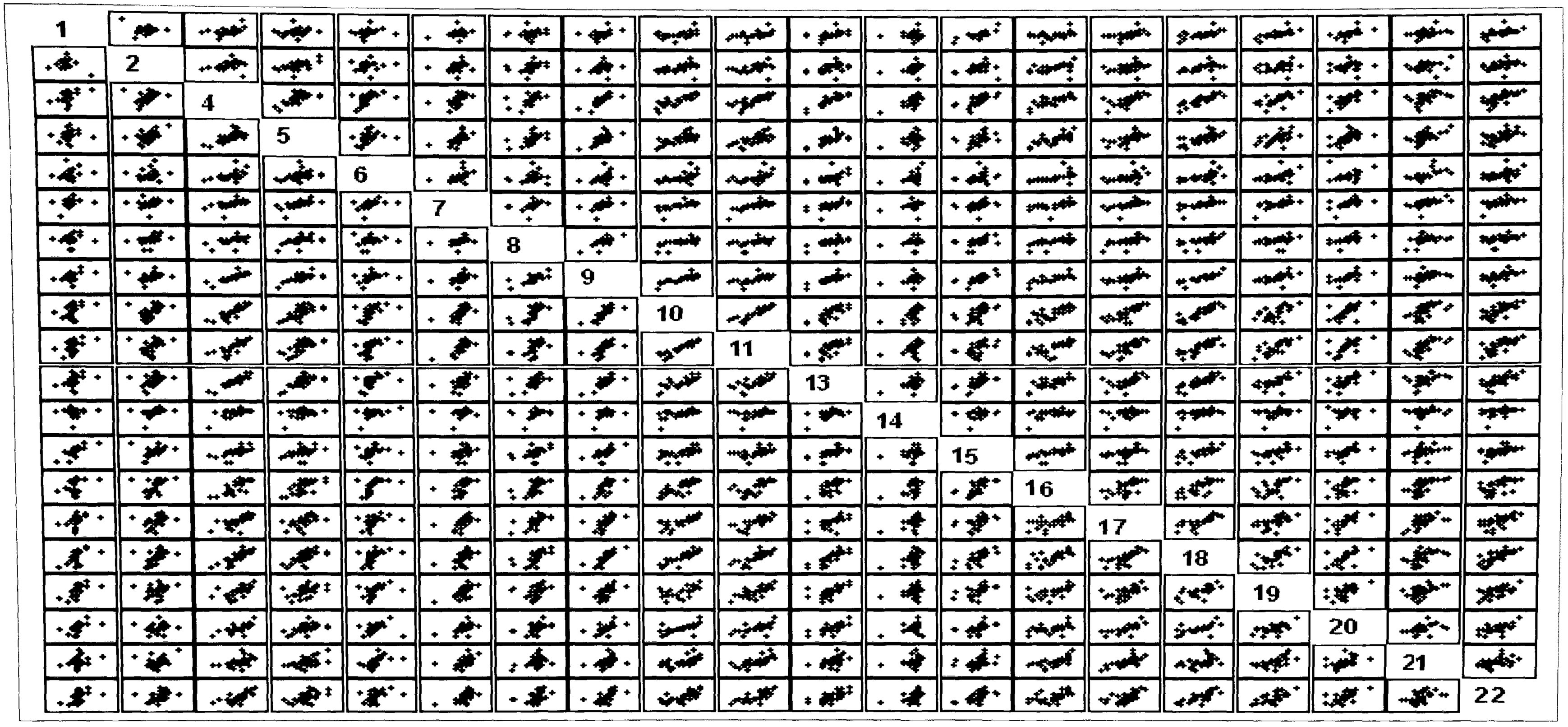


FIGURE 5.7: Matrix-Plot of some endogenous metabolites on Day 3

1. Acetate 2. Acetoacetate 4. Alanine 5. Betaine 6. Choline 7. Citrate. 8. Creatine. 9. Creatine phosphate 10. Creatinine 11. Dimethylamine 13. Glycine 14. Hippurate 15. Lactate 16. Methylamine 17. N,N-dimethylglycine 18. Pyruvate 19. Succinate 20. Trimethylamine 21. Trimethylamine-N-oxide 22. Urea

5.2.3 Correlation-ship of soy - phytoestrogen metabolites and endogenous metabolites

Following the completion of testing each group of metabolites: soy - phytoestrogen and endogenous metabolites individually, there is a need to depict the linkage between these two groups of metabolites, if there are any linkages. On this note, the Pearson correlation-ship of these metabolites has been summarised below in TABLE 5.7.

TABLE 5.7: Pearson correlation of soy - and endogenous – metabolites on Day 3

Metabolites	Daidzein	Equol	Genistein	o-DMA
Acetate	-0.072; 0.691	0.064; 0.725	-0.008; 0.963	-0.153; 0.396
Acetoacetate	-0.015; 0.932	-0.077; 0.672	0.055; 0.761	**-0.359; 0.040***
Acetone	-0.240; 0.179	0.239; 0.180	-0.049; 0.785	-0.127; 0.483
Alanine	-0.110; 0.544	-0.079; 0.663	0.011; 0.953	-0.029; 0.871
Betaine	-0.153; 0.396	-0.117; 0.517	-0.116; 0.520	-0.238; 0.182
Choline	-0.109; 0.546	**-0.457; 0.007***	0.019; 0.915	-0.080; 0.657
Citrate	-0.253; 0.155	-0.023; 0.899	-0.109; 0.548	-0.254; 0.154
Creatine	-0.153; 0.396	-0.110; 0.542	0.015; 0.936	-0.161; 0.371
Creatine-Phosphate	-0.081; 0.654	-0.213; 0.235	-0.102; 0.572	-0.097; 0.591
Creatinine	-0.190; 0.290	-0.169; 0.348	0.006;0.973	-0.215;0.229
Dimethyl-amine	-0.289; 0.103	-0.118; 0.514	-0.059; 0.745	-0.300;0.090
Formate	0.023; 0.900	0.016; 0.928	0.090; 0.618	0.014; 0.938
Glycine	-0.123; 0.494	-0.131; 0.466	-0.084; 0.642	-0.008; 0.967
Hippurate	0.004, 0.981	0.075, 0.680	0.323; 0.066	-0.051; 0.776
Lactate	0.022; 0.903	0.018;0.919	0.112; 0.534	-0.074; 0.682
Methyl-amine	-0.081; 0.653	-0.216;0.226	0.069; 0.704	-0.179; 0.319
Dimethyl-glycine	-0.157; 0.383	0.114; 0.528	0.073; 0.687	-0.131; 0.466
Pyruvate	-0.042; 0.817	-0.142; 0.431	0.039; 0.830	-0.140; 0.437
Succinate	-0.191; 0.288	0.102; 0.573	0.030; 0.870	-0.298; 0.092
TMA	-0.195; 0.277	0.211; 0.237	-0.124; 0.490	-0.104; 0.563
TMAO	-0.200; 0.264	-0.226; 0.206	-0.132; 0.465	-0.247; 0.166
Urea	0.039; 0.828	0.022; 0.902	0.143; 0.427	-0.077; 0.672

Pearson correlation; p-values. *p<0.05

From statistical data generated by Minitab (Student Version 12), only two out of possible 88 pairs of metabolites showed to be linearly correlated. These are aceto-acetate and o-

DMA and choline and equol, both found to be negatively correlated with r >0.390 and p < 0.05.

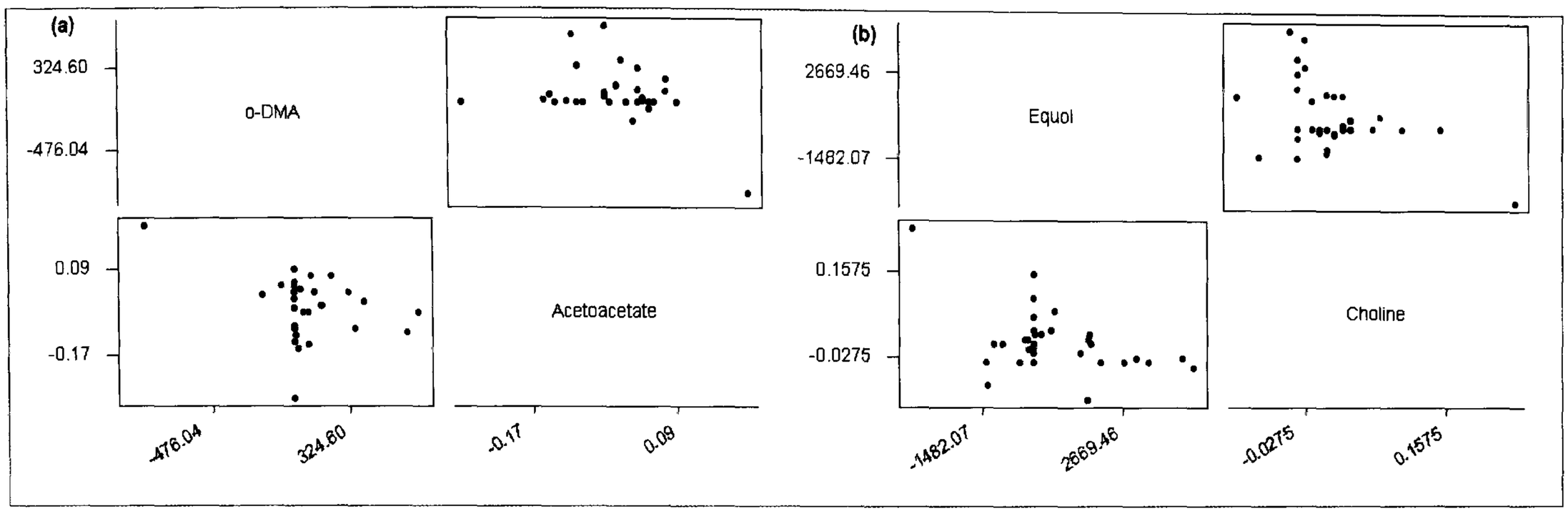


FIGURE 5.8: Matrix Plots of two selected metabolite pairs

(a) o-DMA-acetoacetate (b) equol-choline

Furthermore, on visual interpretation of FIGURE 5.8, it showed that there were no clear linear correlations between each pair of metabolites, thus correlation was influenced by outliers.

Following the location taxonomy employed in previously mentioned in Chapter III and IV, in the NIGERIA cohort, it showed that the same metabolites pairs identified to have good statistical correlation values ($r > 0.3$; $p < 0.05$) in the general cohort, again, showed to have significant correlation values(a. $r = -0.448$, $p = 0.036$; b. $r = -0.603$, $p = 0.008$). In addition, there appeared to be a correlation between equol-methylamine, $r = -0.490$, $p = 0.039$. Similarly, the statistical values seemed not to reflect the true linear correlation of the three pairs of metabolites on the matrix-plot (FIGURE 5.9) except equol-choline, whilst equol-methylamine had the data spread sparingly randomly on the plot and again o-DMA-acetoacetate showed Pearson correlation r - value was influenced by an outlier.

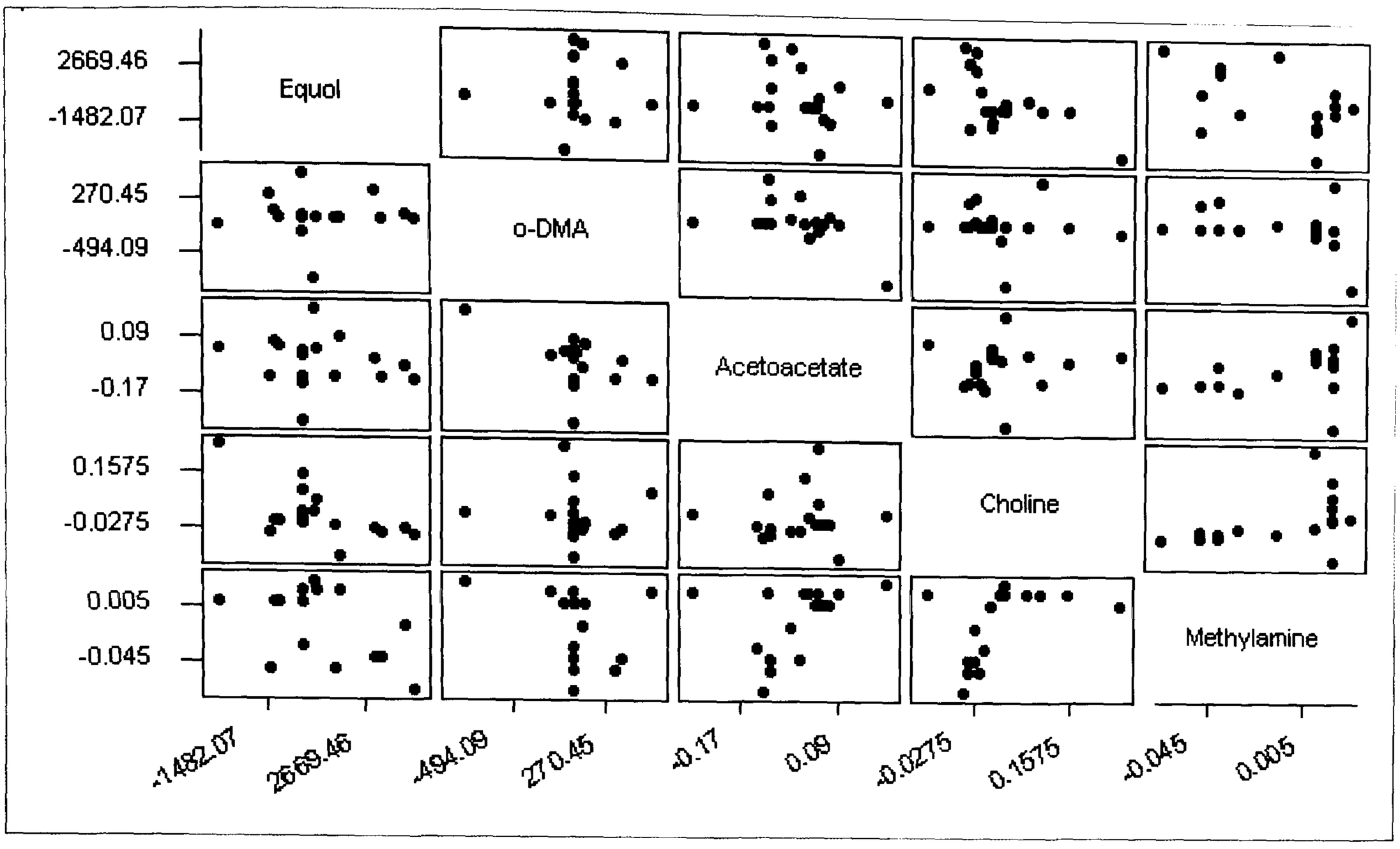


FIGURE 5.9: Matrix plot of some metabolites correlation in NIGERIA cohort

On the other hand in the UK cohort, the only significant correlation was with equol-creatine-phosphate, which was found to be positively significant ($r = 0.548$; $p = 0.034$).

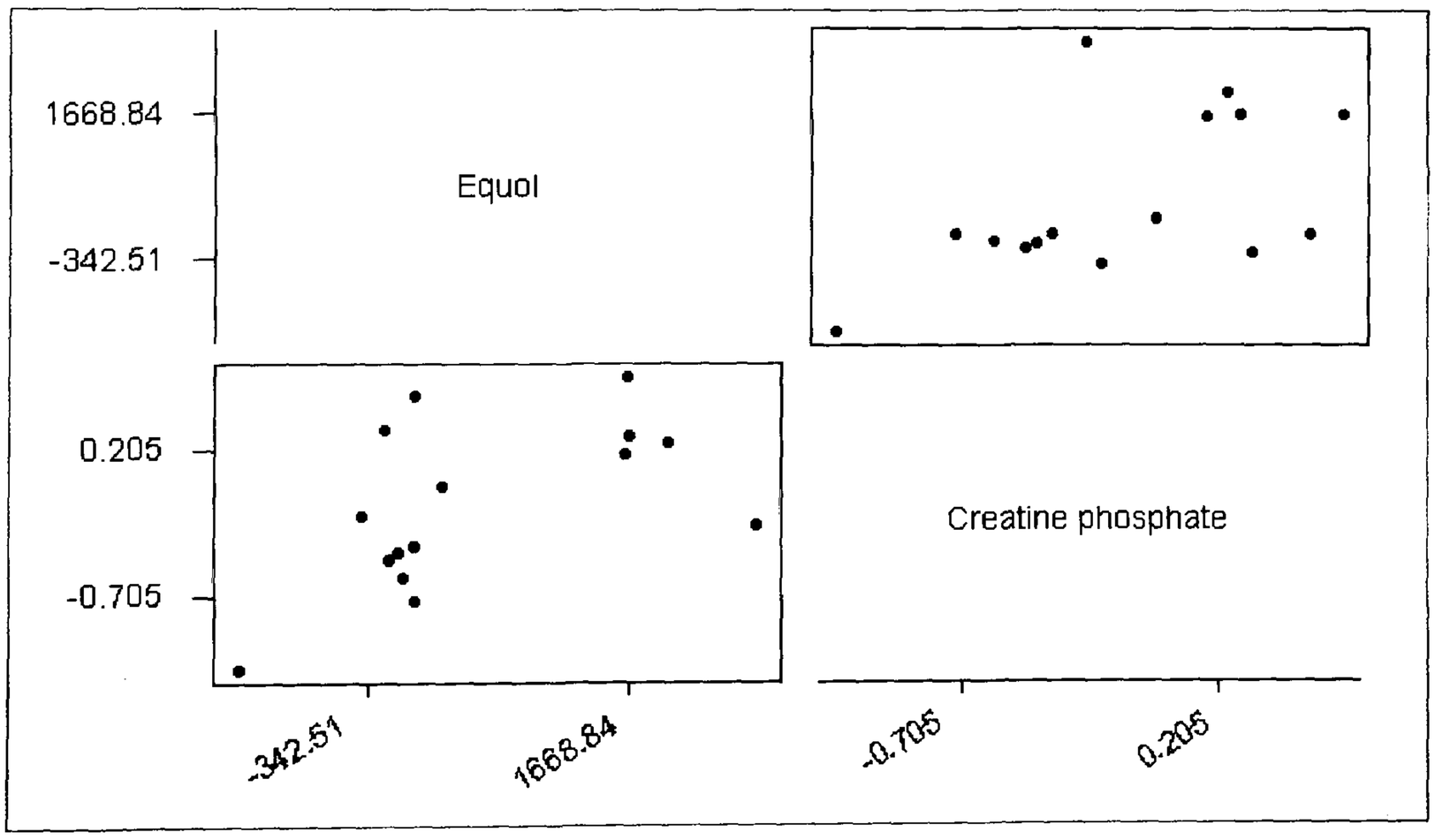


FIGURE 5.10: Matrix-Plot of equol-creatine phosphate in UK cohort

FIGURE 5.10 confirms the correlation of equol and creatinine phosphate within this group.

5.3 Conclusions

There was a general correlation between daidzein and genistein metabolism & excretion. Furthermore, the UK cohort showed to have better correlations in soy - phytoestrogen metabolites than NIGERIA cohort.

Following this, there were distinct correlations in many of the endogenous metabolites, i.e. 157 out of 210 possible matrix pairs. These include alanine, creatine-phosphate, creatinine, dimethylamine, glycine, methylamine, pyruvate and urea.

Little or no significant correlation was observed with soy - phytoestrogen metabolites and endogenous metabolites.

5.4 Experimental

5.4.1 Statistical data analysis

All statistical analyses were performed using Minitab software Student release version 12 (Minitab Ltd, Coventry, United Kingdom). Urinary excretion of soy - metabolites and endogenous metabolites were expressed as pg/μL and mM respectively. Pearson product moment correlations were used to assess the degree of associations between variables, which in this study were urinary excretion levels of soy – metabolisms only, endogenous metabolites only and combined soy – metabolites and endogenous metabolites. This was used to measure the degree of linear associations between the variables, the baseline level was used as a reference to which during - phase (Day 3 & 5) samples were subtracted from, thereby using the changes (Δ) in levels. All p – values are tested based on the hypothesis of zero correlation coefficient. The p – values calculated were two – sided and $p < 0.05$ was considered statistically significant.

Data from *Chapters III* and *IV* were used for these correlation studies.

Conclusions

In this work, seven phytoestrogens - including four isoflavones: biochanin A, daidzein, formononetin, genistein; one coumestan: coumestrol and two lignans: matairesinol and secoisolariciresinol have been identified and quantified in six commonly consumed African foods namely beans, cassava 'garri', plantain, pumpkin leaves 'ugu', rice and yam. Additionally, three novel phytoestrogenic compounds were found in pumpkin leaves 'ugu', identified and characterised by GC-MS and LC-MS. Although, only six African foods were analysed, this research shows the potential of discovery of new phytoestrogenic compounds in these uncommonly studied food samples.

Aside from food analysis, soy - phytoestrogen metabolism in Africans have not been explored, thus, this study however being a fairly small cohort study is the first to report the extent and percentage population to which Africans metabolise isoflavones (daidzein and genistein) present in soy-milk. The data showed that 35% of the African subjects were good responders (both equol and o-DMA producers) to soy - phytoestrogen, in accordance with the generalised 30 - 50% of different world-wide populations reported by other research groups. In this study, 40% were equol producers and 55% were o-DMA producers. In addition to these findings, results showed that two main selective factors were responsible for the production and excretion levels of phytoestrogens (daidzein and genistein) and their metabolites (equol and o-DMA). These selective factors were location and gender. Furthermore, age showed to influence o-DMA production.

Following the metabolism of soy - phytoestrogens, there was the need to understand the biochemical effects of soy phytoestrogens and their metabolites on endogenous metabolites. This was aimed to detect the significance of soy - phytoestrogens with general biochemical processes. Bearing this in mind, investigations and analyses were carried out by ^{1}H-NMR metabonomics: quantitation analyses of twenty-two endogenous metabolites and multivariate analysis. Results showed that there was a reduction in glycolysis, as a result of decrease in lactate and pyruvate. Additionally, as a result of reduced levels of glycine and alanine following soy consumption (Days 3 & 5 - Phases II & III/ during), there was a reduction in collagen. Potentially, there are other biochemical responses that could occur following soy - consumption, thus there is the need to analyse more metabolites for this purpose.

The multivariate analysis results showed location and gender seeming to be the main influencing factors in soy - phytoestrogen metabolism. Based on the MVA pattern

recognition techniques, there is the potential of discovering new biomarkers as well as other factors influencing soy-metabolism. Thus, this would generally encourage more understanding of soy-metabolism in respect to their biochemical influences by employing realistic *in-vivo* rather than *in-vitro* environments in comparison to various vital endogenous metabolites.

Recommendations for further work

The work presented in this research thesis shows the preliminary building blocks 'foundation' of phytoestrogens analyses within Africans based on small scale studies: food, soy-metabolites analyses and endogenous metabolites analyses. With this in mind, further work to be achieved would involve having a larger cohort study, especially investigating the metabolism and biochemical effects of phytoestrogens in Africans. This would entail:

- Introducing a larger cohort to soy-milk or other phytoestrogen-rich diet intervention in the African populations living in Africa and elsewhere in the world.
- Investigating other factors that could encourage or discourage phytoestrogens metabolism, for example, smoking habits, drinking habits, medical background, genetics, diets, BMI, drugs / medications.
- Assessing both quantitative and multivariate analyses of bio-fluids metabonomics (either ^{1}H-NMR or MS) for the biochemical effects of phytoestrogens on a large number of endogenous metabolites. In addition, investigating several possible influencing factors and biomarkers in phytoestrogen-metabolism.

References

1. Knight, D. C.; Eden, J. A., A review of the clinical effects of phytoestrogens. *Obstetrics and Gynecology* **1996,** 87, 897-904.

2. Shu, X. O.; Jin, F.; Dai, Q.; Wen, W. Q.; Potter, J. D.; Kushi, L. H.; Ruan, Z. X.; Gao, Y. T.; Zheng, W., Soyfood intake during adolescence and subsequent risk of breast cancer among Chinese women. *Cancer Epidemiology Biomarkers & Prevention* **2001,** 10, 483-488.

3. Adlercreutz, C. H. T.; Goldin, B. R.; Gorbach, S. L.; Hockerstedt, K. A. V.; Watanabe, S.; Hamalainen, E. K.; Markkanen, M. H.; Makela, T. H.; Wahala, K. T.; Hase, T. A.; Fotsis, T., Soybean Phytoestrogen Intake and Cancer Risk (Vol 125, Pg S757, 1995). *Journal of Nutrition* **1995,** 125, 1960-1960.

4. Somekawa, Y.; Chiguchi, M.; Ishibashi, T.; Aso, T., Soy intake related to menopausal symptoms, serum lipids, and bone mineral density in postmenopausal Japanese women. *Obstetrics and Gynecology* **2001,** 97, 109-115.

5. Vanharanta, M.; Voutilainen, S.; Lakka, T. A.; van der Lee, M.; Adlercreutz, H.; Salonen, J. T., Risk of acute coronary events according to serum concentrations of enterolactone: a prospective population-based case-control study. *Lancet* **1999,** 354, 2112-2115.

6. Kaffashian, F.; Godward, S.; Davies, T.; Solomon, L.; McCann, J.; Duffy, S. W., Socioeconomic effects on breast cancer survival: proportion attributable to stage and morphology. *British Journal of Cancer* **2003,** 89, 1693-1696.

7. Sant, M.; Aareleid, T.; Berrino, F.; Bielska Lasota, M.; Carli, P. M.; Faivre, J.; Grosclaude, P.; Hedelin, G.; Matsuda, T.; Moller, H.; Moller, T.; Verdecchia, A.; Capocaccia, R.; Gatta, G.; Micheli, A.; Santaquilani, M.; Roazzi, P.; Lisi, D., EUROCARE-3: survival of cancer patients diagnosed 1990-94--results and commentary. *Ann Oncol* **2003,** 14 Suppl 5, v61-118.

8. Higuchi, C. M.; Serxner, S. A.; Nomura, A. M. Y.; Stemmermann, G. N., Histopathological predictors of breast-cancer death among Caucasians and Japanese in Hawaii. *Cancer Epidemiology Biomarkers & Prevention* **1993,** 2, 201-205.

9. Shimizu, H.; Ross, R. K.; Bernstein, L.; Yatani, R.; Henderson, B. E.; Mack, T. M., Cancers of the prostate and breast among japanese and white immigrants in los-angeles-county. *British Journal of Cancer* **1991,** 63, 963-966.

10. Adlercreutz, H.; Hamalainen, E.; Gorbach, S.; Goldin, B., Dietary phyto-estrogens and the menopause in japan. *Lancet* **1992,** 339, 1233-1233.

11. Adlercreutz, C. H. T.; Goldin, B. R.; Gorbach, S. L.; Hockerstedt, K. A. V.; Watanabe, S.; Hamalainen, E. K.; Markkanen, M. H.; Makela, T. H.; Wahala, K. T.; Hase, T. A.; Fotsis, T., Soybean phytoestrogen intake and cancer risk. *Journal of Nutrition* **1995,** 125, 1960-1960.

12. Adlercreutz, H.; Markkanen, H.; Watanabe, S., Plasma-concentrations of phyto-estrogens in Japanese men. *Lancet* **1993,** 342, 1209-1210.

13. Adlercreutz, H.; Wahala, K.; Rasku, S.; Salakka, A.; Watanabe, S., Metabolism of isoflavonoids in Japanese men. *American Journal of Clinical Nutrition* **1998,** 68, 1536s-1536s.

14. Adlercreutz, C. H. T.; Goldin, B. R.; Gorbach, S. L.; Hockerstedt, K. A. V.; Watanabe, S.; Hamalainen, E. K.; Markkanen, M. H.; Makela, T. H.; Wahala, K. T.; Hase, T. A.; Fotsis, T. In *Soybean phytoestrogen intake and cancer risk*, 1st International Symposium on the Role of Soy in Preventing and Treating Chronic Disease, Mesa, Az, Feb 20-23, 1994; Amer Inst Nutrition: Mesa, Az, 1994; pp S757-S770.

15. Adlercreutz, H.; Fotsis, T.; Bannwart, C.; Hamalainen, E.; Bloigu, S.; Ollus, A., Urinary estrogen profile determination in young Finnish vegetarian and omnivorous women. *Journal of Steroid Biochemistry and Molecular Biology* **1986,** 24, 289-296.

16. Adlercreutz, H., Estrogen metabolism and excretion in oriental and caucasian women. *Journal of the National Cancer Institute* **1994,** 86, 1076.

17. Baker, W., Robinson, R. , Synthetical experiments in the isoflavone group III. Synthesis of genistein. *Journal of Chemical Society* **1928,** 131, 3115-3118.

18. Waltz, E., Isoflavan und saponin-glucoside is soja hispida. *Justus Liebig's Annual Chemistry* **1931,** 489, 118-155.

19. Marrian, G. F.; Haslewood, G. A., Equol, a new inactive phenol isolated from the ketohydroxyoestrin fraction of mares' urine. *Biochemical Journal* **1932,** 26, 1227-1232.

20. Harworth, R. D., Constituents of natural phenolic resins. *Nature (Lond)* **1941,** 147, 255.

21. Bennetts, H. W.; Underwood, E. J.; Shier, F. L., A specific breeding problem of sheep on subterranean clover pastures in Western Australia. *Australian Veterinary Journal* **1946,** 22, 2-12.

22. Setchell, K. D. R.; Lawson, A. M.; Mitchell, F. L.; Adlercreutz, H.; Kirk, D. N.; Axelson, M., Lignans in man and in animal species. *Nature* **1980,** 287, 740-742.

23. Leopold, A. S.; Erwin, M.; Oh, J.; Browning, B., Phytoestrogens - Adverse effects on reproduction in California quail. *Science* **1976,** 191, 98-100.

24. Common, R. H.; Ainsworth, L., Identification of equol in the urine of the domestic fowl. *Biochim Biophys Acta* **1961,** 53, 403-404.

25. Batterham, T. J.; Hart, N. K.; Lamberton, J. A., Metabolism of oestrogenic isoflavones in sheep. *Nature* **1965,** 206, 509.

26. Axelson, M.; Setchell, K. D. R., The excretion of lignans in rats - evidence for an intestinal bacterial source for this new group of compounds. *Febs Letters* **1981,** 123, 337-342.

27. Axelson, M.; Setchell, K. D. R., Conjugation of lignans in human-urine. *Febs Letters* **1980,** 122, 49-53.

28. Nilsson, A., Demethylation of the plant oestrogen Biochanin A in the rat. *Nature* **1961,** 192, 358-358.

29. Nilsson, A.; Hill, J. L.; Davids, H. L., An in vitro study of formononetin and biochanin A metabolism in rumen fluid from sheep. *Biochim Biophys Acta* **1967,** 148, 92-98.

30. Setchell, K. D.; Lawson, A. M.; Conway, E.; Taylor, N. F.; Kirk, D. N.; Cooley, G.; Farrant, R. D.; Wynn, S.; Axelson, M., The definitive identification of the lignans trans-2,3-bis(3-hydroxybenzy)-gamma-butyrolactone and 2,3-bis(3-hydroxybenzyl)butane-1,4-diol in human and animal urine. *Biochemical Journal* **1981,** 197, 447-458.

31. Axelson, M.; Kirk, D. N.; Farrant, R. D.; Cooley, G.; Lawson, A. M., The identification of the weak oestrogen equol [7-hydroxy-3-(4'-hydroxyphenyl)chroman] in human urine. *Biochemical Journal* **1981,** 201, 353-357.

32. Klyne, W.; Wright, A. A., Steriods and other lipids of pregnant cow's urine. *Journal of Endocrinology* **1959,** 18, 32-45.

33. Klyne, W.; Wright, A. A., Steroids and other lipids of pregnant goat's urine. *Biochemical Journal* **1957,** 66, 92-101.

34. Stitch, S. R.; Toumba, J. K.; Groen, M. B.; Funke, C. W.; Leemhuis, J.; Vink, J.; Woods, G. F., Excretion, Isolation and Structure of a New Phenolic Constituent of Female Urine. *Nature* **1980,** 287, 738-740.

35. Axelson, M.; Sjovall, J.; Gustafsson, B. E.; Setchell, K. D. R., Origin of lignans in mammals and identification of a precursor from plants. *Nature* **1982,** 298, 659-660.

36. Stitch, S. R.; Smith, P. D.; Illingworth, D.; Toumba, K., Cyclic Excretion of a Non-Steroidal Compound in Women. *Journal of Endocrinology* **1980,** 85, P23-P23.

37. Bannwart, C.; Fotsis, T.; Heikkinen, R.; Adlercreutz, H., Identification of the isoflavonic phytoestrogen daidzein in human-urine. *Clinica Chimica Acta* **1984,** 136, 165-172.

38. Bannwart, C.; Adlercreutz, H.; Wahala, K.; Kotiaho, T.; Hesso, A.; Brunow, G.; Hase, T., Identification of the phyto-oestrogen 3',7-dihydroxyisoflavan, an isomer of equol, in human-urine and cows milk. *Biomedical and Environmental Mass Spectrometry* **1988,** 17, 1-6.

39. Setchell, K. D. R.; Bonney, R. C., The excretion of urinary steroids by the owl monkey (Aotus-Trivirgatus) studied using open tubular capillary column gas-chromatography and mass-spectrometry. *Journal of Steroid Biochemistry and Molecular Biology* **1981,** 14, 37-43.

40. Tang, B. Y.; Adams, N. R., Effect of equol on oestrogen receptors and on synthesis of DNA and protein in the immature rat uterus. *Journal of Endocrinology* **1980,** 85, 291-297.

41. Shutt, D. A.; Cox, R. I., Steroid and phytoestrogen binding to sheep uterine receptors *in-vitro*. *Journal of Endocrinology* **1972,** 52, 299-&.

42. Naim, M.; Gestetne.B; Birk, Y.; Bondi, A., Characterization of soybean isoflavones. *Israel Journal of Medical Sciences* **1973,** 9, 555-555.

43. West, L. G. B., P.M.; Pratt, D.E., Separation of the isomeric isoflavones from soybean by High Performance Liquid chromatography. *Journal of Chromatography* **1978,** 150, 266-268.

44. Harworth, R. D., *Annual Rep* **1936,** 33, 270.

45. Naim, M.; Gestetne.B; Kirson, I.; Birk, Y.; Bondi, A., New isoflavone from soya beans. *Phytochemistry* **1973,** 12, 169-170.

46. Adlercreutz, H.; Fotsis, T.; Bannwart, C.; Wahala, K.; Brunow, G.; Hase, T., Isotope-dilution gas-chromatographic mass-spectrometric method for the determination of lignans and isoflavonoids in human urine, including identification of genistein. *Clinica Chimica Acta* **1991,** 199, 263-278.

47. Franke, A. A.; Custer, L. J., High-Performance Liquid-Chromatographic Assay of Isoflavonoids and Coumestrol from Human Urine. *Journal of Chromatography B-Biomedical Applications* **1994,** 662, 47-60.

48. Franke, A. A.; Custer, L. J.; Cerna, C. M.; Narala, K. K., Quantitation of Phytoestrogens in Legumes by Hplc. *Journal of Agricultural and Food Chemistry* **1994,** 42, 1905-1913.

49. Messina, M. J.; Persky, V.; Setchell, K. D. R.; Barnes, S., Soy intake and cancer risk - a review of the *in-vitro* and *in-vivo* data. *Nutrition and Cancer-an International Journal* **1994,** 21, 113-131.

50. Wang, H.; Murphy, P. A., Isoflavone content in commercial soybean foods. *Journal of Agricultural and Food Chemistry* **1994,** 42, 1666-1673.

51. Essex, C., Phytoestrogens and soy based infant formula. *British Medical Journal* **1996,** 313, 507-508.

52. Mazur, W.; Fotsis, T.; Wahala, K.; Ojala, S.; Salakka, A.; Adlercreutz, H., Isotope dilution gas chromatographic mass spectrometric method for the determination of isoflavonoids, coumestrol, and lignans in food samples. *Analytical Biochemistry* **1996,** 233, 169-180.

53. Wu, A. H.; Ziegler, R. G.; HornRoss, P. L.; Nomura, A. M. Y.; West, D. W.; Kolonel, L. N.; Rosenthal, J. F.; Hoover, R. N.; Pike, M. C., Tofu and risk of breast cancer in Asian-Americans. *Cancer Epidemiology Biomarkers & Prevention* **1996,** 5, 901-906.

54. Huang, H. Y.; Hsieh, Y. Z., Determination of puerarin, daidzein, paeoniflorin, cinnamic acid, glycyrrhizin, ephedrine, and [6]-gingerol in Ge-gen-tang by micellar electrokinetic chromatography. *Analytica Chimica Acta* **1997,** 351, 49-55.

55. Slavin, J., More information on phytoestrogens in breast milk. *Clinical Chemistry* **1997,** 43, 548a-549.

56. Slanina, J.; Taborska, E.; Lojkova, L., Lignans in the seeds and fruits of Schisandra chinensis cultured in Europe. *Planta Medica* **1997,** 63, 277-280.

57. Gamache, P. H.; Maher, T. J.; Setchell, K. D. R.; Wu, T. H.; Acworth, I. N., Sensitive analysis of phytoestrogens in animal tissue by HPLC with coulometric array detection. *American Journal of Clinical Nutrition* **1998,** 68, 1536s-1537s.

58. Irvine, C. H. G.; Shand, N.; Fitzpatrick, M. G.; Alexander, S. L., Daily intake and urinary excretion of genistein and daidzein by infants fed soy- or dairy-based infant formulas. *American Journal of Clinical Nutrition* **1998,** 68, 1462s-1465s.

59. Mazur, W., Phytoestrogen content in foods. *Baillieres Clinical Endocrinology and Metabolism* **1998,** 12, 729-742.

60. Liggins, J.; Grimwood, R.; Bingham, S. A., Extraction and quantification of lignan phytoestrogens in food and human samples. *Analytical Biochemistry* **2000,** 287, 102-109.

61. Merken, H. M.; Beecher, G. R., Measurement of food flavonoids by high-performance liquid chromatography: A review. *Journal of Agricultural and Food Chemistry* **2000,** 48, 577-599.

62. de Kleijn, M. J. J.; van der Schouw, Y. T.; Wilson, P. W. F.; Adlercreutz, H.; Mazur, W.; Grobbee, D. E.; Jacques, P. F., Intake of dietary phytoestrogens is low in postmenopausal women in the United States: The Framingham study. *Journal of Nutrition* **2001,** 131, 1826-1832.

63. Dai, Q.; Shu, X. O.; Jin, F.; Potter, J. D.; Kushi, L. H.; Teas, J.; Gao, Y. T.; Zheng, W., Population-based case-control study of soyfood intake and breast cancer risk in Shanghai. *British Journal of Cancer* **2001,** 85, 372-378.

64. Hutabarat, L. S.; Greenfield, I.; Mulholland, M., Isoflavones and coumestrol in soybeans and soybean products from Australia and Indonesia. *Journal of Food Composition and Analysis* **2001,** 14, 43-58.

65. Lampe, J. W.; Skor, H. E.; Li, S.; Wahala, K.; Howald, W. N.; Chen, C., Wheat bran and soy protein feeding do not alter urinary excretion of the isoflavan equol in premenopausal women. *Journal of Nutrition* **2001,** 131, 740-744.

66. Mitchell, J. H.; Cawood, E.; Kinniburgh, D.; Provan, A.; Collins, A. R.; Irvine, D. S., Effect of a phytoestrogen food supplement on reproductive health in normal males. *Clinical Science* **2001,** 100, 613-618.

67. Setchell, K. D. R.; Brown, N. M.; Desai, P.; Zimmer-Nechemias, L.; Wolfe, B. E.; Brashear, W. T.; Kirschner, A. S.; Cassidy, A.; Heubi, J. E., Bioavailability of pure isoflavones in healthy humans and analysis of commercial soy isoflavone supplements. *Journal of Nutrition* **2001,** 131, 1362s-1375s.

68. Boker, L. K.; Van der Schouw, Y. T.; De Kleijn, M. J. J.; Jacques, P. F.; Grobbee, D. E.; Peeters, P. H. M., Intake of dietary phytoestrogens by Dutch women. *Journal of Nutrition* **2002,** 132, 1319-1328.

69. Cao, Y.; Calafat, A. M.; Doerge, D. R.; Umbach, D. M.; Bernbaum, J. C.; Twaddle, N. C.; Ye, X. Y.; Rogan, W. J., Isoflavones in urine, saliva, and blood of infants: data from a pilot study on the estrogenic activity of soy formula. *Journal of Exposure Science and Environmental Epidemiology* **2009,** 19, 223-234.

70. Kuhnle, G. G. C.; Dell'Aquila, C.; Aspinall, S. M.; Runswick, S. A.; Mulligan, A. A.; Bingham, S. A., Phytoestrogen content of beverages, nuts, seeds, and oils. *Journal of Agricultural and Food Chemistry* **2008,** 56, 7311-7315.

71. Penalvo, J. L.; Adlercreutz, H.; Uehara, M.; Ristimaki, A.; Watanabe, S., Lignan content of selected foods from Japan. *Journal of Agricultural and Food Chemistry* **2008,** 56, 401-409.

72. Ward, H.; Chapelais, G.; Kuhnle, G. G. C.; Luben, R.; Khaw, K. T.; Bingham, S., Breast cancer risk in relation to urinary and serum biomarkers of phytoestrogen exposure in the European Prospective into Cancer-Norfolk cohort study. *Breast Cancer Research* **2008,** R32, 10.

73. Adlercreutz, H., ESTROGEN METABOLISM AND EXCRETION IN ORIENTAL AND CAUCASIAN WOMEN (VOL 86, PG 1076, 1994). *Journal of the National Cancer Institute* **1995,** 87, 147-147.

74. Adlercreutz, H.; Hamalainen, E.; Gorbach, S. L.; Goldin, B. R.; Woods, M. N.; Brunson, L. S.; Dwyer, J. T., Association of diet and sex - hormones in relation to breast - cancer. *European Journal of Cancer & Clinical Oncology* **1987,** 23, 1725-1726.

75. Adlercreutz, H.; Hockerstedt, K.; Bannwart, C.; Bloigu, S.; Hamalainen, E.; Fotsis, T.; Ollus, A., Effect of dietary-components, Including lignans and phytoestrogens, on enterohepatic circulation and liver-metabolism of estrogens and on sex-hormone binding globulin (Shbg). *Journal of Steroid Biochemistry and Molecular Biology* **1987,** 27, 1135-1144.

76. Day, A. J., DuPont, M.S., Saxon Ridley, S., Deglycosylation of flavonoid and isoflavonoid glycosides by human small intestine and liver & beta-glucosidase activity *Febs Letters* **1998** 436, 71-75.

77. Allred, C. D.; Allred, K. F.; Ju, Y. H.; Virant, S. M.; Helferich, W. G., Soy diets containing varying amounts of genistein stimulate growth of estrogen-dependent (MCF-7) tumors in a dose-dependent manner. *Cancer Research* **2001,** 61, 5045-5050.

78. Winter, J., Moore, L.H., Dowell, V.R. (Jr), Bokkenheuser, V.D., C-ring cleavage of flavonoids by human intestinal bacteria. *Applied Environmental Microbiology* **1989,** 55, 1203-1208.

79. Winter, J.; Bokkenheuser, V. D., Bacterial metabolism of natural and synthetic sex-hormones undergoing enterohepatic circulation. *Journal of Steroid Biochemistry and Molecular Biology* **1987,** 27, 1145-1149.

80. Mann, J., Conference report: Sixth annual public health forum. *Food Australia* **1996,** 48, A4-A5.

81. Murphy, P. A.; Barua, K.; Hauck, C. C., Solvent extraction selection in the determination of isoflavones in soy foods. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002,** 777, 129-138.

82. Adlercreutz, H., Phyto-oestrogens and cancer. *Lancer Oncology* **2002,** 3, 364-373.

83. Setchell, K. D. R.; Borriello, S. P.; Hulme, P.; Kirk, D. N.; Axelson, M., Nonsteroidal estrogens of dietary origin - possible roles in hormone-dependent disease. *American Journal of Clinical Nutrition* **1984,** 40, 569-578.

84. Yildiz, F., *Phytoestrogens in functional foods*. Taylor & Francis Ltd: 2005; p 3-5.

85. Hwang, C. S.; Kwak, H. S.; Lim, H. J.; Lee, S. H.; Kang, Y. S.; Choe, T. B.; Hur, H. G.; Han, K. O., Isoflavone metabolites and their in vitro dual functions: They can act as an estrogenic agonist or antagonist depending on the estrogen concentration. *Journal of Steroid Biochemistry and Molecular Biology* **2006,** 101, 246-253.

86. Ji, Z. N.; Zhao, W. Y.; Liao, G. R.; Choi, R. C.; Lo, C. K.; Dong, T. T. X.; Tsim, K. W. K., In vitro estrogenic activity of formononetin by two bioassay systems. *Gynecological Endocrinology* **2006,** 22, 578-584.

87. Jones, A. E.; Price, K. R.; Fenwick, G. R., Development and application of a high-performance liquid-chromatographic method for the analysis of phytoestrogens. *Journal of the Science of Food and Agriculture* **1989,** 46, 357-364.

88. Li, J.; Lee, L.; Gong, Y.; Shen, P.; Wong, S. P.; Wise, S. D.; Yong, E. L., Bioassays for estrogenic activity: Development and validation of estrogen receptor (ER alpha/ER beta) and breast cancer proliferation bioassays to measure serum estrogenic activity in clinical studies. *Assay and Drug Development Technologies* **2009,** 7, 80-89.

89. Kinjo, J.; Tsuchihashi, R.; Morito, K.; Hirose, T.; Aomori, T.; Nagao, T.; Okabe, H.; Nohara, T.; Masamune, Y., Interactions of phytoestrogens with estrogen receptors alpha and beta (III). Estrogenic activities of soy isoflavone aglycones and their metabolites isolated from human urine. *Biological & Pharmaceutical Bulletin* **2004,** 27, 185-188.

90. Zava, D. T.; Blen, M.; Duwe, G., Estrogenic activity of natural and synthetic estrogens in human breast cancer cells in culture. *Environmental Health Perspectives* **1997,** 105, 637-645.

91. Wang, C.-C.; Prasain, J. K.; Barnes, S., Review of the methods used in the determination of phytoestrogens. *Journal of Chromatography B* **2002,** 777, 3-28.

92. Kudou, S.; Shimoyamada, M.; Imura, T.; Uchida, T.; Okubo, K., A new isoflavone glycoside in soybean seeds (Glycine-Max merrill), Glycitein 7-O-Beta-D-(6"-O-Acetyl)-glucopyranoside. *Agricultural and Biological Chemistry* **1991,** 55, 859-860.

93. Achouri, A.; Boye, J. I.; Belanger, D., Soybean isoflavones: Efficacy of extraction conditions and effect of food type on extractability. *Food Research International* **2005,** 38, 1199-1204.

94. Rostagno, M. A.; Villares, A.; Guillamon, E.; Garcia-Lafuente, A.; Martinez, J. A., Sample preparation for the analysis of isoflavones from soybeans and soy foods. *Journal of Chromatography A* **2009,** 1216, 2-29.

95. Setchell, K. D. R., Phytoestrogens: the biochemistry, physiology, and implications for human health of soy isoflavones. *American Journal of Clinical Nutrition* **1998,** 68, 1333s-1346s.

96. Schwartz, H.; Sontag, G., Determination of secoisolariciresinol, lariciresinol and isolariciresinol in plant foods by high performance liquid chromatography coupled with coulometric electrode array detection. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2006,** 838, 78-85.

97. Mazur, W. M.; Wahala, K.; Rasku, S.; Salakka, A.; Hase, T.; Adlercreutz, H., Lignan and isoflavonoid concentrations in tea and coffee. *British Journal of Nutrition* **1998,** 79, 37-45.

98. Willfor, S.; Hemming, J.; Reunanen, M.; Eckerman, C.; Holmbom, B., Lignans and lipophilic extractives in Norway spruce knots and stemwood. *Holzforschung* **2003,** 57, 27-36.

99. Slanina, J.; Glatz, Z., Separation procedures applicable to lignan analysis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2004,** 812, 215-229.

100. Wynne-Edwards, K. E., Evolutionary biology of plant defenses against herbivory and their predictive implications for endocrine disruptor susceptibility in vertebrates. *Environmental Health Perspectives* **2001,** 109, 443-448.

101. Hughes, C. L., Phytochemical mimicry of reproductive hormones and modulation of herbivore fertility by phytoestrogens. *Environmental Health Perspectives* **1988,** 78, 171-175.

102. Phillips, D. A., Flavonoids - plant signals to soil microbes. *Phenolic Metabolism in Plants* **1992,** 26, 201-231.

103. Hungria, M.; Stacey, G., Molecular signals exchanged between host plants and rhizobia: Basic aspects and potential application in agriculture. *Soil Biology & Biochemistry* **1997,** 29, 819-830.

104. Rolfe, B. G.; Gresshoff, P. M., Genetic-analysis of legume nodule initiation. *Annual Review of Plant Physiology and Plant Molecular Biology* **1988,** 39, 297-319.

105. Hahlbrock, K.; Scheel, D., Physiology and molecular-biology of phenylpropanoid metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology* **1989,** 40, 347-369.

106. Zubieta, C.; He, X.; Dixon, R. A.; Noel, J. P., Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. *Nature Structural Biology* **2001,** 8, 271 - 279

107. Bradbury, R. B.; White, D. E., Estrogens and related substances in plants. *Vitam Horm* **1954,** 12, 207-33.

108. Umphress, S. T.; Murphy, S. P.; Franke, A. A.; Custer, L. J.; Blitz, C. L., Isoflavone content of foods with soy additives. *Journal of Food Composition and Analysis* **2005,** 18, 533-550.

109. Mazur, M.; Duke, J. A.; Wahala, K.; Rasku, S.; Adlercreutz, H., Phytoestrogens in legumes. *American Journal of Clinical Nutrition* **1998,** 68, 1537s-1537s.

110. Liggins, J.; Bluck, L. J. C.; Runswick, S.; Atkinson, C.; Coward, W. A.; Bingham, S. A., Daidzein and genistein content of fruits and nuts. *Journal of Nutritional Biochemistry* **2000,** 11, 326-331.

111. Kuhnle, G. G. C.; Delcaqulla, C.; Aspinall, S. M.; Runswick, S. A.; Mulligan, A. A.; Bingham, S. A., Phytoestrogen content of foods of animal origin: Dairy products, eggs, meat, fish and seafood. *Journal of Agricultural and Food Chemistry* **2008,** 56, 10099-10104.

112. Mazur, W. M.; Duke, J. A.; Wahala, K.; Rasku, S.; Adlercreutz, H., Isoflavonoids and lignans in legumes: Nutritional and health aspects in humans. *Journal of Nutritional Biochemistry* **1998,** 9, 193-200.

113. Thompson, L. U.; Boucher, B. A.; Liu, Z.; Cotterchio, M.; Kreiger, N., Phytoestrogen content of foods consumed in Canada, including isoflavones, lignans, and coumestan. *Nutrition and Cancer-an International Journal* **2006,** 54, 184-201.

114. Kuhnle, G. G. C.; Dell'Aquila, C.; Low, Y. L.; Kussmaul, M.; Bingham, S. A., Extraction and quantification of phytoestrogens in foods using automated solid-phase extraction and LC/MS/MS. *Analytical Chemistry* **2007,** 79, 9234-9239.

115. Lee, S. J.; Ahn, J. K.; Kim, S. H.; Kim, J. T.; Han, S. J.; Jung, M. Y.; Chung, I. M., Variation in isoflavone of soybean cultivars with location and storage duration. *Journal of Agricultural and Food Chemistry* **2003,** 51, 3382-3389.

116. Hoeck, J. A.; Fehr, W. R.; Murphy, P. A.; Welke, G. A., Influence of genotype and environment on isoflavone contents of soybean. *Crop Science* **2000,** 40, 48-51.

117. Coward, L.; Kirk, M.; Albin, N.; Barnes, S., Analysis of plasma isoflavones by reversed-phase HPLC-multiple reaction ion monitoring-mass spectrometry. *Clinica Chimica Acta* **1996,** 247, 121-142.

118. Espin, J. C.; Garcia-Conesa, M. T.; Tomas-Barberan, F. A., Nutraceuticals: Facts and fiction. *Phytochemistry* **2007,** 68, 2986-3008.

119. Lee, H. P.; Gourley, L.; Duffy, S. W.; Esteve, J.; Lee, J.; Day, N. E., Dietary-effects on breast-cancer risk in Singapore. *Lancet* **1991,** 337, 1197-1200.

120. Hirose, K.; Tajima, K.; Hamajima, N.; Inoue, M.; Takezaki, T.; Kuroishi, T.; Yoshida, M.; Tokudome, S., A large-scale, hospital-based case-central study of risk-factors of breast-cancer according to menopausal status. *Japanese Journal of Cancer Research* **1995,** 86, 146-154.

121. Yamamoto, S.; Sobue, T.; Kobayashi, M.; Sasaki, S.; Tsugane, S., Soy, isoflavones, and breast cancer risk in Japan. *Journal of the National Cancer Institute* **2003,** 95, 906-913.

122. Key, T. J.; Sharp, G. B.; Appleby, P. N.; Beral, V.; Goodman, M. T.; Soda, M.; Mabuchi, K., Soya foods and breast cancer risk: a prospective study in Hiroshima and Nagasaki, Japan. *British Journal of Cancer* **1999,** 81, 1248-1256.

123. Wu, A. H.; Wan, P.; Hankin, J.; Tseng, C. C.; Yu, M. C.; Pike, M. C., Adolescent and adult soy intake and risk of breast cancer in Asian-Americans. *Carcinogenesis* **2002,** 23, 1491-1496.

124. Linseisen, J.; Piller, R.; Hermann, S.; Chang-Claude, J., Dietary phytoestrogen intake and premenopausal breast cancer risk in a German case-control study. *International Journal of Cancer* **2004,** 110, 284-290.

125. Dirx, M. J. M.; van den Brandt, P. A.; Goldbohm, R. A.; Lumey, L. H., Diet in adolescence and the risk of breast cancer: results of the Netherlands cohort study. *Cancer Causes & Control* **1999,** 10, 189-199.

126. Touillaud, M. S.; Thiebaut, A. C. M.; Niravong, M.; Boutron-Ruault, M. C.; Clavel-Chapelon, F., No association between dietary phytoestrogens and risk of premenopausal breast cancer in a french cohort study. *Cancer Epidemiology Biomarkers & Prevention* **2006,** 15, 2574-2576.

127. de Kleijn, M. J. J.; van der Schouw, Y. T.; Wilson, P. W. F.; Grobbee, D. E.; Jacques, P. F., Dietary intake of phytoestrogens is associated with a favorable metabolic cardiovascular risk profile in postmenopausal US women: The Framingham Study. *Journal of Nutrition* **2002,** 132, 276-282.

128. Horn-Ross, P. L.; Barnes, S.; Lee, M.; Coward, L.; Mandel, J. E.; Koo, J.; John, E. M.; Smith, M., Assessing phytoestrogen exposure in epidemiologic studies: development of a database (United States). *Cancer Causes & Control* **2000,** 11, 289-298.

129. Strom, S. S.; Yamamura, Y.; Duphorne, C. M.; Spitz, M. R.; Babaian, R. J.; Pillow, P. C.; Hursting, S. D., Phytoestrogen intake and prostate cancer: A case-control study using a new database. *Nutrition and Cancer-an International Journal* **1999,** 33, 20-25.

130. Dalais, F. S.; Rice, G. E.; Wahlqvist, M. L.; Hsu-Hage, B. H. H.; Wattanapenpaiboon, N., Urinary excretion of isoflavonoid phytoestrogens in Chinese and Anglo-Celtic populations in Australia. *Nutrition Research* **1998,** 18, 1703-1709.

131. Keinan-Boker, L.; van Der Schouw, Y. T.; Grobbee, D. E.; Peeters, P. H. M., Dietary phytoestrogens and breast cancer risk. *American Journal of Clinical Nutrition* **2004,** 79, 282-288.

132. Valsta, L. M.; Kilkkinen, A.; Mazur, W.; Nurmi, T.; Lampi, A. M.; Ovaskainen, M. L.; Korhonen, T.; Adlercreutz, H.; Pietinen, P., Phyto-oestrogen database of foods and average intake in Finland. *British Journal of Nutrition* **2003,** 89, S31-S38.

133. Axelson, M.; Sjovall, J.; Gustafsson, B. E.; Setchell, K. D. R., Soya - a dietary source of the non-steroidal estrogen equol in man and animals. *Journal of Endocrinology* **1984,** 102, 49-56.

134. Tham, D. M.; Gardner, C. D.; Haskell, W. L., Clinical review 97 - Potential health benefits of dietary phytoestrogens: A review of the clinical, epidemiological, and mechanistic evidence. *Journal of Clinical Endocrinology and Metabolism* **1998,** 83, 2223-2235.

135. Adlercreutz, H.; Fotsis, T.; Kurzer, M. S.; Wahala, K.; Makela, T.; Hase, T., Isotope-dilution gas-chromatographic mass-spectrometric method for the determination of unconjugated lignans and isoflavonoids in human feces, with preliminary-results in omnivorous and vegetarian women. *Analytical Biochemistry* **1995,** 225, 101-108.

136. Martin, P. M.; Horwitz, K. B.; Ryan, D. S.; Mcguire, W. L., Phyto-estrogen interaction with estrogen receptors in human breast-cancer cells. *Endocrinology* **1978,** 103, 1860-1867.

137. Farmakalidis, E.; Hathcock, J. N.; Murphy, P. A., Estrogenic potency of genistin and daidzin in mice. *Food and Chemical Toxicology* **1985,** 23, 741-745.

138. Cassidy, A., Physiological effects of phyto-oestrogens in relation to cancer and other human health risks. *Proceedings of the Nutrition Society* **1996,** 55, 399-417.

139. Santell, R. C.; Chang, Y. C.; Nair, M. G.; Helferich, W. G., Dietary genistein exerts estrogenic effects upon the uterus, mammary gland and the hypothalamic/pituitary axis in rats. *Journal of Nutrition* **1997,** 127, 263-269.

140. Ratna, W. N., Inhibition of estrogenic stimulation of gene expression by genistein. *Life Sciences* **2002,** 71, 865-877.

141. Markiewicz, M. A.; Kast, W. M., Advances in immunotherapy for prostate cancer. *Advances in Cancer Research, Vol 87* **2003,** 87, 159-194.

142. Messina, M. J.; Dalais, F. S., Soy foods, soy phytoestrogens and bone health. *Menopause at the Millennium* **2000**, 374-380.

143. Hilakivi-Clarke, L.; Cho, E.; deAssis, S.; Olivo, S.; Ealley, E.; Bouker, K. B.; Welch, J. N.; Khan, G.; Clarke, R.; Cabanes, A., Maternal and prepubertal diet, mammary development and breast cancer risk. *J. Nutr.* **2001,** 131, 154S-157.

144. Adlercreutz, H.; Bannwart, C.; Wahala, K.; Makela, T.; Brunow, G.; Hase, T.; Arosemena, P. J.; Kellis, J. T.; Vickery, L. E., Inhibition of human aromatase by mammalian lignans and isoflavonoid phytoestrogens. *Journal of Steroid Biochemistry and Molecular Biology* **1993,** 44, 147-153.

145. Atkinson, C.; Frankenfeld, C. L.; Lampe, J. W., Gut bacterial metabolism of the soy isoflavone daidzein: Exploring the relevance to human health. *Experimental Biology and Medicine* **2005,** 230, 155-170.

146. Rowland, I. R.; Rumney, C. J.; Pool-Zobel, B. L., Colonic microflora and intestinal function. *Proceedings of the 16th International Congress of Nutrition - Nutrition Montreal 97* **1998**, 181-182.

147. Hawkswor.G; Drasar, B. S.; Hill, M. J., Intestinal bacteria and hydrolysis of glycosidic bonds. *Journal of Medical Microbiology* **1971,** 4, 451-&.

148. Heinonen, S.; Wahala, K.; Adlercreutz, H., Identification of isoflavone metabolites dihydrodaidzein, dihydrogenistein, 6 '-OH-O-dma, and cis-4-OH-equol in human urine by gas chromatography-mass spectroscopy using authentic reference compounds. *Analytical Biochemistry* **1999,** 274, 211-219.

149. Meyer, B. J.; Larkin, T. A.; Owen, A. J., Limited lipid-lowering effects of regular consumption of whole soybean foods *Annals of Nutrition and Metabolism* **2004** 48, 67–78.

150. Rowland, I. R.; Wiseman, H.; Sanders, T. A. B.; Adlercreutz, H.; Bowey, E. A., Interindividual variation in metabolism of soy isoflavones and lignans: Influence of habitual diet on equol production by the gut microflora. *Nutrition and Cancer-an International Journal* **2000,** 36, 27-32.

151. Franke, A. A.; Custer, L. J.; Hundahl, S. A., Determinants for urinary and plasma isoflavones in humans after soy intake. *Nutrition and Cancer-an International Journal* **2004,** 50, 141-154.

152. Watanabe, S.; Yamaguchi, M.; Sobue, T.; Takahashi, T.; Miura, T.; Arai, Y.; Mazur, W.; Wahala, K.; Adlercreutz, H., Pharmacokinetics of soybean isoflavones in plasma, urine and feces of men after ingestion of 60 g baked soybean powder (kinako). *Journal of Nutrition* **1998,** 128, 1710-1715.

153. Hutchins, A. M.; Slavin, J. L.; Lampe, J. W., Urinary isoflavonoid phytoestrogen and lignan excretion after consumption of fermented and unfermented soy products. *Journal of the American Dietetic Association* **1995,** 95, 545-551.

154. Cassidy, A.; Faughnan, M.; Hughes, R.; Fraser, C.; Cathcart, A.; Taylor, N.; Setchell, K.; Bingham, S., Hormonal effects of phytoestrogens in postmenopausal women and middle-aged men. *American Journal of Clinical Nutrition* **1998,** 68, 1531s-1531s.

155. Lampe, J. W.; Karr, S. C.; Hutchins, A. M.; Slavin, J. L., Urinary equol excretion with a soy challenge: Influence of habitual diet. *Proceedings of the Society for Experimental Biology and Medicine* **1998,** 217, 335-339.

156. Rostagno, M. A.; Palma, M.; Barroso, C. G., Ultrasound-assisted extraction of soy isoflavones. *Journal of Chromatography A* **2003,** 1012, 119-128.

157. Klejdus, B.; Vitamvasova, D.; Kuban, V., Reversed-phase high-performance liquid chromatographic determination of isoflavones in plant materials after isolation by solid-phase extraction. *Journal of Chromatography A* **1999,** 839, 261-263.

158. Klejdus, B.; Vacek, J.; Adam, V.; Zehnalek, J.; Kizek, R.; Trnkova, L.; Kuban, V., Determination of isoflavones in soybean food and human urine using liquid chromatography with electrochemical detection. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2004,** 806, 101-111.

159. Griffith, A. P.; Collison, M. W., Improved methods for the extraction and analysis of isoflavones from soy-containing foods and nutritional supplements by reversed-phase high-performance liquid chromatography and liquid chromatography-mass spectrometry. *Journal of Chromatography A* **2001,** 913, 397-413.

160. Lin, F.; Giusti, M. M., Effects of solvent polarity and acidity on the extraction efficiency of isoflavones from soybeans (Glycine max). *Journal of Agricultural and Food Chemistry* **2005,** 53, 3795-3800.

161. Muir, A. D.; Westcott, N. D., Quantitation of the lignan secoisolariciresinol diglucoside in baked goods containing flax seed or flax meal. *Journal of Agricultural and Food Chemistry* **2000,** 48, 4048-4052.

162. Naim, M.; Gestetne.B; Zilkah, S.; Birk, Y.; Bondi, A., Soybean isoflavones characterization, determination, and antifungal activity. *Journal of Agricultural and Food Chemistry* **1974,** 22, 806-810.

163. Adlercreutz, H.; Kiuru, P.; Rasku, S.; Wahala, K.; Fotsis, T., An isotope dilution gas chromatographic-mass spectrometric method for the simultaneous assay of estrogens and phytoestrogens in urine. *Journal of Steroid Biochemistry and Molecular Biology* **2004,** 92, 399-411.

164. Adlercreutz, H.; Fotsis, T.; Lampe, J.; Wahala, K.; Makela, T.; Brunow, G.; Hase, T., Quantitative-determination of lignans and isoflavonoids in plasma of omnivorous and vegetarian women by isotope-dilution gas-chromatography mass-spectrometry. *Scandinavian Journal of Clinical & Laboratory Investigation* **1993,** 53, 5-18.

165. Tswett, M., Physico-chemical studies of chlorophyll: The adsorptions. *Ber. Dutsche Botan. Ges.* **1906,** 24, 316-326.

166. James, A. T.; Martin, A. J. P., Gas-liquid partition ofchromatography: separation and micro-estimation of volatile fatty acids from formic acid to dodecanoic acid. *Biochemical Journal* **1952,** 50, 679-690.

167. Herbert, C. G., *Mass spectrometry: Basics*. CRC Press: p 245.

168. Strobel, H. A.; Heineman, W. R., Chemical Instrumentation: A Systemic Approach. In John Wiley & Sons: New York, 1989.

169. Hamilton, R. J.; Sewell, P. A., *Introduction to high performance liquid chromatography*. 2nd ed. ed.; Chapman and Hall: London, 1982.

170. Kusano, M.; Fukushima, A.; Kobayashi, M.; Hayashi, N.; Jonsson, P.; Moritz, T.; Ebana, K.; Saito, K., Application of a metabolomic method combining one-dimensional and two-dimensional gas chromatography-time-of-flight/mass spectrometry to metabolic phenotyping of natural variants in rice. *Journal of Chromatography B* **2007,** 855, 71-79.

171. Wong, E.; Mortimer, P. T.; Geissman, T. A., Flavonoid constituents of Cicer arietinum. *Phytochemistry* **1965,** 4, 89-95.

172. Axelson, M.; Graham, C. E.; Sjovall, J., Identification and quantitation of steroids in sulfate fractions from plasma of pregnant chimpanzee, orangutan, and rhesus-monkey. *Endocrinology* **1984,** 114, 337-344.

173. Farmakalidis, E.; Murphy, P. A., Isolation of 6"-O-Acetylgenistin and 6"-O-Acetyldaidzin from Toasted Defatted Soyflakes. *Journal of Agricultural and Food Chemistry* **1985,** 33, 385-389.

174. Kitada, Y.; Mizobuchi, M.; Ueda, Y.; Nakazawa, H., Analysis of isoflavones in Puerariae-Radix by high-performance liquid-chromatography with amperometric detection. *Journal of Chromatography* **1985,** 347, 438-442.

175. Adlercreutz, H.; Musey, P. I.; Fotsis, T.; Bannwart, C.; Wahala, K.; Makela, T.; Brunow, G.; Hase, T., Identification of lignans and phytoestrogens in urine of chimpanzees. *Clinica Chimica Acta* **1986,** 158, 147-154.

176. Kitada, Y.; Ueda, Y.; Yamamoto, M.; Ishikawa, M.; Nakazawa, H.; Fujita, M., Determination of isoflavones in soy bean by high-performance liquid-chromatography with ampeometric detection. *Journal of Chromatography* **1986,** 366, 403-406.

177. Setchell, K. D. R.; Welsh, M. B.; Lim, C. K., High-performance liquid-chromatographic analysis of phytoestrogens in soy protein preparations with ultraviolet, electrochemical and thermospray mass-spectrometric detection. *Journal of Chromatography* **1987,** 386, 315-323.

178. Lundh, T. J. O.; Pettersson, H.; Kiessling, K. H., Liquid-chromatographic determination of the estrogens daidzein, formononetin, coumestrol, and equol in bovine blood-plasma and urine. *Journal of the Association of Official Analytical Chemists* **1988,** 71, 938-941.

179. Wang, G.; Kuan, S. S.; Francis, O. J.; Ware, G. M.; Carman, A. S., A simplified HPLC method for the determination of phytoestrogens in soybean and its processed products. *Journal of Agricultural and Food Chemistry* **1990,** 38, 185-190.

180. Lee, N., Phytoestrogens as bioactive ingredients in functional foods: Canadian regulatory update. *Journal of Aoac International* **2006,** 89, 1135-1137.

181. Valentin-Blasini, L.; Blount, B. C.; Caudill, S. P.; Needham, L. L., Urinary and serum concentrations of seven phytoestrogens in a human reference population subset. *Journal of Exposure Analysis and Environmental Epidemiology* **2003,** 13, 276-282.

182. Ward, H.; Chapelais, G.; Kuhnle, G. G. C.; Luben, R.; Khaw, K. T.; Bingham, S., Urinary and serum biomarkers of phytoestrogen exposure are not associated with breast cancer risk in the European Prospective into Cancer Norfolk study. *Breast Cancer Research* **2008,** 10, S28-S28.

183. Wu, Q. L.; Wang, M. F.; Simon, J. E., Analytical methods to determine phytoestrogenic compounds. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2004,** 812, 325-355.

184. Charles, G. D.; Gennings, C.; Tornesi, B.; Kan, H. L.; Zacharewski, T. R.; Gollapudi, B. B.; Carney, E. W., Analysis of the interaction of phytoestrogens and

synthetic chemicals: An in vitro/in vivo comparison. *Toxicology and Applied Pharmacology* **2007,** 218, 280-288.

185. Rice, S.; Whitehead, S. A., Phytoestrogens and breast cancer -promoters or protectors? *Endocr Relat Cancer* **2006,** 13, 995-1015.

186. Saarinen, N. M.; Bingham, C.; Lorenzetti, S.; Mortensen, A.; Makela, S.; Penttinen, P.; Sorensen, I. K.; Valsta, L. M.; Virgili, F.; Vollmer, G.; Warri, A.; Zierau, O., Tools to evaluate estrogenic potency of dietary phytoestrogens: A consensus paper from the eu thematic network "Phytohealth" (QLKI-2002-2453). *Genes and Nutrition* **2006,** 1, 143-158.

187. Gustafsson, J. A., Comments to the paper "Tools to evaluate estrogenic potency Of dietary phytoestrogens: A consensus paper from the EU thematic network "Phytohealth" (QLKI-2002-2453)". *Genes and Nutrition* **2006,** 1, 159-160.

188. Wojcik-Gladysz, A.; Romanowicz, K.; Misztal, T.; Polkowska, J., Oestrogen-like effect of genistein on follicle-stimulating hormone release in ovariectomized ewes. *Journal of Animal and Feed Sciences* **2006,** 15, 579-589.

189. Luine, V.; Attalla, S.; Mohan, G.; Costa, A.; Frankfurt, M., Dietary phytoestrogens enhance spatial memory and spine density in the hippocampus and prefrontal cortex of ovariectomized rats. *Brain Research* **2006,** 1126, 183-187.

190. Misztal, T.; Wafikowska, M.; Gorski, K.; Romanowicz, K., Central estrogen-like effect of genistein on growth hormone secretion in the ewe. *Acta Neurobiologiae Experimentalis* **2007,** 67, 411-419.

191. Chien, H. L.; Huang, H. Y.; Chou, C. C., Transformation of isoflavone phytoestrogens during the fermentation of soymilk with lactic acid bacteria and bifidobacteria. *Food Microbiology* **2006,** 23, 772-778.

192. Lee, S. H.; Jung, B. H.; Kim, S. Y.; Chung, B. C., Determination of phytoestrogens in traditional medicinal herbs using gas chromatography-mass spectrometry. *Journal of Nutritional Biochemistry* **2004,** 15, 452-460.

193. Lee, Y. J.; Adlercreutz, H.; Kwon, H., Quantitative analysis of isoflavones and lignans in sea vegetables consumed in Korea using isotope dilution gas chromatography-mass spectrometry. *Food Science and Biotechnology* **2006,** 15, 102-106.

194. Moors, S.; Blaszkewicz, M.; Bolt, H. M.; Degen, G. H., Simultaneous determination of daidzein, equol, genistein and bisphenol A in human urine by a fast and simple method using SPE and GC-MS. *Molecular Nutrition & Food Research* **2007,** 51, 787-798.

195. Promberger, A.; Dornstauder, E.; Fruhwirth, C.; Schmid, E. R.; Jungbauer, A., Determination of estrogenic activity in beer by biological and chemical means. *Journal of Agricultural and Food Chemistry* **2001,** 49, 633-640.

196. Wu, Q.; Wang, M.; Sciarappa, W. J.; Simon, J. E., LC/UV/ESI-MS analysis of isoflavones in edamame and tofu soybeans. *Journal of Agricultural and Food Chemistry* **2004,** 52, 2763-2769.

197. Kishikawa, N.; Ohyama, K.; Kuroda, N., Human biomonitoring of endocrine disrupting chemicals by HPLC methods. *Current Analytical Chemistry* **2006,** 2, 77-88.

198. Mullner, C.; Sontag, G., Determination of some phytoestrogens in soybeans and their processed products with HPLC and coulometric electrode array detection. *Fresenius Journal of Analytical Chemistry* **1999,** 364, 261-265.

199. Barnes, S.; Coward, L.; Kirk, M.; Sfakianos, J., HPLC mass spectrometry analysis of isoflavones. *Proceedings of the Society for Experimental Biology and Medicine* **1998,** 217, 254-262.

200. Erbs, M.; Hoerger, C. C.; Hartmann, N.; Bucheli, T. D., Quantification of six phytoestrogens at the nanogram per liter level in aqueous environmental samples using C-13(3)-labeled internal standards. *Journal of Agricultural and Food Chemistry* **2007,** 55, 8339-8345.

201. Dinelli, G.; Aloisio, I.; Bonetti, A.; Marotti, I.; Cifuentes, A., Compositional changes induced by UV-B radiation treatment of common bean and soybean seedlings monitored by capillary electrophoresis with diode array detection. *Journal of Separation Science* **2007,** 30, 604-611.

202. Kok, S. J.; Kristenson, E. M.; Gooijer, C.; Velthorst, N. H.; Brinkman, U. A. T., Wavelength-resolved laser-induced fluorescence detection in capillary electrophoresis: Naphthalenesulphonates in river water. *Journal of Chromatography A* **1997,** 771, 331-341.

203. Chu, Q. C.; Wu, T.; Fu, L.; Ye, J. N., Determination and comparison of phytoestrogens in both crude and parched Flos sophorae immaturus by capillary electrophoresis with electrochemical detection. *Microchimica Acta* **2004,** 148, 311-315.

204. Chen, G.; Zhang, J. X.; Ye, J. N., Determination of puerarin, daidzein and rutin in Pueraria lobata (Wild.) Ohwi by capillary electrophoresis with electrochemical detection. *Journal of Chromatography A* **2001,** 923, 255-262.

205. Aramendia, M. A.; Garcia, I.; Lafont, F.; Marinas, J. M., Determination of isoflavones using capillary electrophoresis in combination with electrospray mass-spectrometry. *Journal of Chromatography A* **1995,** 707, 327-333.

206. Pettersson, H.; Kiessling, K. H., Liquid-chromatographic determination of the plant estrogens coumestrol and isoflavones in animal feed. *Journal of the Association of Official Analytical Chemists* **1984,** 67, 503-506.

207. Chu, Q. C.; Fu, L.; Wu, T.; Ye, J. N., Simultaneous determination of phytoestrogens in different medicinal parts of Sophora japonica L. by capillary electrophoresis with electrochemical detection. *Biomedical Chromatography* **2005,** 19, 149-154.

208. Mellenthin, O.; Galensa, R., Analysis of polyphenols using capillary zone electrophoresis and HPLC: Detection of soy, lupin, and pea protein in meat products. *Journal of Agricultural and Food Chemistry* **1999,** 47, 594-602.

209. Iizuka, K.; Aishima, T., Differentiation of soy sauce by pattern recognition analysis of mid- and near-IR spectra. *Journal of Food Composition and Analysis* **1999,** 12, 197-209.

210. Wang, J.; Sporns, P., MALDI-TOF MS analysis of isoflavones in soy products. *Journal of Agricultural and Food Chemistry* **2000,** 48, 5887-5892.

211. Adlercreutz, H.; Wang, G. J. J.; Lapcik, O.; Hampl, R.; Wahala, K.; Makela, T.; Lusa, K.; Talme, M.; Mikola, H., Time-resolved fluoroimmunoassay for plasma enterolactone. *Analytical Biochemistry* **1998,** 265, 208-215.

212. Lapcik, O.; Hampl, R.; Hill, M.; Wahala, K.; Al Maharik, N.; Adlercreutz, H., Radioimmunoassay of free genistein in human serum. *Journal of Steroid Biochemistry and Molecular Biology* **1998,** 64, 261-268.

213. Hampl, R.; Lapcik, O.; Wahala, K.; Starka, L.; Adlercreutz, H., Radioimmunoassay analysis of phytoestrogens of isoflavonoid series. *Chemicke Listy* **1998,** 92, 44-50.

214. Kohen, F.; Gayer, B.; Amir-Zaltsman, Y.; Ben-Hur, H.; Thomas, E.; Lu, L. J. W., A nonisotopic enzyme-based immunoassay for assessing human exposure to genistein. *Nutrition and Cancer-an International Journal* **1999,** 35, 96-103.

215. Pauling, L.; Robinson, A. B.; Teranishi, R.; Cary, P., Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proceedings of the National Academy of Sciences of the United States of America* **1971,** 68, 2374-2376.

216. Gates, S. C.; Sweeley, C. C., Quantitative metabolic profiling based on gas chromatography. *Clin Chem* **1978,** 24, 1663-1673.

217. Pasikanti, K. K.; Ho, P. C.; Chan, E. C. Y., Gas chromatography/mass spectrometry in metabolic profiling of biological fluids. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2008,** 871, 202-211.

218. Pasikanti, K. K.; Ho, P. C.; Chan, E. C. Y., Development and validation of a gas chromatography/mass spectrometry metabonomic platform for the global profiling of urinary metabolites. *Rapid Communications in Mass Spectrometry* **2008,** 22, 2984-2992.

219. Trygg, J.; Lundstedt, T., Chemometric approaches to metabonic studies. In *The Handbook of Metabonomics and Metabolomics* Elsevier Science B.V.: Amsterdam 2007; pp 171-199.

220. Wishart, D. S., Applications of metabolomics in drug discovery and development. *Drugs in R&D* **2008,** 9, 307-322.

221. Guarner, F., Microecology as a target for therapeutic intervention in inflammatory bowel disease. *Idrugs* **2003,** 6, 868-873.

222. Guarner, F.; Malagelada, J. R., Role of bacteria in experimental colitis. *Best Practice & Research in Clinical Gastroenterology* **2003,** 17, 793-804.

223. Tannock, G. W., Microecology of the gastrointestinal tract in relation to lactic acid bacteria. *International Dairy Journal* **1995,** 5, 1059-1070.

224. Xie, Y. S.; Pan, G. X.; Gao, X. M.; Liu, C. X., Application of high performance liquid chromatographic technique in metabonomics studies. *Chinese Journal of Analytical Chemistry* **2006,** 34, 1644-1648.

225. Wishart, D. S.; Lewis, M. J.; Morrissey, J. A.; Flegel, M. D.; Jeroncic, K.; Xiong, Y. P.; Cheng, D.; Eisner, R.; Gautam, B.; Tzur, D.; Sawhney, S.; Bamforth, F.; Greiner, R.; Li, L., The human cerebrospinal fluid metabolome. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2008,** 871, 164-173.

226. Jolliffe, I. T., *Principal component analysis (Series)*. 2nd ed ed.; Springer: NY, 2002; Vol. XXIX, p 487 p.28.

227. Odunsi, K.; Wollman, R. M.; Ambrosone, C. B.; Hutson, A.; McCann, S. E.; Tammela, J.; Geisler, J. P.; Miller, G.; Sellers, T.; Cliby, W.; Qian, F.; Keitz, B.; Intengan, M.; Lele, S.; Alderfer, J. L., Detection of epithelial ovarian cancer using H-1-NMR-based metabonomics. *International Journal of Cancer* **2005,** 113, 782-788.

228. Constantinou, M. A.; Papakonstantinou, E.; Benaki, D.; Spraul, M.; Shulpis, K.; Koupparis, M. A.; Mikros, E., Application of nuclear magnetic resonance spectroscopy combined with principal component analysis in detecting inborn errors of metabolism using blood spots: a metabonomic approach. *Analytica Chimica Acta* **2004,** 511, 303-312.

229. Carty, D. M.; Delles, C.; Dominiczak, A. F., Novel biomarkers for predicting pre-eclampsia. *Trends in Cardiovascular Medicine* **2008,** 18, 186-194.

230. Solanky, K. S.; Bailey, N. J.; Beckwith-Hall, B. M.; Bingham, S.; Davis, A.; Holmes, E.; Nicholson, J. K.; Cassidy, A., Biofluid H-1 NMR-based metabonomic techniques in nutrition research metabolic effects of dietary isoflavones in humans. *Journal of Nutritional Biochemistry* **2005,** 16, 236-244.

231. Solanky, K. S.; Bailey, N. J. C.; Beckwith-Hall, B. M.; Davis, A.; Bingham, S.; Holmes, E.; Nicholson, J. K.; Cassidy, A., Application of biofluid H-1 nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile. *Analytical Biochemistry* **2003,** 323, 197-204.

232. Nicholson, J. K.; Buckingham, M. J.; Sadler, P. J., High-resolution H-1-NMR studies of vertebrate blood and plasma. *Biochemical Journal* **1983,** 211, 605-615.

233. Bales, J. R.; Higham, D. P.; Howe, I.; Nicholson, J. K.; Sadler, P. J., Use of high-resolution proton nuclear magnetic-resonance spectroscopy for rapid multi-component analysis of urine. *Clinical Chemistry* **1984,** 30, 426-432.

234. Bales, J. R.; Sadler, P. J.; Nicholson, J. K.; Timbrell, J. A., Urinary-excretion of acetaminophen and its metabolites as studied by proton NMR-spectroscopy. *Clinical Chemistry* **1984,** 30, 1631-1636.

235. Fiehn, O., Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology* **2002,** 48, 155-171.

236. Robertson, D. G., Metabonomics in toxicology: A review. *Toxicological Sciences* **2005,** 85, 809-822.

237. Lindon, J. C.; Holmes, E.; Bollard, M. E.; Stanley, E. G.; Nicholson, J. K., Metabonomics technologies and their applications in physiological monitoring, drug safety assessment and disease diagnosis. *Biomarkers* **2004,** 9, 1-31.

238. Lindon, J. C.; Nicholson, J. K.; Everett, J. R., NMR spectroscopy of biofluids. *Annual Reports on Nmr Spectroscopy, Vol 38* **1999,** 38, 1-88.

239. van der Greef, J.; Stroobant, P.; van der Heijden, R., The role of analytical sciences medical systems biology. *Current Opinion in Chemical Biology* **2004,** 8, 559-565.

240. Garrod, S.; Humpfer, E.; Spraul, M.; Connor, S. C.; Polley, S.; Connelly, J.; Lindon, J. C.; Nicholson, J. K.; Holmes, E., High-resolution magic angle spinning H-1 NMR spectroscopic studies on intact rat renal cortex and medulla. *Magnetic Resonance in Medicine* **1999,** 41, 1108-1118.

241. Shen, Q.; Li, X.; Qiu, Y. P.; Su, M. M.; Liu, Y. M.; Li, H. K.; Wang, X. Y.; Zou, X. Y.; Yan, C. H.; Yu, L.; Li, S.; Wan, C. L.; He, L.; Jia, W., Metabonomic and metallomic profiling in the amniotic fluid of malnourished pregnant rats. *Journal of Proteome Research* **2008,** 7, 2151-2157.

242. Zhang, Q.; Wang, G. J.; Du, Y.; Zhu, L. L.; Jiye, A., GC/MS analysis of the rat urine for metabonomic research. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2007,** 854, 20-25.

243. Xin, H.; Yi-Fei, G.; Ke, Y.; Yi-Yu, C., Gas chromatography-mass spectrometry based on metabonomic study of carbon tetrachloride-induced acute liver injury in mice. *Chinese Journal of Analytical Chemistry* **2007,** 35, 1736-1740.

244. Yuan, K. L.; Kong, H. W.; Guan, Y. F.; Yang, J.; Xu, G. X., A GC-based metabonomics investigation of type 2 diabetes by organic acids metabolic profile. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2007,** 850, 236-240.

245. Wiklund, S.; Johansson, E.; Sjostrom, L.; Mellerowicz, E. J.; Edlund, U.; Shockcor, J. P.; Gottfries, J.; Moritz, T.; Trygg, J., Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry* **2008,** 80, 115-122.

246. Sieber, M.; Dekant, W., An integrated GC-MS method for metabonomics of biofluids. *Naunyn-Schmiedebergs Archives of Pharmacology* **2007,** 375, 98-98.

247. Sieber, M.; Wagner, S.; Rached, E.; Amberg, A.; Mally, A.; Dekant, W., A combined GC-MS, LC-MS and NMR metabonomics approach for early detection of ochratoxin a nephrotoxicity. *Naunyn-Schmiedebergs Archives of Pharmacology* **2008,** 377, 74-74.

248. Ni, Y.; Su, M. M.; Qiu, Y. P.; Chen, M. J.; Liu, Y. M.; Zhao, A. H.; Jia, W., Metabolic profiling using combined GC-MS and LC-MS provides a systems understanding of aristolochic acid-induced nephrotoxicity in rat. *Febs Letters* **2007,** 581, 707-711.

249. Welthagen, W.; Shellie, R. A.; Spranger, J.; Ristow, M.; Zimmermann, R.; Fiehn, O., Comprehensive two-dimensional gas chromatography-time-of-flight mass

spectrometry (GC x GC-TOF) for high resolution metabolomics: biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice. *Metabolomics* **2005,** 1, 65-73.

250. Oh, C.; Huang, X.; Regnier, F. E.; Buck, C.; Zhang, X., Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry peak sorting algorithm. *Journal of Chromatography A* **2008,** 1179, 205-215.

251. Huang, X. D.; Regnier, F. E., Differential metabolomics using stable isotope labeling and two-dimensional gas chromatography, with time-of-flight mass spectrometry. *Analytical Chemistry* **2008,** 80, 107-114.

252. Idborg, H.; Zamani, L.; Edlund, P. O.; Schuppe-Koistinen, I.; Jacobsson, S. P., Metabolic fingerprinting of rat urine by LC/MS Part 1. Analysis by hydrophilic interaction liquid chromatography-electrospray ionization mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2005,** 828, 9-13.

253. Chen, Y. H.; Chen, C. H.; Lin, C. J.; Chen, C. C., Metabonomic Study with a High Performance Liquid Chromatography Coupling to a Triple Quadruple Mass Spectrometer to Identify Biomarkers from Urine of High-fat Fed and Streptozotocin Treated Rats. *Journal of Food and Drug Analysis* **2009,** 17, 28-35.

254. Guy, P. A.; Tavazzi, I.; Bruce, S. J.; Ramadan, Z.; Kochhar, S., Global metabolic profiling analysis on human urine by UPLC-TOFMS: Issues and method validation in nutritional metabolomics. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2008,** 871, 253-260.

255. Wong, M. C. Y.; Lee, W. T. K.; Wong, J. S. Y.; Frost, G.; Lodge, J., An approach towards method development for untargeted urinary metabolite profiling in metabonomic research using UPLC/QToF MS. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2008,** 871, 341-348.

256. Robertson, D. G.; Reily, M. D.; Cantor, G. H.; John, C. L.; Jeremy, K. N.; Elaine, H., Metabonomics in preclinical pharmaceutical discovery and development. In *The Handbook of Metabonomics and Metabolomics* Elsevier Science B.V.: Amsterdam, 2007; pp 241-277.

257. Weljie, A. M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C. M., Targeted profiling: Quantitative analysis of H-1 NMR metabolomics data. *Analytical Chemistry* **2006,** 78, 4430-4442.

258. Penders, J.; Fiers, T.; Giri, M.; Wuyts, B.; Ysewyn, L.; Delanghe, J. R., Quantitative measurement of ketone bodies in urine using reflectometry. *Clinical Chemistry and Laboratory Medicine* **2005,** 43, 724-729.

259. Gibney, M. J.; Walsh, M.; Brennan, L.; Roche, H. M.; German, B.; van Ommen, B., Metabolomics in human nutrition: opportunities and challenges. *American Journal of Clinical Nutrition* **2005,** 82, 497-503.

260. Watkins, S. M.; German, J. B., Toward the implementation of metabolomic assessments of human health and nutrition. *Current Opinion in Biotechnology* **2002,** 13, 512-516.

261. German, J. B.; Roberts, M. A.; Fay, L.; Watkins, S. M., Metabolomics and individual metabolic assessment: The next great challenge for nutrition. *Journal of Nutrition* **2002,** 132, 2486-2487.

262. Wishart, D. S., Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology* **2008,** 19, 482-493.

263. Fardet, A.; Canlet, C.; Gottardi, G.; Lyan, B.; Llorach, R.; Remesy, C.; Mazur, A.; Paris, A.; Scalbert, A., Whole-grain and refined wheat flours show distinct metabolic profiles in rats as assessed by a H-1 NMR-based metabonomic approach. *Journal of Nutrition* **2007,** 137, 923-929.

264. Fardet, A.; Llorach, R.; Martin, J. F.; Besson, C.; Pujos, E.; Scalbert, A., Characterization of the metabolic effects of catechin in rats fed hyperlipidic diets by using a LC-QTof-based metabonomic approach. *Annals of Nutrition and Metabolism* **2007,** 51, 382-383.

265. Solanky, K. S.; Bailey, N. J. C.; Holmes, E.; Lindon, J. C.; Davis, A. L.; Mulder, T. P. J.; Van Duynhoven, J. P. M.; Nicholson, J. K., NMR-based metabonomic studies on the biochemical effects of epicatechin in the rat. *Journal of Agricultural and Food Chemistry* **2003,** 51, 4139-4145.

266. Solanky, K. S.; Bailey, N. J. C.; Holmes, E.; Lindon, J. C.; Nicholson, J. K.; Davis, A. L.; Mulder, T. P. J.; Van Duynhoven, J., NMR and pattern recognition studies on the biochemistry of black tea-derived polyphenolic compounds in the rat. *Biologically-Active Phytochemicals in Food* **2001**, 140-142.

267. Law, W. S.; Huang, P. Y.; Ong, E. S.; Ong, C. N.; Li, S. F. Y.; Pasikanti, K. K.; Chan, E. C. Y., Metabonomics investigation of human urine after ingestion of green tea with gas chromatography/mass spectrometry, liquid chromatography/mass spectrometry and H-1 NMR spectroscopy. *Rapid Communications in Mass Spectrometry* **2008,** 22, 2436-2446.

268. Watkins, S. M.; Reifsnyder, P. R.; Pan, H.; German, J. B.; Leiter, E. H., Lipid metabolome-wide effects of the PPAR gamma agonist rosiglitazone. *Journal of Lipid Research* **2002,** 43, 1809-1817.

269. Zheng, X.; Shen, J.; Liu, Q.; Wang, S.; Cheng, Y.; Qu, H., Plasma fatty acids metabolic profiling analysis of coronary heart disease based on GC-MS and pattern recognition. *Journal of Pharmaceutical and Biomedical Analysis* **2009,** 49, 481-486.

270. Brindle, J. T.; Antti, H.; Holmes, E.; Tranter, G.; Nicholson, J. K.; Bethell, H. W. L.; Clarke, S.; Schofield, P. M.; McKilligin, E.; Mosedale, D. E.; Grainger, D. J., Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using H-1-NMR-based metabonomics. *Nature Medicine* **2002,** 8, 1439-1444.

271. Sanchez-Moreno, C.; Cano, M. P.; de Ancos, B.; Plaza, L.; Olmedilla, B.; Granado, F.; Martin, A., Consumption of High-Pressurized Vegetable Soup Increases Plasma Vitamin C and Decreases Oxidative Stress and Inflammatory Biomarkers in Healthy Humans. *J. Nutr.* **2004,** 134, 3021-3025.

272. Castilla, P.; Echarri, R.; Davalos, A.; Cerrato, F.; Ortega, H.; Teruel, J. L.; Lucas, M. F.; Gomez-Coronado, D.; Ortuno, J.; Lasuncion, M. A., Concentrated red grape juice

exerts antioxidant, hypolipidemic, and antiinflammatory effects in both hemodialysis patients and healthy subjects. *American Journal of Clinical Nutrition* **2006,** 84, 252-262.

273. Thompson, H. J.; Heimendinger, J.; Diker, A.; O'Neill, C.; Haegele, A.; Meinecke, B.; Wolfe, P.; Sedlacek, S.; Zhu, Z. J.; Jiang, W. Q., Dietary botanical diversity affects the reduction of oxidative biomarkers in women due to high vegetable and fruit intake. *Journal of Nutrition* **2006,** 136, 2207-2212.

274. Bogani, P.; Galli, C.; Villa, M.; Visioli, F., Postprandial anti-inflammatory and antioxidant effects of extra virgin olive oil. *Atherosclerosis* **2007,** 190, 181-186.

275. Garg, M. L.; Blake, R. J.; Wills, R. B. H.; Clayton, E. H., Macadamia nut consumption modulates favourably risk factors for coronary artery disease in hypercholesterolemic subjects. *Lipids* **2007,** 42, 583-587.

276. Schauer, N.; Steinhauser, D.; Strelkov, S.; Schomburg, D.; Allison, G.; Moritz, T.; Lundgren, K.; Roessner-Tunali, U.; Forbes, M. G.; Willmitzer, L.; Fernie, A. R.; Kopka, J., GC-MS libraries for the rapid identification of metabolites in complex biological samples. *Febs Letters* **2005,** 579, 1332-1337.

277. Xu, X.; Duncan, A. M.; Merz, B. E.; Kurzer, M. S., Effects of soy isoflavones on estrogen and phytoestrogen metabolism in premenopausal women. *Cancer Epidemiology Biomarkers & Prevention* **1998,** 7, 1101-1108.

278. deRooij, B. M.; Boogaard, P. J.; Rijksen, D. A.; Commandeur, J. N. M.; Vermeulen, N. P. E., Urinary excretion of N-acetyl-S-allyl-L-cysteine upon garlic consumption by human volunteers. *Archives of Toxicology* **1996,** 70, 635-639.

279. Wang, Y.; Tang, H.; Nicholson, J. K.; Hylands, P. J.; Sampson, J.; Holmes, E., A metabonomic strategy for the detection of the metabolic effects of chamomile (Matricaria recutita L.) ingestion. *Journal of Agricultural and Food Chemistry* **2005,** 53, 191-196.

280. Lawrence, J. R.; Peter, R.; Baxter, G. J.; Robson, J.; Graham, A. B.; Paterson, J. R., Urinary excretion of salicyluric and salicylic acids by non-vegetarians, vegetarians, and patients taking low dose aspirin. *Journal of Clinical Pathology* **2003,** 56, 651-653.

281. Mullen, W.; Boitier, A.; Stewart, A. J.; Crozier, A., Flavonoid metabolites in human plasma and urine after the consumption of red onions: analysis by liquid chromatography with photodiode array and full scan tandem mass spectrometric detection. *Journal of Chromatography A* **2004,** 1058, 163-168.

282. Schneider, H.; Ma, L.; Glatt, H., Extractionless method for the determination of urinary caffeine metabolites using high-performance liquid chromatography coupled with tandem mass spectrometry. *Journal of Chromatography B* **2003,** 789, 227-237.

283. Smith, D.; Ismail, K. M. K.; Diskin, A. M.; Chapman, G.; Magnay, J. L.; Spanel, P.; O'Brien, S., Increase of acetone emitted by urine in relation to ovulation. *Acta Obstetricia Et Gynecologica Scandinavica* **2006,** 85, 1008-1011.

284. Stryer, L., *Biochemistry*. Third Edition ed.; W.H. Freeman and company: New York, 1988.

285. Munir, S.; Bhatty, N.; Almas, K.; Nawaz, M.; Manzoor, F.; Husain, S. J., Estimation of protein metabolites endogenous creatinine and urea in the blood of male volunteers. *Pakistan Journal of Medical Research* **2003,** 42, 157-159.

286. Valltton, M. B., Kidney function and concentration of BUN in human body. In *Division of Endocrinology*, Department of Medicine. University Cantonal Hospital. pp 89-91: Geneva, 1998.

287. Dzurik, R., Hippurate participates in the correction of metabolic acidosis. In *Kidney International 59*, 2001; pp S278-281.

288. Malyusz, M., Renal physiology. *Biochemistry* **1994,** 17, 307-315.

289. Walker, A. R. P.; Adam, F. I.; Walker, B. F., Breast cancer in black African women: a changing situation. *Journal of the Royal Society for the Promotion of Health* **2004,** 124, 81-85.

290. Wang, H. J.; Murphy, P. A., Isoflavone composition of American and Japanese soybeans in Iowa: effects of variety, crop year and location. *Journal of Agricultural and Food Chemistry* **1994,** 42, 1674-1677.

291. Lu, L. J. W.; Anderson, K. E., Sex and long-term soy diets affect the metabolism and excretion of soy isoflavones in humans. *American Journal of Clinical Nutrition* **1998,** 68, 1500s-1504s.

292. Bowey, E. A.; Rowland, I. R.; Adlercreutz, H.; Sanders, T. A. B.; Wiseman, H., Interindividual variation in absorption, metabolism and excretion of phytoestrogens. *Cost 916 - Polyphenols in Food* **1998**, 81-81.

293. Chukwumah, Y. C.; Walker, L. T.; Verghese, M.; Bokanga, M.; Ogutu, S.; Alphonse, K., Comparison of extraction methods for the quantification of selected phytochemicals in peanuts (Arachis hypogaea). *Journal of Agricultural and Food Chemistry* **2007,** 55, 285-290.

294. Choi, Y. B.; Rhee, J. S.; Lee, Y. B.; Nam, S. Y.; Kim, K. S., Extraction of isoflavones from soybean hypocotyl using aqueous ethanol. *Food Science and Biotechnology* **2004,** 13, 719-723.

295. Cheng, W. Y.; Ku, Y. H.; Huang, C. J., Isolation and identification of novel estrogenic compounds in yam tuber (Dioscorea alata cv. Tainung No. 2). *Journal of Agricultural and Food Chemistry* **2007,** 55, 7350-7358.

296. Cherdshewasart, W.; Sriwatcharakul, S., Major isoflavonoid contents of the 1-year-cultivated phytoestrogen-rich herb, Pueraria mirifica. *Bioscience Biotechnology and Biochemistry* **2007,** 71, 2527-2533.

297. Fernandes, A. M. A. P.; Prado, A. L.; Barata, L. E. S.; Paulo, M. Q.; Azevedo, N. R.; Ferri, P. H., A method to separate lignoids from Virola leaves. *Phytochemical Analysis* **1997,** 8, 18-21.

298. Hsu, Y. T.; Wu, C. J.; Chen, J. M.; Yang, Y. C.; Wang, S. Y., The presence of three isoflavonoid compounds in Psoralea corylifolia. *Journal of Chromatographic Science* **2001,** 39, 441-444.

299. Klejdus, B.; Mikelova, R.; Adam, V.; Zehnalek, J.; Vacek, J.; Kizek, R.; Kuban, V., Liquid chromatographic-mass spectrometric determination of genistin and daidzin in soybean food samples after accelerated solvent extraction with modified content of extraction cell. *Analytica Chimica Acta* **2004,** 517, 1-11.

300. Lapcik, O.; Klejdus, B.; Davidova, M.; Kokoska, L.; Kuban, V.; Moravcova, J., Isoflavonoids in the Rutaceae family: Fortunella obovata, Murraya paniculata and four citrus species. *Phytochemical Analysis* **2004,** 15, 293-299.

301. Antonelli, M. L.; Faberi, A.; Pastorini, E.; Samperi, R.; Lagana, A., Simultaneous quantitation of free and conjugated phytoestrogens in Leguminosae by liquid chromatography-tandem mass spectrometry. *Talanta* **2005,** 66, 1025-1033.

302. Klejdus, B.; Mikelova, R.; Petrlova, J.; Potesil, D.; Adam, V.; Stiborova, M.; Hodek, P.; Vacek, J.; Kizek, R.; Kuban, V., Evaluation of isoflavone aglycon and glycoside distribution in soy plants and soybeans by fast column high-performance liquid chromatography coupled with a diode-array detector. *Journal of Agricultural and Food Chemistry* **2005,** 53, 5848-5852.

303. Klejdus, B.; Mikelova, R.; Petrlova, J.; Potesil, D.; Adam, V.; Stiborova, M.; Hodek, P.; Vacek, J.; Kizek, R.; Kuban, V., Determination of isoflavones in soy bits by fast column high-performance liquid chromatography coupled with UV-visible diode-array detection. *Journal of Chromatography A* **2005,** 1084, 71-79.

304. Krenn, L.; Unterrieder, I.; Ruprechter, R., Quantification of isoflavones in red clover by high-performance liquid chromatography. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002,** 777, 123-128.

305. Wu, A. H.; Stanczyk, F. Z.; Seow, A.; Lee, H. P.; Yu, M. C., Soy intake and other lifestyle determinants of serum estrogen levels among postmenopausal chinese women in Singapore. *Cancer Epidemiology Biomarkers & Prevention* **2002,** 11, 844-851.

306. van Erp-Baart, M. A. J.; Brants, H. A. M.; Kiely, M.; Mulligan, A.; Turrini, A.; Sermoneta, C.; Kilkkinen, A.; Valsta, L. M., Isoflavone intake in four different European countries: the VENUS approach. *British Journal of Nutrition* **2003,** 89, S25-S30.

307. Ritchie, M. R.; Cummings, J. H.; Morton, M. S.; Michael Steel, C.; Bolton-Smith, C.; Riches, A. C., A newly constructed and validated isoflavone database for the assessment of total genistein and daidzein intake. *British Journal of Nutrition* **2006,** 95, 204-213.

308. Bhakta, D.; Higgins, C. D.; Sevak, L.; Mangtani, P.; Adlercreutz, H.; McMichael, A. J.; Silva, I. D., Phyto-oestrogen intake and plasma concentrations in South Asian and native British women resident in England. *British Journal of Nutrition* **2006,** 95, 1150-1158.

309. Adlercreutz, H.; Fotsis, T.; Watanabe, S.; Lampe, J.; Wahala, K.; Makela, T.; Hase, T., Determination of Lignans and Isoflavonoids in Plasma by Isotope-Dilution Gas-Chromatography Mass-Spectrometry. *Cancer Detection and Prevention* **1994,** 18, 259-271.

310. Maziya-Dixon, B.; Akinyele, I. O.; Oguntona, E. B.; Nokoe, S.; Sanusi, R. A.; Harris, E., In Nigeria Food Consumption and Nutrition Survey 2001-2003 Summary, International Institute of Tropical Agriculture, Ibadan, Nigeria: 2004.

311. Afolabi, W. A. O.; Addo, A. A.; Sonibare, M. A., Activity pattern, energy intake and obesity among Nigerian urban market women. *International Journal of Food Sciences and Nutrition* **2004,** 55, 85-90.

312. Song, T.; Barau, K., *American Journal of Clinical Nutrition* **1998,** 68, 1474S.

313. Rauha, J. P.; Vuorela, H.; Kostiainen, R., Effect of eluent on the ionization efficiency of flavonoids by ion spray, atmospheric pressure chemical ionization, and atmospheric pressure photoionization mass spectrometry. *Journal of Mass Spectrometry* **2001,** 36, 1269-1280.

314. Eldridge, A. C.; Kwolek, W. F., Soybean isoflavones - effect of environment and variety on composition. *Journal of Agricultural and Food Chemistry* **1983,** 31, 394-396.

315. Kudou, S.; Fleury, Y.; Welti, D.; Magnolato, D.; Uchida, T.; Kitamura, K.; Okubo, K., Malonyl isoflavone glycosides in soybean seeds (Glycine-Max merrill). *Agricultural and Biological Chemistry* **1991,** 55, 2227-2233.

316. Liggins, J.; Bluck, L. J. C.; Coward, W. A.; Bingham, S. A., Extraction and quantification of daidzein and genistein in food. *Analytical Biochemistry* **1998,** 264, 1-7.

317. Grace, P. B.; Taylor, J. I.; Low, Y.-L.; Luben, R.; Mulligan, A.; Botting, N. P.; Dowsett, M.; Welch, A. A.; Khaw, K.-T.; Wareham, N. J.; Day, N. E.; Bingham, S. A., Phytoestrogen concentrations in serum and spot urine as biomarkers for dietary phytoestrogen intake and their relation to breast cancer risk in European prospective investigation of cancer and nutrition - Norfolk. *Cancer Epidemiology Biomarkers & Prevention* **2004,** 13, 698-708.

318. Seow, A.; Shi, C. Y.; Franke, A. A.; Hankin, J. H.; Lee, H. P.; Yu, M. C., Isoflavonoid levels in spot urine are associated with the frequency of dietary soy intake in a population-based sample of middle-aged and older Chinese in Singapore. *Cancer Epidemiology Biomarkers & Prevention* **1998,** 7, 135-140.

319. Chen, Z.; Zheng, W.; Custer, L. J., Usual dietary consumption of soy foods and its correlation with the excretion rate of isoflavonoids in overnight urine samples among Chinese women in Shanghai. *Nutrition and Cancer* **1999,** 33, 82-87.

320. Arai, Y.; Uehara, M.; Sato, Y., Comparison of isoflavones among dietary intake, plasma concentration and urinary excretion of accurate estimation of phytoestrogen intake. *Journal of Epidemiology* **2000,** 10, 127-135.

321. Liu, R.; Ye, M.; Guo, H.; Bi, K.; Guo, D., Liquid chromatography/ electrospray ionisation mass spectrometry for the characterisation of twenty-three flavonoids in the extract of Dalbergia odorifera. *Rapid Communications in Mass Spectrometry* **2005,** 19, 1557-1565.

322. Webb, A. L.; McCullough, M. L., Dietary lignans: Potential role in cancer prevention. *Nutrition and Cancer-an International Journal* **2005,** 51, 117-131.

323. Schabath, M. B.; Hernandez, L. M.; Wu, X. F.; Pillow, P. C.; Spitz, M. R., Dietary phytoestrogens and lung cancer risk. *Jama-Journal of the American Medical Association* **2005,** 294, 1493-1504.

324. Messina, M.; Messina, V., The 2nd golden-age of nutrition - phytochemicals and disease prevention. *Food Phytochemicals for Cancer Prevention I* **1994,** 546, 382-387.

325. Hedlund, T. E.; Maroni, P. D.; Ferucci, P. G.; Dayton, R.; Barnes, S.; Jones, K.; Moore, R.; Ogden, L. G.; Wahala, K.; Sackett, H. M.; Gray, K. J., Long-term dietary habits affect soy isoflavone metabolism and accumulation in prostatic fluid in Caucasian men. *Journal of Nutrition* **2005,** 135, 1400-1406.

326. Izumi, T.; Piskula, M. K.; Osawa, S.; Obata, A.; Tobe, K.; Saito, M.; Kataoka, S.; Kubota, Y.; Kikuchi, M., Soy isoflavone aglycones are absorbed faster and in higher amounts than their glucosides in humans. *Journal of Nutrition* **2000,** 130, 1695-1699.

327. Zubik, L.; Meydani, M., Bioavailability of soybean isoflavones from aglycone and glucoside forms in American women. *American Journal of Clinical Nutrition* **2003,** 77, 1459-1465.

328. Richelle, M.; Pridmore-Merten, S.; Bodenstab, S.; Enslen, M.; Offord., E. A., Hydrolysis of isoflavone glycosides to aglycones by B-glycosidase does not alter plasma and urine isoflavone pharmacokinetics in postmenopausal women. *Journal of Nutrition* **2002,** 132, 2587-2592.

329. Kano, M.; Takayanagi, T.; Harada, K.; Sawada, S.; Ishikawa, F., Bioavailability of isoflavones after ingestion of soy beverages in healthy adults. *Journal of Nutrition* **2006,** 136, 2291-2296.

330. Chen, L. J.; Zhao, X.; Fang, L. Y.; Games, D. E., Quantitative determination of acetyl glucoside isofiavones and their metabolites in human urine using combined liquid chromatography-mass spectrometry. *Journal of Chromatography A* **2007,** 1154, 103-110.

331. Holder, C. L.; Churchwell, M. I.; Doerge, D. R., Quantification of soy isoflavones, genistein and daidzein, and conjugates in rat blood using LC/ES-MS. *Journal of Agricultural and Food Chemistry* **1999,** 47, 3764-3770.

332. Lapcik, O.; Klejdus, B.; Kokoska, L.; Davidova, M.; Afandi, K.; Kuban, V.; Hampl, R., Identification of isoflavones in *Acca sellowiana* and two *Psidium* species (Myrtaceae). *Biochemical Systematics and Ecology* **2005,** 33, 983-992.

333. Kuhnle, G. G. C.; Dell'Aquila, C.; Low, Y. L.; Kussmaul, M.; Bingham, S. A., Extraction and quantification of phytoestrogens in foods using automated solid-phase extraction and LC/MS/MS. *Analytical Chemistry* **2007,** 79, 9234-9239.

334. Antignac, J.-P.; Cariou, R.; Le Bizec, B.; Andre, F., New data regarding phytoestrogens content in bovine milk. *Food Chemistry* **2004,** 87, 275-281.

335. Kelly, G. E.; Joannou, G. E.; Reeder, A. Y.; Nelson, C.; Waring, M. A., The variable metabolic response to dietary isoflavones in humans. *Proceedings of Society Experimental Biology & Medicine* **1995,** 208, 40-43.

336. Nicholson, J. K.; Wilson, I. D., Understanding 'global' systems biology: Metabonomics and the continuum of metabolism. *Nature Reviews Drug Discovery* **2003,** 2, 668-676.

337. Nicholson, J. K.; Wilson, I. D., High-resolution proton magnetic-resonance spectroscopy of biological-fluids. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1989,** 21, 449-501.

338. Coen, M.; Holmes, E.; Lindon, J. C.; Nicholson, J. K., NMR-based metabolic profiling and metabonomic approaches to problems in molecular toxicology. *Chemical Research in Toxicology* **2008,** 21, 9-27.

339. Wiback, S. J.; Palsson, B. O., Extreme pathway analysis of human red blood cell metabolism. *Biophysical Journal* **2002,** 83, 808-818.

340. Tikkanen, M. J.; Adlercreutz, H., Dietary soy-derived isoflavone phytoestrogens - Could they have a role in coronary heart disease prevention? *Biochemical Pharmacology* **2000,** 60, 1-5.

341. Claridge, T. D. W., High resolution NMR techniques in organic chemistry In *High resolution NMR techniques in organic chemistry* Pergamon, Ed. Elsevier: 1999; Vol. 19, pp 111-146.