An investigation into the detection of fraud during poultry egg production: a metabonomic approach using liquid chromatography - mass spectrometry

Amy Elizabeth Johnson

Thesis submitted for the degree of Doctor of Philosophy

March 2020

Keele University

Abstract

Food fraud is becoming an increasingly difficult challenge, with the ongoing global expansion of food supply chains. Previously developed methods have been successful in the detection of some instances of food fraud, however they struggle to detect more subtle examples of fraud. Eggs are a food product that are susceptible to fraud, as the differences between authentic and fraudulent eggs are extremely subtle and difficult to detect. The intention of this work was to reveal how a metabonomic approach, using liquid chromatography quadrupole-time-of-flight mass spectrometry followed by robust statistical analysis, can be used to observe how different conditions and factors in the production of poultry eggs affect the metabolite profiles of the yolk and albumen of the eggs. It aimed to show how the observed differences between the eggs can lead to the detection and identification of compounds that have potential as markers capable of detecting instances of fraud.

The first study, and some aspects of the following two studies, conducted during this research involved some preliminary work, and were carried out in order to optimise the experimental designs for the remaining studies carried out throughout this research. The optimum storage temperature and condition of metabolite extracts was determined, and it was concluded that the age and diet of the laying birds should be kept the same both within and between experimental groups in the same study.

The metabolite profile of the egg, particularly the albumen, was found to be affected by the age of the bird, and a compound found in the albumen that showed potential as a marker of hen age was putatively identified. It was observed that different diets of the hen affect the metabolite profile of the egg differently, and that they affect the egg at varying rates. Choline was discovered to be a suitable marker of egg age when eggs are stored at 23 °C, however refrigerated egg storage was found to inhibit its use as a marker. Some putatively identified compounds were determined to have potential use as markers

of liquid yolk storage time at 5 °C. The cage stocking density of birds was found to have an effect on the metabolite profile of egg yolk, which was independent of the effect of cage population size. A compound that showed some potential as a marker of cage stocking density was putatively identified. Several potential lipids were found to have a higher abundance in barn egg yolk compared to cage egg yolk, and one of these potential lipids was putatively identified. Little difference was observed between the albumen metabolite profiles of eggs from different housing systems.

The research presented in this thesis reveals that a metabonomic approach, using liquid chromatography-mass spectrometry, can be successful in uncovering subtle differences between eggs from different backgrounds. It shows how this metabonomic approach, and the observed differences, can be applied to the development of methods for the potential detection of fraud in the production and marketing of poultry eggs.

Contents

List of Figures xvi

List of Tables xxxii

1 Introduction **1**

1.1 Aims 1

1.2 Food Fraud 1

1.3 Metabonomics 4

1.3.1 Metabonomics in Food Fraud 5

1.4 Poultry Eggs 6

1.4.1 Fraud in Poultry Egg Production 6

1.4.2 Structure and Formation 8

1.4.2.1 Yolk 10

1.4.2.2 Albumen 11

1.4.2.3 Shell 12

1.5 Lipids 13

1.6 Summary 15

1.7 References 16

2 Techniques and Instrumentation **21**

2.1 Aims 21

2.2 Metabonomics Workflow 21

2.2.1 Experimental Design 22

2.2.2 Sample Collection 23

2.2.3 Sample Preparation and Metabolite Extraction 23

2.2.4 Chemical Analysis 25

2.2.5 Data Pre-processing and Analysis 28

2.2.6 Compound Identification, Targeted Studies, and Biochemical Interpretation . . . 29
2.3 Chromatography . . . 29
2.3.1 High Performance Liquid Chromatography . . . 30
2.4 Mass Spectrometry . . . 33
2.4.1 Ionisation . . . 33
2.4.1.1 ElectroSpray Ionisation . . . 34
2.4.2 Mass Analysers . . . 35
2.4.2.1 Quadrupole Mass Analysers . . . 35
2.4.2.2 Time-of-Flight Mass Analysers . . . 36
2.4.2.3 Hybrid Mass Analysers . . . 39
2.4.2.3.1 Quadrupole-Time-of-Flight Mass Analysers . . . 39
2.4.2.4 Tandem Mass Spectrometry . . . 40
2.5 Statistical Analysis . . . 42
2.5.1 Data Pre-processing . . . 42
2.5.1.1 *XCMS Online* . . . 42
2.5.2 Multivariate Analysis . . . 43
2.5.2.1 Principal Component Analysis . . . 43
2.5.3 Univariate Analysis . . . 44
2.5.3.1 Normal Distribution . . . 44
2.5.3.2 Homogeneity of Variance . . . 45
2.5.3.3 Student's *t*-test . . . 46
2.5.3.4 ANOVA/Welch test . . . 46
2.5.3.5 Post-hoc tests . . . 47
2.6 Summary . . . 48
2.7 References . . . 48

3 Experimental **55**

3.1 Aims . . . 55

3.2 Experimental Design 55
3.3 Materials 57
3.4 Extraction Method Development 57
3.4.1 Aqueous Metabolite Extraction 58
3.4.2 Organic Metabolite Extraction 58
3.5 Chromatographic Method Development 59
3.5.1 Albumen Organic Extract 60
3.5.2 Albumen Aqueous Extract 63
3.5.3 Yolk Organic Extract 67
3.5.4 Yolk Aqueous Extract 70
3.6 Mass Spectrometry 76
3.7 Quality Control 76
3.8 Data Pre-processing 80
3.9 Data Analysis 80
3.9.1 Feature Identification 83
3.10 Summary 83
3.11 References 84

4 The effect of sample storage temperature and condition on the metabolite stability of egg yolk and albumen organic extracts 87

4.1 Introduction 87
4.2 Aims and Objectives 91
4.3 Experimental 92
4.4 Yolk Organic Extracts: 6 weeks of Storage 95
4.4.1 Results and Discussion 95
4.4.1.1 Quality Control Analysis 95
4.4.1.2 Metabolite Profiling 97
4.4.1.2.1 Comparison of storage temperature 97

4.4.1.2.2 Comparison of temperature in each storage condition . . . 98
4.4.1.2.3 Comparison of storage condition . . . 100
4.4.1.2.4 Comparison of storage condition at each temperature . . . 101
4.4.1.3 Multivariate Statistics . . . 104
4.4.1.4 Univariate Statistics . . . 105
4.4.1.4.1 Top compounds: Multivariate Statistics . . . 108
4.4.1.5 Summary . . . 110
4.5 Yolk Organic Extracts: 12 weeks of Storage . . . 111
4.5.1 Results and Discussion . . . 111
4.5.1.1 Quality Control Analysis . . . 111
4.5.1.2 Metabolite Profiling . . . 113
4.5.1.2.1 Comparison of storage temperature . . . 113
4.5.1.2.2 Comparison of temperature at each storage condition . . . 114
4.5.1.2.3 Comparison of storage condition . . . 117
4.5.1.2.4 Comparison of storage condition at each storage temperature . . . 118
4.5.1.3 Multivariate Statistics . . . 121
4.5.1.4 Univariate Statistics . . . 123
4.5.1.4.1 Top compounds: Multivariate Statistics . . . 127
4.5.1.5 Summary . . . 129
4.6 Yolk Organic Extracts: Comparison of compounds between 6 and 12 weeks of storage . . . 130
4.6.1 Results and Discussion . . . 130
4.6.1.1 Top statistically significant compounds following six weeks of storage . . . 130

4.6.1.2 Top statistically significant compounds following twelve weeks of storage . . . 131
4.7 Albumen Organic Extracts: 44 weeks of storage . . . 133
4.7.1 Results and Discussion . . . 133
4.7.1.1 Quality Control Analysis . . . 133
4.7.1.2 Metabolite Profiling . . . 135
4.7.1.2.1 Comparison of storage temperature . . . 135
4.7.1.2.2 Comparison of temperature at each storage condition . . . 136
4.7.1.2.3 Comparison of storage condition . . . 138
4.7.1.2.4 Comparison of storage condition at each storage temperature . . . 139
4.7.1.3 Multivariate Statistics . . . 141
4.7.1.4 Univariate Statistics . . . 142
4.7.1.4.1 Top compounds: Multivariate Statistics . . . 145
4.7.1.5 Summary . . . 147
4.8 Conclusions . . . 148
4.9 References . . . 149

5 The effect of hen age on the metabolite profiles of egg yolk and albumen 151
5.1 Introduction . . . 151
5.2 Aims and Objectives . . . 156
5.3 Experimental . . . 157
5.3.1 Yolk Organic Extracts . . . 157
5.3.2 Albumen Organic Extracts . . . 158
5.4 Hen Age 21-26 weeks: Yolk Organic Extracts . . . 159
5.4.1 Results and Discussion . . . 159
5.4.1.1 Quality Control Analysis . . . 159
5.4.1.2 Metabolite Profiling . . . 161

5.4.1.3 Multivariate Statistics 162
5.4.1.4 Univariate Statistics 164
5.5 Hen Age 26-30 weeks: Yolk Organic Extracts 173
5.5.1 Results and Discussion 173
5.5.1.1 Quality Control Analysis 173
5.5.1.2 Metabolite Profiling 175
5.5.1.3 Multivariate Statistics 176
5.5.1.4 Univariate Statistics 178
5.6 Hen Age 30-38 weeks: Yolk Organic Extracts 186
5.6.1 Results and Discussion 186
5.6.1.1 Quality Control Analysis 186
5.6.1.2 Metabolite Profiling 188
5.6.1.3 Multivariate Statistics 189
5.6.1.4 Univariate Statistics 190
5.7 Hen Age 38-44 weeks: Yolk Organic Extracts 194
5.7.1 Results and Discussion 194
5.7.1.1 Quality Control Analysis 194
5.7.1.2 Metabolite Profiling 196
5.7.1.3 Multivariate Statistics 197
5.7.1.4 Univariate Statistics 198
5.7.1.4.1 Top compounds: Multivariate Statistics 198
5.8 Hen Age 21-44 weeks (21, 26, 30, 38, 44 weeks): Yolk Organic Extracts . 200
5.8.1 Results and Discussion 200
5.8.1.1 Quality Control Analysis 200
5.8.1.2 Metabolite Profiling 202
5.8.1.3 Multivariate Statistics 203
5.8.1.4 Univariate Statistics 205

5.8.1.5 Statistically significant compounds from smaller hen age ranges over a wider hen age range . . . 213
5.9 Hen Age 21-44 weeks (21, 26, 30, 38, 44 weeks): Albumen Organic Extracts 216
5.9.1 Results and Discussion . . . 216
5.9.1.1 Quality Control Analysis . . . 216
5.9.1.2 Metabolite Profiling . . . 218
5.9.1.3 Multivariate Statistics . . . 219
5.9.1.4 Univariate Statistics . . . 221
5.10 Summary & Conclusions . . . 227
5.10.1 Summary . . . 227
5.10.2 Conclusions . . . 229
5.11 References . . . 231

6 The effect of hen diet on the metabolite profiles of egg yolk and albumen 235

6.1 Introduction . . . 235
6.2 Aims and Objectives . . . 238
6.3 Experimental . . . 239
6.4 2 weeks post diet change: Yolk Organic Extracts . . . 242
6.4.1 Results and Discussion . . . 242
6.4.1.1 Quality Control Analysis . . . 242
6.4.1.2 Metabolite Profiling . . . 244
6.4.1.3 Multivariate Statistics . . . 245
6.4.1.4 Univariate Statistics . . . 246
6.4.1.4.1 Top compounds: Multivariate Statistics . . . 247
6.4.1.4.2 Comparison between 2 weeks post diet change and 1 week before diet change . . . 248
6.4.1.5 Summary . . . 249
6.5 5 weeks post diet change: Yolk Organic Extracts . . . 250
6.5.1 Results and Discussion . . . 250

6.5.1.1 Quality Control Analysis 250
6.5.1.2 Metabolite Profiling 252
6.5.1.3 Multivariate Statistics 253
6.5.1.4 Univariate Statistics 254
6.5.1.4.1 Top compounds: Multivariate Statistics 255
6.5.1.4.2 Comparison between 5 weeks post diet change and 1 week before diet change 256
6.5.1.5 Summary 257
6.6 5 weeks post diet change: Albumen Organic Extracts 258
6.6.1 Results and Discussion 258
6.6.1.1 Quality Control Analysis 258
6.6.1.2 Metabolite Profiling 260
6.6.1.3 Multivariate Statistics 261
6.6.1.4 Univariate Statistics 262
6.6.1.4.1 Top compounds: Multivariate Statistics 263
6.6.1.4.2 Comparison between 5 weeks post diet change and 1 week before diet change 264
6.6.1.5 Summary 265
6.7 Summary & Conclusions 266
6.7.1 Summary 266
6.7.2 Conclusions 267
6.8 References 269

7 Differences in metabolite profile between shell egg yolk and liquid yolk stored for different lengths of time 273

7.1 Introduction 273
7.2 Aims and Objectives 280
7.3 Experimental 281

7.3.1 Non-targeted Egg Ageing Study 281
7.3.1.1 Non-targeted Egg Ageing Study: 23 °C 282
7.3.1.1.1 Compound Identification 282
7.3.1.2 Non-targeted Egg Ageing Study: 5 °C 282
7.3.1.3 Non-targeted Egg Ageing Study: Comparison of eggs stored at 5 °C and 23 °C for five weeks 283
7.3.2 Targeted Egg Ageing Study 285
7.3.2.1 Chemical Analysis 285
7.3.3 Non-targeted Liquid Egg Yolk Ageing Study 287
7.4 Eggs Stored at 23 °C for 5 weeks: Yolk Organic Extracts 288
7.4.1 Results and Discussion 288
7.4.1.1 Quality Control Analysis 288
7.4.1.2 Metabolite Profiling 290
7.4.1.3 Multivariate Statistics 291
7.4.1.4 Univariate Statistics 293
7.4.1.5 Choline Identification 298
7.4.1.6 Choline 299
7.4.2 Choline Targeted Study 302
7.5 Eggs Stored at 5 °C for 5 weeks: Yolk Organic Extracts 306
7.5.1 Results and Discussion 306
7.5.1.1 Quality Control Analysis 306
7.5.1.2 Metabolite Profiling 308
7.5.1.3 Multivariate Statistics 309
7.5.1.4 Univariate Statistics 311
7.5.1.5 Choline in refrigerated eggs: non-targeted 318
7.5.2 Choline Targeted Study 320
7.6 Comparison of Eggs Stored at 23 °C and 5 °C for 5 weeks: Yolk Organic Extracts 324

7.6.1 Results and Discussion 324
7.6.1.1 Quality Control Analysis 324
7.6.1.2 Metabolite Profiling 326
7.6.1.3 Multivariate Statistics 327
7.6.1.4 Univariate Statistics 330
7.7 Liquid Eggs Stored at 5 °C for 96 hours: Yolk Organic Extracts 333
7.7.1 Results and Discussion 333
7.7.1.1 Quality Control Analysis 333
7.7.1.2 Metabolite Profiling 335
7.7.1.3 Multivariate Statistics 336
7.7.1.4 Univariate Statistics 338
7.7.1.4.1 Top compounds: Multivariate Statistics 348
7.8 Summary & Conclusions 349
7.8.1 Summary 349
7.8.2 Conclusions 351
7.9 References 352

8 Differences in metabolite profile between the yolks of eggs from hens kept in cages of different stocking densities 355

8.1 Introduction 355
8.2 Aims and Objectives 357
8.3 Experimental 358
8.3.1 Cage Population Size 358
8.3.2 Stocking Density 359
8.4 Cage Population Size (9 Weeks): Yolk Organic Extracts 362
8.4.1 Results and Discussion 362
8.4.1.1 Quality Control Analysis 362
8.4.1.2 Metabolite Profiling 364
8.4.1.3 Multivariate Statistics 365

8.4.1.4 Univariate Statistics 366
8.4.1.4.1 Comparison of top compounds between eggs collected after 9 weeks and control eggs collected after 3 weeks of birds living in cages of different population size 366
8.5 Stocking Density (2 weeks): Yolk Organic Extracts 368
8.5.1 Results and Discussion 368
8.5.1.1 Quality Control Analysis 368
8.5.1.2 Metabolite Profiling 370
8.5.1.3 Multivariate Statistics 371
8.5.1.4 Univariate Statistics 372
8.5.1.4.1 Top compounds: Multivariate Statistics 375
8.5.1.4.2 Comparison of top compounds in eggs collected after 2 weeks of birds living in cages of different stocking densities, with control eggs collected after 4 days, and eggs laid by birds from cages with different population size 377
8.6 Summary & Conclusions 379
8.6.1 Summary 379
8.6.2 Conclusions 381
8.7 References 382

9 Differences in metabolite profile between the yolk and albumen of eggs laid by hens from different housing systems 383

9.1 Introduction 383
9.2 Aims and Objectives 388
9.3 Experimental 389
9.4 Cage and Barn eggs: Yolk Organic Extracts 392
9.4.1 Results and Discussion 392

9.4.1.1 Quality Control Analysis . . . 392
9.4.1.2 Metabolite Profiling . . . 394
9.4.1.3 Multivariate Statistics . . . 395
9.4.1.4 Univariate Statistics . . . 396
9.5 Cage and Barn eggs: Albumen Organic Extracts . . . 399
9.5.1 Results and Discussion . . . 399
9.5.1.1 Quality Control Analysis . . . 399
9.5.1.2 Metabolite Profiling . . . 401
9.5.1.3 Multivariate Statistics . . . 402
9.5.1.4 Univariate Statistics . . . 403
9.6 Cage, Barn, Free Range and Organic Eggs . . . 404
9.6.1 Yolk Organic Extracts . . . 404
9.6.2 Albumen Organic Extracts . . . 408
9.7 Summary & Conclusions . . . 410
9.7.1 Summary . . . 410
9.7.2 Conclusions . . . 411
9.8 References . . . 413

10 Conclusions and Future work 417
10.1 Conclusions . . . 417
10.2 Future work . . . 422
10.3 References . . . 426

Appendix A for Chapter 4 428

Appendix B for Chapter 5 436

Appendix C for Chapter 6 458

Appendix D for Chapter 7 470

Appendix E for Chapter 8 **494**

Appendix F for Chapter 9 **504**

List of Figures

1.1 Diagram showing "omics"disciplines involved in systems biology 4
1.2 Structure of a bird egg 8
1.3 Diagram showing the formation of egg yolk in the ovarian follicle 9
1.4 Diagram showing lipid classes 13

2.1 Diagram showing the typical workflow of a non-targeted metabonomic study 21
2.2 Diagram showing a simplified HPLC system 30
2.3 Diagram showing the separation of compounds in a HPLC column 32
2.4 Diagram showing ESI modes of action 35
2.5 Diagram showing a representation of a quadrupole mass analyser 36
2.6 Diagram showing a representation of a Time-of-Flight mass analyser . . . 37
2.7 Diagram showing a representation of a Time-of-Flight mass analyser with a reflectron 38
2.8 Diagram showing a representation of a Q-ToF mass spectrometer 39
2.9 Diagram showing the different scan modes for tandem mass spectrometers 41

3.1 TIC for albumen organic extract, analysed using initial 2 hour method, using positive ionisation 60
3.2 TIC for albumen organic extract, analysed using final method, using positive ionisation 61
3.3 TIC for albumen organic extract, analysed using initial 2 hour method, using negative ionisation 62
3.4 TIC for albumen organic extract, analysed using final method, using negative ionisation 62
3.5 TIC for albumen aqueous extract, analysed using initial 2 hour method, using positive ionisation 63

3.6 TIC for albumen aqueous extract, analysed using final method, using positive ionisation 64
3.7 Pressure curves showing increasing pressure in the HPLC system with repeated injections of albumen aqueous sample 65
3.8 TIC for albumen aqueous extract, analysed using initial 2 hour method, using negative ionisation 66
3.9 TIC for albumen aqueous extract, analysed using final method, using negative ionisation 66
3.10 TIC for yolk organic extract, analysed using initial 2 hour method, using positive ionisation 67
3.11 TIC for yolk organic extract, analysed using final method, using positive ionisation 68
3.12 TIC for yolk organic extract, analysed using initial 2 hour method, using negative ionisation 69
3.13 TIC for yolk organic extract, analysed using final method, using negative ionisation 69
3.14 TIC for yolk aqueous extract, analysed using initial 2 hour method, using positive ionisation 70
3.15 a) TIC for a blank sample, analysed using initial 2 hour method, positive ionisation. b) Overlaid TICs for a blank sample and yolk aqueous extract, analysed using initial 2 hour method, positive ionisation 71
3.16 TIC for yolk aqueous extract, analysed using initial HILIC method, using positive ionisation 72
3.17 TIC for yolk aqueous extract, analysed using final HILIC method, using positive ionisation 73
3.18 a) TIC for yolk aqueous extract, analysed using initial 2 hour method, negative ionisation. b) TIC for yolk aqueous extract, analysed using initial HILIC method, negative ionisation 74

3.19 TIC for yolk aqueous extract, analysed using final HILIC method, using negative ionisation 75
3.20 Overlaid TICs for the first ten injections of QC sample for a) albumen organic extract and b) yolk organic extract 77
3.21 a) Total Ion Chromatogram for albumen organic extract analysis showing the six peaks that were chosen for quality control monitoring. b) Total Ion Chromatogram for yolk organic extract analysis showing the six peaks that were chosen for quality control monitoring 79
3.22 Diagram showing the statistical workflow that was implemented in this research 81

4.1 Diagram showing the experimental design for the yolk organic extract storage study 92
4.2 Diagram showing the experimental design for the albumen organic extract storage study 93
4.3 Overlaid average TICs of yolk organic extracts that were stored for 6 weeks at -25 ℃, -46 ℃, and -80 ℃ 97
4.4 Overlaid TICs for yolk organic extracts stored at all three temperatures a) re-dissolved in methanol and b) dried, for 6 weeks 99
4.5 Overlaid average TICs of yolk organic extracts that were stored dried and re-dissolved in methanol for 6 weeks 101
4.6 Overlaid TICs for dried and re-dissolved yolk organic extracts stored at a) -25 ℃ b) -46 ℃ and c) -80 ℃ for 6 weeks 102
4.7 PCA scores plot showing PC3 vs PC4 for yolk organic extracts that were stored dried and re-dissolved at -25 ℃, -46 ℃, and -80 ℃ for 6 weeks, including all compounds with CV%<30% 105
4.8 PCA scores plot showing PC2 vs PC3 for yolk organic extracts that were stored dried and re-dissolved at -25 ℃, -46 ℃, and -80 ℃ for 6 weeks, including the top 44 statistically significant compounds 109

4.9 Overlaid average TICs of yolk organic extracts that were stored for 12 weeks at -25 °C, -46 °C, and -80 °C 113
4.10 Overlaid TICs for yolk organic extracts stored at all three temperatures a) re-dissolved in methanol and b) dried, for 12 weeks 115
4.11 Overlaid average TICs of yolk organic extracts that were stored dried or re-dissolved in methanol for 12 weeks 117
4.12 Overlaid TICs for dried and re-dissolved yolk organic extracts stored at a) -25 °C b) -46 °C and c) -80 °C for 12 weeks 119
4.13 PCA scores plot showing PC2 vs PC4 for yolk organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 12 weeks, including all compounds with CV%<30% 122
4.14 PCA scores plot showing PC2 vs PC3 for yolk organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 12 weeks, including the top 46 statistically significant compounds 127
4.15 Overlaid average TICs of albumen organic extracts that were stored at -25 °C, -46 °C, and -80 °C for 44 weeks 135
4.16 Overlaid TICs for albumen organic extracts stored at all three temperatures a) re-dissolved in methanol and b) dried, for 44 weeks 137
4.17 Overlaid average TICs of albumen organic extracts that were stored dried and re-dissolved for 44 weeks 138
4.18 Overlaid TICs for dried and re-dissolved albumen organic extracts stored at a) -25 °C b) -46 °C and c) -80 °C for 44 weeks 140
4.19 PCA scores plot showing PC3 vs PC4 for albumen organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 44 weeks, including all compounds with CV%<30% 142
4.20 PCA scores plot showing PC3 vs PC4 for albumen organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 44 weeks, including all compounds with CV%<30% 146

5.1 Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 21 weeks and 26 weeks old 161
5.2 PCA scores plot showing PC2 vs PC4 for yolk organic extracts of eggs laid by hens from 21 weeks to 26 weeks old, including all compounds with CV%<30% 162
5.3 Line graph showing the trend in changing abundance of compound *m/z* 331.2847 in yolk across eggs laid by hens from 21-26 weeks old 165
5.4 Line graph showing the trend in changing abundance of compound *m/z* 496.3412 in yolk across eggs laid by hens from 21-26 weeks old 166
5.5 Line graph showing the trend in changing abundance of compound *m/z* 648.6307 in yolk across eggs laid by hens from 21-26 weeks old 167
5.6 Line graph showing the trend in changing abundance of compound *m/z* 703.5761 in yolk across eggs laid by hens from 21-26 weeks old 168
5.7 Line graph showing the trend in changing abundance of compound *m/z* 716.5243 in yolk across eggs laid by hens from 21-26 weeks old 169
5.8 Line graph showing the trend in changing abundance of compound *m/z* 734.5709 in yolk across eggs laid by hens from 21-26 weeks old 170
5.9 Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 26 weeks and 30 weeks old 175
5.10 PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 26 weeks to 30 weeks old, including all compounds with CV%<30% 176
5.11 Line graph showing the trend in changing abundance of compound *m/z* 331.2847 in yolk across eggs laid by hens from 26-30 weeks old 179
5.12 Line graph showing the trend in changing abundance of compound *m/z* 338.3420 in yolk across eggs laid by hens from 26-30 weeks old 180
5.13 Line graph showing the trend in changing abundance of compound *m/z* 454.2932 in yolk across eggs laid by hens from 26-30 weeks old 181

5.14 Line graph showing the trend in changing abundance of compound *m/z* 496.3405 in yolk across eggs laid by hens from 26-30 weeks old 182
5.15 Line graph showing the trend in changing abundance of compound *m/z* 703.5758 in yolk across eggs laid by hens from 26-30 weeks old 183
5.16 Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 30 weeks and 38 weeks old 188
5.17 PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 30 weeks to 38 weeks old, including all compounds with CV%<30% 189
5.18 Line graph showing the trend in changing abundance of compound *m/z* 331.2848 in yolk across eggs laid by hens from 30-38 weeks old 191
5.19 Line graph showing the trend in changing abundance of compound *m/z* 338.3412 in yolk across eggs laid by hens from 30-38 weeks old 192
5.20 Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 38 weeks and 44 weeks old 196
5.21 PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by hens from 38 weeks to 44 weeks old, including all compounds with CV%<30% 197
5.22 PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 38 weeks to 44 weeks old, including all compounds with CV%<30% 199
5.23 Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 21 weeks and 44 weeks old 202
5.24 PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 21 weeks to 44 weeks old, including all compounds with CV%<30% 203
5.25 Bar chart showing the trend in changing abundance of compound *m/z* 331.2851 in yolk across eggs laid by hens from 21-44 weeks old 206

5.26 Bar chart showing the trend in changing abundance of compound *m/z* 454.2937 in yolk across eggs laid by hens from 21-44 weeks old 207
5.27 Bar chart showing the trend in changing abundance of compound *m/z* 524.3717 in yolk across eggs laid by hens from 21-44 weeks old 208
5.28 Bar chart showing the trend in changing abundance of compound *m/z* 703.5762 in yolk across eggs laid by hens from 21-44 weeks old 209
5.29 Overlaid average TICs of albumen organic extracts of eggs that were laid by hens of 21 weeks and 44 weeks old 218
5.30 PCA scores plot showing PC2 vs PC4 for albumen organic extracts of eggs laid by hens from 21 weeks to 44 weeks old, including all compounds with CV%<30% 219
5.31 Bar chart showing the trend in changing abundance of compound *m/z* 310.3104 in albumen across eggs laid by hens from 21-44 weeks old 222
5.32 Bar chart showing the trend in changing abundance of compound *m/z* 331.2840 in albumen across eggs laid by hens from 21-44 weeks old 223
5.33 Bar chart showing the trend in changing abundance of compound *m/z* 338.3418 in albumen across eggs laid by hens from 21-44 weeks old 224

6.1 Diagram showing the experimental design for the change in laying hen diet study 239
6.2 Overlaid average TICs of organic extracts of yolks of eggs laid by hens that were fed four different diets, two weeks post diet change 244
6.3 PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds that were fed four different diets, including all compounds with CV%<30% 245
6.4 PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by birds that were fed four different diets, including the top 43 statistically significant compounds 247

6.5 Overlaid average TICs of organic extracts of the yolk of eggs laid by hens that were fed four different diets, five weeks post diet change 252
6.6 PCA scores plot showing PC2 vs PC4 for yolk organic extracts of eggs laid by birds that were fed four different diets, including all compounds with CV%<30% 253
6.7 PCA scores plot showing PC2 vs PC3 for organic extracts of the yolks of eggs laid by birds that were fed four different diets, including the top 30 statistically significant compounds 255
6.8 Overlaid average TICs of organic extracts of albumen of eggs laid by hens that were fed four different diets, five weeks post diet change 260
6.9 PCA scores plot showing PC3 vs PC4 for organic extracts of the albumen of eggs laid by birds that were fed four different diets, including all compounds with CV%<30% 261
6.10 PCA scores plot showing PC3 vs PC4 for organic extracts of the albumen of eggs laid by birds that were fed four different diets, including the top 37 statistically significant compounds 263

7.1 Diagram showing the statistical workflow that was implemented in the non-targeted egg ageing study: 23 ℃ 283
7.2 Diagram showing the statistical workflow that was implemented in the non-targeted egg ageing study: comparison of eggs stored at 23 ℃ and 5 ℃ for five weeks 284
7.3 Overlaid average TICs of yolk organic extracts from fresh eggs and eggs that were five weeks old, stored at 23 ℃ 290
7.4 PCA scores plot showing PC3 vs PC4 for eggs up to five weeks old stored at 23 ℃, including all compounds with CV%<30% 291
7.5 PCA scores plot showing PC3 vs PC4 for eggs up to five weeks old stored at 23 ℃, including only compounds with CV%<30% and p<0.01 292

7.6 Line graph showing the trend in changing abundance of compound *m/z* 104.1070 across eggs of six different ages 295
7.7 Line graph showing the trend in changing abundance of compound *m/z* 331.2860 across eggs of six different ages 297
7.8 Overlaid EICs of compounds with *m/z* 104.1000 for choline chloride standard and QC sample 298
7.9 Comparison of mass spectra at 10 V, 20 V, 40 V collision energy, between a QC sample and a choline chloride standard, for compound *m/z* 104.1000 299
7.10 Example catabolic pathway of the hydrolysis of phosphatidylcholine to choline (circled) 301
7.11 Standard and drift calibration curves for choline chloride standards ranging from 0.2-2.6 µg/mL 302
7.12 Average EICs of choline $[M+H]^+$ at *m/z* 104.1 in yolk extracts for eggs stored for 0, 5, 9, and 12 weeks 304
7.13 Line graph showing the trend in changing abundance of choline in eggs ranging from fresh to twelve weeks old 304
7.14 Overlaid average TICs of yolk organic extracts from fresh eggs and eggs that were five weeks old, stored at 5 ℃ 308
7.15 PCA scores plot showing PC2 vs PC3 for eggs up to five weeks old stored at 5 ℃, including all compounds with CV%<30% 309
7.16 Line graph showing the trend in changing abundance of compound *m/z* 331.2857 across eggs of six different ages 312
7.17 Line graph showing the trend in changing abundance of compound *m/z* 338.3435 across eggs of six different ages 313
7.18 Line graph showing the trend in changing abundance of compound *m/z* 454.2945 across eggs of six different ages 314
7.19 Line graph showing the trend in changing abundance of compound *m/z* 496.3420 across eggs of six different ages 315

7.20 Line graph showing the trend in changing abundance of choline across eggs of six different ages for eggs stored at 5 °C for five weeks 319
7.21 Standard and drift calibration curves for choline chloride standards ranging from 0.2-0.8 µg/mL 321
7.22 Line graph showing the trend in changing yolk choline concentration across eggs of different ages, which were stored at 5 °C for up to five weeks . . . 322
7.23 Line graph showing the trend in changing yolk choline concentration across eggs of different ages, which were stored at 5 °C for up to twelve weeks . 323
7.24 Overlaid average TICs of yolk organic extracts from eggs that were stored at 5 °C and 23 °C for five weeks 326
7.25 PCA scores plot showing PC2 vs PC3 for eggs stored at 5 °C and 23 °C for five weeks, including all compounds with CV%<30% 328
7.26 PCA scores plot showing PC2 vs PC4 for eggs stored at 5 °C and 23 °C for five weeks, including all compounds with CV%<30% and p<0.01 . . . 329
7.27 Bar chart showing the difference in abundance of the compound putatively identified as 1-monopalmitin between eggs stored at 23 °C and 5 °C for five weeks 331
7.28 Bar chart showing the difference in abundance of choline between eggs stored at 23 °C and 5 °C for five weeks 332
7.29 Overlaid average TICs of organic extracts from fresh egg yolk and liquid egg yolk that was stored at 5 °C for 96 hours 335
7.30 PCA scores plot showing PC3 vs PC4 for egg yolks stored at 5 °C up to 96 hours, including all compounds with CV%<30% 337
7.31 Bar chart showing the trend in changing abundance of compound m/z 310.3103 across liquid egg yolk stored for up to 96 hours 340
7.32 Bar chart showing the trend in changing abundance of compound m/z 331.2840 across liquid egg yolk stored for up to 96 hours 341

7.33 Bar chart showing the trend in changing abundance of compound *m/z* 338.3416 across liquid egg yolk stored for up to 96 hours 342
7.34 Bar chart showing the trend in changing abundance of compound *m/z* 525.3743 across liquid egg yolk stored for up to 96 hours 343
7.35 PCA scores plot showing PC2 vs PC3 for egg yolks stored at 5 °C up to 96 hours, including the top 50 compounds 348

8.1 Diagram showing the experimental design for the cage population size study 358
8.2 Diagram showing the experimental design for the stocking density study 360
8.3 Overlaid average TICs of organic extracts of yolks of eggs laid by hens that were kept in cages of different population size 364
8.4 PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds that were kept in cages of four different population sizes, including all compounds with CV%<30% 365
8.5 Overlaid average TICs of organic extracts of yolks of eggs laid by hens that were kept in cages of different stocking densities 370
8.6 PCA scores plot showing PC2 vs PC4 for yolk organic extracts of eggs laid by birds that were kept in cages of four different stocking densities, including all compounds with CV%<30% 371
8.7 Bar chart showing the trend in abundance of compound *m/z* 565.4040 between the yolks of eggs laid by hens kept in cages with different stocking densities of 2, 4, 6, and 8 birds per cage 374
8.8 PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds that were kept in cages of four different stocking densities, including the top 59 statistically significant compounds 376
8.9 Bar chart showing the difference in abundance of compound *m/z* 565.4040 between the yolks of eggs laid by hens kept in cages with different stocking densities of 2, 4, 6, and 8 birds per cage, after just four days of living in these cages 378

9.1 Diagram showing the data analysis workflow that was implemented in the study comparing barn and cage eggs 390
9.2 Overlaid average TICs of organic extracts of yolks of eggs laid by hens from cage and barn housing systems 394
9.3 PCA scores plot showing PC2 vs PC3 for yolk organic extracts of cage and barn eggs, including all compounds with CV%<30% 395
9.4 Bar chart showing the difference in abundance of compound *m/z* 734.5699 between the yolks of cage and barn eggs 397
9.5 Overlaid average TICs of organic extracts of albumen of eggs laid by hens from cage and barn housing systems 401
9.6 PCA scores plot showing PC3 vs PC4 for albumen organic extracts of cage and barn eggs, including all compounds with CV%<30% 402
9.7 Overlaid average TICs of organic extracts of yolks of eggs laid by hens from four different housing systems 404
9.8 PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds from four different housing systems, including all compounds with CV%<30% 405
9.9 Overlaid average TICs of organic extracts of albumen of eggs laid by hens from four different housing systems 408
9.10 PCA scores plot showing PC3 vs PC4 for albumen organic extracts of eggs laid by birds from four different housing systems, including all compounds with CV%<30% 409

B.1 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds 1-monopalmitin and *m/z* 331.2847 438
B.2 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PC(16:0/0:0) and *m/z* 496.3412 439

B.3 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds nervonic ceramide and *m/z* 648.6307 440

B.4 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds SM(18:1/16:0) and *m/z* 703.5761 441

B.5 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PE(16:0/18:2) and *m/z* 716.5243 442

B.6 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PC(16:0/16:0) and *m/z* 734.5709 443

B.7 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds docosenamide and *m/z* 338.3420 446

B.8 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PE(16:0/0:0) and *m/z* 454.2932 447

B.9 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PAF C-16 and *m/z* 524.3717 454

B.10 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds oleoylethylamide and *m/z* 310.3104 456

D.1 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds choline and *m/z* 104.1070 472

D.2 Comparison between the mass spectrum provided by METLIN, and the mass spectrum resulting from analysis for the compounds 1-monopalmitin and *m/z* 331.2860 473

D.3 474

D.4 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds docosenamide and *m/z* 338.3435 482

D.5 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PE(16:0/0:0) and *m/z* 454.2945 483

D.6 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PC(16:0/0:0) and *m/z* 496.3420 484

D.7 Comparison between the mass spectrum provided by METLIN, and the mass spectrum resulting from analysis for the compounds oleoyl ethyl amide and *m/z* 310.3103 492

D.8 Comparison between the mass spectrum provided by METLIN, and the mass spectrum resulting from analysis for the compounds PAF C-16 and *m/z* 525.3743 493

E.1 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds canthaxanthin and *m/z* 565.4040 500

F.1 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PC(16:0/16:0) and *m/z* 734.5699 506

F.2 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds PE(16:0/18:2) and m/z 716.5227 507
F.3 Comparison between the mass spectrum provided by METLIN and the mass spectrum resulting from analysis for the compounds SM(18:1/16:0) and m/z 725.5556 508

List of Tables

3.1 Table showing the solvent gradient of the final method for the analysis of albumen organic extract 60

3.2 Table showing the solvent gradient of the final method for the analysis of albumen aqueous extract 63

3.3 Table showing the solvent gradient of the final method for the analysis of yolk organic extract 67

3.4 Table showing the solvent gradient of the method for the analysis of yolk aqueous extract using a HILIC column 72

3.5 Table showing the parameters for the mass spectrometry analysis 76

4.1 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for six weeks 95

4.2 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for six weeks 96

4.3 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for 12 weeks 111

4.4 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for 12 weeks 112

4.5 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts that had been stored for forty-four weeks 133

4.6 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts that had been stored for forty-four weeks 134

5.1 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 21-26 weeks old 159

5.2 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 21-26 weeks old 160

5.3 Table showing the putative identification of compound m/z 331.2847 and the p-values resulting from ANOVA and Tukey tests 165

5.4 Table showing the putative identification of compound m/z 496.3412 and the p-values resulting from ANOVA and Tukey tests 166

5.5 Table showing the putative identification of compound m/z 648.6307 and the p-values resulting from ANOVA and Tukey tests 167

5.6 Table showing the putative identification of compound m/z 703.5761 and the p-values resulting from ANOVA and Tukey tests 168

5.7 Table showing the putative identification of compound m/z 716.5243 and the p-values resulting from the Welch test and Games-Howell tests 169

5.8 Table showing the putative identification of compound m/z 734.5709 and the p-values resulting from ANOVA and Tukey tests 170

5.9 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 26-30 weeks old 173

5.10 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 26-30 weeks old 174

5.11 Table showing the putative identification of compound m/z 331.2847 and the p-values resulting from ANOVA and Tukey tests 179
5.12 Table showing the putative identification of compound m/z 338.3420 and the p-values resulting from ANOVA and Tukey tests 180
5.13 Table showing the putative identification of compound m/z 454.2932 and the p-values resulting from the Welch test and Games-Howell tests 181
5.14 Table showing the putative identification of compound m/z 454.2932 and the p-values resulting from ANOVA and Tukey tests 182
5.15 Table showing the putative identification of compound m/z 454.2932 and the p-values resulting from ANOVA and Tukey tests 183
5.16 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 30-38 weeks old 186
5.17 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 30-38 weeks old 187
5.18 Table showing the putative identification of compound m/z 331.2848 and the p-values resulting from ANOVA and Tukey tests 191
5.19 Table showing the putative identification of compound m/z 331.2848 and the p-values resulting from ANOVA and Tukey tests 192
5.20 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 38-44 weeks old 194
5.21 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 38-44 weeks old 195

5.22 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 21-44 weeks old 200
5.23 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 21-44 weeks old 201
5.24 Table showing the putative identification of compound *m/z* 331.2851 and the *p*-values resulting from ANOVA and Tukey tests 206
5.25 Table showing the putative identification of compound *m/z* 454.2937 and the *p*-values resulting from ANOVA and Tukey tests 207
5.26 Table showing the putative identification of compound *m/z* 524.3717 and the *p*-values resulting from the Welch test and Games-Howell tests 208
5.27 Table showing the putative identification of compound *m/z* 703.5762 and the *p*-values resulting from ANOVA and Tukey tests 209
5.28 Table showing the *p*-values resulting from ANOVA and Tukey tests for the compound *m/z* 338.3420 213
5.29 Table showing the *p*-values resulting from ANOVA and Tukey tests for the compound *m/z* 734.5708 214
5.30 Table showing the *p*-values resulting from ANOVA and Tukey tests for the compound *m/z* 716.5243 215
5.31 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts from eggs laid by hens 21-44 weeks old 216
5.32 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts from eggs laid by hens 21-44 weeks old 217
5.33 Table showing the putative identification of compound *m/z* 310.3104 and the *p*-values resulting from ANOVA and Tukey tests 222

5.34 Table showing the putative identification of compound m/z 331.2840 and the p-values resulting from ANOVA and Tukey tests 223
5.35 Table showing the putative identification of compound m/z 338.3418 and the p-values resulting from ANOVA and Tukey tests 224

6.1 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens fed four different diets, two weeks post diet change 242
6.2 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens fed four different diets, two weeks post diet change 243
6.3 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens fed four different diets, five weeks post diet change 250
6.4 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens fed four different diets, five weeks post diet change 251
6.5 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens fed four different diets, five weeks post diet change 258
6.6 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens fed four different diets, five weeks post diet change 259

7.1 Table showing the chromatographic method for the targeted analyses in the egg ageing study 286

7.2 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 23 ℃ 288

7.3 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 23 ℃ 289

7.4 Table showing the putative identification of compound *m/z* 104.1070 and the *p*-values resulting from ANOVA and Tukey tests 294

7.5 Table showing the putative identification of compound *m/z* 331.2860 and the *p*-values resulting from ANOVA and Tukey tests 296

7.6 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic samples of eggs stored at 5 ℃ 306

7.7 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 5 ℃ 307

7.8 Table showing the putative identification of compound *m/z* 331.2857 and the *p*-values resulting from the Welch test and Games-Howell tests 312

7.9 Table showing the putative identification of compound *m/z* 338.3435 and the *p*-values resulting from the Welch test and Games-Howell tests 313

7.10 Table showing the putative identification of compound *m/z* 454.2945 and the *p*-values resulting from the ANOVA and Tukey tests 314

7.11 Table showing the putative identification of compound *m/z* 496.3420 and the *p*-values resulting from the ANOVA and Tukey tests 315

7.12 Table showing the *p*-values resulting from the ANOVA and Tukey tests for choline in the yolks of eggs stored at 5 ℃ for five weeks. 319

7.13 Table showing peak areas, SDs, means, and CV%s for six peaks in five QC samples that were analysed throughout the analytical run for yolk organic samples of eggs stored at 5 ℃ and 23 ℃ for five weeks 324
7.14 Table showing RTs, SDs, means, and CV%s for six peaks in five QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 5 ℃ and 23 ℃ for five weeks 325
7.15 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for organic extracts of liquid egg yolk stored at 5 ℃ 333
7.16 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for organic extracts of liquid egg yolk stored at 5 ℃ 334
7.17 Table showing the putative identification of compound *m/z* 310.3103 and the *p*-values resulting from the Welch test and Games-Howell tests 340
7.18 Table showing the putative identification of compound *m/z* 331.2840 and the *p*-values resulting from ANOVA and Tukey tests 341
7.19 Table showing the putative identification of compound *m/z* 338.3416 and the *p*-values resulting from ANOVA and Tukey tests 342
7.20 Table showing the putative identification of compound *m/z* 525.3743 and the *p*-values resulting from ANOVA and Tukey tests 343

8.1 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different population size . . . 362
8.2 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different population size . . . 363

8.3 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different stocking density . . 368
8.4 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different stocking density . . 369
8.5 Table showing the putative identification of compound *m/z* 565.4040 and the *p*-values resulting from ANOVA and Tukey tests 373

9.1 Table showing the main requirements for the four different housing systems 384
9.2 Table showing the age and breed of birds from the four different housing systems 389
9.3 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens from cage and barn housing systems 392
9.4 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens from cage and barn housing systems 393
9.5 Table showing the putative identification of compound *m/z* 734.5699 and the *t*-test *p*-value 398
9.6 Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens from cage and barn housing systems . . 399
9.7 Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens from cage and barn housing systems . . 400
9.8 Table showing the putative identification of compound *m/z* 734.5699 and the *p*-values resulting from ANOVA and Tukey tests 406

9.9 Table showing the putative identification of compound m/z 716.5227 and the p-values resulting from ANOVA and Tukey tests 407

9.10 Table showing the putative identification of compound m/z 725.5556 and the p-values resulting from Welch test and Games-Howell tests 407

10.1 Summary table showing putatively identified potential markers in each chapter 421

A.1 Table showing the significantly different features between yolk organic extracts that were stored either dried or re-dissolved at -25 °C, -46 °C, and -80 °C for six weeks 428

A.2 Table showing the significantly different features between yolk organic extracts that were stored either dried or re-dissolved at -25 °C, -46 °C, and -80 °C for twelve weeks 430

A.3 Table showing the significantly different features between yolk organic extracts that were stored under different conditions and at different temperatures for six weeks, as well as twelve weeks of storage 432

A.4 Table showing the significantly different features between yolk organic extracts that were stored under different conditions and at different temperatures for twelve weeks, as well as six weeks of storage 433

A.5 Table showing the significantly different features between albumen organic extracts that were stored either dried or re-dissolved at -25 °C, -46 °C, and -80 °C for forty-four weeks 434

B.1 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (21-26 weeks old) . 436

B.2 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (26-30 weeks old) . 444

B.3 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (30-38 weeks old) . 448

B.4 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (38-44 weeks old) . 450

B.5 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (21-44 weeks old) . 452

B.6 Table showing the significantly different features in albumen organic extracts between eggs that were laid by hens of different ages (21-44 weeks old) 455

C.1 Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed four different diets, two weeks post diet change 458

C.2 Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed the four different diets two weeks post diet change, as well as one week prior to diet change 461

C.3 Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed four different diets, five weeks post diet change 462

C.4 Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed the four different diets five weeks post diet change, as well as one week prior to diet change 465

C.5 Table showing the significantly different features in albumen organic extracts between eggs laid by birds that were fed four different diets, five weeks post diet change 466

C.6 Table showing the significantly different features in albumen organic extracts between eggs laid by birds that were fed the four different diets five weeks post diet change, as well as one week prior to diet change 469

D.1 Table showing the significantly different features in yolk organic extracts between eggs of different ages stored at 23 ℃ 470

D.2 Table showing concentration of choline in yolk organic extracts, the yolks themselves, and the mean choline concentration in egg yolks from eggs of different ages, stored at 23 ℃ 475
D.3 Table showing p-values resulting from Tukey tests for yolk choline concentration between all pairwise egg ages 478
D.4 Table showing the significantly different features in yolk organic extracts between eggs of different ages stored at 5 ℃ 480
D.5 Table showing concentration of choline in yolk organic extracts, the yolks themselves, and the mean choline concentration in egg yolks from eggs of different ages, stored at 5 ℃ 485
D.6 Table showing the significantly different features in yolk organic extracts between eggs stored at 5 ℃ and 23 ℃ for five weeks 488
D.7 Table showing the significantly different features in organic extracts between egg yolks stored at 5 ℃ for different lengths of time 490

E.1 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages with different numbers of birds 494
E.2 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages of different population size for nine weeks, as well as three weeks 496
E.3 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages with different stocking densities 498
E.4 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages of different stocking densities for two weeks, as well as four days 502

F.1 Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens from cage and barn housing systems 504

F.2 Table showing the 12 compounds that produced potential metabolite matches through METLIN 505

F.3 Table showing the significantly different features in albumen organic extracts between eggs that were laid by hens from cage and barn housing systems 507

Acknowledgements

I would like to take this opportunity to thank my supervisor, Dr David Thompson, for his help, support, and friendship over the past few years. Reaching the end of my PhD with the skills and confidence I have gained is down to you, and for that I am extremely grateful.

I would also like to thank Dr Vladimir Zholobenko, Dr Vasil Pirgozliev, and Professor Tony Edge for their continued help and support throughout my PhD.

Thanks to Steven Wye for his technical support, and his excellent provision of snacks over the past few years. You are a star.

A special thanks to Kate Sidwick and Sue Shemilt for making the past few years so enjoyable, and for supporting me through both personal and PhD-related life. Thanks to Kate for teaching me the ways of the Q-ToF, endlessly discussing my experiments with me, and for our morning chats in an empty office. And thanks to Sue for always keeping me calm, and for being the voice of reason I've needed to hear on several occasions (and undoubtedly will in the future). You are both amazing.

A huge thank you to my wonderful partner, Kieran Brayford, for his unwavering support, and for always believing in me, even when I didn't. I could not have done this without you.

Thank you especially to my mum Margot Caine, step-dad Ed Caine, and sibling Ash Johnson, for their unending love and support always. You were right mum - turns out I do enjoy research!

And finally, thank you to my dad, Ray Johnson, for instilling in me the importance of education, always asking questions, and never throwing away old notes! You may not have been around to see me complete my degree and start my PhD journey, but I hope you'd be proud of me. This one's for you.

1. Introduction

1.1 Aims

- *To introduce the concepts of food fraud and metabonomics, and how metabonomics can be applied to food authentication studies.*
- *To discuss how poultry eggs are affected by food fraud, and the structure and formation of eggs.*
- *To introduce and discuss different classes of lipids.*

1.2 Food Fraud

Food fraud is the act of intentionally deceiving consumers regarding food products, and most often occurs for purposes of financial gain [1][2][3]. There are three main categories of food fraud: adulteration, substitution, and misrepresentation of foodstuffs, and these can affect a wide variety of products such as meat, dairy, fresh produce, alcohol, and eggs, amongst others [2][3].

Food adulteration is the act of deliberately adjusting a food product in some way; by either the use of extraneous ingredients, or the substitution of a current, high quality ingredient for another, which may be inferior or substandard [1][4]. The European horsemeat scandal of 2013, where beef products were found to contain horsemeat, is a good example of food adulteration [2][4]. As well as being a moral issue, food adulteration

is also a food safety concern, as these extra ingredients in the food products may be allergens, which could be dangerous for allergy sufferers.

Food substitution involves completely replacing one product with another, inferior one, but advertising and selling it as the original product [5]. For example, it is possible for bar staff to refill an empty bottle of an expensive brand of liquor with a less expensive brand and continue selling it to consumers at the higher price, as though it were still the original brand. This enables them to increase their profit through fraudulent practices.

Food misrepresentation can present in several different forms, but is essentially misleading the consumer regarding the origins of the product [1][6]. Some food products may be advertised as originating from certain countries, making them particularly desirable, when they were actually grown or produced elsewhere [6]. It may be less expensive to import these products from different countries, tempting companies to mislead consumers about the origins of the products. As organic produce can be sold at a higher price than conventionally grown/farmed produce, due to the lower yields and costs associated with organic certification [7], food misrepresentation also occurs with the mislabelling of conventional products as organic [6][7]. In fact, this type of food fraud has increased in occurrence over recent years, due to the increasing demand of consumers for organic produce [5]. The supply of organic produce from the agricultural industry is struggling to keep up with the demand from consumers and so companies are importing products from less reliable sources, which may mislead these companies about the agricultural origins of their products [5].

Food fraud is an age-old problem that can be traced back thousands of years to the ancient Roman and Greek civilisations, which had laws regarding the adulteration of wine [5]. Evidence has been found of fraudulent wine and olive oil dating from the ancient Roman times [1], and fraudulent practices have continued through the Middle Ages and into modern times [5]. Although food fraud is not a new issue, due to the modern globalisation of food supply chains, the problem of food fraud has expanded in recent

years, increasing the vastness of its impact [1][3][8].

As consumer awareness of food fraud is increasing [1][9], particularly following the horsemeat scandal throughout Europe in 2013 [3][10], the trust that consumers have in Food Business Operators (FBOs), and their produce, is decreasing [2][3]. This can result in a lack of sales, causing negative economic implications not only for FBOs, but also for the government due to a loss of value added tax (VAT) from sales [2]. Economic issues can also arise from product recalls and authenticity testing of products following the discovery of fraudulence [3]. It has been estimated that globally, food fraud costs approximately $US49billion a year [11].

The European Union has various laws and regulations in place in order to control the production and marketing of food products, and FBOs are responsible for ensuring that these are followed within all areas of the food business that they govern [12]. The EC Regulation 178/2002 lays down the principles and requirements of food law, and states that the aim of food law is to prevent food adulteration, misleading or deceptive practices, and any type of fraudulence with respect to food products [12].

In the United Kingdom, the Food Safety Act of 1990 covers the legislation regarding food fraud. In this act there are two sections dedicated to consumer protection. Section 14 states that it is an offence to sell any food products that are not of the nature, or quality expected by the consumer, and section 15 makes it an offence to falsely describe any food products that are for sale for human consumption, whether it be mislabelling the products or false advertisement [13]. In the Unites States (US) the Federal Food, Drug and Cosmetic Act prohibits the adulteration and misbranding of food products [14].

As the issue of food fraud is becoming increasingly problematic, and there are laws and regulations aimed at preventing food fraud, it is important to have a way of detecting instances where it has occurred. Having a method that is capable of uncovering cases of food fraud, that does not simply rely on inspections and identifying inconsistencies

in paperwork, would help to prevent instances of food fraud from occurring. It would potentially result in the detection of fraud prior to selling the products to the consumer, and it would also act as a deterrent due to the increased risk of being caught. Metabonomics is an increasingly popular technique that is being employed in food fraud studies, to develop methods of detection of food fraud affecting various different food products [15].

1.3 Metabonomics

Metabonomics is the comprehensive profiling of low molecular weight species (<1000Da) [16] in organic tissues and biofluids, and the observation of how these profiles are affected by both intrinsic and extrinsic factors [17][18]. Metabonomics is often confused with, or referred to interchangeably with, metabolomics. However, these two disciplines are slightly different; **metabolomics** is the study of the complete metabolome, i.e. all of the metabolites present in a biological system, whereas **metabonomics** is the study of how the metabolite profile *changes* in response to internal and external stimuli [19][20].

Systems biology involves the integration of a range of different "omics"disciplines, all aimed at learning and understanding more about biological systems at increasingly lower levels [16][21]. *Geno*mics is the study of genes and DNA, *transcripto*mics is the study of RNA, *proteo*mics is the study of proteins, and the most recent "omics"technology, *metabono*mics, is the study of the metabolites that make up the metabolome [16][18].

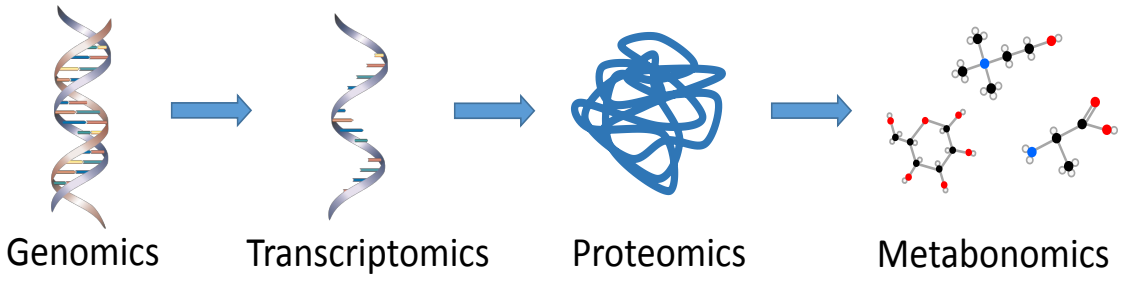


Figure 1.1: *Diagram showing "omics"disciplines involved in systems biology*

Examples of some biomolecules that make up the metabolome are: amino acids, carbohydrates, lipids, and vitamins, as well as various other metabolites that are involved in biosynthesis and biodegradation pathways [21].

There are two broad categories of metabonomic studies; targeted and non-targeted [19][22]. Targeted studies can be useful when particular biological pathways are known to be affected by the factor that is being investigated, as the experimental methods can be tailored to particular compounds or compound classes that are found in that pathway [19][22]. Non-targeted metabonomic studies are preferred when it is unknown which biological pathways may be affected, as they aim to profile as many metabolites, or small biomolecules, as possible, in order to get a more holistic view of the metabolome [19][22]. Non-targeted studies generate large amounts of data, providing information on thousands of compounds found in the samples. From this data it is possible, through the use of statistical workflows, to discover compounds that show a significant trend between the different sample sets in the study [22]. If just one of these compounds is identified, following further work and larger, targeted metabonomic studies, it could become a known biomarker relating to the factor under investigation.

1.3.1 Metabonomics in Food Fraud

Metabonomics is a relatively recent discipline that is becoming increasingly popular, particularly within areas of research such as: disease diagnostics, toxicology, environmental research, as well as food authentication studies [15][18][23]. It is particularly well suited to authentication studies which investigate the detection of food fraud, as it is capable of exploring the more subtle differences in food products in order to detect instances of fraudulent activity. Most cases of food fraud are undetectable by the human senses; the colour, smell, and taste of the fraudulent products are almost indistinguishable from true authentic products [24]. Therefore, various scientific methods have been developed that are capable of detecting the different types of fraud affecting various food products

[15][25].

However, some cases of food fraud can be particularly difficult to detect, for example the adulteration of olive oil with hazelnut oil [15], as the chemical compositions of the fraudulent and authentic products are very similar. Other cases of food fraud can be even more subtle than this; food misrepresentation can be particularly difficult to identify as the chemical and physical characteristics of the fraudulent products are the same as the authentic products, however the origins of the products may not be the same as what is advertised on the label [1][6]. There is no way for the consumer to identify whether any misrepresentation has taken place, and as the differences between true, authentic products and the misrepresented products are so subtle, it is difficult to detect inauthenticity, even through the use of scientific methods. However, as non-targeted metabonomic studies aim to gain information about as many different compounds and compound classes in the samples as possible, they are likely to be capable of uncovering some compound differences between authentic and fraudulent samples.

1.4 Poultry Eggs

1.4.1 Fraud in Poultry Egg Production

Eggs are an example of food products that are susceptible to fraudulent misrepresentation [2]. In England and Wales, it is the responsibility of Egg Marketing Inspectors (EMIs) from the Animal and Plant Health Agency (APHA) to carry out inspections and enforce legislation outlined in the APHA Guidance on Legislation Covering the Marketing of Eggs [26].

There are four main farming methods of egg production: cage, barn, free range, and organic [27]. The average U.K. farm-gate price of eggs is approximately 30p greater for free range eggs than it is for cage eggs [28], and by comparing the supermarket prices of eggs it can be observed that cage eggs are the cheapest to buy, whilst organic eggs

fetch the highest price. Therefore, there is the potential for eggs to be mislabelled with a false farming method, in order to sell them at a higher price and make an increased profit. Laying hens that are used in organic egg production must be fed an organic diet, which consists of different ingredients to the conventional diets that are fed to other, non-organic birds [29]. However, this organic feed is more expensive, creating the temptation for farmers to feed the organic birds a conventional diet, in order to avoid the higher costs of organic feed.

Eggs must be labelled with a producer code, to enable the eggs to be traced back to their production site, and even to a specific flock of birds at that site [30][31]. However, this allows cases of misrepresentation, such as labelling the eggs with incorrect farming methods, or falsely labelling them as "locally sourced", to be easily discovered. Therefore, producers may be tempted to mislabel the eggs with false producer codes, to avoid detection.

In 2012 the legislation regarding the cage size for laying hens changed, and stated that all cages must be "enriched"cages, with a minimum of 750 cm^2 of cage area per hen, of which 600 cm^2 must be usable [30]. This created the potential for fraud, as egg producers now require more space to house the same number of birds, and therefore may be tempted to continue housing laying hens in the old battery cages.

Eggs have a sell-by date of three weeks post-lay, and a best before date of four weeks post-lay [32]. The date of minimum durability, or best before date, must be clearly labelled on the packaging [32]. However, as there is no way for the consumer to confirm that these dates are correct, there is again potential for the eggs to be mislabelled with false sell-by and best before dates, in order to give them a longer shelf life and sell them before they reach their falsely labelled sell-by date.

1.4.2 Structure and Formation

The type of egg laid by hens, and in fact all birds, is known as a cleidoic egg, meaning that the interior environment of the egg is almost completely separate to the exterior [33]. The egg is the largest single cell found in animals and contains all of the nutrients essential for the development of an embryo outside of its mother's body [33][34][35], making it extremely compound-rich. Bird eggs consist of various components: the yolk (including the germinal disc), vitelline membrane, thick inner albumen, thin outer albumen, chalazae, shell membranes and the shell [33], see Figure 1.2. The structure of the egg is similar for all bird species, with the main differences being the proportions of the different parts. The proportions of the parts are affected by the age of the bird, the diet of the bird, and the bird's environment [33].

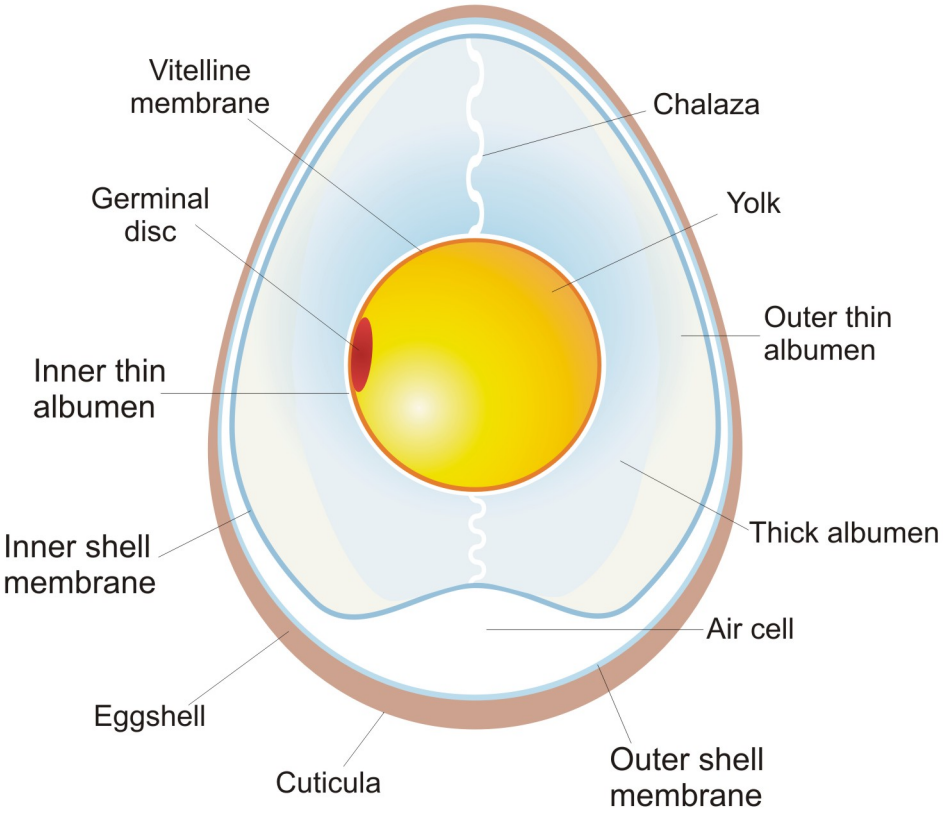


Figure 1.2: *Structure of a bird egg. Adapted from Nys and Guyot [33].*

The formation of an egg is controlled by steroid and pituitary hormones produced and secreted by the ovary and the pituitary gland [33]. The components of the egg yolk are produced by the liver of the bird and are then transported to the ovary via the bloodstream, where they are incorporated into the largest ovarian follicle in order to form the yolk, which is then released from the follicle in the oviduct during ovulation [33]. Once the yolk precursors reach the ovary, they then have to penetrate several layers of the follicle wall before they can be incorporated into the yolk during formation [36], see Figure 1.3.

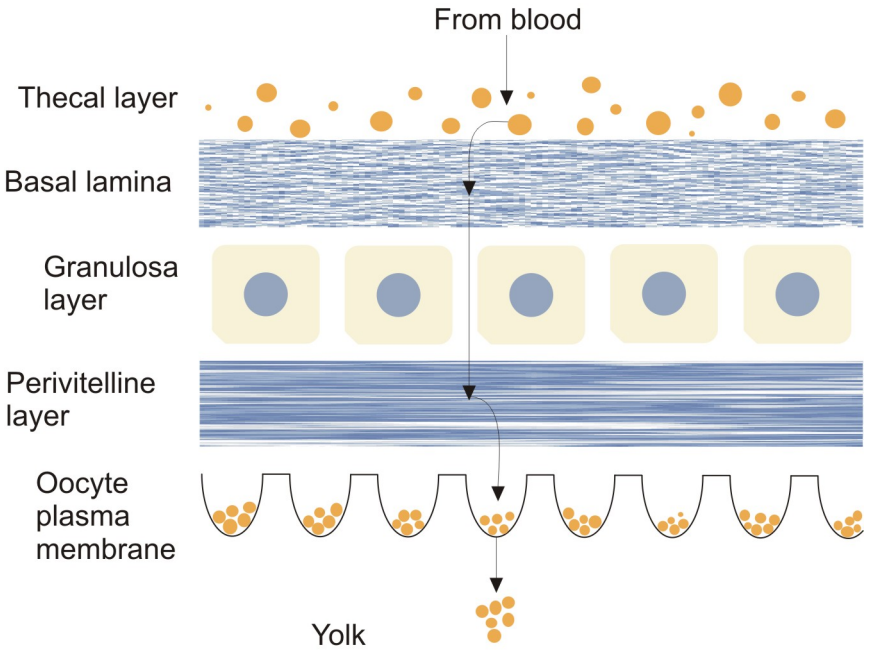


Figure 1.3: *Diagram showing the formation of egg yolk in the ovarian follicle. Adapted from Griffin [36].*

The outermost layer of the follicle, the thecal layer, is highly vascularised with highly permeable capillaries, allowing the plasma to easily leak through [36]. The yolk precursors then have to penetrate the basal lamina, which is made up of connective tissue and filters out any large particles in the plasma that are trying to pass through. They then penetrate the granulosa cell layer and the perivitelline layer before binding to receptors on the oocyte plasma membrane [36]. When enough receptors have been activated by the binding of these precursors, endocytic vesicles are formed, allowing incorporation

of these components into the yolk [36]. Following ovulation, the oviduct then produces constituents of the other egg components; the vitelline membrane, albumen, shell membranes, and shell which are deposited around the yolk to form the egg [33].

An egg consists of approximately 50-60% albumen, 30-35% yolk and 10-15% shell [33][34], with the edible portion of the egg (yolk and albumen) consisting of 74.4% water, 12.3% protein and 11.6% lipids. Eggs also contain all vitamins except vitamin C, as well as a wide range of minerals and trace elements [33].

1.4.2.1 Yolk

The yolk is found in the centre of the egg, surrounded by the vitelline membrane. This is a very thin membrane that provides a barrier between the yolk and albumen in order to prevent any exchange of material between these two egg components, and also acts as a barrier against bacterial penetration [33]. On the surface of the yolk is a small clear disc known as the blastodisc, or the germinal disc, which contains the female chromosomes [33]. The yolk is held in the centre of the egg by two chalazae, which are spiral filaments joining the opposite sides of the yolk to either end of the egg, in order to keep the germinal disc in a stable position [33][37].

The yolk is made up of approximately 48% water, 32-35% lipids and 15-16% proteins, with a very low percentage of carbohydrates [35]. All lipids that are present in the egg are found in the yolk, with neutral lipids making up 65% of the total lipid content, phospholipids 31%, and cholesterol 4% [33][35][37]. Most of the lipid content is made up of fatty acids, with the major fatty acids consisting of both saturated and unsaturated fatty acids: oleic acid (43.6%), palmitic acid (25.1%), linoleic acid (13.4%), stearic acid (8.6%), palmitoleic acid (3.6%), docosahexaenoic acid (1.8%) and arachidonic acid (1.7%) [35]. There are four main groups of phospholipids found in the yolk: phosphatidylcholine (PC), phosphatidylethanolamine (PE), sphingomyelin and lysophosphatidylcholine (LPC), making up approximately 84%, 12%, 2% and 2% of the total phospholipid content respectively,

in addition to other minor components [35].

Proteins that are present in the yolk are mostly bound to lipids, to form low and high-density lipoproteins [33]. Of the small percentage of carbohydrates found in the yolk, most are bound to proteins or lipids to form glycoproteins and glycolipids respectively [33]. Phosphorous is present in large amounts in the yolk, mainly incorporated into phospholipids, and most iron that is present in an egg is located in the yolk. All fat-soluble vitamins that are found in eggs are present in the yolk (due to the high lipid content) as well as many water-soluble vitamins [33]. The colour of an egg yolk is due to the presence of pigments called carotenoids, particularly xanthophylls as well as small amounts of carotenes [33][37]. Carotenoids are not endogenous to hens, and so the total carotenoid content of an egg comes from the diet of the bird and is dependent on the absorption of carotenoids by the hen and the subsequent transfer into eggs [37].

1.4.2.2 Albumen

The primary role of the albumen is to surround and protect the yolk within the shell, both from physical and bacterial trauma [37]. The albumen consists of two parts; thin albumen and thick albumen. There is a small amount of thin albumen surrounding the yolk, which itself is surrounded by the thick albumen, which reaches to either end of the egg [33]. Surrounding this is the remaining thin albumen. The proportions of thin to thick albumen change over an increasing egg storage time, and with an increased weight of the egg [33].

The egg albumen consists mainly of water (84-89%, although this varies between thin and thick albumen) [37], in addition to protein, which makes up around 90% of the dry matter of the albumen, as well as minerals and vitamins [33]. Proteins that are present in the albumen exist in the free form, as well as bound to lipids and carbohydrates to form lipoproteins and glycoproteins [34]. It is the high content of a glycoprotein called ovomucin in the thick albumen, which gives it a higher viscosity than the thin albumen

[37]. Carbohydrates also make up a very small percentage of albumen, with half existing in free form, mainly glucose [33]. All minerals that are essential for embryonic growth and development are also found in the albumen [33], but it contains no fat-soluble vitamins [37] and only water-soluble vitamins of the B group [33]. The albumen also contains an antibacterial enzyme, lysozyme, which helps protect the yolk from any bacteria that may penetrate through the shell [35].

1.4.2.3 Shell

The shell is the part of the egg that is formed last during egg development, and consists of various layers [33]. Shell membranes (an inner and an outer membrane) made up of proteins and glycoproteins are the first layers of the shell, and they surround the albumen, holding it in place, and determine the shape of the egg before the shell itself is formed [33]. The actual shell consists of solely inorganic minerals, including a large amount of calcium carbonate [33][37][38], with a very small percentage of water, and low levels of minerals and trace elements [33]. It is made up of three layers: the mammillary layer, the palisade layer (which is the thickest), and the vertical crystal layer [33][37]. On the surface of the shell is the cuticle, or cuticula, which is an organic layer in contrast to the mineral layers of the actual shell, consisting mostly of proteins (mainly glycoproteins) as well as small amounts of polysaccharides and lipids [38]. There are thousands of pores penetrating through the eggshell, which allow for the transfer of water and various gases between the internal and external environments of the egg [37][38]. The cuticle controls this exchange of substances by plugging the entry to the pores, and thus acts as a barrier against bacterial penetration [37][38]. Once an egg has been laid, the liquid components begin to lose water by evaporation through the pores in the eggshell, resulting in the formation of an air cell [39].

As lipids make up such a large component of egg yolk, it is important to have an understanding of the basic biochemistry behind this class of compounds.

1.5 Lipids

Lipids are a class of compounds that share chemical and physical properties due to their structural similarities [40]. There are two main groups of lipids; those that consist of fatty acid components, with polar head groups and long non-polar tails, and a group of lipids known as isoprenoids, such as steroids, lipid vitamins, and terpenes, which all have structures related to the molecule isoprene [40][41]. Figure 1.4 shows the structural relationships of the main classes of lipids.

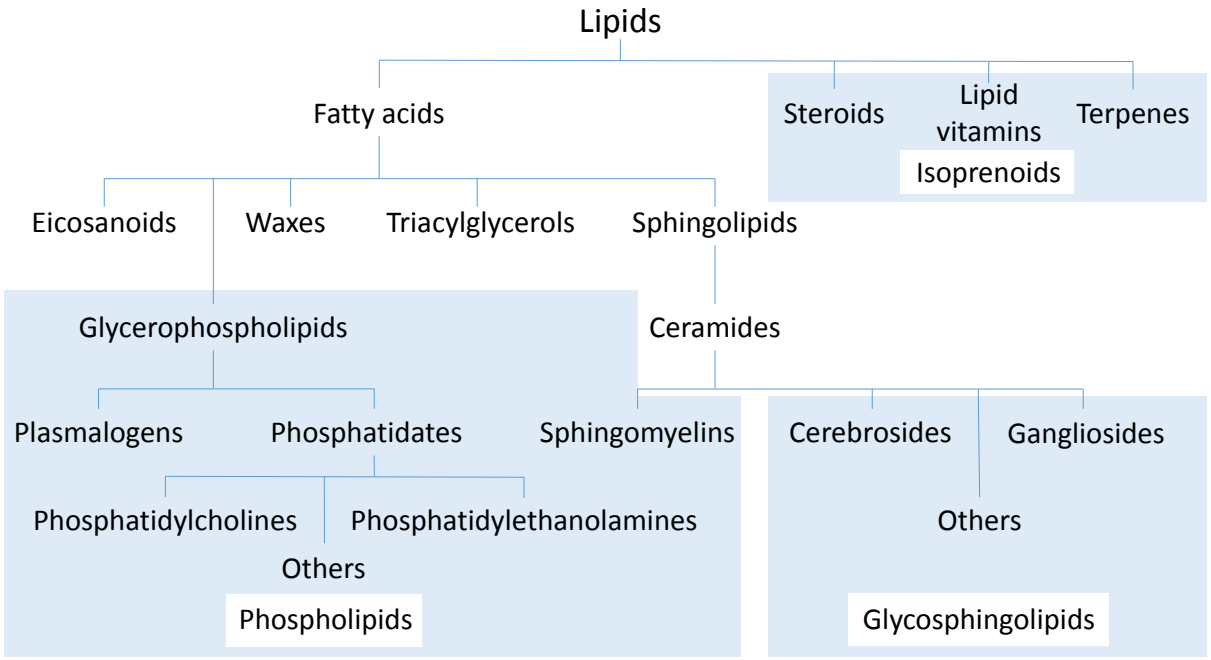


Figure 1.4: *Diagram showing lipid classes. Adapted from Figure 1, Chapter 9, in Principles of Biochemistry [41]*

The simplest lipids are fatty acids, which are amphipathic molecules, consisting of a hydrophilic carboxyl group and a hydrophobic hydrocarbon chain [40]. Saturated fatty acids have a long hydrocarbon chain with only single bonds, and unsaturated fatty acids have at least one (monounsaturated) or more (polyunsaturated) double bonds throughout the hydrocarbon chain [40][41]. The oxidation of fatty acids is the main source of energy in lipid catabolism [42].

Most fatty acids do not exist in the free form, and are mainly present as components of other, more complex lipids [40][41]. Triacylglycerols, or triglycerides, consist of three fatty acids joined to a glycerol backbone by ester bonds. The main role of triacylglycerols is to act as a store of fatty acids [40][41]. Lipase enzymes catalyse the hydrolysis of the acyl groups, breaking the ester bonds between the fatty acids and the glycerol, in order to release the fatty acids ready for oxidation [40]. Monoacylglycerols have just one fatty acid attached to the glycerol, and diacylglycerols have two fatty acids attached to the glycerol.

Glycerophospholipids have a similar structure to triacylglycerols, but one of the three fatty acids attached to the alcohol groups of the glycerol is replaced with a phosphate molecule through an ester linkage [40]. When only one fatty acid is attached to the glycerol backbone of a phospholipid, the resulting molecule is a lysophospholipid. The most simple glycerophospholipids are phosphatidates, which consist of two fatty acids attached to glycerol-3-phosphate, with no other groups forming ester linkages to the phosphate group [40][41]. These compounds are only present in small amounts, as intermediates in the biosynthesis or degradation of more complex glycerophospholipids [41].

However, in most glycerophospholipids, the phosphate group is also joined via an ester bond to other compounds, such as choline or ethanolamine, in order to produce phosphatidylcholines or phosphatidylethanolamines respectively [40][41]. There are various different glycerophospholipids that exist, due to the many different fatty acids that can join to the glycerol backbone. Phospholipase enzymes catalyse the hydrolysis of the acyl groups, breaking the ester bonds between the fatty acids and the glycerol-3-phosphate molecule, releasing the fatty acids [40][41][43]. There are also a group of phospholipids known as plasmalogens, which are similar to glycerophospholipids, but which have a hydrocarbon chain attached to the first carbon of the glycerol via a vinyl ether bond [41].

Sphingolipids are a group of lipids that are built upon a sphingosine backbone [40][41]. Ceramides, which have a fatty acid linked to the amino group of the sphingosine molecule

by an amide bond [40][41], are the precursors to all sphingolipids and can be split into three main groups: sphingomyelins, cerebrosides, and gangliosides [41]. Sphingomyelins have a phosphocholine molecule (phosphate esterified to choline) attached to the C-1 of a ceramide [40][41], making them the only group of sphingolipids to also be classed as phospholipids [41]. Cerebrosides consist of a monosaccharide linked to the C-1 of a ceramide, whilst gangliosides have a polysaccharide chain attached to the C-1 of a ceramide. Both of these groups of compounds are classed as glycosphingolipids [41].

The two remaining main groups of lipids containing fatty acids are waxes and eicosanoids. Waxes are non-polar esters of long-chain fatty acids and long-chain alcohols, and eicosanoids are the oxygenated derivatives of C_{20} polyunsaturated fatty acids [41].

1.6 Summary

Food fraud is an issue which is becoming increasingly prevalent, necessitating the development of analytical techniques capable of detecting even subtle cases of fraud, in order to reduce the occurrence and impact of fraudulent practices. Metabonomics, the profiling of small compounds/metabolites in organic samples, can be employed as a non-targeted approach to discover subtle differences between authentic and fraudulent products. Poultry eggs are particularly susceptible to food fraud, as the potential for gaining an increased profit is greater than the likelihood of being caught. Egg yolk contains a range of different lipids, particularly those consisting of fatty acid components. Therefore, metabonomic investigations into the detection of fraud within poultry eggs are likely to uncover differences in lipid profile between egg yolk from authentic and fraudulent eggs.

1.7 References

[1] J. Spink and D. C. Moyer, "Defining the Public Health Threat of Food Fraud," *Journal of Food Science*, vol. 76, pp. 157–163, 2011.

[2] S. Tähkäpää, R. Maijala, H. Korkeala, and M. Nevas, "Patterns of food frauds and adulterations reported in the EU rapid alert system for food and feed and in Finland," *Food Control*, vol. 47, pp. 175–184, 2015.

[3] J. Spink, D. L. Ortega, C. Chen, and F. Wu, "Food fraud prevention shifts the food risk focus to vulnerability," *Trends in Food Science & Technology*, vol. 62, pp. 215–220, 2017.

[4] D. J. Armstrong, "Food chemistry and US food regulations," *Journal of Agricultural and Food Chemistry*, vol. 57, pp. 8180–8186, 2009.

[5] P. Shears, "Food fraud: A current issue but an old problem," *British Food Journal*, vol. 112, pp. 198–213, 2010.

[6] L. Manning, "Food fraud: Policy and food chain," *Current Opinion in Food Science*, vol. 10, pp. 16–21, 2016.

[7] E. Capuano, R. Boerrigter-Eenling, G. V. D. Meer, and S. M. V. Ruth, "Analytical authentication of organic products: An overview of markers," *Journal of the Science of Food and Agriculture*, vol. 93, pp. 12–28, 2013.

[8] L. Manning and J. Mei, "Developing systems to control food adulteration," *Food Policy*, vol. 49, pp. 23–32, 2014.

[9] E. Hong, Y. Lee, Y. Jeong, M. Park, B. H. Kim, and H. Sook, "Modern analytical methods for the detection of food fraud and adulteration by food category," *Journal of the Science of Food and Agriculture*, vol. 97, pp. 3877–3896, 2017.

[10] G. M. Fiorino, C. Garino, M. Arlorio, A. F. Logrieco, I. Losito, and L. Monaci, "Overview on untargeted methods to combat food frauds: A focus on fishery products," *Journal of Food Quality*, vol. 2018, pp. 1–13, 2018.

[11] (2018). The public health and safety organization, [Online]. Available: `http://www.nsf.org/services/by-industry/food-safety-quality/food-fraud` (visited on 07/2019).

[12] *Regulation of the European Parliament and Council (EC) 178/2002 of 28 january 2002, Laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety*, Official Journal of the European Union, L31/1-24, 2002.

[13] *Food Safety Act*, Chapter 16: Food Safety and Consumer Protection, 1990.

[14] *Food Drug and Cosmetic Act*, 21 US Code Chapter 9, 1938.

[15] E. Cubero-Leon, R. Pealver, and A. Maquet, "Review on metabolomics for food authentication," *Food Research International*, vol. 60, pp. 95–107, 2014.

[16] J.-F. Xia, Q.-L. Liang, P. Hu, Y.-M. Wang, and G.-A. Luo, "Recent trends in strategies and methodologies for metabonomics," *Chinese journal of analytical chemistry*, vol. 37, pp. 136–143, 2008.

[17] I. Wilson, R. Plumb, J. Granger, H. Major, R. Williams, and E. Lenz, "HPLC-MS-based methods for the study of metabonomics," *Journal of Chromatography B*, vol. 817, pp. 67–76, 2005.

[18] H. R. Tang and Y. L. Wang, "Metabonomics: A revolution in progress," *Progress in Biochemistry and Biophysics*, vol. 33, pp. 401–417, 2006.

[19] B. Zhou, J. Xiao, L. Tuli, and H. Ressom, "LC-MS-based metabolomics," *Molecular BioSystems*, vol. 8, pp. 470–481, 2012.

[20] J. K. Nicholson and J. C. Lindon, "Metabonomics," *Nature*, vol. 455, pp. 1054–1056, 2008.

[21] P. Guy, I. Tavazzi, S. Bruce, Z. Ramadan, and S. Kochhar, "Global metabolic profiling analysis on human urine by UPLC-TOFMS: Issues and method validation in nutritional metabolomics," *Journal of Chromatography B*, vol. 871, pp. 253–260, 2008.

[22] H. Gika, G. Theodoridis, R. Plumb, and I. Wilson, "Current practice of liquid chromatography mass spectrometry in metabolomics and metabonomics," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 87, pp. 12–25, 2014.

[23] S. Moco, R. J. Bino, R. C. H. de Vos, and J. Vervoort, "Metabolomics technologies and metabolite identification," *Trends in analytical chemistry*, vol. 26, pp. 855–866, 2007.

[24] D. S. Wishart, "Metabolomics: Applications to food science and nutrition research," *Trends in Food Science & Technology*, vol. 19, pp. 482–493, 2008.

[25] S. Esslinger, J. Riedl, and C. Fauhl-Hassek, "Potential and limitations of non-targeted fingerprinting for authentication of food in official control," *Food Research International*, vol. 60, pp. 189–204, 2014.

[26] *Guidance on Legislation Covering the Marketing of Eggs*, Department for Environment, Food and Rural Affairs, Animal and Plant Health Agency, 2016.

[27] *Commission Regulation (EC) 589/2008 of 23 June 2008, Laying down detailed rules for implementing Council Regulation (EC) No 1234/2007 as regards marketing standards for eggs*, Official Journal of the European Union, L163/6-23, 2008.

[28] *United Kingdom egg statistics - Quarter 4, 2018*, Department for Environment Food & Rural Affairs, 2018.

[29] *Commission Regulation (EC) No 889/2008 of 5 September 2008, laying down detailed rules for the implementation of Council Regulation (EC) No 834/2007 on organic production and labelling of organic products with regard to organic production, labelling and control*, Official Journal of the European Union, L250/1-84, 2008.

[30] *Council Directive 1999/74/EC of 19 July 1999, Laying down minimum standards for the protection of laying hens*, Official Journal of the European Communities, L203/53-57, 1999.

[31] *Council Regulation (EC) No 1234/2007 of 22 October 2007, establishing a common organisation of agricultural markets and on specific provisions for certain agricul-*

tural products (Single CMO Regulation, Official Journal of the European Union, L299/1-149, 2007.

[32] *Regulation of the European Parliament and Council No 853/2004, Laying down specific hygiene rules for on the hygiene of foodstuffs*, Official Journal of the European Union, L139/55, 2004.

[33] Y. Nys and N. Guyot, "Improving the safety and quality of eggs and egg products," in. Woodhead Publishing Limited, 2011, vol. 1: Egg Chemistry, Production and Consumption, ch. 4: Egg formation and chemistry, pp. 83–132.

[34] Y. Mine and I. D'Silva, "Egg bioscience and biotechnology," in. John Wiley & Sons Inc, 2008, ch. 4: Bioactive Components in Egg White, pp. 141–184.

[35] H. Hatta, M. P. Kapoor, and L. R. Juneja, "Egg bioscience and biotechnology," in. John Wiley & Sons Inc, 2008, ch. 4: Bioactive Components in Egg Yolk, pp. 185–237.

[36] H. D. Griffin, "Manipulation of egg yolk cholesterol: A physiologist's view," *World's Poultry Science Journal*, vol. 48, pp. 101–112, 1992.

[37] E. C. Y. Li-Chan and H. O. Kim, "Egg bioscience and biotechnology," in. John Wiley & Sons Inc, 2008, ch. 1: Structure and chemical composition of eggs, pp. 185–237.

[38] A. B. Rodríguez-navarro, N. Domínguez-gasca, A. Muñoz, and M. Ortega-Huertas, "Change in the chicken eggshell cuticle with hen age and egg freshness," *Poultry Science*, vol. 92, pp. 3026–3035, 2013.

[39] J. P. Mortola and V. Gaonac'h-Lovejoy, "The cooling time of fertile chicken eggs at different stages of incubation," *Journal of Thermal Biology*, vol. 55, pp. 7–13, 2016.

[40] M. Campbell and S. Farrell, "Biochemistry," in, 8th. Cengage Learning, 2015, ch. 8: Lipids and Proteins Are Associated in Biological Membranes, pp. 189–195.

[41] L. Moran, R. Horton, G. Scrimgeour, and M. Perry, "Principles of Biochemistry," in, 5th. Pearson Education Limited, 2014, ch. 9: Lipids and Membranes, pp. 308–320.

[42] M. Campbell and S. Farrell, "Biochemistry," in, 8th. Cengage Learning, 2015, ch. 21: Lipids Metabolism, pp. 601–607.

[43] M. I. Gurr and J. L. Harwood, *Lipid Biochemistry: An Introduction*, 4th ed. London: Chapman & Hall, 1991.

2. Techniques and Instrumentation

2.1 Aims

- *To introduce the components of a metabonomic workflow.*
- *To introduce and discuss the analytical techniques that were used throughout this research.*
- *To introduce and discuss the statistical analysis techniques that were implemented in this research.*

2.2 Metabonomics Workflow

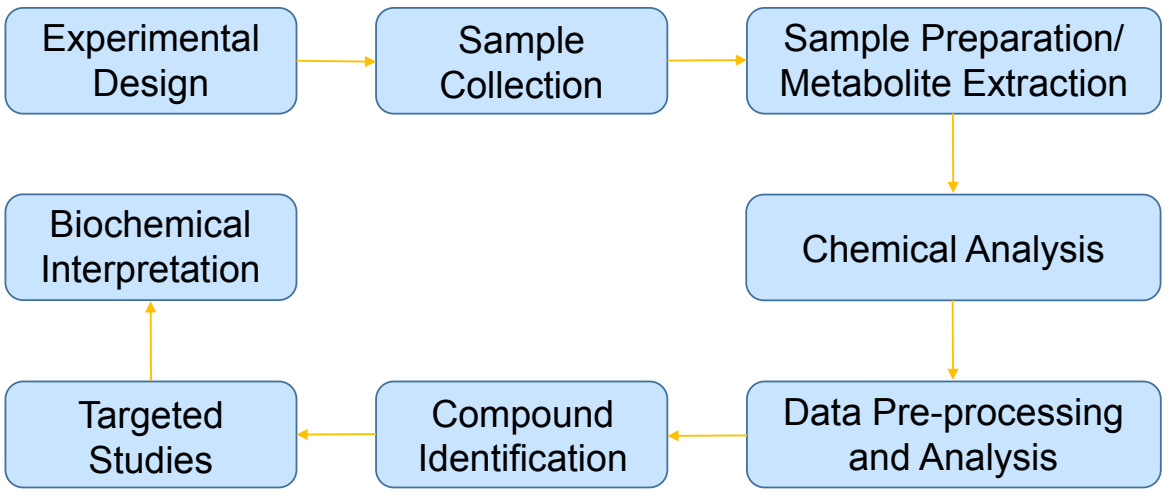


Figure 2.1: *Diagram showing the typical workflow of a non-targeted metabonomic study*

As in any kind of scientific study, the workflow that is undertaken in non-targeted metabonomic studies consists of several stages, from experimental design and sample collection, to compound identification and follow-up targeted studies. For the study to be successful, each stage must be carefully planned to ensure that high quality results can be obtained, and that all interpretations are meaningful. Figure 2.1 shows a diagram of the workflow in a typical non-targeted metabonomic study.

2.2.1 Experimental Design

The experimental design is possibly the most vital aspect of a study to get right. If the initial design and set-up of an experiment is flawed, then any results will be meaningless and any interpretations can be disregarded. It is important to identify which are the independent and dependent variables in the study, and then ensure that other potential variables are controlled, in order to be confident that any observed results are due to the factor that is under investigation, rather than an uncontrolled variable [1]. Another important factor to consider is the number of biological replicates required for each sample set [2]. A minimum of three biological replicates has been proposed, with five stated as being preferable [3]. The experimental design should also include the addition of quality control samples during the analysis [2], to act as a standard in order to monitor drift throughout the analysis and ensure that results are reliable [4].

Quality control (QC) samples in metabonomic studies are made up of equal aliquots of each individual sample extract, and are analysed at the start and end of a sequence of analyses, as well as periodically throughout the analytical sequence [4][5]. The results of the QC sample analyses can then be compared to observe any large changes that may have occurred throughout the analytical sequence, in order to give an indication of the quality and reliability of the resulting data [5]. The precision of the results of the QC samples can then be measured and compared against reference criteria to ensure that the data is reliable [4][5].

2.2.2 Sample Collection

Sample collection, and the treatment of samples prior to metabolite extraction, are also important aspects of the metabonomic workflow that can have a great impact on the results of a study. Samples should all be handled in the same manner, and exposed to the same materials, e.g. containers, and temperatures, in order to prevent any external factors from influencing differences in metabolite profile between the samples [1]. Samples should all ideally be collected at the same time, so that any changes to the metabolite profile that occur prior to extraction affect all samples equally. When this is not possible, and if metabolite extraction is not to be carried out immediately after sample collection, storage of samples at -80 °C or in liquid nitrogen, as soon as possible, would be suitable to achieve minimal metabolite degradation [2].

2.2.3 Sample Preparation and Metabolite Extraction

Following sample collection the preparation of samples and the extraction of metabolites takes place. The main purpose of metabolite extraction is to release the metabolites that are present in the samples, whilst also removing any interferences, e.g. proteins, and making the extract compatible with the proposed analytical technique [2][6]. During the extraction procedure, the metabolites can be concentrated if necessary. In targeted studies, the metabolite extraction procedure can be adapted to make it more selective to the particular compound, or class of compounds, that is being investigated. However, in non-targeted metabonomic studies the aim is to gain information about as many small molecules present in the samples as possible, therefore the extraction procedure is aimed at removing large molecules such as proteins, whilst releasing and retaining as many small molecules as possible [6].

The type of sample preparation or metabolite extraction that is carried out depends on the nature of the study, the sample type, and on the type of analysis that is to be carried

out [2][7]. For example, infra-red spectroscopy may not require sample preparation or metabolite extraction, as it could be carried out directly on the samples themselves.

One of the most common sample preparation techniques in non-targeted metabonomic studies is protein precipitation, via the addition of an organic solvent to the sample [6][7]. Following centrifugation of the sample-solvent mix, the supernatant contains many low molecular weight compounds, and no macromolecules such as proteins, which are found in the resulting pellet [7][8]. The supernatant is then retained as the metabolite extract. Deproteinisation can also be achieved through denaturing the proteins by applying heat or with the addition of an acid to lower the pH, or through ultrafiltration to filter out the protein molecules [9][10].

A liquid extraction is commonly carried out on both solid samples; solid-liquid extraction (SLE), and liquid samples; liquid-liquid extraction (LLE) [2][6]. Prior to any extraction, solid samples must be broken down and converted into a homogeneous substance that is suitable for a liquid extraction; e.g. by using a mortar and pestle [6]. Metabolites in liquid samples can also be extracted using solid phase extraction (SPE), or solid phase microextraction (SPME) [2][6]. However, both SPE and SPME are used for more selective metabolite extractions, with SPME used for targeting more volatile metabolites, and are therefore more widely used as metabolite extraction techniques in targeted metabonomic studies, rather than non-targeted studies [6].

In both LLE and SLE, a solvent or solvent mix is added to the samples in order to extract the metabolites from the sample medium into the solvent [2][6]. The polarity of the solvent used in the extraction determines the classes of compounds that are extracted. The extraction solvent also has the secondary purpose of helping to remove proteins through protein precipitation [6]. Following centrifugation of the sample-solvent mix the supernatant, which is protein-free, is removed and retained as a metabolite extract [7][8]. If a relatively polar solvent or solvent mix, such as methanol or methanol/water, is used as the initial solvent, a further extraction can then be carried out with the addition of

a more non-polar solvent or solvent mix, such as dichloromethane/methanol, in order to extract any remaining compounds that are not extracted into the first solvent [8]. These sequential extractions can take place in order to obtain extracts containing metabolites of a range of different polarities; the initial addition of a more aqueous solvent or solvent mix to the sample targeting the more polar molecules, followed by an extraction with a more organic solvent, or solvent mix, targeting the remaining less polar molecules [8].

2.2.4 Chemical Analysis

Once sample preparation and metabolite extraction have taken place, the sample extracts undergo chemical analysis. There are various different analytical techniques that have been employed in metabonomic studies, ranging from spectroscopic methods, such as Infra-Red (IR) and Nuclear Magnetic Resonance (NMR) spectroscopy, to Mass Spectrometry (MS) and chromatographic techniques. The ideal analytical technique in metabonomic studies would involve very little or no sample preparation, be rapid and high throughput, and have a high sensitivity that is equal to all compound classes. It would produce reproducible, quantitative results with rich molecular information to enable metabolite identification. It would also be inexpensive, and the analysis would be non-destructive to allow the sample to undergo further investigation [11].

Vibrational spectroscopy techniques, such as IR and Raman spectroscopy, are less commonly used in metabonomic studies compared to other analytical techniques [1]. This is due not only to their low sensitivity [12], but also due to the lack of molecular information provided [13], resulting in no identification of metabolites. However, these techniques do provide rapid, high throughput spectral metabolic fingerprints of biological samples, with very little to no sample preparation required [12]. They are also non-destructive, meaning the samples can then undergo further analysis [12].

The two main analytical techniques that are used in metabonomic studies are NMR and MS, with and without prior chromatographic separation [1][13][14][15]. NMR is

a non-destructive spectroscopic technique that provides highly reproducible, quantitative results, that are rich in molecular information to enable metabolite identification [12][16][17]. It is non-selective and can detect multiple compound classes, without any specific pre-selection of analytical conditions or parameters [11][16]. This technique requires very little to no sample preparation [12][18] and the analysis itself is relatively fast, resulting in NMR being a high throughput analytical technique [11][17]. However, NMR has a lower sensitivity than MS, and is therefore only capable of detecting a limited number of metabolites during the analysis, compared to the high numbers of metabolites detected by MS [16][17].

Mass spectrometry can be employed in metabonomic studies either on its own, through direct infusion of the sample into the mass spectrometer, or preceded by a chromatographic separation technique [15][18]. Direct infusion (DI) MS is a rapid, high throughput technique [15][18][19] with reproducible results that provide information to enable the identification of metabolites [12]. However, this direct infusion of the sample, with no prior separation of the compounds, can result in ion suppression or enhancement, affecting the precision and accuracy of the analysis [15][18][19]. Therefore, although the sensitivity of MS analysis is much greater than NMR, and the number of metabolites detected during the analysis is much higher [15], this technique may not be equally sensitive to all compounds, due to the potential of ion suppression and enhancement [12]. Although DIMS is a high throughput technique [12][15][18], it requires more sample preparation than NMR, resulting in a greater total analysis time [11][17]. It is also a destructive technique, therefore the samples cannot be retained for further analysis [11][15][17]. Mass spectrometry can be used for quantification, however appropriate standards are needed, making quantification less simple and straightforward than in NMR [17]. More prior knowledge of the samples is required in MS compared to NMR in order to optimise the parameters for the analysis [11], and the cost of running an experiment is higher when using MS [11][17].

The two most common separation techniques that precede MS analysis are Gas Chromatography (GC) and Liquid Chromatography (LC), or High Performance Liquid Chromatography (HPLC) [1][16]. The use of chromatography to separate analytes prior to MS analysis can help to reduce ion suppression and enhancement [15], as there are fewer analytes entering the ionisation chamber at any one time, lowering the chances of certain molecules inhibiting, or enhancing, the ionisation of others. However, it does increase the analysis time due to the additional time taken to separate the analytes [11][15]. Combining chromatography with MS analysis also affects the selectivity of the analysis for different metabolite classes; GC is only suitable for the analysis of volatile, thermally stable compounds of a low molecular weight [20][12], and the use of columns with different stationary phases affects which types of compounds will be separated, eluted, and detected with LC [12].

The identification of metabolites is easier with GC-MS compared to DIMS, as there are extensive spectral libraries and databases that can be searched to find a match to spectra of particular compounds [12][18][19]. However, sample preparation is more involved in GC analysis than LC or DIMS, as the samples must be derivatised in order to ensure that the compounds are volatile [1][12][15]. Both GC and LC analysis require a greater prior knowledge of the compounds that may be present in samples than DIMS, as parameters such as the column stationary phase, and the mobile phase, should be optimised in order to analyse as many compounds from as many compound classes in the sample extracts as possible [16]. The introduction of Ultra High Performance Liquid Chromatography (UHPLC) has improved throughput, peak resolution, sensitivity, and reproducibility in LC-MS, therefore its application in metabonomic studies is on the rise [15].

The use of GC and LC as independent analytical techniques is uncommon in metabonomic studies as they provide no molecular information, and so no information regarding the identification of the metabolites can be determined [11][16]. Other separation techniques, Capillary Electrophoresis (CE) and Supercritical Fluid Chromatography (SFC)

have also been used in combination with MS for the analysis of metabolite extracts in metabonomic studies, however they are not currently as widely used as GC or LC, which are well established, robust techniques in the metabonomic field [20].

2.2.5 Data Pre-processing and Analysis

The chemical analysis of sample extracts results in the production of a large amount of raw data. This data must then be pre-processed, in order to convert it into a format that is suitable for data analysis [21][22]. The required pre-processing is different for data resulting from the different types of analysis, but typically consists of: peak alignment, baseline correction, normalisation, signal filtering and noise removal, and feature detection [21][22]. Data analysis can then be carried out on the processed data.

In metabonomic studies, both univariate and multivariate statistical analyses are often used to analyse the data [23][24]. Multivariate analysis (MVA) is often carried out on large data sets in order to analyse all variables simultaneously and observe any patterns that are present in the data [18][24]. In metabonomic studies, each variable is an individual compound. There are two main categories of multivariate statistical analyses: supervised and unsupervised, and the two most popular of these analyses are Partial Least Squares Discriminant Analysis (PLS-DA) and Principal Component Analysis (PCA) respectively [22][25]. The purpose of unsupervised MVA is to explore and discover patterns and trends within the data, without the labelling of sample sets prior to analysis, whilst supervised MVA is used to make predictions about samples from unknown sample sets, based on the variables in samples from known sample sets [22].

Univariate analysis, where only one variable is analysed at a time, can then be carried out. This involves comparing each individual variable, or compound, between different sample sets, using tests such as t-tests and ANOVA, to see whether or not there is a statistically significant difference between the sample sets [18][23]. This helps to determine which compounds are of significant interest in discriminating between sample sets, and

therefore show potential as biomarkers of the particular factor that is under investigation.

2.2.6 Compound Identification, Targeted Studies, and Biochemical Interpretation

Following the use of multivariate and univariate analyses to explore the data and search for variables of statistical significance, the compounds of interest must then be identified, before targeted studies are carried out focussing on just these compounds [2][18]. The initial identification of these compounds is based on the molecular information obtained from the chemical analysis of the samples, therefore metabolite identification is easier following some analytical techniques than others. This identification is confirmed through the use of an analytical standard [2]. Targeted studies can then be carried out, in order to confirm the statistical significance of the compounds, and obtain more quantitative information [18]. The final stage of a metabonomic study is the interpretation of the results, regarding the metabolic pathways involved in the biosynthesis and degradation of the compounds of interest [2][25].

Further studies can then be carried out in order to validate the use of the compounds as biomarkers of the particular factor that is being investigated [26].

2.3 Chromatography

Chromatography is a technique which facilitates the separation of compounds in a sample due to their different affinities for two phases; a stationary phase, which is typically immobilised within a column, and a mobile phase which travels through the column [27]. The different affinities of the compounds for the stationary and mobile phase causes the differential migration of the components through the column, resulting in the separation of the compounds within the sample [27]. Compounds with a stronger affinity for the

stationary phase take longer to pass through the column than those with a stronger affinity for the mobile phase, which pass through the column more quickly [28]. There are two main types of chromatography; Gas Chromatography (GC) and Liquid Chromatography (LC) or High Performance Liquid Chromatography (HPLC), as well as others including Supercritical Fluid Chromatography (SFC) [27][29][30].

2.3.1 High Performance Liquid Chromatography

A HPLC instrument consists of five major components: a solvent reservoir, a pump, an injector, a column, a detector, and a data acquisition system (usually a computer) [27][29][30], as shown in Figure 2.2.

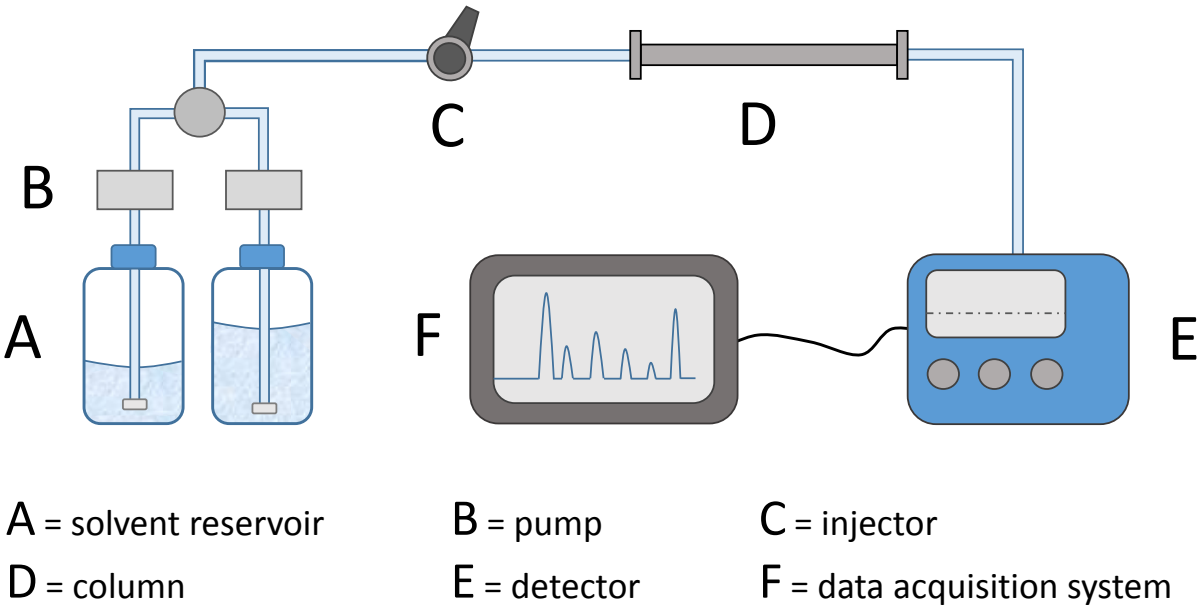


Figure 2.2: *Diagram showing a simplified HPLC system. Adapted from Higson and Meyer [28][30].*

The solvent reservoir stores the mobile phase solvents, and the pump forces these solvents out of the reservoir, through a mixing chamber, along the HPLC system and through the column, which is packed with a solid stationary phase [27]. The sample is introduced through the injector, which consists of a high pressure valve and a sample loop. The sample is injected onto the sample loop whilst the valve is in the load position,

and the mobile phase flows straight through the column, bypassing the sample loop. The valve then rotates, allowing the mobile phase to flow through the sample loop, carrying the sample out and through the column [27][29][30]. The sample components are then separated on the column, before eluting and reaching the detector at different times. When an analyte reaches the detector, of which there are several different types, the data acquisition system registers this and a peak appears on the resulting chromatogram [27].

In HPLC the mobile phase consists of a liquid solvent, or combination of liquid solvents, and the column is packed with a solid stationary phase [27][28]. If the mobile phase remains at a fixed composition of solvent, it is known as isocratic elution, and if the composition of solvents in the mobile phase changes throughout the analysis, it is referred to as an elution gradient [27][29][30]. The use of an elution gradient can help to elute compounds, or speed up the elution of compounds, that would otherwise remain on the column or take a long time to pass through [30]. The stationary phase is typically composed of spherical particles of silica gel, which ensure a homogeneous packing of the column [27][29]. This results in a polar stationary phase, suitable for the separation of more polar compounds. Chromatographic separation using this polar stationary phase is referred to as Normal Phase (NP) HPLC [27][29]. In Reversed-Phase (RP) HPLC the opposite is true; non-polar groups, typically C4, C8, and C18 hydrocarbon chains, are bonded onto the silica resulting in a non-polar stationary phase, which is suitable for the separation of more non-polar compounds [27][29][30].

As the stationary phase in RP-HPLC is non-polar, the more non-polar compounds in a sample have a stronger affinity for the stationary phase than the polar compounds, resulting in a longer retention time of the non-polar compounds compared to the polar compounds [27][30]. Therefore in Figure 2.3, for RP-HPLC, the analytes that elute first, resulting in the earlier peaks in the chromatogram, are the more polar compounds, and the analytes that elute later, resulting in the subsequent peaks in the chromatogram, are the more non-polar analytes. In an elution gradient for RP-HPLC, the initial mobile phase

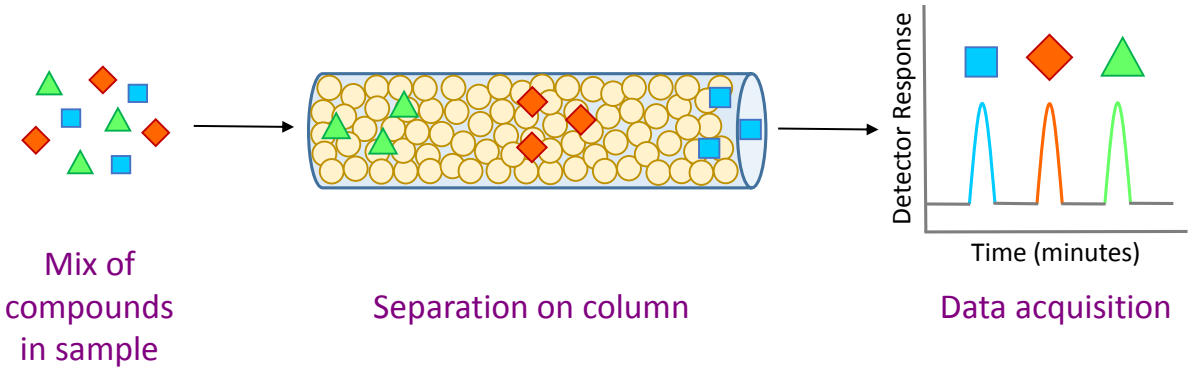


Figure 2.3: *Diagram showing the separation of compounds in a HPLC column. Based on a diagram from Meyer [30].*

composition is highly polar, consisting mainly of the polar solvent, most commonly water, which elutes the polar analytes that do not have a strong affinity for the stationary phase [27][30]. In order to elute the more non-polar analytes, which have a stronger affinity for the stationary phase, the proportion of the organic, non-polar solvent (commonly methanol or acetonitrile) increases throughout the analysis, increasing the eluotropic strength of the mobile phase [27][30].

An alternative to NP-HPLC for the separation of polar analytes is Hydrophilic Interaction Liquid Chromatography (HILIC). The stationary phase can be either bare silica, as in most NP stationary phases, or have polar groups bonded onto the silica [31][32]. The more polar analytes in a sample are retained on the column for longer than the non-polar analytes, and so the elution order of analytes is opposite to that observed with RP [32]. In Figure 2.3 for HILIC separations the analytes eluting first, resulting in the earlier peaks in the chromatogram, are the more non-polar analytes whilst those that elute later are the polar analytes. The mobile phases used in HILIC are similar to RP-HPLC, except the eluotropic strength of the mobile phase increases with a higher proportion of aqueous solvent (water). Therefore, an elution gradient in HILIC would start with a high proportion of organic solvent in the mobile phase, and increase the proportion of aqueous solvent throughout the analysis [31][32].

The aim of HPLC analysis is to achieve the greatest separation of the compounds

within a sample in the minimum amount of time [30]. Once a column with an appropriate stationary phase, and a suitable mobile phase have been chosen, the chromatographic separation can be optimised by developing an elution gradient which results in a constant elution of separated compounds over a short amount of time. The flow rate of the mobile phase, and temperature of the column, may also be optimised to achieve better, more efficient, separation [30].

HPLC is often coupled to another technique, such as Mass Spectrometry (MS).

2.4 Mass Spectrometry

Mass Spectrometry (MS) is a technique that is used to provide information about compounds in a sample, following their ionisation. The analytes are introduced into the ionisation source, where they are either positively or negatively ionised before being directed towards the mass analyser. The analyte ions then pass through the mass analyser, which separates the ions based on their mass to charge ratios (m/z), before reaching the detector. Once the ions have made contact with the detector, a data acquisition system, typically a computer, records the signals generated by the detector, and produces a mass spectrum, showing a plot of the relative abundances of ions of different m/z [33][34].

2.4.1 Ionisation

There are several different ionisation methods that can be employed in MS analysis, and the method of choice depends on the nature of the analytes and the phase state of the sample [27]. When LC is used to separate compounds prior to MS analysis, the sample is in the liquid phase, and therefore Atmospheric Pressure Ionisation (API) techniques, in an ionisation chamber outside of the vacuum of the MS instrument, are required to ionise the analytes [33][35][36]. One of the common types of API that is used in MS is ElectroSpray Ionisation (ESI).

2.4.1.1 ElectroSpray Ionisation

In ESI, analytes in a liquid sample are subjected to a high voltage as they pass through a narrow steel capillary, producing a spray of charged droplets of the same polarity as that of the applied voltage as they exit the capillary [33][37][38][39]. When the ionised analytes in solution reach the end of the capillary, the analyte ions of the chosen polarity accumulate at the surface of the liquid solution at the tip of the capillary, producing a Taylor cone [33][37][40]. The surface tension of the Taylor cone is too low to keep the cone stable, and the ionised analyte solution is dispersed from the tip of the Taylor cone to form charged droplets of analyte solution [33][37][40][41], as seen in Figure 2.4 - **A**. The liquid solvent of the charged droplets of analyte ions evaporates, aided by a heated drying gas, decreasing the size of the droplets, and increasing their charge density [34][37][40][41] (Figure 2.4 - **B**). There are then two different potential modes of action resulting in the production of free, gaseous analyte ions: **C1** and **D1** or **C2** and **D2** in Figure 2.4. One potential mode of action is further desolvation of the droplet solvent, resulting in the distortion of the droplets, due to charge repulsion, to form a Taylor cone (**C1**), which then releases analyte ions [37] (**D1**). Another potential mode of action is an increase in the charge density of the droplets, due to the desolvation of the droplet solvent, resulting in the droplets disintegrating into smaller droplets when the charge repulsion between analyte ions exceeds the surface tension of the droplets [37][40] (**C2**). This is known as a Coulomb explosion or fission. This then continues, and the droplets disintegrate into smaller and smaller droplets until the free gaseous analyte ions are formed [37][40] (**D2**).

ESI is a soft ionisation technique that produces little fragmentation [37][41]. Ions are typically formed by protonation or deprotonation to form $[M+H]^+$ or $[M-H]^-$ ions respectively, although adduct formation and multiply charged ions are common. Ions can also be formed by the attachment of cations or anions to the analytes, to produce other positively or negatively charged species [37]. Identification can be difficult due to the lack of fragmentation patterns and molecular ion (M^+) in the resulting mass spectrum.

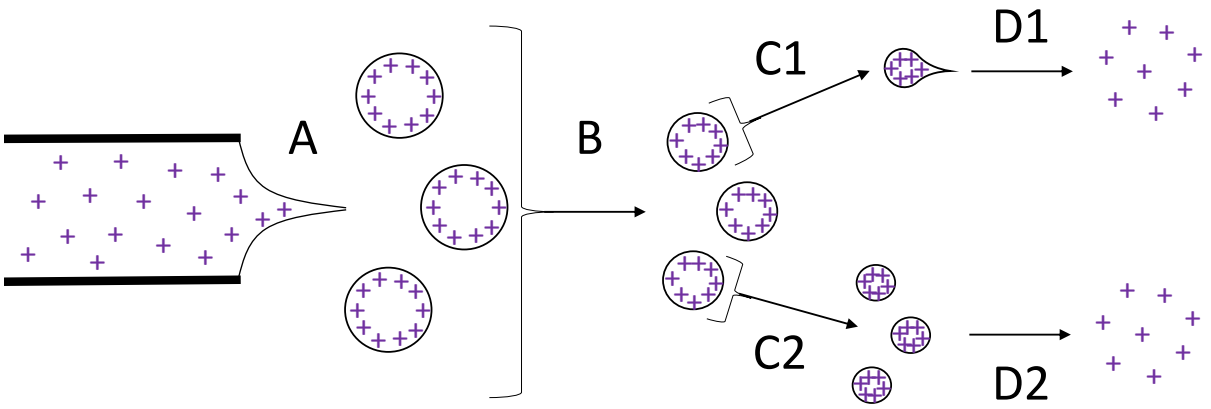


Figure 2.4: *Diagram showing ESI modes of action. Adapted from Ho, Greaves, and Awad [34][37][38].*

Following ionisation, the gaseous analyte ions are transferred into the mass analyser by passing through several ion optics, which focus the ions into a beam with a stable trajectory [37][40].

2.4.2 Mass Analysers

Mass analysers separate ions in the mass spectrometer based on their mass-to-charge ratios (m/z). They can be divided into two broad categories; scanning mass analysers which transmit ions of different m/z successively through time, and mass analysers which transmit all ions simultaneously [33].

2.4.2.1 Quadrupole Mass Analysers

A single quadrupole mass analyser is one of the most common scanning mass analysers. It consists of four parallel conductive rods: one pair of rods have a positive direct current (DC) potential applied to them, with a superimposed alternating radio frequency (RF) potential, and the other pair have a negative DC potential applied, again with a superimposed alternating RF potential [34][42]. The resulting electric field creates a stable trajectory for ions within a certain m/z range, allowing them to pass through the

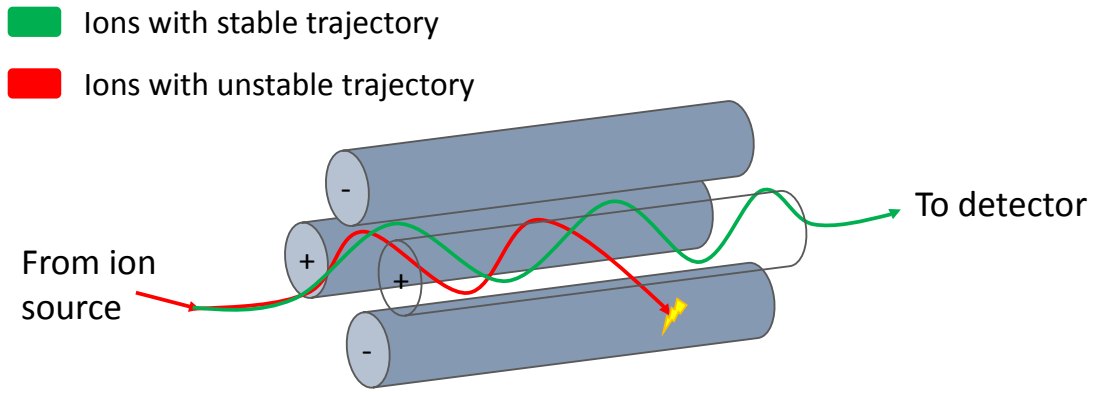


Figure 2.5: *Diagram showing a representation of a quadrupole mass analyser. Adapted from University of Bristol [43].*

quadrupole and reach the detector. Analyte ions not within this m/z range will have an unstable trajectory and will collide with one of the rods, discharge and become neutralised, failing to reach the detector [34][42]. By progressively changing the DC and RF voltages that are applied to these four rods, all of the ions of different m/z, within the total desired mass range, can be transmitted through the quadrupole over time, via a stable trajectory, to reach the detector. This results in ions of different m/z reaching the detector at different times [37][42]. Figure 2.5 shows a diagram representation of a quadrupole mass analyser.

2.4.2.2 Time-of-Flight Mass Analysers

A Time-of-Flight (ToF) mass analyser transmits all ions simultaneously, and separates ions of different m/z based on the time it takes for them to travel the length of a field-free drift region of the analyser, to reach the detector [33][42]. An ion beam that has been focussed by the ion optics enters the ion modulator region of the ToF analyser, and the ions are then pushed into an ion acceleration region, where they acquire a greater energy and are accelerated into the field-free region, known as the flight tube, where the separation of the ions occurs [33][44]. All ions acquire the same kinetic energy in the acceleration region, and therefore the velocity of the ions as they pass through the flight tube is dependent upon their mass [33][37]. Ions with the same charge but

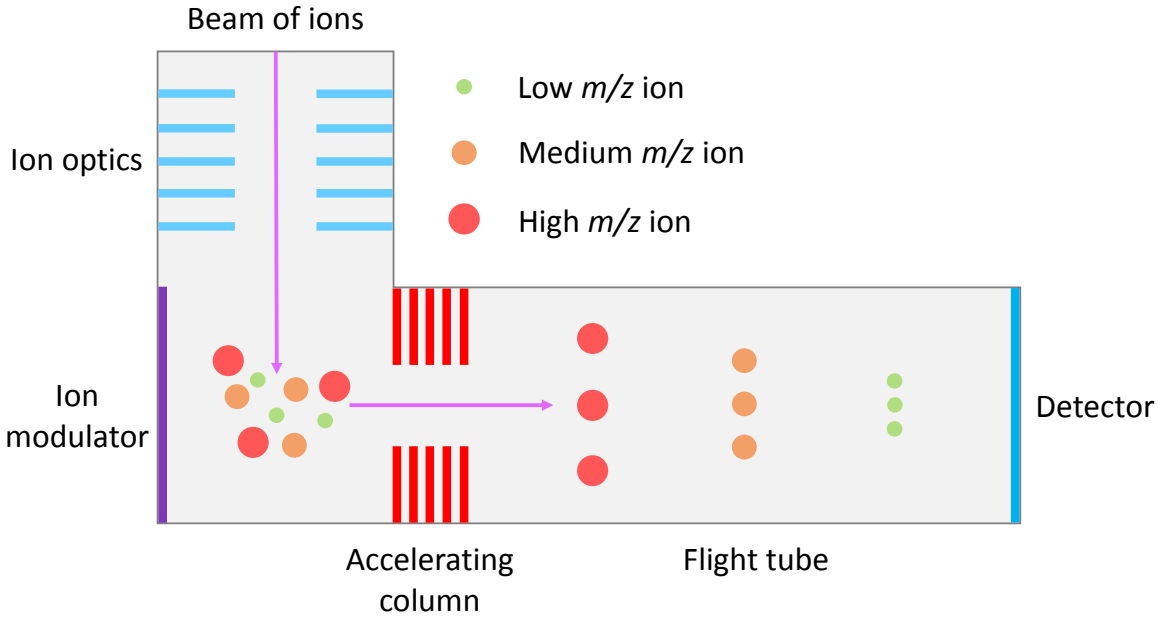


Figure 2.6: *Diagram showing a representation of a ToF mass analyser. Adapted from Rubinson et al. and Chernushevich et al. [29][44].*

different masses, and therefore different m/z, will travel at different velocities and reach the detector at different times; ions of a lower m/z will have a greater velocity and will therefore reach the detector first [37][42]. The m/z of the ions reaching the detector are determined by the length of time it takes for them to travel the length of the flight tube [33]. Figure 2.6 shows a simple diagram representation of a ToF mass analyser.

The ToF mass analyser in Figure 2.6 is a linear ToF, where ions travel linearly through the flight tube to reach the detector. However, linear ToF mass analysers can suffer from poor mass resolution due to a short path length of ions, and small kinetic energy differences between ions of the same m/z [33][37]. These small differences in kinetic energy cause ions of the same m/z to travel at slightly different velocities through the flight tube, reaching the detector at slightly different times, resulting in a loss of resolution between ions of different m/z [33][37]. Reflectrons, or ion mirrors, are used to improve the mass resolution of ToF mass analysers by increasing the path length, and correcting the kinetic energy dispersion between ions. The reflectron is located at the opposite end of the flight tube to where the ions enter, in the same position as the detector in a linear ToF mass analyser, and has a voltage applied of the same polarity as the ions, in order

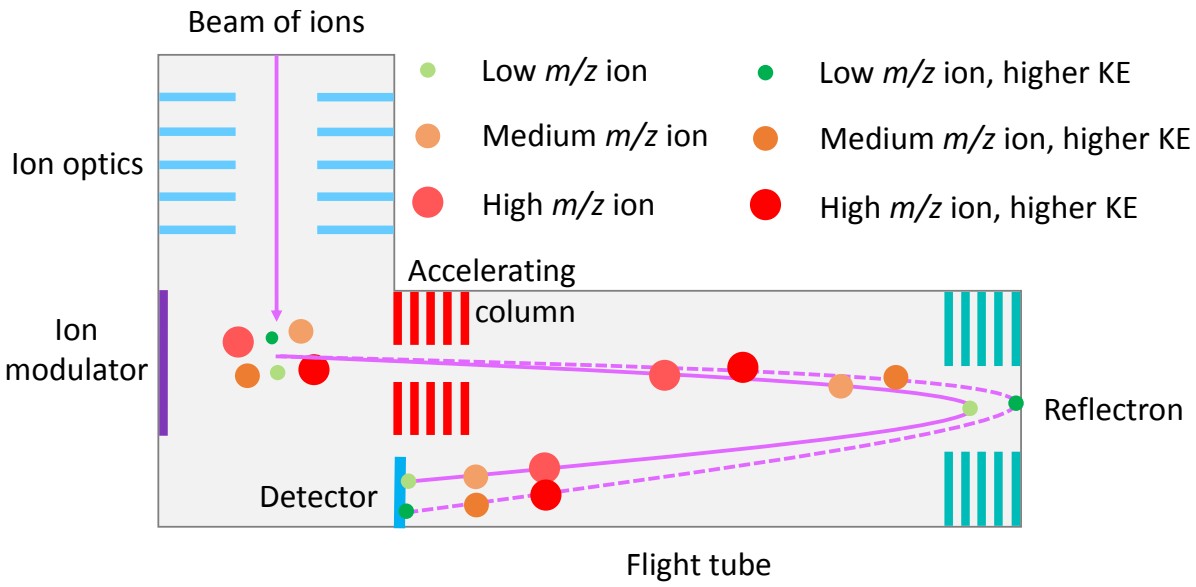


Figure 2.7: *Diagram showing a representation of a ToF mass analyser with a reflectron. The dashed line shows the flight path of ions with a slightly higher kinetic energy (KE). Adapted from Rubinson et al. and Chernushevich et al. [29][44].*

to reflect them back down the flight tube towards the detector, which is located at the same end of the flight tube that the ions enter from [33][37]. Figure 2.7 shows a diagram representation of a ToF mass analyser with a reflectron.

The ions travel the length of the flight tube towards the reflectron, becoming separated based on their different velocities due to their different m/z, as described previously [33][44]. Ions of the same m/z that have acquired slightly different kinetic energies, and therefore travel at different velocities, reach the reflectron at slightly different times. The ions with a higher velocity reach the reflectron first; however, they penetrate further into the reflectron than the ions travelling at a lower velocity, before being reflected back along the flight tube, and therefore experience a longer path length before reaching the detector. This results in the ions with the same m/z, but slightly different kinetic energies, reaching the detector at the same time, correcting the kinetic energy dispersion and improving the mass resolution of the ToF mass analyser [29][33][37].

2.4.2.3 Hybrid Mass Analysers

Some mass spectrometers combine multiple mass analysers, such as a triple quadrupole which consists of three quadrupoles in tandem, or multiple *types* of mass analyser such as a quadrupole-time-of-flight (Q-ToF), which consists of two quadrupoles followed by a ToF mass analyser. Those that combine multiple types of mass analyser are known as hybrid instruments. Better performance can be achieved from hybrid instruments as they combine the strengths of the two mass analysers, whilst avoiding their weaknesses [33].

2.4.2.3.1 Quadrupole-Time-of-Flight Mass Analysers

In a Q-ToF mass spectrometer there are two quadrupoles situated in tandem, Q1 and q2, followed by a ToF mass analyser [33][42] as shown in Figure 2.8.

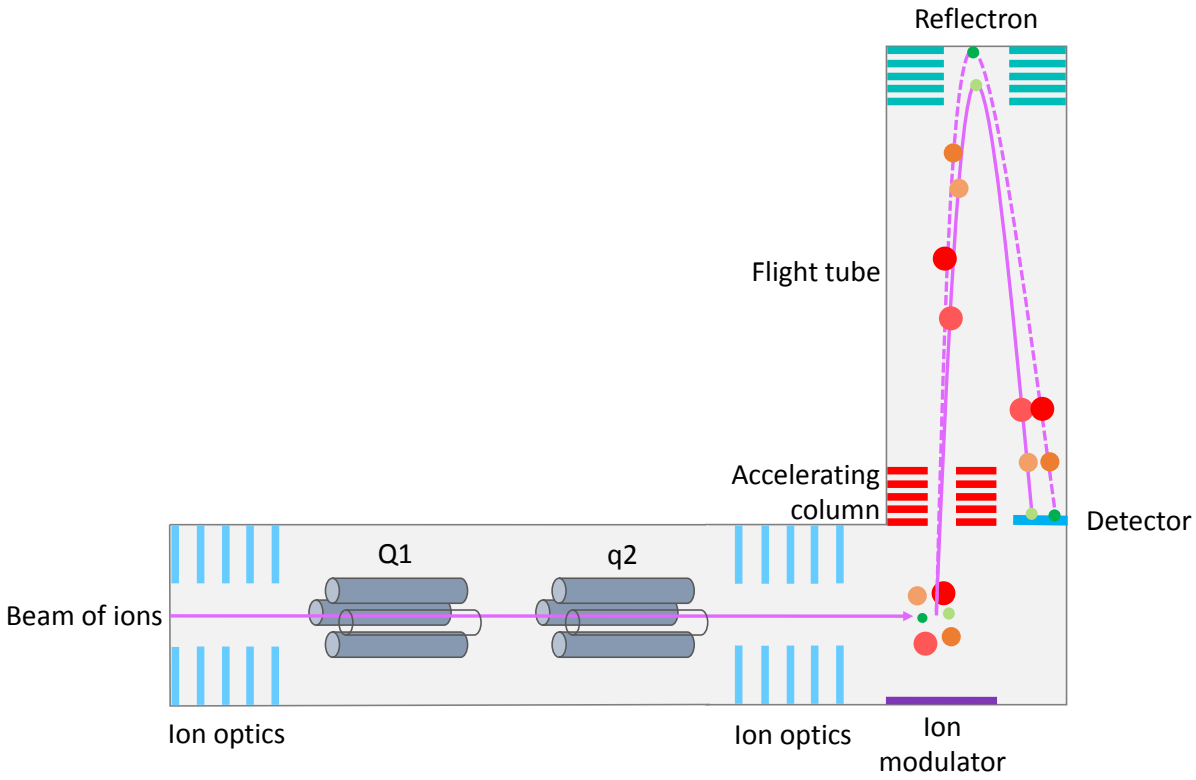


Figure 2.8: *Diagram showing a representation of a Q-ToF mass spectrometer. Based on diagrams from Rubinson et al., Hoffman and Stroobant, and Chernushevich et al. [29][33][44].*

In MS mode, the two quadrupoles operate in RF-only mode and act simply as ion guides, transmitting the ions to the ToF mass analyser [33]. However, in MS/MS mode Q1 is operated in the mass filter mode and selects only ions of certain *m/z* to be transmitted through to q2, which acts as a collision cell. The ions that enter q2 then undergo fragmentation due to collisions with neutral gas molecules, before passing through ion optics which focus the ions into a beam [33][44]. The ion beam then enters the ToF mass analyser, which acts as described previously in Section 2.4.2.2.

2.4.2.4 Tandem Mass Spectrometry

Mass spectrometers with multiple mass analysers allow for a type of analysis known as tandem mass spectrometry, which involves two stages of mass analysis [33]. There are four main scan modes of tandem mass spectrometry: product ion scan, precursor ion scan, neutral loss scan, and selected reaction monitoring [33][34], see Figure 2.9.

In product ion scan mode, the first mass analyser selects ions of a certain *m/z*, and allows them to pass through into a collision cell where they are fragmented before reaching the final mass analyser, which analyses the masses of all of the fragment ions [33][34]. Precursor ion scans are used to determine the precursor ions that produce a certain product ion following fragmentation. The first mass analyser scans a range of *m/z* greater than the *m/z* of the product fragment ion, and transmits all ions with an *m/z* above this value through to a collision cell where they are fragmented. The final mass analyser then selects ions with the *m/z* of the product fragment ion to be transmitted to the detector [33]. In a neutral loss scan, both mass analysers scan a mass range, with the final mass analyser set to scan a lower mass range at an offset equal to the mass of a neutral fragment, *f*, in order to detect the fragmentations that lead to the loss of the neutral fragment [33][34]. In selected reaction monitoring, both mass analysers are set to select ions of certain *m/z*, however the ions selected by the first mass analyser are only detected if they produce ions of the specific mass selected by the final mass analyser [33].

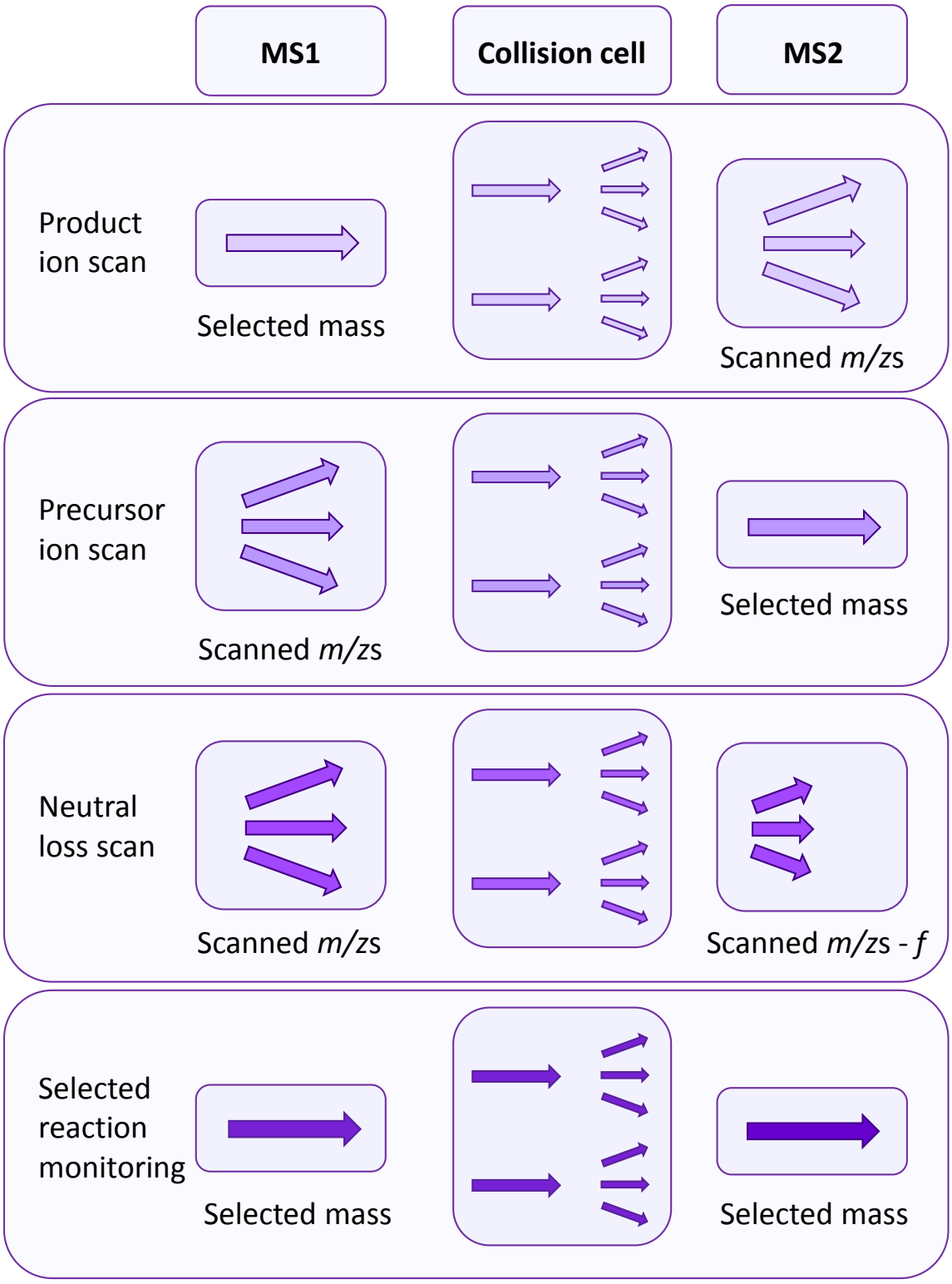


Figure 2.9: *Diagram showing the different scan modes for tandem mass spectrometers. Adapted from Hoffmann and Stroobant [33].*

2.5 Statistical Analysis

There are several stages involved in the statistical analysis workflow that is applied to data generated from metabonomic studies; ranging from data pre-processing, to multivariate and univariate analysis [45][46].

2.5.1 Data Pre-processing

Data pre-processing is the transformation of the raw data resulting from chemical analysis, into a format that is suitable for data analysis [46]. There are many different software programs that can be used for the pre-processing of LC-MS data, ranging from commercial products that are compatible only with certain data file types produced by instruments from the vendor, to freely available programs that are compatible with data files produced by instruments from a variety of different vendors, such as *XCMS Online* [4][45].

2.5.1.1 *XCMS Online*

The web-based platform *XCMS Online* is capable of carrying out the whole metabonomic statistical workflow, from data pre-processing to data analysis and metabolite identification [47]. However, following the data pre-processing stage, it produces a feature table which contains the peak areas of all compounds in each sample, with compounds represented by their *m/z* and retention time. Therefore, *XCMS Online* can also be used only for data pre-processing, and the feature table can be opened in *Microsoft Excel*, ready for data analysis.

The data pre-processing carried out by *XCMS Online* broadly involves peak detection, peak matching and retention time correction, and chromatogram alignment [48][49]. *XCMS Online* has default parameters that have been optimised to pre-process data re-

sulting from analysis using various different instruments [47]. There is the option to customise these parameters; for example the signal-to-noise threshold can be changed depending on the noise level in the mass spectra, and the mass tolerance can be adjusted for the mass identification. There are also various methods that *XCMS* can use for feature detection, and retention time correction and alignment, which can be changed from the default settings [47].

2.5.2 Multivariate Analysis

Multivariate analysis (MVA) involves the analysis of all variables in the data simultaneously, and can be either supervised or unsupervised [22][23]. An example of an unsupervised MVA technique is Principal Component Analysis.

2.5.2.1 Principal Component Analysis

Principal Component Analysis (PCA) is a data reduction technique, which is used to reduce the size of a large data set containing many correlated variables, into a much smaller data set containing just a few, uncorrelated variables known as principal components [22][25]. There is an eigenvalue associated with each principal component, which gives the amount of variance within the data set that is explained by that principal component. The first principal component accounts for the most variation within a data set, the second component accounts for the next highest amount of variation etc. [50]. The values of each of these new variables, or principal components, for each of the samples within the data set are known as *scores* [51]. The scores of two principal components, for all of the samples, can be plotted against each other to produce a *scores plot*. This plot is then used to discover groupings between the samples and sample sets [22].

There are also values known as *loadings* that exist between all variables in the data set and all principal components following PCA. Each loading value describes how much

the variable contributes to the variation accounted for by the principal component [51].

When the variables included in the PCA have considerably different variances, the variables with the larger variances can dominate the first principal component with very large loading values. When this is the case the data can be standardised prior to PCA, resulting in all variables having an equal weighting, regardless of their variances [50][52].

2.5.3 Univariate Analysis

Whilst multivariate analysis is the simultaneous analysis of all variables within a data set, univariate analysis is the analysis of just one variable within a data set at a time. Univariate analysis uses significance tests to compare the means of the measurements of one particular variable at a time, between different sample sets [23]. When carrying out univariate analysis, some tests assume certain conditions, such as normality of the data and homogeneity of variance [50].

2.5.3.1 Normal Distribution

When the variance associated with the measurement of a variable in a set of data is due to random processes, rather than any additional factors which may influence and bias the data, the probability distribution of the measurements of the variable follow what is known as the *normal distribution*. When plotted on a graph, the normal distribution shows a characteristic symmetrical bell-shaped curve, where the measurements around the mean have a high probability of occurrence, and those further away from the mean have a lower probability of occurrence [53]. Normality of the data refers to when the data follows this normal distribution, and can be assumed when only one independent variable is changed between sample sets, and the remaining independent variables are well controlled, and when it can be certain that no bias has affected the samples during chemical analysis.

2.5.3.2 Homogeneity of Variance

Homogeneity of variance simply means that the variances of all sample sets in a data set are equal [54]. When the sample sets are of different sizes, the variance is unequal, therefore appropriate significance tests that account for this unequal variance must be used. However, when the sample sets are of equal size it is sensible to investigate whether the variance between the sample sets is equal or not, by carrying out certain tests such as the F-test or the Levene's test.

The F-test compares the variances between two sample sets to determine whether they are equal or not. The F value is a ratio of the two variances: $F = \frac{s^2_{max}}{s^2_{min}}$, where $s^2_{\max}$ is the highest variance and $s^2_{\min}$ is the lowest variance [54][55]. When computer software is used to carry out an F-test a p-value is reported, which represents the probability of observing an F value equal to or greater than the observed F value if the null hypothesis, that the variances are equal, was true [54]. Values of $P<0.05$ (i.e. a 5% chance of observing an F value equal to or greater than the observed F value if the null hypothesis was true) are typically seen as significant, $P<0.01$ as highly significant, and $P<0.001$ as very highly significant.

A Levene's test is used to compare the variances between multiple sample sets to determine whether the assumption of equal variances is true [54]. The absolute deviations of values within each sample set, from the mean or median of the sample set, are used as data values in a one-way ANOVA (see Section 2.5.3.4). This one-way ANOVA then determines whether the variance between the means of the absolute deviations of all of the sample sets is statistically significant, and hence whether there is a significant difference in variance between the sample sets [54].

2.5.3.3 Student's *t*-test

A two-sample student's t-test can be used to compare the means of two sample sets of normally distributed data, and determine whether the null hypothesis, that there is no significant difference between the sample sets, is true [54]. If the alternative hypothesis is that the mean of one sample set is significantly different in a particular direction (e.g. greater) than the mean of the other sample set, then a *one-tailed* t-test is carried out. If the alternative hypothesis is that the mean of one sample set is significantly different to that of the other sample set, irrespective of direction, then a *two-tailed* t-test is carried out [54].

The t-statistic for a two-tailed t-test is: $t = \frac{|\bar{X}_1 - \bar{X}_2|}{SE}$ where $\bar{X}_1$ and $\bar{X}_2$ are the means of sample sets 1 and 2 respectively, and SE is the standard error, derived from the pooled standard deviation of the two sample sets [53][54]. The two-tailed t-test can be carried out for data with both equal and unequal variances, however the calculation of the standard error is different when the variances are unequal. When computer software is used to carry out a two-tailed t-test a p-value is reported, which represents the probability of observing a t value equal to the calculated t value, or further away from *zero* than the calculated t value, if the null hypothesis was true [54].

2.5.3.4 ANOVA/Welch test

When the homogeneity of variances is found to be true, ANOVA can be used to compare the means between several sample sets to test for any significant difference between them [50][54]. However, when the variances are found to be unequal between sample sets, the Welch test can be used instead [56].

When just a single factor is changed between the sample sets that are being compared, a *one-way* ANOVA is carried out [54][57]. ANOVA tests whether the variation between the mean values of several sample sets is likely to be caused by random chance (the null

hypothesis), or whether the changing of the controlled factor between the sample sets is more likely to be responsible for the variation. It does this by separating and estimating the variation due to random chance (M_w - *within group* variation) and the variation caused by the controlled factor (M_b - *between group* variation), where M is the mean square value [50][54]. The F-statistic is used in ANOVA to measure the ratio of between group variation (variation between the means of different sample sets) to within group variation (variation between samples within a sample set): $F = \frac{M_b}{M_w}$ [54]. Again, when computer software is used to carry out ANOVA a p-value is reported, which represents the probability of observing an F value equal to or greater than the observed F value, if the null hypothesis was true.

The Welch test is used instead of the one-way ANOVA when there are unequal variances between the sample sets. It modifies the F-statistic that is used in ANOVA to correct for this heterogeneity of variances [56].

2.5.3.5 Post-hoc tests

Tests such as ANOVA or Welch tests, that find a significant difference between the means of several sample sets within a set of data, can not provide information regarding where the statistical significance lies, and which combinations of sample sets are significantly different to one another. Therefore, additional post-hoc tests are carried out in order to provide a pairwise comparison of all of the sample sets included in the analysis, to determine which particular sample sets are significantly different to each other [58][59].

It is possible to carry out a series of individual two-sample t-tests between all of the sample sets, however the higher the number of pairwise t-tests, the greater the error rate for a type 1 error (the rejection of the null hypothesis when it is actually true). Post-hoc tests control the error rate of type 1 errors when carrying out pairwise comparisons, and therefore these tests are preferable to a series of pairwise t-tests [58][59].

The Tukey test commonly follows ANOVA for the pairwise comparison of sample sets

of equal variances, and is very similar to the two-sample t-test [58][60]. When the sample sets have unequal variance, a Games-Howell test is used for the pairwise comparison of sample sets, following the Welch test. This test is very similar to the two-sample t-test for unequal variances [58].

2.6 Summary

There are many stages involved in a metabonomics study; from experimental design, through sample and data analysis, to biochemical interpretation. Common sample preparation/metabolite extraction techniques are protein precipitation and liquid extraction, which are often followed by LC-MS analysis of the resulting sample extract. The collected data is then pre-processed and analysed using both multivariate and univariate analyses, and the results interpreted to gain some biochemical understanding.

2.7 References

[1] S. Barnes, H. Benton, K. Casazza, S. Cooper, X. Cui, X. Du, J. Engler, J. Kabarowski, S. Li, W. Pathmasiri, J. Prasain, M. Renfrow, and H. Tiwari, "Training in metabolomics research. i. Designing the experiment, collecting and extracting samples and generating metabolomics data," *Journal of Mass Spectrometry*, vol. 51, pp. 461–475, 2016.

[2] B. Zhou, J. Xiao, L. Tuli, and H. Ressom, "LC-MS-based metabolomics," *Molecular BioSystems*, vol. 8, pp. 470–481, 2012.

[3] L. Sumner, A. Amberg, D. Barrett, M. Beale, R. Beger, C. Daykin, T. Fan, O. Fiehn, R. Goodacre, J. Griffin, T. Hankemeier, N. Hardy, and J. Harnly, "Proposed minimum reporting standards for chemical analysis Chemical Analysis Working

Group (CAWG) Metabolomics Standards Initiative (MSI)," *Metabolomics*, vol. 3, pp. 211–221, 2007.

[4] G. Theodoridis, H. Gika, and I. Wilson, "LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics," *Trends in Analytical Chemistry*, vol. 27, pp. 251–260, 2008.

[5] T. Sangster, H. Major, R. Plumb, J. A. Wilson, and I. D. Wilson, "A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis," *The Analyst*, vol. 131, pp. 1075–1078, 2006.

[6] B. Álvarez-Sánchez, F. Priego-Capote, and M. D. L. de Castro, "Metabolomics analysis ii. Preparation of biological samples prior to detection," *Trends in Analytical Chemistry*, vol. 29, pp. 120–127, 2010.

[7] R. Wawrzyniak, A. Kosnowska, S. Macioszek, R. Bartoszewski, and M. Markuszewski, "New plasma preparation approach to enrich metabolome coverage in untargeted metabolomics: Plasma protein bound hydrophobic metabolite release with proteinase k," *Scientific Reports*, vol. 8, pp. 1–10, 2018.

[8] P. Masson, A. Alves, T. Ebbels, J. Nicholson, and E. Want, "Optimization and evaluation of metabolite extraction protocols for untargeted metabolic profiling of liver samples by UPLC-MS," *Analytical Chemistry*, vol. 82, pp. 7779–7786, 2010.

[9] S. Tulipani, R. Llorach, M. Urpi-Sarda, and C. Andres-Lacueva, "Comparative analysis of sample preparation methods to handle the complexity of the blood fluid metabolome: When less is more," *Analytical Chemistry*, vol. 85, pp. 341–348, 2013.

[10] Y. Chen, J. Xu, R. Zhang, and Z. Abliz, "Methods used to increase the comprehensive coverage of urinary and plasma metabolomes by MS," *Bioanalysis*, vol. 8, 2016.

[11] H. R. Tang and Y. L. Wang, "Metabonomics: A revolution in progress," *Progress in Biochemistry and Biophysics*, vol. 33, pp. 401–417, 2006.

[12] W. B. Dunn, N. J. C. Bailey, and H. E. Johnson, "Measuring the metabolome: Current analytical technologies," *Analyst*, vol. 130, pp. 606–625, 2005.

[13] J. Lindon and J. Nicholson, "Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery," *Trends in Analytical Chemistry*, vol. 27, pp. 194–204, 2008.

[14] D. Vuckovic, "Improving metabolome coverage and data quality: Advancing metabolomics and lipidomics for biomarker discovery," *Chemical Communications*, vol. 54, pp. 6728–6749, 2018.

[15] Á. González-Domínguez, E. Durán-Guerrero, Á. Fernández-Recamales, A. Lechuga-Sancho, A. Sayago, M. Schwarz, C. Segundo, and R. González-Domínguez, "An overview on the importance of combining complementary analytical platforms in metabolomic research," *Current Topics in Medicinal Chemistry*, vol. 17, pp. 1–7, 2017.

[16] H. Issaq, Q. Van, T. Waybright, G. Muschik, and T. Veenstra, "Analytical and statistical approaches to metabolomics research," *Journal of Separation Science*, vol. 32, pp. 2183–2199, 2009.

[17] D. Yoon, M. Lee, S. Kim, and S. Kim, "Applications of NMR spectroscopy based metabolomics: A review," *Journal of the Korean Magnetic Resonance Society*, vol. 17, pp. 1–10, 2013.

[18] S. Putri, S. Yamamoto, H. Tsugawa, and E. Fukusaki, "Current metabolomics: Technological advances," *Journal of Bioscience and Bioengineering*, vol. 116, pp. 9–16, 2013.

[19] E. M. Lenz and I. D. Wilson, "Analytical strategies in metabonomics," *Journal of Proteome Research*, vol. 6, pp. 443–458, 2007.

[20] H. J. Issaq, E. Abbott, and T. D. Veenstra, "Utility of separation science in metabolomic studies," *Journal of Separation Science*, vol. 31, pp. 1936–1947, 2008.

[21] K. Liland, "Multivariate methods in metabolomics - from pre-processing to dimension reduction and statistical analysis," *Trends in Analytical Chemistry*, vol. 30, pp. 827–841, 2011.

[22] S. Ren, A. Hinzman, E. Kang, R. Szczesniak, and L. Lu, "Computational and statistical analysis of metabolomics data," *Metabolomics*, vol. 11, pp. 4192–1513, 2015.

[23] E. Saccenti, H. Hoefsloot, A. Smilde, J. Westerhuis, and M. Hendriks, "Reflections on univariate and multivariate analysis of metabolomics data," *Metabolomics*, vol. 10, pp. 361–374, 2014.

[24] D. Barupal, S. Fan, and O. Fiehn, "Integrating bioinformatics approaches for a comprehensive interpretation of metabolomics datasets," *Current Opinion in Biotechnology*, vol. 54, pp. 1–9, 2018.

[25] S. Barnes, H. Benton, K. Casazza, S. Cooper, X. Cui, X. Du, J. Engler, J. Kabarowski, S. Li, W. Pathmasiri, J. Prasain, M. Renfrow, and H. Tiwari, "Training in metabolomics research. ii. processing and statistical analysis of metabolomics data, metabolite identification, pathway analysis, applications of metabolomics and its future," *Journal of Mass Spectrometry*, vol. 51, pp. 535–548, 2016.

[26] S. Naz, M. Vallejo, A. García, and C. Barbas, "Method validation strategies involved in non-targeted metabolomics," *Journal of Chromatography A*, vol. 1353, pp. 99–105, 2014.

[27] F. Rouessac and A. Rouessac, *Chemical Analysis: Modern Instrumentation, Methods and Techniques*. Chichester: John Wiley & Sons Ltd, 2000.

[28] S. Higson, *Analytical Chemistry*. New York: Oxford University Press Inc., 2003.

[29] K. Rubinson and J. Rubinson, *Contemporary Instrumental Analysis*. New Jersey: Pentice-Hall Inc., 2000.

[30] V. Meyer, *Practical High-Performance Liquid Chromatography*, 5th ed. Chichester: John Wiley & Sons Ltd, 2010.

[31] B. Buszewski and S. Noga, "Hydrophilic interaction liquid chromatography (HILIC) - a powerful separation technique," *Analytical and Bioanalytical Chemistry*, vol. 402, pp. 231–247, 2012.

[32] *HILIC separations technical guide*, Thermo Scientific, 2014.

[33] E. de Hoffman and V. Stroobant, *Mass Spectrometry Principles and Applications*, 3rd ed. Chichester: Wiley, 2007.

[34] C. Ho, M. Chan, R. Cheung, L. Law, L. Lit, K. Ng, M. Suen, and H. Tai, "Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications," *The Clinical Biochemist Reviews*, vol. 24, pp. 3–12, 2003.

[35] Shimadzu. (2016). Introduction to LC-MS: Shimadzu, [Online]. Available: `http://www.shimadzu.com/an/hplc/support/lib/lctalk/47/47intro.html` (visited on 07/2019).

[36] T. R. Covey, B. A. Thomson, and B. B. Schneider, "Atmospheric pressure ion sources," *Mass Spectrometry Reviews*, vol. 28, pp. 870–897, 2009.

[37] J. Greaves and J. Roboz, *Mass Spectrometry for the novice*. Florida: CRC Press, 2014.

[38] H. Awad, M. M. Khamis, and A. El-Aneed, "Mass spectrometry, review of the basics: Ionization," *Applied Spectroscopy Reviews*, vol. 50, pp. 158–175, 2015.

[39] S. J. Gaskell, "Electrospray: Principles and practice," *Journal of Mass Spectrometry*, vol. 32, pp. 677–688, 1997.

[40] S. Banerjee and S. Mazumdar, "Electrospray Ionization Mass Spectrometry: A technique to access the information beyond the molecular weight of the analyte," *International Journal of Analytical Chemistry*, vol. 2012, pp. 1–29, 2012.

[41] T. C. Rohner, N. Lion, and H. H. Girault, "Electrochemical and theoretical aspects of electrospray ionisation," *Physical Chemistry Chemical Physics*, vol. 6, pp. 3056–3068, 2004.

[42] A. El-Aneed, A. Cohen, and J. Banoub, "Mass spectrometry, review of the basics: Electrospray, MALDI, and commonly used mass analyzers," *Applied Spectroscopy Reviews*, vol. 44, pp. 210–230, 2009.

[43] (2014). Mass Spectrometry Facility:Quadrupole. University of Bristol School of Chemistry, [Online]. Available: `http://www.chm.bris.ac.uk/ms/quadrupole.xhtml` (visited on 07/2019).

[44] I. V. Chernushevich, A. V. Loboda, and B. A. Thomson, "An introduction to quadrupole-time-of-flight mass spectrometry," *Journal of Mass Spectrometry*, vol. 36, pp. 849–865, 2001.

[45] G. A. Theodoridis, H. G. Gika, E. J. Want, and I. D. Wilson, "Liquid chromatography-mass spectrometry based global metabolite profiling: A review," *Analytica Chimica Acta*, vol. 711, pp. 7–16, 2012.

[46] M. Katajamaa and M. Orešič, "Data processing for mass spectrometry-based metabolomics," *Journal of Chromatography A*, vol. 1158, pp. 318–328, 2007.

[47] R. Tautenhahn, G. J. Patti, D. Rinehart, and G. Siuzdak, "XCMS Online: A web-based platform to process untargeted metabolomic data," *Analytical Chemistry*, vol. 84, pp. 5035–5039, 2012.

[48] H. P. Benton, J. Ivanisevic, N. G. Mahieu, M. E. Kurczy, C. H. Johnson, L. Franco, D. Rinehart, E. Valentine, H. Gowda, B. K. Ubhi, R. Tautenhahn, A. Gieschen, M. W. Fields, G. J. Patti, and G. Siuzdak, "Autonomous Metabolomics for Rapid Metabolite Identification in Global Profiling," *Analytical Chemistry*, vol. 87, pp. 884–891, 2015.

[49] C. Smith, E. Want, G. O'Maille, R. Abagyan, and G. Siuzdak, "XCMS: Processing Mass Spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification," *Analytical Chemistry*, vol. 78, pp. 779–787, 2006.

[50] J. Miller and J. Miller, *Statistics and chemometrics for analytical chemistry*, 6th ed. Essex: Pearson, 2010.

[51] H. Abdi and L. Williams, "Principal component analysis," *Computational Statistics*, vol. 2, pp. 433–459, 2010.

[52] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, pp. 37–52, 1987.

[53] C. Adam, *Essential mathematics and statistics for forensic science*. Chichester: Wiley-Blackwell, 2010.

[54] S. Ellison, V. Barwick, and T. Farrant, *Practical statistics for the forensic scientist*, 2nd ed. Cambridge: RSC, 2009.

[55] D. Kealey and P. Haines, *Analytical Chemistry*. Oxford: BIOS, 2002.

[56] S. Jan and G. Shieh, "Sample size determinations for welch's test in one-way heteroscedastic ANOVA," *British Journal of Mathematical and Statistical Psychology*, vol. 67, pp. 72–93, 2013.

[57] R. Ennos, *Statistical and Data Handling Skills in Biology*, 3rd ed. Pearson Education UK, 2012.

[58] S. Lee and D. Lee, "What is the proper way to apply the multiple comparison test?" *Korean Journal of Anaesthesiology*, vol. 71, pp. 353–360, 2018.

[59] H.-Y. Kim, "Analysis of variance (ANOVA) comparing means of more than two groups," *Restorative Dentistry & Endodontics*, vol. 39, pp. 74–77, 2014.

[60] M. McHugh, "Multiple comparison analysis testing in (ANOVA)," *Biochemia Medica*, vol. 21, pp. 203–209, 2011.

3. Experimental

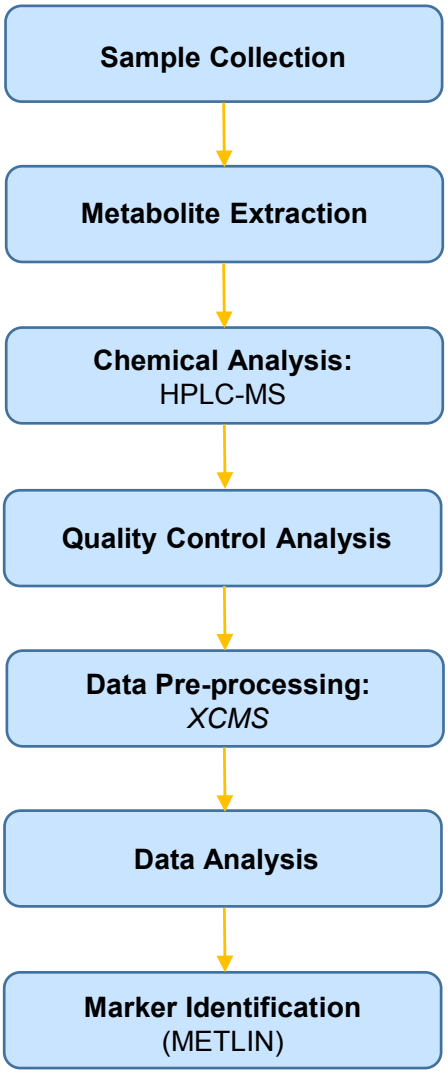
3.1 Aims

- *To display the experimental workflow that was carried out throughout this research.*
- *To show the development of chromatographic methods for the analysis of metabolite extracts.*
- *To outline the statistical analysis workflow that was used throughout this research.*

3.2 Experimental Design

Throughout the various studies presented in this research, every effort has been made to reduce the number of factors that could affect the experiment. The age and breed of the laying hens was considered, as was the diet fed to the birds, and the stocking density and number of birds per cage. All eggs used in the same experiment/data set were laid on the same day, and extractions were also carried out on the same day to ensure that there was no difference in the age of the eggs (except for the study investigating the effects of egg age on metabolite profile). All samples were randomised prior to analysis, to ensure that any instrumental drift that may have occurred did not bias the results.

Other than the initial experimental set-up and sample collection, the design of all of the studies presented in this research followed the same workflow, described in the following diagram.



In all studies throughout this research, each sample set consisted of six extracts. The metabolite extraction was carried out identically in all studies, followed by the chemical analysis of sample extracts using HPLC-Q-ToF-MS, using the same method in all non-targeted studies. Quality control analysis was carried out, and the data was then pre-processed using *XCMS Online*, which produced a feature table that was then transferred into *Microsoft Excel* for further analysis. Following data analysis, identification of potential biomarkers took place.

3.3 Materials

Methanol (HPLC, isocratic grade) and dichloromethane (stabilised with 0.002% 2-methyl-2-butene) were purchased from VWR (Radnor, PA). Ultra-pure water (at 18.2 MΩ/cm) was purified in-house using a Milli-Q system from Elga (High Wycombe, U.K.). Formic acid (90%, laboratory-reagent grade) and ammonium acetate were purchased from Fisher Scientific (Loughborough, U.K.). ESI-L low-concentration tuning mix and API-ToF reference-mass solution were purchased from Sigma-Aldrich (St. Louis, MO). Analysis was carried out using a Thermo Scientific Accucore RP-MS (100 X 2.1 mm, 2.6 µm particle size) on an Agilent Technologies 1260 Infinity Binary HPLC system, coupled to an Agilent Technologies 6530 Accurate-Mass Quadrupole-Time-of-Flight mass spectrometer.

3.4 Extraction Method Development

Aqueous and organic solvent liquid extractions were carried out to extract metabolites from both the albumen and yolk of eggs, based on the metabolite extraction method described in Thompson et al. [1]. This approach has been shown to efficiently and reproducibly extract metabolites from biological matrices [2]. The homogenisation and sonication steps described in this method were excluded, as egg is already homogeneous and is a single cell which does not require lysis through sonication. The aqueous and organic extraction stages were not carried out sequentially as there was enough sample to carry out separate extractions, and it allowed the organic solvent mixture to extract some more polar metabolites that may otherwise have been previously extracted by the aqueous solvent mixture.

Egg yolk and albumen were separated using a stainless steel egg separator, which was washed with ultra-pure water and methanol between uses. Approximately 50 mg yolk

and 50 mg albumen were weighed out into separate 1.5 mL microcentrifuge tubes, ready for addition of the aqueous and organic extraction solvents.

3.4.1 Aqueous Metabolite Extraction

An aqueous extraction solvent mixture (methanol:water, 1:1) was added to the yolk and albumen samples, 1 mL of solvent per 50 mg of sample, or part thereof, which were then briefly vortexed to ensure thorough extraction. The vortexed samples were then centrifuged at 16,100 rcf for 20 minutes, before 0.75 mL of supernatant was removed and transferred into fresh microcentrifuge tubes. The sample extracts were then stored at -80 ℃, prior to analysis.

3.4.2 Organic Metabolite Extraction

An organic extraction solvent mixture (dichloromethane:methanol, 3:1) was added to the yolk and albumen samples, 1 mL of solvent per 50 mg of sample, or part thereof, which were then briefly vortexed to ensure thorough extraction. The samples were then centrifuged at 16,100 rcf for 20 minutes, before 0.75 mL of supernatant was removed and transferred into fresh microcentrifuge tubes. The supernatant was allowed to evaporate under ambient conditions, and the dried sample extracts were then stored at -80 ℃. The extracts were re-dissolved in an equal volume of methanol immediately prior to analysis and briefly vortexed to ensure thorough dissolution.

3.5 Chromatographic Method Development

Chromatographic methods were developed for all resulting sample extracts; albumen organic and aqueous extracts, and yolk organic and aqueous extracts, in both positive and negative ionisation mode. However, not all of these were successful, and so the only methods that were used for analysis in this research were for albumen and yolk organic extracts in positive ionisation mode.

All sample extracts were analysed using an Agilent 1260 Infinity Binary HPLC system and a Thermo Scientific Accucore RP-MS column (100 x 2.1 mm, 2.6 µm particle size) kept at a temperature of 40 ℃. The injection volume was 5 µl and the flow rate was 0.3 mL/min. The mobile phase consisted of solvents A, 0.1% formic acid, and B, methanol with 0.1% formic acid. Methanol was chosen as the organic mobile phase as the aqueous samples were extracted using a methanol and water extraction solvent mix, and the organic samples were re-dissolved in methanol following evaporation of the extraction solvent. Formic acid was added to both mobile phase solvents to promote positive ionisation of molecules in the MS analysis. During the method development stage, 5mM ammonium acetate was added to both solvents A and B to prevent build up of sample on the column. The initial chromatographic method for all extracts was a 90 minute gradient from 95% solvent A (0.1% formic acid), 5% solvent B (methanol, 0.1% formic acid), to 100% solvent B, which was then held isocratically for 30 minutes, resulting in a two hour total run time. This initial method was then adapted and developed for each sample extract, to result in a shorter analysis time, with compounds eluting throughout the analysis. A 5 minute post-run time was added to the final methods for all extracts, to allow the HPLC system to return to starting conditions and stabilise.

3.5.1 Albumen Organic Extract

An albumen organic extract was initially analysed in positive ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.1.

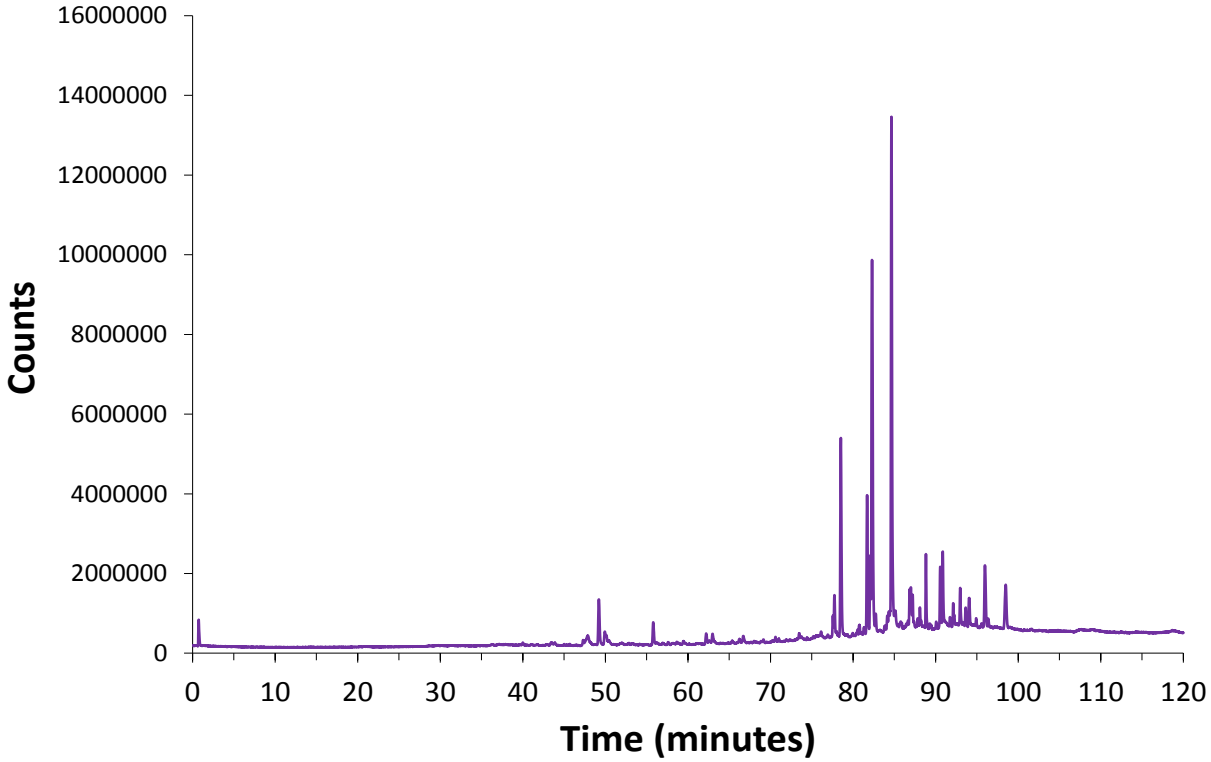


Figure 3.1: *TIC for albumen organic extract, analysed using initial 2 hour method, using positive ionisation*

This method was then adapted and developed, resulting in the final method:

Table 3.1: *Table showing the solvent gradient of the final method for the analysis of albumen organic extract*

Time (minutes)	Solvent A%	Solvent B%
0	20	80
5	16	84
20	0	100
30	0	100
31	20	80

The resulting chromatogram can be seen in Figure 3.2.

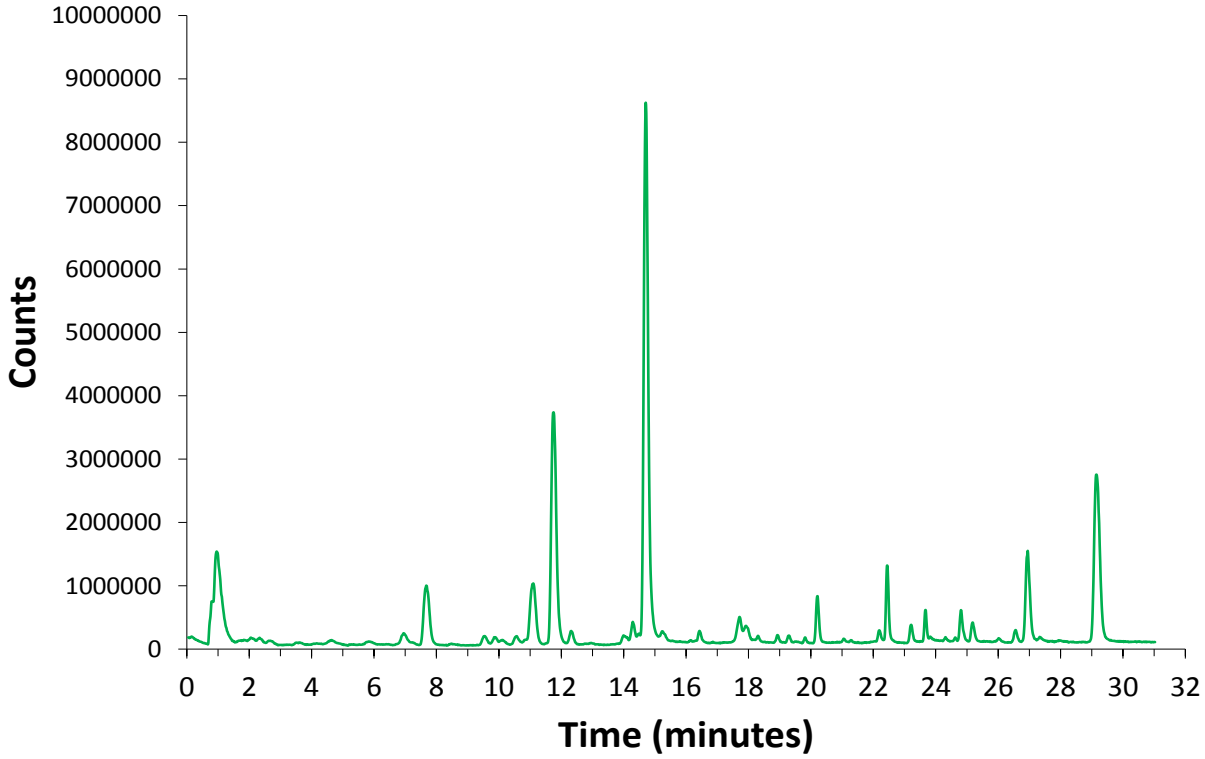


Figure 3.2: *TIC for albumen organic extract, analysed using final method, using positive ionisation*

This method was then used to analyse albumen organic extracts in positive ionisation mode throughout this research.

An albumen organic extract was then analysed in negative ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.3. From this, it appears that there are not many negatively ionised compounds detected in the albumen organic extract using this method.

However, an albumen organic extract was then analysed in negative ionisation mode using the final method developed in positive ionisation mode, resulting in the chromatogram that can be seen in Figure 3.4. Again, it appears that there are very few negatively ionised compounds detected in the albumen organic extract using this method, and so negative ionisation was not used to analyse any of these sample extracts throughout this research.

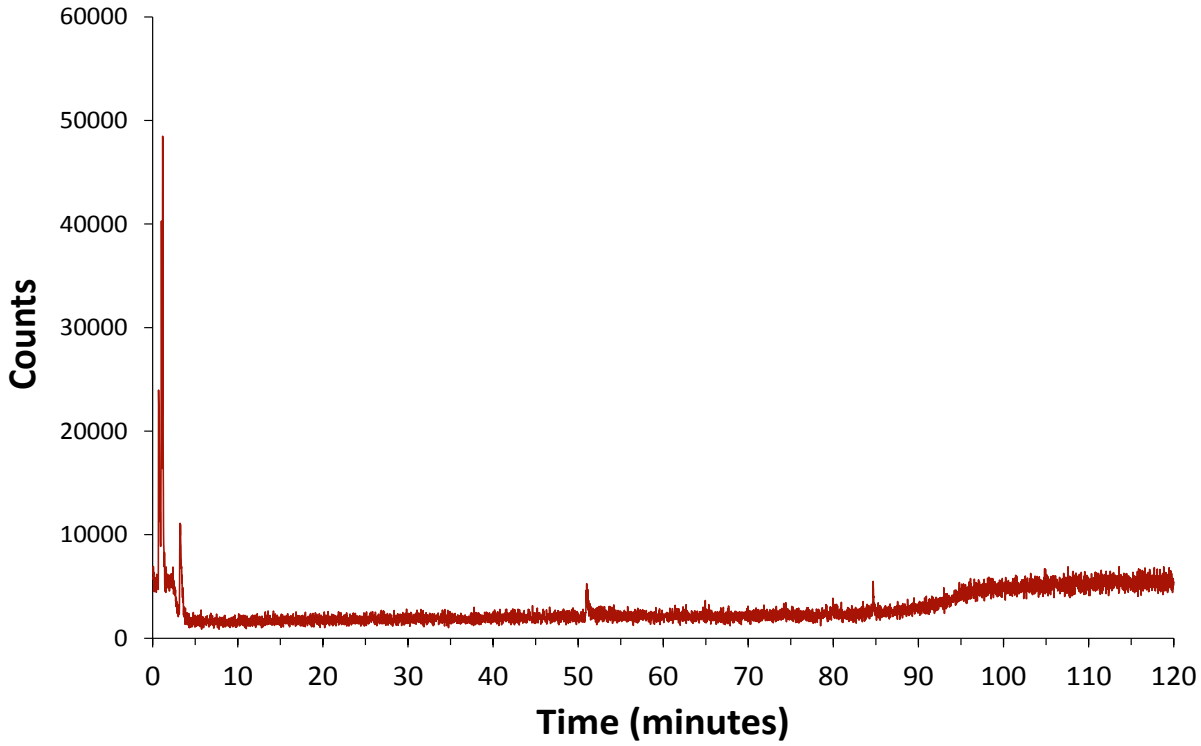


Figure 3.3: *TIC for albumen organic extract, analysed using initial 2 hour method, using negative ionisation*

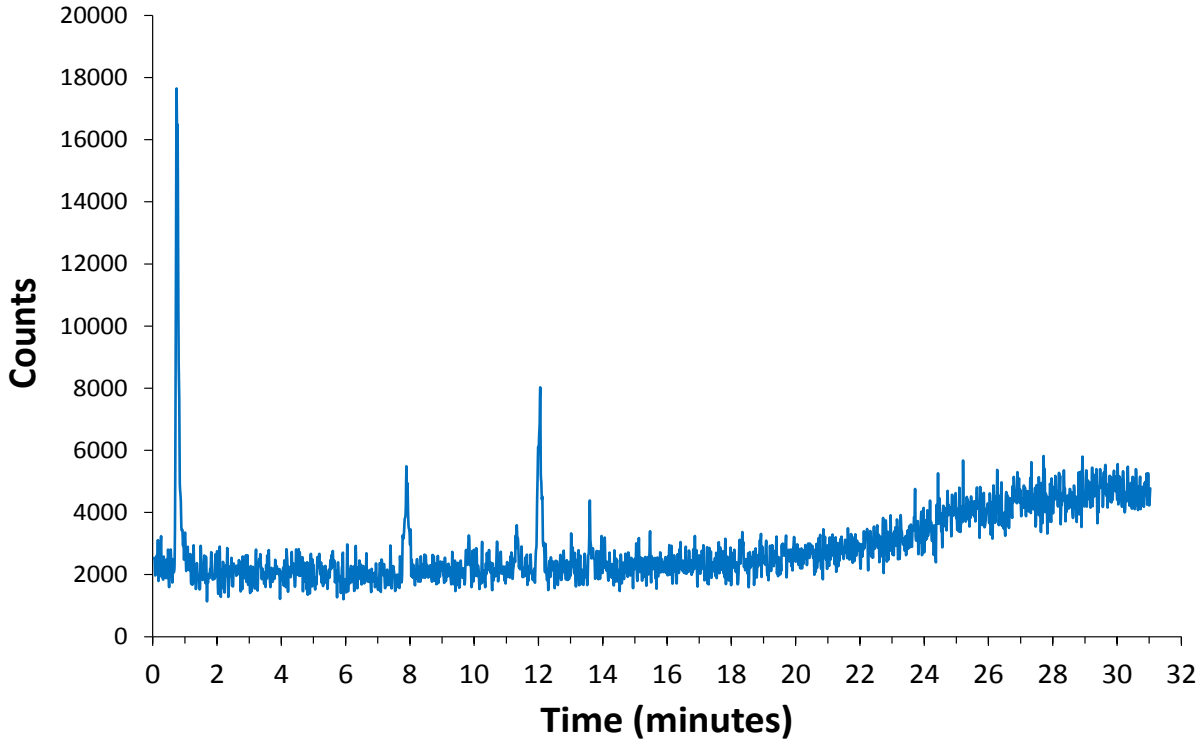


Figure 3.4: *TIC for albumen organic extract, analysed using final method, using negative ionisation*

3.5.2 Albumen Aqueous Extract

An albumen aqueous extract was initially analysed in positive ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.5.

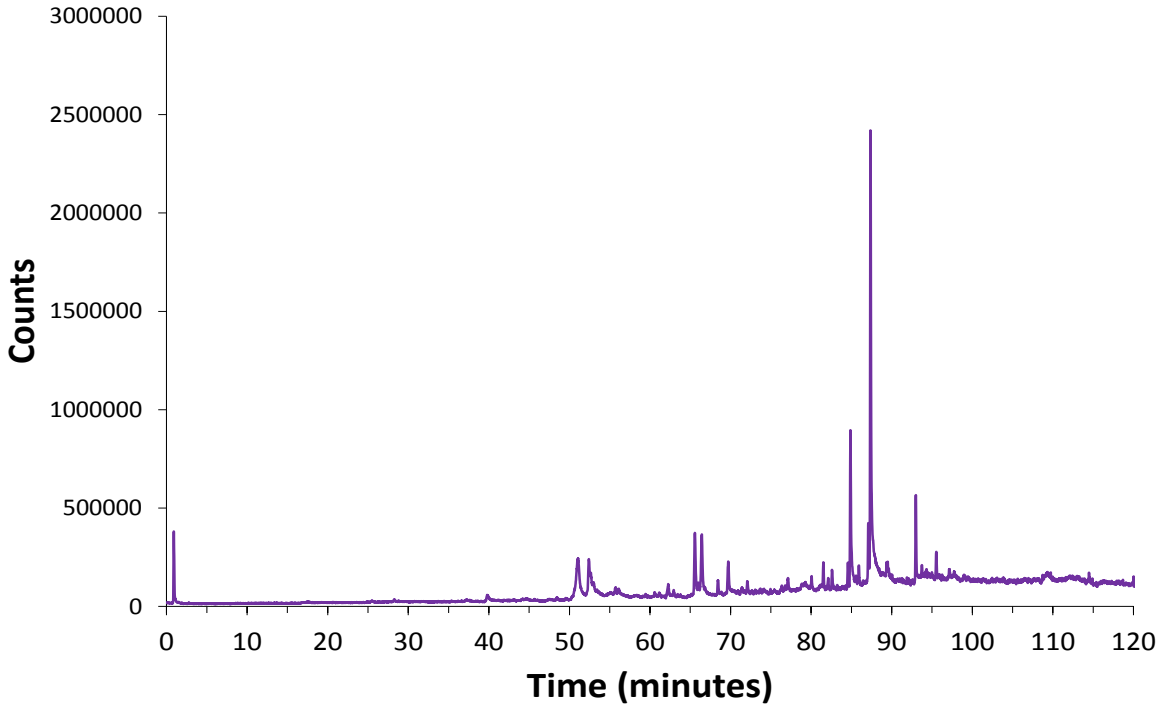


Figure 3.5: *TIC for albumen aqueous extract, analysed using initial 2 hour method, using positive ionisation*

This method was then adapted and developed, resulting in the final method:

Table 3.2: *Table showing the solvent gradient of the final method for the analysis of albumen aqueous extract*

Time (minutes)	Solvent A%	Solvent B%
0	35	65
17	15	85
25	0	100
40	0	100
45	35	65

The resulting chromatogram can be seen in Figure 3.6. This chromatogram looks to

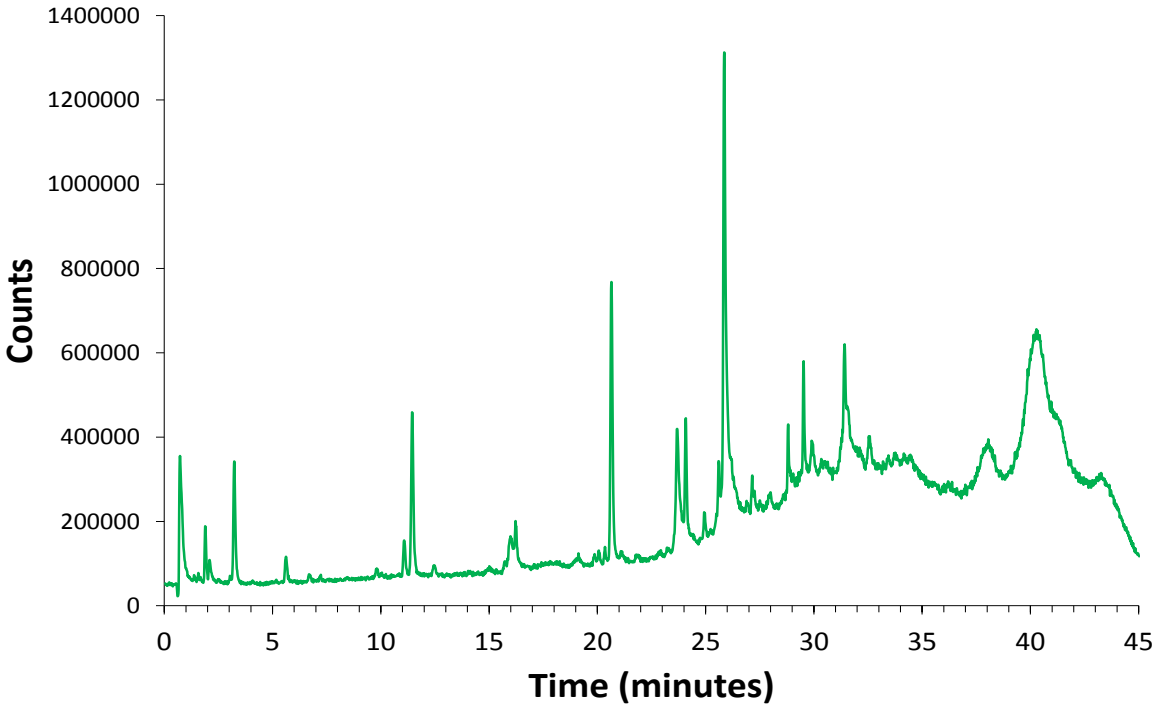


Figure 3.6: *TIC for albumen aqueous extract, analysed using final method, using positive ionisation*

be interesting, showing several peaks spread throughout the analysis time, indicating that many compounds are present in albumen aqueous extracts.

However, when repeated injections of these extracts were analysed over a longer period of time, the back pressure of the HPLC began to increase, eventually reaching the point where the HPLC stopped the analysis to prevent the pressure from increasing too high. Figure 3.7 shows the pressure curves for 10 injections of a pooled albumen aqueous extract sample.

It was thought that the cause for this increase in pressure could be due to the albumen extract precipitating out when the mobile phase reaches a certain ratio of methanol and water. A solvent mix of the two mobile phases at different ratios was then added to different aliquots of albumen aqueous extract to determine at which point the extract precipitated out. It was observed that with closer to 50:50 ratios of water:methanol the albumen extract began to precipitate out, which is likely to be the cause of this increase of pressure within the system during analysis. Various methods were attempted, including

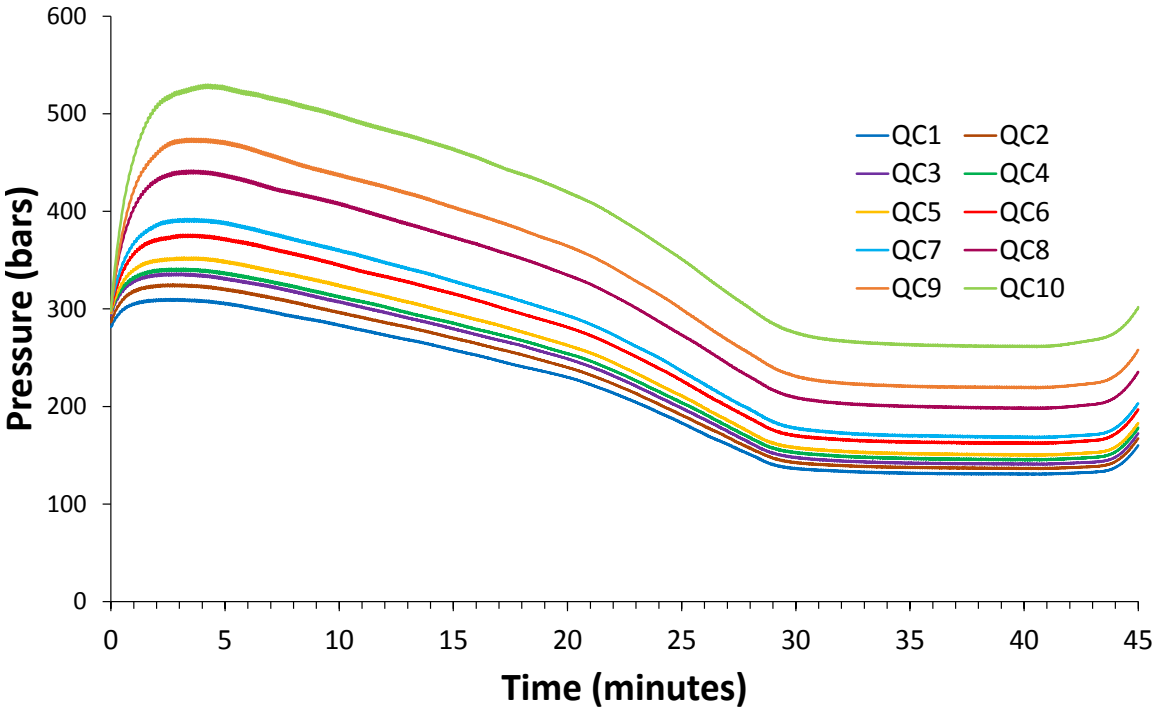


Figure 3.7: *Pressure curves showing increasing pressure in the HPLC system with repeated injections of albumen aqueous sample*

substituting methanol for acetonitrile, both with and without ammonium acetate in the mobile phase, however the same response was observed each time. It was concluded that albumen aqueous extracts are not suitable for HPLC analysis using the column and mobile phases that were used throughout this research.

An albumen aqueous extract was also analysed in negative ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.8. From this, it appears that there are not many negatively ionised compounds detected in the albumen aqueous extract using this method.

However, an albumen aqueous extract was then analysed in negative ionisation mode, using the final method developed in positive ionisation mode, resulting in the chromatogram that can be seen in Figure 3.9. Again, it appears that there are very few negatively ionised compounds detected in the albumen aqueous extract using this method, and so negative ionisation was not used to analyse any of these sample extracts throughout this research.

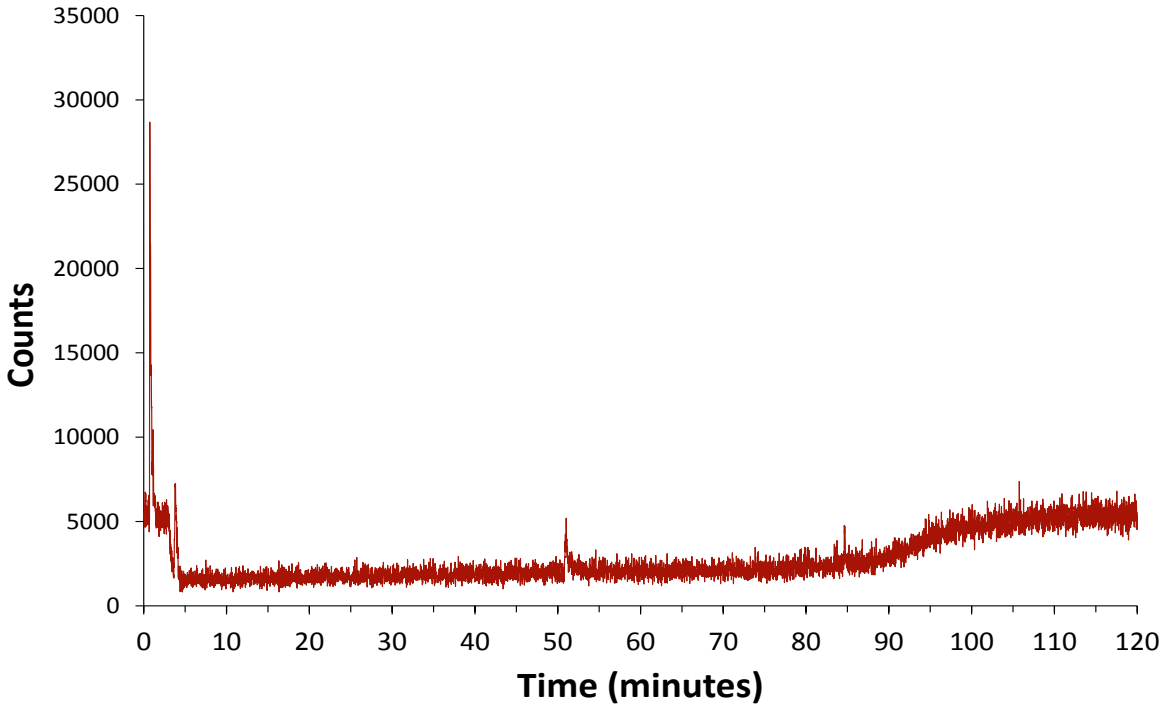


Figure 3.8: *TIC for albumen aqueous extract, analysed using initial 2 hour method, using negative ionisation*

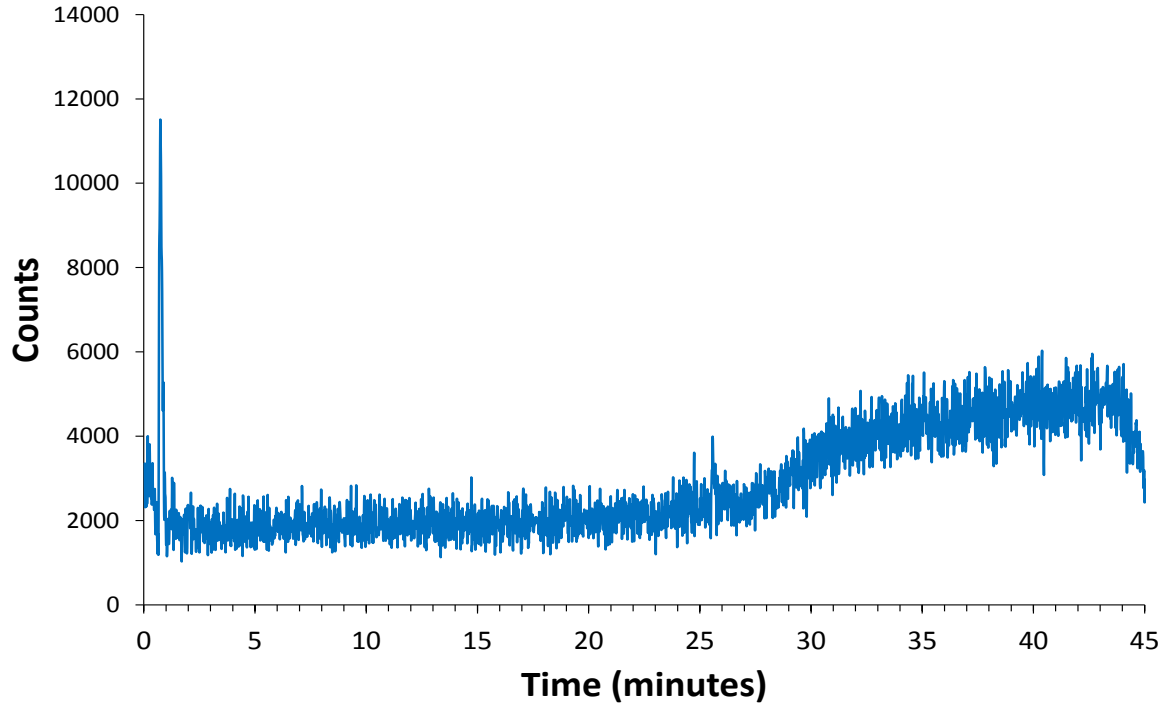


Figure 3.9: *TIC for albumen aqueous extract, analysed using final method, using negative ionisation*

3.5.3 Yolk Organic Extract

A yolk organic extract was initially analysed in positive ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.10.

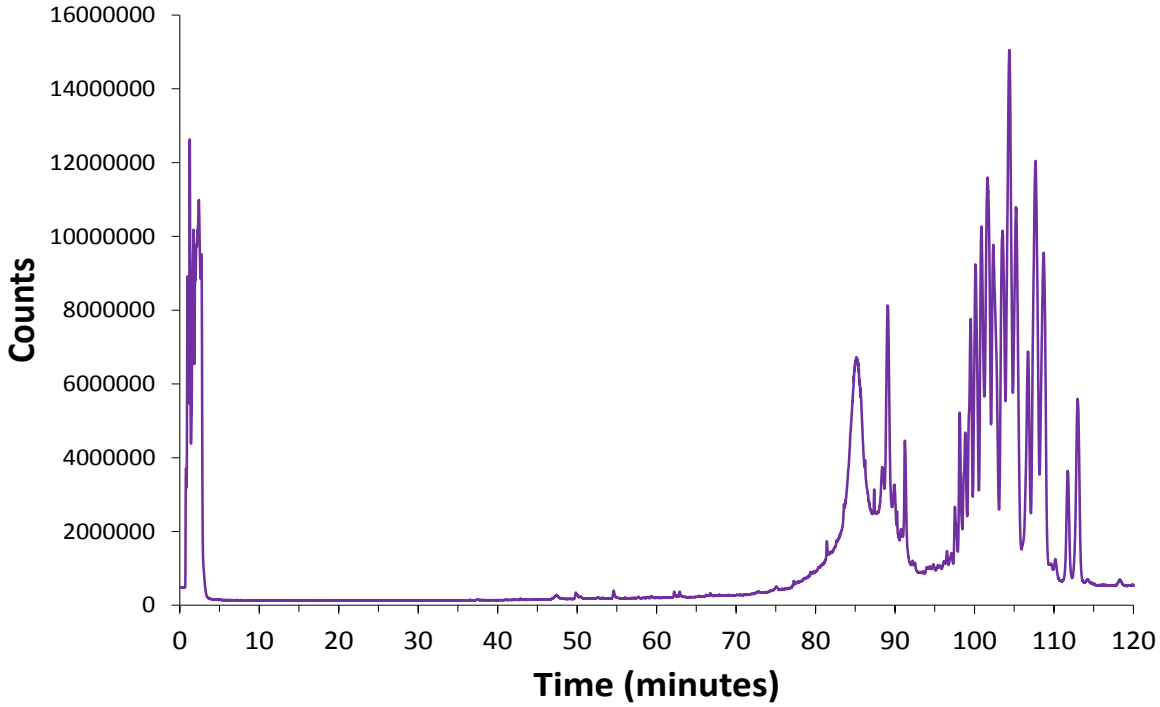


Figure 3.10: *TIC for yolk organic extract, analysed using initial 2 hour method, using positive ionisation*

This method was then adapted and developed, resulting in the final method:

Table 3.3: *Table showing the solvent gradient of the final method for the analysis of yolk organic extract*

Time (minutes)	Solvent A%	Solvent B%
0	25	75
20	19	81
21	10	90
31	10	90
61	0	100
81	0	100
85	25	75

The resulting chromatogram can be seen in Figure 3.11. This method was then used to analyse yolk organic extracts in positive ionisation mode throughout this research.

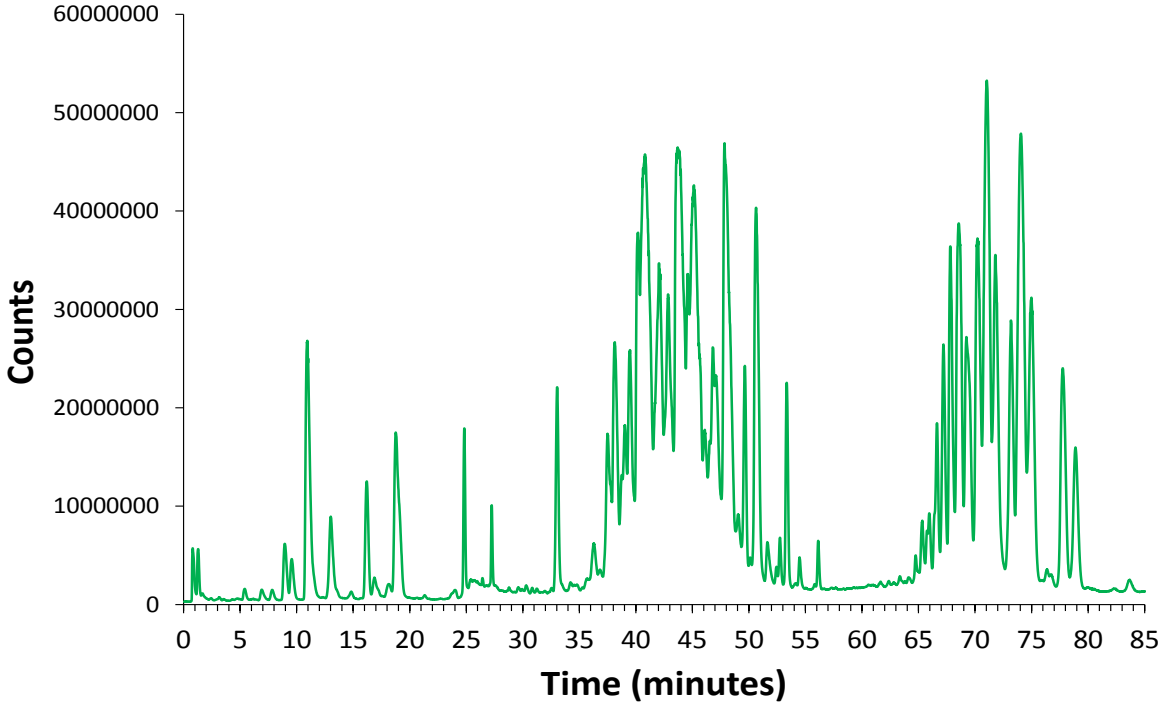


Figure 3.11: *TIC for yolk organic extract, analysed using final method, using positive ionisation*

A yolk organic extract was then analysed in negative ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.12. From this, it appears that there are not many negatively ionised compounds detected in the yolk organic extract using this method.

However, a yolk organic extract was then analysed in negative ionisation mode, using the final method developed in positive ionisation mode, resulting in the chromatogram that can be seen in Figure 3.13. This chromatogram contains several peaks, which indicates that there are several negatively ionised compounds detected in this sample extract that elute using this chromatographic method. However, the intensities of the peaks are not very high, and as much more information can be gained from analysing these samples using positive ionisation, negative ionisation mode was not used to analyse any of these sample extracts throughout this research.

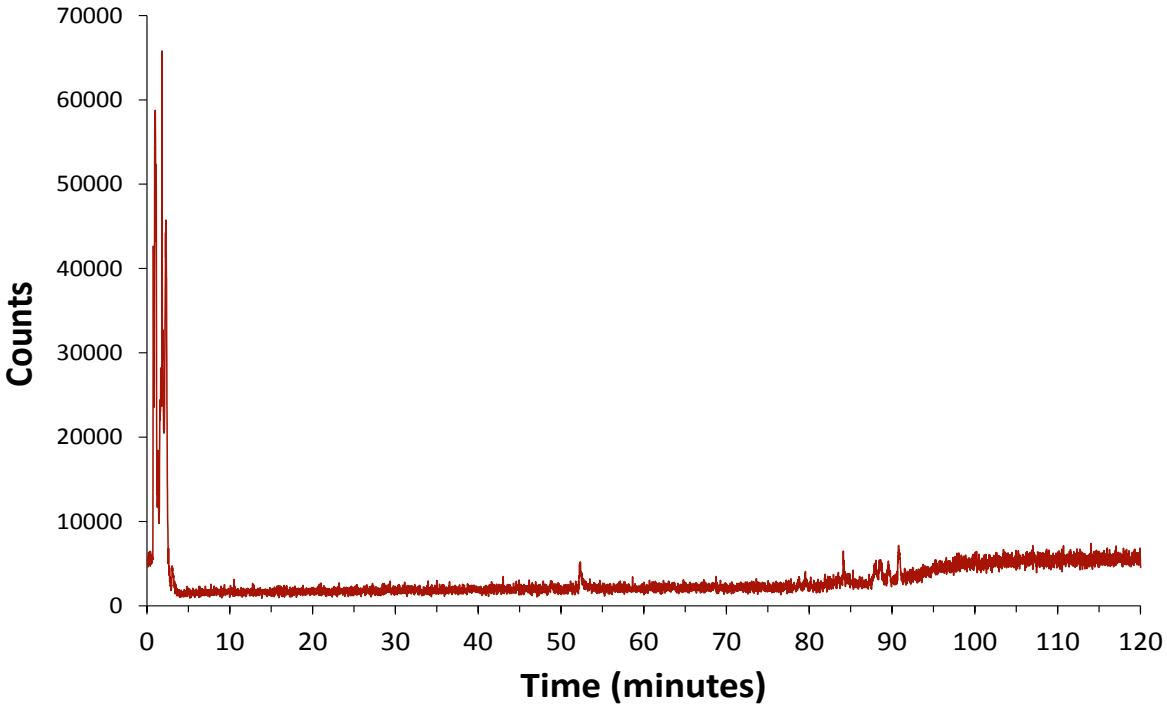


Figure 3.12: *TIC for yolk organic extract, analysed using initial 2 hour method, using negative ionisation*

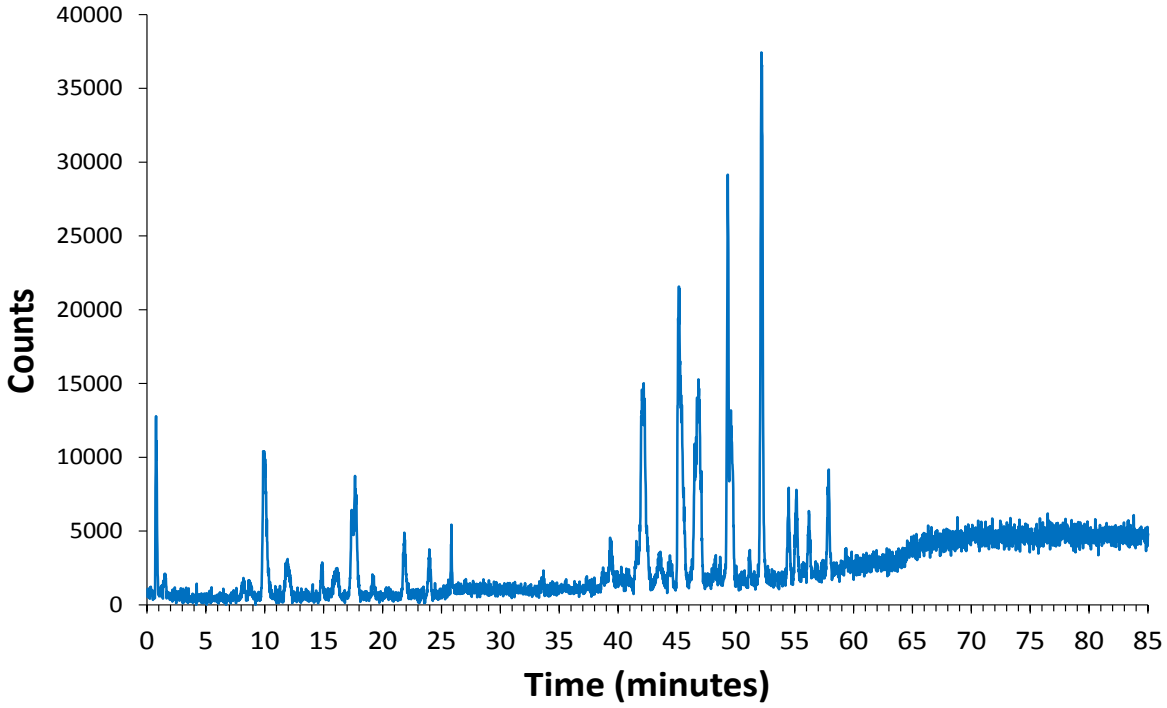


Figure 3.13: *TIC for yolk organic extract, analysed using final method, using negative ionisation*

3.5.4 Yolk Aqueous Extract

A yolk aqueous extract was initially analysed in positive ionisation mode using the two hour chromatographic method, resulting in the chromatogram that can be seen in Figure 3.14.

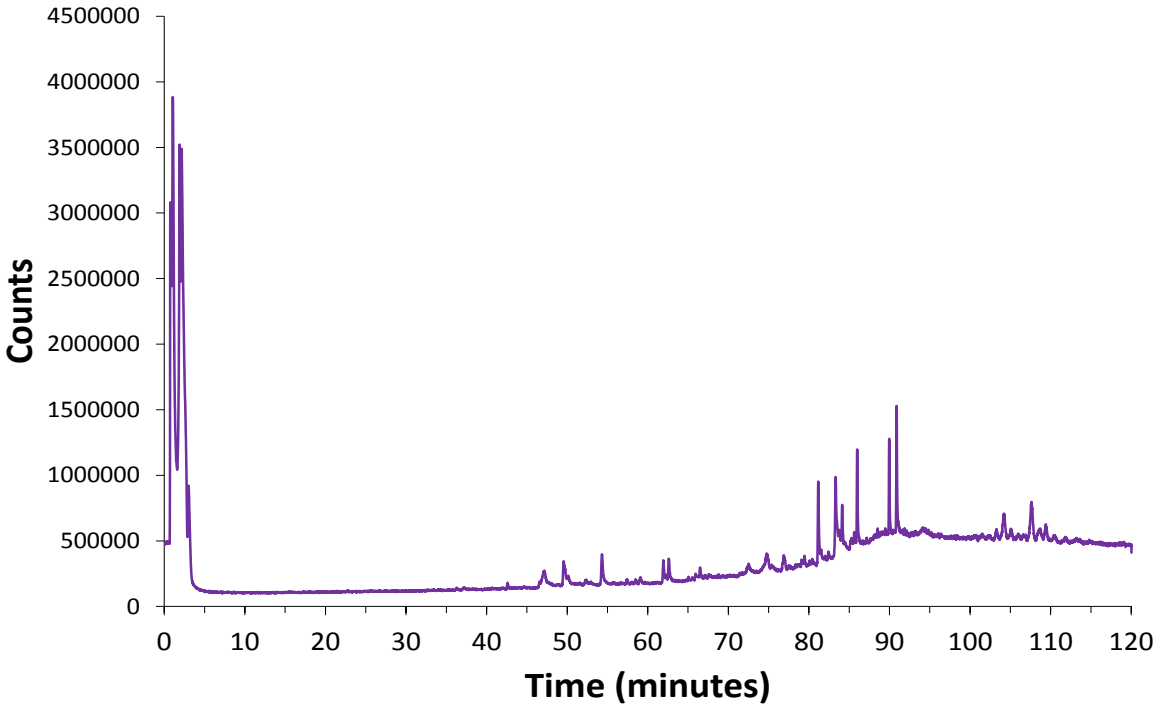
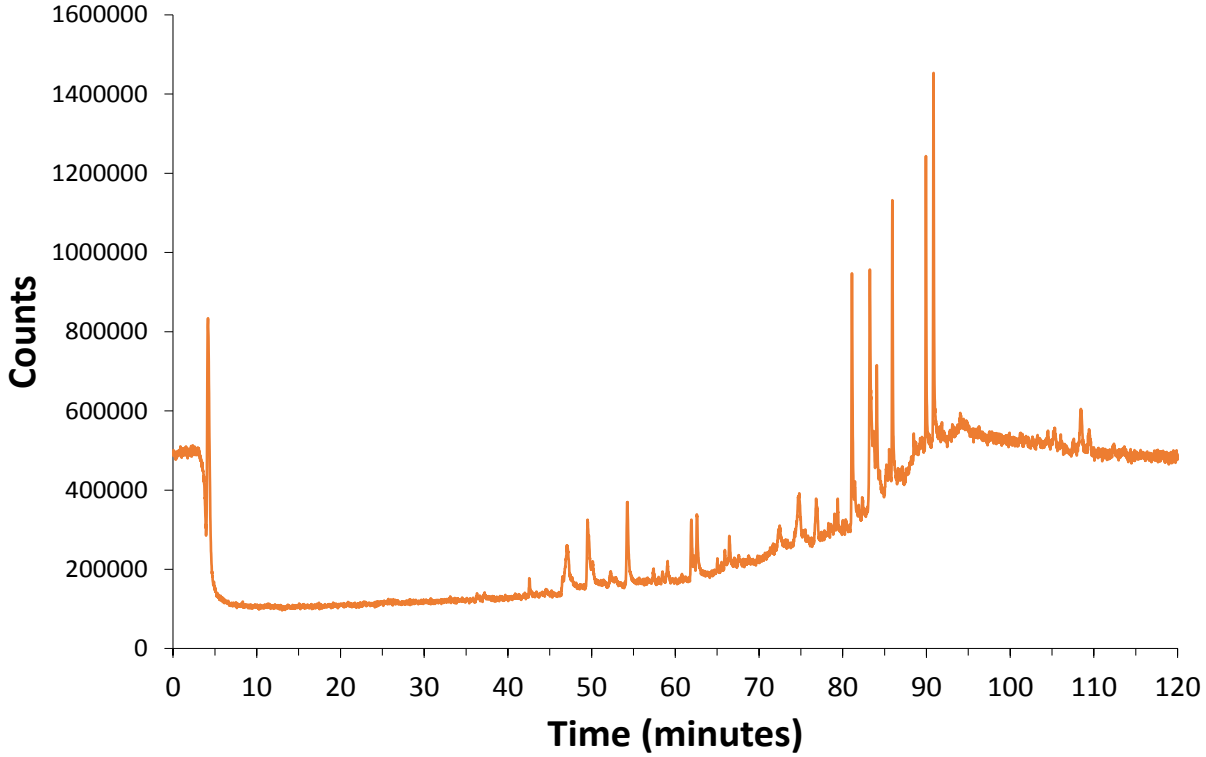
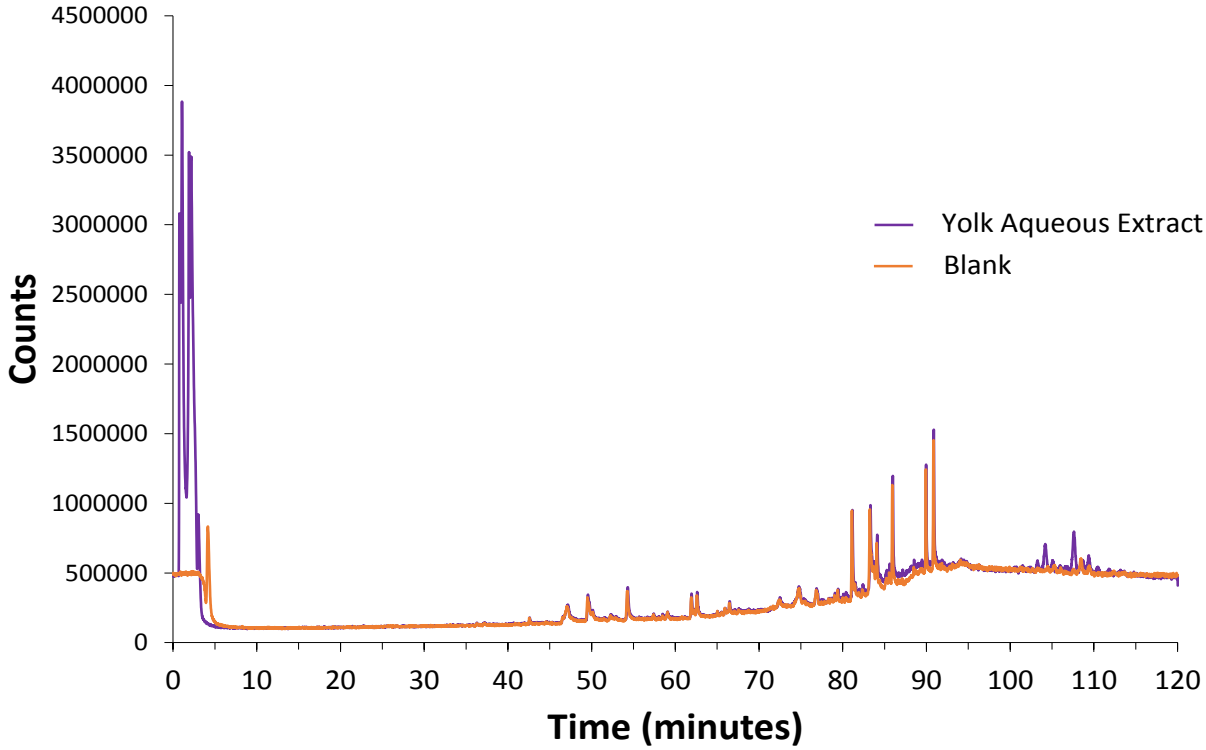


Figure 3.14: *TIC for yolk aqueous extract, analysed using initial 2 hour method, using positive ionisation*

Although this chromatogram appears to show lots of peaks, when compared to the chromatogram resulting from the analysis of a blank sample in Figure 3.15, it can be seen that most of these peaks are due to elution of compounds that have been retained on the column from previous analyses. Most of the peaks with higher intensities, resulting from the analysis of the yolk aqueous sample, are during the first five minutes of the analysis, with a few more towards the end of the analysis. This indicates that most of the more highly abundant compounds present in the yolk aqueous extract are very polar, and so would be more suited to HPLC analysis using a HILIC column.



(a) *TIC for a blank sample, analysed using initial 2 hour method, using positive ionisation*



(b) *Overlaid TICs for a blank sample and yolk aqueous extract, analysed using initial 2 hour method, using positive ionisation*

Figure 3.15

A yolk aqueous extract was then analysed using the same mobile phase solvents and other chromatographic parameters as previously, but using a Thermo Accucore HILIC column (100 mm x 2.1 mm, 2.6 µm particle size). The mobile phase composition started at 90% solvent B for the first 30 minutes, then decreased to 50% over 30 minutes, where it was held for a further 30 minutes, before returning to 90% in 5 minutes. The resulting chromatogram can be seen in Figure 3.16.

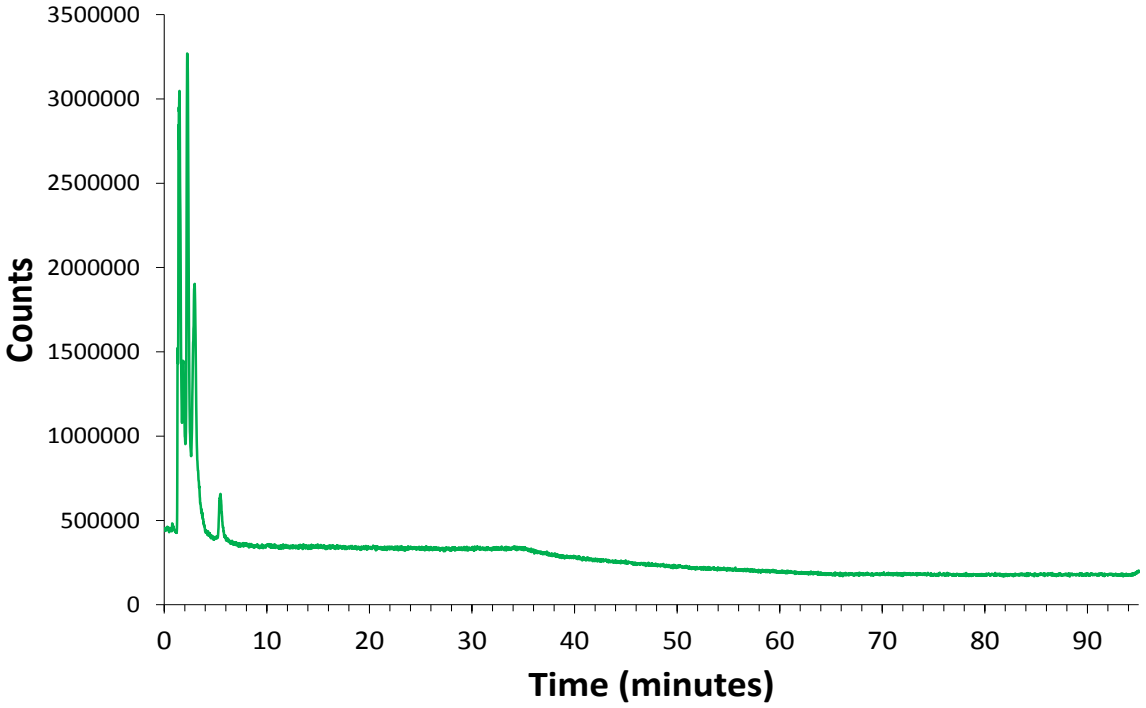


Figure 3.16: *TIC for yolk aqueous extract, analysed using initial HILIC method, using positive ionisation*

This method was then developed and adapted, and solvent B in the mobile phase was changed to acetonitrile with 0.1% formic acid and no ammonium acetate, resulting in the following method:

Table 3.4: *Table showing the solvent gradient of the method for the analysis of yolk aqueous extract using a HILIC column*

Time (minutes)	Solvent A%	Solvent B%
0	3	97
2	40	60
12	40	60
15	3	97

The resulting chromatogram can be seen in Figure 3.17.

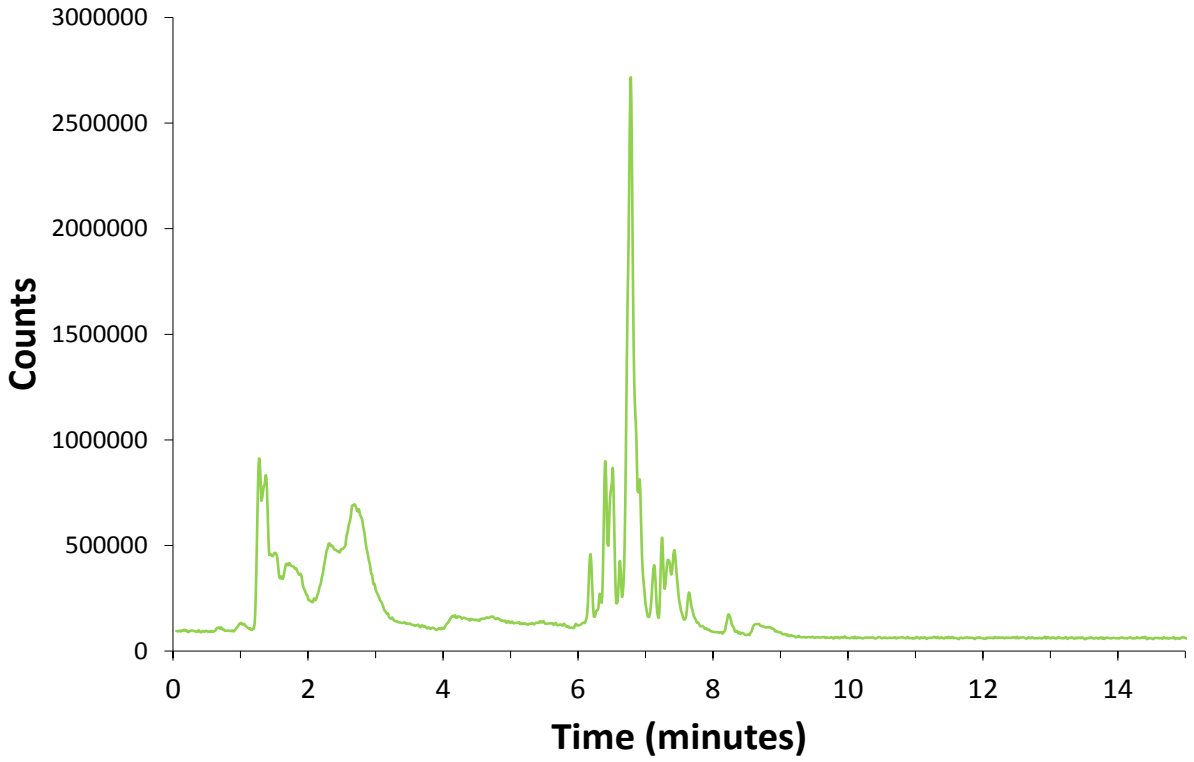
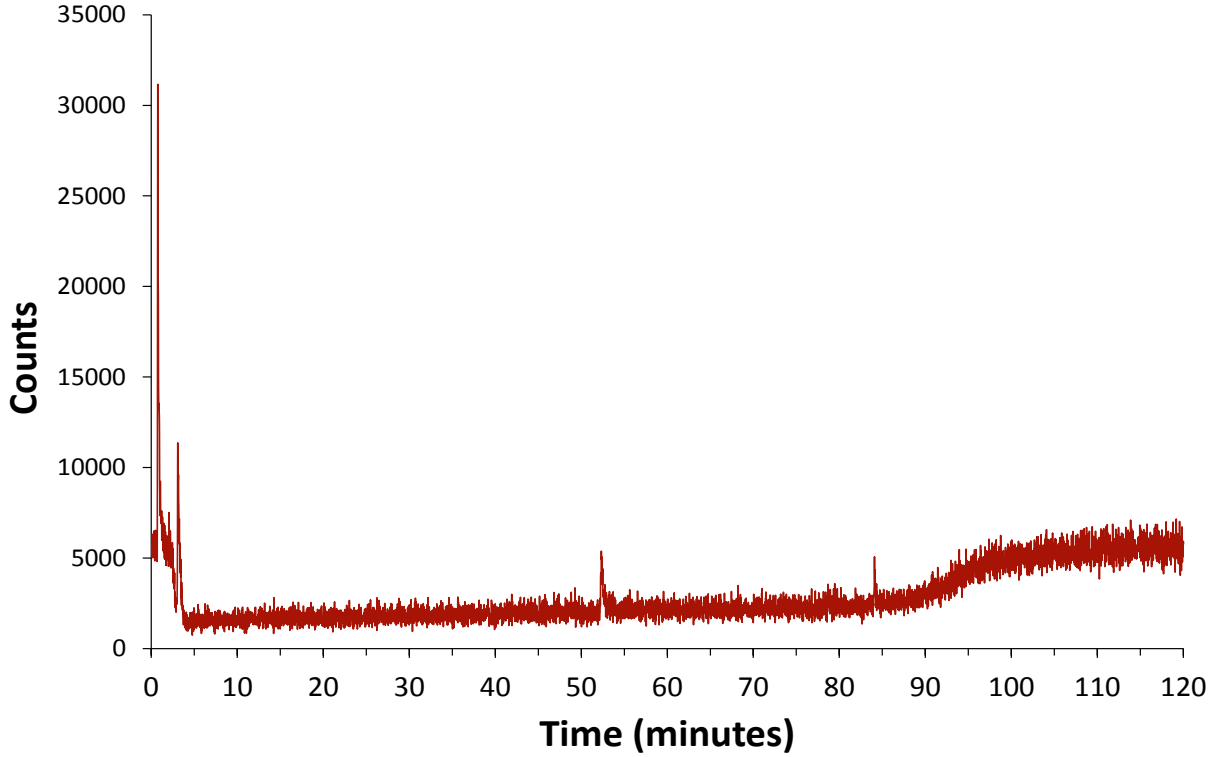


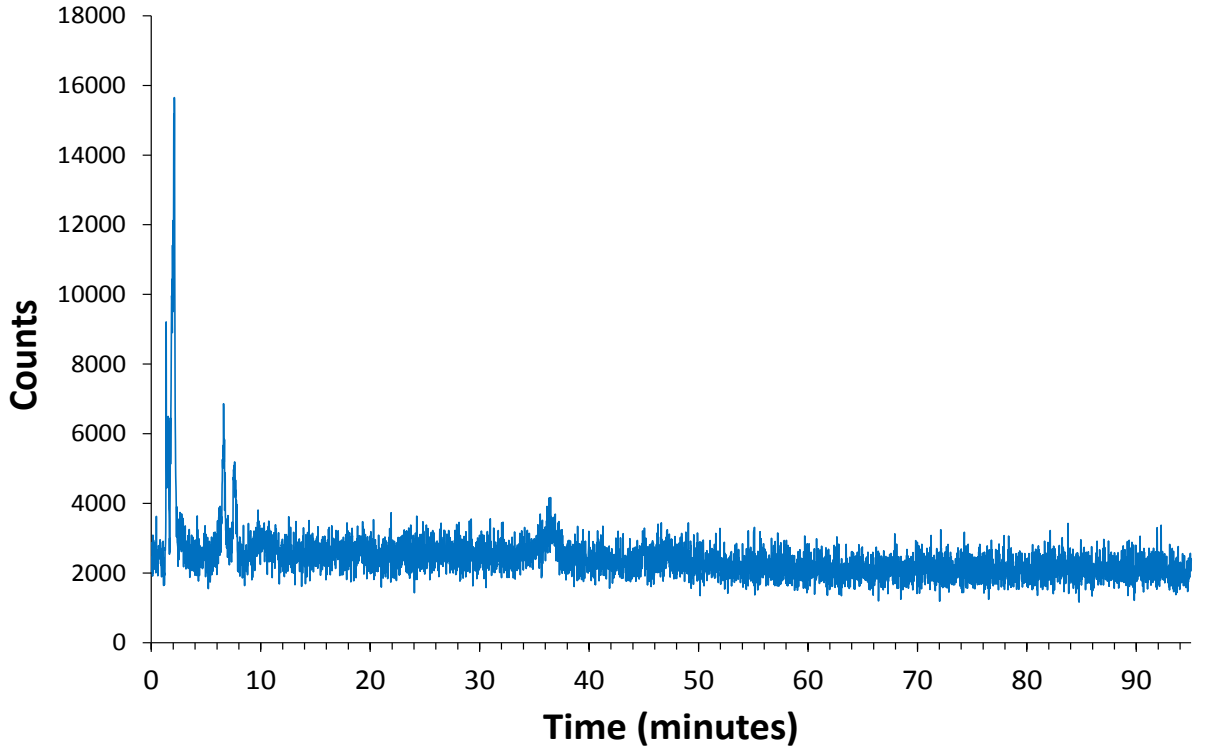
Figure 3.17: *TIC for yolk aqueous extract, analysed using final HILIC method, using positive ionisation*

This method was the product of much method development and the resulting chromatogram still does not display good peak shape or resolution, and so yolk aqueous extracts were not analysed in positive ionisation mode in any of the experiments throughout this research.

A yolk aqueous extract was then analysed in negative ionisation mode using the two hour chromatographic method, and the initial HILIC method, resulting in the chromatograms that can be seen in Figure 3.18.



(a) *TIC for yolk aqueous extract, analysed using initial 2 hour method, using negative ionisation*



(b) *TIC for yolk aqueous extract, analysed using initial HILIC method, using negative ionisation*

Figure 3.18

From these chromatograms, it appears that there are not many negatively ionised compounds detected in the yolk aqueous extract using these methods.

A yolk aqueous extract was then analysed in negative ionisation using the final HILIC method developed in positive ionisation, resulting in the chromatogram in Figure 3.19. Again, this chromatogram shows few peaks, indicating that there are few negatively ionised compounds detected in the yolk aqueous extract using this method. Therefore, analysis in negative ionisation mode was not carried out for yolk aqueous extract samples throughout this research.

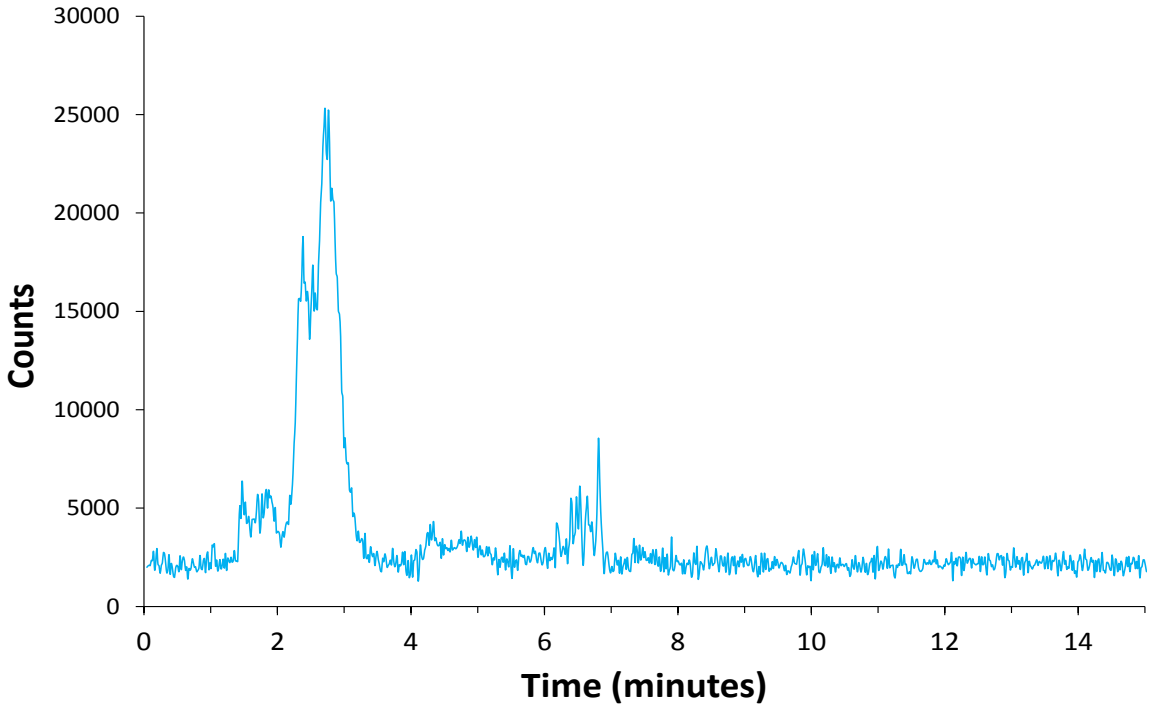


Figure 3.19: *TIC for yolk aqueous extract, analysed using final HILIC method, using negative ionisation*

Following method development, it was concluded that yolk organic extracts result in the best chromatogram, with many compounds separated, eluted, and detected throughout the chromatographic analysis, indicating a range of different compound types in the samples. Therefore, yolk organic extracts were analysed in all studies in this research. Albumen organic extracts also resulted in interesting chromatograms, and thus these samples were also analysed in some of the studies in this research.

3.6 Mass Spectrometry

All samples were anaylsed using an Agilent 6530 Accurate Mass Quadrupole-Time-of-Flight mass spectrometer in positive ionisation mode using an electrospray ion source. The parameters used were as follows:

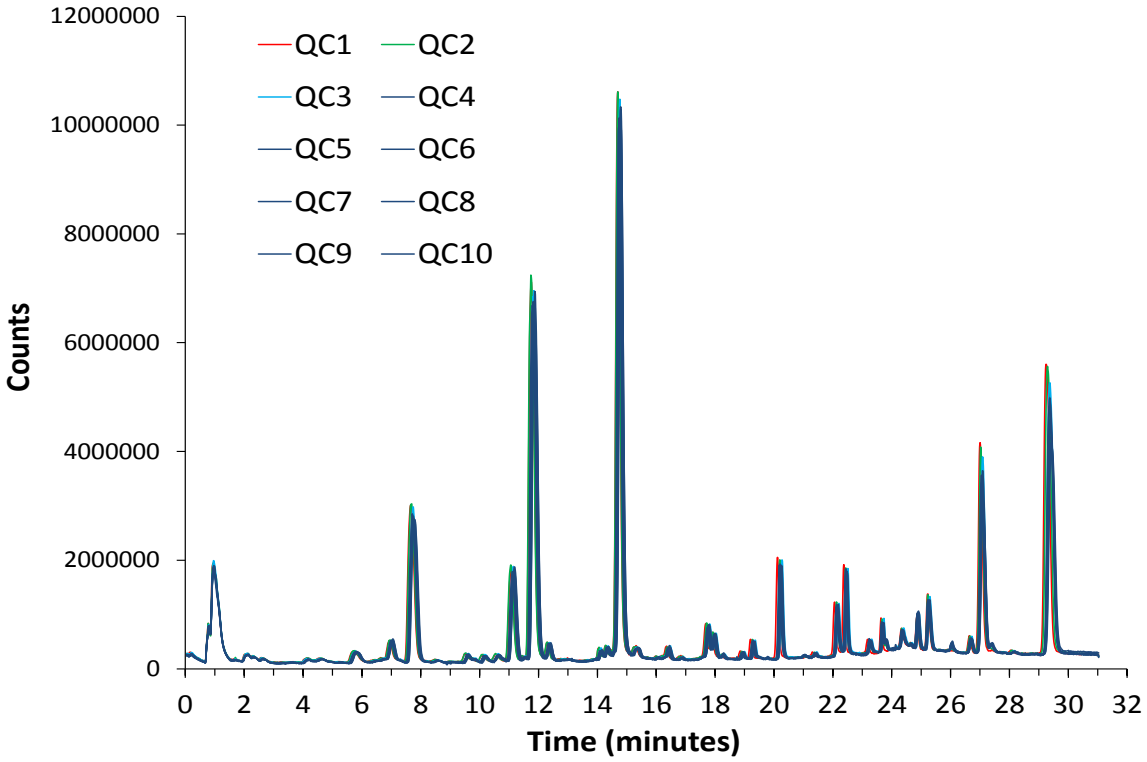
Table 3.5: *Table showing the parameters for the mass spectrometry analysis*

Drying gas temperature	Drying gas flow rate	Capillary voltage	Nebuliser pressure	Fragmentor voltage	Skimmer voltage
300 ℃	8 L/min	3500 V	35 psi	175 V	65 V

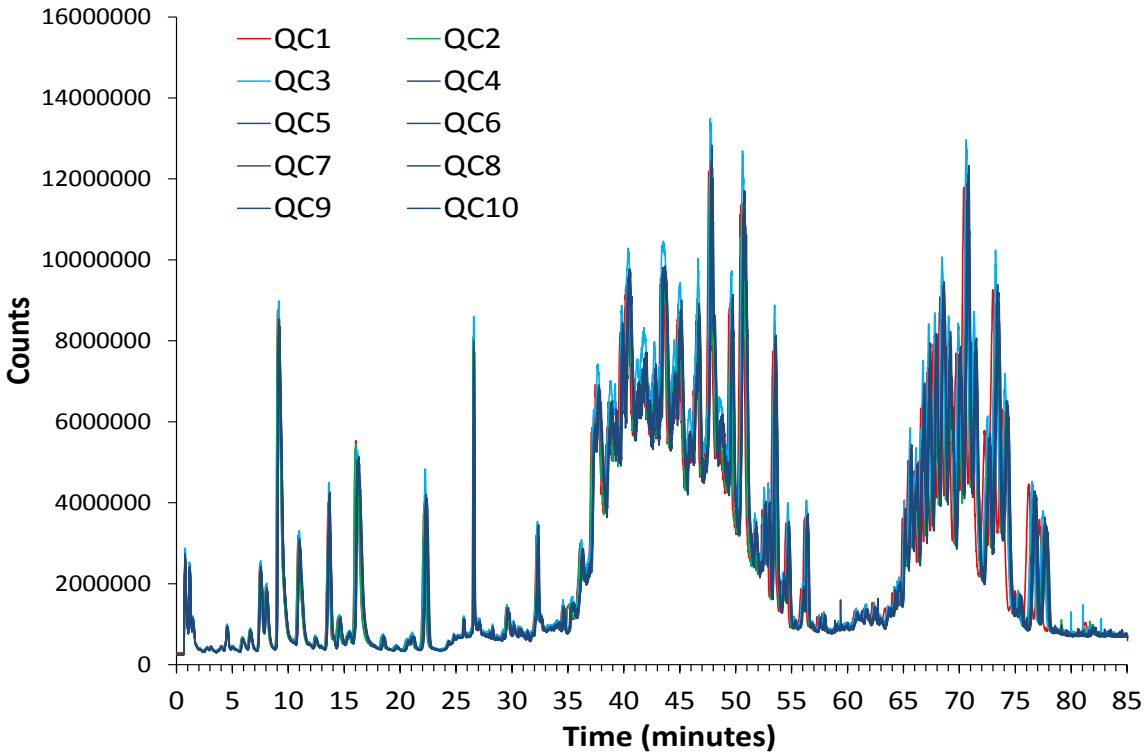
An ESI-L low concentration tuning mix was used to calibrate the instrument prior to analysis in order to ensure the mass accuracy, and an API-ToF reference mass solution, consisting of purine (m/z 121.0509) and hexakis (1H,1H,3H-tetrafluoropropoxy)phosphazine (m/z 922.0098) was used throughout the analytical run to maintain this mass accuracy. The mass range for all analyses was 100-1000 m/z.

3.7 Quality Control

Quality control is an extremely important aspect of metabonomic studies; if the analysis is not robust, then the results are not reliable. Quality control (QC) samples were produced for each study by pooling equal aliquots of all sample extracts together, in accordance with published guidelines [3]. These samples were then injected and analysed ten times prior to the analysis of actual samples in all studies, in order to condition the column ready for analysis. It has been shown in several studies that the first few injections of sample during an analytical run are not stable or reliable [4][5]. Therefore, there is a need to inject several QC samples, the exact number depending on the particular matrix and analytical conditions, before starting the actual analytical sequence, in order to condition the column ready for analysis [4][5]. Figure 3.20 shows overlaid TICs of the first ten QC sample injections for both albumen and yolk organic extracts.



(a) *Overlaid TICs for first ten injections of albumen organic extract QC sample*



(b) *Overlaid TICs for first ten injections of yolk organic extract QC sample*

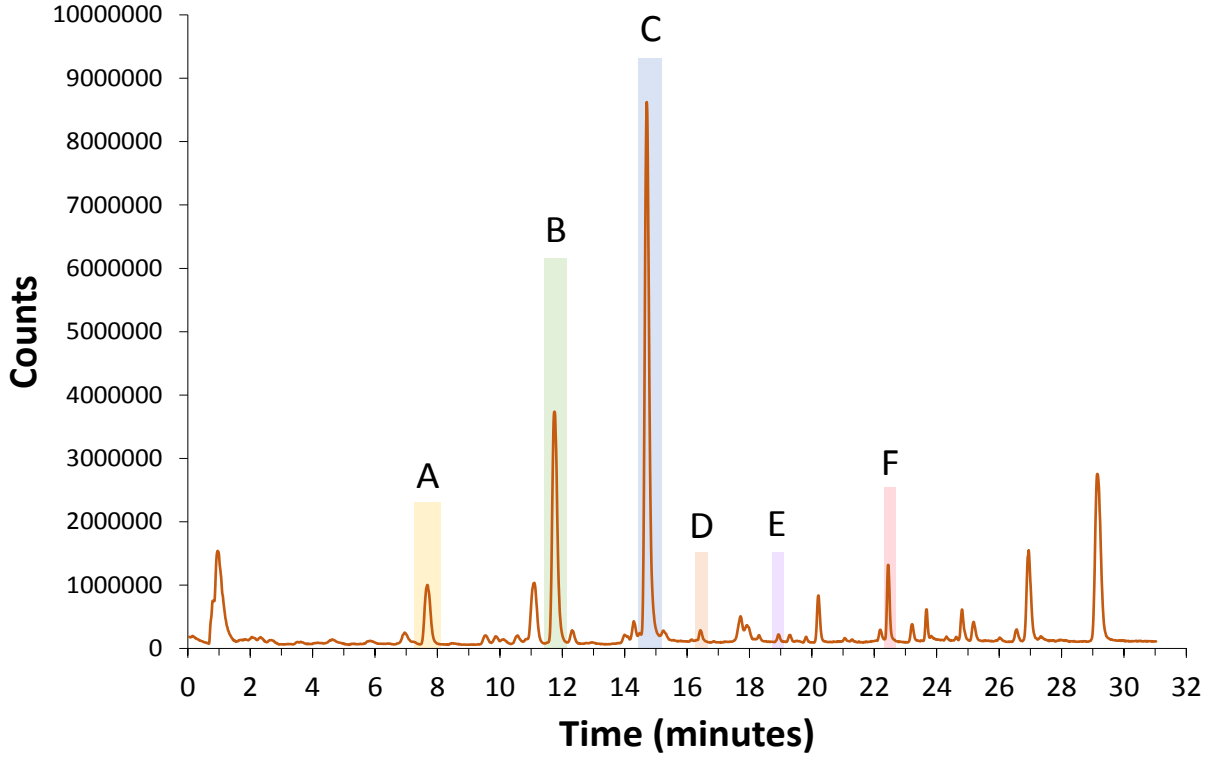
Figure 3.20

In both of the chromatograms in Figure 3.20, it can be seen that there is some retention time drift over the first three QC samples that were injected, as well as some instability in peak intensity. However, by the fourth injection the chromatograms for all QC sample injections remained stable, with no retention time drift or peak intensity instability. Although stability in retention time and peak intensity was reached at the fourth QC sample injection, ten samples were injected before each analytical run in order to ensure that the system was completely stable prior to sample extract analysis.

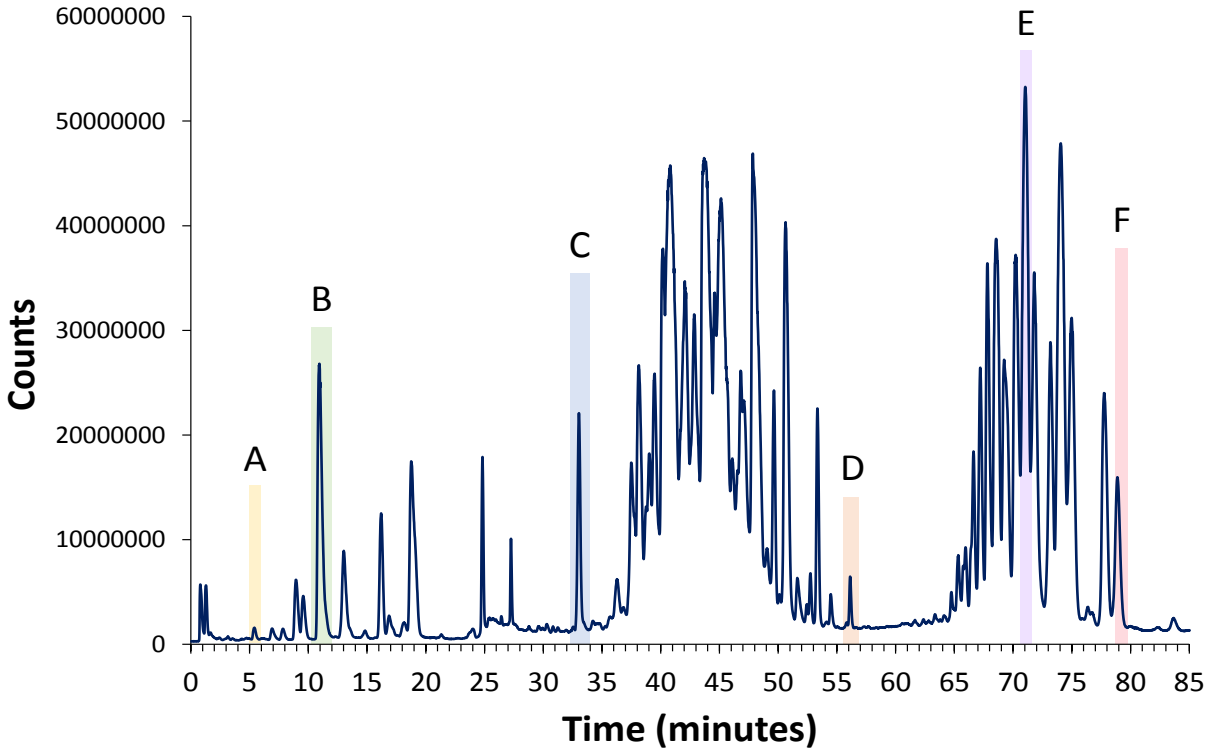
The QC sample for each study was also injected and analysed regularly throughout the analytical sequence, to monitor any drift in retention time and peak intensity that occurred throughout the analysis. This peak intensity drift can occur due to instrumental factors, such as column ageing, or temperature fluctuations [6], and will therefore be referred to throughout this research as instrumental drift. The drift in retention time can occur due to gradual column degradation and contamination [7], and will be referred to as retention time drift.

Quality control studies were carried out on the resulting data following chemical analysis, in order to ensure that it was reliable. Six peaks, in the Total Ion Chromatograms (TICs) of the QC samples analysed throughout the analytical runs, were chosen based on their intensity and retention time, ensuring that a range of intensities and retention times were monitored. Coefficient of variance percentages (CV%s) were calculated for both the intensities and the retention times of the peaks, in order to measure the precision and ensure that the analyses were reproducible throughout the analytical run. Acceptable levels of precision were taken from Theodoridis et al. [8]; CV%<30% for peak intensities and CV%<2% for retention times. Figure 3.21 shows the six peaks that were chosen to be monitored in both the albumen and yolk organic extract analyses.

During data analysis, when PCA scores plots are produced, the QC samples should be clustered more tightly together on the plots than the samples in other sample sets, indicating that there was little instrumental drift occurring throughout the analysis [3].



(a) *Total Ion Chromatogram for albumen organic extract analysis showing the six peaks that were chosen for quality control monitoring*



(b) *Total Ion Chromatogram for yolk organic extract analysis showing the six peaks that were chosen for quality control monitoring*

Figure 3.21

3.8 Data Pre-processing

Following chemical analysis of samples, and the quality control studies, the data was pre-processed using *XCMS Online*. Although this software has integrated data pre-processing and data analysis into one step, it was simply used throughout this research as a pre-processing tool, to produce a feature table containing the peak areas of all compounds in each sample, with compounds represented by their m/z and retention time [9].

XCMS Online has default parameters that have been optimised to pre-process data resulting from analysis on various different instruments [10]. As all of the analysis in this research was carried out using HPLC-Q-ToF-MS, the HPLC/QToF default parameters were used for pre-processing the data using this software.

3.9 Data Analysis

Following the initial QC analysis and the pre-processing of data by *XCMS Online*, resulting in the production of a feature table, the feature table was transferred into *Microsoft Excel* for data analysis to be carried out. All data were assumed to be normally distributed, as only one independent variable was changed between sample sets in each study, and the remaining independent variables were well controlled. The order sequence for the analysis of samples was randomised to prevent any instrumental bias, and quality control measures were implemented to monitor any changes throughout the analysis that could create bias. Therefore, data analysis was carried out assuming normality of the data. The following diagram on the next page describes the data analysis workflow:

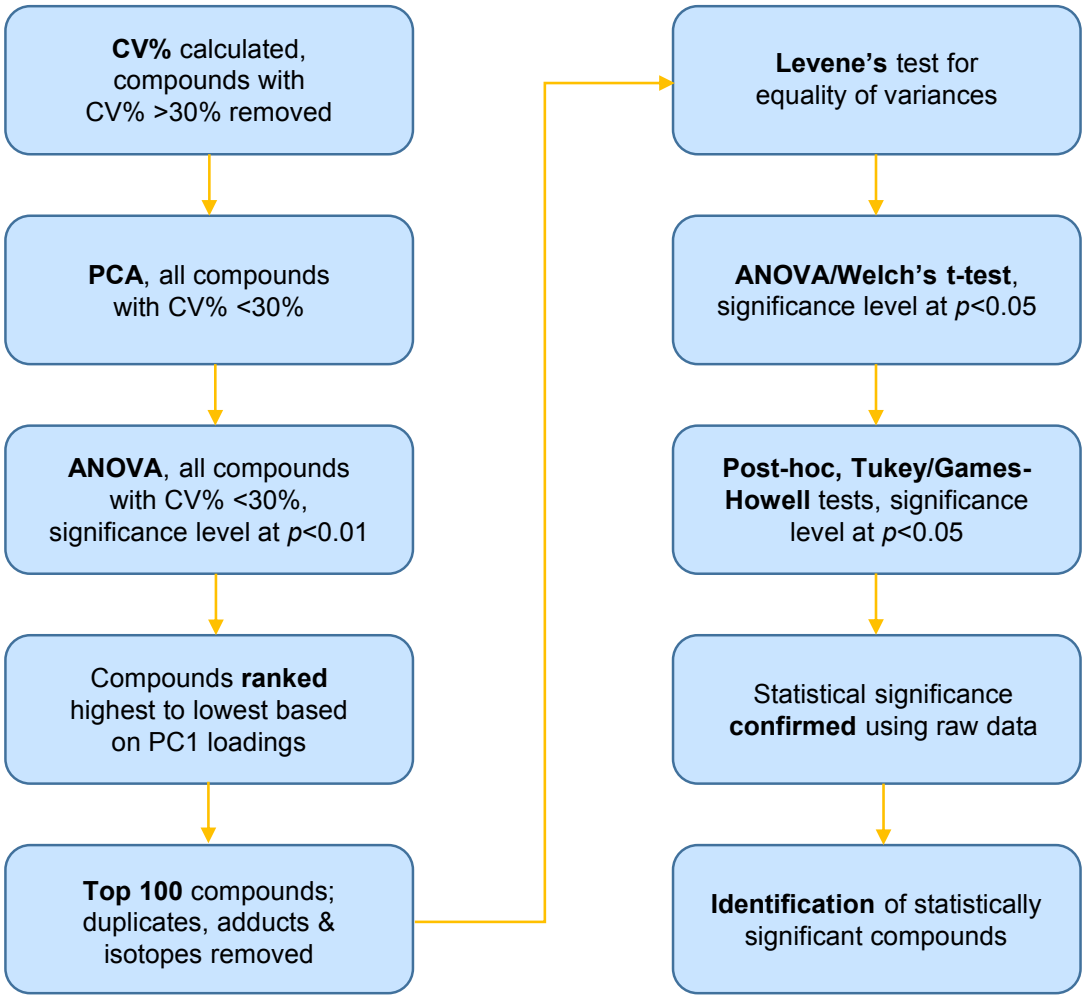


Figure 3.22: *Diagram showing the statistical workflow that was implemented in this research*

The first step of data analysis is more quality control; CV%s were calculated for all compounds in the feature table based on their peak areas in the QC samples. Compounds with a CV% >30% were removed from further analysis in accordance with published guidelines [8]; this is because compounds with such low precision in regards to detector response are not robust enough to act as reliable biomarkers.

Following this, Principal Component Analysis (PCA) was carried out for all remaining compounds, using an *Excel* Multivariate Analysis add-in which standardised the data and included six principal components. Scores plots were then produced in order to display the separation between the different sample sets. An initial ANOVA, assuming equal variances, was then carried out with the significance level at p<0.01. Any compounds

with $p>0.01$ were removed. This high threshold of statistical significance was used to narrow down the number of compounds in the data set, as there were still many that showed this high significance. The loadings from PC1 were used to rank the remaining compounds from the highest to the lowest, and the top 100 features were taken for further analysis. Any duplicates, adducts, or isotopes that were found in these top 100 compounds were removed from the analysis. *IBM SPSS Statistics* software was used to carry out a Levene's test for equality of variances on the remaining compounds, followed by either an ANOVA or Welch test, and post-hoc Tukey or Games-Howell tests, depending on the results of the Levene's test. The significance level was $p<0.05$. This lower threshold of statistical significance was used as it was not necessary to cut down the number of compounds any further, and some compounds with unequal variances, that may have had a p-value<0.01 in the initial ANOVA, may not be as highly significant when the more appropriate Welch test is used. Compounds that were still found to show significant differences between sample sets then had their statistical significance confirmed through the raw data. Extracted Ion Chromatograms (EICs) were produced in *Agilent Technologies' MassHunter Qualitative Analysis* software, and the peak areas from the integrated peaks in the EICs were used to carry out the same tests as before; Levene's test followed by ANOVA/Welch test and Tukey/Games-Howell tests. Identification of those compounds that were still found to be statistically significant then took place.

The loadings from PC1 were used to rank the compounds in most studies, regardless of which PCs best described the separation between sample sets. This is because PC1 describes the greatest amount of variation between all samples, and there was rarely just a single PC that best described the variation in the data; it was a combination of different PCs that showed the best separation between sample sets on the scores plots. At the stage of ranking the compounds, most compounds that remained in the data set were statistically significant due to the initial ANOVA that was carried out, so whichever PC was chosen to rank the compounds would result in statistically significant features being in the top 100.

3.9.1 Feature Identification

For compounds that were still found to be statistically significant when analysing the raw data, attempts were made to identify them. *Agilent Technologies' MassHunter Qualitative Analysis* software was used to predict potential molecular formulae for the compounds, based on the monoisotopic mass, isotope abundances, and isotopic peak spacings in the mass spectra [11]. Each predicted formula also had an associated probability score, based on how closely the isotope abundance ratios in the mass spectrum match those expected from the predicted formula [11]. All predicted formulae with a probability score of 95 or over were searched on METLIN, an online metabolite database, to find potential metabolite matches. For those compounds that did produce matches, the mass spectrum generated by the HPLC-MS analysis was compared with the mass spectrum provided by METLIN to see if they were comparable, and thus whether they could be putatively identified [12]. A difficulty that arose using this approach is that most mass spectra provided by METLIN are produced using higher collision energies, resulting in more fragmentation, whereas the MS analysis in these studies used 0 V collision energy, resulting in less fragmentation, thus it was not always possible to compare the mass spectra from the analysis with those on METLIN.

3.10 Summary

Although both aqueous and organic liquid extractions were carried out on the egg yolk and albumen, successful chromatographic methods were only developed for the analysis of organic extracts of the yolk and albumen. Quality control studies were developed in order to monitor any instrumental or retention time drift that may occur throughout the analysis, and ensure that they are within acceptable limits. A data pre-processing and analysis workflow was then developed to enable the discovery and potential identification of compounds capable of discriminating between authentic and fraudulent eggs.

3.11 References

[1] D. F. Thompson, F. Michopoulos, C. J. Smith, C. J. Duckett, R. W. Wilkinson, P. Jarvis, and I. D. Wilson, "Phosphorous and sulfur metabonomic profiling of tissue and plasma obtained from tumour-bearing mice using ultra-performance liquid chromatogrphy/inductively coupled plasma mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 27, pp. 2539–2545, 2013.

[2] P. Masson, A. Alves, T. Ebbels, J. Nicholson, and E. Want, "Optimization and evaluation of metabolite extraction protocols for untargeted metabolic profiling of liver samples by UPLC-MS," *Analytical Chemistry*, vol. 82, pp. 7779–7786, 2010.

[3] T. Sangster, H. Major, R. Plumb, J. A. Wilson, and I. D. Wilson, "A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis," *The Analyst*, vol. 131, pp. 1075–1078, 2006.

[4] H. Gika, G. Theodoridis, J. Wingate, and I. Wilson, "Within-Day Reproducibility of an HPLC-MS-based Method for metabonomic analysis: Application to Human Urine," *Journal of Proteome Research*, vol. 6, pp. 3291–3303, 2007.

[5] H. Gika, E. Macpherson, G. Theodoridis, and I. Wilson, "Evaluation of the repeatability of ultra-performance liquid chromatography-TOF-MS for global metabolic profiling of human urine samples," *Journal of Chromatography B*, vol. 871, pp. 299–305, 2008.

[6] F. Fernández-Albert, R. Llorach, M. Garcia-Aloy, A. Ziyatdinov, C. Andres-Lacueva, and A. Perera, "Intensity drift removal in LC/MS metabolomics by common variance compensation," *Bioinformatics*, vol. 30, pp. 2899–2905, 2014.

[7] M. Berg, M. Vanaerschot, A. Jankevics, B. Cuypers, R. Breitling, and J.-C. Dujardin, "LC-MS metabolomics from study design to data analysis - using a versatile pathogen as a test case," *Computational and Structural Biotechnology Journal*, vol. 4, 2013.

[8] G. A. Theodoridis, H. G. Gika, E. J. Want, and I. D. Wilson, "Liquid chromatography-mass spectrometry based global metabolite profiling: A review," *Analytica Chimica Acta*, vol. 711, pp. 7–16, 2012.

[9] H. P. Benton, J. Ivanisevic, N. G. Mahieu, M. E. Kurczy, C. H. Johnson, L. Franco, D. Rinehart, E. Valentine, H. Gowda, B. K. Ubhi, R. Tautenhahn, A. Gieschen, M. W. Fields, G. J. Patti, and G. Siuzdak, "Autonomous Metabolomics for Rapid Metabolite Identification in Global Profiling," *Analytical Chemistry*, vol. 87, pp. 884–891, 2015.

[10] R. Tautenhahn, G. J. Patti, D. Rinehart, and G. Siuzdak, "XCMS Online: A web-based platform to process untargeted metabolomic data," *Analytical Chemistry*, vol. 84, pp. 5035–5039, 2012.

[11] T. R. Sana, J. C. Roark, X. Li, K. Waddell, and S. M. Fischer, "Molecular Formula and METLIN Personal Metabolite Database Matching Applied to the Identification of Compounds Generated by LC/TOF-MS," *Journal of Biomolecular Techniques*, vol. 19, pp. 258–266, 2008.

[12] W. B. Dunn, A. Erban, R. J. M. Weber, D. J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, and M. R. Viant, "Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics," *Metabolomics*, vol. 9, S44–S66, 2013.

4. The effect of sample storage temperature and condition on the metabolite stability of egg yolk and albumen organic extracts

4.1 Introduction

The objective of metabonomic studies is to uncover differences between the metabolite profiles of organic tissues or biofluids that have been exposed to different endogenous or exogenous factors [1], e.g. chicken eggs that have originated from different housing systems. Non-targeted metabonomic studies aim to obtain as wide a metabolite profile as possible for the samples, increasing the potential of discovering compounds, or compound classes, of interest [2]. However, in order to ensure that any differences in metabolite profile that are observed are due to true biological differences, it is important that the storage of extracts following metabolite extraction, and prior to chemical analysis, minimises the potential for any further change in metabolite profile.

There has been considerable research into the effects of temperature and different storage conditions on the stability of extracts. A study into the effects of storage on a mix of metabolite and amino acid standards, found that when the compound mix was stored

at -20 °C, there was little distinction between samples that were stored for 3, 7, 14, 21, 28, or 42 days, indicating that the mix of standards was mostly stable at -20 °C [3]. Some metabolites and amino acids in the mix experienced a significant change in abundance over the 42 days of storage, however they did not appear to follow any particular trend over the storage time. Not all of the compounds were affected by an increasing storage time, indicating that different compounds are affected in different ways.

Another study investigated the effect of storage temperature and duration on the retention of carotenoids, and other secondary metabolites, in marigold petal extracts [4]. Extracts were stored at an ambient temperature, 4 °C, and -20 °C for 20, 40, and 60 days. It was found that all compounds under investigation, after 60 days of storage at all three storage temperatures, had decreased in abundance, indicating that increased storage time leads to greater metabolic change in the extracts. However, the extracts that were stored at -20 °C showed the least change, and the highest retention of compounds, over the 20, 40, and 60 days of storage, indicating that the storage of extracts at -20 °C improves the stability of the compounds, compared to at 4 °C and ambient temperatures.

A study was carried out into the modifications of the ^{1}H NMR metabolite profile of mullet roes when stored for 7 months at room temperature, 3 °C, and -20 °C [5]. This study found that following storage, extracts that were stored at -20 °C experienced the least amount of change compared to those that were analysed immediately after extraction. Extracts that were stored at 3 °C began to show relevant changes in metabolite profile after 5 months of storage, and those that were stored at room temperature showed an increasing difference in metabolite profile throughout the whole storage period. This corroborates the previously described results [4]; that storing extracts at -20 °C improves the stability of compounds. Another study using ^{1}H NMR to monitor the metabolite profiles of samples over a month of storage investigated urine samples, not metabolite extracts, however similar results were obtained [6]. Samples that were stored at 22 °C experienced a significant change in the concentrations of metabolites, whilst those that

were stored at 4 °C were slightly less affected, and those that were stored at -80 °C best reflected the original metabolite concentrations in the samples. Another, similar study, produced comparable results, with urine samples that were stored at -25 °C and -80 °C for 26 weeks experiencing no metabolite profile change [7]. This study also found that storing the samples at -80 °C had no benefit over sample storage at -25 °C, which was also concluded by other researchers when they investigated the stability of urine samples at -20 °C and -80 °C for 6 months [8].

Another study investigated the effect of storage temperature on the metabolite profiles of human plasma samples [9]. Again, the samples were not extracted prior to storage, but the results are comparable. Samples were stored at room temperature, 4 °C, -30 °C, and -80 °C for various lengths of time. It was observed that the samples that were stored at room temperature showed the greatest change in metabolite abundance over increasing storage time, compared to samples that were stored at -80 °C, followed by those that were stored at 4 °C. Samples that were stored at -30 °C showed very little change compared to those that were stored at -80 °C, indicating that lower storage temperatures improve the stability of the metabolite profiles of the samples, and that little benefit is gained from storing samples at temperatures lower than -30 °C. Another study also investigated the stability of human plasma samples at different storage temperatures [10]. Samples were stored at room temperature, -20 °C, and -80 °C, and changes in the concentrations of some metabolites in the samples were observed following just 2.5 hours of storage at room temperature. Samples remained relatively unchanged up to 7 days of storage at -20 °C, but for storage periods of one month or longer, it was recommended that samples be stored at -80 °C to minimise any metabolic changes in the samples.

One study investigated the effects of storage temperature and condition on the steroid metabolite extracts of baboon faeces over 50 weeks of storage [11]. However, this study produced contrasting results to the previous studies described, as it was determined that after 24 weeks the metabolite profiles of the liquid extracts, and extracts that were dried

onto SPE cartridges, both stored at 30 °C, showed a similar stability to the liquid extracts that were stored at -20 °C. However, following 50 weeks of storage, the liquid extracts that were stored at 30 °C did show a significant difference in metabolite concentration compared to liquid extracts that were frozen at -20 °C. The authors postulate that this is due to the extract evaporating slightly during storage, and suggest that storage at slightly lower temperatures, or drying extracts prior to storage would prevent this. However, the extracts that were stored dried down at 30 °C showed the most instability, and the greatest metabolic change throughout the 50 weeks of storage. The authors discussed possible explanations for this, suggesting that as the concentrations of metabolites were lower for the dried down extracts compared to the others, it could partly be due to insufficient recovery from the walls of the sample tubes during reconstitution of the extracts in solvent. As the concentration of metabolites in the dried down extracts appeared to decrease over an increasing storage time, it was postulated that perhaps the metabolites became bound to the polypropylene wall of the tube, and that the cohesiveness of the binding increased over time, resulting in a lower recovery of metabolites during reconstitution at each sampling point.

Although the final study mentioned is quite contrasting to the others, most studies appear to agree that storing samples and extracts under frozen conditions, rather than refrigerated or at ambient temperatures, improves the stability of the metabolites. Some studies determined that there is no benefit to storing extracts at -80 °C compared to -25 °C [7][8], but another study concluded that whilst -20 °C results in a high metabolite stability over a short storage time, for periods of storage over a month, -80 °C is recommended [10]. The effect of storage temperature on the stability of the extracts appears to be slightly different for the different sample types studied, therefore it is important to investigate the impact of temperature and storage condition on the stability of the metabolite profiles of the sample extracts that are being studied.

4.2 Aims and Objectives

As it is crucial in metabonomic studies to ensure that any observed differences in metabolite profile between sample extracts are due to true biological differences, it is important to ensure that the storage conditions of the extracts, prior to analysis, minimise any change in metabolite profile of the extracts during storage.

This work aimed to use the same metabonomic workflow that is used throughout the rest of this research, to determine whether storage condition; dried or re-dissolved in methanol, and storage temperature; -25 °C, -46 °C, or -80 °C, affect the resulting metabolite profiles of yolk and albumen organic extracts, and to conclude which storage conditions result in the highest stability of metabolite profiles of extracts.

4.3 Experimental

For this study, an organic metabolite extraction was carried out on the yolk and albumen of just one egg, from a batch of eggs that were collected from Oaklands Farm Eggs Ltd., Shrewsbury, U.K. Laying hens were of the Hy-line brown breed, kept in enriched cages of 20 birds per cage, and were 39 weeks old at the point of lay. Just one egg was used, in order to minimise the amount of biological variation between extracts, to ensure that the main differences between extracts were due to their storage conditions.

The yolk and albumen were separated and each was split into six aliquots, a–f.

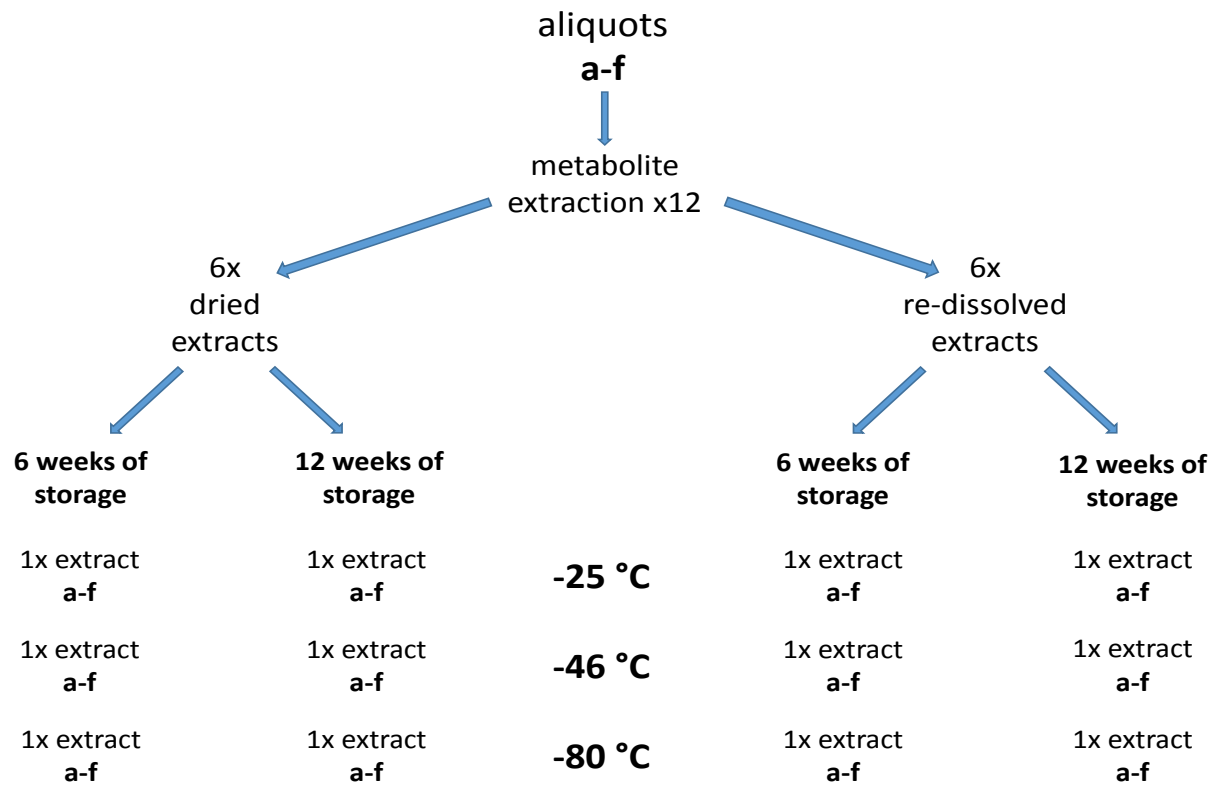


Figure 4.1: *Diagram showing the experimental design for the yolk organic extract storage study*

Twelve organic metabolite extractions were then carried out per aliquot of yolk, resulting in a total of seventy-two yolk organic extracts. Six of the metabolite extractions carried out per aliquot were completed with the re-dissolving of the dried extract in

methanol, and six extracts were retained in their dried state. Two dried and re-dissolved extracts from each aliquot were then stored in freezers kept at three different temperatures; -25 °C, -46 °C, and -80 °C. Once the experimental set-up was complete, each freezer stored twenty-four extracts; two dried, and two re-dissolved extracts per aliquot, a–f. Following 6 weeks of storage, one dried and one re-dissolved organic yolk extract for each aliquot, a–f, from each freezer were defrosted, and the dried extracts were re-dissolved in methanol. These thirty-six extracts were then analysed using HPLC-MS. This was then repeated following 12 weeks of yolk extract storage. Figure 4.1 on the previous page shows the experimental design for this study.

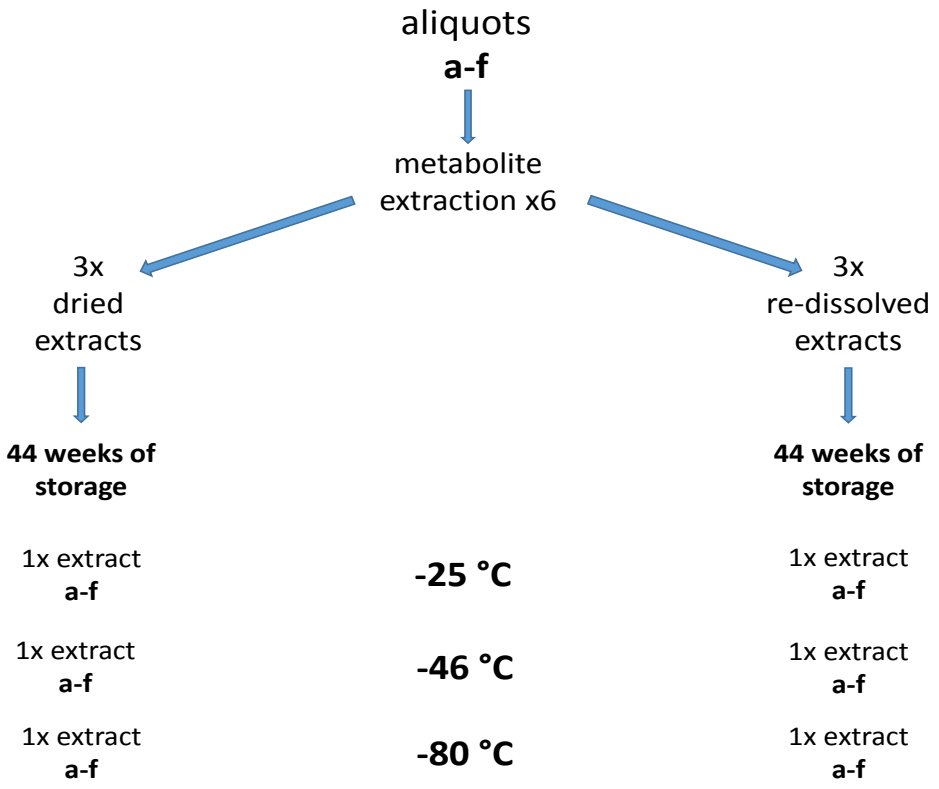


Figure 4.2: *Diagram showing the experimental design for the albumen organic extract storage study*

Six organic metabolite extractions were carried out per aliquot of albumen, resulting in a total of thirty-six albumen organic extracts. Three of the metabolite extractions carried out per aliquot were completed with the re-dissolving of the dried extract in methanol, and three extracts were retained in their dried state. One of the dried and

re-dissolved extracts from each aliquot were then stored in freezers kept at three different temperatures; -25 °C, -46 °C, and -80 °C. Once the experimental set-up was complete, each freezer stored twelve extracts; one dried, and one re-dissolved extract per aliquot, a–f. After forty-four weeks of storage, all albumen organic extracts from each freezer were defrosted, and the dried extracts were re-dissolved in methanol. These thirty-six extracts were then analysed using HPLC-MS. Figure 4.2 shows the experimental design for this study.

Organic metabolite extraction of the yolk and albumen, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. Data analysis for all three experiments (yolk extracts after 6 and 12 weeks of storage, and albumen extracts after 44 weeks of storage) was then carried out as described in Chapter 3, Section 3.9, with the addition of a second PCA, using only the top statistically significant compounds following ANOVA/Welch tests. No identification of compounds took place, as the aim of the study was to understand *if* and *how* the compounds that make up the metabolite profiles of the extracts are affected by storage temperature and condition, rather than *which* compounds are significantly affected.

4.4 Yolk Organic Extracts: 6 weeks of Storage

Yolk organic extracts that were stored either dried or re-dissolved for 6 weeks, at three different temperatures, were compared to see whether storage condition and temperature affect the metabolite profile of an extract.

4.4.1 Results and Discussion

4.4.1.1 Quality Control Analysis

Table 4.1 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 4.1: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for six weeks*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	785042	119413170	44026133	29171841	323227222	81307383
QC2	732145	119729042	46545772	31063133	345779678	71893451
QC3	945162	126470319	48390361	30859982	351324044	66975924
QC4	733781	114969574	45951088	30502164	341964294	56560805
QC5	686170	108880211	45047815	31836032	335824304	58481910
QC6	575895	112695186	44334359	29883792	335545978	57061490
QC7	463816	109084612	42788591	28817781	321482161	51646230
SD	153198	6397898	1846263	1075406	11096353	10447392
Mean	703144	115891731	45297731	30304961	336449668	63418170
CV%	21.79	5.52	4.08	3.55	3.30	16.47

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 4.2, shows the retention times (RTs) for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 4.2: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for six weeks*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	5.088	9.122	32.410	56.529	70.971	77.129
QC2	5.072	9.256	32.345	56.381	70.623	76.533
QC3	5.085	9.169	32.425	56.593	71.002	77.177
QC4	5.135	9.235	32.474	56.610	71.068	77.193
QC5	5.069	9.236	32.558	56.627	70.869	76.861
QC6	5.052	9.185	32.407	56.510	70.894	76.944
QC7	5.082	9.232	32.487	56.623	71.048	77.206
SD	0.026	0.048	0.069	0.089	0.152	0.248
Mean	5.083	9.205	32.444	56.553	70.925	77.006
CV%	0.51	0.52	0.21	0.16	0.21	0.32

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

4.4.1.2 Metabolite Profiling

4.4.1.2.1 Comparison of storage temperature

Comparisons were made between extracts that were stored at the three different temperatures, by studying the Total Ion Chromatograms (TICs) of the extracts following HPLC-MS analysis. Figure 4.3 shows overlaid TICs for the average chromatograms of extracts (both dried and re-dissolved) stored at the three different temperatures.

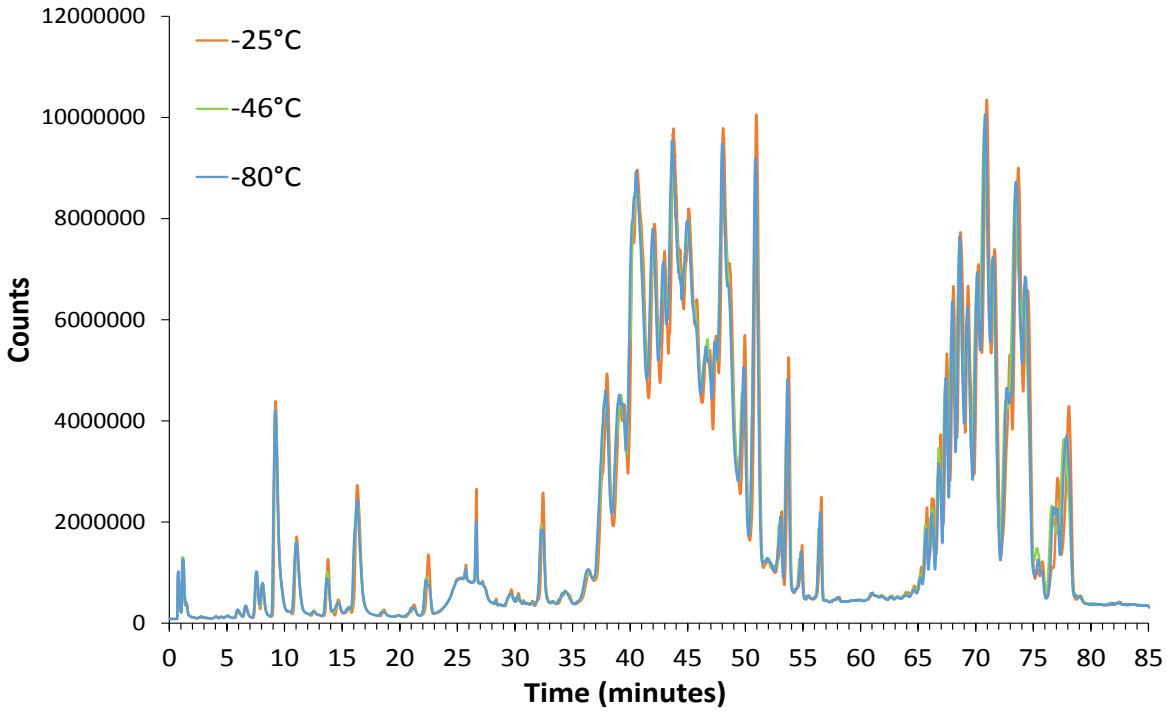


Figure 4.3: *Overlaid average TICs of yolk organic extracts that were stored for 6 weeks at -25 °C, -46 °C, and -80 °C*

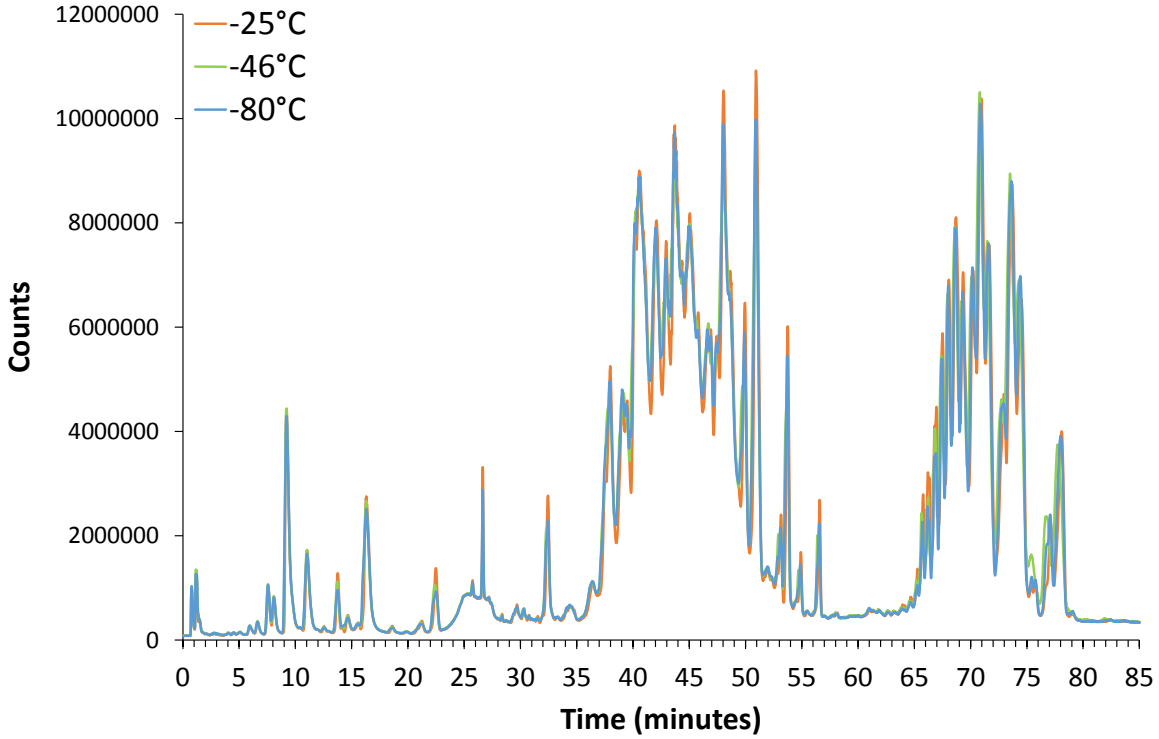
Observing these overlaid TICs, a slight difference can be seen between the chromatograms for the extracts stored at the different temperatures. There is almost no observable difference in peak intensity between the chromatograms of extracts that were stored at -46 °C and -80 °C, however there is a slightly higher peak intensity in the TIC of extracts that were stored at -25 °C, compared to the TICs of extracts that were stored at the other two temperatures. This indicates that there is a greater stability in metabolite profile between extracts that are stored at different temperatures, when the storage temperatures are lower.

4.4.1.2.2 Comparison of temperature in each storage condition

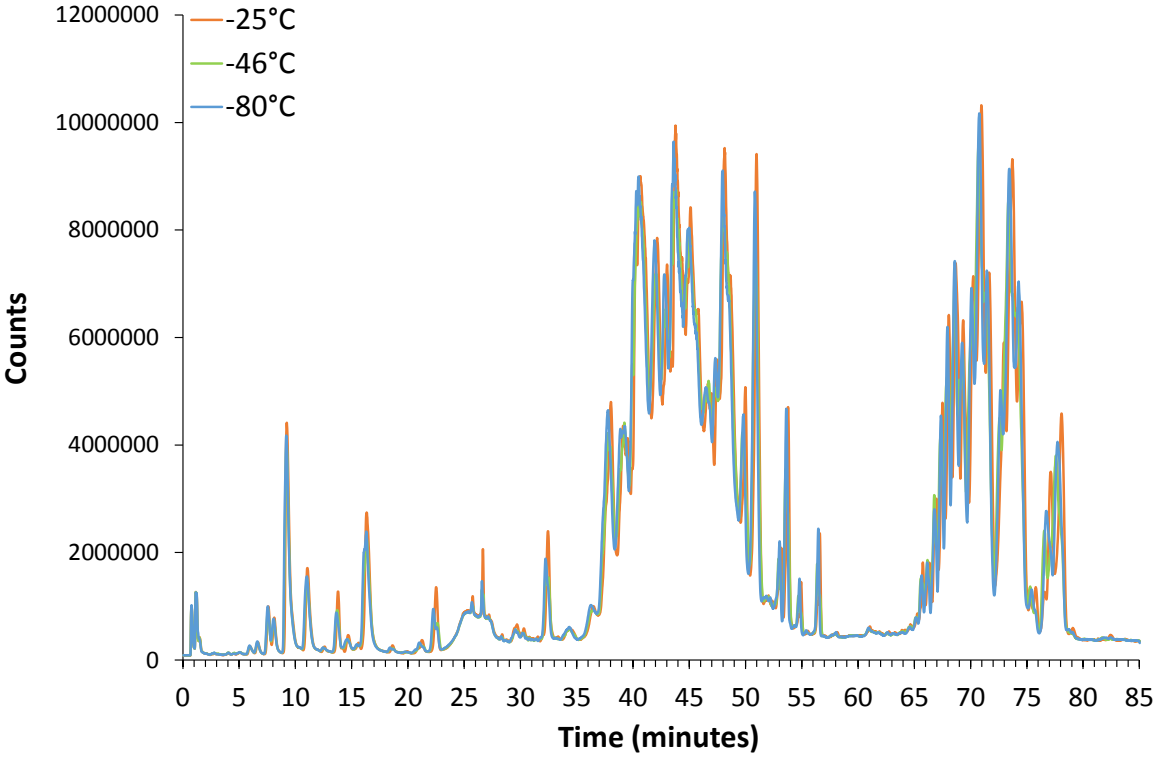
The chromatograms in Figure 4.4 show comparisons between the TICs of re-dissolved (Figure 4.4 a)) and dried (Figure 4.4 b)) extracts stored at all three temperatures. The chromatogram in Figure 4.4 a) shows that between 13-70 minutes, there is a considerably higher peak intensity for the TIC of re-dissolved extracts that were stored at -25 °C, compared to those of the re-dissolved extracts that were stored at the other two temperatures, which showed very little difference in peak intensity to each other throughout the chromatogram. Before 13 minutes, and after 70 minutes, there is very little difference in peak intensity between the TICs of re-dissolved extracts stored at all three temperatures, with the TIC of extracts that were stored at -25 °C showing very slightly lower peak intensities compared to the TICs of the extracts stored at the other two temperatures.

This changing difference in peak intensity throughout the chromatogram could be because the later eluting, more non-polar compounds are likely to be larger lipid molecules which undergo degradation during storage, breaking down to produce more of the polar compounds. This degradation may occur at a lower rate in extracts that have been stored at lower temperatures, resulting in the non-polar compounds having a higher abundance in extracts stored at the lower temperatures compared to the higher temperatures. The higher abundance of the earlier eluting, less non-polar compounds in the TIC of extracts stored at -25 °C is likely to be due to the higher rate of degradation of the more non-polar compounds at the higher storage temperature resulting in the increased production of these less non-polar compounds, compared to at lower storage temperatures. This suggests that storing extracts at lower temperatures improves the metabolite stability.

The chromatogram in Figure 4.4 b) shows that the TIC of dried extracts that were stored at -25 °C has a higher peak intensity than those of dried extracts stored at the other two temperatures, which have a similar peak intensity to each other, throughout the whole of the analysis. The lack of change in peak intensity difference due to compound polarity suggests that there was a reduced rate of, or no degradation of, metabolites in



(a) *Overlaid TICs for yolk organic extracts re-dissolved in methanol*



(b) *Overlaid TICs for yolk organic dried extracts*

Figure 4.4: *Overlaid TICs for yolk organic extracts stored at all three temperatures a) re-dissolved in methanol and b) dried, for 6 weeks*

dried extracts during storage, indicating that storing the extracts dried provides greater metabolite stability, compared to re-dissolving extracts in methanol prior to storage. The reason that the TIC of dried extracts that were stored at -25 °C has a higher peak intensity throughout the analysis compared to the TICs of dried extracts that were stored at the other two temperatures, could be due to a better recovery of analytes during the re-dissolving of the dried extracts that were stored at -25 °C in methanol, than in the re-dissolving of the extracts that were stored at the two lower temperatures. The similarities in peak intensity between dried extracts that were stored at -46 °C and -80 °C indicate that storing the extracts at these two temperatures results in a similar metabolite recovery yield.

4.4.1.2.3 Comparison of storage condition

Comparisons were also made between extracts that were stored dried and re-dissolved in methanol, by studying the TICs of the extracts following HPLC-MS analysis. Figure 4.5 shows overlaid TICs for the average chromatograms of extracts that were stored dried and re-dissolved in methanol, at all three temperatures.

There is a clear difference in peak intensity between the chromatograms of dried extracts, and extracts that were re-dissolved in methanol; there is a higher peak intensity for most peaks in the TIC of re-dissolved extracts, compared to that of dried extracts. This shows that the storage condition of extracts, i.e. dried or re-dissolved, does have an impact on the resulting metabolite profile of the extracts after storage. This could be due to, as postulated by Kalbitzer and Heistermann [11], an insufficient recovery of analytes from the wall of the sample tube during the re-dissolving of the dried extracts in methanol following storage.

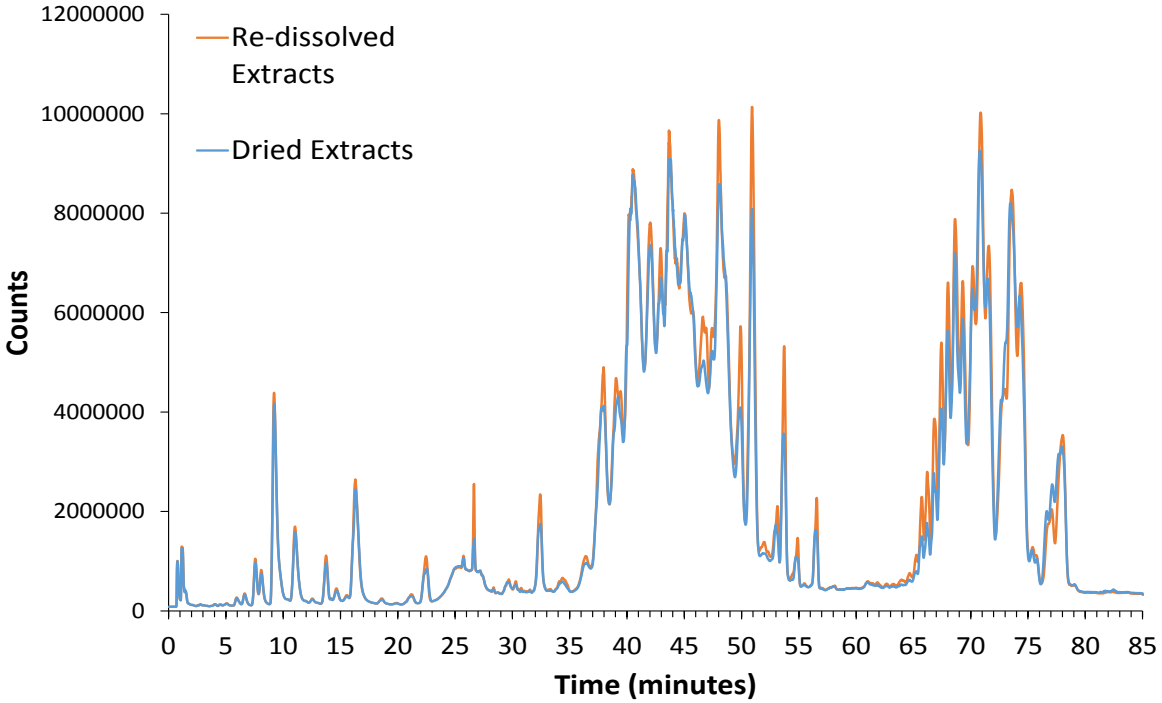
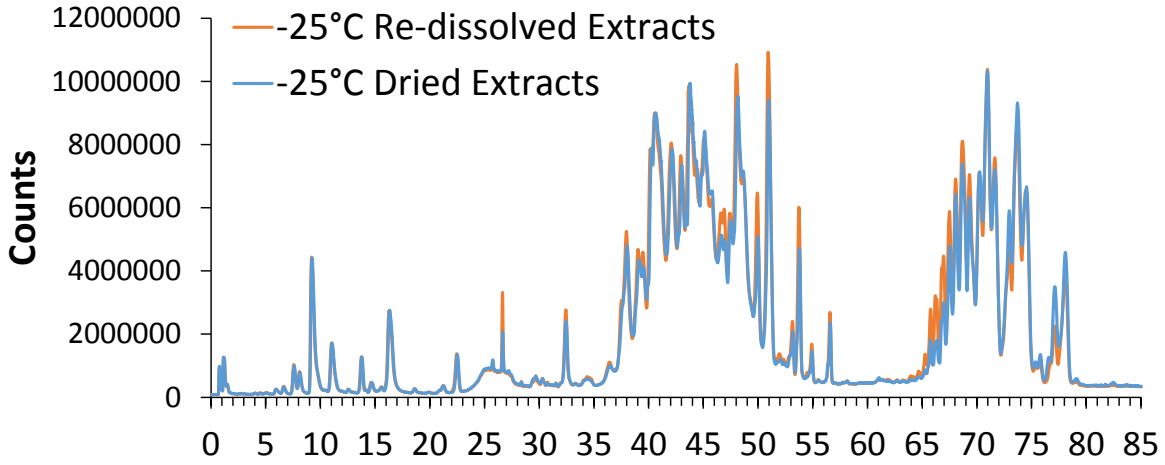


Figure 4.5: *Overlaid average TICs of yolk organic extracts that were stored dried and re-dissolved in methanol for 6 weeks*

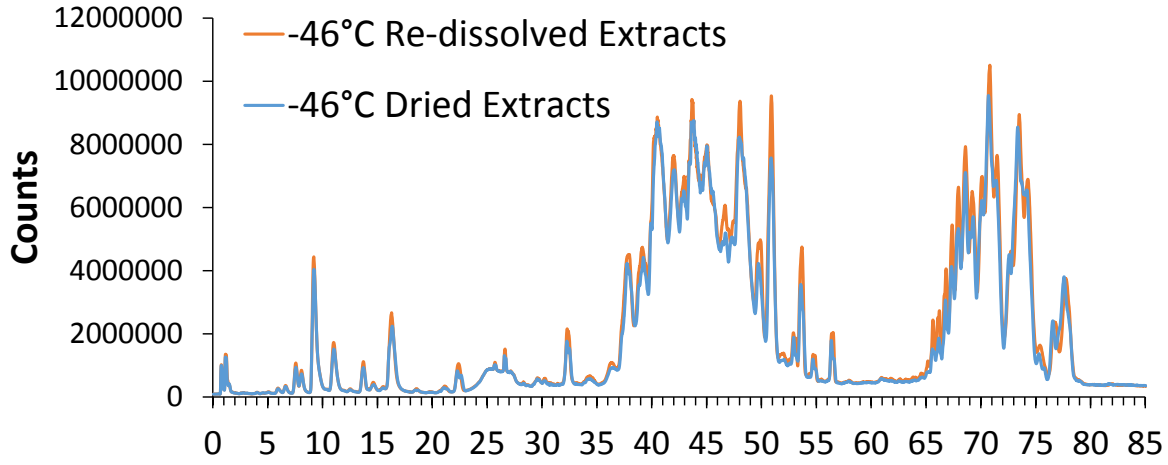
4.4.1.2.4 Comparison of storage condition at each temperature

Comparisons were also made between dried and re-dissolved extracts at each of the three different storage temperatures. Figure 4.6 shows the overlaid TICs for the average chromatograms of extracts that were stored dried and re-dissolved in methanol at -25 °C, -46 °C, and -80 °C.

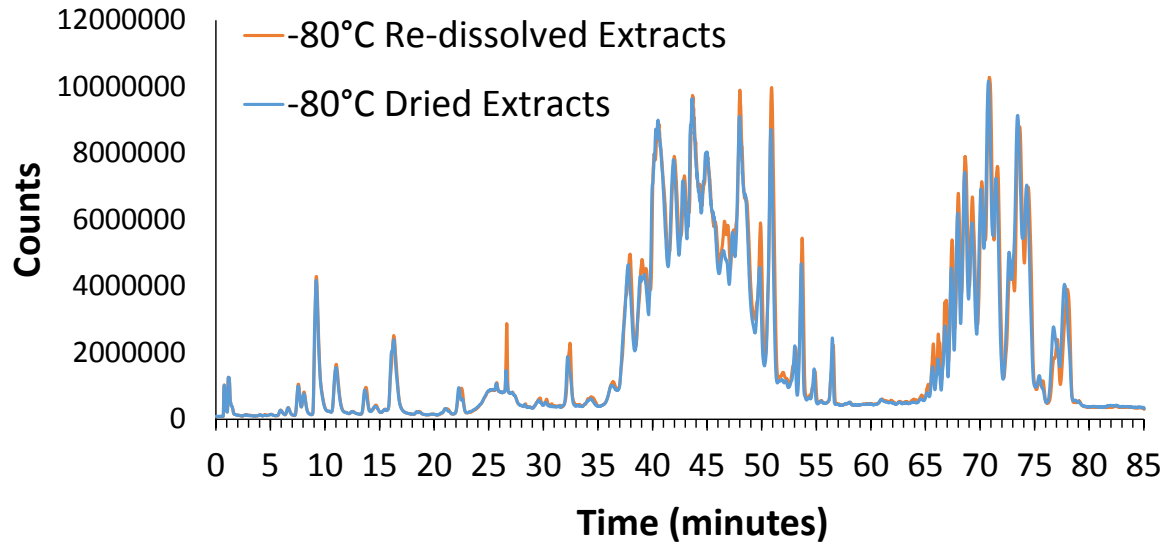
These overlaid TICs show that there is some difference between the average chromatograms of dried and re-dissolved extracts that have been stored at all three temperatures for 6 weeks. Figures 4.6 a) and c) show that, for extracts that were stored at -25 °C and -80 °C, there is very little difference in peak intensity between the TICs of dried and re-dissolved extracts during the first 25 minutes of analysis. This indicates that storage condition does not have much of an effect on the resulting metabolite profile of the less non-polar compounds, following 6 weeks of storage at these temperatures. However, the chromatograms of extracts stored at -80 °C in Figure 4.6 c) do show a very slightly



(a)



(b)



(c)

Figure 4.6: *Overlaid TICs for dried and re-dissolved yolk organic extracts stored at a) -25 ˚C b) -46 ˚C and c) -80 ˚C for 6 weeks*

higher peak intensity for the TIC of re-dissolved extracts compared to that of dried extracts during the first 25 minutes of analysis. Between 25 and 70 minutes of analysis, the chromatograms of extracts stored at both -25 °C and -80 °C show that some peaks are of a considerably higher intensity in the TICs of re-dissolved extracts, compared to those of dried extracts. This suggests that the storage condition of extracts does have some effect on the metabolite profile of the more non-polar compounds. After 70 minutes of analysis, the opposite trend is observed, and there are several peaks which have a higher peak intensity in the TICs of dried extracts, compared to those of re-dissolved extracts. This difference in peak intensity is greater between dried and re-dissolved extracts that were stored at -25 °C, than between the extracts that were stored at -80 °C, suggesting that lower storage temperatures provide greater stability of extracts.

This greater peak intensity for the TICs of dried extracts towards the end of the analysis indicates that the storage condition of the extracts affects the more non-polar compounds in an opposite manner to the slightly more polar compounds, following 6 weeks of storage at these two temperatures. This could be because the more non-polar compounds undergo degradation during storage, producing more of the polar compounds. This degradation may occur at a higher rate in extracts that have been re-dissolved in methanol prior to storage, resulting in the non-polar compounds having a lower abundance, and the polar compounds having a higher abundance, in re-dissolved extracts compared to dried extracts following 6 weeks of storage. This suggests that a better metabolite stability is achieved through storing extracts dried, rather than re-dissolved in methanol.

The chromatogram in Figure 4.6 b) shows a greater difference in peak intensity between the chromatograms of dried and re-dissolved extracts during the first 25 minutes of analysis, than was observed between the TICs of dried and re-dissolved extracts that were stored at -80 °C, with the TIC for re-dissolved extracts again having a higher peak intensity. This trend is observed throughout the chromatograms of extracts stored at

-46 °C, in Figure 4.6 b); there is a higher peak intensity for almost all peaks in the TIC of re-dissolved extracts compared to that of dried extracts. This indicates that at this temperature, the polarity of the compounds does not influence how they are affected by the storage condition of the extracts. It could be that following 6 weeks of storage at this temperature the metabolite recovery yield, resulting from the re-dissolving of extracts in methanol, has a greater effect on the resulting metabolite profile, than the difference in degradation rates between the extracts stored under different conditions.

4.4.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the extracts that were stored either dried or re-dissolved at the three different temperatures. The scores plot in Figure 4.7 shows that the greatest difference between extracts was between those that were stored dried and re-dissolved, rather than between those that were stored at different temperatures. It shows most dried extracts to be on the right side of the plot, and most re-dissolved extracts to be on the left side of the plot, divided by the purple dashed line. The only exceptions to this are one dried extract that was stored at -25 °C, and one re-dissolved extract that was stored at -46 °C, which are circled on the plot. It is not known why these two extracts show so much difference to the others within their sample sets, particularly the re-dissolved extract that was stored at -46 °C. Nothing different was noted during the extraction procedure that could explain this, however it could be that an error occurred during the metabolite extraction that was not noticed at the time. There is no obvious difference in the chromatogram for the dried extract stored at -25 °C, compared to the other chromatograms for the other extracts stored in the same conditions, however the chromatogram for the re-dissolved extract that was stored at -46 °C shows a much higher peak intensity for the penultimate peak in the TIC, compared to the other chromatograms for the remaining extracts stored in these conditions.

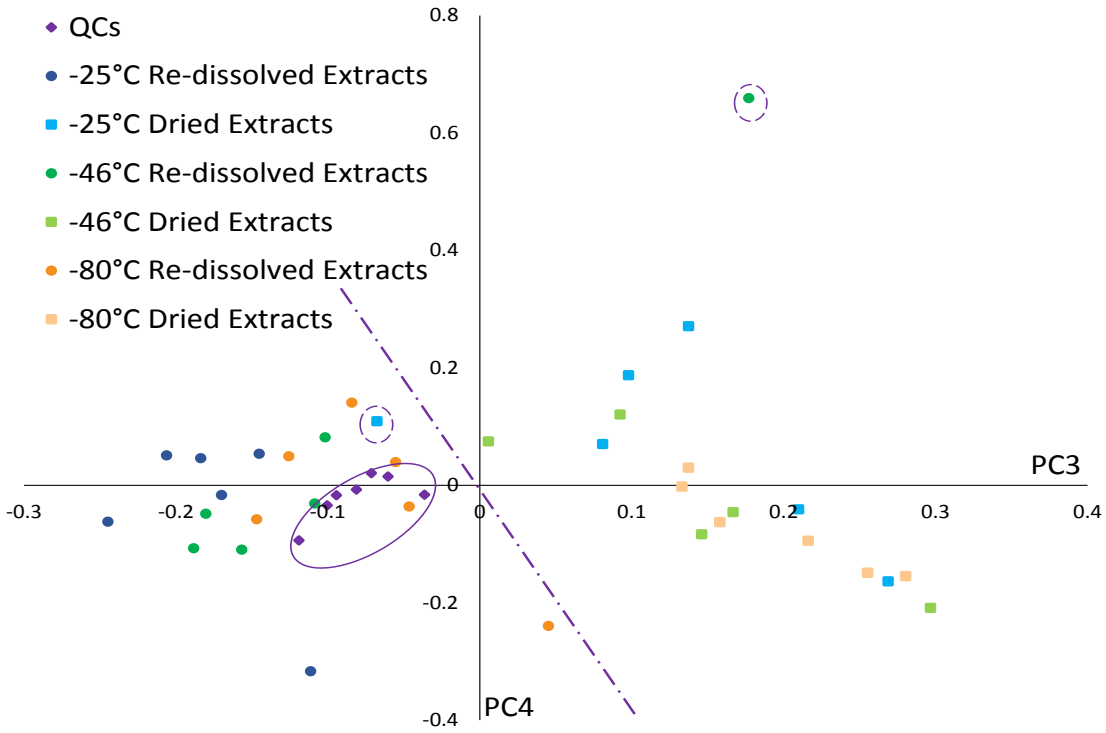


Figure 4.7: *PCA scores plot showing PC3 vs PC4 for yolk organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 6 weeks, including all compounds with CV%<30%. PC3 explains 0.29% of the variance, and PC4 explains 0.10% of the variance. Dashed line separates dried and re-dissolved extracts. Circled: samples on opposite sides of the plot to what is expected.*

The separation between dried and re-dissolved extracts on the scores plot in Figure 4.7 is mostly across PC3, which only accounts for 0.29% of the total variance between samples in the PCA, proving that the differences in metabolite profile between extracts stored under different conditions are extremely subtle. Although the QC samples are not particularly tightly clustered, they do form a tighter grouping than the other sample sets, which shows that the differences between samples and sample sets are due to true biological differences, rather than instrumental drift.

4.4.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 44 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, all 44 were found to be statistically

significant.

There are 27 compounds of the top 44 that have been found to show significant differences between dried and re-dissolved extracts at at least one storage temperature, one compound that has been found to show a significant difference between extracts stored in the same condition but at different temperatures, and 5 compounds that have been found to show significant differences between dried and re-dissolved extracts at at least one storage temperature, as well as between extracts stored in the same condition but at different temperatures. The remaining 11 compounds have only been found to be significantly different between extracts of different storage temperatures *and* conditions. This, again, indicates that the storage condition of the extracts, i.e. dried or re-dissolved, has a greater effect on the resulting metabolite profiles of the extracts following 6 weeks of storage, than the storage temperature. Of the 32 compounds that show a significant difference in abundance between dried and re-dissolved extracts; 25 are significantly different between the two storage conditions at -25 °C, 18 show a significant difference between the storage conditions at -46 °C, and 19 are significantly different between the storage conditions at -80 °C. This shows that the storage condition of the extracts has a slightly reduced effect when extracts are stored at lower temperatures, indicating that it is best to store extracts at lower temperatures to lessen any changes to metabolite profile that may occur during storage. Of the 6 compounds that have been found to show statistical significance when comparing extracts stored under the same conditions but at different temperatures, only one of these shows a significant difference in abundance between dried extracts stored at different temperatures; the other compounds only show differences between re-dissolved extracts stored at different temperatures. This indicates that storing extracts dried, rather than re-dissolved in solvent, reduces the effect of temperature on the resulting metabolite profile after extract storage.

The 33 compounds that are statistically significant when comparing extracts from the same storage condition but different temperatures, or the same temperature but

different storage conditions, can be found in Table A.1 in Appendix A. This table shows their CV%s, p-values resulting from ANOVA/Welch test, and figures representing the difference in the mean abundance of the compounds between the extracts. Those in bold were also found to be statistically significant when confirmed with the raw data.

As can be seen in Table A.1, there are various trends in the different abundances of the compounds between extracts stored under different conditions and at different temperatures. Some compounds are clearly of a higher abundance in all re-dissolved extracts, compared to in dried extracts, whereas others show the opposite trend and are of a higher abundance in all dried extracts. One compound, m/z 734.5704, has a higher abundance in the re-dissolved extracts that were stored at -25 °C compared to the dried extracts, but a higher abundance in the dried extracts that were stored at -46 °C and -80 °C compared to the re-dissolved extracts. However, this compound has not been found to show any significant differences between dried and re-dissolved extracts; it is only significantly different when comparing re-dissolved extracts stored at -25 °C and those stored at -46 °C and -80 °C, which explains this lack of consistency regarding differences in abundance between dried and re-dissolved extracts at different temperatures.

There are 5 compounds that have been found to be significantly different between dried and re-dissolved extracts at at least one storage temperature, as well as between extracts stored in the same condition but at different temperatures. Of these 5 compounds, m/z 744.5544, 844.7413, and 896.7739, show a correlation between abundance and storage temperature in re-dissolved extracts. Only one of these 3 compounds, m/z 744.5544, also shows a correlation between abundance and storage temperature in dried extracts, indicating that extract storage temperature has a greater effect on the resulting metabolite profile of extracts that were re-dissolved in methanol prior to storage, than those that were dried. The remaining 2 compounds of the 5 that show statistical significance due to both storage condition and temperature, m/z 610.5415 and 870.7593, do not show a correlation between abundance and storage temperature in either dried

or re-dissolved extracts; they both experience just one instance of a large difference in abundance between extracts stored at one temperature and condition compared to the others. These large differences in abundance for both compounds are due to re-dissolved extracts, indicating that the metabolite profiles of dried extracts remain more stable than those of extracts that are re-dissolved in methanol prior to storage.

4.4.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 44 statistically significant compounds, and scores plots were again produced to display the variation between yolk organic extracts that were stored either dried or re-dissolved in methanol, at the three different temperatures.

The scores plot in Figure 4.8 shows a similar variation between extracts to the previous scores plot in Figure 4.7. All of the dried extracts are on the right side of the plot, and the re-dissolved extracts are on the left side of the plot, divided by the purple dashed line. This, again, indicates that the storage condition of the extracts, i.e. dried or re-dissolved, has a greater effect on the resulting metabolite profiles of the extracts after 6 weeks of storage, than the storage temperature. The re-dissolved extracts show some slight separation between those stored at the three different temperatures, indicating that the top statistically significant compounds describe the variation between re-dissolved extracts due to temperature, better than when all of the compounds with a CV%<30% are included in the PCA. The extracts stored at -80 °C are the closest to the centre of the plot, those stored at -46 °C are slightly further out, and those stored at -25 °C are the furthest to the left of the plot. As the re-dissolved extracts that were stored at -80 °C are the closest to the centre of the plot, and the QC samples, it indicates that they are the most representative of all of the organic yolk extracts, and therefore, if extracts are to be stored re-dissolved in methanol, they should be stored at -80 °C to minimise any changes to the metabolite profile. However, the dried extracts show no separation between those

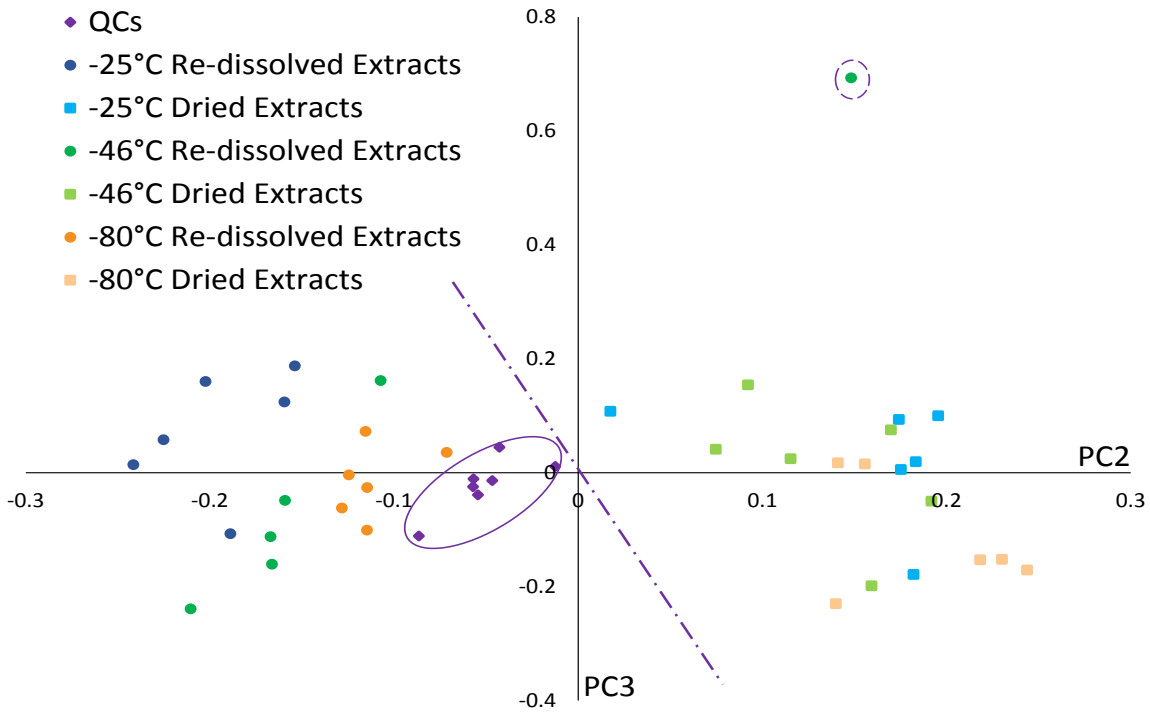


Figure 4.8: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 6 weeks, including the top 44 statistically significant compounds. PC2 explains 0.68% of the variance, and PC3 explains 0.16% of the variance. Dashed line separates dried and re-dissolved extracts. Circled: sample on opposite side of the plot to what is expected.*

stored at different temperatures, indicating that temperature has very little effect on the metabolite profiles of dried extracts. This shows that storing extracts dried, rather than re-dissolved in methanol, reduces the effect of temperature on the metabolite profiles of the extracts during storage, suggesting that it may be best to store extracts dried, rather than re-dissolved in solvent.

The separation between dried and re-dissolved extracts on this scores plot is across PC2, which describes 0.68% of the total variance between samples, compared to 0.29% described by PC3 in the previous plot in Figure 4.13. This shows that when just the top statistically significant compounds are included in the PCA, there is a higher percentage of variance describing the difference between dried and re-dissolved extracts. There is again one re-dissolved extract (circled) that was stored at -46 °C that appears to be an outlier and is very separate to the rest of the extracts. This is the only sample that does

not follow the trend of dried and re-dissolved extracts on opposite sides of the scores plot. The QC samples are again more tightly grouped than the sample sets, indicating that there was little instrumental drift affecting the analysis, and thus that the analysis was robust.

4.4.1.5 Summary

From the overlaid TICs in Figures 4.3-4.6, it could be seen that both storage temperature and condition have an effect on the resulting metabolite profiles of yolk organic extracts, and that lower storage temperatures, and storing extracts dried rather than re-dissolved in methanol, result in a greater stability of metabolite profile. The scores plots in Figures 4.7 and 4.8 showed that the storage condition of extracts has a greater impact on the resulting metabolite profile than the storage temperature, and that dried extracts are more stable and less affected by storage temperature than those re-dissolved in methanol. The results of statistical tests in Table A.1 corroborated this. More compounds were statistically significant in abundance between extracts that were stored dried and re-dissolved, than between extracts that were stored at different temperatures, and most of these significant differences were found when extracts were stored at -25 °C. Most significant differences in compound abundance between extracts stored at different temperatures, were between extracts that were re-dissolved in methanol prior to storage, rather than those that were dried.

4.5 Yolk Organic Extracts: 12 weeks of Storage

Yolk organic extracts that were stored either dried or re-dissolved for 12 weeks, at three different temperatures, were compared to see whether storage condition affects the metabolite profile of an extract.

4.5.1 Results and Discussion

4.5.1.1 Quality Control Analysis

Table 4.3 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 4.3: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for 12 weeks*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	1195324	148814801	53806582	32840278	414676396	102728803
QC2	990121	147680808	47501712	23980003	321106340	60268247
QC3	1023021	96474264	35011757	22620912	279308663	43073610
QC4	729523	98113161	36854710	25312134	341966080	90328982
QC5	908633	106145274	36052987	28047237	352900642	62708273
QC6	969705	92755263	30139702	23867153	293959217	71670335
QC7	730983	84016867	58979142	31676917	470655274	67542800
SD	165478	26569881	10852033	4041002	67946189	19817018
Mean	935330	110571491	42620942	26906376	353510373	71188721
CV%	17.69	24.03	25.46	15.02	19.22	27.84

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 4.4 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 4.4: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts that had been stored for 12 weeks*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	5.323	9.771	33.359	57.717	72.750	80.219
QC2	5.388	9.787	33.408	57.842	72.948	80.367
QC3	5.339	9.821	33.492	57.926	73.131	80.734
QC4	5.400	9.815	33.519	57.885	73.109	80.662
QC5	5.388	9.770	33.308	57.759	72.881	80.235
QC6	5.389	9.754	33.292	57.737	72.899	80.302
QC7	5.386	9.818	33.539	58.020	73.427	81.113
SD	0.030	0.027	0.102	0.111	0.223	0.332
Mean	5.373	9.791	33.417	57.841	73.021	80.519
CV%	0.55	0.28	0.30	0.19	0.31	0.41

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

4.5.1.2 Metabolite Profiling

4.5.1.2.1 Comparison of storage temperature

Comparisons were made between extracts that were stored at the three different temperatures, by studying the TICs of the extracts following HPLC-MS analysis. Figure 4.9 shows overlaid TICs for the average chromatograms of extracts (both dried and re-dissolved) stored at the three different temperatures.

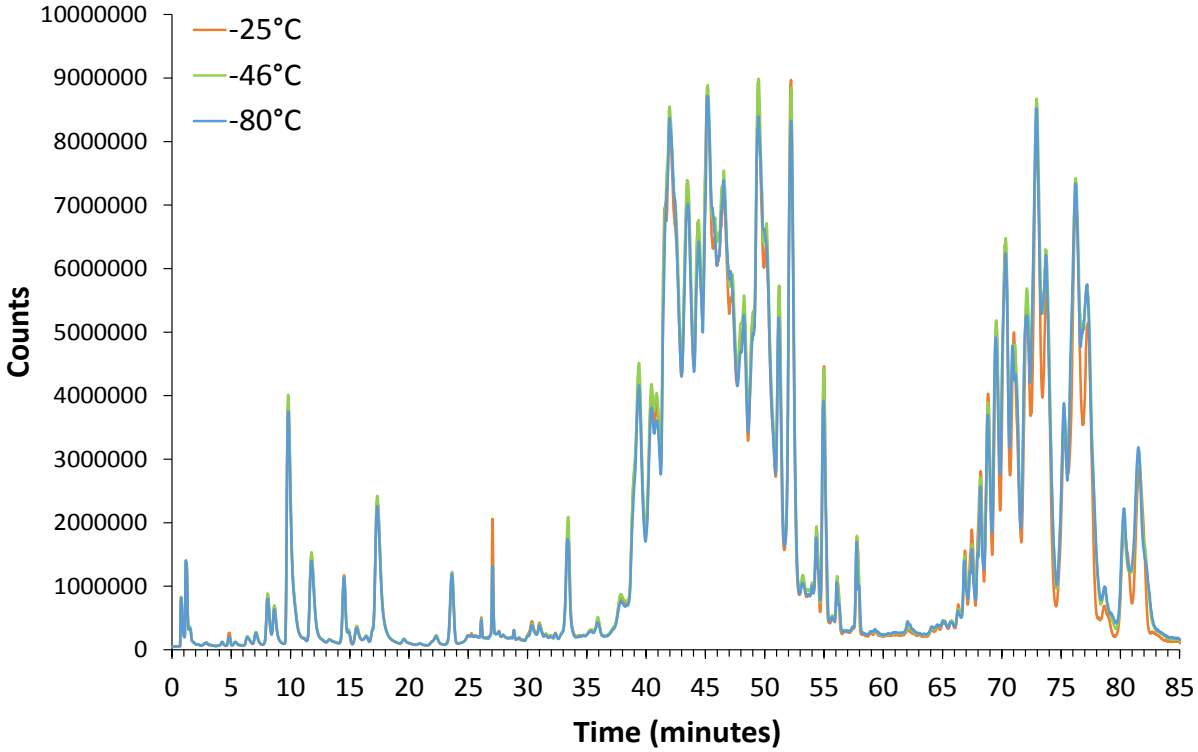


Figure 4.9: *Overlaid average TICs of yolk organic extracts that were stored for 12 weeks at -25 °C, -46 °C, and -80 °C*

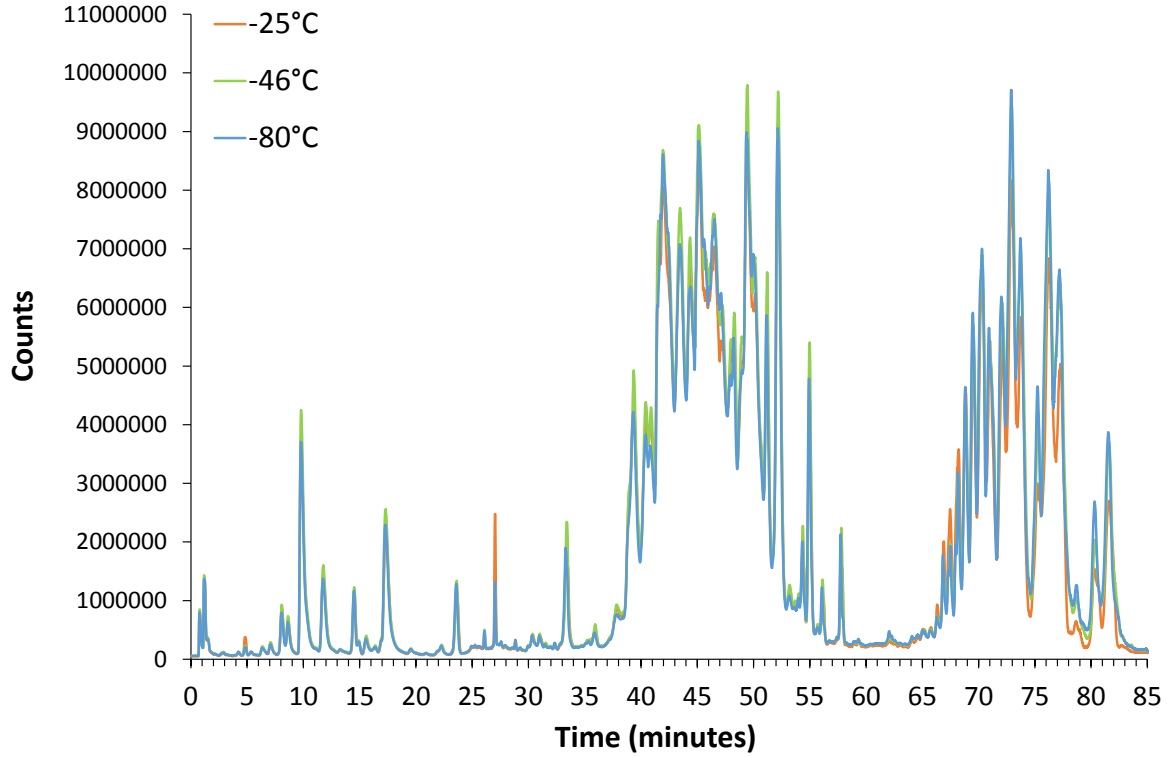
When observing these overlaid TICs, a slight difference can be seen between the chromatograms of the extracts that were stored at different temperatures for 12 weeks. Throughout the chromatograms, the peak intensity is greater for most peaks in the TIC of extracts that have been stored at -46 °C, than in the TICs of extracts that were stored at the other two temperatures. The peak intensity for most of the peaks up until approximately 70 minutes is greater in the TIC of extracts that were stored at -25 °C than in the TIC of those that were stored at -80 °C, at which point it becomes lower. This

chromatogram indicates that over 12 weeks of storage, the storage temperature of the extracts does have some effect on the resulting metabolite profiles.

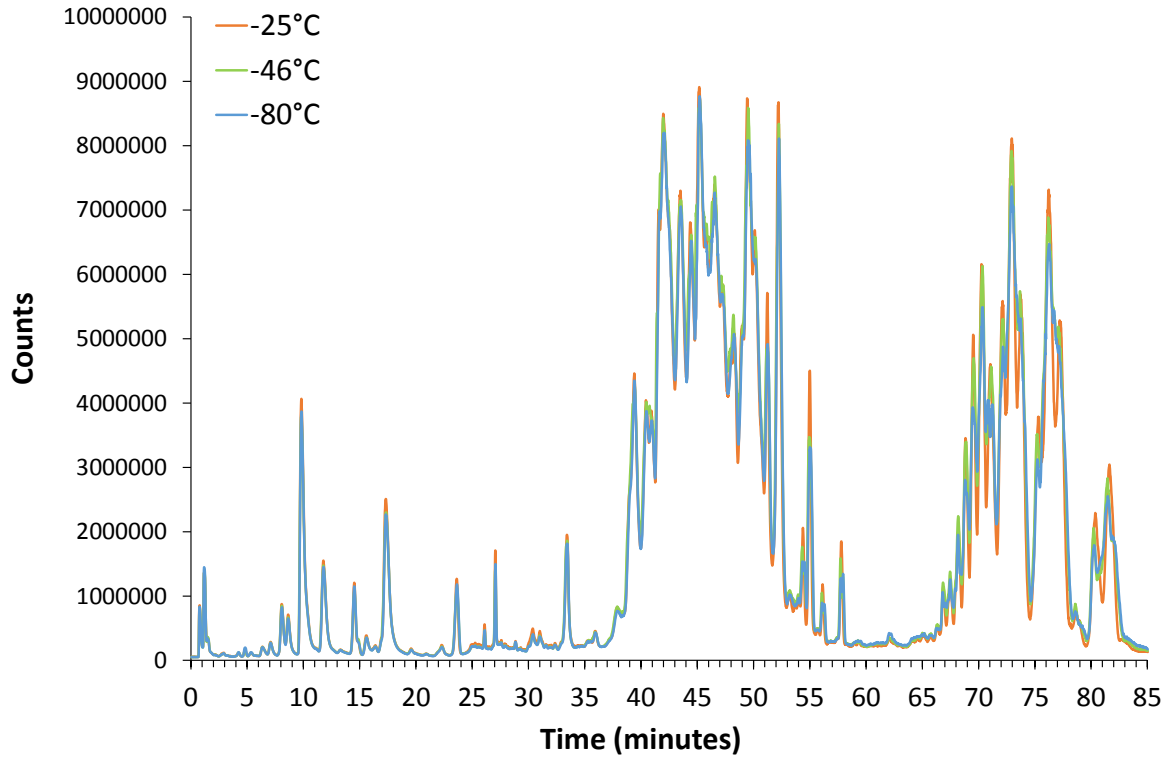
4.5.1.2.2 Comparison of temperature at each storage condition

The chromatograms in Figure 4.10 show comparisons between the TICs for re-dissolved (Figure 4.10 a)) and dried (Figure 4.10 b)) extracts stored at all three temperatures. The chromatogram in Figure 4.10 a) shows that until around 60 minutes of analysis, there is a higher peak intensity for most peaks in the TIC of re-dissolved extracts that were stored at -46 °C, compared to those of the extracts that were stored at the other two temperatures. The TIC of re-dissolved extracts that were stored at -25 °C shows the lowest peak intensity for most peaks, except for those at 5 and 27 minutes, which have the greatest peak intensity. After 60 minutes of analysis, the TIC of re-dissolved extracts that were stored at -80 °C shows the highest peak intensity for most peaks, compared to the TICs of extracts that were stored at the other two temperatures. The TIC of re-dissolved extracts that were stored at -25 °C again shows the lowest peak intensity for most peaks, except for those between 65 and 68 minutes, which have the highest peak intensity.

This greater peak intensity for the later-eluting, more non-polar compounds in the re-dissolved extracts that were stored at -80 °C, could be because these compounds may be larger lipid molecules which degrade during storage, producing more polar compounds. This degradation may occur at a lower rate in extracts that have been stored at lower temperatures, resulting in the non-polar compounds having a higher abundance in extracts stored at the lower temperatures compared to the higher temperatures. The higher abundance of the less non-polar compounds in the chromatograms of extracts that were stored at -46 °C could then be due to the higher rate of degradation of the more non-polar compounds at the higher storage temperature, resulting in the increased production of these less non-polar compounds compared to at -80 °C. This suggests that lower storage



(a) *Overlaid TICs for yolk organic extracts re-dissolved in methanol*



(b) *Overlaid TICs for yolk organic dried extracts*

Figure 4.10: *Overlaid TICs for yolk organic extracts stored at all three temperatures a) re-dissolved in methanol and b) dried, for 12 weeks*

temperatures provide better metabolite stability of extracts than higher storage temperatures. The peak intensity for the TIC of extracts that were stored at -25 °C may be the lowest throughout most of the analysis, as over the second 6 week storage period following the initial 6 week study the less non-polar compounds, as well as the more non-polar compounds, in the extracts stored at this temperature may have undergone further degradation, thereby decreasing in abundance.

The chromatogram in Figure 4.10 b) shows that the TIC of dried extracts that were stored at -25 °C has a higher peak intensity for most peaks throughout the entire chromatogram, compared to the TICs of the dried extracts stored at the other two temperatures. This could be due to a higher recovery yield of analytes following the re-dissolving of the dried extracts that were stored at -25 °C in methanol, than following the re-dissolving of the extracts that were stored at the other two temperatures. As there is no change in peak intensity difference due to compound polarity throughout the chromatogram, between the TICs of dried extracts stored at the three different temperatures, it indicates that storing the extracts dried, rather than re-dissolved in methanol, prevents, or slows down, the degradation of metabolites during storage. Therefore, storing extracts dried provides better metabolite stability than re-dissolving them in methanol prior to storage.

During the first 35 minutes of analysis, the TICs of dried extracts that were stored at -46 °C and -80 °C show a very similar peak intensity to each other, indicating that storing extracts at these two temperatures provides similar metabolite recovery. However, throughout the final 50 minutes of analysis, there is a more observable difference in peak intensity between the TICs of the dried extracts that were stored at these two temperatures, with the TIC of the extracts that were stored at -80 °C showing the lowest peak intensity. This suggests that the more non-polar compounds, that elute later on in the analysis, are more affected by storage temperature than the less non-polar compounds that elute earlier on in the analysis. However, by observing the latter 50 minutes of the TIC of dried extracts that were stored at -80 °C, it can be seen that the peak shape is

different compared to the other TICs. This is because there was some slight retention time drift throughout the analytical sequence, which affected the average chromatogram that was produced for dried extracts that were stored at -80 °C. Therefore, this lower peak intensity for the TIC of the dried extracts that were stored at -80 °C, compared to that of the dried extracts that were stored at -46 °C, could be due to the poor peak shape of the average chromatogram, rather than a reflection of compound abundance due to storage temperature.

4.5.1.2.3 Comparison of storage condition

Comparisons were also made between extracts that were stored dried and re-dissolved in methanol, by studying the TICs of the extracts following HPLC-MS analysis. Figure 4.11 shows overlaid TICs for the average chromatograms of extracts that were stored dried and re-dissolved in methanol at all three temperatures.

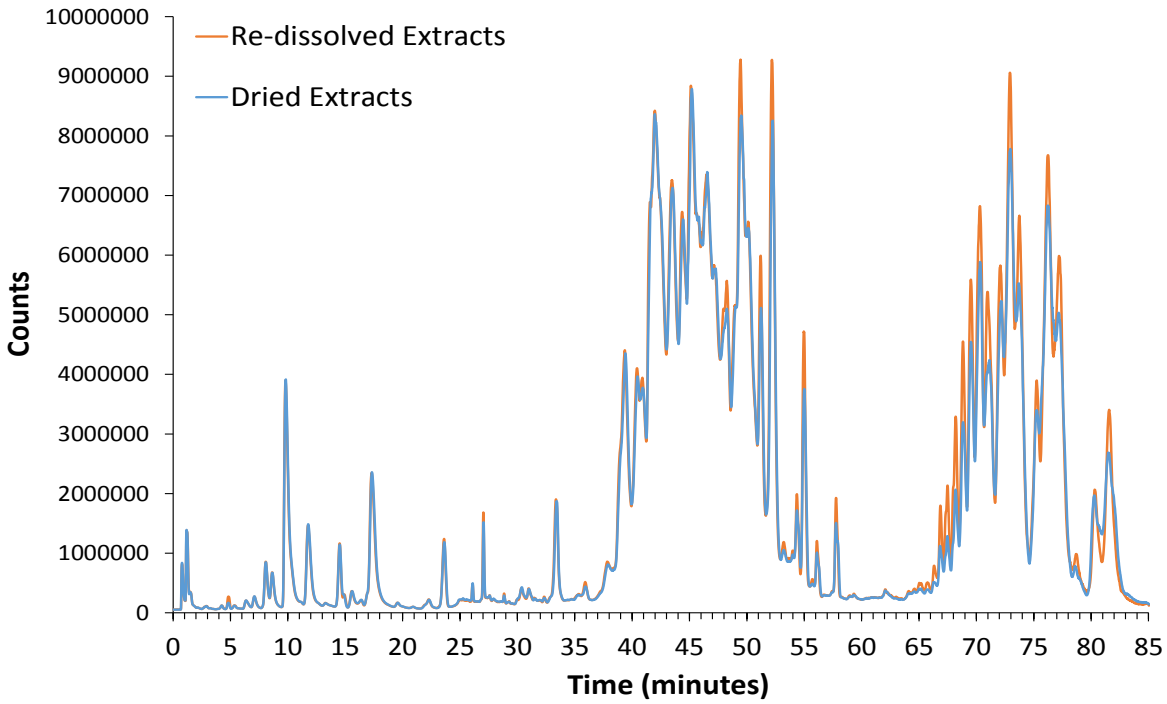


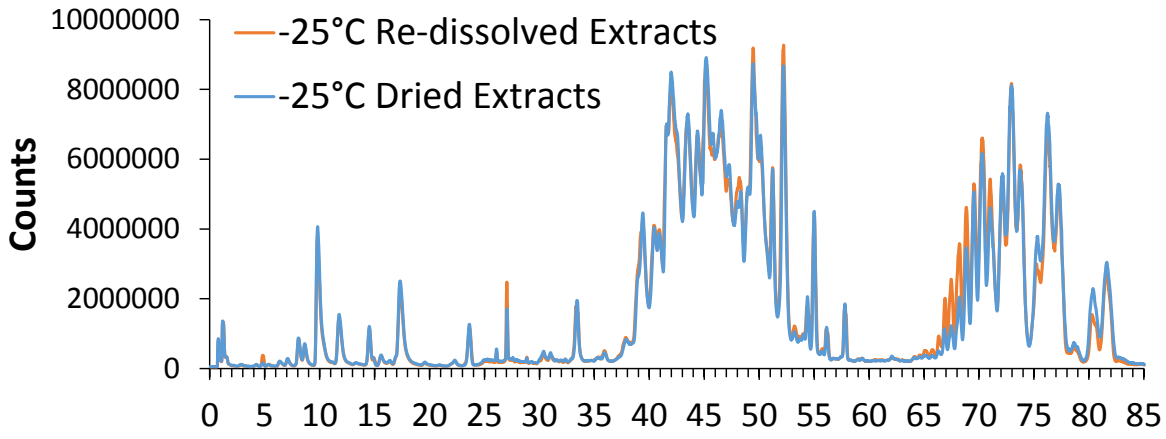
Figure 4.11: *Overlaid average TICs of yolk organic extracts that were stored dried or re-dissolved in methanol for 12 weeks*

This chromatogram shows that there is a clear difference in metabolite profile between extracts that were stored under the different conditions. The peak intensity is greater for almost all peaks after 20 minutes, in the TIC of extracts that were stored re-dissolved in methanol for 12 weeks, compared to that of extracts that were stored dried for 12 weeks. This could be due to incomplete recovery of the compounds in the dried extracts during re-dissolving, following 12 weeks of storage. As it is the peaks later on in the chromatogram that show the largest difference in intensity between extracts stored under the two different conditions, it indicates that it is the more non-polar compounds that are more greatly affected by storage condition following 12 weeks of storage.

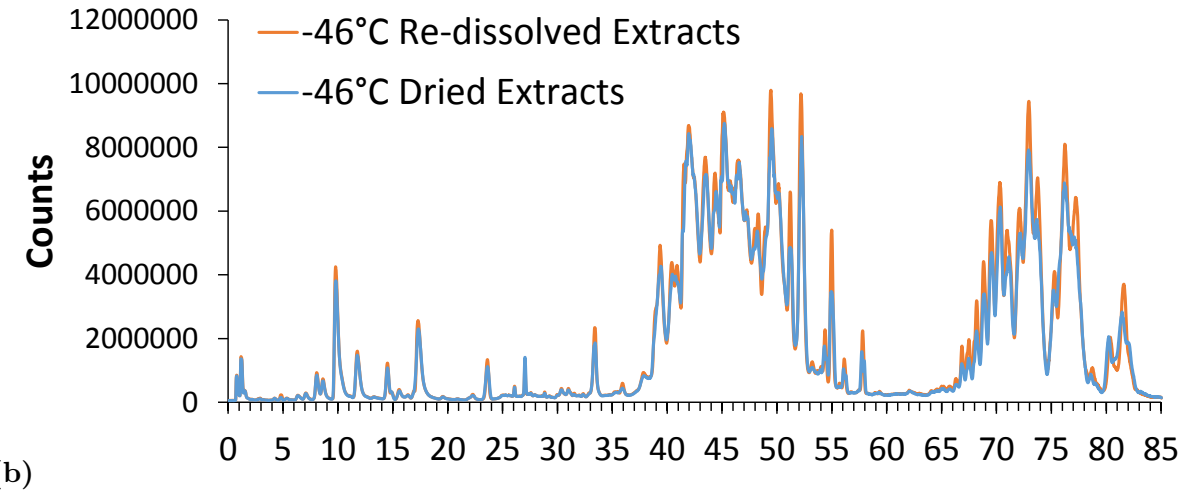
4.5.1.2.4 Comparison of storage condition at each storage temperature

Comparisons were also made between dried and re-dissolved extracts at each of the three different storage temperatures. Figure 4.12 shows the overlaid TICs for the average chromatograms of extracts that were stored dried and re-dissolved in methanol at -25 °C, -46 °C, and -80 °C for 12 weeks.

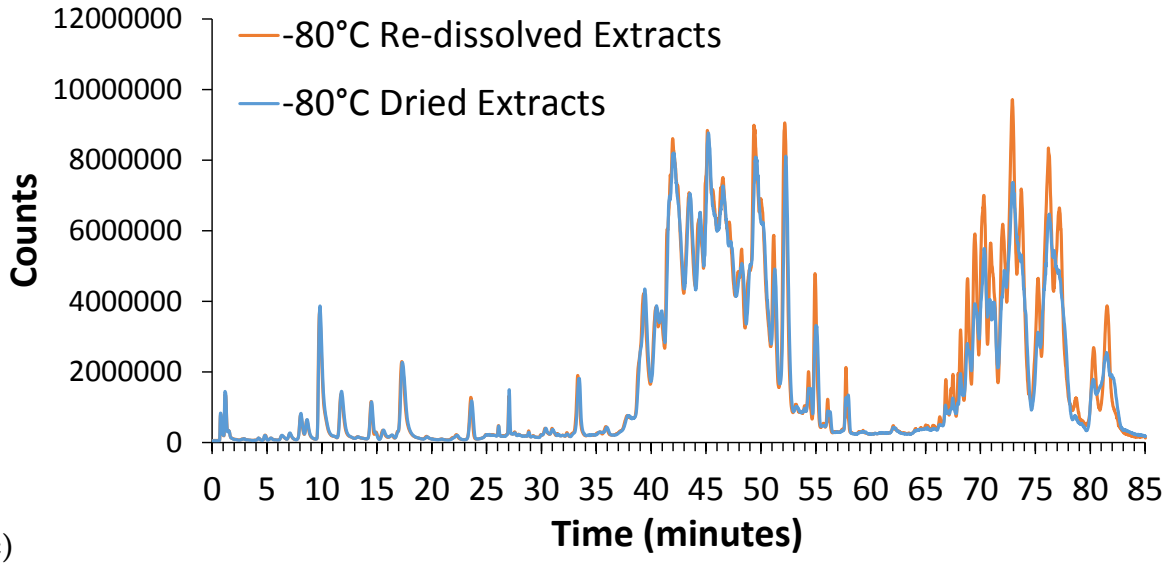
Figure 4.12 shows that there is some difference between the average chromatograms of dried and re-dissolved extracts that have been stored at all three temperatures for 12 weeks. Figures 4.12 a) and c) show that for extracts that were stored at -25 °C and -80 °C, there is very little difference in peak intensity for most peaks, between the chromatograms of dried and re-dissolved extracts during the first 45 minutes of analysis. This is similar to what was observed for extracts stored at these temperatures for just 6 weeks, when there was little difference between the chromatograms of dried and re-dissolved extracts during the first 25 minutes of analysis. This indicates that storage condition does not have much of an effect on the resulting metabolite profile of the less non-polar compounds, following 12 weeks of storage at these temperatures. Between 45-72 minutes of analysis, the chromatograms of extracts that were stored at -25 °C show a higher peak intensity in the TIC of re-dissolved extracts for most peaks, and after 72 minutes the opposite trend



(a)



(b)



(c)

Figure 4.12: *Overlaid TICs for dried and re-dissolved yolk organic extracts stored at a) -25 °C b) -46 °C and c) -80 °C for 12 weeks*

is observed, and there is a higher peak intensity in the TIC of dried extracts.

The greater peak intensity observed towards the end of the TIC of dried extracts that were stored at -25 °C compared to that of re-dissolved extracts, indicates that the storage condition of the extracts affects the more non-polar compounds in an opposite manner to the slightly more polar compounds, following 12 weeks of storage at this temperature, as it did following just 6 weeks of storage. This again could be due to the lower rate of degradation of the more non-polar compounds in dried extracts during storage, indicating that dried extracts experience greater metabolite stability compared to extracts that are re-dissolved in methanol prior to storage.

However, the chromatograms of extracts that were stored at -80 °C show a higher peak intensity in the TIC of re-dissolved extracts that were stored at -80 °C for almost all peaks after 45 minutes, compared to that of dried extracts. This trend was also observed in the chromatograms of extracts that were stored at -46 °C, in Figure 4.12 b), which showed a higher peak intensity for almost all peaks in the TIC of re-dissolved extracts compared to dried extracts. This, again, indicates that at a storage temperature of -46 °C, the polarity of the compounds does not influence how they are affected by the storage condition of the extracts. It could be that following 12 weeks of storage at this temperature, the metabolite recovery yield has a greater effect on the resulting metabolite profile than the difference in rates of degradation between dried and re-dissolved extracts.

The higher peak intensity that is observed throughout the second half of the TICs of re-dissolved extracts, compared to dried extracts when stored at -80 °C in Figure 4.12 c) was not expected, as based on previous results it was anticipated that at lower temperatures, the peak intensity difference between the TICs of extracts that were stored dried and re-dissolved would be less, and that the intensity of the peaks relating to more non-polar compounds would be higher for dried extracts compared to re-dissolved extracts, as the rate of degradation would be lower. However, it can be seen that the peaks in the TIC of dried extracts at -80 °C, as well as at -46 °C, have a different peak shape compared

to the other chromatograms. This is because retention time drift occurred throughout the analytical sequence, affecting the average chromatograms that were produced. This could have affected the observed peak intensities in the chromatograms, explaining the surprising result regarding the lower peak intensity in the TIC of dried extracts at -80 °C.

4.5.1.3 Multivariate Statistics

Following PCA, scores plots were produced to display the variation between samples, to observe whether there was any separation between the sets of extracts that were stored under different conditions, at different temperatures, for 12 weeks.

Figure 4.13 shows a scores plot of PC2 vs PC4. There is a clear separation between extracts that were stored dried, and those that were re-dissolved in methanol prior to storage, as indicated by the purple dashed line, and less variation between extracts that were stored at different temperatures. This indicates that over 12 weeks of storage, the storage condition of the extracts, i.e. dried or re-dissolved, has a greater impact on the resulting metabolite profile of the extracts, than temperature. The separation between the dried and re-dissolved extracts is across PC4, which describes 0.13% of the variance between all samples, showing the extreme subtlety of the differences between these extracts.

This plot also shows that there is some separation between the re-dissolved extracts that were stored at different temperatures. Although there is not a clear separation, and there is some variation within the extracts stored under the same condition and at the same temperature, three groups can be observed for the re-dissolved extracts stored at the three different temperatures. This indicates that temperature has a greater effect on the metabolite profiles of extracts that are re-dissolved in methanol prior to storage, compared to those that are stored dried, when stored for 12 weeks. The re-dissolved extracts that were stored at -80 °C are the closest to the centre of the plot, followed by

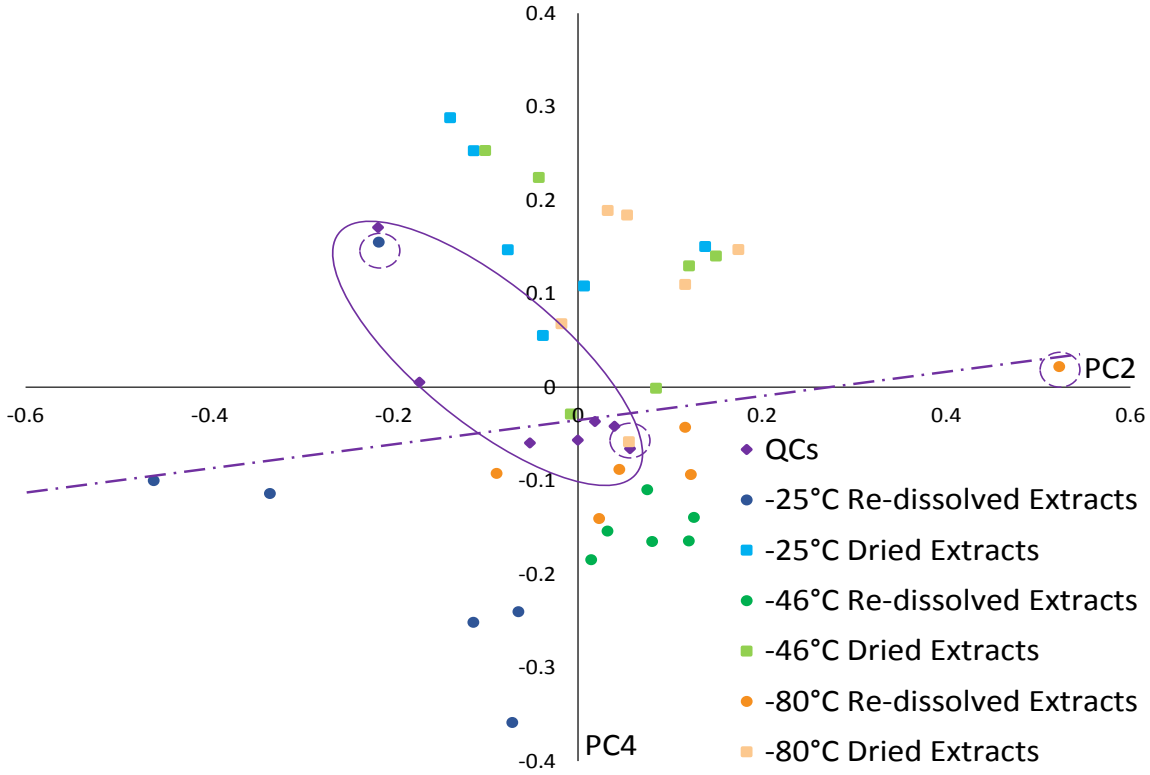


Figure 4.13: *PCA scores plot showing PC2 vs PC4 for yolk organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 12 weeks, including all compounds with CV%<30%. PC2 explains 0.29% of the variance, and PC4 explains 0.13% of the variance. Dashed line separates dried and re-dissolved extracts. Circled: samples separate to the rest of their sample sets.*

those that were stored at -46 °C, whilst those that were stored at -25 °C are the furthest out from the centre. As the re-dissolved extracts that were stored at -80 °C are the closest to the centre of the plot, and the QC samples, it indicates that they are the most representative of all of the yolk organic extracts. Therefore, if extracts are re-dissolved in methanol prior to storage, they should be stored at -80 °C to minimise any changes to the metabolite profile.

The plot in Figure 4.13 shows that there is most variation within the re-dissolved extracts that were stored at -25 °C for 12 weeks. This indicates that extracts stored under this condition, and at this temperature, are the most susceptible to metabolite degradation during storage. There are two extract samples that are observed on the opposite side of the plot to what would be expected (circled); one dried extract that was

stored at -80 °C, and one re-dissolved extract that was stored at -25 °C. These samples are grouped more closely to the extracts that were stored at the same temperature but under different conditions, than the extracts that were stored under the same conditions at the same temperature. There is also one re-dissolved extract that was stored at -80 °C (circled) that is quite separate to the rest of the extracts stored under the same condition at the same temperature. It is not known why these extracts show more separation from the other extracts in the same sample set, as nothing unusual was noticed during the extraction procedure or during the analysis.

The QC samples in this scores plot in Figure 4.13 are quite spread out, and show more variation than some of the sample sets themselves. This indicates that the analysis may not have been as robust as it could have been, and that there may have been some instrumental drift occurring throughout the analysis. This spread in QC samples could also be due to the subtlety of the differences between the extract sample sets; as the differences are so small, indicated by the amount of variation described by PC2 and PC4 (0.29% and 0.13% respectively), the PCA has found variation within the QC samples as well as between the sample sets. The two QC samples that show this spread of variation are the QC sample that was analysed immediately prior to sample analysis, and the final QC sample that was analysed immediately following the final extract sample. However, the analysis was still considered robust, as the QC analysis in section 4.5.1.1 showed the CV%s of the peak areas of all six peaks that were monitored throughout the seven QC samples to be within acceptable limits. The samples were also completely randomised prior to analysis, so any instrumental drift that may have been present would not have caused any bias in the analysis of the samples.

4.5.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 49 compounds remaining, after the removal of any duplicates, adducts or isotopes from the

top 100 compounds based on PC1 loadings, 46 were found to be statistically significant.

There are 22 compounds, of the top 46, that have been found to show significant differences between dried and re-dissolved extracts at at least one storage temperature, 11 compounds that have been found to show significant differences between extracts stored in the same condition but at different temperatures, and 10 compounds that have been found to show significant differences between dried and re-dissolved extracts at at least one storage temperature, as well as between extracts stored in the same condition but at different temperatures. The remaining 3 compounds have only been found to be significantly different between extracts stored under different temperatures *and* conditions. This again indicates that the storage condition of the extracts, i.e. dried or re-dissolved, has a greater effect on the resulting metabolite profile of the extracts following storage, than the storage temperature.

Of the 32 compounds that show a significant difference in abundance between dried and re-dissolved extracts; 28 are significantly different between the two storage conditions at -25 °C, 9 show a significant difference between the storage conditions at -46 °C, and 13 are significantly different between the storage conditions at -80 °C. This shows that the storage condition of the extracts has a reduced effect when extracts are stored at lower temperatures, indicating that it is best to store extracts at lower temperatures, to lessen any metabolite profile changes that may occur during storage. Of the 21 compounds that have been found to show a significant difference between extracts stored under the same conditions but at different temperatures, only one compound has been found to be significantly different between dried extracts stored at different temperatures; the other 20 compounds are all only statistically significant for re-dissolved extracts. This indicates that storage temperature has less of an effect on the resulting metabolite profile of an extract when it has been stored dried for 12 weeks, rather than re-dissolved in methanol.

The 43 compounds that are statistically significant when comparing extracts of the same storage condition but different temperatures, or the same temperature but different

storage conditions, can be found in Table A.2 in Appendix A. This table shows their CV%s, p-values resulting from ANOVA/Welch test, and figures representing the difference in the mean abundance of the compounds between the extracts. Those in bold were also found to be statistically significant when confirmed with the raw data.

As can be seen in Table A.2, there are various trends in the different abundances of the compounds between extracts stored under different conditions and at different temperatures. Some compounds have a much higher abundance in dried extracts, whereas others have a much higher abundance in re-dissolved extracts. This appears to be the greatest trend that is observed throughout the compounds; a difference in compound abundance between extracts stored dried and re-dissolved in methanol, rather than between extracts stored at different temperatures. This again indicates that the storage condition of the extract has a greater impact on the resulting metabolite profile, than the storage temperature.

The trends in the different abundances of these compounds are quite erratic. Several compounds show a large difference in abundance between dried and re-dissolved extracts at one storage temperature, then show an opposite difference in abundance between the two storage conditions at other temperatures. However, some of these compounds are only statistically significant due to different storage temperatures, which is why the storage condition appears to have no consistent impact on the differences in abundance between extracts. The rest of these compounds have been found to show significant differences in abundance between extracts due to both storage temperature and condition; clearly for these compounds, storage temperature had a greater effect on their abundance than the storage condition, explaining why there is no consistent difference in abundance between dried and re-dissolved extracts at different storage temperatures.

Of the 11 compounds that have been found to show significant differences in abundance between extracts due only to storage temperature, 10 of these show a correlation, mostly statistically significant, between abundance in re-dissolved extracts and storage temper-

ature. A correlation between abundance and storage temperature can be seen for only 6 of these compounds, and only one significantly so, in dried extracts. This again indicates that temperature has a stronger impact on the resulting metabolite profile of extracts that are stored re-dissolved in methanol, rather than dried. The remaining compound of these 11, *m/z* 905.8284, shows a large difference in abundance between re-dissolved extracts that were stored at -25 °C and those that were stored at -46 °C and -80 °C, but not between extracts that were stored at -46 °C and -80 °C. This indicates not only that re-dissolved extracts are more susceptible to metabolite profile changes due to different storage temperatures, but also that metabolite profiles of extracts remain more stable between different storage temperatures, when these temperatures are lower.

For the 10 compounds that have been found to show significant differences between extracts due to both storage condition and temperature, all of the observed trends in compound abundance due to storage temperature affect the re-dissolved extracts, rather than those that were dried prior to storage, again indicating that re-dissolved extracts are more affected by storage temperature than dried extracts. Two of these compounds; *m/z* 339.2891 and *m/z* 928.8331 show a large difference in abundance between the re-dissolved extracts stored at -25 °C and those stored at -46 °C and -80 °C, whilst the compound abundances in the dried extracts appear to remain similar to one another regardless of the storage temperature. This indicates not only that dried extracts are less affected by temperature than re-dissolved extracts, but also that the lower the storage temperature, the smaller the difference in compound abundance between extracts stored at different temperatures. Three of these compounds; *m/z* 860.7736, 871.7624, and 894.7560, all show a large difference in abundance between re-dissolved extracts that were stored at -25 °C and all other extracts. This indicates that for these compounds, storing the extracts re-dissolved in methanol at -25 °C results in a large difference in abundance compared to the other storage temperatures and conditions, making it the least stable method of extract storage.

4.5.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 46 statistically significant compounds, and scores plots were again produced to display the variation between yolk organic extracts that were stored either dried or re-dissolved in methanol, at the three different temperatures. Figure 4.14 shows a scores plot of PC2 vs PC3.

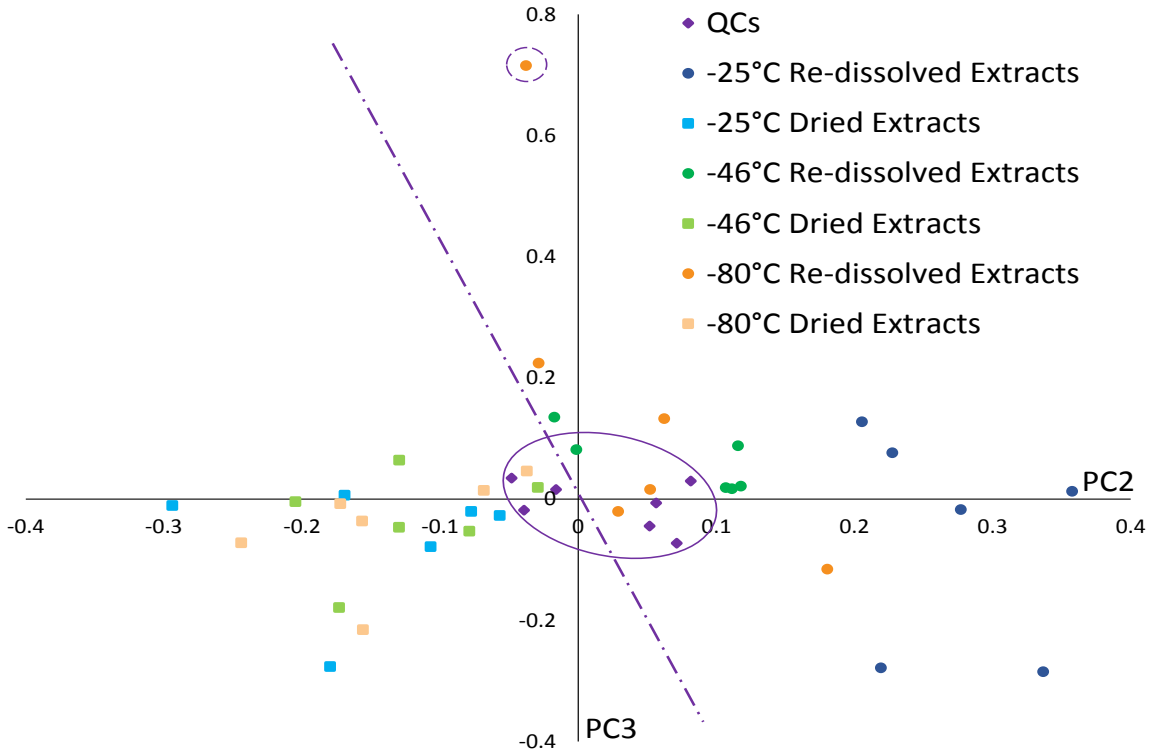


Figure 4.14: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 12 weeks, including the top 46 statistically significant compounds. PC2 explains 0.78% of the variance, and PC3 explains 0.24% of the variance. Dashed line separates dried and re-dissolved extracts. Circled: sample separate to rest of sample set.*

This scores plot shows a similar separation between extracts that were stored under different conditions, at different temperatures, to the plot observed in Figure 4.13, following PCA including all compounds with a CV% <30%. There is a clear separation between the dried extracts and those that were stored re-dissolved in methanol for 12 weeks, as indicated by the purple dashed line, with no overlap of extracts between the two conditions. The separation between dried and re-dissolved extracts is across PC2,

which describes 0.78% of the total variance between samples, compared to only 0.13% described by PC4 in the previous plot, showing that when PCA is carried out using only the top statistically significant compounds, there is a higher percentage of variance describing the difference between dried and re-dissolved extracts. However, in contrast to the previous scores plot, there is only a very slight separation within the re-dissolved extracts between extracts stored at the three different temperatures. This indicates that for the top 46 compounds included in this PCA, the amount of variation between dried and re-dissolved extracts is much greater than the variation between re-dissolved extracts stored at different temperatures, so this trend is not observed as clearly.

There is only one sample extract in this scores plot that shows separation to the other extracts stored under the same condition and at the same temperature (circled), and it is the same re-dissolved extract that was stored at -80 °C, that showed separation to other extracts in the previous plot in Figure 4.13. This indicates that there was potentially an anomaly in the metabolite extraction procedure that affected this extract sample. The other extracts that were observed to be on opposite sides of the plot to what was expected in Figure 4.13, are grouped with the rest of the extracts stored under the same condition and at the same temperature in this plot in Figure 4.14. This shows that by carrying out PCA using only the top statistically significant compounds, the difference between dried and re-dissolved extracts that is observed on the resulting scores plot is accentuated.

The QC samples in this plot in Figure 4.14 are grouped together much more closely than in the previous scores plot. This confirms that part of the reason why the QC samples were so spread out in the previous plot is due to the subtlety in the differences between the extract sample sets; when only the top statistically significant compounds are used, the difference between the sample sets is accentuated, and therefore the PCA does not find as much variation within the QC samples.

4.5.1.5 Summary

The overlaid TICs in Figures 4.9-4.12 showed that both storage temperature and storage condition have an effect on the resulting metabolite profiles of extracts, and that storing extracts dried, at lower temperatures, provides the best metabolite stability. The scores plots in Figures 4.13 and 4.14 showed a greater variation between extracts due to storage condition, rather than storage temperature. They showed some variation between re-dissolved extracts due to temperature, but no temperature based variation between dried extracts. The results of statistical tests seen in Table A.2 confirmed these observations. They showed that more compounds were statistically significant in abundance when comparing extracts stored under different conditions, rather than at different temperatures, particularly when extracts were stored at -25 °C. Most compounds that were significantly different in abundance between extracts that were stored at different temperatures, were found in extracts that were re-dissolved in methanol, rather than dried. Fewer compounds were statistically significant in abundance when comparing extracts that were stored at -46 °C and -80 °C, than when comparing extracts that were stored at these temperatures with those that were stored at -25 °C.

4.6 Yolk Organic Extracts: Comparison of compounds between 6 and 12 weeks of storage

The significant differences of the top statistically significant compounds between the yolk organic extracts stored under different conditions at the three different temperatures were compared between 6 weeks and 12 weeks of storage, to observe how the significance changed with increasing storage time.

4.6.1 Results and Discussion

4.6.1.1 Top statistically significant compounds following six weeks of storage

Table A.3 in Appendix A shows the ANOVA/Welch test p-values of the top 27 compounds that were found to be significantly different, following confirmation using the raw data, between the yolk organic extracts of eggs stored under different conditions and at different temperatures following 6 weeks of storage, as well as the p-values of these compounds after 12 weeks of storage.

Of these 27 compounds, 15 were found to not be significantly different between yolk organic extracts that were stored under different conditions and at different temperatures for 12 weeks. This could be because the metabolite degradation affecting the abundances of these compounds may have occurred at different rates in the extracts that were stored under different conditions and at different temperatures. After just 6 weeks of storage, this resulted in a significant difference in compound abundance. However, by 12 weeks of storage the degradation processes occurring at faster rates may have slowed down due to a reduced number of molecules, and those occurring at slower rates may have eventually reached the same point, resulting in these compounds being present in a similar abundance between all extracts.

Twelve of these compounds were found to be significantly different between yolk or-

ganic extracts that were stored under different conditions and at different temperatures for 12 weeks, as well as for 6 weeks. This indicates that for these compounds, an increased storage time from 6 to 12 weeks does not appear to have an effect on their relative abundances between extracts that were stored under different conditions and at different temperatures. This could be because these compounds have a different stability to the other compounds, when present in extracts stored under different conditions. Under some conditions and temperatures their abundances may have been affected by metabolite degradation during storage, whereas under other storage conditions and temperatures they may have remained stable and experienced little to no difference in abundance due to metabolite degradation. Therefore after 12 weeks of storage, as well as 6 weeks of storage, the differences in abundance of these compounds between the extracts stored under different conditions and temperatures remained statistically significant.

4.6.1.2 Top statistically significant compounds following twelve weeks of storage

Table A.4 in Appendix A shows the ANOVA/Welch test p-values of the top 12 compounds that were found to be significantly different, following confirmation using the raw data, between the yolk organic extracts of eggs stored under different conditions and at different temperatures following 12 weeks of storage, as well as the p-values of these compounds after just 6 weeks of storage.

Only one of these compounds was found to not be statistically significant when comparing yolk organic extracts that were stored under different conditions and at different temperatures for just 6 weeks. This suggests that for this compound, the length of time of extract storage does affect the relative abundance between extracts stored under different conditions and at different temperatures. This could be because after only 6 weeks of extract storage, there may have been very little metabolite degradation affecting the abundance of this compound in any of the extracts. However, following a further 6 weeks

of storage, there may have been more metabolite degradation in some of the extracts stored under certain conditions and temperatures, resulting in a significant difference in the abundance of this compound between these extracts and those that were stored under other conditions and temperatures.

The rest of these compounds were statistically significant after only 6 weeks of extract storage, as well as after the full 12 weeks. This suggests that the metabolite degradation affecting the abundances of these compounds had already begun in extracts stored under some conditions and temperatures, after just 6 weeks of storage. This indicates that these compounds are not stable during even short term extract storage under certain conditions and temperatures.

4.7 Albumen Organic Extracts: 44 weeks of storage

Albumen organic extracts that were stored either dried or re-dissolved, at three different temperatures, for 44 weeks were compared to observe if and how long term storage, under different temperatures and conditions, affects the metabolite profile of an extract.

4.7.1 Results and Discussion

4.7.1.1 Quality Control Analysis

Table 4.5 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 4.5: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts that had been stored for forty-four weeks*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	13050477	16697134	33108106	1620097	1794590	5442613
QC2	12615777	17654602	31493708	1834811	1671667	6936308
QC3	14403817	17697206	34338316	1701872	1778483	6506407
QC4	13449121	17210864	32774138	1801302	1648707	6545390
QC5	13292821	17117644	33183207	1933109	1533665	6060444
QC6	13904515	17165798	32658038	1995869	1627966	6171043
QC7	14005298	17831378	35091813	2151064	1662589	6996809
SD	613494	403683	1173401	180330	89487	541254
Mean	13531689	17339232	33235332	1862589	1673952	6379859
CV%	4.53	2.33	3.53	9.68	5.35	8.48

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 4.6 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 4.6: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts that had been stored for forty-four weeks*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	8.210	12.357	15.286	16.969	19.441	22.875
QC2	8.203	12.383	15.287	16.979	19.451	22.885
QC3	8.220	12.401	15.304	16.979	19.451	22.885
QC4	8.202	12.399	15.303	16.978	19.466	22.917
QC5	8.222	12.403	15.325	16.998	19.470	22.937
QC6	8.265	12.462	15.380	17.057	19.529	22.963
QC7	8.304	12.468	15.388	17.080	19.535	22.986
SD	0.038	0.041	0.043	0.044	0.038	0.043
Mean	8.232	12.410	15.325	17.006	19.478	22.921
CV%	0.46	0.33	0.28	0.26	0.20	0.19

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

4.7.1.2 Metabolite Profiling

4.7.1.2.1 Comparison of storage temperature

Comparisons were made between extracts that were stored at the three different temperatures, by studying the TICs of the extracts following HPLC-MS analysis. Figure 4.15 shows overlaid TICs for the average chromatograms of extracts (both dried and re-dissolved) stored at the three different temperatures.

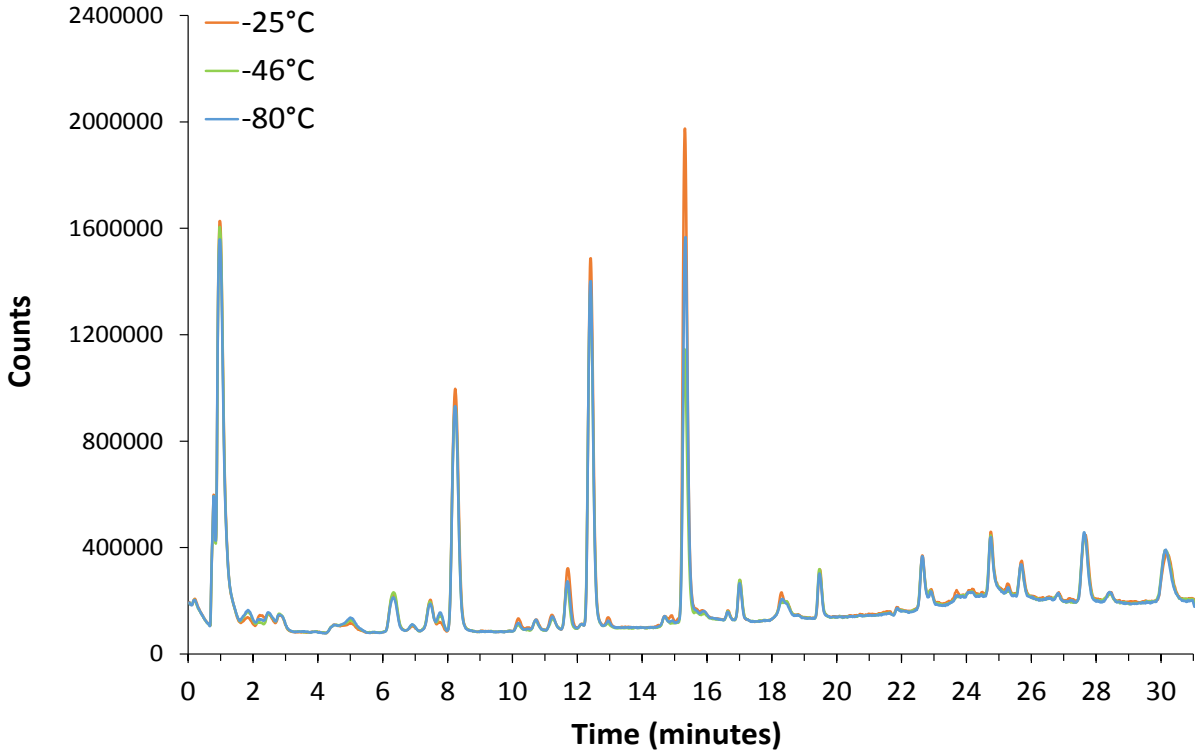


Figure 4.15: *Overlaid average TICs of albumen organic extracts that were stored at -25 °C, -46 °C, and -80 °C for 44 weeks*

By observing these overlaid TICs in Figure 4.15, a clear difference in metabolite profile can be seen between albumen organic extracts that were stored at the three different temperatures. The peak intensity for most peaks is highest in the TIC of extracts that were stored at -25 °C, particularly for the four largest peaks, which show the greatest difference in peak intensity. There is some slight difference in peak intensity between the TICs of extracts that were stored at -46 °C and -80 °C; some peaks appear to have a higher intensity in the TIC of extracts that were stored at -46 °C, whilst others have a higher

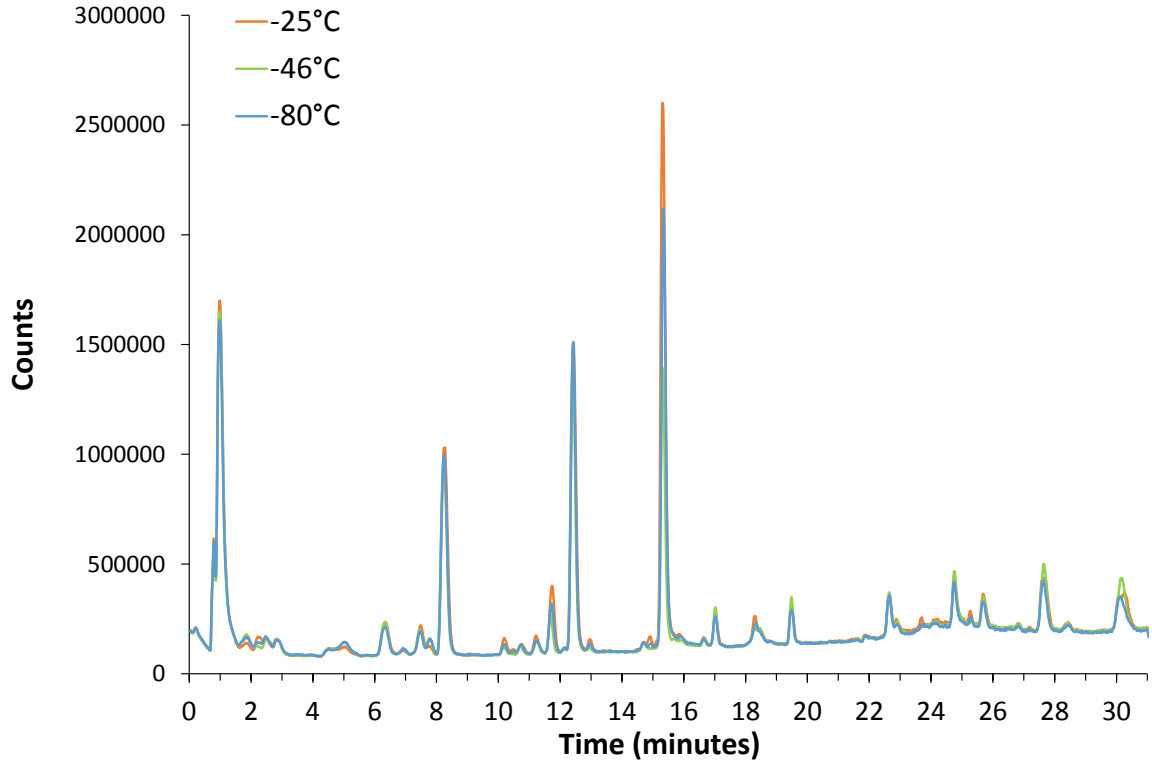
intensity in the TIC of extracts that were stored at -80 °C. As the greatest difference that can be observed is between the TICs of extracts that were stored at -25 °C, and those that were stored at -46 °C and -80 °C, it indicates that the metabolite profiles of extracts become more stable between extracts stored at different temperatures, as the storage temperature decreases.

4.7.1.2.2 Comparison of temperature at each storage condition

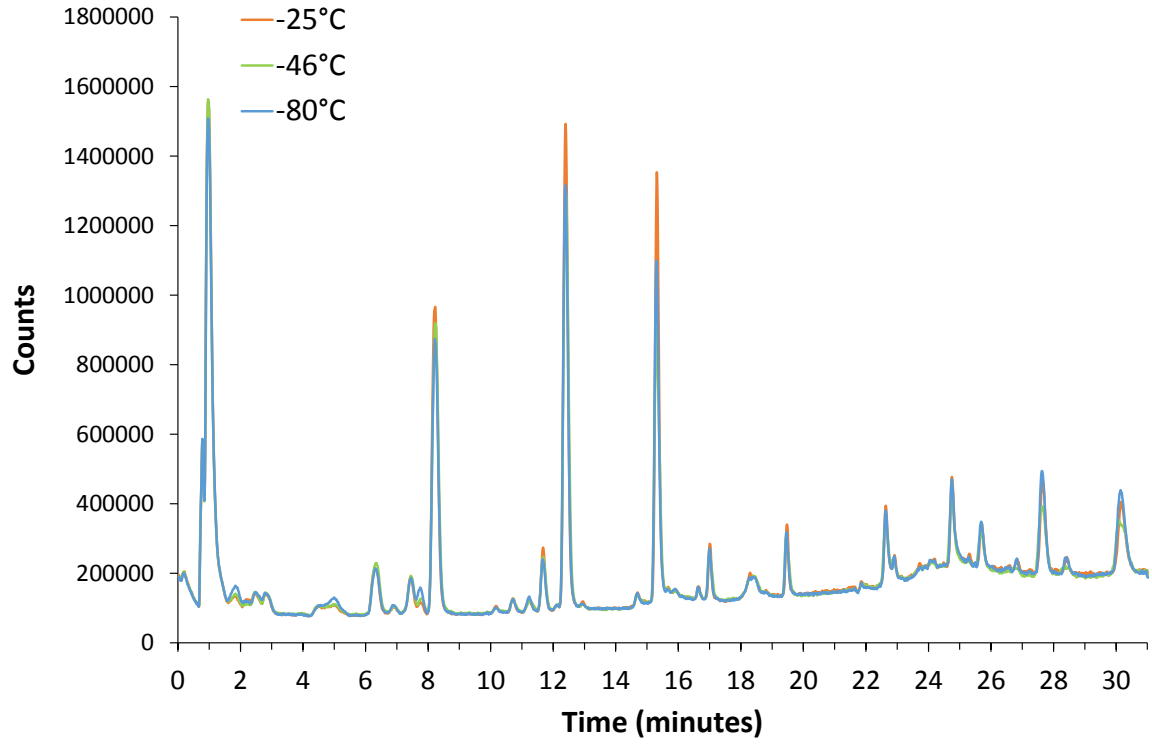
The chromatograms in Figure 4.16 show comparisons between the TICs of re-dissolved (Figure 4.16 a)) and dried (Figure 4.16 b)) extracts stored at all three temperatures. The chromatogram in Figure 4.16 a) shows a higher peak intensity for most peaks in the TIC of re-dissolved extracts that were stored at -25 °C, in the first 16 minutes of analysis, compared to the TICs of the extracts that were stored at -46 °C and -80 °C. The peak intensity for some of the peaks is greater in the TIC of extracts that were stored at -46 °C compared to that of extracts that were stored at -80 °C, whilst other peaks show a higher intensity in the TIC of extracts that were stored at -80 °C. Throughout the remaining analysis time, there is a higher intensity for most of the peaks in the TIC of extracts that were stored at -46 °C, with the TIC of extracts that were stored at -25 °C showing a higher intensity for some peaks.

A similar trend was observed in the chromatogram in Figure 4.16 b), however the higher peak intensity for most peaks in the TIC of dried extracts that were stored at -25 °C was observed up until 25 minutes, rather than 16 minutes. The final few peaks all show a very slightly higher peak intensity for the TICs of extracts that were stored at -80 °C, compared to those of the extracts stored at the other two temperatures.

The reason for the TICs of extracts that were stored at -25 °C showing a lower intensity for the peaks relating to the later eluting, more non-polar compounds compared to the TICs of extracts that were stored at the other two temperatures, could be due to these compounds undergoing a higher rate of degradation during storage, compared to the



(a) *Overlaid TICs for albumen organic extracts re-dissolved in methanol*



(b) *Overlaid TICs for albumen organic dried extracts*

Figure 4.16: *Overlaid TICs for albumen organic extracts stored at all three temperatures a) re-dissolved in methanol and b) dried, for 44 weeks*

compounds in extracts that were stored at the lower temperatures. This would result in a lower abundance of these compounds, and therefore a higher abundance of the earlier eluting, less non-polar compounds, which was observed by the higher peak intensity for the peaks relating to these less non-polar compounds. This suggests that the storage of extracts at lower temperatures provides a greater metabolite stability than storage at higher temperatures.

4.7.1.2.3 Comparison of storage condition

Comparisons were also made between extracts that were stored dried, and those that were re-dissolved prior to storage, by studying the TICs of the extracts following HPLC-MS analysis. Figure 4.17 shows overlaid TICs for the average chromatograms of the dried and re-dissolved extracts.

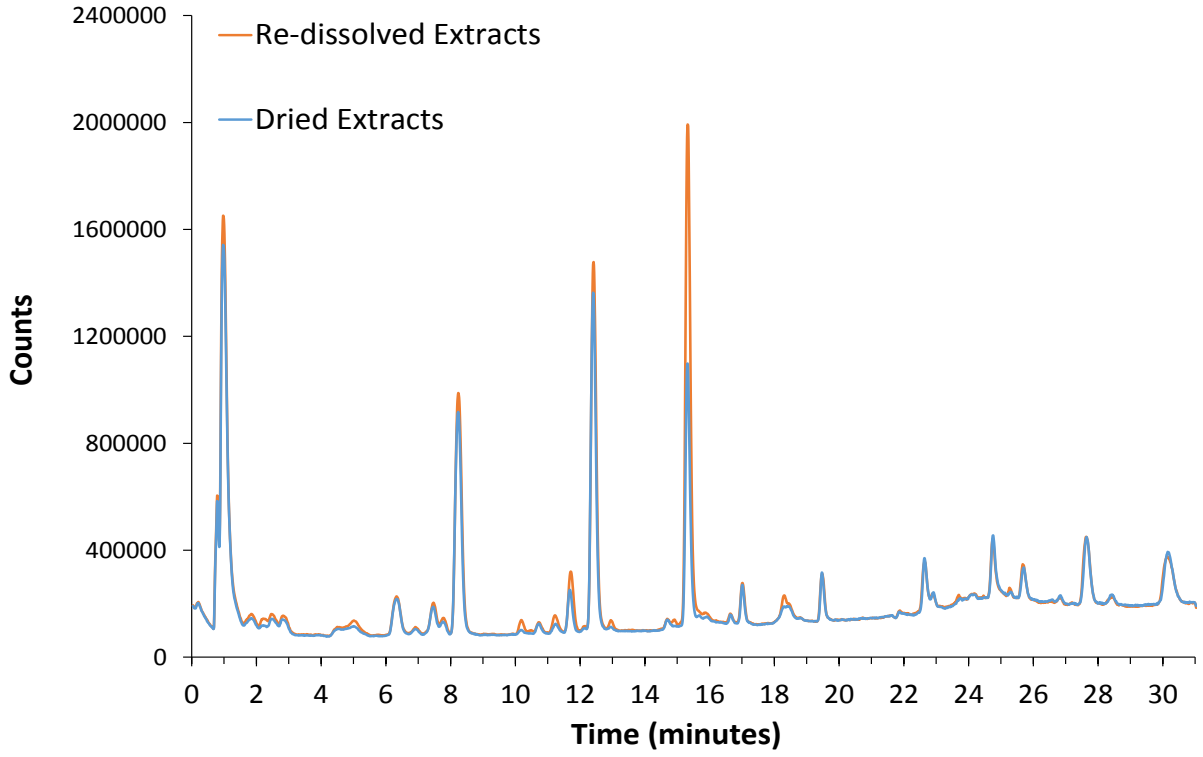


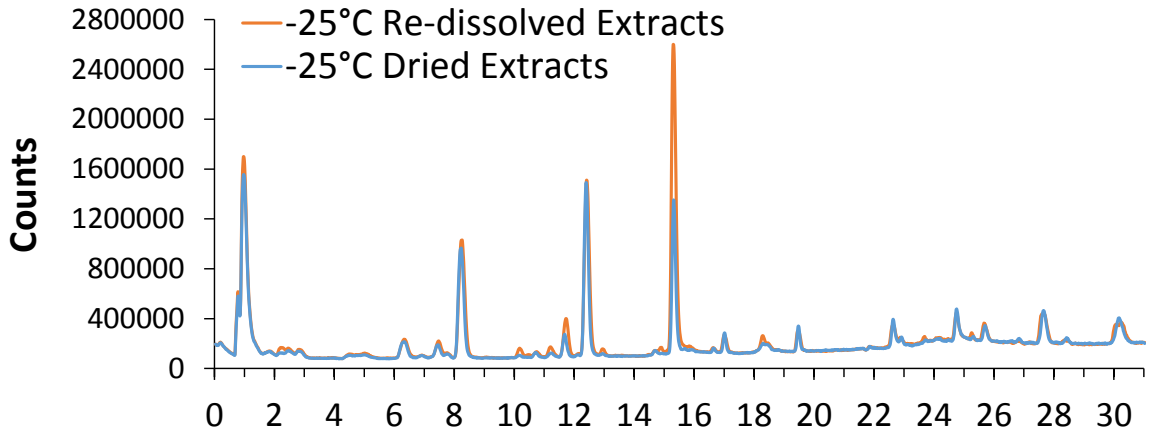
Figure 4.17: *Overlaid average TICs of albumen organic extracts that were stored dried and re-dissolved for 44 weeks*

It can be seen from the chromatograms in Figure 4.17, that there is a clear difference in metabolite profile between albumen organic extracts that were stored dried, and those that were re-dissolved prior to storage. The peak intensity is greater for most peaks in the TIC of re-dissolved extracts, compared to that of dried extracts, again particularly for the four largest peaks in the chromatograms, which show the greatest difference in intensity between dried and re-dissolved extracts. This could be due to incomplete compound recovery during the re-dissolving of dried extracts in methanol, following storage. There is less difference in peak intensity for the peaks in the latter half of the chromatograms, indicating that the more non-polar compounds are less affected by storage condition, than the more polar compounds.

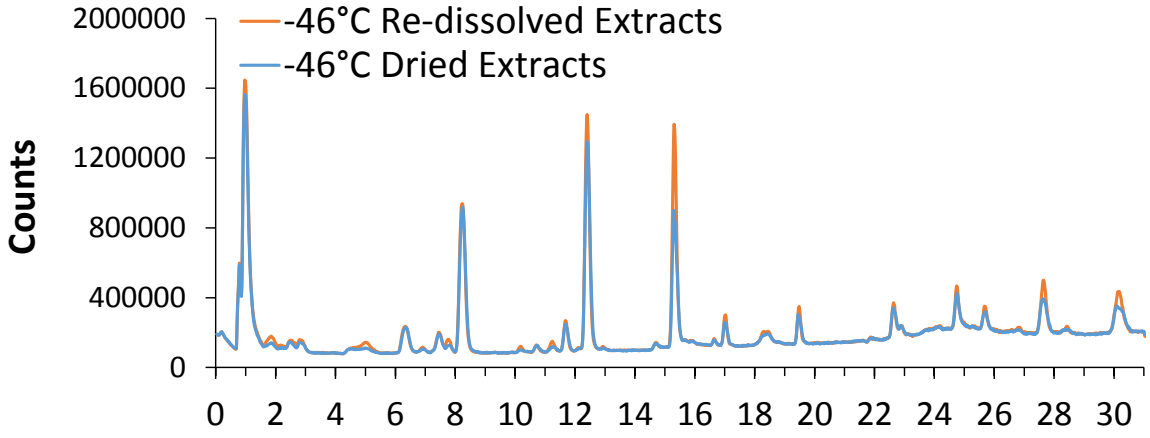
4.7.1.2.4 Comparison of storage condition at each storage temperature

Comparisons were also made between dried and re-dissolved extracts at each of the three different storage temperatures. Figure 4.18 shows the overlaid TICs for the average chromatograms of extracts that were stored dried and re-dissolved in methanol at -25 °C, -46 °C, and -80 °C for 44 weeks.

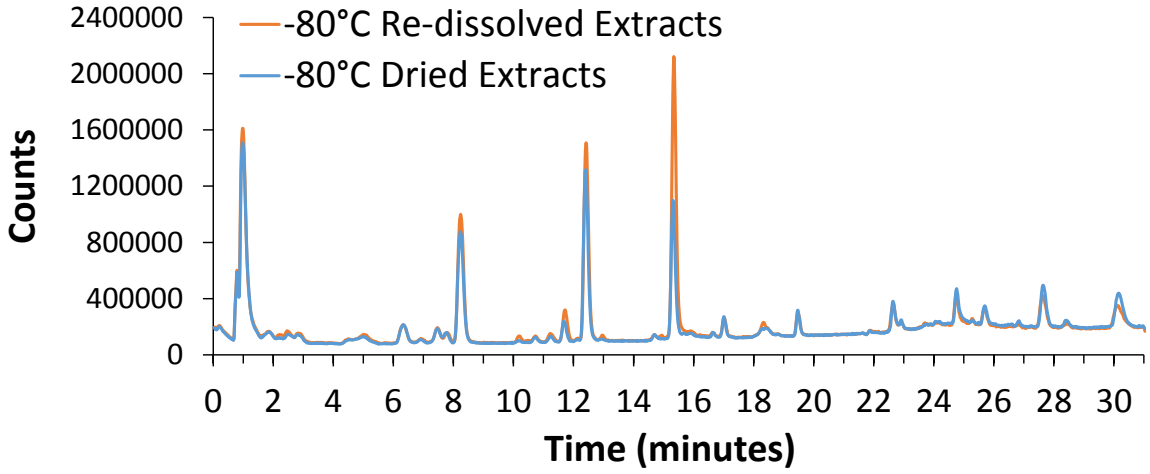
The chromatograms in Figures 4.18 a) and c) for extracts that were stored at -25 °C and -80 °C, show very similar differences in peak intensity between the TICs of dried and re-dissolved extracts. During the first 16 minutes of analysis there is a higher peak intensity for most peaks in the TIC of re-dissolved extracts compared to dried extracts, however throughout the remaining analysis time there is a higher peak intensity in the TICs of dried extracts. This suggests that the later eluting, more non-polar compounds were degraded at a faster rate in the re-dissolved extracts than the dried extracts, resulting in a lower abundance of these compounds, as evidenced by the lower peak intensity. This will have resulted in the observed higher abundance of the less non-polar, earlier eluting compounds in the re-dissolved extracts, as these will be the degradation products of the more non-polar compounds. This suggests that storing extracts dried, rather than



(a)



(b)



(c)

Figure 4.18: *Overlaid TICs for dried and re-dissolved albumen organic extracts stored at a) -25 °C b) -46 °C and c) -80 °C for 44 weeks*

re-dissolved in methanol, provides a better metabolite stability.

The chromatogram in Figure 4.18 b) shows a higher peak intensity for most peaks throughout the whole analysis in the TIC of re-dissolved extracts compared to dried extracts. This indicates that at this temperature, the polarity of the compounds does not influence how they are affected by the storage condition of the extracts. It could be that at this storage temperature, the difference in recovery yield of compounds during the re-dissolving of extracts in methanol has a greater effect on the resulting metabolite profile than the difference in rates of degradation between dried and re-dissolved extracts.

4.7.1.3 Multivariate Statistics

Following PCA, scores plots were produced to display the variation between samples, to observe whether there was any separation between the sets of albumen organic extracts that were stored under different conditions and at different temperatures, for 44 weeks. Figure 4.19 shows a scores plot of PC3 vs PC4.

As can be seen in Figure 4.19, contradictory to what was observed in the chromatograms in 4.15 and 4.17, there is no separation on the scores plot between extracts that were stored under different conditions and at different temperatures. This indicates that the differences between random extracts are greater than any differences that were observed in the previous chromatograms, between extracts that were stored under different conditions and at different temperatures. This shows that for albumen organic extracts, the effect that storage condition and temperature have on the resulting metabolite profiles is very subtle.

The QC samples on the scores plot in Figure 4.20 are grouped together more tightly than the samples from other sample sets, indicating that the analysis was robust and that there was no instrumental drift affecting the analysis.

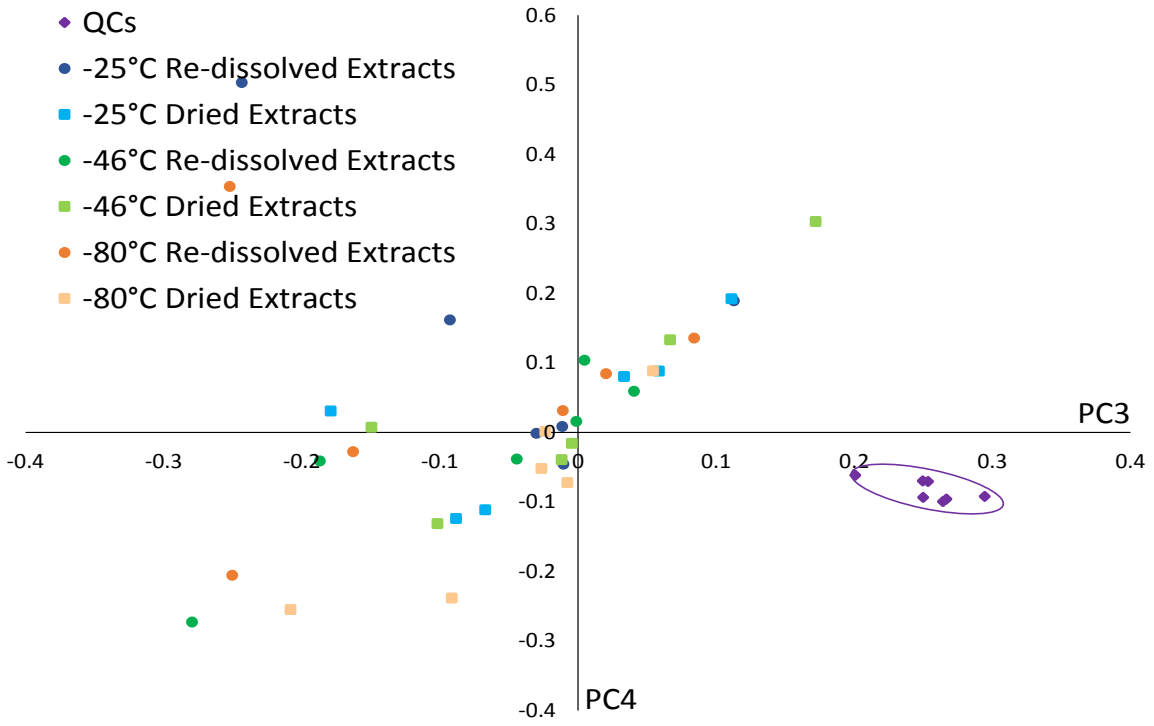


Figure 4.19: *PCA scores plot showing PC3 vs PC4 for albumen organic extracts that were stored dried and re-dissolved at -25 °C, -46 °C, and -80 °C for 44 weeks, including all compounds with CV%<30%. PC3 explains 0.34% of the variance, and PC4 explains 0.23% of the variance.*

4.7.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 39 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, 33 were found to be statistically significant.

There are 8 compounds, of the top 33, that have been found to show significant differences between dried and re-dissolved extracts at at least one storage temperature, 3 compounds that have been found to show significant differences between extracts that were stored in the same condition but at different temperatures, and 12 compounds that have been found to show significant differences both between dried and re-dissolved extracts at at least one storage temperature, as well as between extracts stored in the same condition but at different temperatures. The remaining 10 compounds have only been

found to be significantly different between extracts from different storage temperatures *and* conditions. This indicates that storage condition, i.e. dried or re-dissolved, has a slightly greater effect on the resulting metabolite profiles of extracts than temperature, and that most compounds, if they are significantly affected by storage temperature, are likely to also be affected by storage condition. As there are a high proportion of these top statistically significant compounds that do not show any statistical significance due to storage condition *or* temperature, it indicates that the differences in metabolite profile, between extracts stored under different conditions, and at different temperatures, are very small.

Of the 20 compounds that show a significant difference in abundance between dried and re-dissolved extracts; 13 are significantly different between dried and re-dissolved extracts when stored at -25 °C, one compound shows statistical significance when extracts are stored at -46 °C, and 8 are significantly different between dried and re-dissolved extracts when stored at -80 °C. This indicates that storage condition, i.e. dried or re-dissolved, has a greater effect on the compound abundances in extracts when they have been stored at -25 °C, compared to at the two lower temperatures, indicating that it is better to store extracts at a lower storage temperature to improve stability. Of the 15 compounds that have been found to show a significant difference between extracts stored under the same conditions but at different temperatures, only one of these compounds is statistically significant for dried extracts; the other 14 compounds only show significant differences between re-dissolved extracts stored at the three different temperatures. This indicates that storing extracts dried, rather than re-dissolved in methanol, improves the stability of the extracts across different storage temperatures, resulting in more uniform metabolite profiles between the extracts.

The 23 compounds that are statistically significant when comparing extracts stored under the same storage condition but different temperatures, or the same temperature but different storage conditions, can be found in Table A.5 in Appendix A. This table

shows their CV%s, p-values resulting from ANOVA/Welch test, and figures representing the difference in the mean abundance of the compounds between the extracts. Those in bold were also found to be statistically significant when confirmed with the raw data.

As can be seen in Table A.5, there are various trends in the different abundances of the compounds between extracts stored under different conditions and at different temperatures. Some compounds have a much higher abundance in dried extracts, whereas others have a much higher abundance in re-dissolved extracts, and some compounds show a large difference in abundance between dried and re-dissolved extracts at one storage temperature, then show an opposite difference in abundance between the two storage conditions at the other temperatures. Of the 13 compounds that show an opposite trend in abundance between dried and re-dissolved extracts at the different temperatures, 3 of these were only statistically significant due to storage temperature, whilst the other 10 were significantly different between extracts that were stored under different conditions, as well as between extracts stored at different temperatures. Clearly, for these 10 compounds, the effect of storage temperature had a greater impact on compound abundance than the storage condition, explaining the lack of consistent difference in abundance between dried and re-dissolved extracts at the different storage temperatures.

Of the 15 compounds that have been found to show significant differences in abundance between extracts due to storage temperature, 11 of these show a correlation between abundance and storage temperature in re-dissolved extracts. This includes one compound, m/z 934.6395, which shows an opposite correlation between abundance and storage temperature in dried extracts compared to re-dissolved extracts. One of the remaining 4 compounds, m/z 617.5126, shows a correlation between abundance and storage temperature in dried extracts, but not in re-dissolved extracts. As most correlation is observed between compound abundance and storage temperature in re-dissolved extracts, this indicates that storage temperature has a stronger impact on the resulting metabolite profile of extracts that are stored re-dissolved in methanol, rather than dried. Although

the compound *m/z* 617.5126 shows no correlation between abundance and storage temperature for re-dissolved extracts, it is highly significantly different in abundance between re-dissolved extracts that were stored at -25 °C and -46 °C, and those that were stored at -80 °C. The remaining 3 compounds, *m/z* 294.1547, 328.1362, and 371.3144, also show no correlation between abundance and storage temperature in re-dissolved extracts, or in dried extracts, but they do show a large significant difference in abundance between re-dissolved extracts that were stored at -25 °C, and all other dried and re-dissolved extracts. This indicates that re-dissolving albumen organic extracts in methanol, then storing them at -25 °C, results in a large difference in abundance for some compounds, compared to all other methods of storage explored in this experiment, and thus is the least reliable storage condition.

4.7.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 33 statistically significant compounds, and scores plots were again produced to display the variation between albumen organic extracts that were stored either dried or re-dissolved in methanol, at the three different temperatures.

The scores plot in Figure 4.20 shows a very slight separation between the extracts stored under different conditions and at different temperatures. There is no separation between dried and re-dissolved extracts, however there is some slight separation between the re-dissolved extracts that were stored at the three different temperatures. The re-dissolved extracts that were stored at -80 °C are the closest to the QC samples, and are therefore the most representative of the albumen organic extracts, indicating that if extracts are re-dissolved in methanol prior to storage, then they should be stored at -80 °C. Re-dissolved extracts that were stored at -46 °C are slightly further away from the QCs, and show more variation between samples compared to those stored at -80 °C, and re-dissolved extracts that were stored at -25 °C are the furthest from the QC samples,

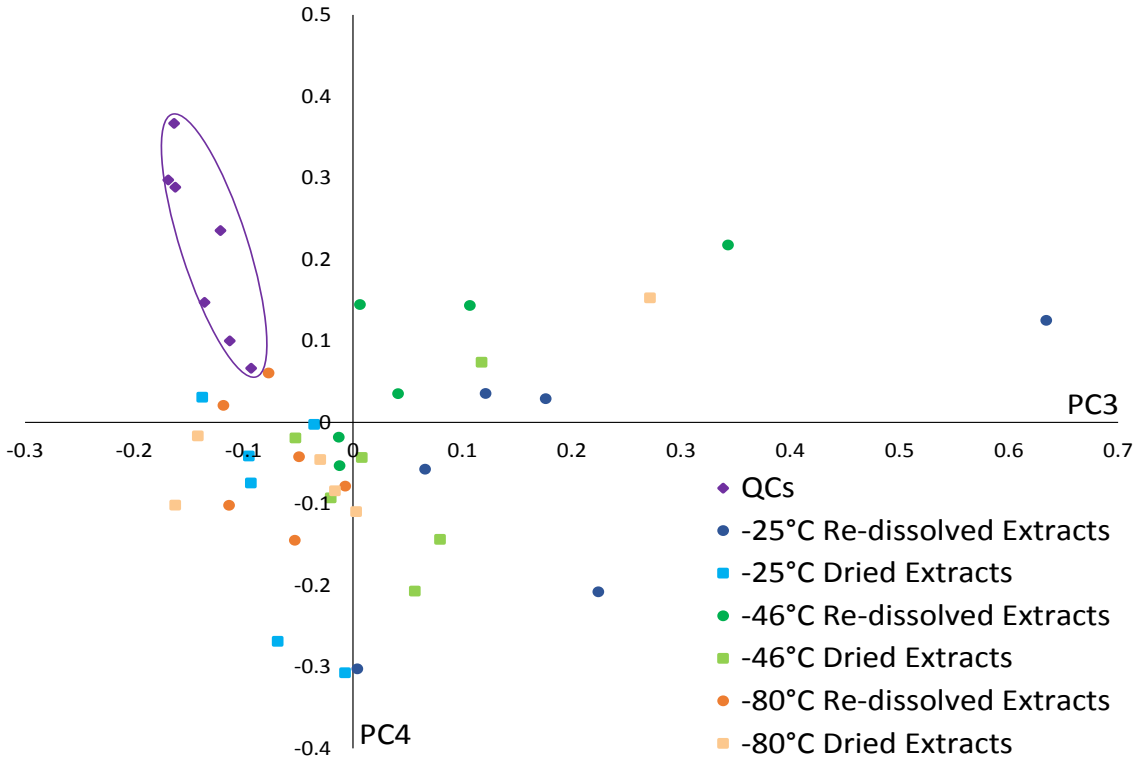


Figure 4.20: *PCA scores plot showing PC3 vs PC4 for albumen organic extracts that were stored dried and re-dissolved at -25 ˚C, -46 ˚C, and -80 ˚C for 44 weeks, including all compounds with CV%<30%. PC3 explains 0.01% of the variance, and PC4 explains 0.006% of the variance.*

and are the most spread out, showing the most variation between samples. This indicates that the higher the storage temperature of extracts, the less stable the metabolite profiles are following 44 weeks of storage, resulting in extracts that were stored under the same conditions and temperatures, having varying metabolite profiles. The separation between re-dissolved extracts that were stored at different temperatures is across PC3, which describes just 0.01% of the variance between all of the samples, highlighting the extreme subtlety of the differences between extracts that were stored at different temperatures. As the dried extracts show no separation due to storage temperature, this indicates that storing extracts dried, rather than re-dissolved in methanol, reduces the effect of temperature on the resulting metabolite profiles. The scores plot also shows that when albumen organic extracts are stored for 44 weeks, storage temperature has more of an effect on the metabolite profiles of the extracts, than storage condition.

4.7.1.5 Summary

From the overlaid TICs in Figures 4.15-4.18, it could be seen that both storage temperature and condition have an effect on the resulting metabolite profiles of extracts, and that lower temperatures result in a greater stability in metabolite profile, as does storing extracts dried rather than re-dissolved in methanol. Although the scores plot in Figure 4.19 showed no variation between extracts stored under different conditions and temperatures, proving how subtle the effects are, the plot in Figure 4.20 showed some variation between re-dissolved extracts stored at the three different temperatures. The results of the statistical tests in Table A.5 showed that more compounds were statistically significant in abundance between extracts that were stored under different conditions, rather than at different temperatures. There were also more statistically significant compounds found when comparing re-dissolved extracts that were stored at the three different temperatures, rather than dried extracts.

The differences observed in metabolite profile between albumen organic extracts that were stored under different conditions and at different temperatures for 44 weeks, were less than the differences observed between yolk organic extracts that were stored under different conditions and at different temperatures for 6 and 12 weeks. This could be because over a longer storage period, the changes that occurred in metabolite profile earlier on in the extracts that were stored under less stable conditions may have occurred later on, over a longer period of time, during storage in the extracts that were stored under more stable conditions, resulting in similar metabolite profiles. It could also suggest that storage temperature and conditions have a lower impact on the metabolite profiles of albumen organic extracts than yolk organic extracts.

4.8 Conclusions

It was observed that when yolk organic extracts were stored under different conditions, i.e. dried or re-dissolved in methanol, and at three different temperatures, -25 °C, -46 °C, and -80 °C, there was a difference between the resulting metabolite profiles of the extracts, after 6 weeks. As the results showed that re-dissolved yolk organic extracts are more susceptible to changes in metabolite profile due to temperature than dried extracts, and that lower storage temperatures result in less difference between the metabolite profiles of extracts stored at different temperatures, it was concluded that yolk organic extracts are best stored dried, at -80 °C, when being stored for a 6 week period.

Similar results were observed when yolk organic extracts were stored under different conditions and at different temperatures for 12 weeks. The results again showed that re-dissolved yolk organic extracts are more susceptible to changes in metabolite profile due to temperature, than dried extracts, and that the differences in metabolite profile between extracts stored at different temperatures, are smaller between extracts stored at lower temperatures. Therefore, it was again concluded, that when storing yolk organic extracts for a 12 week period, they should be stored dried, at -80 °C, in order to achieve the highest stability.

Very similar trends occurred over both 6 and 12 weeks of extract storage, with dried extracts stored at lower temperatures appearing to be the most stable. Therefore, yolk organic extracts that are being stored for up to 12 weeks, should be stored dried, and at -80 °C, to attain minimum metabolite profile change.

When albumen organic extracts were stored under different conditions and at different temperatures, there was less difference between the resulting metabolite profiles of the extracts after 44 weeks, than was observed over 6 and 12 weeks of yolk organic extract storage. However, the differences that were observed follow the same trend as for yolk organic extracts; the metabolite profiles of re-dissolved albumen organic extracts are more

affected by storage temperature than dried extracts, and lower storage temperatures result in a smaller difference between extracts stored at different temperatures. Therefore, it was also concluded that when storing albumen organic extracts for an extended period of time, it is best to store them dried, and at -80 °C, to achieve the best stability.

4.9 References

[1] I. Wilson, R. Plumb, J. Granger, H. Major, R. Williams, and E. Lenz, "HPLC-MS-based methods for the study of metabonomics," *Journal of Chromatography B*, vol. 817, pp. 67–76, 2005.

[2] H. Gika, G. Theodoridis, R. Plumb, and I. Wilson, "Current practice of liquid chromatography mass spectrometry in metabolomics and metabonomics," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 87, pp. 12–25, 2014.

[3] R. Culp-Hill, J. Reisz, K. Hansen, and A. D'Alessandro, "Investigation of the effects of storage and freezing on mixes of heavy-labeled metabolite and amino acid standards," *Rapid Communications in Mass Spectrometry*, vol. 31, pp. 2030–2034, 2017.

[4] H. Akshaya, Namita, K. Singh, S. Saha, S. Panwar, and C. Bharadwaj, "Standardization of storage conditions of marigold (*Tagetes* sp.) petal extract for retention of carotenoid pigments and their antioxidant activities," *Indian Journal of Agricultural Sciences*, vol. 87, pp. 765–775, 2017.

[5] P. Scano, A. Rosa, E. Locci, G. Manzo, and M. Dessì, "Modifications of the ^{1}H NMR metabolite profile of processed mullet (*Mugil cephalus*) roes under different storage conditions," *Magnetic Resonance in Chemistry*, vol. 50, pp. 436–442, 2012.

[6] E. J. Saude and B. D. Sykes, "Urine stability for metabolomic studies: Effects of preparation and storage," *Metabolomics*, vol. 3, pp. 19–27, 2007.

[7] M. Lauridsen, S. Hansen, J. Jaroszewski, and C. Cornett, "Human Urine as Test Material in ^{1}H NMR-Based Metabonomics: Recommendations for Sample Preparation and Storage," *Analytical Chemistry*, vol. 79, pp. 1181–1186, 2007.

[8] H. Gika, G. Theodoridis, and I. Wilson, "Liquid chromatography and ultra-performance liquid chromatography-mass spectrometry fingerprinting of human urine. Sample stability under different handling and storage conditions for metabonomics studies," *Journal of Chromatography A*, vol. 1189, pp. 314–322, 2008.

[9] T. Moriya, Y. Satomi, and H. Kobayashi, "Intensive determination of storage condition effects on human plasma metabolomics," *Metabolomics*, vol. 12, pp. 178–187, 2016.

[10] J. Pinto, M. Rasario, M. Domingues, E. Galhano, C. Pita, M. do Céu Almeida, I. Carreira, and A. Gil, "Human plasma stability during handling and storage: Impact on NMR metabolomics," *Analyst*, vol. 139, pp. 1168–1177, 2014.

[11] U. Kalbitzer and M. Heistermann, "Long-term storage effects in steroid metabolite extracts from baboon (*Papio* sp.) faeces - a comparison of three commonly applied storage methods," *Methods in Ecology and Evolution*, vol. 4, pp. 493–500, 2013.

5. The effect of hen age on the metabolite profiles of egg yolk and albumen

5.1 Introduction

Age is a variable that has an effect on the metabolite profile of the subject [1][2], and if not controlled or accounted for, can greatly affect the results of a metabonomic study [3][4]. Therefore, it is important to understand how the age of a laying hen affects the metabolite profiles of the yolk and albumen of the egg, so that the experimental design of metabonomic studies can account for this, and any differences that are observed in the studies can be confidently concluded to be as a result of the factor that is under investigation, rather than due to any difference in the age of the laying hens.

All eggs must be labelled with a producer code, either at the site of production or at the first packing centre that they are delivered to, to enable eggs to be traced back to their production site, and in fact to the particular flock of birds from which they originate [5][6]. However, it would be easy for producers or packing centre operatives to mislabel the eggs with incorrect producer codes, in order to avoid detection of fraud.

Food businesses may falsely advertise their eggs as being locally sourced, in order to

make them more desirable. Having the producer code on the egg would make it very easy to determine the true origin of the eggs, and therefore the fraud would be detected immediately. It would be easy for the producers to mislabel the eggs with false producer codes that correspond with egg production facilities that are truly local to the food business, to give the false impression that the eggs were locally sourced.

However, it is not likely that the birds from the true egg production site, or flock, and the birds from the falsely labelled egg production site, or flock, would be the same age, even if the fraudsters were savvy enough to ensure that they were the same breed. Therefore, being able to discriminate between eggs laid by birds of different ages, and being able to predict the age of the laying bird based on the egg itself, would help to identify and prevent cases of fraud. Clearly, there would be a benefit to having a robust scientific method capable of this.

As hens age, it is not only themselves that undergo changes, but the eggs that they produce are also affected by the increase in age. One of the most obvious changes in eggs laid by older hens compared to younger hens is their size and weight. As the age of the bird increases, so does the size and weight of the eggs that they produce [7]–[11]. The shape of eggs also changes as birds age, becoming more elongated as the age of the bird increases. Young birds are more likely to produce double yolked eggs or overly small eggs lacking any albumen, as well as other abnormalities such as soft shelled or shell-less eggs. Egg defects associated with older birds include deposits or pimples on the shells. Brown eggshells become paler in colour with an increase in hen age, due to the increasing size of the egg, without an accompanying increase in the amount of pigment produced [11].

The increase in yolk size with increasing hen age is greater than the increase of the size of the egg as a whole, and so the proportion of yolk relative to the rest of the egg is greater in eggs laid by older birds [8][12][13]. Alongside the increasing size of the yolk in eggs produced by older birds, the vitelline membrane becomes weaker, probably due to this increase in yolk size [11]. The yolk index (ratio of yolk height to diameter) has

been observed to peak when birds reach the age at which they experience 50% hen-day production, at which point it decreases with increasing bird age, meaning that the yolk is less rounded and more flat in eggs from older birds [8]. This is likely to be as a result of the weaker vitelline membrane becoming less able to hold the shape of the yolk in place.

Eggs laid by younger hens have thicker shells and longer pores compared to those produced by older birds [14]. This means that as the hen ages the shells become thinner with shorter pores, making them more porous, and thus the rate of diffusion of carbon dioxide and water through the shell increases in eggs laid by older birds [14]. The loss of carbon dioxide and water through the eggshell results in the thinning of the albumen [14], and so eggs produced by older birds have a thinner albumen than those from younger birds [14][15]. This then means that the albumen height, measured by Haugh unit (based on the relationship between the weight of the intact egg, and the albumen height once the egg has been broken [16]), is decreased due to this thinning of the albumen [7][8][11][17]. One study found that the pH of the albumen was significantly different between eggs from birds of different ages, but the change of pH did not follow any particular trend [7].

The physical changes of an egg in response to an increase in the age of the laying bird seem to be agreed on between researchers; however, the chemical changes of an egg are more in dispute, with different studies producing contrasting results.

One study found that the concentration of albumen solids decreased with increasing bird age [11], which is in agreement with another study which found that albumen solids had the highest concentration in eggs laid by young birds (28 weeks old) [18]. However, this first study also found that yolk solids remained constant with increasing hen age [11], whilst the second study found that yolk solids had the highest concentration in eggs laid by old birds (97 weeks old) [18]. The difference here could be due to the researchers referring to different ages of hen.

Two studies found that the total lipid content of egg yolk increased with increasing bird age [11][19]. It was posited that this could be due to the higher feed intake of the older

birds, with no change of the rate of lay, meaning that the lipids are more concentrated in the same number of eggs [19]. However, another study found that egg yolk lipids experienced a small increase followed by a decrease in concentration with increasing bird age [20], whilst two other studies found that the lipid content of the yolk remained stable with an increasing bird age [18][21].

The age of the laying bird has been found to affect the rate of transfer of fatty acids into the egg, and thus the content of some fatty acids in eggs can differ with the age of the hen [11]. Another study confirmed this when the percentages of certain fatty acids were found to be higher in lipid profiles of the egg yolk of younger birds, compared to older birds [21]. The results of one study showed that concentrations of palmitic and oleic acid were not affected by the age of the bird, but that stearic acid had a lower concentration in eggs from older birds whereas linoleic acid had a higher concentration in eggs from older birds [20]. This results in a significant difference in the ratio of unsaturated to saturated fatty acids and of monounsaturated to polyunsaturated fatty acids in the yolk between eggs laid by birds of different ages.

The effect of bird age on the cholesterol levels of an egg seems to be in dispute, with different researchers reaching different conclusions based on their studies. One study found that the cholesterol levels of egg yolk were not affected by the age of the laying hen [21], whereas another found that the cholesterol levels in egg yolk decreased with an increase in bird age, which was hypothesised to be due to the inability of younger birds to metabolise cholesterol efficiently [19]. Another study found that although the cholesterol concentration in the yolk decreased with an increase in hen age, the total cholesterol content of the egg increased [10]. This was likely due to the increase in the size of the yolk in eggs laid by older birds, meaning that the changing size of the egg yolk has a greater effect on the total egg cholesterol content than the change in concentration of cholesterol itself in the yolk.

One group of researchers admitted that there is some discrepancy around how choles-

terol levels are affected by laying bird age, and carried out some research in order to reach their own conclusions [22]. The results of this study revealed that there were significant differences in the cholesterol concentration between eggs laid by birds of different ages and that the cholesterol concentration in the yolk, and the egg as a whole, exhibited almost identical trends. However, it was also observed that although there were significant differences in the concentration of cholesterol between eggs laid by birds of different ages, there was no correlation between bird age and cholesterol concentration.

The effect of the age of the laying bird on other, non-lipid, chemical variables of the egg has also been examined [12]. It was found that glucose, uridine and uric acid were not affected by hen age, but that pyroglutamic acid was significantly affected by an increase in bird age. The concentration of pyroglutamic acid initially increased significantly with increasing hen age up to around 40-50 weeks of age, before dramatically decreasing again to levels similar to those found in the eggs laid by younger birds.

Various studies have produced contrasting results regarding the effect of laying hen age on the chemical composition of eggs. The use of a metabonomic approach would enable the observation of how a wide range of compounds and compound classes, in egg yolk and albumen, are affected by an increase in laying hen age. This would also help to develop an understanding of how other metabonomic studies carried out on eggs may be affected by the age of the laying birds.

5.2 Aims and Objectives

When carrying out metabonomic studies, it is important to consider how the age of the laying hen affects the metabolite profiles of the yolk and albumen of the eggs.

The accurate prediction of the age of the laying hen, from the egg itself, could help to detect cases of fraud where eggs have been labelled with false producer codes, corresponding to incorrect production sites, or flocks.

This work aimed to use a non-targeted metabonomic workflow to observe how a wide range of compounds and compound classes, in the yolk and albumen of eggs, are affected by hen age. It then aimed to determine whether any compounds are suitable to act as markers capable of predicting hen age, in order to identify and prevent cases of fraud. The observations of how the metabolite profiles of egg yolk and albumen change in response to laying hen age, could then be used to optimise the experimental design, and to correctly interpret the results of future metabonomic studies.

5.3 Experimental

Eggs were collected on the day of lay from the National Institute of Poultry Husbandry (Harper Adams University, Newport, U.K.). Laying hens were of the Hy-line brown breed, fed the same diet, and kept in enriched cages with 80 birds per cage. The first set of six eggs were collected when the laying hens were 21 weeks old, and six eggs were collected every week for the first nine weeks of the study, every other week up to seventeen weeks, and every third week up to twenty-three weeks of the study, when laying hens were 44 weeks old. All six eggs underwent metabolite extraction on the day of lay, and the resulting extracts were stored at -80 °C until sample collection, after twenty-three weeks, was complete and chemical analysis could take place.

5.3.1 Yolk Organic Extracts

In order to analyse yolk organic extracts of eggs laid by hens of increasingly higher ages in the same batch, several analytical runs took place, with yolk extracts from overlapping hen ages in each batch; 21-26 weeks old (every week), 26-30 weeks old (every week), 30-38 weeks old (every other week), and 38-44 weeks old (every third week). The data from these analyses were all analysed separately, to observe how metabolite profiles change over a short period of increasing hen age. A final analytical run was carried out, with extracts from eggs laid by hens of 21 weeks, 26 weeks, 30 weeks, 38 weeks, and 44 weeks of age, in order to observe how the metabolite profiles change over a longer period of increasing hen age.

Metabolite extraction, chemical analysis, quality control analysis, and data pre-process -ing and analysis were carried out as described in Chapter 3. For hens ranging from 38 to 44 weeks old in Section 5.7, a second PCA was included using just the top statistically significant compounds following ANOVA/Welch tests. In Section 5.8 for hens ranging from 21 to 44 weeks old, following the usual data analysis workflow additional

ANOVA/Welch tests, and post-hoc tests, were carried out for any previously putatively identified compounds that were found to show a significant difference in abundance between eggs laid by the youngest and oldest birds in the smaller age ranges that were studied. This was done in order to confirm the significance of these compounds between eggs laid by birds of these ages, and to determine whether they showed any statistical significance, or correlation between abundance and hen age, over a wider hen age range.

5.3.2 Albumen Organic Extracts

For the albumen organic extracts, only those from eggs laid by hens of 21 weeks, 26 weeks, 30 weeks, 38 weeks, and 44 weeks of age were analysed in the same analytical run, in order to observe how the metabolite profiles change over a longer period of increasing hen age.

Metabolite extraction, chemical analysis, quality control analysis, and data pre-process -ing and analysis were carried out as described in Chapter 3.

5.4 Hen Age 21-26 weeks: Yolk Organic Extracts

Yolk organic extracts of eggs laid by hens ranging from 21 weeks to 26 weeks old were compared, to see whether hen age affects the metabolite profile of an egg.

5.4.1 Results and Discussion

5.4.1.1 Quality Control Analysis

Table 5.1 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 5.1: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 21-26 weeks old*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	2205874	367763234	184829806	29671734	966714125	47512887
QC2	2544736	353379168	174368304	29014898	932457139	39602874
QC3	2550947	339375492	165664998	25636233	888948721	36545277
QC4	2458735	339927200	153523553	25967550	848864012	33974178
QC5	3124837	333983190	142295641	24051675	825570218	33207934
QC6	2102558	325477715	138819842	23805930	810107555	36383193
QC7	2132169	320257410	148597601	23775944	825415240	33031730
SD	3549289	16263760	17169892	2458324	59999850	5114349
Mean	2445694	340023344	158299964	25989138	871153859	37179725
CV%	14.51	4.78	10.85	9.46	6.89	13.76

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 5.2 shows the retention times (RTs) for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 5.2: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 21-26 weeks old*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.674	9.737	33.407	57.476	72.300	79.487
QC2	4.759	9.855	33.476	57.562	72.402	79.540
QC3	4.692	9.705	33.326	57.396	72.219	79.257
QC4	4.719	9.682	33.270	57.289	72.096	78.968
QC5	4.722	9.768	33.339	57.458	72.298	79.369
QC6	4.707	9.637	33.291	57.410	72.183	79.105
QC7	4.675	9.754	33.342	57.494	72.267	79.338
SD	0.030	0.070	0.070	0.087	0.098	0.204
Mean	4.707	9.734	33.350	57.441	72.252	79.295
CV%	0.64	0.72	0.21	0.15	0.14	0.26

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

5.4.1.2 Metabolite Profiling

Comparisons were made between yolk organic extracts of eggs laid by hens that were 21 weeks and 26 weeks old, by studying the Total Ion Chromatograms (TICs) of the extracts following HPLC-MS analysis. Figure 5.1 shows overlaid TICs for the average chromatograms of extracts from eggs laid by hens of these two ages.

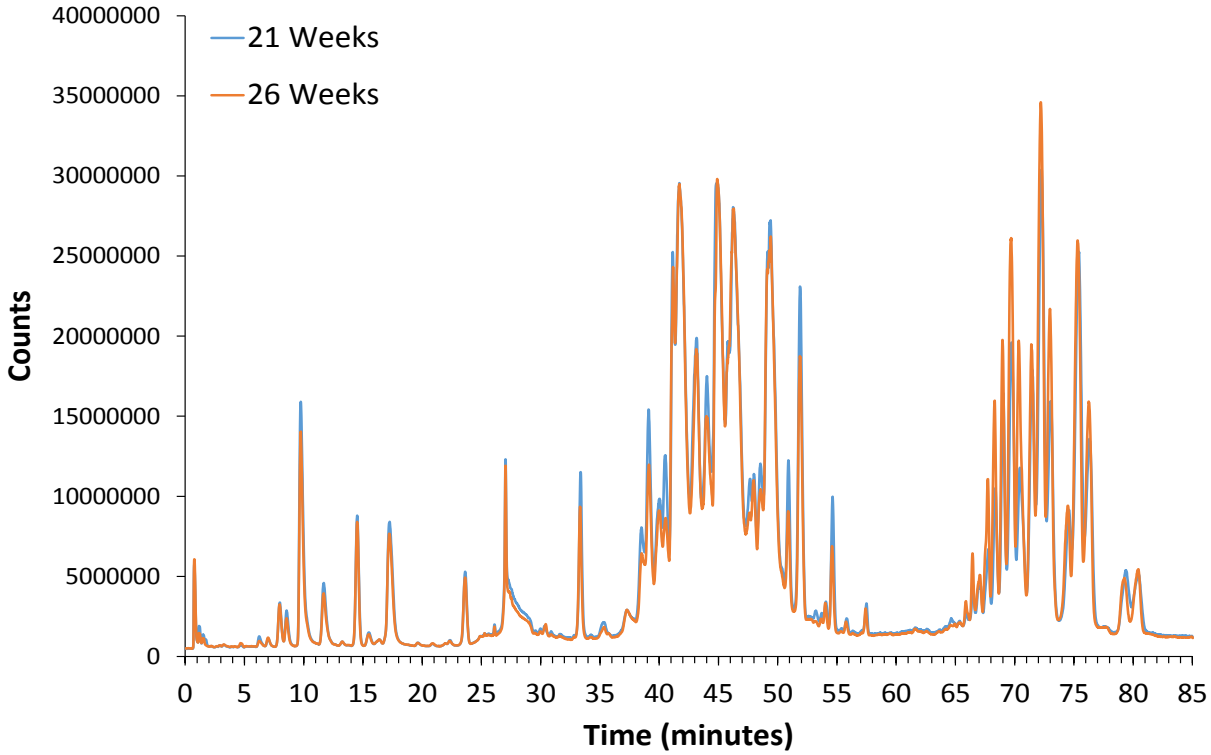


Figure 5.1: *Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 21 weeks and 26 weeks old*

Observing these overlaid TICs, a slight difference in peak intensity can be seen between the chromatograms of the yolk extracts of eggs laid by hens of different ages. Throughout the first part of the chromatograms, up to approximately 65 minutes, most peaks show a higher peak intensity in the TIC of the yolks of eggs laid by hens that were 21 weeks old, compared to the TIC of the yolks of eggs laid by hens that were 5 weeks older. This suggests that these earlier eluting, less non-polar compounds are more abundant in eggs laid by younger hens, compared to those laid by older hens. However after approximately 65 minutes, most peaks show a higher intensity in the TIC of the yolks of eggs laid

by hens that were 26 weeks old. As the compounds that elute towards the end of the chromatographic analysis are the more non-polar compounds, they could be larger lipid molecules such as phospholipids, which are highly abundant in egg yolk. If these later eluting compounds are phospholipids, then the overlaid TICs indicate that this class of compound is more abundant in eggs laid by older hens, compared to those laid by younger hens.

5.4.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the eggs laid by hens of different ages.

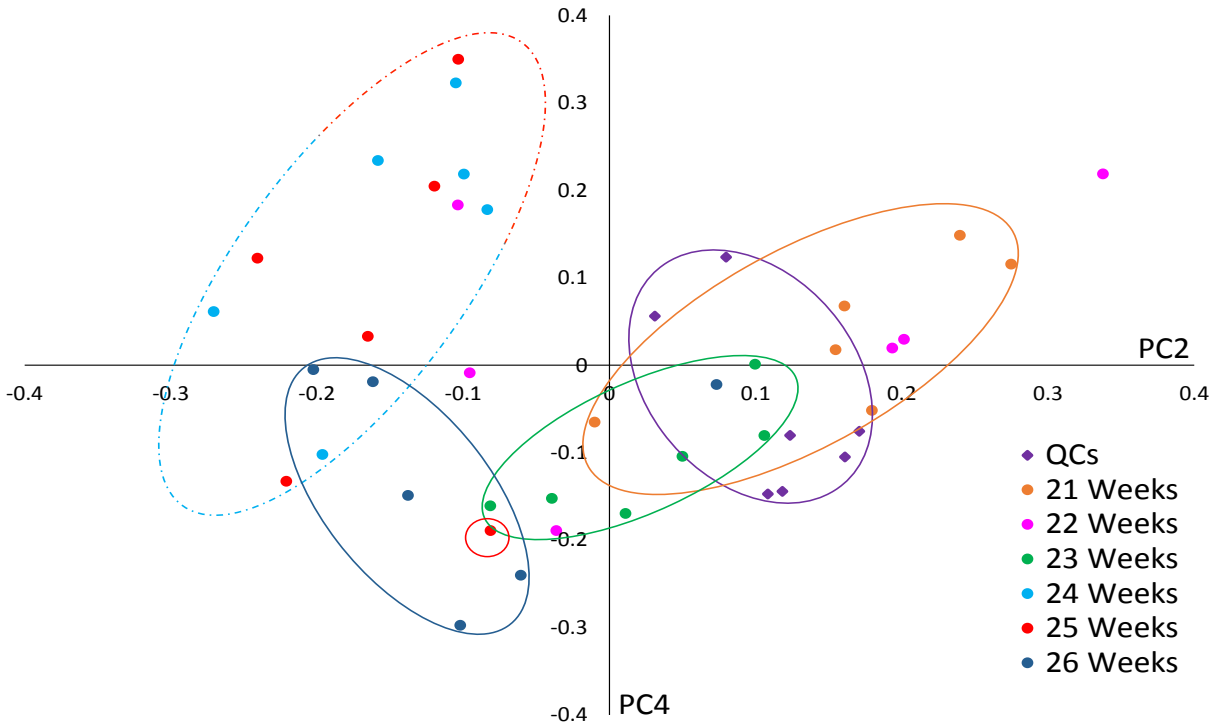


Figure 5.2: *PCA scores plot showing PC2 vs PC4 for yolk organic extracts of eggs laid by hens from 21 weeks to 26 weeks old, including all compounds with CV%<30%. PC2 explains 0.65% of the variance, and PC4 explains 0.13% of the variance.*

The scores plot in Figure 5.2 shows some separation between eggs laid by different ages of hen. There is a large amount of separation between eggs laid by hens that were 21 and 23 weeks old, and those laid by hens that were 24 and 25 weeks old, indicating

that there is a clear difference in yolk metabolite profile between eggs laid by birds of these ages. There is also significant separation between eggs laid by hens that were 21 weeks old, and those laid by hens that were 26 weeks old, again indicating that there is a clear difference in metabolite profile between the yolks of eggs laid by birds of these ages. Eggs laid by hens that were 24 and 25 weeks old appear to be grouped together, suggesting that there is not much difference in yolk metabolite profile between eggs laid by birds of these two ages. However, there is one egg laid by a hen that was 25 weeks old (circled on the scores plot), which shows some variation to the other eggs within its sample set. It is not known why this is, as no anomalies were noted during the metabolite extraction or chemical analysis. Eggs laid by birds that were 22 weeks old have quite a lot of variation and are very spread out over the scores plot. This indicates that for this set of eggs, most of the variation described by PC2 and PC4 is between random eggs within this sample set, resulting in them being highly spread throughout the plot.

The QC samples on this plot are not particularly tightly clustered, however the analysis was robust, as observed in the quality control analysis in Section 5.4.1.1. The reason for the spread in QC samples could be because the differences between the yolk samples are very subtle. When this is the case, the PCA exploits differences between QC samples, as there is such a small amount of variation between sample sets. Most of the separation between eggs laid by hens of different ages appears to be across PC2, which describes only 0.65% of the total variation between the yolk samples, showing just how subtle the differences in metabolite profile are between eggs laid by birds of different ages.

5.4.1.4 Univariate Statistics

There were 46 compounds remaining following the removal of any duplicates, adducts and isotopes from the top 100 compounds, and all 46 of these were found to be statistically significant. These compounds, and their CV%s and p-values, can be seen in Table B.1 in Appendix B; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

Of the 42 compounds that were still found to be statistically significant when confirmed using the raw data, six were putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The observed m/z for all compounds was due to the $[M+H]^+$ adduct. The compounds m/z 331.2847, 496.3412, 648.6307, 703.5761, 716.5243, and 734.5709 were putatively identified as 1-monopalmitin, 1-palmitoyl-glycero-3-phosphatidylcholine (PC(16:0/0:0)), nervonic ceramide, 1-palmitoyl-sphingomyelin (SM(18:1/16:0)), 1-palmitoyl-2-linoleoyl-glycero-3-phosphatidylethanolamine (PE(16:0/18:2)), and 1,2-dipalmitoyl-glycero-3-phosphatidylcholine (PC(16:0/16:0)) respectively. The comparisons between mass spectra provided by METLIN and those resulting from the analysis in this study can be seen in Appendix B, Figures B.1-B.6.

Tables 5.3-5.8, and Figures 5.3-5.8, show the putative identifications of these compounds, the p-values resulting from ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests, as well as the trends in the changing abundance of the compounds in the yolks of eggs laid by birds of increasing age from 21 to 26 weeks old.

Table 5.3: *Table showing the putative identification of compound m/z 331.2847 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
331.2847	$C_{19}H_{38}O_4$	99.35	1-monopalmitin	<0.001

Post-hoc test *p*-values

	21 Weeks	**22 Weeks**	**23 Weeks**	**24 Weeks**	**25 Weeks**	**26 Weeks**
21 Weeks		0.468	0.011	0.001	<0.001	0.982
22 Weeks			<0.001	<0.001	<0.001	0.153
23 Weeks				0.964	0.619	0.008
24 Weeks					0.619	0.008
25 Weeks						<0.001
26 Weeks						

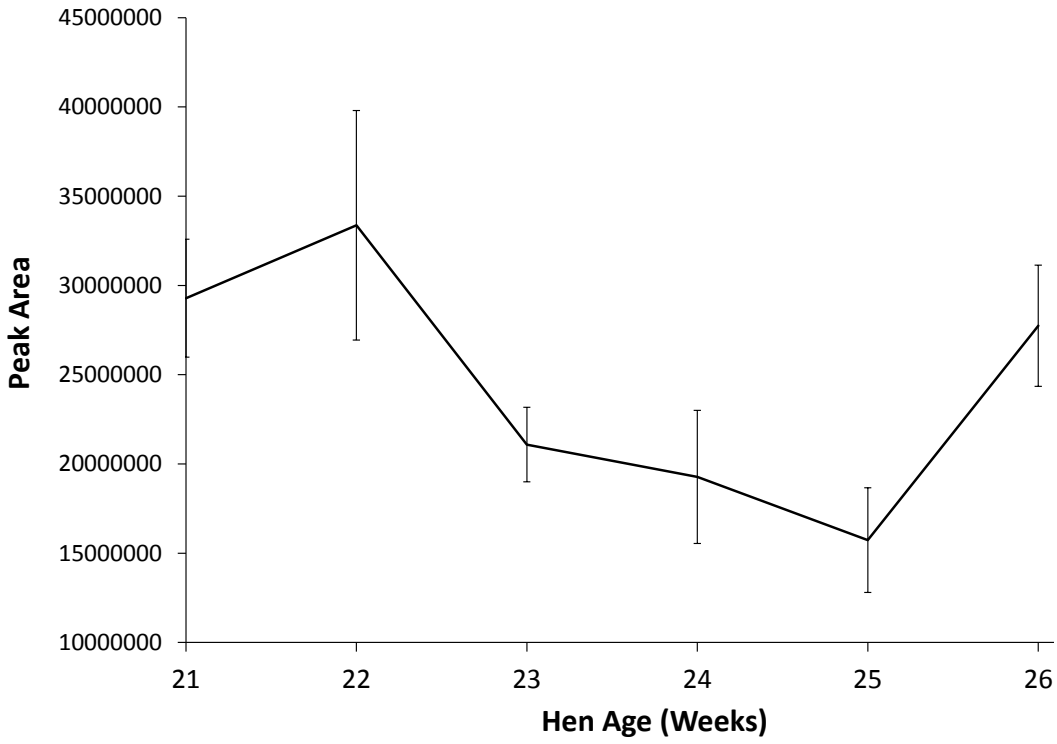


Figure 5.3: *Line graph showing the trend in changing abundance of compound m/z 331.2847 in yolk across eggs laid by hens from 21-26 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.4: *Table showing the putative identification of compound m/z 496.3412 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
496.3412	$C_{24}H_{50}NO_7P$	96.25	PC(16:0/0:0)	<0.001

Post-hoc test *p*-values

	21 Weeks	**22 Weeks**	**23 Weeks**	**24 Weeks**	**25 Weeks**	**26 Weeks**
21 Weeks		1	0.512	0.001	0.003	0.121
22 Weeks			0.668	0.001	0.005	0.196
23 Weeks				0.058	0.162	0.948
24 Weeks					0.996	0.316
25 Weeks						0.604
26 Weeks						

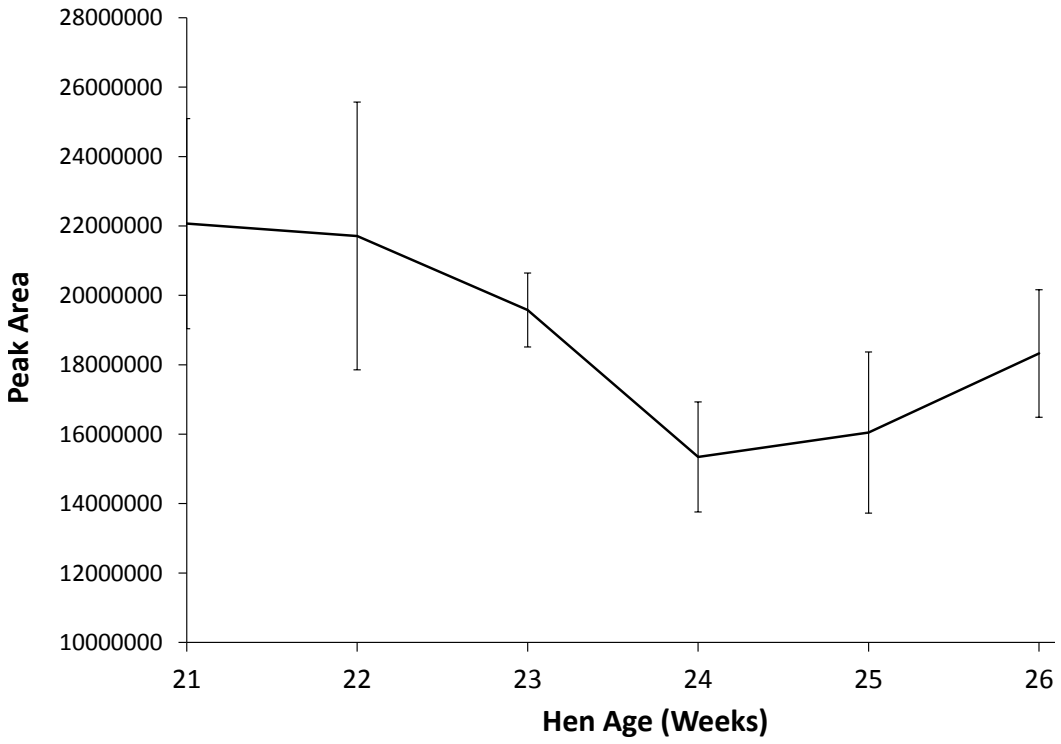


Figure 5.4: *Line graph showing the trend in changing abundance of compound m/z 496.3412 in yolk across eggs laid by hens from 21-26 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.5: *Table showing the putative identification of compound m/z 648.6307 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
648.6307	$C_{42}H_{81}NO_3$	95.61	Nervonic Ceramide	0.007

Post-hoc test *p*-values

	21 Weeks	**22 Weeks**	**23 Weeks**	**24 Weeks**	**25 Weeks**	**26 Weeks**
21 Weeks		0.991	1	0.112	0.059	0.186
22 Weeks			0.983	0.325	0.195	0.471
23 Weeks				0.093	0.048	0.158
24 Weeks					1	1
25 Weeks						0.993
26 Weeks						

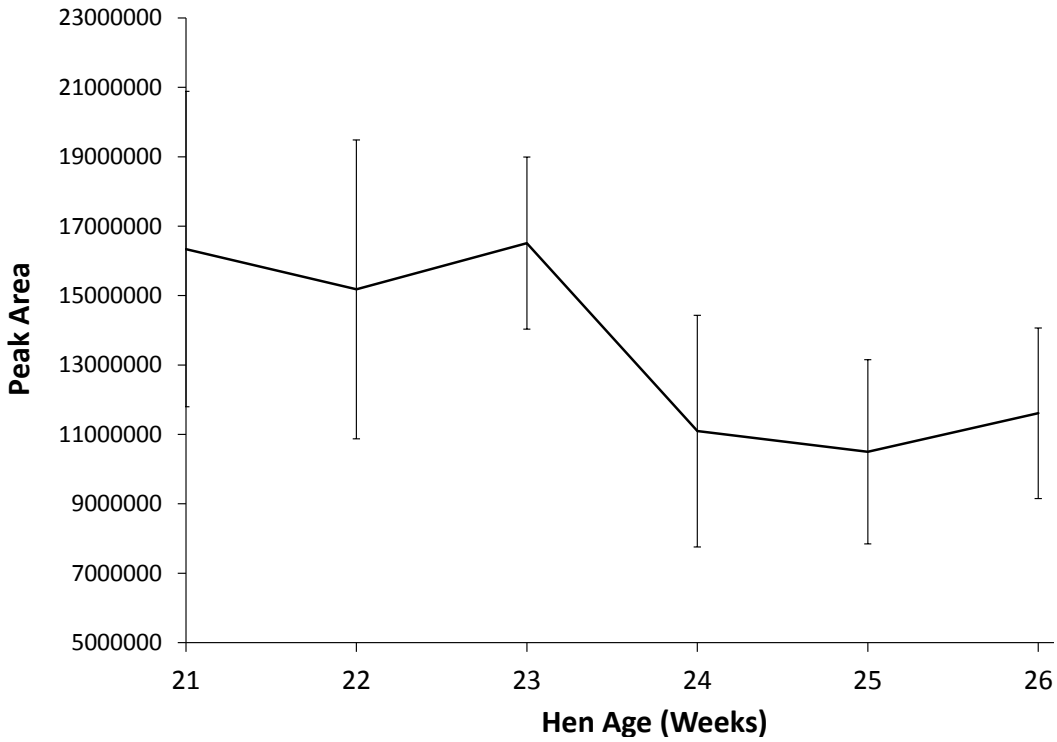


Figure 5.5: *Line graph showing the trend in changing abundance of compound m/z 648.6307 in yolk across eggs laid by hens from 21-26 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.6: *Table showing the putative identification of compound m/z 703.5761 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
703.5761	$C_{39}H_{79}N_2O_6P$	95.76	SM(18:1/16:0)	<0.001

Post-hoc test *p*-values

	21 Weeks	22 Weeks	23 Weeks	24 Weeks	25 Weeks	26 Weeks
21 Weeks		1	0.456	0.001	<0.001	0.001
22 Weeks			0.582	0.002	0.001	0.002
23 Weeks				0.110	0.047	0.095
24 Weeks					0.999	1
25 Weeks						0.999
26 Weeks						

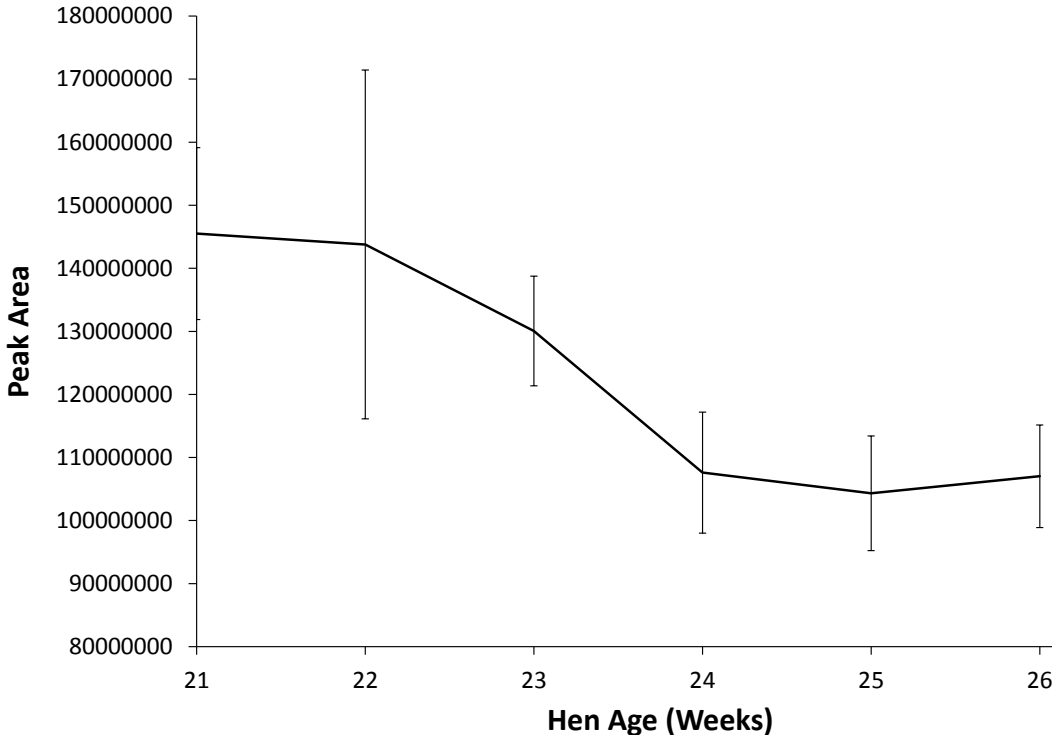


Figure 5.6: *Line graph showing the trend in changing abundance of compound m/z 703.5761 in yolk across eggs laid by hens from 21-26 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.7: *Table showing the putative identification of compound m/z 716.5243 and the p-values resulting from Welch test and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	Welch *p*-value
716.5243	$C_{39}H_{74}NO_8P$	95.32	PE(16:0/18:2)	<0.001

Post-hoc test *p*-values

	21 Weeks	22 Weeks	23 Weeks	24 Weeks	25 Weeks	26 Weeks
21 Weeks		0.969	0.018	0.008	0.031	0.001
22 Weeks			0.607	0.025	0.081	0.017
23 Weeks				0.084	0.254	0.031
24 Weeks					0.998	0.998
25 Weeks						1
26 Weeks						

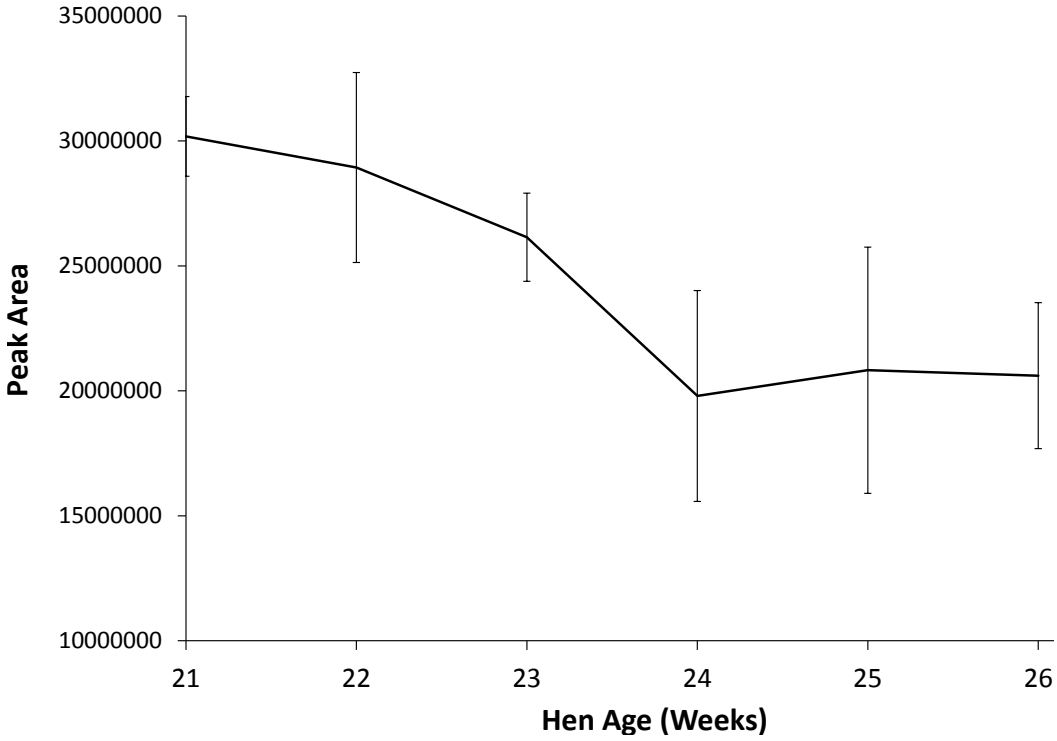


Figure 5.7: *Line graph showing the trend in changing abundance of compound m/z 716.5243 in yolk across eggs laid by hens from 21-26 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.8: *Table showing the putative identification of compound m/z 734.5709 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
734.5709	$C_{40}H_{80}NO_8P$	96.59	PC(16:0/16:0)	<0.001

Post-hoc test *p*-values

	21 Weeks	22 Weeks	23 Weeks	24 Weeks	25 Weeks	26 Weeks
21 Weeks		0.849	0.129	<0.001	<0.001	0.001
22 Weeks			0.709	<0.001	0.001	0.014
23 Weeks				0.001	0.031	0.286
24 Weeks					0.747	0.169
25 Weeks						0.880
26 Weeks						

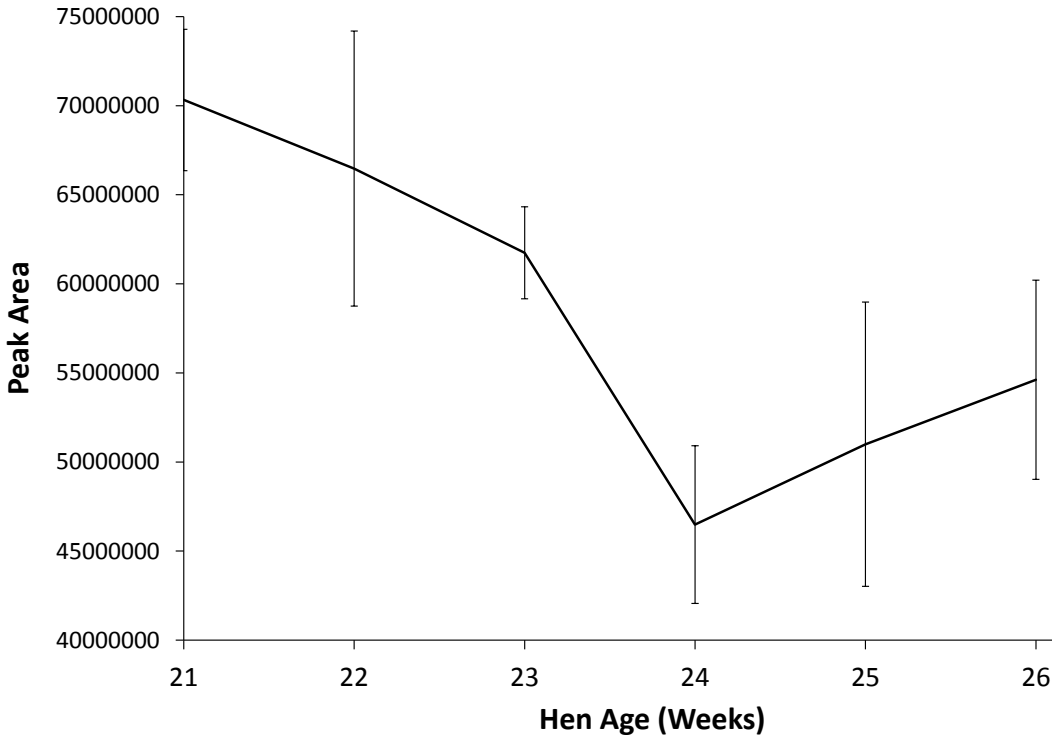


Figure 5.8: *Line graph showing the trend in changing abundance of compound m/z 734.5709 in yolk across eggs laid by hens from 21-26 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

The probability scores for the predicted molecular formulae for all of these compounds are above 95, indicating that there is a strong likelihood that the predicted formulae correspond to the compounds of interest. The p-values from ANOVA and Welch tests show that the abundances of these compounds are statistically significant when comparing them between eggs laid by birds of different ages, ranging from 21 to 26 weeks old, and the post-hoc test p-values mostly corroborate this. However, the compound that was putatively identified as nervonic ceramide, which is a ceramide with nervonic acid bonded to the amino group of the sphingosine molecule, shows only one instance of very slight statistical significance, when comparing eggs laid by birds that were 23 and 25 weeks old, as can be seen in Table 5.5. In addition to not showing much significant difference through the post-hoc test p-values, this compound does not show any correlation between abundance in egg yolk and laying hen age, as seen in Figure 5.5, and so is unsuitable as a marker of hen age.

The compound that was putatively identified as 1-monopalmitin, a monoglyceride with a palmitic acid chain attached to the first carbon of the glycerol backbone, also does not show any correlation between abundance in egg yolk and increasing laying hen age, as can be seen in Figure 5.3. However, this compound does show many instances of high statistical significance in abundance between eggs laid by birds of different ages, through the post-hoc test p-values seen in Table 5.3. Although the changing abundance of this compound is highly significantly different between eggs laid by birds of increasing age, as there is no trend in this changing abundance, it is not suitable as a marker of hen age.

The remaining putatively identified compounds: PC(16:0/0:0), SM(18:1/16:0), PE(16:0 /18:2), and PC(16:0/16:0) all show a similar trend in decreasing abundance between eggs laid by birds of increasing age, as seen in Figures 5.4, 5.6, 5.7, and 5.8 respectively. Both PC(16:0/0:0) and PC(16:0/16:0), as well as PE(16:0/18:2), experience a negative correlation, and decrease in abundance with increasing laying hen age from 21 to 24 weeks old, before increasing in abundance with increasing hen age over the next two

weeks. This is likely to be because these compounds are very similar; PC(16:0/16:0) is a phosphatidylcholine, with a palmitic acid group attached to the first two carbons of the glycerol backbone, and (PC(16:0/0:0)) is a lysophosphatidylcholine, with just one palmitic acid group attached to the first carbon of the glycerol backbone. PE(16:0/18:2) is a phosphatidylethanolamine with a palmitic acid group attached to the first carbon and a linoleic acid group attached to the second carbon of the glycerol backbone. These compounds all belong to the glycerophospholipid metabolic pathway, and follow very similar pathways in their production from diacylglycerol, and their catabolism to glycerol-3-phosphate [23]. As very similar enzymes are involved in the metabolic pathways associated with all three compounds, the rate of metabolism will be very similar between these compounds in birds of different ages, resulting in similar abundances in the eggs laid by the birds, and thus similar trends in the changing abundance of these compounds in eggs laid by birds of increasing age. The compound that was putatively identified as SM(18:1/16:0), which is a sphingomyelin with a palmitic acid group attached, shows a similar trend in changing abundance, but continues decreasing in abundance up to 25 weeks of hen age, before increasing slightly in eggs laid by birds that were 26 weeks old. All four of these compounds also show high statistical significance in the p-values from the post-hoc tests, seen in Tables 5.4, 5.6, 5.7, and 5.8, and so could have potential as markers of hen age between 21 and 24 or 25 weeks of age.

Some of the standard deviations of these compound abundances are quite large, resulting in wide error bars on the line graphs. Some error bars only appear large due to the scaling of the y-axis, but some of them are quite wide due to the fact that there is biological variation within the sets of eggs laid by birds of different ages, because the eggs were likely laid by different birds. Although the variables regarding the birds were kept as controlled and uniform as possible, there was naturally some variation between them, and thus between the eggs that they laid. Larger sample sizes would likely reduce the standard deviations of the abundances of the compounds within eggs laid by birds of the same age.

5.5 Hen Age 26-30 weeks: Yolk Organic Extracts

Yolk organic extracts of eggs laid by hens ranging from 26 weeks to 30 weeks old were compared, to see whether hen age affects the metabolite profile of an egg.

5.5.1 Results and Discussion

5.5.1.1 Quality Control Analysis

Table 5.9 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 5.9: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 26-30 weeks old*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	2792029	387789304	158759218	28565846	948201985	145902974
QC2	3067092	379290823	149003792	26674067	896256874	130209929
QC3	2660938	375157980	143993416	27410483	890458521	127447915
QC4	2618668	371076672	141845443	25780768	908356257	125439761
QC5	2489738	375273600	143097360	25479742	921333649	138084240
QC6	2973919	377319310	151689036	27084169	920486340	128730403
QC7	2634185	368842436	149374674	26370111	899443609	115317173
SD	207669	6153371	5913125	1044538	19801088	9683548
Mean	2748081	376392875	148251848	26766455	912076748	130161771
CV%	7.56	1.63	3.99	3.90	2.17	7.44

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 5.10 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 5.10: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens aged 26-30 weeks old*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.675	9.688	33.209	57.278	71.919	78.625
QC2	4.673	9.586	33.140	57.243	71.817	78.606
QC3	4.675	9.705	33.276	57.312	71.886	78.708
QC4	4.626	9.788	33.376	57.462	72.185	79.141
QC5	4.690	9.687	33.258	57.360	72.018	78.899
QC6	4.624	9.637	33.191	57.244	71.918	78.697
QC7	4.589	9.702	33.207	57.309	71.983	78.822
SD	0.037	0.062	0.076	0.077	0.118	0.188
Mean	4.650	9.685	33.237	57.315	71.961	78.785
CV%	0.80	0.65	0.23	0.13	0.16	0.24

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

5.5.1.2 Metabolite Profiling

Comparisons were made between yolk organic extracts of eggs laid by hens that were 26 weeks and 30 weeks old, by studying the TICs of the extracts following HPLC-MS analysis. Figure 5.9 shows overlaid TICs for the average chromatograms of extracts from eggs laid by hens of these two ages.

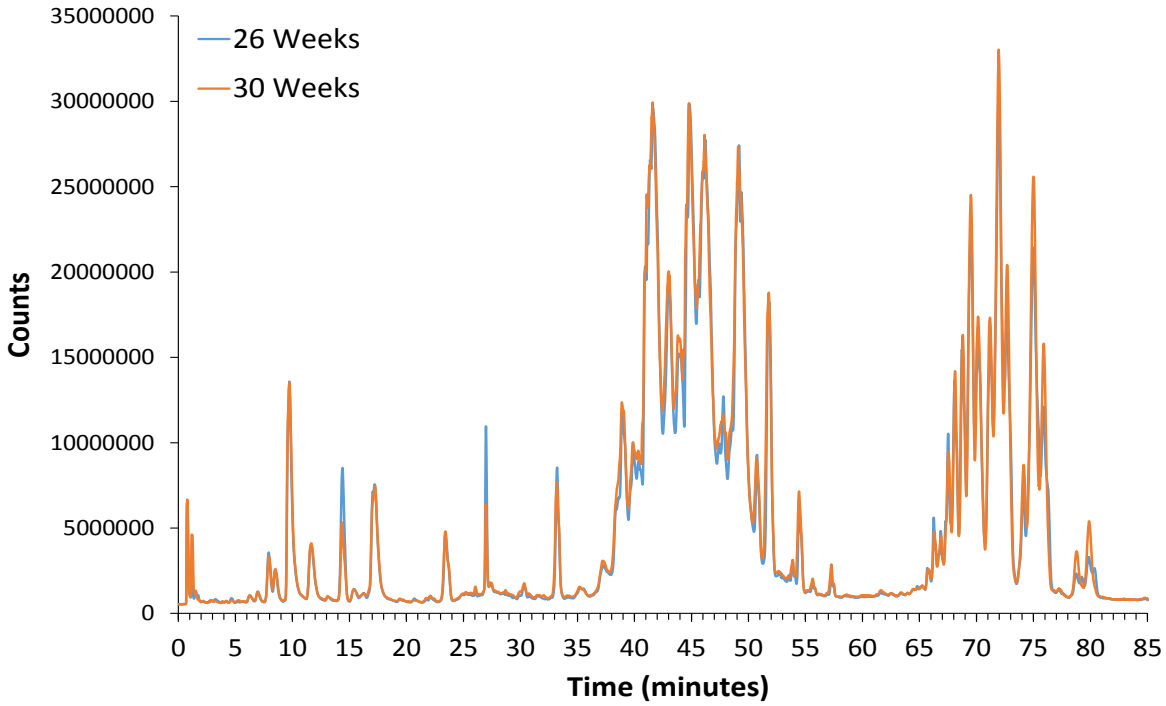


Figure 5.9: *Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 26 weeks and 30 weeks old*

There is a slight visible difference in metabolite profile seen in Figure 5.9, between the yolks of eggs laid by hens that were 26 weeks and 30 weeks old. During the first 35 minutes of analysis there is very little difference in peak intensity between the two chromatograms, with most peaks being of equal intensity, with a few peaks showing a higher intensity in the TIC of the yolks of eggs laid by hens that were 26 weeks old. Between 35 and 65 minutes of analysis, there is again very little difference in peak intensity between chromatograms for eggs laid by birds of the two different ages, but most peaks have a slightly higher intensity in the TIC of eggs laid by birds that were 30 weeks old. After 65 minutes of analysis, most peaks have a higher intensity in the TIC of the yolks of eggs

laid by hens that were 30 weeks old. This indicates that the later eluting, more non-polar molecules have a higher compound abundance in eggs laid by 30 week old birds, compared to eggs laid by 26 week old birds. These non-polar compounds may be larger lipid molecules, such as phospholipids. If this is the case then this suggests, as before, that the yolks of eggs laid by older birds have a higher abundance of these phospholipid compounds.

5.5.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the eggs laid by hens of different ages.

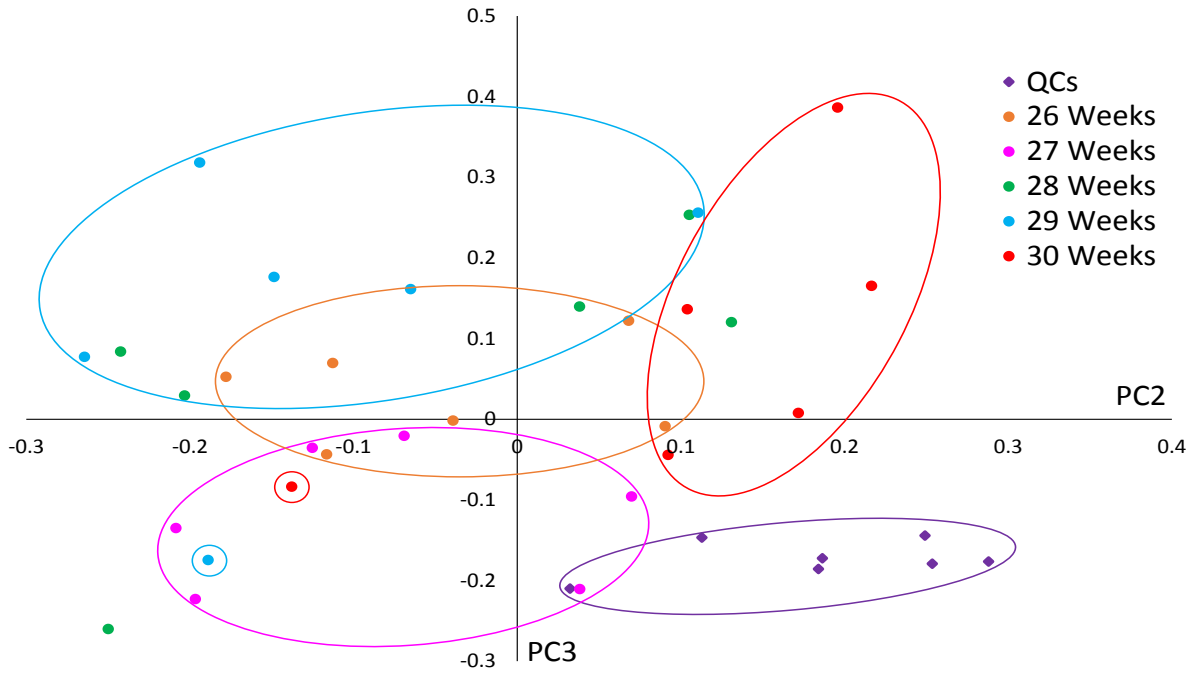


Figure 5.10: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 26 weeks to 30 weeks old, including all compounds with CV%<30%. PC2 explains 0.35% of the variance, and PC3 explains 0.19% of the variance.*

The scores plot in Figure 5.10 shows almost complete separation between eggs laid by birds that were 27 and 29 weeks old, with eggs laid by birds that were 26 weeks old forming a group overlapping both of these sample sets. However, one sample (circled)

from the eggs laid by birds that were 29 weeks old does overlap into the group of eggs laid by birds that were 27 weeks old. The eggs laid by birds that were 30 weeks old form a group that shows some separation to the sets of eggs laid by birds that were 26 and 29 weeks old, and is almost completely separate to eggs laid by birds that were 27 weeks old. However, again, just one egg (circled) from the set of eggs laid by birds of this age overlaps into the set of eggs laid by birds that were 27 weeks old. As there are two eggs, laid by birds that were 29 and 30 weeks old, that overlap into the set of eggs laid by birds that were 27 weeks old, it suggests that the differences between eggs laid by birds of these ages are very subtle. The eggs laid by birds that were 28 weeks old show a large amount of variation and are very spread out across the plot.

The separation between eggs laid by birds that were 26, 27, and 29 weeks old is across PC3 which accounts for just 0.19% of the total variation between samples, indicating that the differences between the eggs laid by birds of these ages is extremely subtle. The separation between eggs laid by birds that were 26, 27, and 29 weeks old, and eggs laid by birds that were 30 weeks old, is across PC2 which accounts for only 0.35% of the total variation, again showing the subtlety of the differences between eggs laid by birds of these ages.

The QC samples on the scores plot are quite widely spread across PC2. However, as the separation between most sample sets is across PC3, any instrumental drift that may have caused the spread across PC2 for the QC samples is not responsible, and it is due to true biological differences. Although the separation between eggs laid by birds that were 30 weeks old, and eggs laid by birds of other ages is across PC2, as the samples were randomised prior to analysis, any drift should not have affected the results.

Although there is some spread in the QC samples, the analysis was still robust, as the QC analysis in Section 5.5.1.1 shows. This spread in QC samples observed on the scores plot shows just how subtle the differences between eggs due to hen age are, as the PCA has exploited variation within the QC samples due to there being such little variation

between all of the samples.

5.5.1.4 Univariate Statistics

There were 46 compounds remaining following the removal of any duplicates, adducts and isotopes from the top 100 compounds, and 45 of these were found to be statistically significant. These compounds, and their CV%s and p-values, can be seen in Table B.2 in Appendix B; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

Of the 31 compounds that were still found to be statistically significant when confirmed using the raw data, five were putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The observed m/z for all compounds was due to the $[M+H]^+$ adduct. The compounds m/z 331.2847, 338.3420, 454.2932, 496.3405, and 703.5758 were putatively identified as 1-monopalmitin, docosenamide (erucamide), 1-palmitoyl-glycero-3-phosphatidylethanolamine (PE(16:0/0:0)), 1-palmitoyl-glycero-3-phosphatidylcholine (PC(16:0/0:0)), and 1-palmitoyl-sphingomyelin (SM(18:1/16:0)), respectively. The compounds m/z 331.2847, 496.3405, and 703.5758 were previously putatively identified when found to be in the top statistically significant compounds for eggs laid by hens aged between 21 and 26 weeks old, and their comparisons between mass spectra provided by METLIN and mass spectra resulting from chemical analysis can be seen in Figures B.1, B.2, and B.4 in Appendix B. The mass spectra comparisons for the remaining compounds can be seen in Appendix B, Figures B.7 and B.8.

Tables 5.11-5.15, and Figures 5.11-5.15, show the putative identifications of all of these compounds, the p-values resulting from ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests, as well as the trends of changing abundance of the compounds in the yolks of eggs laid by birds of increasing age from 26 to 30 weeks old.

Table 5.11: *Table showing the putative identification of compound m/z 331.2847 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
331.2847	$C_{19}H_{38}O_4$	99.49	1-monopalmitin	<0.001

Post-hoc test *p*-values					
	26 Weeks	**27 Weeks**	**28 Weeks**	**29 Weeks**	**30 Weeks**
26 Weeks		<0.001	<0.001	<0.001	<0.001
27 Weeks			1	0.064	0.007
28 Weeks				0.060	0.006
29 Weeks					0.859
30 Weeks					

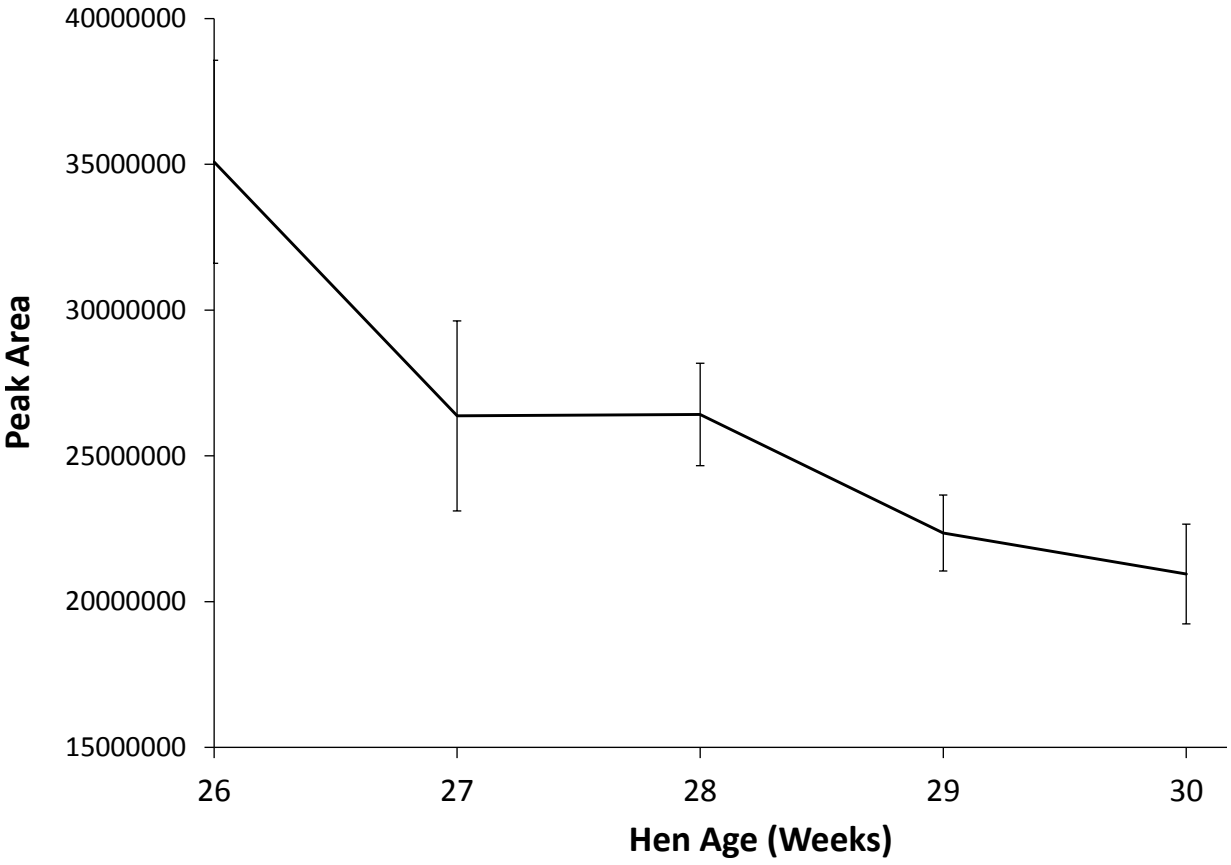


Figure 5.11: *Line graph showing the trend in changing abundance of compound m/z 331.2847 in yolk across eggs laid by hens from 26-30 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.12: *Table showing the putative identification of compound m/z 338.3420 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
338.3420	$C_{22}H_{43}NO$	99.13	Erucamide	<0.001

Post-hoc test *p*-values

	26 Weeks	**27 Weeks**	**28 Weeks**	**29 Weeks**	**30 Weeks**
26 Weeks		<0.001	<0.001	<0.001	<0.001
27 Weeks			0.998	0.962	0.654
28 Weeks				0.996	0.826
29 Weeks					0.955
30 Weeks					

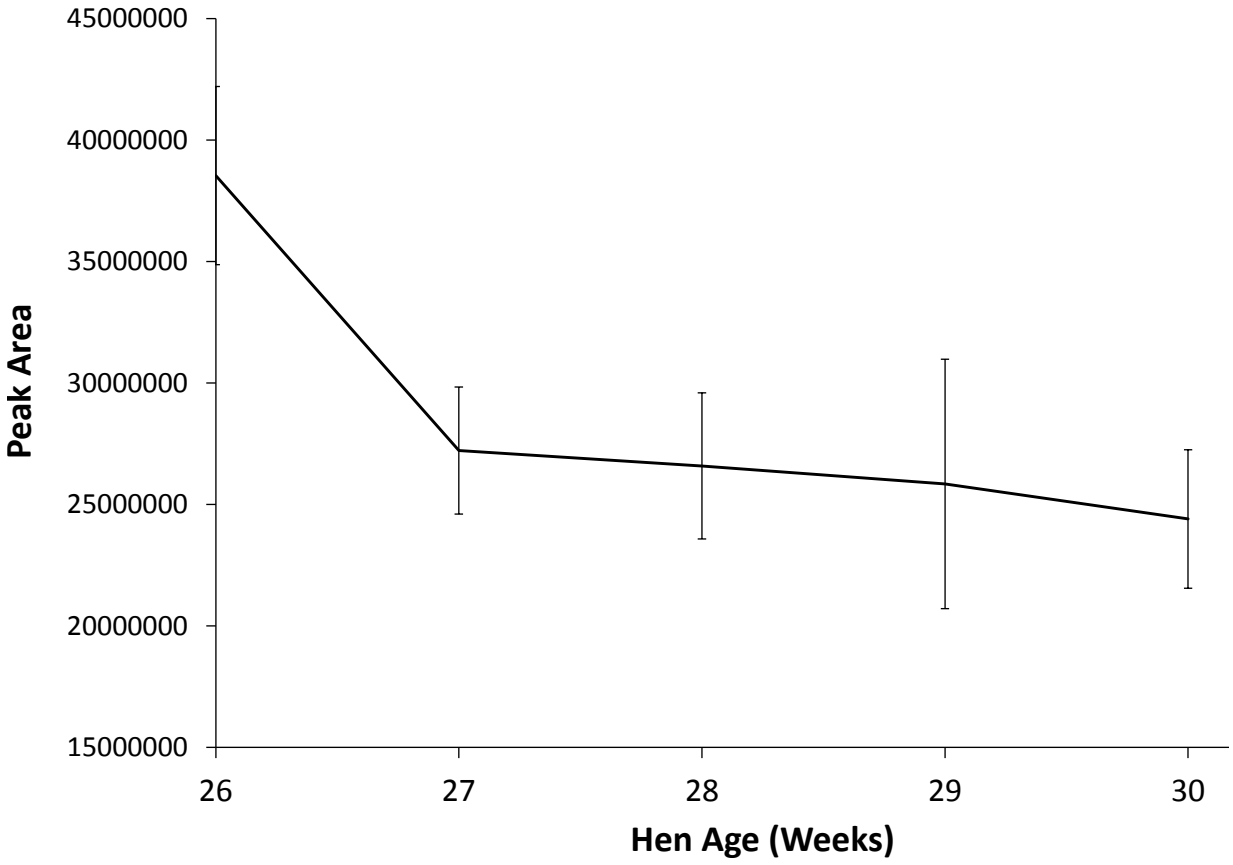


Figure 5.12: *Line graph showing the trend in changing abundance of compound m/z 338.3420 in yolk across eggs laid by hens from 26-30 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.13: *Table showing the putative identification of compound m/z 454.2932 and the p-values resulting from Welch test and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**Welch *p*-value**
454.2932	$C_{21}H_{44}NO_7P$	99.35	PE(16:0/0:0)	0.001

Post-hoc test *p*-values

	26 Weeks	**27 Weeks**	**28 Weeks**	**29 Weeks**	**30 Weeks**
26 Weeks		1	0.718	0.023	0.206
27 Weeks			0.959	0.376	0.612
28 Weeks				0.048	0.041
29 Weeks					0.003
30 Weeks					

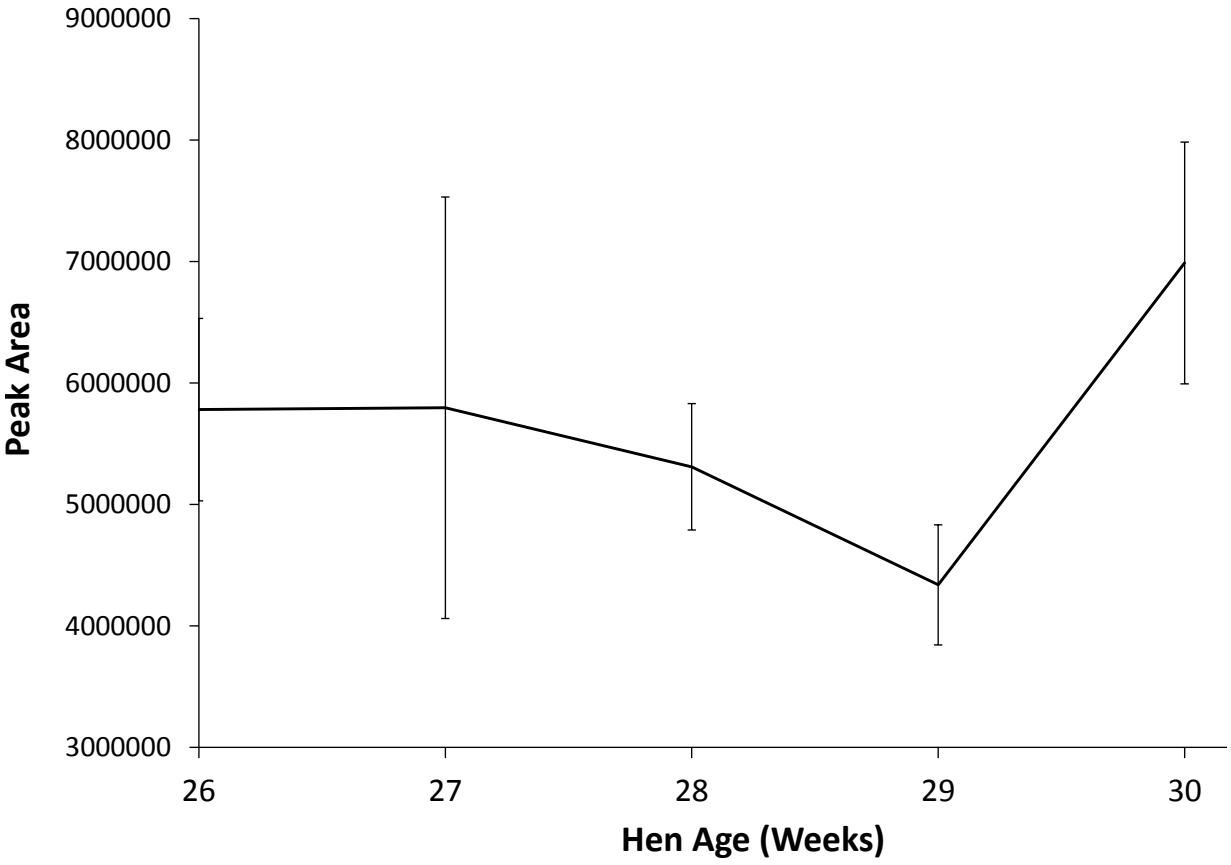


Figure 5.13: *Line graph showing the trend in changing abundance of compound m/z 454.2932 in yolk across eggs laid by hens from 26-30 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.14: *Table showing the putative identification of compound m/z 496.3405 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
496.3405	$C_{24}H_{50}NO_7P$	97.97	PC(16:0/0:0)	0.001

Post-hoc test *p*-values					
	26 Weeks	**27 Weeks**	**28 Weeks**	**29 Weeks**	**30 Weeks**
26 Weeks		0.978	0.324	0.033	0.606
27 Weeks			0.653	0.111	0.288
28 Weeks				0.760	0.018
29 Weeks					0.001
30 Weeks					

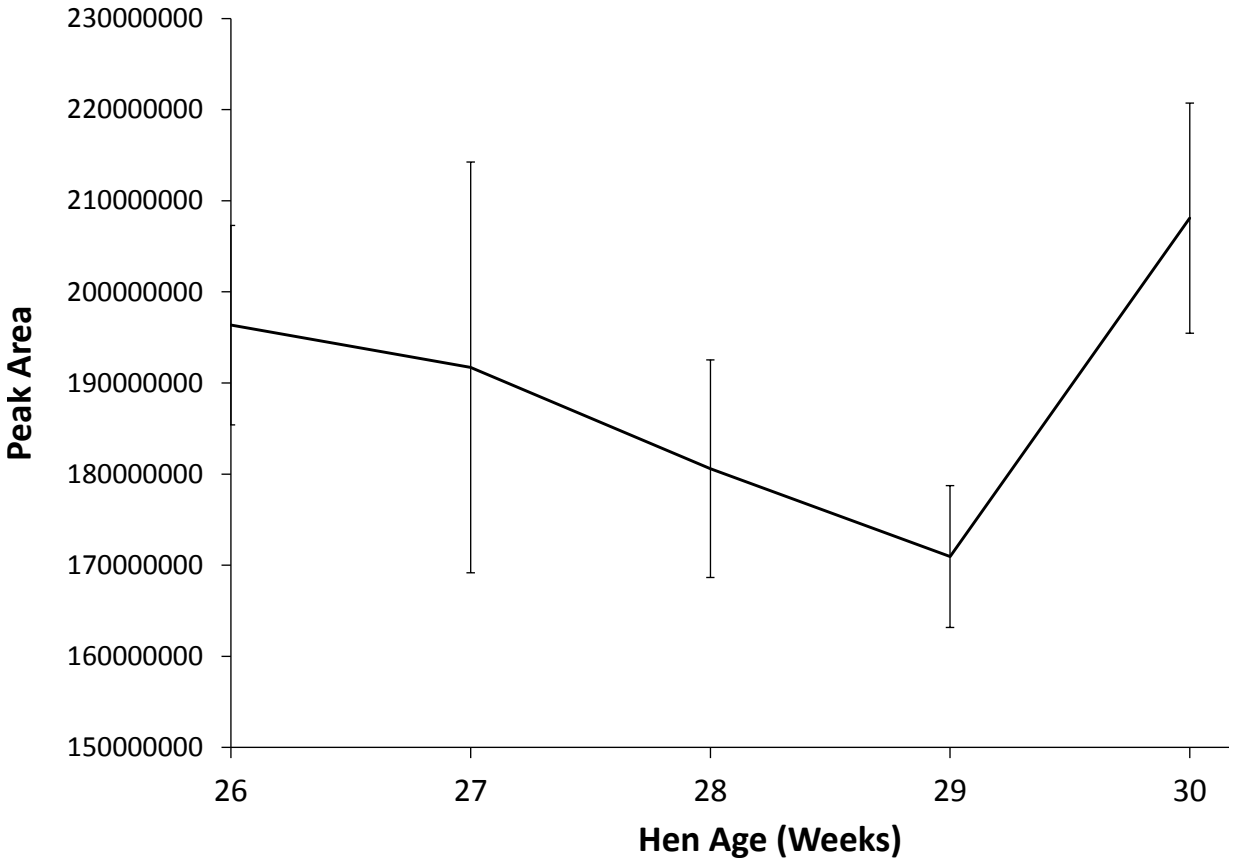


Figure 5.14: *Line graph showing the trend in changing abundance of compound m/z 496.3405 in yolk across eggs laid by hens from 26-30 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.15: *Table showing the putative identification of compound m/z 703.5758 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
703.5758	$C_{39}H_{79}N_2O_6P$	96.48	SM(18:1/16:0)	0.003

Post-hoc test *p*-values					
	26 Weeks	**27 Weeks**	**28 Weeks**	**29 Weeks**	**30 Weeks**
26 Weeks		1	0.237	0.994	0.020
27 Weeks			0.276	0.987	0.025
28 Weeks				0.114	0.751
29 Weeks					0.008
30 Weeks					

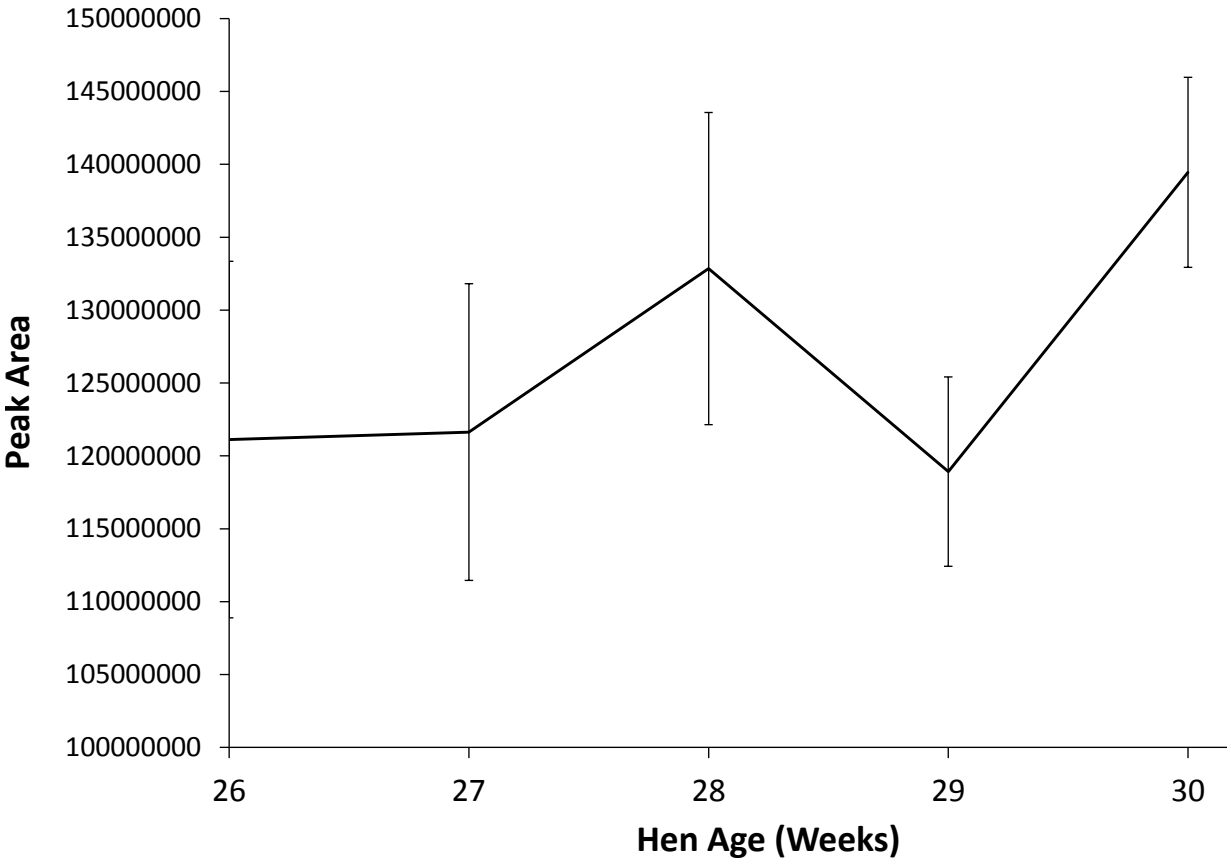


Figure 5.15: *Line graph showing the trend in changing abundance of compound m/z 703.5758 in yolk across eggs laid by hens from 26-30 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

The probability scores for the predicted molecular formulae for all of these compounds are above 95, indicating that there is a strong likelihood that the predicted formulae correspond to the compounds of interest. The p-values from ANOVA and Welch tests show that the abundances of these compounds are statistically significant when comparing them between eggs laid by birds of different ages, ranging from 26 to 30 weeks old, and the post-hoc test p-values corroborate this. Although all of these compounds show a high statistical significance in the p-values from the post-hoc tests, not all of them show a correlation between abundance in egg yolk and laying hen age, and therefore are not suitable as markers of hen age within this age range.

The compound that was putatively identified as SM(18:1/16:0), has some instances of statistical significance in abundance when comparing eggs laid by birds of different ages, as can be seen in Table 5.15. However, contrary to what was observed for this compound previously for hens aged between 21 and 26 weeks old, there is no correlation associated with these changes of abundance, as can be seen in Figure 5.15, and so the compound is not suitable as a marker of hen age for birds aged between 26 and 30 weeks old.

The compound that was putatively identified as erucamide, an amide derivative of the fatty acid, erucic acid, has also been determined to be unsuitable as a marker of hen age. Although this compound does decrease in abundance in egg yolk with increasing laying hen age, as seen in Figure 5.12, the only points of statistical significance are between eggs laid by birds that were 26 weeks old, and those laid by birds of all other ages, as shown in Table 5.12. This is due to the large decrease in abundance that can be observed in the yolks of eggs laid by birds that were 27 weeks old, compared to those laid by birds that were 26 weeks old. This limits the use of this compound as a marker of hen age.

The compound that was putatively identified as 1-monopalmitin, again has several instances of highly significant differences in abundance between eggs laid by birds of different ages, as seen in Table 5.11. However, contrary to what was observed previously, for laying hens ranging from 21 to 26 weeks old, a negative correlation was observed

between the abundance of this compound in egg yolk, and increasing laying hen age, as shown in Figure 5.11, making it potentially useful as a marker of hen age for birds ranging from 26 to 30 weeks old, but not for birds younger than this.

The remaining two putatively identified compounds, PE(16:0/0:0), a lysophosphatidylethanolamine with a palmitic acid group attached to the first carbon of the glycerol backbone, and PC(16:0/0:0), both show some statistical significance in abundance between the yolks of eggs laid by birds of different ages, as shown in Tables 5.13 and 5.14, and show very similar trends in changing abundance between eggs laid by birds of different ages, as seen in Figures 5.13 and 5.14. These compounds both show a decreasing abundance in eggs laid by birds ranging from 26 to 29 weeks old, followed by a significant increase in abundance in eggs laid by birds just one week older. This is likely to be because these compounds are very similar to each other, and both follow very similar metabolic pathways in their production from diacylglycerol, and their catabolism to glycerol-3-phosphate [23]. The rate of metabolism will be very similar between both compounds in birds of different ages, as very similar enzymes are involved in the metabolic pathways associated with these compounds. This results in eggs having similar abundances of both compounds, and similar trends being observed in the changing abundance of these compounds in eggs laid by birds of increasing age. Although these compounds show a trend in decreasing abundance with increasing hen age from 26 to 29 weeks old, there is not much statistical significance in the changing abundance within this age range, and so these compounds are not very suitable as markers of hen age.

Some of the standard deviation error bars on these line graphs are again quite wide. Again, some only appear large due to the scaling of the y-axis, but others are due to a larger standard deviation, because of the biological variation within the sets of eggs laid by birds of different ages, caused by the fact that they were likely laid by different birds. Larger sample sizes would reduce the standard deviations of the compound abundances in sets of eggs laid by birds of the same age.

5.6 Hen Age 30-38 weeks: Yolk Organic Extracts

Yolk organic extracts of eggs laid by hens ranging from 30 weeks to 38 weeks old were compared, to see whether hen age affects the metabolite profile of an egg.

5.6.1 Results and Discussion

5.6.1.1 Quality Control Analysis

Table 5.16 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 5.16: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 30-38 weeks old*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	1924651	216736268	70604441	27715741	605217644	128342976
QC2	1676742	194846964	65985602	26017109	554332604	120110223
QC3	1911792	180024727	64674638	25083677	541555242	112168260
QC4	1870664	159060662	59448575	21921311	492549583	96945950
QC5	1714219	156121292	60551088	22214854	512929982	95769405
QC6	1636161	168617245	61317467	22627169	504651293	92039800
QC7	1746380	150062789	53888075	20229767	458924411	89814339
SD	117800	23923860	5339818	2644852	47512754	15128532
Mean	1782944	175067135	62352841	23687090	524308680	105027279
CV%	6.61	13.67	8.56	11.17	9.06	14.40

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 5.17 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 5.17: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 30-38 weeks old*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.690	9.504	32.992	57.178	71.885	78.524
QC2	4.707	9.537	32.992	57.144	71.751	78.358
QC3	4.643	9.540	32.962	57.130	71.771	78.378
QC4	4.174	9.519	32.957	57.126	71.750	78.323
QC5	4.113	9.507	32.946	57.114	71.805	78.378
QC6	4.178	9.572	33.027	57.212	71.803	78.493
QC7	4.127	9.489	32.894	57.062	71.603	78.144
SD	0.286	0.028	0.042	0.048	0.086	0.124
Mean	4.376	9.524	32.967	57.138	71.767	78.371
CV%	0.65	0.03	0.01	0.01	0.01	0.02

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

5.6.1.2 Metabolite Profiling

Comparisons were made between yolk organic extracts of eggs laid by hens that were 30 weeks and 38 weeks old, by studying the TICs of the extracts following HPLC-MS analysis. Figure 5.16 shows overlaid TICs for the average chromatograms of extracts from eggs laid by hens of these two ages.

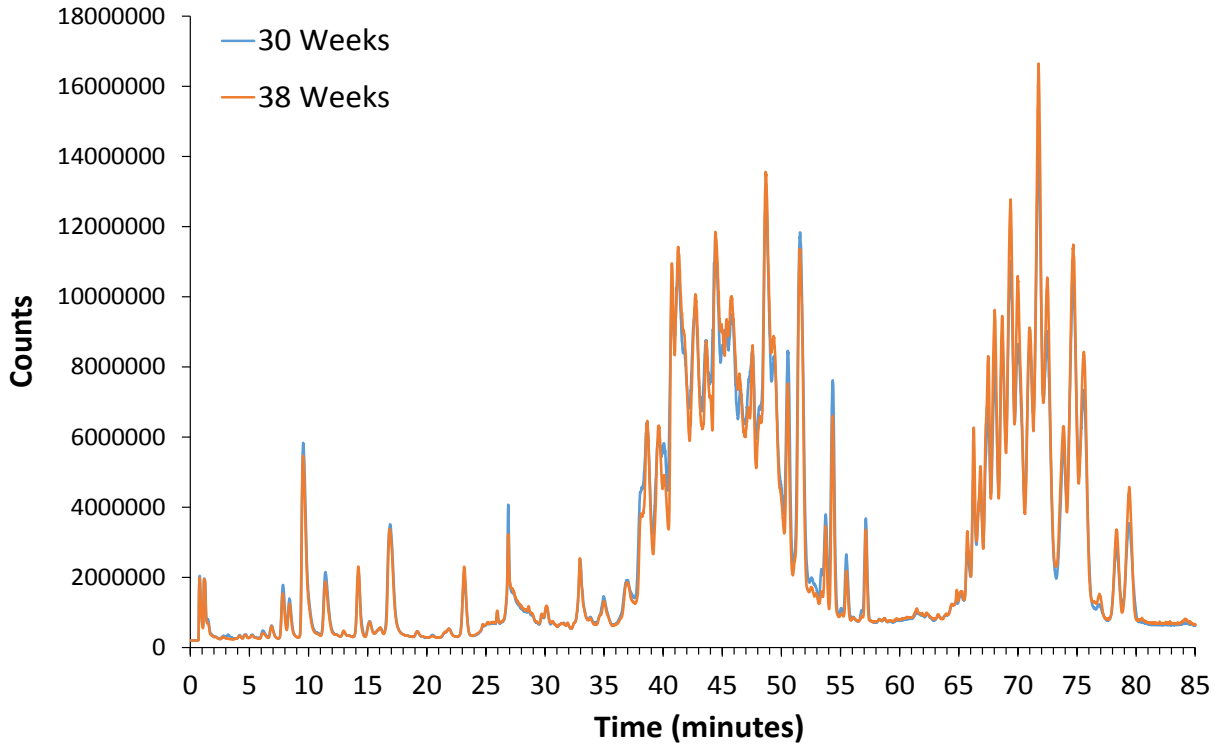


Figure 5.16: *Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 30 weeks and 38 weeks old*

There is again a slight visible difference in metabolite profile between eggs laid by birds that were 30 weeks old, and those laid by birds that were 38 weeks old, as can be seen in Figure 5.16. During the first 35 minutes of analysis, a slightly higher peak intensity can be observed in the TIC of the yolks of eggs laid by birds that were 30 weeks old, indicating that the earlier eluting compounds are more highly abundant in the yolks of eggs laid by younger birds. Later on in the analysis, between 35 and 65 minutes, approximately half of the peaks are of a higher intensity in the TIC of the yolks of eggs laid by the younger birds, and half are of a higher intensity in the TIC of the yolks of

eggs laid by the older birds. After 65 minutes there is a higher peak intensity for the TIC of the yolks of eggs laid by hens that were 38 weeks old. This indicates that the later eluting, non-polar molecules, such as potential phospholipids, are of a higher abundance in eggs laid by older birds, which corresponds with what was observed previously, for other hen age ranges.

5.6.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the eggs laid by hens of different ages.

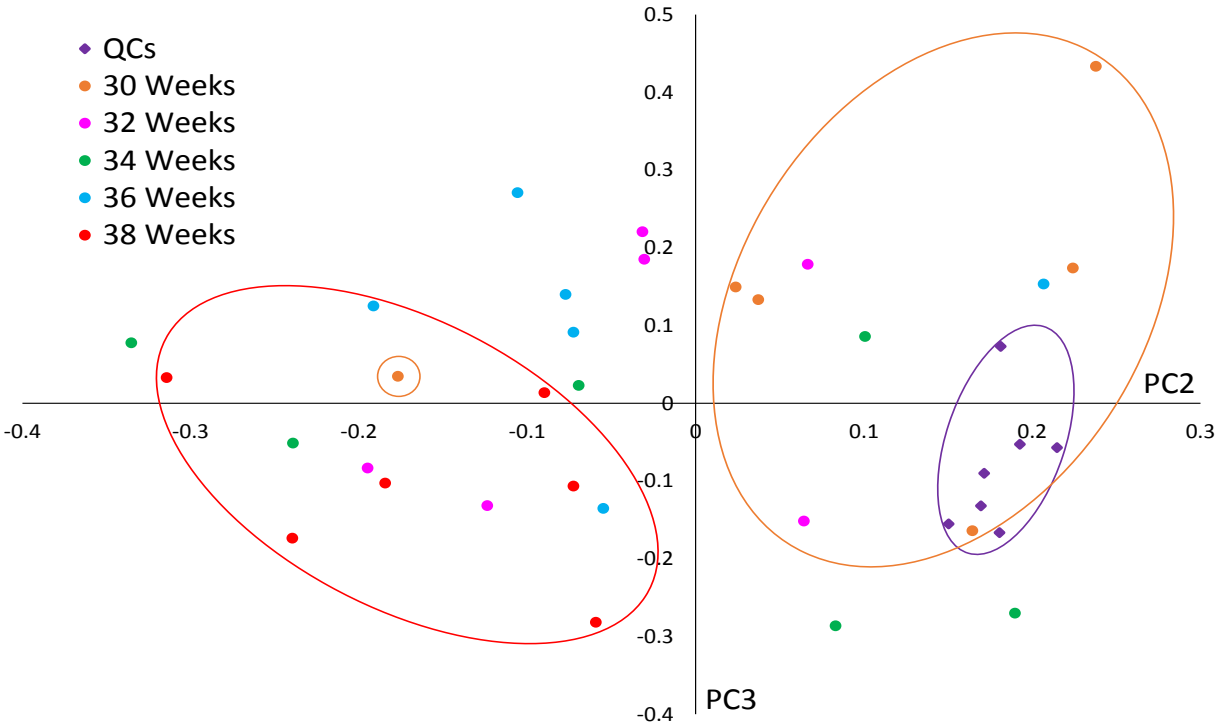


Figure 5.17: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 30 weeks to 38 weeks old, including all compounds with CV%<30%. PC2 explains 0.32% of the variance, and PC3 explains 0.22% of the variance.*

The scores plot in Figure 5.17 shows quite a large variation within each set of eggs laid by birds of different ages, with samples from each set spread across the whole plot. The eggs laid by birds that were 38 weeks old show the least amount of variation and only spread across two quadrants of the plot. These samples form a group that is almost

completely separate to the group of eggs laid by birds that were 30 weeks old at the point of lay, however one of the eggs (circled) laid by a 30 week old bird is quite separate to the rest of the group, and sits within the group of eggs laid by birds that were 38 weeks old. It is not known why this is, however it does show how subtle the differences are between sets of eggs laid by birds of these ages. The QC samples form a relatively tight group, showing that the analysis was robust and that the differences observed between samples on the scores plot are due to true biological differences, rather than instrumental drift.

5.6.1.4 Univariate Statistics

There were 43 compounds remaining following the removal of any duplicates, adducts and isotopes from the top 100 compounds, and 41 of these were found to be statistically significant. These compounds, and their CV%s and p-values, can be seen in Table B.3 in Appendix B; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

Of the 28 compounds that were still found to be statistically significant when confirmed using the raw data, two were putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The observed m/z for both compounds was due to the $[M+H]^+$ adduct. The compounds m/z 331.2848 and m/z 338.3412 were putatively identified as 1-monopalmitin and erucamide respectively. Both compounds were previously putatively identified when found to be in the top statistically significant compounds between eggs laid by birds within other ranges of hen age, and their comparisons between mass spectra provided by METLIN and mass spectra resulting from chemical analysis can be seen in Figures B.1 and B.7 in Appendix B.

Tables 5.18 and 5.19, and Figures 5.18 and 5.19, show the putative identifications of these compounds, the p-values resulting from ANOVA and post-hoc Tukey tests, as well as the trends of changing abundance of the compounds in eggs laid by birds of increasing age from 30 to 38 weeks old.

Table 5.18: *Table showing the putative identification of compound m/z 331.2848 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
331.2848	$C_{19}H_{38}O_4$	99.45	1-monopalmitin	0.001

Post-hoc test *p*-values

	30 Weeks	32 Weeks	34 Weeks	36 Weeks	38 Weeks
30 Weeks		0.413	0.011	0.001	0.166
32 Weeks			0.379	0.053	0.978
34 Weeks				0.814	0.718
36 Weeks					0.166
38 Weeks					

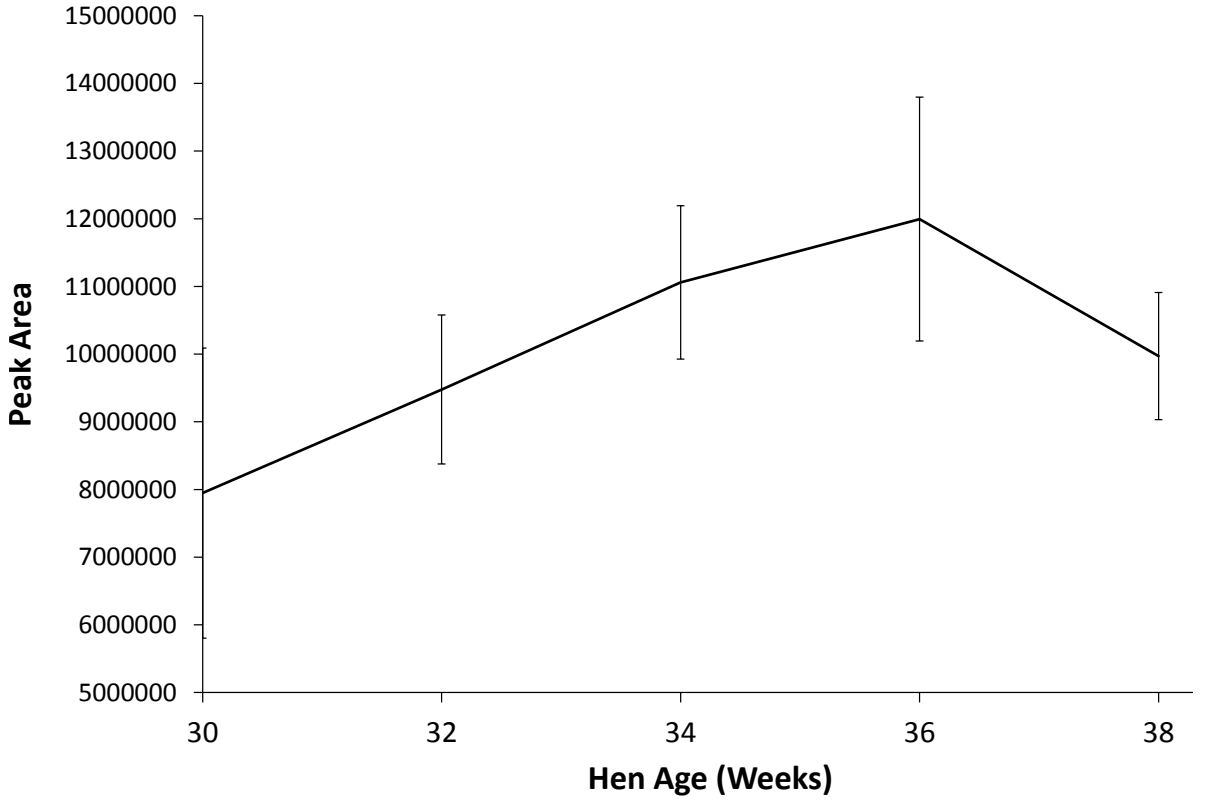


Figure 5.18: *Line graph showing the trend in changing abundance of compound m/z 331.2848 in yolk across eggs laid by hens from 30-38 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.19: *Table showing the putative identification of compound m/z 338.3412 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
338.3412	$C_{22}H_{43}NO$	90.09	Erucamide	0.003

Post-hoc test *p*-values

	30 Weeks	**32 Weeks**	**34 Weeks**	**36 Weeks**	**38 Weeks**
30 Weeks		0.728	0.153	0.998	0.004
32 Weeks			0.783	0.883	0.075
34 Weeks				0.264	0.509
36 Weeks					0.009
38 Weeks					

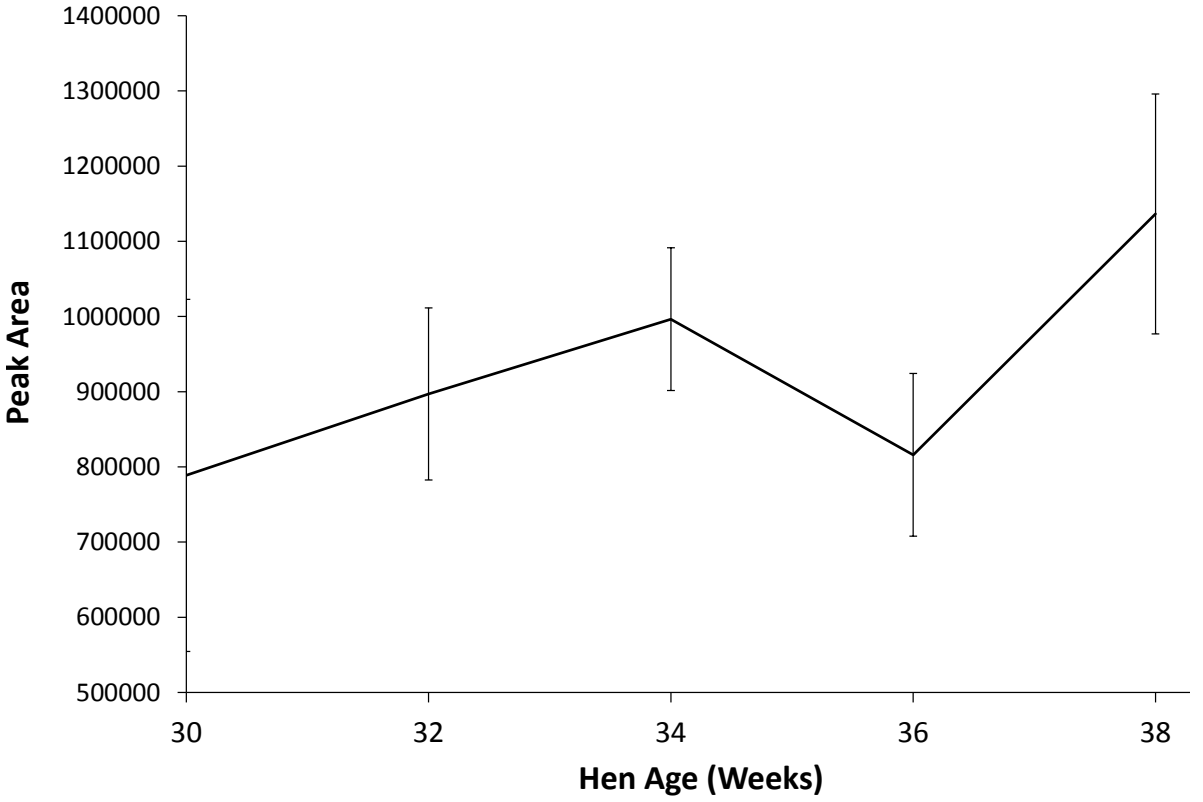


Figure 5.19: *Line graph showing the trend in changing abundance of compound m/z 338.3412 in yolk across eggs laid by hens from 30-38 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

The ANOVA p-values show that the abundances of these compounds are statistically significant when comparing them between eggs laid by birds ranging from 30 to 38 weeks old. However, the compound that was putatively identified as 1-monopalmitin only shows one instance of high statistical significance, and one of moderate statistical significance, in the p-values from the post-hoc tests shown in Table 5.18, meaning that it may not be useful as a marker of hen age for laying hens within this age range. Although there are not many instances of statistical significance for this compound between eggs laid by birds ranging from 30 to 38 weeks old, there is a positive correlation between the abundance of this compound in egg yolk and laying hen age, up to an age of 36 weeks, as can be seen in Figure 5.18. This is opposite to the trend that was observed for the abundance of this compound in eggs laid by hens ranging from 26 to 30 weeks old, indicating that this compound is quite erratic in its changing abundance with increasing hen age, suggesting that it may not be useful as a marker of hen age over a wider age range.

The compound that was putatively identified as erucamide has just two instances of high statistical significance shown by the post-hoc test p-values in Table 5.19, and does not show a correlation between abundance in egg yolk and laying hen age, as can be seen in Figure 5.19, therefore it is not very useful as a marker of hen age within this age range. Contrary to what was observed in the previous study for hens aged between 26 and 30 weeks old, it shows an overall increase in abundance with increasing hen age. This indicates that the changing abundance is quite erratic between eggs laid by birds of different ages, even over wider age ranges.

There is again quite a high amount of variation within the sample sets. Although the scaling of the y-axis makes the error bars look larger, they are still quite wide which is probably due to the fact that the eggs were likely laid by different birds. Larger sample sizes would reduce the standard deviations of the compound abundances within sets of eggs laid by birds of the same age.

5.7 Hen Age 38-44 weeks: Yolk Organic Extracts

Yolk organic extracts of eggs laid by hens ranging from 38 weeks to 44 weeks old were compared, to see whether hen age affects the metabolite profile of an egg.

5.7.1 Results and Discussion

5.7.1.1 Quality Control Analysis

Table 5.20 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 5.20: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 38-44 weeks old*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	2030876	166427120	58481476	30957559	442511518	156220416
QC2	2002013	165578586	59459806	29891511	442026394	166301625
QC3	1864841	157068754	54624220	29826917	437674783	151674572
QC4	1924249	165769955	53862239	29488680	415085737	150185602
QC5	1893351	160187803	56816698	29553300	461377202	141156195
QC6	1924067	152026397	58145270	28438801	411032390	143113124
QC7	2144994	152542858	54522037	29161724	420826408	132059410
SD	97027	6238496	2230783	767297	18031112	11123963
Mean	1969199	159943068	56558821	29616927	432933490	148672992
CV%	4.93	3.90	3.94	2.59	4.16	7.48

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that any instrumental drift occurring throughout the analysis was minimal.

Table 5.21 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 5.21: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 38-44 weeks old*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.639	9.685	33.174	57.309	71.983	78.722
QC2	4.770	9.700	33.205	57.357	72.031	78.837
QC3	4.773	9.670	3.175	57.344	72.051	78.823
QC4	4.777	9.707	33.312	57.497	72.204	79.109
QC5	4.739	9.703	33.274	57.476	72.183	79.088
QC6	4.789	9.719	33.257	57.426	72.150	78.955
QC7	4.754	9.701	33.206	57.391	72.082	78.887
SD	0.051	0.016	11.363	0.070	0.083	0.142
Mean	4.749	9.698	28.943	57.400	72.098	78.917
CV%	1.07	0.16	0.16	0.12	0.12	0.18

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

5.7.1.2 Metabolite Profiling

Comparisons were made between yolk organic extracts of eggs laid by hens that were 38 weeks and 44 weeks old, by studying the TICs of the extracts following HPLC-MS analysis. Figure 5.20 shows overlaid TICs for the average chromatograms of extracts from eggs laid by hens of these two ages.

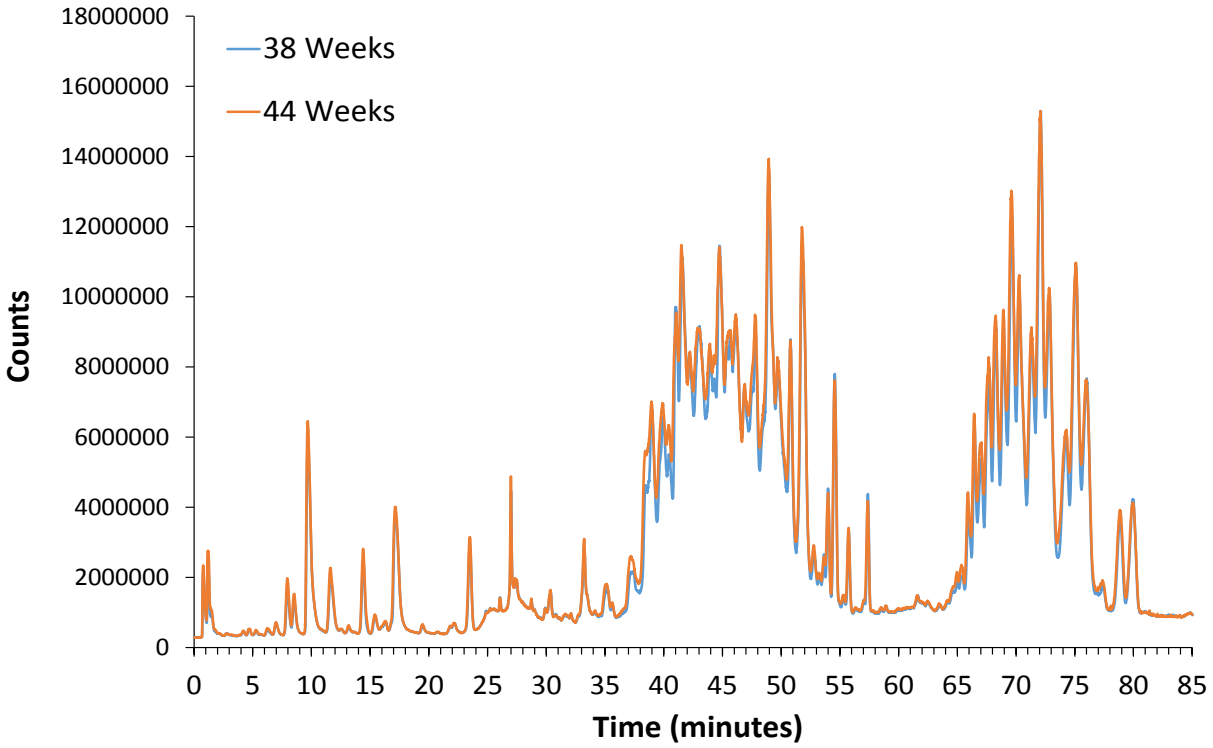


Figure 5.20: *Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 38 weeks and 44 weeks old*

There is again a slight visible difference in metabolite profile between the yolks of eggs laid by birds that were 38 weeks old, and those laid by birds that were 44 weeks old, as can be seen in Figure 5.20. These overlaid TICs show that there is a higher peak intensity for most peaks, throughout the whole analysis, in the chromatogram of yolk extracts from eggs that were laid by the older birds, compared to that of yolk extracts from eggs laid by birds that were 38 weeks old. This indicates that within this age range, there is a greater abundance of most compounds obtained from an organic extraction of the yolk, in eggs laid by the older birds than the younger birds, compared to just the

higher abundance of non-polar, potential phospholipid molecules in eggs laid by older birds, which was observed when studying previous hen age ranges within this research.

5.7.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the eggs laid by hens of different ages.

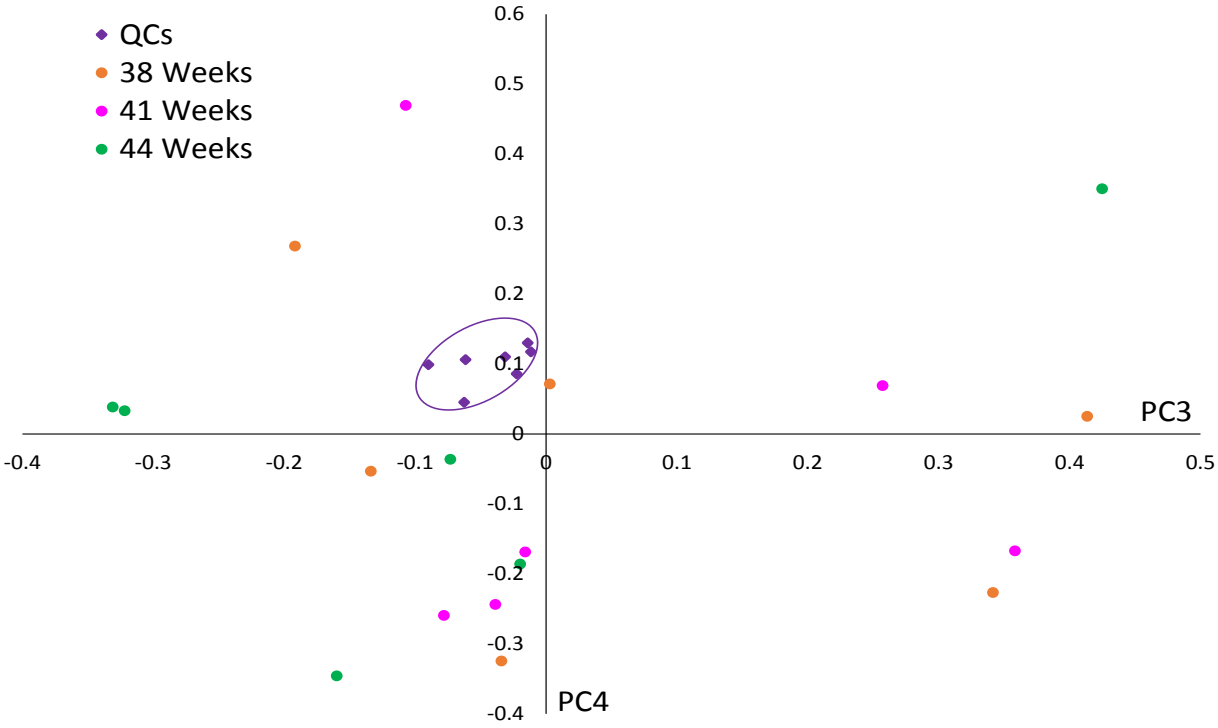


Figure 5.21: *PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by hens from 38 weeks to 44 weeks old, including all compounds with CV%<30%. PC3 explains 0.23% of the variance, and PC4 explains 0.16% of the variance.*

The scores plot in Figure 5.21, contrary to what was observed in the overlaid TICs in Figure 5.20, shows no separation between the eggs laid by birds of different ages, with all three sets of eggs spread throughout the whole plot. This could be because the more highly abundant compounds that contribute to the observed differences in the overlaid TICs may not contribute as much to the variance explained by PC3 and PC4 as the less abundant compounds, which may experience more variation between random egg samples, rather than between eggs laid by birds of different ages.

The lack of separation indicates that the variance described by PC3 (0.23%) and PC4 (0.16%) is greater between random eggs than between eggs laid by birds of different ages, showing that the differences in metabolite profile between eggs laid by birds of different ages, within this age range, are very subtle. The QC samples are clustered tightly together towards the centre of the plot, showing that there was little instrumental drift and that the separation between samples on the plot is due to true biological differences, rather than instrument bias.

5.7.1.4 Univariate Statistics

There were 57 compounds remaining following the removal of any duplicates, adducts and isotopes from the top 100 compounds, and all 57 of these were found to be statistically significant. These compounds, and their CV%s and p-values, can be seen in Table B.4 in Appendix B; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

Of the 44 compounds that were still found to be statistically significant when confirmed using the raw data, none were successfully putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis.

5.7.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using just the top 44 compounds, in order to observe whether this resulted in any separation between eggs laid by different ages of bird, as all of these compounds were found to be statistically significant when comparing eggs laid by birds of different ages.

The scores plot in Figure 5.22 shows complete separation between eggs laid by birds that were 38, 41, and 44 weeks old. This indicates that the statistical significance of the top 44 compounds is between eggs laid by birds of all three ages. The separation

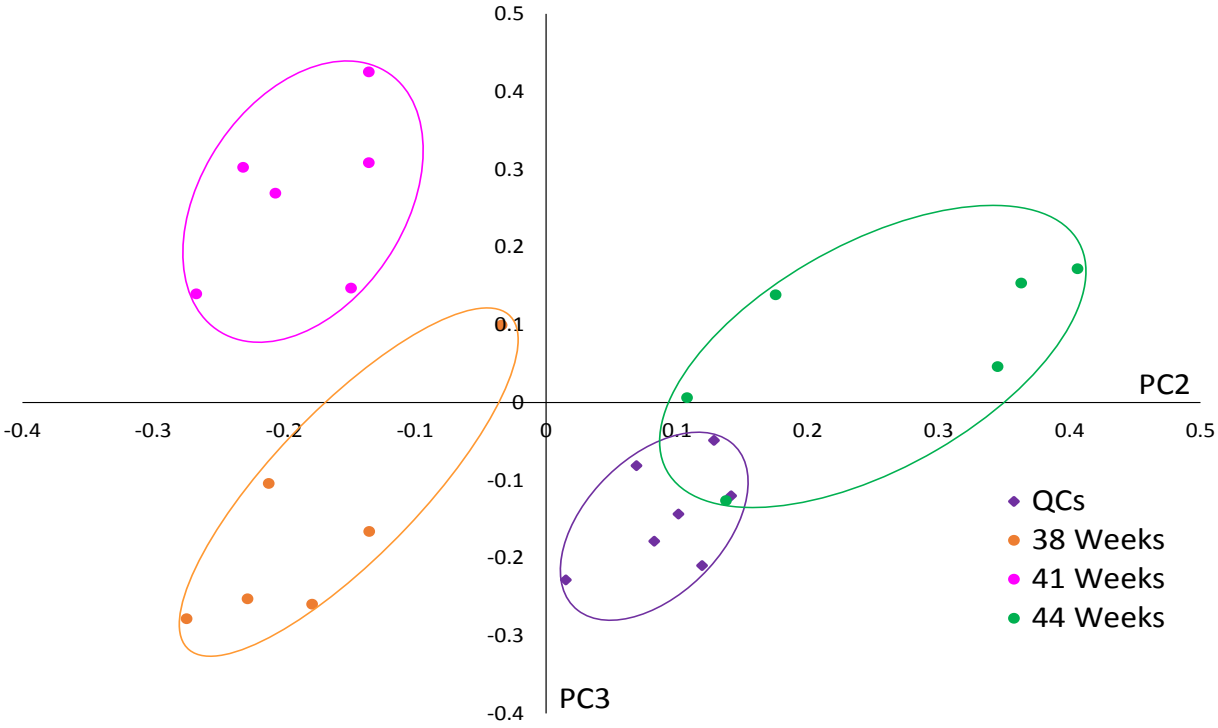


Figure 5.22: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 38 weeks to 44 weeks old, including all compounds with CV%<30%. PC2 explains 1.07% of the variance, and PC3 explains 0.38% of the variance.*

between all three sets of eggs laid by birds of different ages is across both PC2 and PC3, which describe 1.07% and 0.38% of the total variance respectively. This shows that, although the variance described by both PCs, and the separation between sample sets, is greater when only the top statistically significant compounds are included in the PCA, the differences between eggs laid by birds of different ages within this age range are still very subtle.

The QC samples are grouped more closely together than the other sample sets, indicating that the differences observed between samples are true biological differences, and not caused by instrumental drift.

5.8 Hen Age 21-44 weeks (21, 26, 30, 38, 44 weeks): Yolk Organic Extracts

5.8.1 Results and Discussion

5.8.1.1 Quality Control Analysis

Table 5.22 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 5.22: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 21-44 weeks old*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	9533445	226629258	56638645	48459694	328725577	94625879
QC2	9211706	219789419	54750441	47350638	334120744	83765914
QC3	9256130	226239364	58539237	48420121	347502651	85363385
QC4	9510922	219068800	55035906	48134354	345170326	85398671
QC5	9626331	225806875	56471569	48377649	334998135	79800066
QC6	9329760	226224449	58883221	49024893	342997547	80287370
QC7	9903791	229295444	56607194	48545140	317764689	73038655
SD	241342	3798678	1569697	508888	10450491	6629218
Mean	9481726	224721944	56703745	48330356	335897096	83182849
CV%	2.55	1.69	2.77	1.05	3.11	7.97

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 5.23 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 5.23: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens 21-44 weeks old*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.526	9.141	32.280	56.333	70.592	76.501
QC2	4.524	9.172	32.344	56.480	70.838	76.864
QC3	4.525	9.156	32.262	56.331	70.623	76.549
QC4	4.567	9.232	32.371	56.473	70.782	76.791
QC5	4.570	9.218	32.391	56.460	70.818	76.827
QC6	4.562	9.226	32.416	56.554	70.893	76.985
QC7	4.545	9.193	32.349	56.485	70.810	76.786
SD	0.021	0.036	0.056	0.083	0.113	0.173
Mean	4.546	9.191	32.345	56.445	70.765	76.758
CV%	0.46	0.39	0.17	0.15	0.16	0.23

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

5.8.1.2 Metabolite Profiling

Comparisons were made between yolk organic extracts of eggs laid by hens that were 21 weeks and 44 weeks old, by studying the TICs of the extracts following HPLC-MS analysis. Figure 5.23 shows overlaid TICs for the average chromatograms of extracts from eggs laid by hens of these two ages.

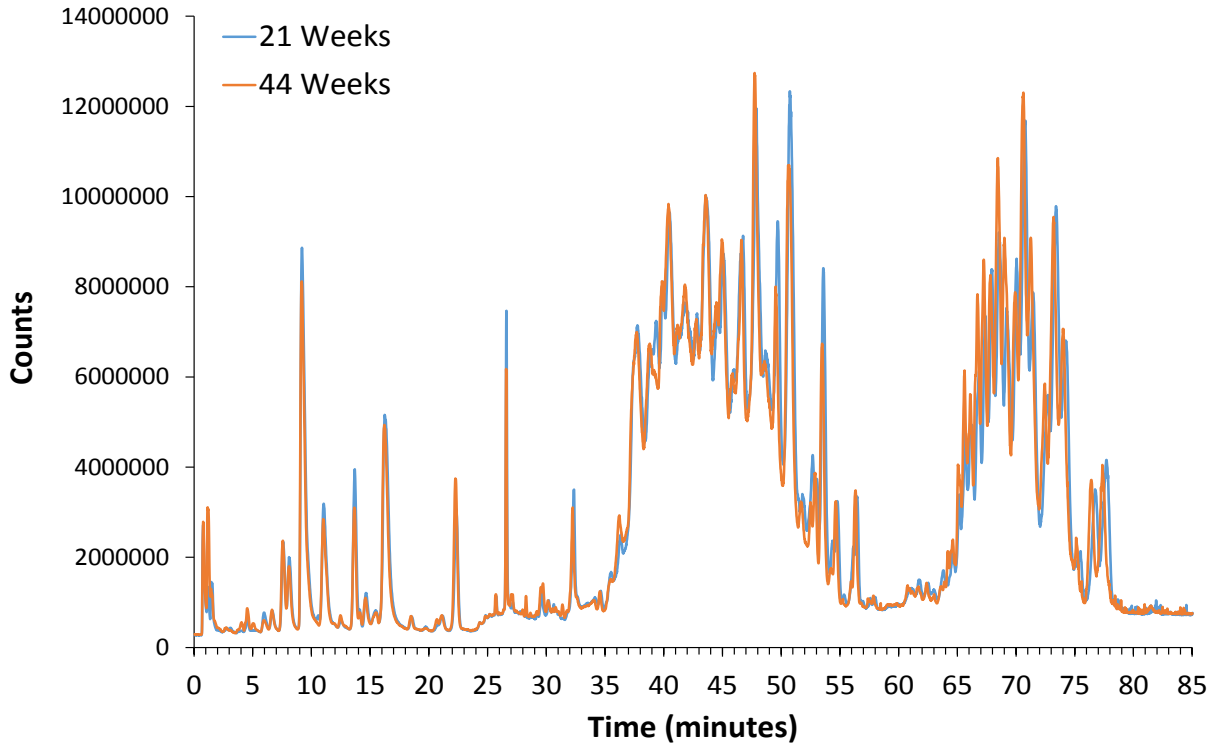


Figure 5.23: *Overlaid average TICs of yolk organic extracts of eggs that were laid by hens of 21 weeks and 44 weeks old*

The overlaid TICs in Figure 5.23 display a clear difference in metabolite profile between eggs laid by birds that were 21 weeks old, and those laid by birds that were 44 weeks old, showing that there is a difference in metabolite profile between eggs laid by birds of different ages over a wider age range than has previously been studied in this work. Over the first 35 minutes of analysis there is a higher peak intensity for most peaks in the TIC of the yolks of eggs laid by birds that were 21 weeks old at the point of lay, indicating that the smaller, and less non-polar compounds in an organic extract of egg yolk, are of a higher abundance in eggs laid by younger birds compared to those laid

by older birds. Between 35 and 65 minutes of analysis, there is little difference in peak intensity between the two TICs; the peak intensity is similar for most peaks throughout both chromatograms, and approximately half of the peaks that do show a difference are of a higher intensity in the TIC of the yolks of eggs laid by the younger birds, and half are of a higher intensity in the TIC of the yolks of eggs laid by the older birds. Throughout the last 20 minutes of the chromatograms, most peaks have a higher peak intensity in the TIC of the yolks of eggs laid by birds that were 44 weeks old at the point of lay. This, again, indicates that the more non-polar molecules, such as potential phospholipids, that elute later on in the analysis, are of a higher abundance in eggs laid by older birds.

5.8.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the eggs laid by hens of different ages.

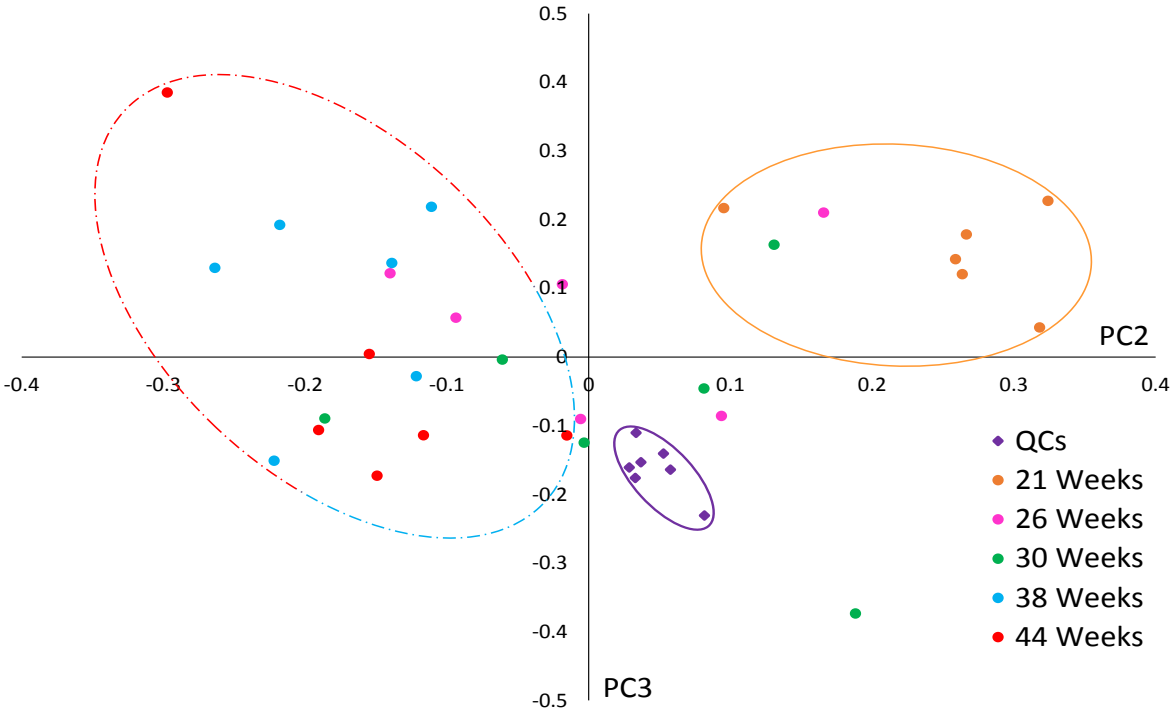


Figure 5.24: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by hens from 21 weeks to 44 weeks old, including all compounds with CV%<30%. PC2 explains 0.49% of the variance, and PC3 explains 0.24% of the variance.*

The scores plot in Figure 5.24 shows complete separation between eggs laid by young birds that were 21 weeks old, and those laid by older birds that were 38 and 44 weeks old, confirming what was observed in Figure 5.23, that there is a difference in the metabolite profiles of eggs laid by birds at the start of their laying cycle, and those that are several months older. The eggs laid by birds that were 38 and 44 weeks old are grouped together on the plot, indicating that the differences between eggs laid by birds of these two ages are less than the differences between eggs laid by birds of these ages, and eggs laid by birds that were 21 weeks old. The eggs laid by birds that were 26 and 30 weeks old at the point of lay are widely spread throughout the plot, indicating that the differences between random eggs within these sample sets are greater than any difference caused by laying hen age.

The QC samples form a tight cluster towards the centre of the plot, indicating that the analysis was robust, and that the differences observed on the scores plot between random eggs, and eggs laid by birds of different ages, are due to true biological differences, not instrumental drift.

5.8.1.4 Univariate Statistics

There were 58 compounds remaining following the removal of any duplicates, adducts and isotopes from the top 100 compounds, and all 58 of these were found to be statistically significant. These compounds, and their CV%s and *p*-values, can be seen in Table B.5 in Appendix B; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

Of the 51 compounds that were still found to be statistically significant when confirmed using the raw data, four were putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The observed *m/z* for all compounds was due to the $[M+H]^+$ adduct. The compounds *m/z* 331.2851, 454.2937, 524.3717, and 703.5762 were putatively identified as 1-monopalmitin, 1-palmitoyl-glycero-3-phosphatidylethanolamine PE(16:0/0:0), 1-palmitoyl-2-acetyl-glycero-3-phosphatidylcholine (Platelet Activating Factor (PAF) C-16), and 1-palmitoyl-sphingomyelin (SM(18:1/16:0)) respectively. The compounds *m/z* 331.2851, 454.2937, and 703.5762 were previously putatively identified in this work, and their comparisons between mass spectra provided by METLIN and mass spectra resulting from chemical analysis can be seen in Figures B.1, B.8, and B.4 respectively, in Appendix B. The mass spectra comparison for the remaining compound, *m/z* 524.3717, can be seen in Appendix B, Figure B.9.

Tables 5.24-5.27, and Figures 5.25-5.28, show the putative identifications of all of these compounds, the *p*-values resulting from ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests, as well as the trends of changing abundance of the compounds in eggs laid by birds of increasing age from 21 to 44 weeks old.

Table 5.24: *Table showing the putative identification of compound m/z 331.2851 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
331.2851	$C_{19}H_{38}O_4$	97.14	1-monopalmitin	<0.001

Post-hoc test *p*-values					
	21 Weeks	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		0.994	<0.001	<0.001	0.005
26 Weeks			<0.001	<0.001	0.002
30 Weeks				0.036	<0.001
38 Weeks					0.400
44 Weeks					

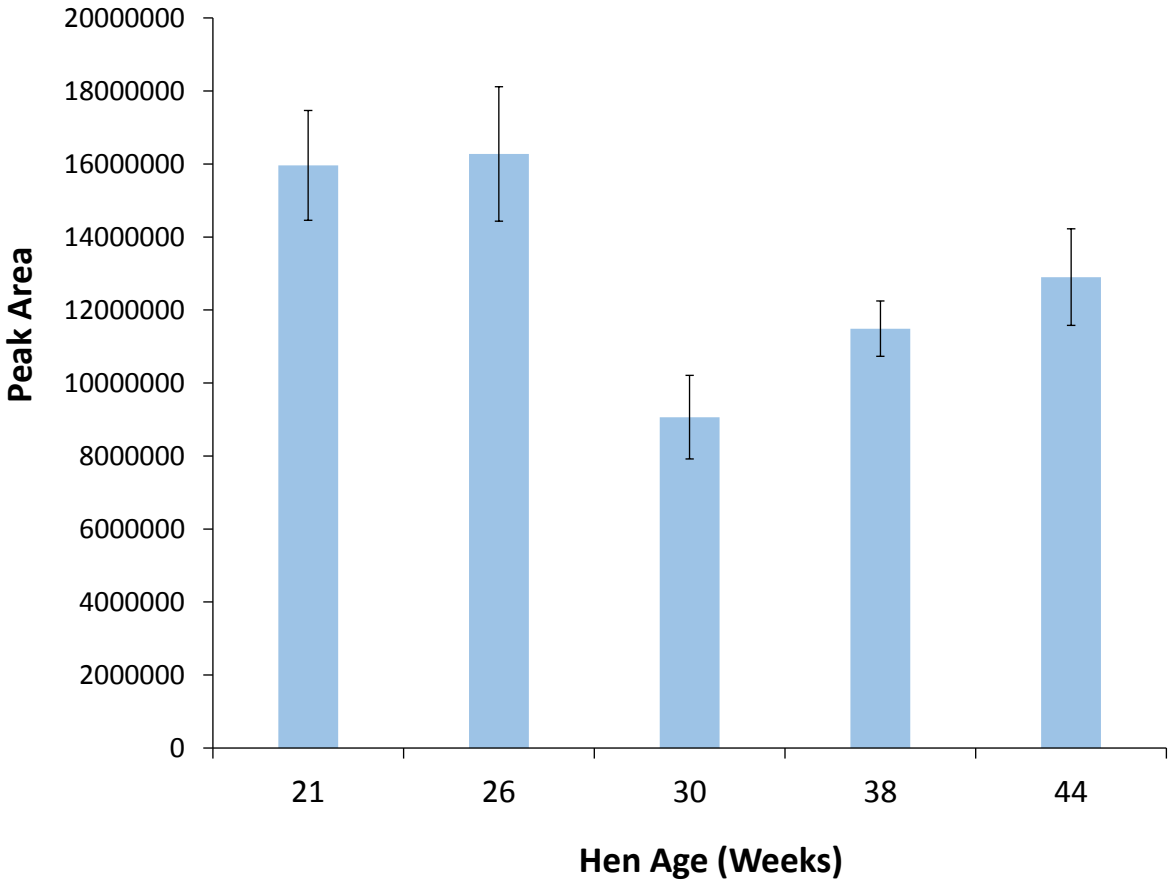


Figure 5.25: *Bar chart showing the trend in changing abundance of compound m/z 331.2851 in yolk across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.25: *Table showing the putative identification of compound m/z 454.2937 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
454.2937	$C_{21}H_{44}NO_7P$	98.56	PE(16:0/0:0)	0.001

Post-hoc test *p*-values					
	21 Weeks	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		0.002	0.015	0.002	0.006
26 Weeks			0.915	1	0.992
30 Weeks				0.902	0.994
38 Weeks					0.989
44 Weeks					

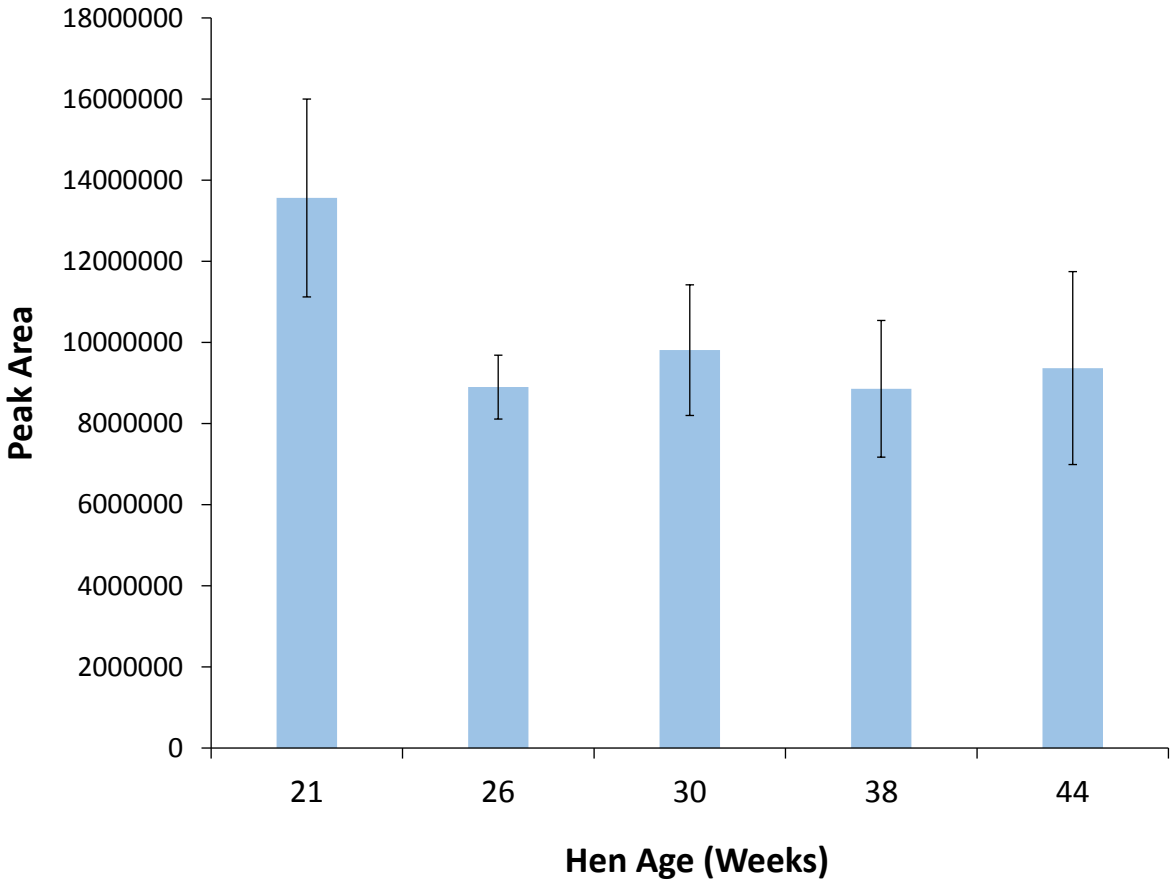


Figure 5.26: *Bar chart showing the trend in changing abundance of compound m/z 454.2937 in yolk across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.26: *Table showing the putative identification of compound m/z 524.3717 and the p-values resulting from Welch test and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	Welch *p*-value
524.3717	$C_{26}H_{54}NO_7P$	99.42	PAF C-16	0.005

Post-hoc test *p*-values

	21 Weeks	26 Weeks	30 Weeks	38 Weeks	44 Weeks
21 Weeks		0.873	0.411	0.022	0.089
26 Weeks			0.582	0.027	0.063
30 Weeks				0.734	0.997
38 Weeks					0.691
44 Weeks					

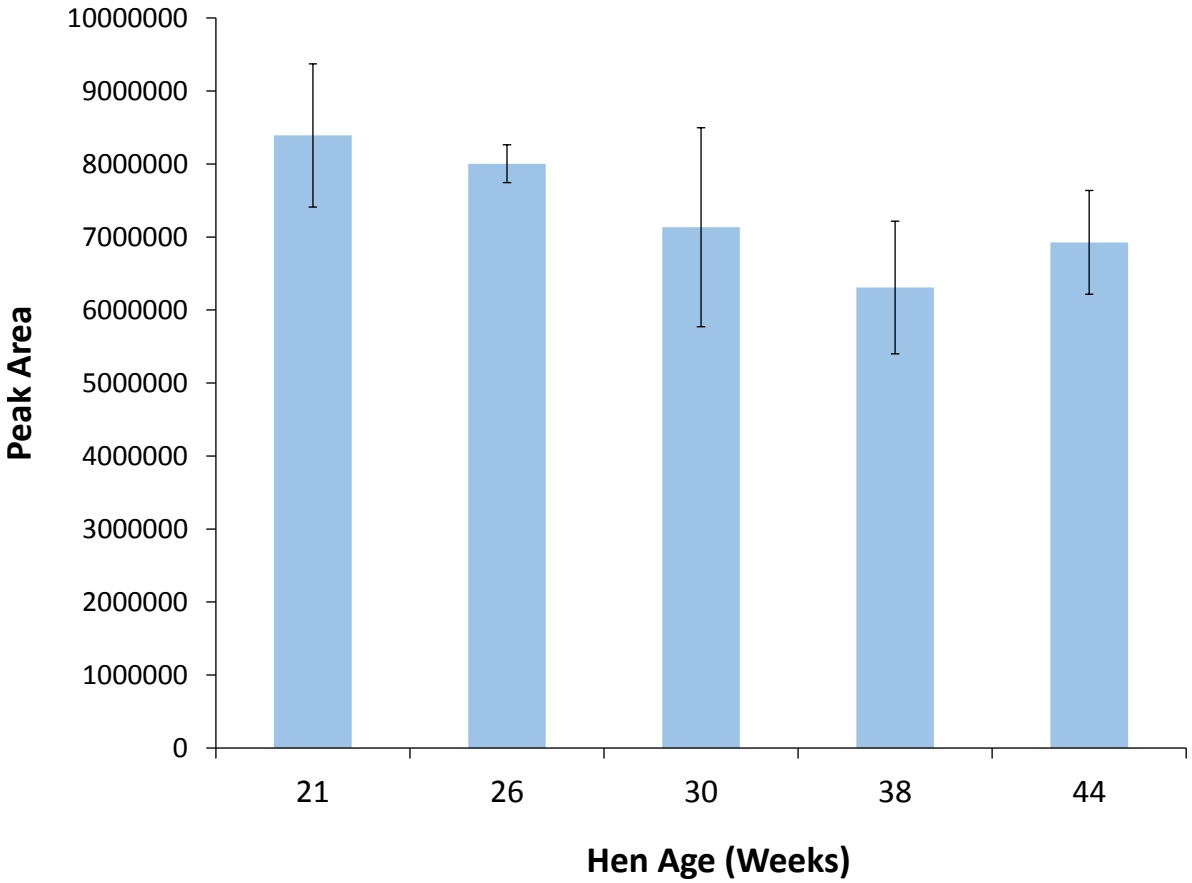


Figure 5.27: *Bar chart showing the trend in changing abundance of compound m/z 524.3717 in yolk across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.27: *Table showing the putative identification of compound m/z 703.5762 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
703.5762	$C_{39}H_{79}N_2O_6P$	95.38	SM(18:0/16:0)	<0.001

Post-hoc test *p*-values					
	21 Weeks	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		<0.001	0.006	0.003	0.001
26 Weeks			0.493	0.658	0.962
30 Weeks				0.999	0.868
38 Weeks					0.956
44 Weeks					

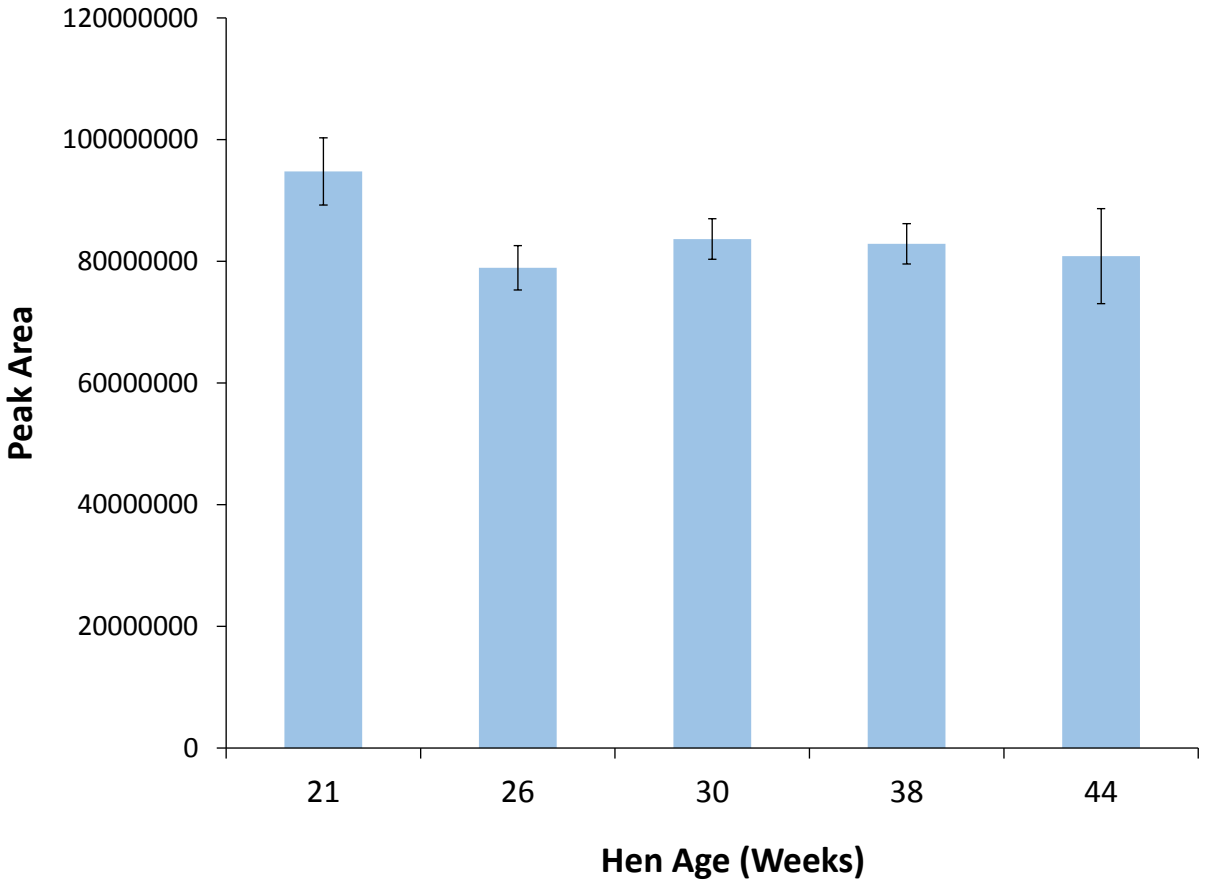


Figure 5.28: *Bar chart showing the trend in changing abundance of compound m/z 703.5762 in yolk across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

The probability scores for the predicted molecular formulae are all above 95, therefore there is a high likelihood that the predicted formulae correspond to the compounds of interest. The p-values from ANOVA and Welch tests show that the abundances of these compounds are statistically significant when comparing them between eggs laid by birds of different ages, ranging from 21 weeks old to 44 weeks old, and the post-hoc test p-values tend to corroborate this. Three of these compounds have been previously putatively identified in this work, when studying metabolite profile differences in egg yolk between eggs laid by birds within smaller age ranges. Not only did these compounds show significant differences between eggs laid by birds of different ages on a small timescale, but they also are statistically significant over a larger timescale of increasing bird age.

The compound that was putatively identified as 1-monopalmitin was in the top statistically significant compounds in several of the previous ranges of hen age that were studied. There was some trend in the changing abundance of this compound over some of these age ranges; however there is no correlation between abundance and hen age over the total range of hen age, as can be seen in Figure 5.25. Therefore, this compound is not a suitable marker of hen age for laying hens ranging from 21 to 44 weeks old. However, Table 5.24 shows that there is a high statistical significance in the abundance of this compound between eggs laid by birds that were 26 and 30 weeks old, and Figure 5.25 shows that there is a decrease in abundance in egg yolk between eggs laid by birds of these two ages. This confirms what was observed previously in Table 5.11 and Figure 5.11 in Section 5.5, therefore this compound may be suitable as a marker of hen age within this smaller range. The post-hoc test p-values in Table 5.24 show several points of high statistical significance between eggs laid by birds of different ages, and Figure 5.25 shows a gradual increase in the abundance of this compound in eggs laid by birds increasing from 30 to 44 weeks of age, which was also observed over the smaller hen age range of 30-38 weeks in Section 5.6. This suggests that this compound may have some use as a marker of hen age for laying hens between 30 and 44 weeks of age. However, although the abundance of this compound was statistically significant between eggs laid by birds of 30

and 38 weeks old when studied over the whole hen age range, as seen in Table 5.24, no significant difference was observed between eggs laid by birds of these ages in the smaller hen age range study in Section 5.6. This suggests that the statistical significance of this compound between eggs laid by birds of different ages is not stable and reproducible between analytical runs, and therefore this compound may not be suitable as a marker of hen age.

The post-hoc test p-values for the compound that was putatively identified as PE(16:0/0:0), in Table 5.25, show several points of strong significant difference in compound abundance, but only between eggs laid by birds that were 21 weeks old, and those laid by birds that were 26 weeks old and above, indicating that the greatest difference in compound abundance is between eggs laid by birds of these two ages. This is confirmed by observing the graph in Figure 5.26, which shows a large difference in abundance between eggs laid by birds that were 21 and 26 weeks old, and little difference and no correlation between abundance and hen age, for eggs laid by birds of other ages. This compound was previously determined to not be suitable as a marker of hen age for hens in the age range of 26 to 30 weeks old, due to low statistical significance, and can now also be determined to be unsuitable as a marker of hen age over a wider age range of laying bird.

The compound PAF C-16 is a unique ether glycerophosphatidylcholine, with a palmitoyl group attached to the first carbon of the glycerol backbone by an ether linkage, and an acetyl group attached to the second carbon by an ester bond. The Welch test p-value for this compound, seen in Table 5.26, shows a high statistical significance in compound abundance between eggs laid by birds of different ages; however, there are only two instances of moderate statistical significance shown by the post-hoc test p-values in Table 5.26. Although the graph in Figure 5.27 shows some trend in decreasing abundance of this compound between eggs laid by birds that were 21 weeks old and those laid by birds that were 38 weeks old, because there is only a slight statistical significance between the

sets of eggs, this compound is not useful as a marker of hen age.

The compound that was putatively identified as SM(18:1/16:0) was previously found to be in the top statistically significant compounds when studying smaller hen age ranges of 21 to 26 weeks and 26 to 30 weeks old, where it was found to be significantly different between eggs laid by the youngest and oldest birds in the age range. It was determined to be suitable as a marker of hen age for laying hens ranging from 21 to 26 weeks, as it showed a statistically significant negative correlation between its abundance in yolk and laying hen age. However, it was determined to not be suitable as a marker of hen age for hens between 26 to 30 weeks old, as no correlation was observed between its abundance in egg yolk and hen age within this age range. The significant difference in the abundance of this compound between eggs laid by birds that were 21 and 26 weeks old was confirmed when studying the whole range of hen age, as can be seen in Table 5.27. However, both the post-hoc test p-values in Table 5.27, and the graph in Figure 5.28, show that over the whole hen age range, the greatest difference in abundance is between eggs laid by birds of these two ages, and there is little statistical significance or correlation between compound abundance and hen age for eggs laid by birds of other ages, limiting its use as a marker of hen age. The abundance of this compound was not found to be significantly different between eggs laid by birds that were 26 and 30 weeks old when studying the whole range of hen age, which also indicates that this compound may not be suitable as a marker of hen age, as its statistical significance is not stable and reproducible between analytical runs.

The standard deviations for these compounds, in eggs laid by birds of different ages, are again quite large, resulting in wide error bars. This indicates that there was a relatively high amount of variation within the sample sets due to eggs being laid by different birds. Larger sample sizes would reduce these standard deviations.

5.8.1.5 Statistically significant compounds from smaller hen age ranges over a wider hen age range

In addition to the compounds that were putatively identified as 1-monopalmitin and SM(18:1/16:0), which were found in the top 58 compounds when comparing eggs laid by birds ranging from 21 to 44 weeks of age, there were three other previously putatively identified compounds that were found to be significantly different between eggs laid by the youngest and oldest birds when studied over smaller hen age ranges.

Table 5.28: *Table showing the p-values resulting from ANOVA and Tukey tests for the compound m/z 338.3420. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

ANOVA *p*-value	**Post-hoc test *p*-values**				
0.021	**21 Weeks**	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		0.525	0.722	0.984	0.473
26 Weeks			0.997	0.252	0.025
30 Weeks				0.411	0.052
38 Weeks					0.782
44 Weeks					

The compound that was putatively identified as erucamide was previously found to be highly significantly different between eggs laid by birds that were 26 and 30 weeks old, and between those laid by birds that were 30 and 38 weeks old. The ANOVA *p*-value given for this compound in Table 5.28 shows that the abundance of this compound was found to be statistically significant when comparing eggs laid by birds ranging from 21 to 44 weeks of age. However, the post-hoc test *p*-values for this compound in Table 5.28, show only one instance of statistical significance, that is over a wider hen age range than the previously observed significant differences. The post-hoc test *p*-values for the comparisons between eggs laid by birds that were 26 and 30 weeks old, and between those laid by birds that were 30 and 38 weeks old, are both very non-significant, which suggests that the statistical significance of this compound is not stable or replicable between analytical runs, making

it unsuitable as a marker of hen age.

Table 5.29: *Table showing the p-values resulting from ANOVA and Tukey tests for the compound m/z 734.5708. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

ANOVA *p*-value	Post-hoc test *p*-values				
0.034	**21 Weeks**	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		1	0.914	0.989	0.039
26 Weeks			0.957	0.998	0.055
30 Weeks				0.996	0.211
38 Weeks					0.106
44 Weeks					

The compound that was putatively identified as PC(16:0/16:0) was previously found to be highly significantly different between eggs laid by birds that were 21 and 26 weeks old. The ANOVA *p*-value given for this compound in Table 5.29 shows that the abundance of this compound was found to be statistically significant when comparing eggs laid by birds ranging from 21 to 44 weeks of age. However, the post-hoc test *p*-values for this compound in Table 5.29 show only one instance of statistical significance, that is again over a wider hen age range than the previously observed significant difference. The post-hoc test *p*-value for the comparison between eggs laid by birds that were 21 and 26 weeks old shows a complete lack of difference in compound abundance, which indicates that the significant difference of this compound is not reproducible between analytical runs, making it unsuitable as a marker of hen age.

The compound that was putatively identified as PE(16:0/18:2) was previously found to be highly statistically significant between eggs laid by birds that were 21 and 26 weeks old. However, no statistical significance in the abundance of this compound was found between eggs laid by birds of different ages, across the whole laying hen age range of 21-44 weeks, as can be seen in Table 5.30. Therefore, this compound is not suitable as a marker of hen age.

Table 5.30: *Table showing the p-values resulting from ANOVA and Tukey tests for the compound m/z 716.5243. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

ANOVA *p*-value	**Post-hoc test *p*-values**				
0.05	**21 Weeks**	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		0.050	0.816	0.998	0.987
26 Weeks			0.359	0.093	0.134
30 Weeks				0.936	0.976
38 Weeks					1
44 Weeks					

5.9 Hen Age 21-44 weeks (21, 26, 30, 38, 44 weeks): Albumen Organic Extracts

5.9.1 Results and Discussion

5.9.1.1 Quality Control Analysis

Table 5.31 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 5.31: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts from eggs laid by hens 21-44 weeks old*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	37844340	85094524	106142272	2175472	837007	9630662
QC2	37118341	82820154	105794644	1940579	905615	9602381
QC3	38715435	86464744	104565589	2095544	780009	10026394
QC4	25602799	60493292	79365449	1420942	735923	7026979
QC5	38292672	83606497	106944407	2207249	690366	9967612
QC6	35964283	78754275	99862886	1976375	877906	9309659
QC7	39301450	83836784	107139551	2215328	782314	9874763
SD	4764175	8993775	10028210	279431.8	76832.47	1053157
Mean	36119903	80152896	101402114	2004498	801306	9348350
CV%	13.19	11.22	9.89	13.94	9.59	11.27

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 5.32 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 5.32: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for albumen organic extracts from eggs laid by hens 21-44 weeks old*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	7.793	11.874	14.794	16.453	18.975	22.475
QC2	8.027	12.208	15.127	16.770	19.258	22.725
QC3	8.096	12.244	15.186	16.806	19.294	22.745
QC4	8.175	12.322	15.258	16.901	19.422	22.939
QC5	8.141	12.322	15.241	16.834	19.339	22.806
QC6	8.147	12.294	15.197	16.806	19.295	22.762
QC7	8.144	12.275	15.211	16.837	19.325	22.792
SD	0.133	0.158	0.160	0.147	0.141	0.140
Mean	8.075	12.220	15.145	16.772	19.273	22.749
CV%	1.65	1.29	1.06	0.87	0.73	0.61

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

5.9.1.2 Metabolite Profiling

Comparisons were made between albumen organic extracts of eggs laid by hens that were 21 weeks and 44 weeks old, by studying the TICs of the extracts following HPLC-MS analysis. Figure 5.29 shows overlaid TICs for the average chromatograms of extracts from eggs laid by hens of these two ages.

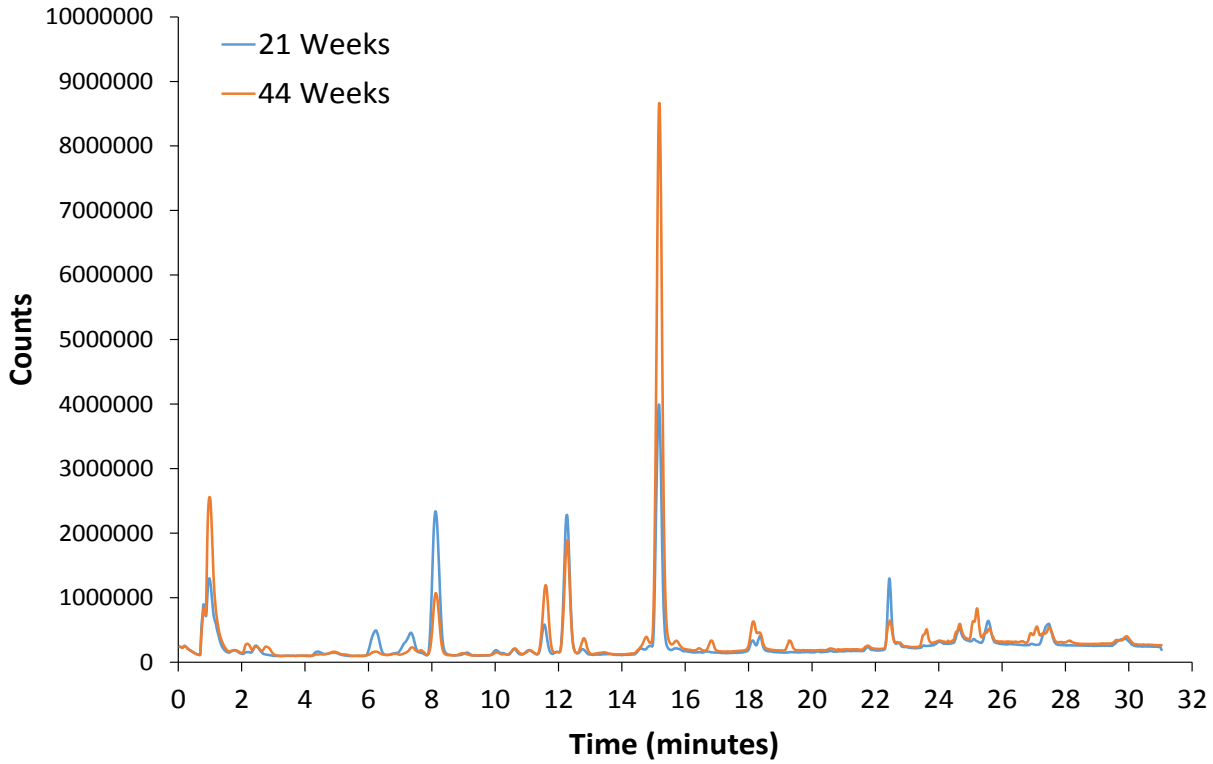


Figure 5.29: *Overlaid average TICs of albumen organic extracts of eggs that were laid by hens of 21 weeks and 44 weeks old*

The overlaid TICs in Figure 5.29 display a clear difference in metabolite profile between eggs laid by birds that were 21 weeks old, and those laid by birds that were 44 weeks old. There does not appear to be a particular trend regarding the difference in peak intensity between the two chromatograms; some peaks are of a higher intensity in the chromatogram of eggs laid by younger hens, and some are of a higher intensity in the chromatogram of eggs laid by the older birds.

5.9.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the eggs laid by hens of different ages.

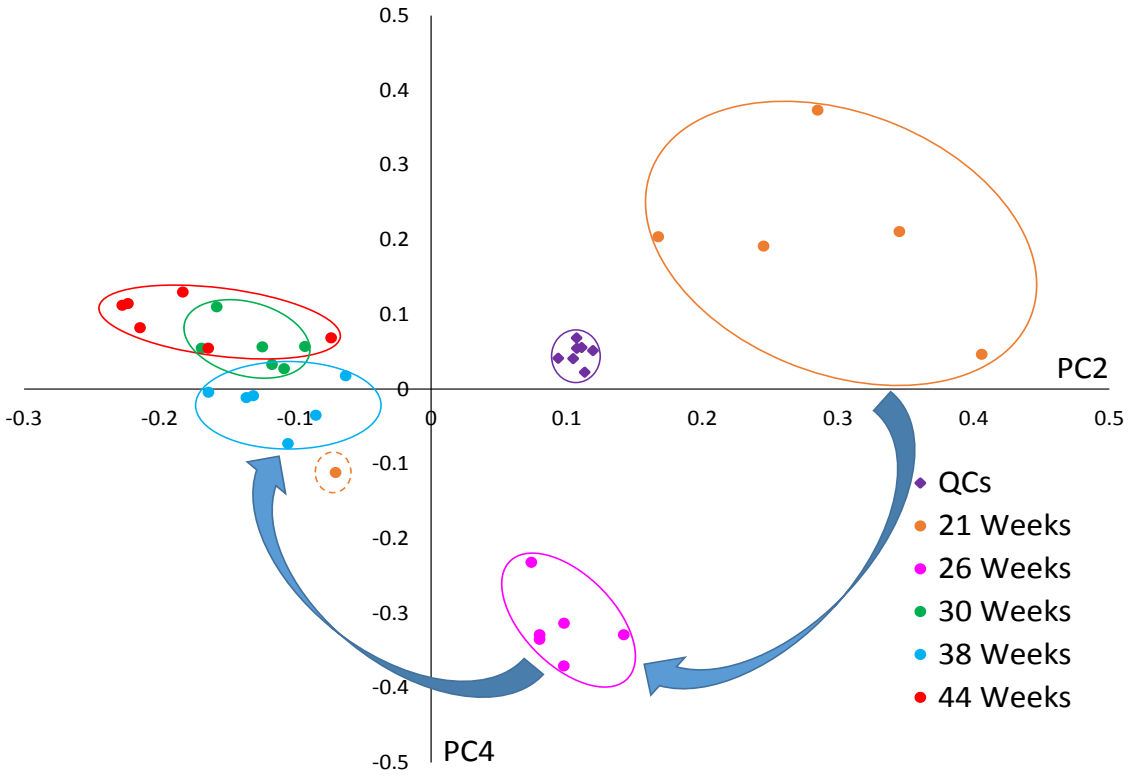


Figure 5.30: *PCA scores plot showing PC2 vs PC4 for albumen organic extracts of eggs laid by hens from 21 weeks to 44 weeks old, including all compounds with CV%<30%. PC2 explains 5.87% of the variance, and PC4 explains 0.43% of the variance. Circled: sample separate to rest of sample set.*

The scores plot in Figure 5.30 shows that eggs laid by birds that were 21 weeks old, and those laid by hens that were 26 weeks old, form their own groups separate to any other sets of eggs laid by birds of other ages. This indicates that a large change in the metabolite profile of the egg albumen occurs between a laying hen age of 21 weeks and 26 weeks, and between a laying hen age of 26 weeks and 30 weeks. There is one egg laid by a bird that was 21 weeks old (circled) that is grouping more closely with eggs laid by birds that were 38 weeks old. It is not known why this is, as nothing unusual was noted during the extraction or analysis, compared to the other eggs laid by birds of the

same age. However, eggs laid by birds ranging from 30 weeks to 44 weeks old all group together on the opposite side of the plot, showing very little separation. This suggests that when the birds are later on in their laying cycle, the difference in metabolite profile between eggs laid by birds of different ages decreases.

The separation between eggs laid by birds that were 21 and 26 weeks old, and those laid by the older birds is across PC2, which explains 5.87% of the variance. This indicates that there is a relatively large amount of variation between these sets of eggs laid by younger and older birds, compared to the previous variation observed in PCA scores plots when comparing the metabolite profiles of the yolks of eggs laid by birds of different ages. The very slight separation between eggs laid by birds that were 38 and 44 weeks old is across PC4, which explains only 0.43% of the variance, showing just how subtle these differences are. Both PC2 and PC4 are responsible for the separation between eggs laid by birds that were 26 weeks old, and eggs laid by birds of other ages, with most of the separation across PC4. Although PC4 does completely separate eggs laid by birds that were 26 weeks old from the other sets of eggs laid by birds of other ages, as it only describes such a small amount of variation, it again shows the subtlety of the differences between eggs laid by birds of different ages.

5.9.1.4 Univariate Statistics

There were 35 compounds remaining following the removal of any duplicates, adducts and isotopes from the top 100 compounds, and all 35 of these were found to be statistically significant. These compounds, and their CV%s and p-values, can be seen in Table B.6 in Appendix B; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

All 35 compounds were still found to be statistically significant when confirmed using the raw data, and of these 35, three were putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The observed m/z for all compounds was due to the $[M+H]^+$ adduct. The compounds m/z 310.3104, 331.2840, and 338.3418 were putatively identified as oleoyl ethyl amide, 1-monopalmitin, and docosenamide (erucamide) respectively. The compounds m/z 331.2840 and 338.3418 were previously putatively identified in this research, and their comparisons between mass spectra provided by METLIN and mass spectra resulting from the chemical analysis in this study can be seen in Figures B.1 and B.7 respectively, in Appendix B. The mass spectra comparison for the remaining compound, m/z 310.3104, can be seen in Appendix B, Figure B.10.

Tables 5.33-5.35, and Figures 5.31-5.33, show the putative identifications of all of these compounds, the p-values resulting from ANOVA and post-hoc Tukey tests, as well as the trends of changing abundance of the compounds in eggs laid by birds of increasing age from 21 to 44 weeks old.

Table 5.33: *Table showing the putative identification of compound m/z 310.3104 and the p-values resulting from Welch tests and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**Welch *p*-value**
310.3104	$C_{20}H_{39}NO$	97.31	Oleoyl ethyl amide	<0.001

Post-hoc test *p*-values					
	21 Weeks	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		0.020	0.001	<0.001	<0.001
26 Weeks			0.629	0.042	0.008
30 Weeks				0.506	0.175
38 Weeks					0.954
44 Weeks					

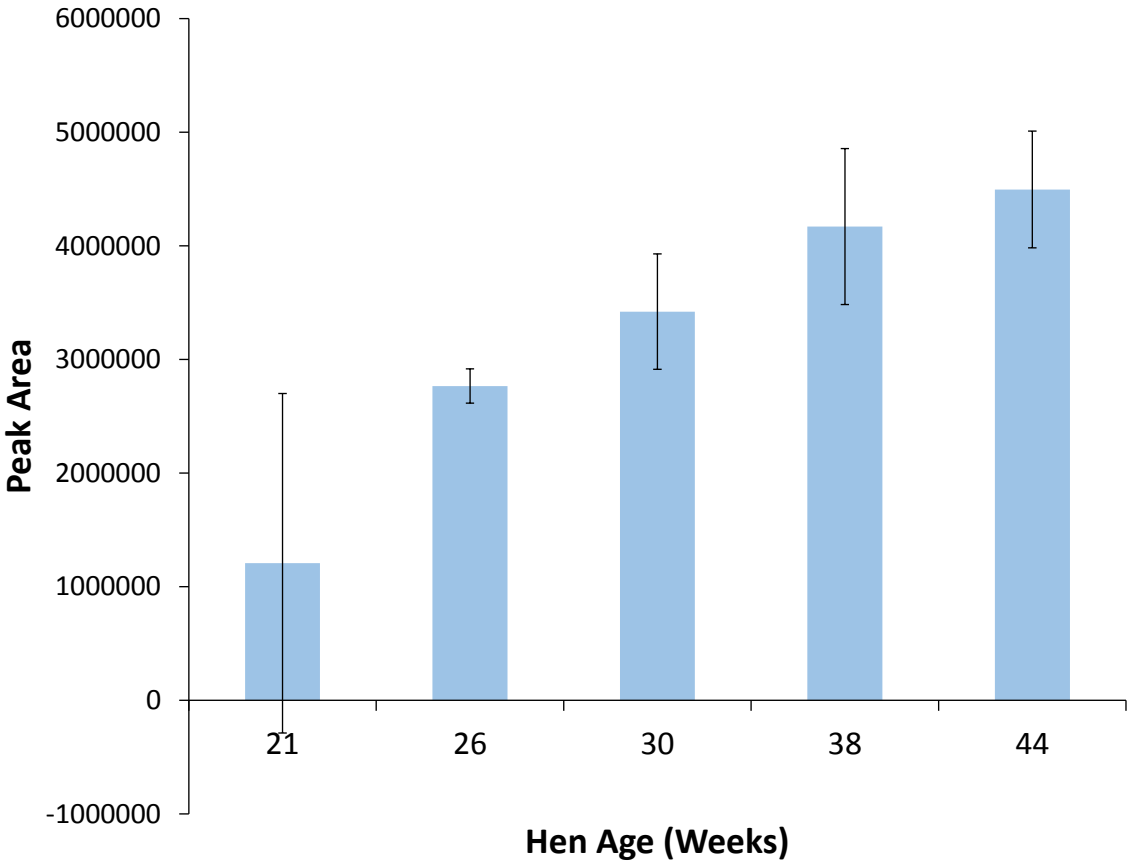


Figure 5.31: *Bar chart showing the trend in changing abundance of compound m/z 310.3104 in albumen across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.34: *Table showing the putative identification of compound m/z 331.2840 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
331.2840	$C_{19}H_{38}O_4$	99.39	1-monopalmitin	<0.001

Post-hoc test *p*-values

	21 Weeks	26 Weeks	30 Weeks	38 Weeks	44 Weeks
21 Weeks		0.036	0.001	<0.001	<0.001
26 Weeks			<0.001	<0.001	<0.001
30 Weeks				0.403	0.712
38 Weeks					0.039
44 Weeks					

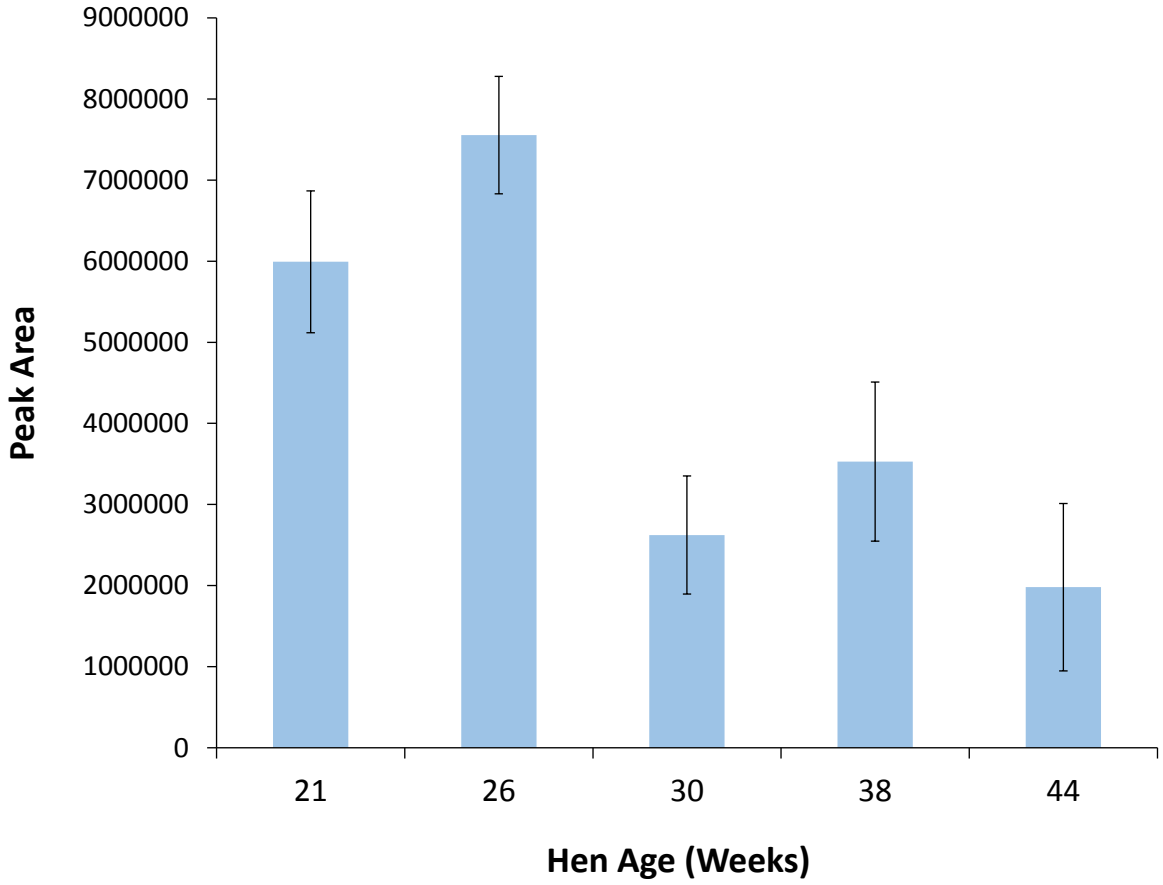


Figure 5.32: *Bar chart showing the trend in changing abundance of compound m/z 331.2840 in albumen across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

Table 5.35: *Table showing the putative identification of compound m/z 338.3418 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
338.3418	$C_{22}H_{43}NO$	99.85	Erucamide	<0.001

Post-hoc test *p*-values					
	21 Weeks	**26 Weeks**	**30 Weeks**	**38 Weeks**	**44 Weeks**
21 Weeks		0.103	0.005	<0.001	<0.001
26 Weeks			0.658	0.078	0.014
30 Weeks				0.652	0.235
38 Weeks					0.937
44 Weeks					

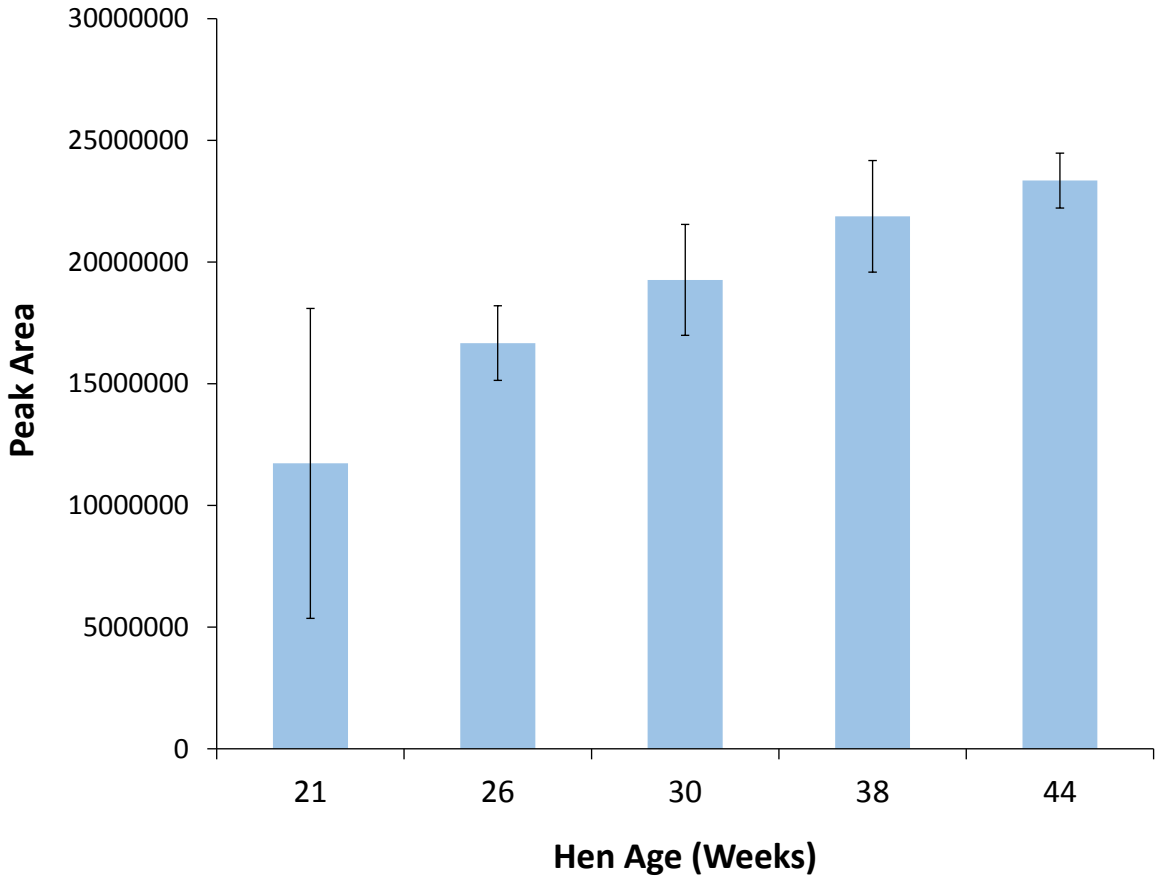


Figure 5.33: *Bar chart showing the trend in changing abundance of compound m/z 338.3418 in albumen across eggs laid by hens from 21-44 weeks old. n=6. Measured as mean values of peak area for each hen age, with error bars ±1 standard deviation.*

The probability scores for the predicted molecular formulae are all above 95, therefore there is a high likelihood that the predicted formulae correspond to the compounds of interest. The p-values from ANOVA/Welch tests show that the abundances of all of these compounds are highly statistically significant when comparing them between eggs laid by birds of different ages, ranging from 21 weeks old to 44 weeks old, and the post-hoc test p-values confirm this.

The post-hoc test p-values in Table 5.33 show several points of high statistical significance for the abundance of the compound that was putatively identified as oleoyl ethyl amide, in eggs laid by birds of different ages, ranging from 21 to 44 weeks old. This compound is an amide derivative of the fatty acid oleic acid, with an ethyl group attached to the nitrogen of the amide group. As well as being highly statistically significant, this compound shows a steadily increasing abundance in eggs laid by birds of increasing age from 21 to 44 weeks, as can be seen in Figure 5.31, which makes it potentially suitable as a marker of hen age over this age range. The standard deviation error bar for eggs laid by birds that were 21 weeks old is very wide, indicating a large amount of variance in the abundance of this compound in eggs laid by birds of this age. This could be because the birds were quite young and very early in their laying cycle, and therefore the metabolic state of these hens was not homogeneous between the birds. This suggests that it may be difficult to predict laying hen age for younger birds, based on compound abundance in eggs, as there may be too much variation.

The compound that was putatively identified as 1-monopalmitin was also found in the top statistically significant compounds when comparing the yolks of eggs laid by birds of different ages within the same age range. However, when found in the yolk previously, as well as in the albumen presently, it has been determined to be unsuitable as a marker of hen age for birds within this age range. Although several of the post-hoc test p-values in Table 5.34 are highly significant, Figure 5.32 shows that there was no correlation between the abundance of this compound in egg albumen and laying hen age, indicating that this

compound would not be suitable for predicting hen age.

The compound that was putatively identified as erucamide can be seen in Figure 5.33 to gradually increase in abundance in egg albumen with increasing laying hen age, which suggests that this compound may be suitable as a marker of hen age. Table 5.35 shows several points of high statistical significance when comparing eggs laid by birds of different ages, which confirms the suitability of this compound as a hen age marker. However, most of the statistical significance lies between eggs laid by birds that were 21 weeks old, and those laid by birds that were 30 weeks and older, suggesting that the increase in compound abundance is mostly over the first few weeks of increasing hen age, before becoming more gradual between eggs laid by older birds. This could limit the use of this compound as a marker of hen age over this age range.

The pattern of the changing abundance for this compound is very similar to that of the compound that was putatively identified as oleoyl ethyl amide. This is likely to be because these compounds are very similar to each other; they are both amide derivatives of fatty acids, which only differ by a chain length of four carbon atoms. Therefore, these compounds are likely to experience similar metabolic processes, explaining why their changing abundances in the albumen of eggs laid by birds of increasing age are so similar. However, the compound that was putatively identified as erucamide does not show as much potential as a marker of hen age as the compound that was putatively identified as oleoyl ethyl amide, as it does not experience quite as significant a difference in abundance in the albumen of eggs laid by birds of increasing age. Similarly to what was observed for the compound that was putatively identified as oleoyl ethyl amide, there is quite a wide standard deviation error bar for eggs laid by birds that were 21 weeks old, indicating that there was a large amount of variance in the abundance of this compound in eggs laid by birds of this age. This again suggests that predicting laying hen age based on the abundance of this compound may not be possible for younger birds, as there may be too much variation to make an accurate prediction.

5.10 Summary & Conclusions

5.10.1 Summary

The overlaid chromatograms in each of the yolk organic extract studies within this work showed a slight difference in the metabolite profiles of egg yolk between eggs laid by birds of different ages. These chromatograms showed a higher peak intensity for later eluting, non-polar compounds that may be phospholipids, in the TICs of yolk extracts from eggs laid by older birds, indicating that compounds within this class may be present in a higher abundance in the yolks of eggs laid by older birds, compared to in the yolks of eggs laid by younger birds. The PCA scores plots in these studies also showed some separation between sets of eggs laid by birds of different ages, again indicating that there is some difference in the metabolite profile of yolk between eggs laid by birds of different ages. These observations were noted for the studies over small hen age ranges, as well as over a wider hen age range from 21 to 44 weeks of age.

Several statistically significant compounds were putatively identified in the yolk organic extract studies, but were determined to not be suitable as markers of hen age. Some compounds were found to have potential as markers of hen age, but only between certain ages of birds within the smaller age ranges, as there was only a correlation between abundance and hen age up to a certain age of hen within the age range. Therefore, these compounds were found to not be suitable as markers of hen age over wider age ranges. However, other compounds were determined to be completely unsuitable as markers of hen age, as they either did not show any correlation between abundance and hen age, or their statistical significance observed through post-hoc test p-values was not very strong. Some compounds, which were putatively identified and found in the top statistically significant compounds for more than one laying hen age range, were determined to have some potential as markers of hen age within one age range and not others, and therefore were found to not be suitable as markers of hen age over a longer age range.

One compound, which was putatively identified as 1-monopalmitin, appeared in the top statistically significant compounds for several different hen age ranges that were studied, including the wider hen age range from 21-44 weeks old. This compound was determined to have potential as a marker of hen age within one of the smaller hen age ranges, but not others. When observed over a wider hen age range, from 21 to 44 weeks of age, this compound showed some potential as a marker of hen age, due to the positive correlation of its abundance with increasing hen age, and statistical significance between eggs laid by birds that were 30 to 44 weeks old. However, the results observed over this wider hen age range did not corroborate what was observed in the smaller hen age range studies; therefore this compound was determined to not be suitable as a marker of hen age as its relative abundance, and therefore statistical significance, was not reproducible between analytical runs. The remaining putatively identified compounds that appeared in the top statistically significant compounds over the wider hen age range did not show any potential as markers of hen age for laying hens between 21 and 44 weeks old.

Other putatively identified compounds that were found to be significantly different between eggs laid by the youngest and oldest birds in the individual hen age ranges that were studied, were not found to be significantly different between eggs laid by birds of these ages when observed in the data for eggs laid by birds ranging from 21 to 44 weeks old. This indicates that these compounds are not suitable as markers of hen age as they are unreliable and do not show repeatability between analytical runs.

The results of the albumen organic extract study, over a laying hen age range of 21 to 44 weeks, were similar to what was observed from analysing the yolk. The overlaid chromatograms showed a difference in metabolite profile between eggs laid by younger and older birds, but there appeared to be no trend in which types of compound were of a higher or lower abundance in eggs laid by younger or older birds. The PCA scores plot showed some separation between eggs laid by birds of different ages, particularly between the eggs laid by the younger birds.

One statistically significant compound was again putatively identified as 1-monopalmitin, but this compound did not show any potential as a marker of hen age when found in albumen. Another compound was putatively identified as erucamide, which was determined to not be suitable as a marker of hen age when found in the yolk. However, when found in the albumen, this compound showed some limited potential as a marker of hen age for birds ranging from 21 to 44 weeks old. The compound that was putatively identified as oleoyl ethyl amide was only found in the top statistically significant compounds when studying the albumen, and was found to show potential as a marker of hen age due to its gradually increasing abundance and statistical significance over an increasing hen age from 21-44 weeks.

5.10.2 Conclusions

It can be concluded from this research that the age of the laying hen does have an effect on the resulting metabolite profiles of the yolk and albumen of the egg. The overlaid TICs and PCA scores plots through the various hen age ranges that were studied, have shown that there are slight differences in metabolite profile between eggs laid by birds of different ages. Although the influence that laying hen age has on the metabolite profiles is only subtle, the effect of other variables that may be investigated in future metabonomic studies may also be very slight; this could result in the impact of hen age on the metabolite profile being relatively large. Therefore, it is important to not only be aware of the effects that different hen ages may have on the results of a metabonomic study into eggs, but also to try and keep the age of the birds as similar as possible when designing and carrying out the study.

This research has shown that a non-targeted metabonomic study is capable of observing differences in both the yolk and albumen of eggs laid by birds of different ages. It has shown that although some statistically significant compounds show a correlation between abundance and hen age, they are not always suitable as markers of hen age; the correla-

tion may only exist for eggs laid by birds between certain ages and may change erratically throughout the laying cycle of a hen, and the compounds may not show repeatability in relative abundance and statistical significance between analytical runs. Therefore these compounds are not suitable as markers of hen age to help detect cases of fraud, as they would not be able to accurately, and reliably, predict the age of the laying bird. However, this research found that one statistically significant compound with a correlation between its abundance in albumen and laying hen age, which was putatively identified as oleoyl ethyl amide, does show potential as a marker of hen age.

Egg albumen appeared to be more greatly affected than the yolk by the increasing age of the laying hen, suggesting that analysis of the albumen may produce better results when detecting cases of potential fraud. This also indicates that it is even more important to be aware of the effect that hen age can have on the metabolite profile of the albumen, than the yolk, when carrying out a metabonomic study into eggs.

Further work would involve observing the abundance in albumen of the compound that was putatively identified as oleoyl ethyl amide, over smaller ranges of hen age, to determine whether it would be useful when predicting a more accurate hen age, and whether the statistical significance of this compound is reproducible. Observing the changing abundance of this compound over an even wider range of hen age would also be useful in determining at which hen age the correlation between abundance and laying hen age stops, and it is no longer possible to use this compound to discriminate between eggs laid by birds of different ages. If this compound was found to be reproducibly capable of predicting hen age, future studies could then determine whether these results are replicable with different breeds of bird, and birds kept in different housing systems. If this compound was found to be successful in predicting an accurate laying hen age, and the results were replicable between different breeds of bird, and birds from different housing systems, then it could be used to detect cases of fraud regarding false producer codes on eggs.

5.11 References

[1] T. Sangster, H. Major, R. Plumb, J. A. Wilson, and I. D. Wilson, "A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis," *The Analyst*, vol. 131, pp. 1075–1078, 2006.

[2] H. Issaq, Q. Van, T. Waybright, G. Muschik, and T. Veenstra, "Analytical and statistical approaches to metabolomics research," *Journal of Separation Science*, vol. 32, pp. 2183–2199, 2009.

[3] H. Gika, G. Theodoridis, and I. Wilson, "Liquid chromatography and ultra-performance liquid chromatography-mass spectrometry fingerprinting of human urine. Sample stability under different handling and storage conditions for metabonomics studies," *Journal of Chromatography A*, vol. 1189, pp. 314–322, 2008.

[4] S. Ren, A. Hinzman, E. Kang, R. Szczesniak, and L. Lu, "Computational and statistical analysis of metabolomics data," *Metabolomics*, vol. 11, pp. 4192–1513, 2015.

[5] *Council Directive 1999/74/EC of 19 July 1999, Laying down minimum standards for the protection of laying hens*, Official Journal of the European Communities, L203/53-57, 1999.

[6] *Council Regulation (EC) No 1234/2007 of 22 October 2007, establishing a common organisation of agricultural markets and on specific provisions for certain agricultural products (Single CMO Regulation*, Official Journal of the European Union, L299/1-149, 2007.

[7] F. G. Silversides and T. A. Scott, "Effect of storage and layer age on quality of eggs from two lines of hens," *Poultry Science*, vol. 80, pp. 1240–1245, 2001.

[8] B. Y. Dikmen, A. Ipek, Ü. Şahan, A. Sözcü, and S. C. Baycan, "Impact of different housing systems and age of layers on egg quality characteristics," *Turkish Journal of Veterinary Animal Science*, vol. 41, pp. 77–84, 2017.

[9] G. B. Anderson, W. Bolton, R. M. Jones, and M. H. Draper, "Effect of age of the laying hen on the composition of the egg," *British Food Journal*, vol. 19, pp. 741–745, 1978.

[10] Z. Jiang and J. S. Sim, "Research note: Egg cholesterol values in relation to the age of laying hens and to egg and yolk weights," *Poultry Science*, vol. 70, pp. 1838–1841, 1991.

[11] A. Travel, Y. Nys, and M. Bain, "Improving the safety and quality of eggs and egg products," in. Woodhead Publishing Limited, 2011, vol. 1: Egg Chemistry, Production and Consumption, ch. 13: Effect of hen age, moult, laying environment and egg storage on egg quality, pp. 300–329.

[12] M. Rossi and C. Pompei, "Changes in some egg components and analytical values due to hen age," *Poultry Science*, vol. 74, pp. 152–160, 1995.

[13] D. L. Fletcher, W. M. Britton, A. P. Rahn, and S. I. Savage, "The influence of layer flock age on egg component yields and solids content," *Poultry Science*, vol. 60, pp. 983–987, 1981.

[14] H. Akyurek and A. A. Okur, "Effect of storage time, temperature and hen age on egg quality in free- range layer hens," *Journal of Animal and Veterinary Advances*, vol. 8, pp. 1953–1957, 2009.

[15] A. B. Rodríguez-navarro, N. Domínguez-gasca, A. Muñoz, and M. Ortega-Huertas, "Change in the chicken eggshell cuticle with hen age and egg freshness," *Poultry Science*, vol. 92, pp. 3026–3035, 2013.

[16] R. Karoui, B. Kemps, F. Bamelis, B. deKetelaere, E. Decuypere, and J. deBaerdemaeker, "Methods to evaluate egg freshness in research and industry: A review," *European Food Research and Technology*, vol. 222, pp. 727–732, 2006.

[17] P. C. deMenezes, E. R. deLima, J. P. deMedeiros, W. M. K. deOliveira, and J. Evencio-Neto, "Egg quality of laying hens in different conditions of storage, ages and housing densities," *Revista Brasileira de Zootecnia*, vol. 41, pp. 2064–2069, 2012.

[18] D. U. Ahn, S. M. Kim, and H. Shu, "Effect of egg size and strain and age of hens on the solids content of chicken eggs," *Poultry Science*, vol. 76, pp. 914–919, 1997.

[19] R. Oloyo, "Effect of age on total lipid and cholesterol of hen eggs," *Indian Journal of Animal Science*, vol. 73, pp. 93–100, 2003.

[20] T. M. Shafey, "The relationship between age and egg production, egg components and lipoprotein, lipids and fatty acids of the plasma and eggs of laying hens," *Journal of Applied Animal Research*, vol. 10, pp. 155–162, 1996.

[21] H. Nielsen, "Hen age and fatty acid composition of egg yolk lipid," *British Poultry Science*, vol. 39, pp. 53–56, 1998.

[22] L. Zemková, J. Simeonovová, M. Lichovníková, and K. Somerlíková, "The effects of housing systems and age of hens on the weight and cholesterol concentration of the egg," *Czech Journal of Animal Science*, vol. 4, pp. 110–115, 2007.

[23] (2019). KEGG PATHWAY: Glycerophospholipid metabolism, [Online]. Available: `https://www.kegg.jp/kegg-bin/show_pathway?map00564` (visited on 07/2019).

6. The effect of hen diet on the metabolite profiles of egg yolk and albumen

6.1 Introduction

Diet is a variable that affects the metabolite profile of a subject [1][2][3], and if not controlled, can greatly affect the results of a metabonomic study [4][5]. Therefore, it is important to be aware of the effects that the diet of the laying hen may have on the metabolite profiles of the yolk and albumen of the eggs.

Laying hens that are kept for organic egg production must be provided with an organic feed as their diet [6], whilst non-organic birds are given a diet of conventional feed. Organic feed contains different ingredients to those that are present in conventional feed, and there are regulations regarding which ingredients and additives may be allowed in organic feed [7]. However, organic feed is more expensive than conventional feed [8], which would potentially tempt organic egg producers to provide their organic poultry with conventional, rather than organic, feed in order to cut costs. The organic protein supply is not currently great enough to meet the nutritional requirements of poultry, therefore the organic feed of laying hens in organic egg production in the European Union is allowed to contain just 5% non-organic protein feed [6]. This regulation is currently

in place until December 2020, at which point it will be reviewed. If this regulation is changed, and it is deemed unnecessary for organic feed to contain non-organic protein feed, then it could encourage cases of fraud, as producers may not wish to switch to a full, 100% organic feed, and may stick to the current feed composition.

It would therefore be useful to have a scientific method that is capable of discriminating between different laying hen diets, based on the analysis of the egg, in order to detect cases of fraud. However, if the testing of eggs to detect the incorrect feeding of organic poultry was implemented, egg producers may pre-empt an investigation, and correct the diet of the laying hens. Therefore, it is also important to determine how soon after changing the diet of a bird, an effect can be observed in the egg. The quicker the egg is affected by a diet change, the sooner it fully reflects the effect of the new diet on the egg, and the less time there is to detect the fraud.

Although it is not possible to change the total lipid content of an egg yolk through dietary modifications, the fatty acid profile of the yolk is strongly influenced by diet [9]. There has been much research into how the diet of a laying hen affects the resulting fatty acid (FA) profile in the yolks of eggs laid by the bird, as there are various human health implications regarding the lipid content of the egg yolk. It has been determined that the dietary supply of fatty acids is the greatest influence on the FA profile of egg yolk, particularly for omega-3 FAs [10]. One study found that the omega-6 to omega-3 FA ratio can be decreased with the addition of different oils to the diet [11], which is similar to what was observed by other researchers, who found that the omega-3 FA content of egg yolk can be increased with the addition of fish oil and flaxseed to the diet [12]. Other researchers have also found that the addition of fish oil to diets decreases the ratio of omega-6 to omega-3 FAs by increasing the content of omega-3 FAs [13], whilst others have determined that the addition of linseed oil to laying hen diet results in good omega-6 to omega-3 ratios in the yolk [14].

There has also been some research into how the laying hen diet affects the cholesterol

levels of egg yolk. However, it is difficult to modify the level of cholesterol in the yolk, as the cholesterol content of the egg is dependent on the hepatic synthesis of lipoproteins, rather than the blood cholesterol level of the hen [9]. The proportion of the lipid and protein components of very low density lipoproteins (VLDLs), which are the precursors to egg yolk components, is highly stable. Therefore, it is difficult to modify the cholesterol level in VLDLs without negatively impacting the physiology of the hen, and thus it is difficult to modify the cholesterol level of the yolk [9]. Some studies have reported successful modifications of cholesterol levels of egg yolk by altering the laying hen diet, whilst others have been less successful [12]. One study found that yolk cholesterol levels were not affected by the fatty acid composition of the laying hen diet [15], and another found that flaxseed and antioxidant supplementation of laying hen diet also had no effect on the cholesterol content of the egg yolk [16]. One group of researchers found that diets containing higher percentages of chia seed, compared to a control diet with no chia seed, resulted in a lower cholesterol content of egg yolks [17], whilst others found that laying hen diets supplemented with increasing amounts of garlic resulted in decreasing levels of cholesterol in egg yolk [18].

The effects of dietary zinc, iron and copper in bird feed on the distribution of these elements in eggs have been studied [19]. It was found that supplementing the diet with iron increased the concentration of iron in both the yolk and albumen of eggs, and that when the diet was supplemented with all three elements, the concentration of iron increased even further in both egg components. However, little effect was observed on the enrichment of eggs with copper and zinc with the supplemented diets.

Although there is much research regarding the effect of laying hen diet on egg yolk lipid content, there does not appear to be much research regarding how laying hen diet can affect other compounds in the egg yolk, or how the diet impacts the albumen of the eggs. The use of a metabonomic approach, using HPLC-MS, would enable the observation of how a wide range of compounds and compound classes, in egg yolk and albumen, are

affected by different laying hen diets. This would not only help with the detection of fraud regarding the diets of organic laying hens, but would help to develop an understanding of how the laying hen diet can affect the results of other metabonomic studies on eggs.

6.2 Aims and Objectives

When carrying out metabonomic studies, it is important to understand how the diet of the laying hen affects the resulting metabolite profiles of the yolk and albumen of the egg. This work aimed to show how the metabolite profiles of both egg yolk and albumen are influenced by the diet of the laying hen.

In order to detect cases of fraud where hens in organic egg production systems are fed a conventional diet, it is important to be able to discriminate between eggs laid by birds that are fed different diets, and to determine how soon after switching the diet of the laying hen the effects can be seen in the metabolite profile of the yolk and albumen of the eggs.

This work aimed to perform a preliminary study involving the use of a non-targeted metabonomic workflow to uncover the differences in metabolite profile between eggs laid by four sets of birds that were fed four different diets, both two and five weeks after the diets were changed. This work also aimed to observe the rates at which different laying hen diets have an effect on the metabolite profiles of the yolk and albumen of eggs.

6.3 Experimental

Eggs were collected on the day of lay from the National Institute of Poultry Husbandry (Harper Adams University, Newport, U.K.). Laying hens were of the Hy-line brown breed, 21 weeks old at the start of the study, and kept in enriched cages with 80 birds per cage. Initially, all birds were fed the same diet, Diet A.

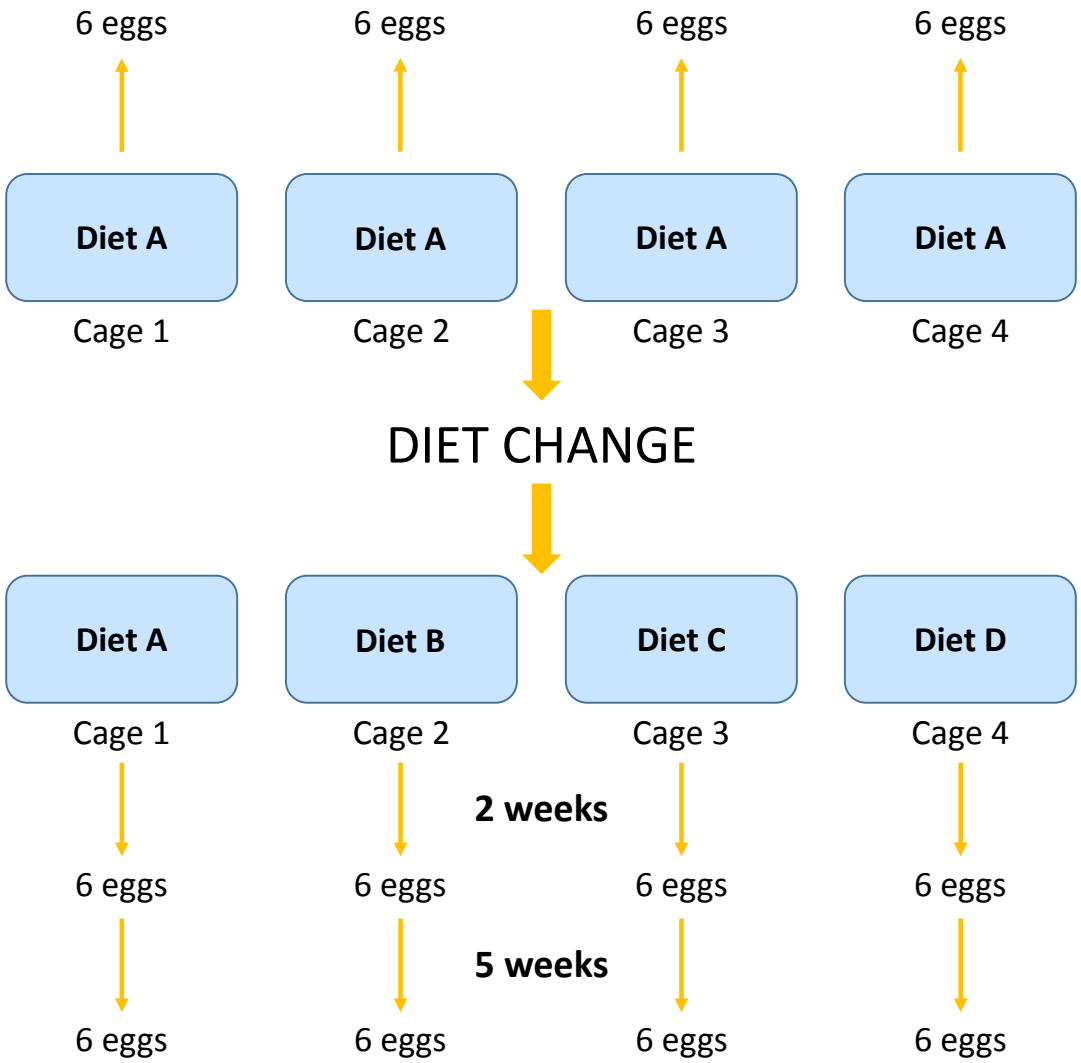


Figure 6.1: *Diagram showing the experimental design for the change in laying hen diet study*

Six eggs were collected from four different cages one week prior to the diet change, in order to act as a control. The diets of three of these cages of birds were then changed,

resulting in four cages of birds being fed four different diets; Diets A, B, C, and D. The compositions of all four diets were different, but unknown. The next set of eggs were collected two weeks after this diet change, and six eggs were collected per cage/diet. Another set of six eggs per cage/diet was collected three weeks later, five weeks following the initial diet change. All eggs underwent metabolite extraction on the day of lay, and the resulting extracts were stored at -80 °C until chemical analysis took place. Figure 6.1 shows the experimental design for this study.

Analysis of the yolks of eggs laid by birds that were fed the four different diets took place both two weeks, and five weeks after the diet change, in order to observe how any resulting changes in metabolite profile develop over an increasing time post diet change. The metabolite profiles of the yolks of eggs laid by birds which were fed diets B, C, and D were compared to the profile of the yolks of eggs laid by birds which were fed diet A, in order to observe which diets had the greatest and quickest effects on the yolk metabolite profile following the change from diet A.

Analysis of the albumen of eggs laid by birds that were fed the four different diets took place five weeks after the diet change, to determine whether a difference in metabolite profile can be observed in the albumen of eggs laid by birds which were fed different diets. Again, the metabolite profiles of the yolks of eggs laid by birds which were fed diets B, C, and D were compared to the profile of the yolks of eggs laid by birds which were fed diet A, in order to observe which diets had the greatest impact on albumen metabolite profile following the change from diet A.

Organic metabolite extraction of the yolk and albumen, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. Data analysis for all three experiments was then carried out as described in Chapter 3, Section 3.9, with the addition of a second PCA, using only the top statistically significant compounds following ANOVA/Welch tests. The p-values for the top statistically significant compounds were compared to p-values

resulting from ANOVA/Welch tests for these compounds between eggs laid by the birds from the different cages one week prior to the diet change. This was to ensure that any statistically significant compounds were significant due to the effects of the different diets, and not any environmental differences that may exist between the sets of birds.

No identification of compounds took place, as the compounds that are significantly affected by a change of diet will be different depending on the composition of the initial diet, and the compositions of the new diets. In terms of understanding the effect of hen diet on the metabolite profiles of yolk and albumen in metabonomic studies, it is more important to determine *if* and *how* the metabolite profiles of the yolk and albumen are affected, than *which* compounds are affected.

6.4 2 weeks post diet change: Yolk Organic Extracts

Organic extracts of the yolk of eggs laid by four sets of hens which were fed four different diets were compared, two weeks after the diets of three sets of hens were changed, to see whether a difference in metabolite profile could be observed between the yolks of eggs laid by hens that were fed the four different diets.

6.4.1 Results and Discussion

6.4.1.1 Quality Control Analysis

Table 6.1 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 6.1: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens fed four different diets, two weeks post diet change*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	1835145	131801593	56711557	26597296	465408900	82038292
QC2	1567475	130955535	56300801	26785943	452967309	80809174
QC3	1631501	129758396	55854682	25692277	452194634	77984124
QC4	1445203	129795696	55345965	26005762	450897595	77977511
QC5	1431086	130320776	53717693	25941661	447066406	74956238
QC6	1755498	132596886	57164116	25331402	447435075	69195314
QC7	1520903	133516651	62406170	26066318	452557830	70798831
SD	152888	1446035	2718020	499790	6117582	4855229
Mean	1598116	131249362	56785855	26060094	452646821	76251355
CV%	9.57	1.10	4.79	1.92	1.35	6.37

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 6.2 shows the retention times (RTs) for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 6.2: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens fed four different diets, two weeks post diet change*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.524	9.139	32.511	56.762	71.420	77.744
QC2	4.557	9.188	32.527	56.762	71.536	77.910
QC3	4.558	9.206	32.511	56.813	71.570	77.994
QC4	4.452	9.233	32.555	56.856	71.597	78.099
QC5	4.537	9.218	32.557	56.858	71.648	78.122
QC6	4.593	9.241	32.646	57.014	71.854	78.494
QC7	4.477	9.274	32.696	57.047	71.987	78.709
SD	0.049	0.043	0.072	0.115	0.196	0.337
Mean	4.528	9.214	32.572	56.873	71.659	78.153
CV%	1.08	0.47	0.22	0.20	0.27	0.43

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

6.4.1.2 Metabolite Profiling

Comparisons were made between the yolks of eggs laid by hens that were fed the four different diets, by studying the Total Ion Chromatograms (TICs) of the extracts following HPLC-MS analysis. Figure 6.2 shows overlaid TICs for the average chromatograms of organic extracts of yolk from eggs laid by hens that were fed these four diets.

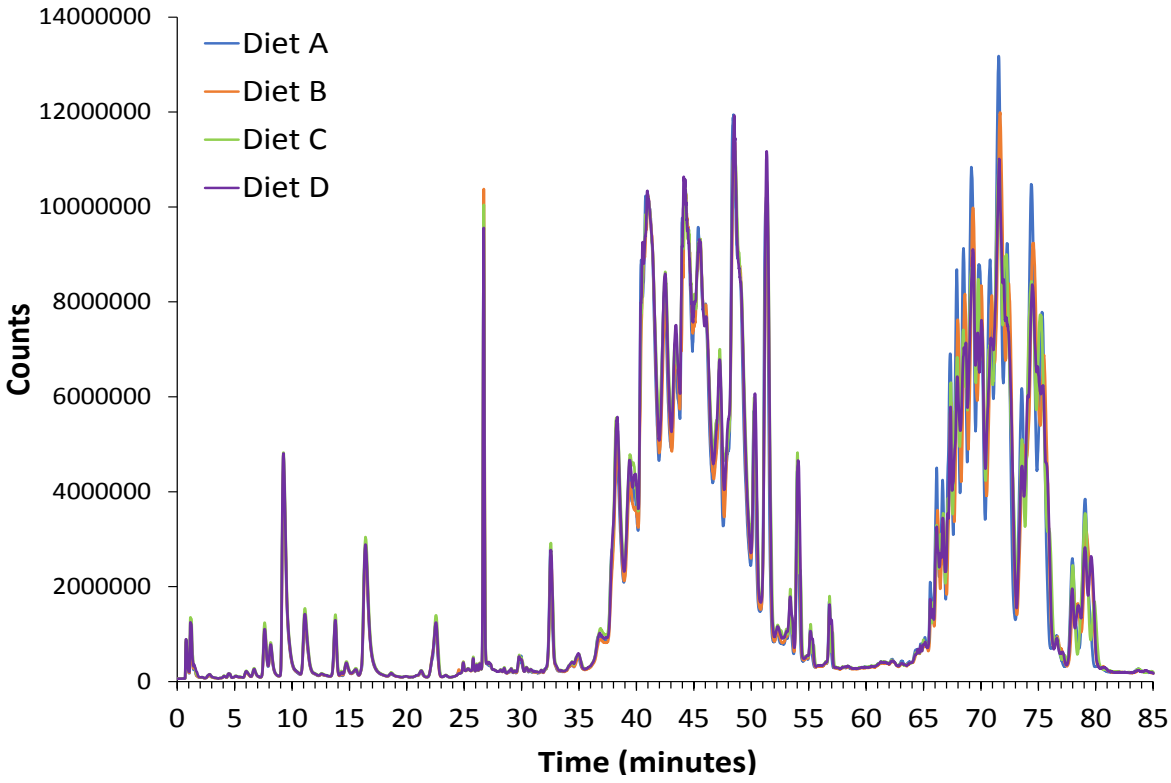


Figure 6.2: *Overlaid average TICs of organic extracts of yolks of eggs laid by hens that were fed four different diets, two weeks post diet change*

As can be seen from the overlaid TICs in Figure 6.2, there is a subtle difference in metabolite profile between the yolks of eggs laid by birds that were fed the four different diets. There is a slightly higher peak intensity for most peaks in the first 25 minutes of analysis, in the TIC of yolk from eggs laid by birds which were fed diet C, indicating either that this diet may be rich in less non-polar compounds, or that it results in an increased abundance of these compounds in the yolk. Between 35 and 65 minutes of analysis there is very little difference in peak intensity between the TICs of the yolks of eggs laid by birds which were fed the four different diets. However, the peaks that do

show some difference in intensity, again have a higher intensity in the TIC of the yolks of eggs laid by birds which were fed diet C. It is difficult to compare the peak intensities after 65 minutes of analysis, as there was some retention time drift throughout the last 20 minutes of each analytical run, which has affected the average TICs that were produced, resulting in a poor peak shape. However, this drift was still within accepted limits, as seen in the QC analysis in Section 6.4.1.1.

6.4.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the yolks of eggs laid by hens which were fed the four different diets.

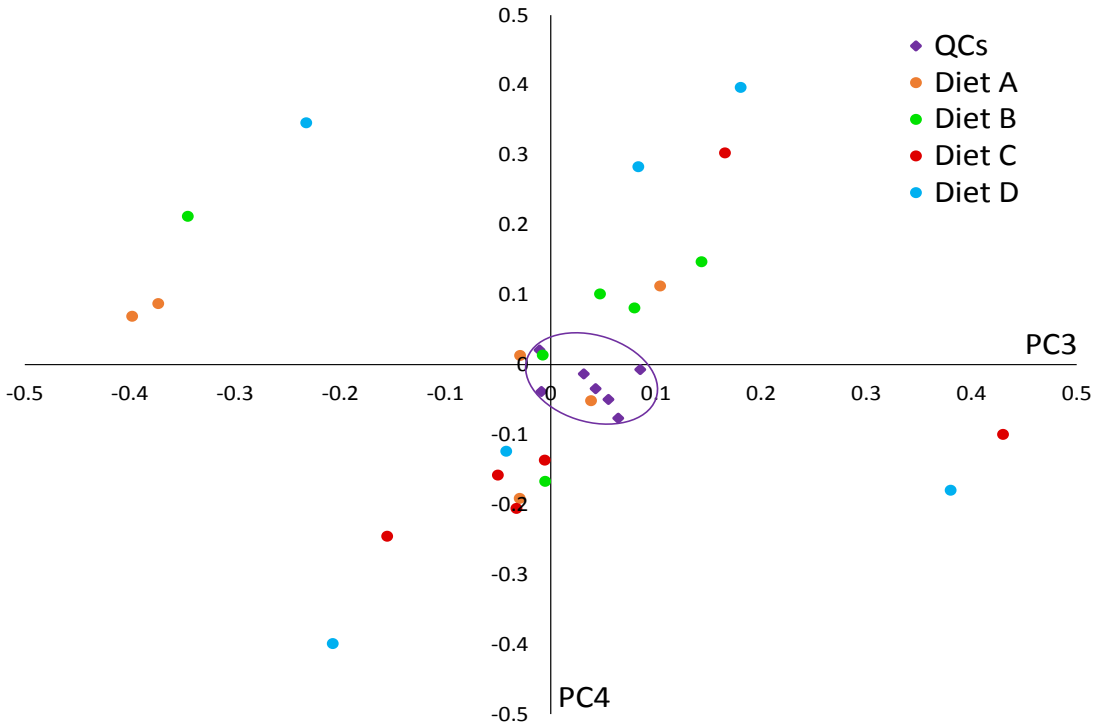


Figure 6.3: *PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds that were fed four different diets, two weeks post diet change, including all compounds with CV%<30%. PC3 explains 0.38% of the variance, and PC4 explains 0.15% of the variance.*

The scores plot in Figure 6.3 does not display the subtle differences in metabolite profile that were observed in the overlaid TICs in Figure 6.2. The only variation that is

present is between the yolks of random eggs, rather than between the yolks of eggs laid by birds which were fed the different diets. This suggests that the differences between the metabolite profiles which were observed in Figure 6.2, are not great enough for the PCA to uncover. The QC samples form a group on the scores plot that is more tightly clustered than the other sample sets, indicating that the analysis was robust and that there was little instrumental drift affecting the analysis.

6.4.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 43 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, all 43 were found to be statistically significant. These compounds, their CV%s, p-values resulting from ANOVA/Welch tests, and trends in abundance between eggs laid by birds which were fed the four different diets, can be seen in Table C.1 in Appendix C. Those in bold were also found to be statistically significant following confirmation using the raw data.

The trends in the abundances of the compounds between the yolks of eggs laid by hens which were fed the four different diets can be seen in Table C.1 in Appendix C. Most of the compounds in this table have a similar abundance in the yolks of eggs laid by birds which were fed diet B, to the abundance in the yolks of eggs laid by birds which were fed diet A. The abundances of these compounds in the yolks of eggs laid by birds which were fed diets C and D are different to this, but similar to each other. The yolks of eggs laid by birds which were fed diet C appear to show the most difference in compound abundance compared to the yolks of eggs laid by birds which were fed diet A, whilst the yolks of eggs laid by birds which were fed diet D show slightly less difference.

6.4.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 43 statistically significant compounds, and scores plots were again produced to display the variation between yolks of eggs that were laid by birds which were fed the four different diets.

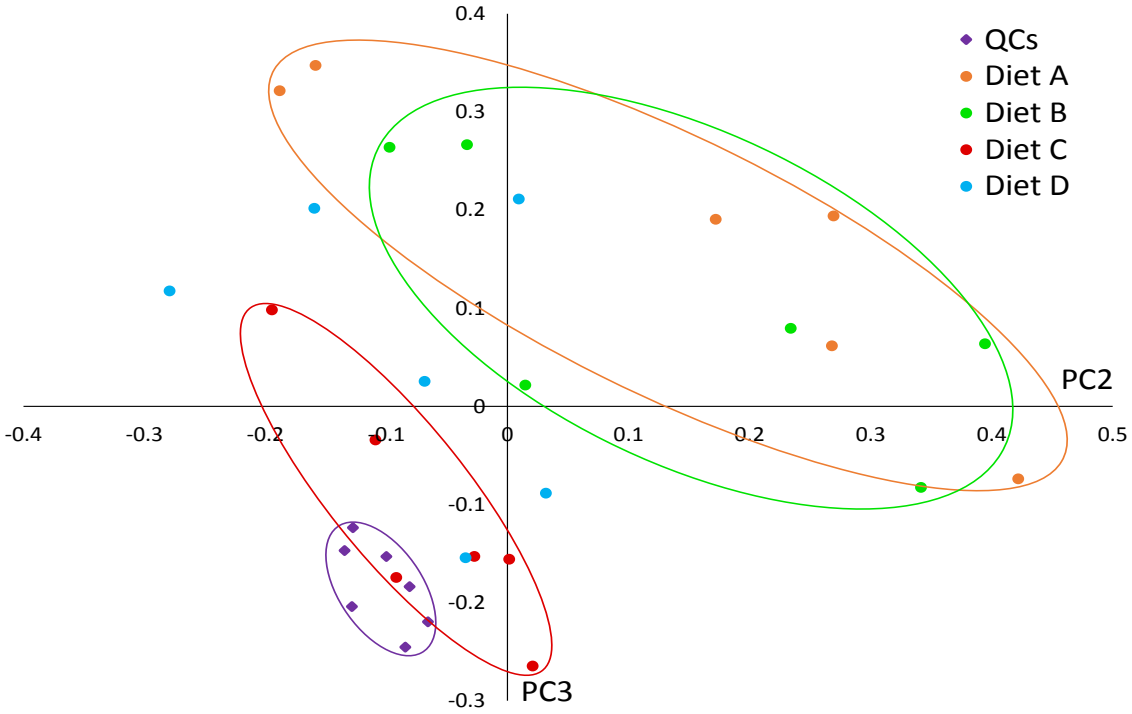


Figure 6.4: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by birds that were fed four different diets, two weeks post diet change, including the top 43 statistically significant compounds. PC2 explains 0.65% of the variance, and PC3 explains 0.004% of the variance.*

The scores plot in Figure 6.4 shows some separation between the yolks of eggs laid by birds which were fed the different diets, indicating that the differences in the metabolite profiles of the yolks are stronger when only the top statistically significant compounds are included in the PCA. The yolks of eggs laid by birds which were fed diet B show no separation to those laid by birds which were fed diet A, which corroborates what was observed from the trends in the abundance of the top statistically significant compounds in Table C.1 in Appendix C. The yolks of eggs laid by birds which were fed diet C show complete separation from the yolks of eggs laid by birds which were fed both diets A

and B, but not from the yolks of eggs laid by birds which were fed diet D, which are spread throughout the plot. This again confirms what was observed from the trends in the abundances of the compounds in Table C.1. The variation between the yolk samples on the scores plot is across both PC2 and PC3, which describe just 0.65% and 0.004% of the variation respectively, showing just how subtle these differences are. Again, the QC samples are grouped more tightly together than the samples, showing that the differences between samples are due to true biological differences, rather than any instrumental drift.

6.4.1.4.2 Comparison between 2 weeks post diet change and 1 week before diet change

The significant differences of the top statistically significant compounds between the yolks of eggs laid by birds that were fed the four different diets, were compared between two weeks after the diets were changed, and one week before the diets were changed.

Table C.2 in Appendix C shows the ANOVA/Welch test p-values of the top 23 compounds that were found to be significantly different, following confirmation using the raw data, between the yolks of eggs laid by birds that were fed the four different diets, two weeks post diet change, as well as the p-values of these compounds one week prior to the change of diet. Those highlighted in red were found to be significantly different between the eggs laid by birds kept in the four different cages prior to any diet change, and so may not have been affected by the change in diet, but by different environmental conditions between the cages. There are more statistically significant compounds that showed a significant difference in abundance before the diets were changed, than did not, indicating that the differences in metabolite profile between the yolks of eggs laid by birds which were fed the four different diets are very subtle, and that other variables between the different cages of birds are responsible for some of the observed differences.

6.4.1.5 Summary

The overlaid TICs in Figure 6.2 showed that there was a very subtle difference in metabolite profile between the yolks of eggs laid by birds that were fed the four different diets. Although these differences were too subtle to show any significant variation on the first scores plot in Figure 6.3, the scores plot in Figure 6.4 did confirm the subtle differences in metabolite profile between the yolks of eggs laid by birds which were fed the four different diets. It showed no difference between the yolks of eggs laid by birds which were fed diet B and the yolks of eggs laid by birds which were fed diet A, which confirmed what was observed from the trends in the abundances of the top statistically significant compounds in Table C.1 in Appendix C. This could indicate either that diet B is very similar to diet A, or has a similar effect on the metabolite profile of the egg yolk to diet A, or that the influence of diet B on the metabolite profile of the egg yolk takes longer than two weeks to have an effect.

The yolks of eggs laid by birds which were fed diet C were shown to be different to the yolks of eggs laid by birds which were fed diets A and B on the scores plot in Figure 6.4, as well as from the trends observed in Table C.1. This indicates that either diet C is more different to diet A than diet B is, and therefore has a greater effect on the resulting metabolite profile of the egg yolk, or that is has a faster effect on the metabolite profile than diet B does. Figure 6.4 and Table C.1 also showed that the metabolite profiles of the yolks of eggs laid by birds which were fed diet D were equally similar to the profiles of the yolks of eggs laid by birds that were fed all three of the other diets. This suggests that diet D affects the metabolite profile of the yolk either a greater amount, or faster, than diet B, but not as much, or more slowly than, diet C.

6.5 5 weeks post diet change: Yolk Organic Extracts

Organic extracts of the yolk of eggs laid by four sets of hens which were fed four different diets were compared, five weeks after the diets of three sets of hens were changed, to see whether a difference in metabolite profile could be observed between the yolks of eggs laid by hens that were fed the four different diets.

6.5.1 Results and Discussion

6.5.1.1 Quality Control Analysis

Table 6.3 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 6.3: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens fed four different diets, five weeks post diet change*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	1774166	135298548	59815619	23678811	476135683	110685403
QC2	1644212	128411600	57204277	25186854	464159081	192474929
QC3	1522953	129153748	56353551	24184949	468986940	103194854
QC4	1806457	128025708	58917002	25812746	454093614	100433829
QC5	1608390	132564405	57219567	26104073	483247179	102440445
QC6	1474546	118635111	53214273	25742243	429337959	115920408
QC7	1525359	114628855	56311422	25985960	421503521	127931091
SD	128290	7413776	2118023	951893	23392757	32578274
Mean	1622298	126673996	57005102	25242234	456780568	121868708
CV%	7.91	5.85	3.72	3.77	5.12	26.73

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 6.4, shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 6.4: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for yolk organic extracts from eggs laid by hens fed four different diets, five weeks post diet change*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.690	9.703	33.324	57.642	72.748	79.985
QC2	4.744	9.691	33.262	57.563	72.636	79.856
QC3	4.760	9.690	33.261	57.629	72.651	79.888
QC4	4.745	9.708	33.179	57.580	72.603	79.840
QC5	4.745	9.742	33.346	57.731	72.886	80.223
QC6	4.772	9.719	33.207	57.608	72.597	79.801
QC7	4.792	9.722	33.343	57.727	72.783	80.186
SD	0.032	0.019	0.066	0.066	0.108	0.171
Mean	4.750	9.711	33.275	57.640	72.701	79.968
CV%	0.67	0.19	0.20	0.12	0.15	0.21

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

6.5.1.2 Metabolite Profiling

Comparisons were made between the yolks of eggs laid by hens that were the fed four different diets, by studying the TICs of the extracts following HPLC-MS analysis. Figure 6.5 shows overlaid TICs for the average chromatograms of organic extracts of yolk from eggs laid by hens that were fed these four diets.

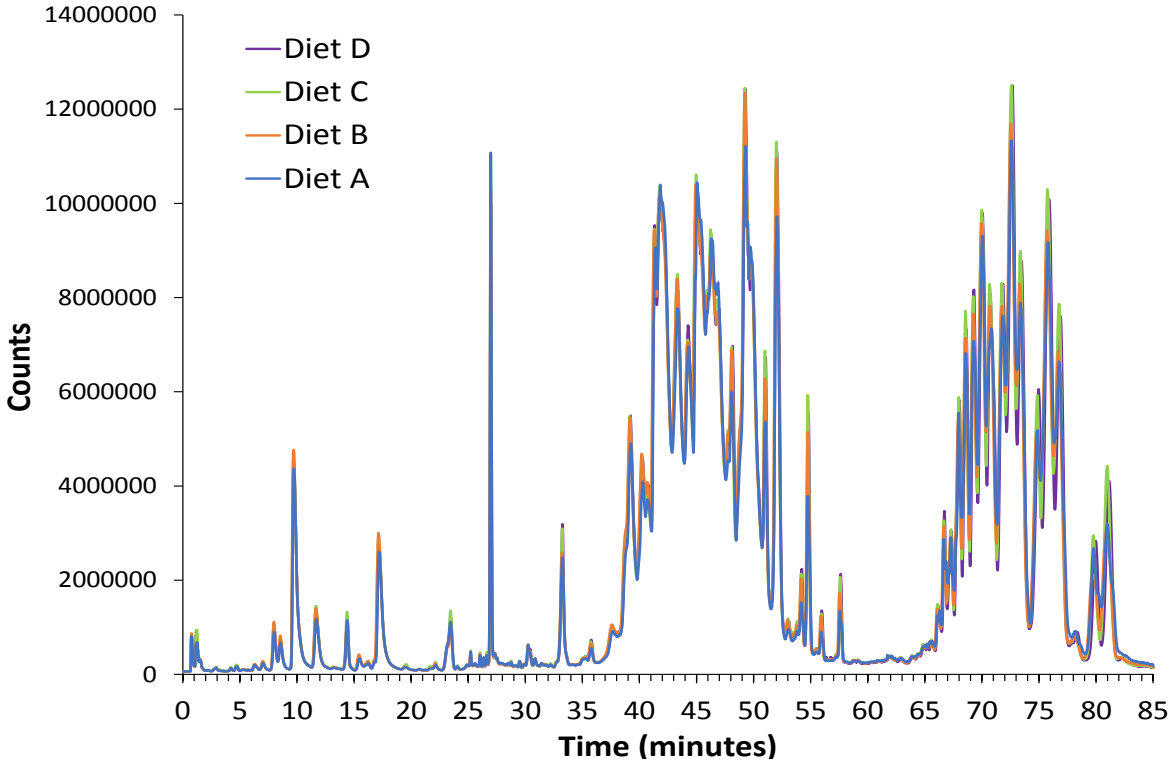


Figure 6.5: *Overlaid average TICs of organic extracts of the yolk of eggs laid by hens that were fed four different diets, five weeks post diet change*

A subtle difference in metabolite profile between the yolks of eggs laid by birds that were fed the four different diets, five weeks post diet change, can be seen by observing the overlaid TICs in Figure 6.5. Over the first 25 minutes of analysis, there appears to be a slightly higher peak intensity in the TICs of yolk from eggs laid by birds which were fed diets B and C. This suggests, similarly to what was observed for diet C when comparing the metabolite profiles just two weeks post diet change, either that these diets are rich in the less non-polar compounds that elute earlier on in the analysis, or that they result in an increased abundance of these compounds in the yolk. The peak intensities for the

TICs of the yolk from eggs laid by birds which were fed diets A and D are similar to each other during these first 25 minutes of analysis. Throughout the rest of the analysis, there appears to be a very similar peak intensity in the TICs of yolk from eggs laid by birds which were fed diets B, C, and D, and a lower peak intensity in the TIC of yolk from eggs laid by birds which were fed diet A. This suggests that diet A may be less rich in the more non-polar compounds than the other three diets, or that it results in a lower abundance of these compounds being present in the yolk.

6.5.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the yolks of eggs laid by hens which were fed the four different diets.

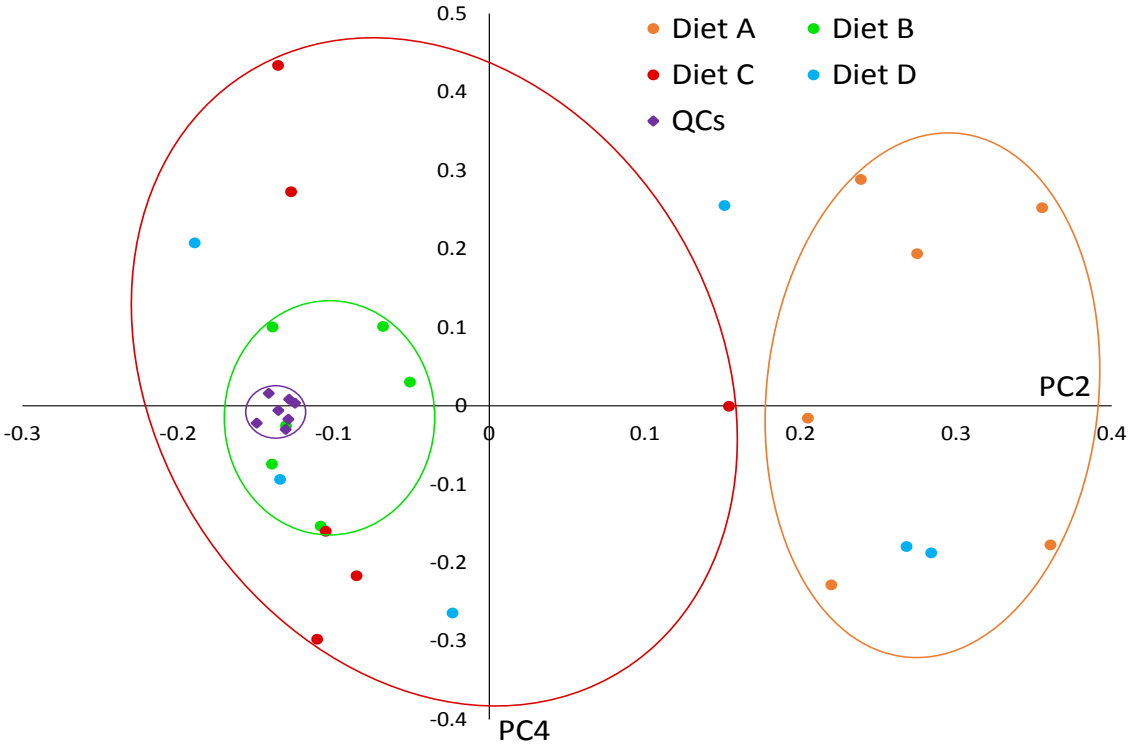


Figure 6.6: *PCA scores plot showing PC2 vs PC4 for yolk organic extracts of eggs laid by birds that were fed four different diets, five weeks post diet change, including all compounds with CV%<30%. PC2 explains 0.85% of the variance, and PC4 explains 0.15% of the variance.*

The scores plot in Figure 6.6 confirms what was observed in the overlaid TICs in Figure 6.5, and shows that the yolks of eggs laid by birds which were fed diets B and C have greater similarities to each other, than to the yolks of eggs laid by birds which were fed diet A. However, the yolks of eggs laid by birds which were fed diet D show a wide variation and are spread throughout the plot, overlapping into the otherwise independent group of yolks of eggs laid by birds which were fed diet A. The separation between the yolks of eggs laid by birds which were fed diet A, and those laid by birds which were fed diets B and C, is across PC2 which accounts for only 0.85% of the variance between samples, indicating that the difference in metabolite profile between these yolks is very subtle. The QC samples on the scores plot in Figure 6.6 are grouped tightly together, confirming that the analysis was robust and that there was little instrumental drift affecting the analysis.

6.5.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 31 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, 30 were found to be statistically significant. These compounds, their CV%s, p-values resulting from ANOVA/Welch tests, and trends between eggs laid by birds which were fed the four different diets, can be seen in Table C.3 in Appendix C. Those in bold were also found to be statistically significant following confirmation using the raw data.

The trends in the abundances of the compounds in the yolks of eggs laid by hens that were fed the four different diets, five weeks post diet change, can be seen in Table C.3 in Appendix C. These trends show that the yolks of eggs laid by birds which were fed diets B, C, and D show the most similarities in compound abundance to each other, but have quite different compound abundances to the yolks of eggs laid by birds which were fed diet A. This confirms what was observed in the overlaid TICs in Figure 6.5 and in the scores plot in Figure 6.6.

6.5.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 30 statistically significant compounds, and scores plots were again produced to display the variation between the yolks of eggs that were laid by birds which were fed the four different diets.

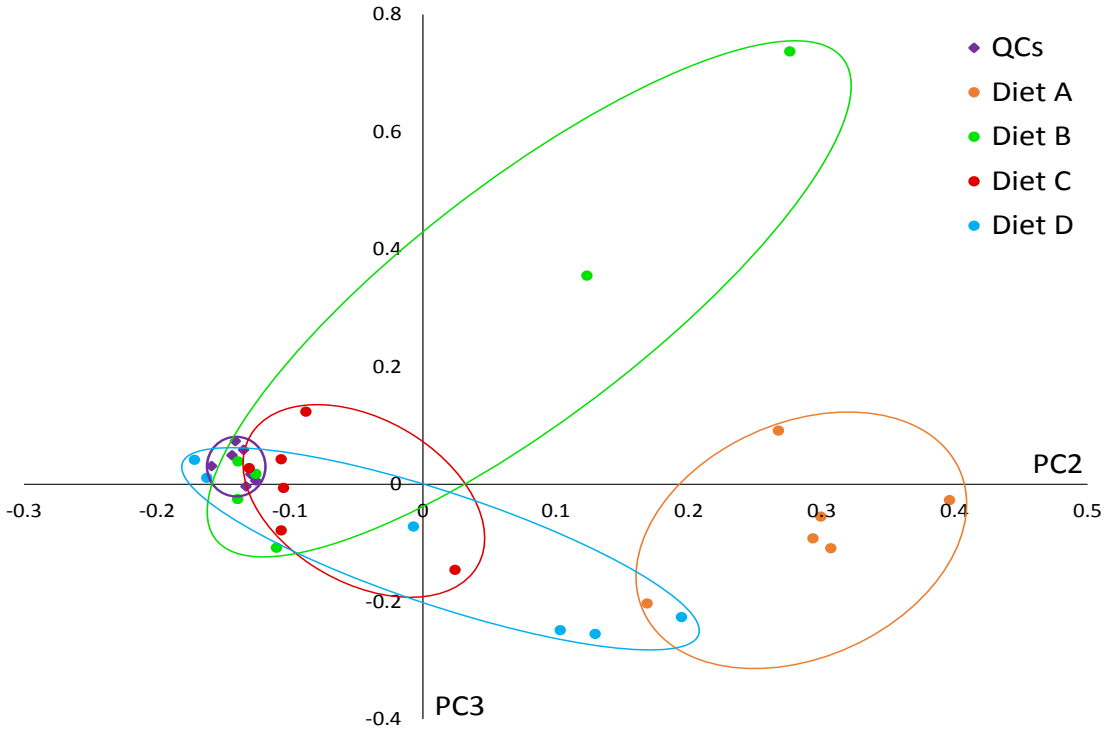


Figure 6.7: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of eggs laid by birds that were fed four different diets, five weeks post diet change, including the top 30 statistically significant compounds. PC2 explains 8.1% of the variance, and PC3 explains 0.36% of the variance.*

The scores plot in Figure 6.7 shows a similar separation between the yolks of eggs laid by birds that were fed the four different diets, to what was observed in Figure 6.6. This shows that in this case, using only the top statistically significant compounds in the PCA does not affect the outcome, indicating that all compounds with a CV%<30% show a similar pattern of variation to the top statistically significant compounds.

The yolks of eggs laid by birds which were fed diets B, C, and D all show some similarities to each other, as they overlap on the plot, but a difference to the yolks of eggs laid by birds which were fed diet A. The yolks of eggs laid by birds which were fed

diets B and C show a particularly strong difference, as they do not overlap at all with the yolks of eggs laid by birds which were fed diet A. However, the yolks of eggs laid by birds which were fed diet D, again also show some similarities to the yolks of eggs laid by birds which were fed diet A. This confirms what was observed from the overlaid TICs in Figure 6.5, the previous scores plot in Figure 6.6, and from the trends in compound abundance in Table C.4.

Again, the separation between the yolks of eggs laid by hens which were fed the different diets is across PC2, which now describes 8.1% of the variance. This increase in variance that is described by PC2 shows that including only the top statistically significant compounds in the PCA strengthens the variance described between the sample sets. The QC samples again form a tight grouping on the plot, indicating that the analysis was robust and that there was little instrumental drift affecting the analysis.

6.5.1.4.2 Comparison between 5 weeks post diet change and 1 week before diet change

The significant differences of the top statistically significant compounds between the yolks of eggs laid by birds that were fed the four different diets, were compared between five weeks after the diets were changed, and one week before the diets were changed.

Table C.4 in Appendix C shows the ANOVA/Welch test p-values of the top 20 compounds that were found to be significantly different following confirmation using the raw data, between the yolks of eggs laid by birds that were fed the four different diets, five weeks post diet change, as well as the p-values of these compounds one week prior to the change of diet. Those highlighted in red were found to be significantly different between the eggs laid by birds kept in the four different cages prior to any diet change and so again, may not be statistically significant as a result of the change in diet, but due to different environmental conditions between the sets of birds. There are fewer statistically significant compounds that showed a significant difference in abundance before the diets

were changed, than did not. This indicates that after a further three weeks following the change in diet of the laying hens, the differences in metabolite profile between the yolks of eggs laid by birds which were fed the four different diets are less subtle than after only two weeks.

6.5.1.5 Summary

The overlaid TICs in Figure 6.5 showed that there was a subtle difference in metabolite profile, between the yolks of eggs laid by birds that were fed the four different diets. The yolks of eggs laid by birds which were fed diets B and C showed a difference in metabolite profile to the yolks of eggs laid by birds which were fed diets A and D during the first 25 minutes of analysis. After 60 minutes of analysis, the yolks of eggs laid by birds which were fed diets B, C, and D appeared to show a significant difference in metabolite profile, compared to the yolks of eggs laid by birds which were fed diet A. These observations were confirmed by the scores plots in Figures 6.6 and 6.7, which showed that the yolks of eggs laid by birds which were fed diets B and C were different to the yolks of eggs laid by birds that were fed diet A, whilst the yolks of eggs laid by birds which were fed diet D showed some difference to the yolks of eggs laid by birds that were fed diets B and C, as well as diet A. The trends in Table C.3 also showed that the compound abundances were the most different between the yolks of eggs laid by birds which were fed diet A and the yolks of eggs laid by birds which were fed the other three diets, confirming what was observed in the overlaid TICs and the scores plots.

These results indicate that diets B and C are similar to each other, or have a similar effect on the resulting metabolite profile of the egg yolk, and are different to diet A, resulting in a different influence on the metabolite profile of the yolk. The results also suggest that diet D is either less different to diet A than diets B and C are, or has a lesser effect on the metabolite profile of the yolk, or that diet D takes longer than diets B and C to have a significant effect on the metabolite profile of the yolk.

6.6 5 weeks post diet change: Albumen Organic Extracts

Organic extracts of the albumen of eggs laid by four sets of hens which were fed four different diets were compared, five weeks after the diets of three sets of hens were changed, to see whether a difference in metabolite profile could be observed between the albumen of eggs laid by hens that were fed the four different diets.

6.6.1 Results and Discussion

6.6.1.1 Quality Control Analysis

Table 6.5 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 6.5: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens fed four different diets, five weeks post diet change*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	15153978	47849665	99245880	935674	660245	9966617
QC2	14342228	44900807	94638795	945605	597374	9552835
QC3	14726559	46622820	97290620	955634	669317	10411850
QC4	15063884	46661897	97225414	1007782	659155	10330385
QC5	18049156	54489247	109740643	1195558	696536	12650132
QC6	21068259	61810095	120583762	1324846	852854	14194073
QC7	19806364	59300344	116700552	1336019	798086	13552932
SD	2733335	6830855	10508311	180545	89030	1892377
Mean	16887204	51662125	105060809	1100160	704795	11522689
CV%	16.19	13.22	10.00	16.41	12.63	16.42

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 6.6, shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 6.6: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens fed four different diets, five weeks post diet change*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	7.698	11.779	14.732	16.491	18.979	22.480
QC2	7.696	11.777	14.713	16.472	18.927	22.461
QC3	7.732	11.796	14.750	16.507	18.996	22.513
QC4	7.779	11.838	14.791	16.533	19.021	22.538
QC5	7.807	11.891	14.841	16.583	19.071	22.588
QC6	7.806	11.904	14.857	16.598	19.087	22.604
QC7	7.818	11.898	14.852	16.594	19.082	22.616
SD	0.053	0.057	0.060	0.052	0.060	0.061
Mean	7.762	11.840	14.791	16.540	19.023	22.543
CV%	0.68	0.48	0.41	0.32	0.32	0.27

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

6.6.1.2 Metabolite Profiling

Comparisons were made between the albumen of eggs laid by hens that were fed the four different diets, by studying the TICs of the extracts following HPLC-MS analysis. Figure 6.8 shows overlaid TICs for the average chromatograms of organic extracts of albumen from eggs laid by hens that were fed these four diets.

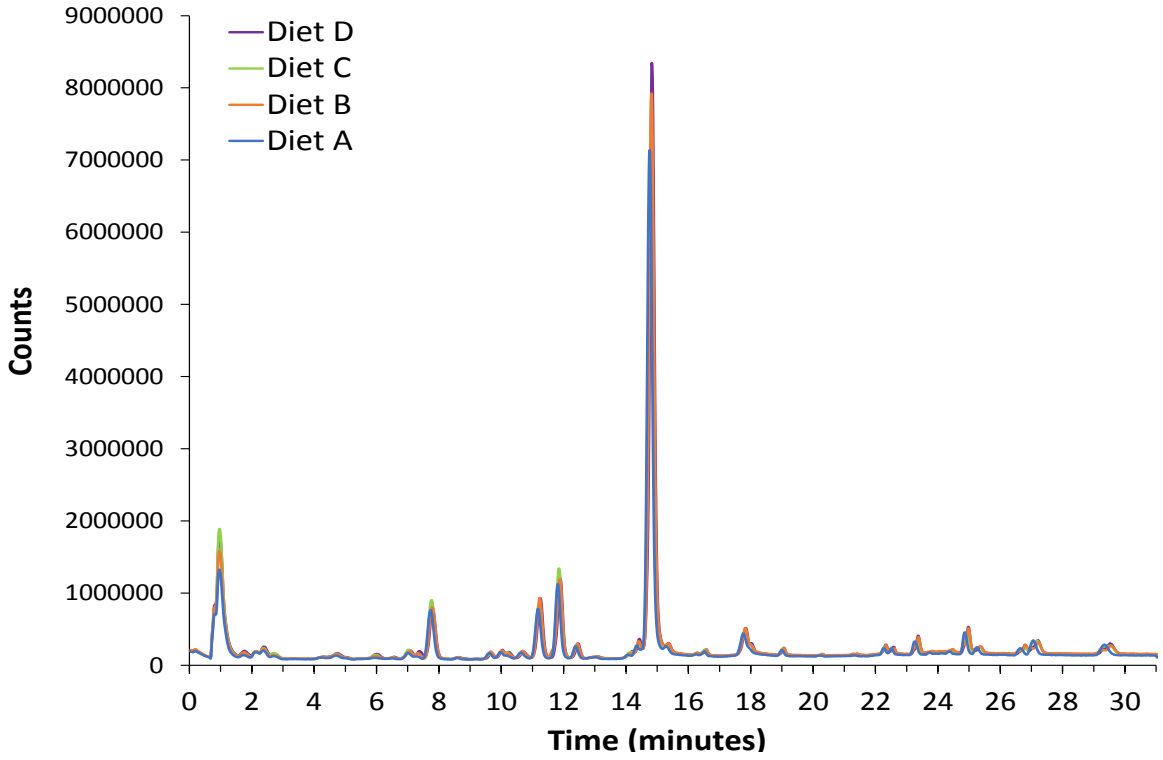


Figure 6.8: *Overlaid average TICs of organic extracts of albumen of eggs laid by hens that were fed four different diets, five weeks post diet change*

A slight difference in metabolite profile between the albumen of eggs laid by birds that were fed the four different diets, five weeks post diet change, can be seen by observing the overlaid TICs in Figure 6.8. The TIC of the albumen of eggs laid by birds which were fed diet A shows a lower peak intensity throughout the analysis, compared to the TICs of the albumen of eggs laid by birds which were fed the other three diets.

6.6.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the albumen of eggs laid by hens which were fed the four different diets.

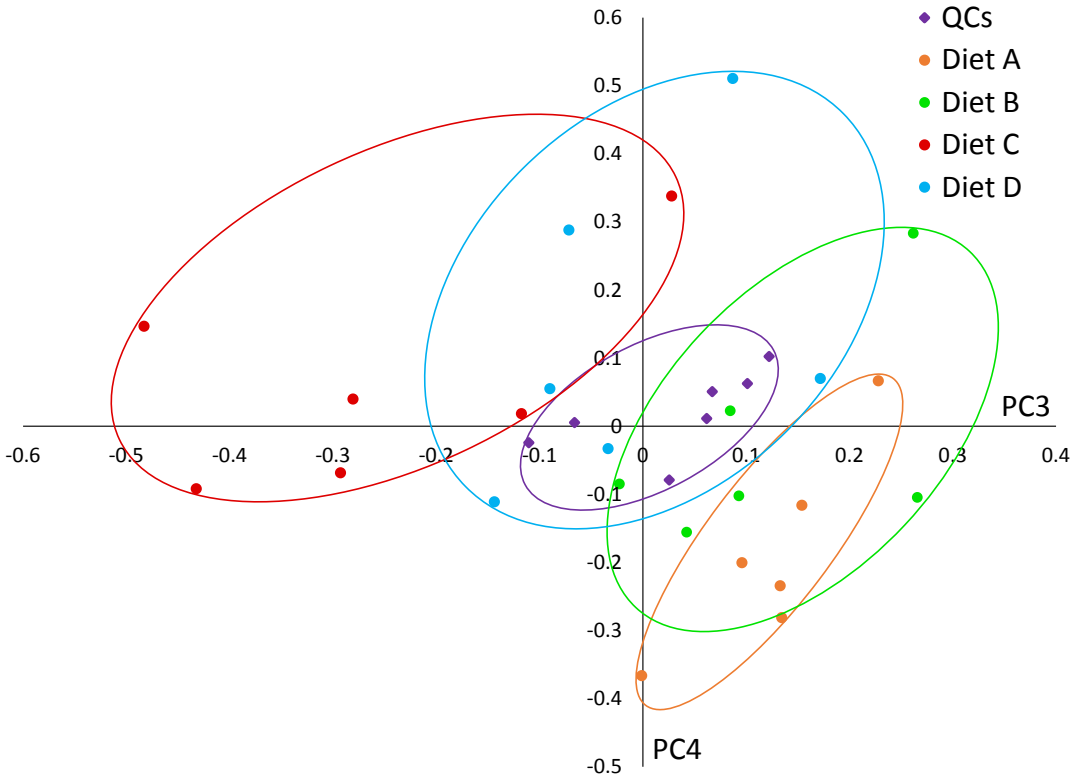


Figure 6.9: *PCA scores plot showing PC3 vs PC4 for albumen organic extracts of eggs laid by birds that were fed four different diets, five weeks post diet change, including all compounds with CV%<30%. PC3 explains 0.06% of the variance, and PC4 explains 0.03% of the variance.*

The scores plot in Figure 6.9 confirms that there are some differences in metabolite profile between the albumen of eggs laid by the birds which were fed the four different diets. The albumen of eggs laid by birds which were fed diet C show a complete separation to the albumen of eggs laid by birds which were fed diet A, and the albumen of eggs laid by birds which were fed diet D also show some separation. However, the plot shows no difference between the albumen of eggs laid by birds which were fed diet B and those which were fed diet A, which is different to what was observed from the TICs in Figure 6.8. This could be because the more highly abundant compounds that result in the

observable difference between the albumen of eggs laid by birds which were fed the four different diets in the overlaid TICs, may not be responsible for as much variation between the samples and sample sets, as some of the lower abundance compounds.

The variation between samples described by PC3 is just 0.06%, and that described by PC4 is 0.03%, showing just how subtle the differences between the albumen of eggs laid by birds which were fed the four different diets are. The QC samples form a tighter group than the other sample sets, showing that the analysis was robust and that there was little instrumental drift affecting the analysis.

6.6.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 37 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, all 37 were found to be statistically significant. These compounds, their CV%s, p-values resulting from ANOVA/Welch tests, and trends between eggs laid by birds which were fed the four different diets, can be seen in Table C.5 in Appendix C. Those in bold were also found to be statistically significant following confirmation using the raw data.

The trends in the abundances of the top compounds in the albumen of eggs laid by hens that were fed the four different diets, five weeks post diet change, can be seen in Table C.5 in Appendix C. These trends show that the albumen of eggs laid by birds which were fed diet B experience the least difference in terms of compound abundance to the albumen of eggs laid by birds which were fed diet A, whilst the albumen of eggs laid by birds which were fed diet C experience the greatest difference. The albumen of eggs laid by birds which were fed diet D show similarities in compound abundance to the albumen of eggs laid by birds which were fed diets A and B, as well as those which were fed diet C. This confirms what was observed from the scores plots in Figure 6.9.

6.6.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 37 statistically significant compounds, and scores plots were again produced to display the variation between the albumen of eggs that were laid by birds which were fed the four different diets.

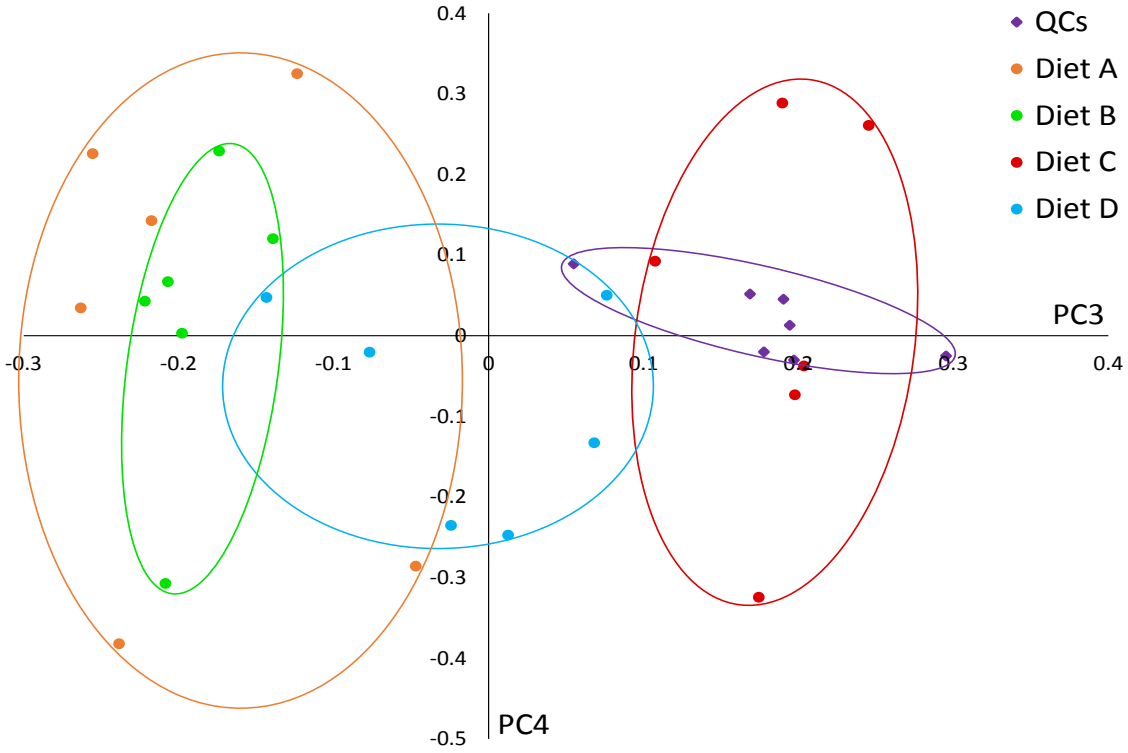


Figure 6.10: *PCA scores plot showing PC3 vs PC4 for albumen organic extracts of eggs laid by birds that were fed four different diets, five weeks post diet change, including the top 37 statistically significant compounds. PC3 explains 0.74% of the variance, and PC4 explains 0.36% of the variance.*

The plot in Figure 6.10 shows a similar separation between sample sets to what was observed previously in the scores plot in Figure 6.9. This shows that in this example, using only the top statistically significant compounds in the PCA does not affect the outcome of the analysis, indicating that including all compounds with a CV%<30% provides comparable results.

The albumen of eggs laid by birds which were fed diet B show no separation to the albumen of eggs laid by birds which were fed diet A, confirming what was observed in the previous scores plot in Figure 6.9 and from the trends in compound abundance in

Table C.5. The albumen of eggs laid by birds which were fed diet C show a complete separation to the albumen of eggs laid by birds that were fed diets A and B, whilst the albumen of eggs laid by birds which were fed diet D are spread across the plot, between the albumen of eggs laid by birds which were fed all three of the other diets. This again confirms what was observed in Figure 6.9 and Table C.6.

The variation between samples described by PC3 is 0.74%, and that described by PC4 is 0.36%, showing the subtlety of the differences between the albumen of eggs laid by birds which were fed the four different diets. Although the QC samples are spread across PC3, there was no instrumental drift occurring throughout the analysis. These QC samples again form a smaller group on the plot than the other sample sets, indicating that the analysis was robust and that there was little instrumental drift affecting the analysis.

6.6.1.4.2 Comparison between 5 weeks post diet change and 1 week before diet change

The significant differences of the top statistically significant compounds between the albumen of eggs laid by birds that were fed the four different diets, were compared between five weeks after the diets were changed, and one week before the diets were changed.

Table C.6 in Appendix C shows the ANOVA/Welch test p-values of the top 26 compounds that were found to be significantly different following confirmation using the raw data, between the albumen of eggs laid by birds that were fed the four different diets, five weeks post diet change, as well as the p-values of these compounds one week prior to the change of diet. Those highlighted in red were found to be significantly different between the eggs laid by birds kept in the four different cages prior to any diet change, and so may not have been significantly different as a result of the change in diet, but due to different environmental conditions between the cages of birds. As only a few compounds showed significant differences in abundance between the albumen of eggs laid by birds kept in the four different cages before the diets were changed, it indicates that the

greatest difference in metabolite profile between the albumen of eggs laid by birds which were fed the different diets, five weeks post diet change, is due to the difference in laying hen diet rather than any environmental differences between the cages of birds.

6.6.1.5 Summary

The overlaid TICs in Figure 6.8 showed a slight difference in metabolite profile between the albumen of eggs laid by birds that were fed the four different diets, particularly between the albumen of eggs laid by birds which were fed diet A, and those laid by birds which were fed the other three diets. The scores plots in Figures 6.9 and 6.10, and the trends in compound abundance in Table C.5 mostly corroborated this, and showed the metabolite profiles of the albumen of eggs laid by birds which were fed diet C to be very different to the profiles of the albumen of eggs laid by birds which were fed diet A, whilst the albumen of eggs laid by birds which were fed diet D showed a lesser difference in metabolite profile. This suggests either that diet C has a greater difference to diet A, than diet D does, or has a greater effect on the resulting metabolite profile of the albumen than diet D, or that diet C affects the metabolite profile of the albumen at a faster rate than diet D.

However, the scores plots and trends in compound abundance also showed that the metabolite profile of the albumen of eggs laid by birds which were fed diet B was very similar to that of the albumen of eggs laid by birds which were fed diet A. These results indicate either that diet B is very similar to the original diet, diet A, or has a similar influence on the resulting metabolite profile, or that the effect of diet B on the metabolite profile of the albumen takes longer than five weeks to become apparent.

6.7 Summary & Conclusions

6.7.1 Summary

Comparing the yolks of eggs laid by four sets of birds which were fed the four different diets, two weeks after the diets of three sets of birds were changed, it is possible to observe differences between the metabolite profiles of the yolks of these eggs. Diet C was shown to have had the greatest effect on the metabolite profile of the yolk, which suggests that this diet is either the most different to the original diet, diet A, or that it has the greatest or quickest impact on the metabolite profile. Diet B appeared to have very little effect on the metabolite profile of the yolk, indicating that either this diet is very similar to diet A, or has a similar effect on the metabolite profile of the yolk, or that this diet takes longer than two weeks to have an observable impact on the metabolite profile. The metabolite profile of the yolks of eggs laid by birds which were fed diet D was affected by the change in diet, but not as much as diet C. This suggests that either diet D is less different to diet A than diet C is, or has a lesser effect on the metabolite profile of the yolk, or that diet D affects the metabolite profile of the yolk at a slower rate than diet C.

Similar results were observed when comparing metabolite profiles of the yolks of eggs laid by birds which were fed the four different diets, five weeks after the diets were changed. Diet C still appeared to have had the greatest impact on the metabolite profile of the yolk, and the profiles of the yolks of eggs laid by birds which were fed diet D were again less affected by the change in diet. However, diet B was shown to have had a greater effect on the metabolite profile of the yolk, which showed a greater similarity to the profile of the yolks of eggs laid by birds which were fed diet C, rather than diet A. This suggests that diet B has a similar impact on the metabolite profile of the yolk to diet C, but that it takes longer to result in any observable effect.

It was expected that the comparison of the albumen of eggs laid by birds which were fed the four different diets, five weeks after the diets were changed, would produce similar

results to the comparison of the yolk five weeks post diet change. However, the results were more comparable to the comparison of the yolks just two weeks after the diets were changed. Diet C was again shown to have had the greatest effect on the metabolite profile of the albumen, whilst diet D appeared to have a slightly lesser effect. However, diet B was shown to have had very little effect on the metabolite profile of the albumen, with the profile showing no obvious difference to the profile of the albumen of eggs laid by birds which were fed diet A. This suggests that either diet B takes longer to influence the metabolite profile of the albumen than the yolk, or that diet B has less of an effect on the compounds found in the albumen than the yolk.

6.7.2 Conclusions

This study has shown that the diet of the laying hen does have a small effect on the resulting metabolite profiles of the yolk and albumen of the egg. Although the influence that diet has on the metabolite profiles is only subtle, the effects of many variables that could be investigated in metabonomic studies may also be very slight; therefore the diet of the laying hen may have a relatively large impact on the results of a metabonomic study. It is therefore important to be aware of the effect that different diets may have on any results, and where possible keep diets the same during metabonomic studies, in order to ensure that any significant results are due to the factor that is under investigation, and not caused by diet acting as an uncontrolled variable.

This work has revealed that a difference in the metabolite profiles of the yolks of eggs laid by four sets of birds which were fed four different diets, can be observed two weeks after the diets of three of the sets of birds were changed. This suggests that it may be difficult to detect cases of fraud using a metabonomic approach, if the fraudulent diet of the laying hens is switched to the correct diet two weeks before an investigation takes place. However, this study has also shown that different diets affect the metabolite profile of the yolk at different rates; therefore the successful detection of fraud depends

on how quickly the metabolite profiles of the yolk and albumen of eggs are affected, after switching the diet of the laying hens to a 100% organic feed. The influence of different diets on metabolite profile may vary between the yolk and albumen; the profile of one component may be affected more quickly, or more greatly, than the other. If one component is affected more quickly than the other by the switching of the diet of the laying hen to 100% organic feed, then it will be more difficult to detect fraud by studying the metabolite profile of this component than the other, which is affected at a slower rate. Likewise, if the metabolite profile of one component is affected to a lesser extent than the other by the switching of the diet of the laying hen to a 100% organic feed, then it would again be difficult to detect fraud; the metabolite profile would be similar regardless of which diet had been fed to the birds, and more success may be had by analysing the metabolite profile of the other component, which is affected to a greater extent.

In order to explore this further, similar studies using 100% organic feed, 95% organic feed with 5% non-organic protein feed, and 100% conventional feed as the original diets, switching to 100% organic feed for all diets, would have to be carried out. A range of varieties of each of the three different feeds would have to be included, as the differences between two varieties of 100% organic feed, or two varieties of 100% conventional feed, may result in different impacts on the eggs. Both the yolk and albumen of the eggs would need to be tested at regular intervals of every one or two days, to identify the time point at which the eggs laid by birds which were originally fed non-100% organic feeds, could be distinguished from eggs laid by birds that were still being fed these diets, and could no longer be successfully discriminated from eggs laid by birds which were always fed 100% organic feed.

6.8 References

[1] F. Courant, J. Antignac, G. Dervilly-Pinel, and B. L. Bizec, "Basics of mass spectrometry based metabolomics," *Proteomics*, vol. 14, pp. 2369–2388, 2014.

[2] H. Issaq, Q. Van, T. Waybright, G. Muschik, and T. Veenstra, "Analytical and statistical approaches to metabolomics research," *Journal of Separation Science*, vol. 32, pp. 2183–2199, 2009.

[3] G. A. Theodoridis, H. G. Gika, E. J. Want, and I. D. Wilson, "Liquid chromatography-mass spectrometry based global metabolite profiling: A review," *Analytica Chimica Acta*, vol. 711, pp. 7–16, 2012.

[4] W. B. Dunn, N. J. C. Bailey, and H. E. Johnson, "Measuring the metabolome: Current analytical technologies," *Analyst*, vol. 130, pp. 606–625, 2005.

[5] E. M. Lenz and I. D. Wilson, "Analytical strategies in metabonomics," *Journal of Proteome Research*, vol. 6, pp. 443–458, 2007.

[6] *Commission implementing Regulation (EU) 2018/1584 of 22 October 2018, amending Regulation (EC) No 889/2008 laying down detailed rules for the implementation of Council Regulation (EC) No 834/2007 on organic production and labelling of organic products with regard to organic production, labelling and control*, Official Journal of the European Union, L250/1-84, 2008.

[7] *Commission Regulation (EC) No 889/2008 of 5 September 2008, laying down detailed rules for the implementation of Council Regulation (EC) No 834/2007 on organic production and labelling of organic products with regard to organic production, labelling and control*, Official Journal of the European Union, L250/1-84, 2008.

[8] A. Tres, G. van der Veer, M. D. Perez-Marin, S. M. van Ruth, and A. Garrido-Varo, "Authentication of organic feed by near-infrared spectroscopy combined with chemometrics: A feasibility study," *Journal of agricultural and food chemistry*, vol. 60, pp. 8129–8133, 2012.

[9] Y. Nys and N. Guyot, "Improving the safety and quality of eggs and egg products," in. Woodhead Publishing Limited, 2011, vol. 1: Egg Chemistry, Production and Consumption, ch. 4: Egg formation and chemistry, pp. 83–132.

[10] R. Poureslami, K. Raes, G. Huyghebaert, A. B. Batal, and S. de Smet, "Egg yolk fatty acid profile in relation to dietary fatty acid concentrations," *Journal of the Science of Food and Agriculture*, vol. 92, pp. 366–372, 2012.

[11] M. C. Milinsk, A. E. Murakami, S. T. M. Gomes, M. Matsushita, and N. E. de Souza, "Fatty acid profile of egg yolk lipids from hens fed diets rich in n-3 fatty acids," *Food Chemistry*, vol. 83, pp. 287–292, 2003.

[12] H. Yalçýn, M. K. Ünal, and H. Basmacýoolu, "The fatty acid and cholesterol composition of enriched egg yolk lipids obtained by modifying hens' diets with fish oil and flaxseed," *Grasa y Aceites*, vol. 58, pp. 372–378, 2007.

[13] M. Omidi, S. Rahimi, and M. A. K. Torshizi, "Modification of egg yolk fatty acids profile by using different oil sources," *Veterinary Research Forum*, vol. 6, pp. 137–141, 2015.

[14] D. D. Oliviera, N. C. Baião, S. V. Cançado, R. Grimaldi, M. R. Souza, L. J. C. Lara, and A. M. Q. Lana, "Effects of lipid sources in the diet of laying hens on the fatty acid profiles of egg yolks," *Poultry Science*, vol. 89, pp. 2484–2490, 2010.

[15] A. Cobos, L. de la Hoz, M. I. Cambero, and J. A. Ordóñez, "Dietary modification and hen strain dependence of egg yolk lipids," *Food Research International*, vol. 28, pp. 71–76, 1995.

[16] Z. Hayat, G. Cherian, T. N. Pasha, F. M. Khattak, and M. A. Jabbar, "Effect of feeding flax and two types of antioxidants on egg production, egg quality, and lipid composition of eggs," *Journal of Applied Poultry Research*, vol. 18, pp. 541–551, 2009.

[17] R. Ayerza and W. Coates, "Dietary levels of chia: Influence on yolk cholesterol, lipid content and fatty acid composition for two strains of hens," *Poultry Science*, vol. 79, pp. 724–739, 2000.

[18] S. R. Chowdhury, S. D. Chowdhury, and T. K. Smith, "Effects of dietary garlic on cholesterol metabolism in laying hens," *Poultry Science*, vol. 81, pp. 1856–1862, 2002.

[19] M. Skřivan, V. Skřivanová, and M. Marounek, "Effects of dietary zinc, iron, and copper in layer feed on distribution of these elements in eggs, liver, excreta, soil and herbage," *Poultry Science*, vol. 84, pp. 1570–1575, 2005.

7. Differences in metabolite profile between shell egg yolk and liquid yolk stored for different lengths of time

7.1 Introduction

Poultry shell eggs must reach the consumer within a maximum time limit of 21 days from lay, and have a best before date of 28 days post-lay, according to European Union (EU) legislation [1]. Not only is the age of eggs being sold to consumers a legal issue, but the sell-by date of eggs was found, by a survey conducted by Fearne and Lavelle into consumer attitudes towards purchasing eggs, to be the most important factor considered by the highest number of respondents, when purchasing eggs [2]. However, there is no way for the consumer to confirm whether the age of the eggs that they have purchased corresponds to what the label states. One method that many people use in their home kitchens to assess the freshness of an egg, is to submerge it in water; if the egg sinks then it is still fresh, but if it floats to the top then it is not. There is, in fact, some science behind this; as an egg ages, water evaporates out through the shell and the egg experiences an increase in volume of the air cell [3]. This means that the density of the

egg decreases with age, and thus older eggs will be less dense and more likely to float when submerged in water. However, this cannot aid in predicting the actual age of the egg, it simply provides an indication of egg freshness. As there is no simple and reliable way for consumers to determine the age of eggs that they purchase, it would be easy for workers, at any point during the process between laying and selling, to be tempted to mislabel the eggs with incorrect lay dates, or mislabel the dates on the packaging, in order to falsely give the eggs a longer period of time before reaching their "sell-by"date. This could result in increased sales, as consumers are more likely to purchase produce with a later best before date, and would also result in less waste, and therefore less lost revenue, as eggs would not be taken off the shelves and disposed of at their true sell-by date.

Eggs are labelled with their lay date by packing centre operatives on the day that they are laid, when received not in a container from units on the same site as the packing centre. However, when the packing centre receives the eggs in a container from an external unit, eggs are labelled within two working days of being received. This happens either during, or immediately after, grading of the eggs [4]. If eggs are packed in a different centre to the one that graded them, then the packing and marking of eggs may be carried out within an additional three days [4].

Part of the legislation outlined in the APHA Guidance on Legislation Covering the Marketing of Eggs, that EMIs from APHA in England and Wales enforce, is the correct labelling of egg packaging with the egg lay date [5]. In January 2016, an egg packer from London was fined a total of £2,515 under the Eggs and Chicks (England) Regulations 2009 [6], for selling eggs incorrectly labelled with false sell-by dates [7]. However, no scientific testing methods were used to reveal this fraud; the eggs themselves had been labelled with an earlier sell-by date than the date on the packaging [7]. In fact, very little scientific analysis of samples takes place following an inspection by the EMIs; the only testing that is carried out on egg samples is to detect any medicine residues [8].

In order to reduce the temptation of fraud by misrepresenting food products, and to deter those who may be tempted, the risk of being uncovered needs to increase; one way of increasing this risk is to implement random testing of food products. With regard to fraudulent egg sell-by and best before date mislabelling, there are various testing methods that could be implemented.

From the moment that an egg is laid, the chemical and physical characteristics of the egg begin to change as the ageing processes take place within the egg [3]. As carbon dioxide and water evaporate out through the shell, the pH of the albumen and the volume of the air cell within the egg increase. This increase of pH causes some structural changes to take place in the albumen, which begins to thin, and the vitelline membrane surrounding the yolk becomes weakened [3]. Water migrates from the albumen through the weakened vitelline membrane, which causes the yolk to become flattened [9]. These physical changes allow for a range of relatively simple testing methods to take place, such as testing of the albumen pH, measuring the air cell height, and calculating the Haugh Unit of the egg [3].

Albumen pH ranges from 8.08-8.30 in freshly laid eggs, but as the carbon dioxide evaporates out, it has been observed to increase to 9.12-9.17 over a period of 8 days [10]. Most of this increase occurs during the first 4 days of storage, after which there is a much slower rate of increase throughout the remaining storage time [10]. This method of age determination would only be useful when measuring the pH over a short storage time, not over a longer time period such as 21 days; however it may be helpful for detecting when eggs have been falsely labelled as fresh, when they are in fact several days post-lay.

The EU regulations state that the air cell height must not exceed 6 mm at the point of packing [4]; measurement of air cell height is carried out in packing centres using a candling technique. This is a method that is used to assess the freshness of an egg [11], not specifically the age, so it is not a robust method of predicting egg age. However, it may be used as a simple, non-destructive indicator of egg age; if the cell height is above

6 mm it may require further testing to assess whether it is within the sell-by date or not.

Haugh Unit (HU) is based on the relationship between the weight of the intact egg, and the albumen height once the egg has been broken [11]. The HU decreases over time as the egg weight decreases due to the evaporation of water, and the albumen height also decreases with time [9][12][13]. However, an issue with using HU as a measure of egg age is the fact that both the age and the breed of the laying hen also affect the HU [14].

The Yolk Index (YI), which is the height of the intact yolk of a broken egg divided by the width of the intact yolk [9], is another parameter that changes over time. As the yolk becomes flattened when water migrates through the vitelline membrane to the yolk from the albumen, the YI decreases with increased storage time [9][13]. However, storage temperature affects this migration of water [15]; refrigeration of an egg at 7 °C has been observed to lower the rate of decrease of YI [9]. Although most studies agree that the YI decreases over egg storage time [9][13][16], one study found that the opposite was true, and that YI increased over an increasing storage time of eggs [15].

As well as methods based on the physical changes that occur in an egg over time, there are ways to test the age of an egg based on chemical changes within the egg. Spectroscopic methods such as Infra-Red (IR) and Visible Transmission spectroscopy have been shown to successfully provide information about the freshness or age of an egg. One study was able to successfully discriminate eggs by age, as well as classify the age of the eggs based on IR spectroscopy of the intact eggs [17]. Another study found that the light transmission spectrum of eggs provides quantitative information about egg freshness [11].

Studies have been carried out to determine whether the change in the furosine concentration of eggs is useful in predicting egg age. It has been observed that in albumen, furosine content increases over time, with higher storage temperatures causing a higher rate of increase, and refrigerated temperatures reducing the rate significantly [3]. It was determined that the furosine content of albumen would be a good indicator of egg freshness in countries where egg refrigeration is not common practice, as there is a steady

increase of furosine content over time and there is low variability between eggs of the same age [3]. The furosine content of yolk is lower than in albumen, but follows a similar trend of increasing over time. However, the use of furosine content in yolk, as well as albumen, to predict egg age in countries that refrigerate eggs is limited, as there is very little change in furosine levels over time at low temperatures [18].

Another compound that may be useful for egg age determination, based on increasing content over time, is uridine. In albumen it has been observed that uridine content increases exponentially; at 5 °C it is constant up to 150 days of storage, at which point it increases rapidly. This makes it useful only over long storage periods at 5 °C, but it may also be useful over shorter storage times at higher temperatures [3]. Uridine has also shown an exponential increase in abundance over time in egg yolk, but is not useful for predicting egg age as it shows a high natural variability [18]. This is also true for pyroglutamic acid, another compound present in egg yolk, which has been found to show a linear increase over storage time [18]. Pyroglutamic acid has also been observed to increase linearly over time in albumen, however the increase is very small and it again shows a high natural variability, so is not suitable for egg age prediction [3].

Ovalbumin is the most abundant protein found in the albumen of eggs, making up approximately 55% of egg protein content. Over time, ovalbumin is converted into its isomer, S-ovalbumin, in an irreversible reaction [12]. It has been found that at 25 °C there is a high correlation between S-ovalbumin content and storage time, with S-ovalbumin increasing from 14.42% (of total ovalbumin and S-ovalbumin) in fresh eggs to 91.86% after 27 days, allowing a prediction model to be created which is able to estimate the age of an egg based on S-ovalbumin content [12].

In addition to legislation regarding the sell-by and best before dates of shell eggs, there is also EU legislation regarding the maximum storage time of liquid egg, which is the liquid contents of the shell egg, i.e. the yolk and the albumen. During the manufacture of egg products, liquid egg must either be frozen, or stored at a temperature of 4 °C or

lower prior to use; if the liquid egg is stored at the refrigerated temperature, it must be stored for no longer than 48 hours [19]. The storage container of the liquid egg must be labelled with the date and hour of shell breaking [19].

In the U.K., egg processors must be approved by the local authority as food business operators for egg products; this approval is subject to them satisfying the conditions laid out in EC Regulation 853/2004 [5]. Inspections of these egg processing facilities may be carried out at any time by EMIs from APHA, who have the power to enforce action in the case of infringements [5]. However, these EMIs do not carry out any testing of the egg; they detect any infringements based on observations and paperwork.

If the paperwork does not show evidence of any legislation breaches, then they are unlikely to be uncovered without any scientific testing. As there is no way to determine how long the liquid egg was stored for once the egg has been incorporated into a product, it is important to uncover breaches of legislation prior to product manufacture. This necessitates the development of a robust technique that can predict the length of storage time of liquid egg that has been refrigerated. If EMIs were able to take random samples of refrigerated liquid egg to be tested, in order to confirm that it had been stored for under 48 hours, it would deter egg processing staff from storing liquid egg for longer than this time period, as the risk of being uncovered would increase.

There appears to be very little existing research into the observation of physicochemical changes within liquid egg during storage. Although some of the changes observed in the yolk and albumen of whole shell eggs during storage may also be observed in liquid egg during storage, some of these changes are due to interactions between the various components of the shell egg, which may not occur when the components are stored separately. This emphasises the need for the development of a method capable of determining the storage time of liquid egg.

Various existing methods have been successful in predicting egg age, which could be used to help detect cases of fraud. However, a metabonomic approach would allow

information to be obtained about how a wide range of compounds and compound classes found in eggs, are affected by an increase in egg age. This could lead to the development of further methods capable of detecting mislabelled dates on eggs, and detecting liquid egg that has been stored for longer than is legally allowed.

7.2 Aims and Objectives

The accurate prediction of egg age, using a robust scientific method, could help to detect cases of fraud caused by mislabelling eggs with false sell-by and best before dates. The ability to predict the age of liquid egg could also help to detect legislation breaches, caused by the liquid egg being stored for longer than is allowed.

This work aimed to use a non-targeted metabonomic workflow to uncover differences between the yolks of eggs of different ages, up to five weeks post-lay, when eggs are stored at 23 °C [†]*. It then aimed, using the same workflow, to observe differences in yolk between eggs of different ages, again up to five weeks post-lay, when eggs are stored at 5 °C.*

Next, it aimed to determine whether storing eggs at a lower temperature can prevent the metabolic changes that are characteristic of egg ageing, in egg yolk, from occurring, and therefore hinder the prediction of egg age based on these observed changes [‡]*.*

Finally, this work aimed to use the same non-targeted metabonomic workflow to observe differences between liquid egg yolks that have been stored at 5 °C for different lengths of time, ranging from fresh yolk, to yolk that has been stored for 96 hours.

[†]This research has been published in *Analytical Chemistry*[*]

[‡]This research has been published in *Food Control*[†]

[*]A. E. Johnson, K. L. Sidwick, V. R. Pirgozliev, A. Edge, and D. F. Thompson, "Metabonomic profiling of chicken eggs during storage using High-Performance Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry", *Analytical Chemistry*, vol. 90, pp. 7489-7494, 2018.

[†]A. E. Johnson, K. L. Sidwick, V. R. Pirgozliev, A. Edge, and D. F. Thompson, "The effect of storage temperature on the metabolic profiles derived from chicken eggs", *Food Control*, vol. 109, 2020.

7.3 Experimental

A non-targeted metabonomic study was carried out to observe how the metabolite profile of egg yolk changes with increasing egg age, for eggs kept at two different storage temperatures: 23 ℃, and 5 ℃. A follow up targeted study, focussing on just one identified compound, was then carried out for eggs stored at both temperatures, in order to gain more quantitative information, and to see if the results were reproducible over a longer period of time. A non-targeted study was carried out in order to compare the metabolite profiles of the yolks of eggs that had been stored at 23 ℃ and 5 ℃ for five weeks, to observe the effects of storage temperature on the metabolite profiles of eggs. It also enabled the determination of whether storing eggs at a lower temperature can prevent the changes to the metabolite profile, that are characteristic of egg ageing, from occurring. A separate study was then carried out to observe how the metabolite profile of liquid egg yolk changes over a 96 hour period, when kept refrigerated at 5 ℃.

7.3.1 Non-targeted Egg Ageing Study

For the initial non-targeted study, eggs were collected on the day of lay from the National Institute of Poultry Husbandry (Harper Adams University, Newport, U.K.). Laying hens were of the Hy-line brown breed, 21 weeks old at the point of lay, fed the same diet, and kept in enriched cages with 8 birds per cage. An organic liquid extraction was carried out on the yolks of six of the eggs on the day of collection (fresh eggs), and half of the remaining eggs were stored at 23 ℃, whilst the other half were kept in refrigerated conditions at 5 ℃. An extraction was carried out on the yolks of six eggs from each storage temperature every seven days, up to a total of five weeks. At the conclusion of the experiment, organic extracts had been obtained from the yolks of six eggs at two different storage temperatures, across six time points.

7.3.1.1 Non-targeted Egg Ageing Study: 23 °C

Organic metabolite extraction of the yolk, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. For the MS/MS analysis in this study, the mass range was lowered to m/z 25-200 and spectra were collected at three different collision energies: 10, 20, 40 V. The data analysis workflow for this study was slightly different to that described in Chapter 3, Section 3.9, as a second PCA was carried out following the initial ANOVA, using only compounds with a CV%<30% and p<0.01. It was the PC1 loadings from this second PCA that were used to rank the compounds highest to lowest, before choosing the top 100 for further analysis. The diagram in Figure 7.1 describes the workflow that was used for the data analysis in this study.

7.3.1.1.1 Compound Identification

In order to confirm the identification of a compound, a chemical standard of choline chloride (≥99%, purchased from Sigma-Aldrich) was used to make a 53 µg/mL standard solution in methanol, which was then analysed alongside an injection of QC sample, in order to compare the retention times and mass spectra of choline and the compound of interest at three different collision energies.

7.3.1.2 Non-targeted Egg Ageing Study: 5 °C

Organic metabolite extraction of the yolk, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, data pre-processing and data analysis were all carried out as described in Chapter 3.

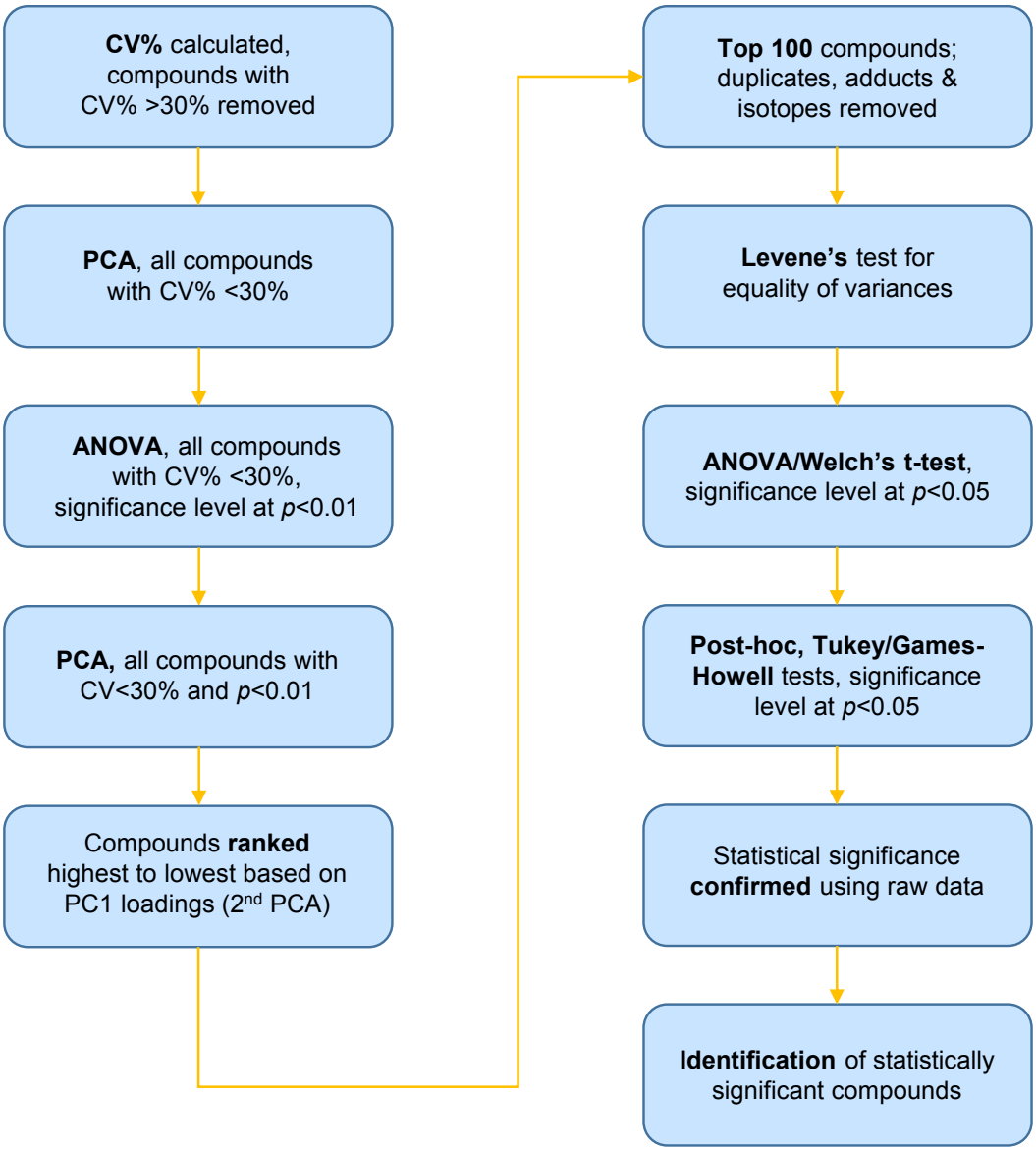


Figure 7.1: *Diagram showing the statistical workflow that was implemented in the non-targeted egg ageing study: 23 °C*

7.3.1.3 Non-targeted Egg Ageing Study: Comparison of eggs stored at 5 °C and 23 °C for five weeks

Organic metabolite extraction of the yolk, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. The data analysis workflow for this study was slightly different to that described in Chapter 3, Section 3.9; two tailed t-tests were carried out rather than

ANOVA as there were only two sample sets, which also negated the need for post-hoc tests. F-tests for equality of variances were carried out prior to any t-tests, so only one set of t-tests was required. A second PCA was carried out following the t-tests, using only compounds with a CV%<30% and p<0.01. It was the PC1 loadings from this second PCA that were used to rank the compounds highest to lowest, before choosing the top 100 for further analysis. The diagram in Figure 7.2 describes the workflow that was used for the data analysis in this study.

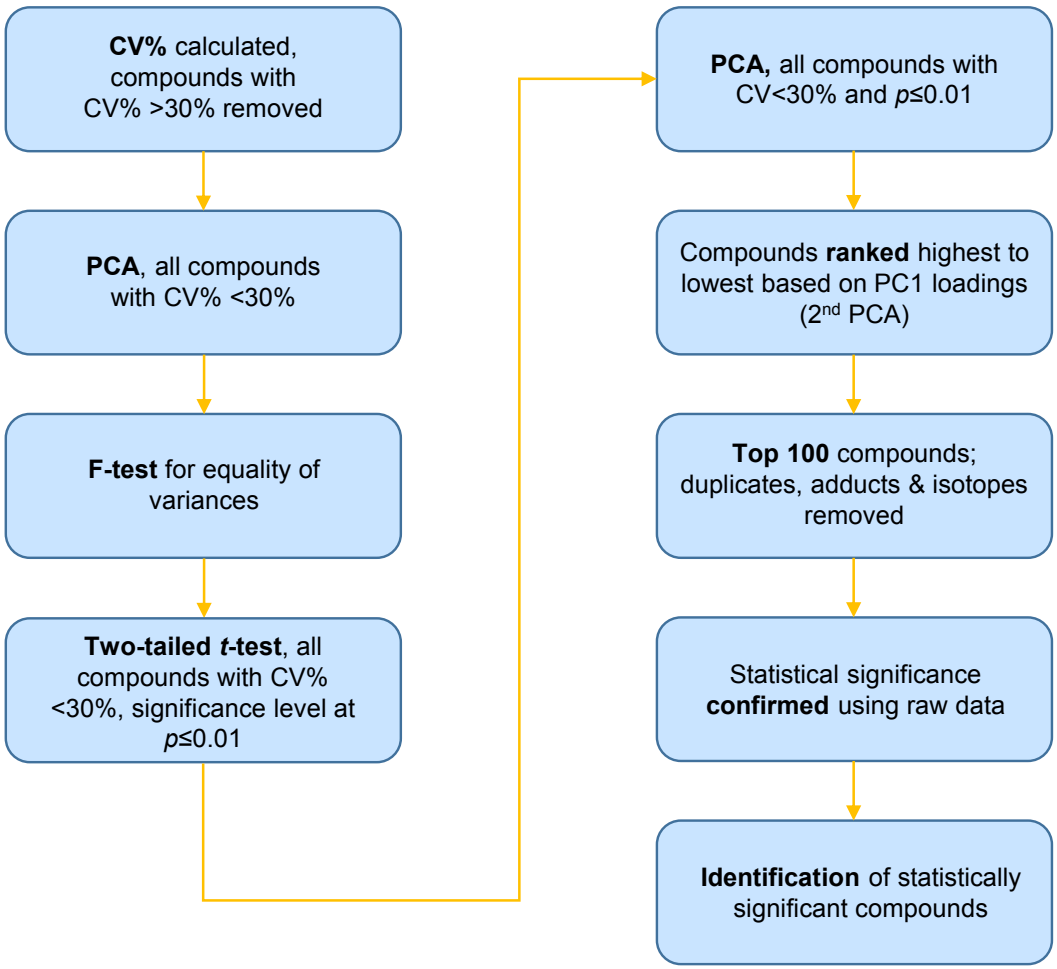


Figure 7.2: *Diagram showing the statistical workflow that was implemented in the non-targeted egg ageing study: comparison of eggs stored at 23 °C and 5 °C for five weeks*

7.3.2 Targeted Egg Ageing Study

The experimental design for the follow up targeted study was very similar. Eggs were collected from Oaklands Farm Eggs Ltd. (Shrewsbury, U.K.); laying hens were of the Hy-line brown breed, 46 weeks old at the point of lay, fed the same diet, and kept in enriched cages of 80 birds per cage. Again, an extraction was carried out on the yolks of six eggs on the day of collection (fresh eggs) and half of the remaining eggs were stored at 23 ℃, whilst the other half were stored at 5 ℃. This study was carried out over a 12 week period, with an extraction carried out on six eggs from each storage temperature every seven days, resulting in sample extracts from thirteen time points ranging from fresh eggs to eggs that were 12 weeks old.

Organic metabolite extraction of the yolk was carried out as described in Chapter 3. External standard solutions of choline chloride in methanol were produced, ranging from 0.2-2.6 µg/mL for eggs stored at 23 °C, and 0.2-0.8 µg/mL for eggs stored at 5 °C. Each of these standards was analysed three times, both before and after the analysis of the sample extracts.

7.3.2.1 Chemical Analysis

As this study was targeted, the chemical analysis was different in order to target the specific compound of interest. The chromatographic parameters were largely the same as described in Chapter 3, however the chromatographic method was much shorter as the targeted compound is polar and therefore eluted very early on in the analysis. The method was as outlined in Table 7.1, with a post-time of two minutes to allow the instrument to return to starting conditions.

The MS parameters again remained largely the same as described in Chapter 3, however the mass range was lowered to m/z 50-150.

Table 7.1: *Table showing the chromatographic method for the targeted analyses in the egg ageing study*

Injection volume: 1.5 µL Flow rate: 0.2 mL/min		
Time (minutes)	**Solvent A%**	**Solvent B%**
0	95	5
1	95	5
1.1	90	10
2.5	95	5
3	95	5

The choline chloride standards of 1.4 µg/mL and 0.5 µg/mL were used as the QC samples in the targeted analyses for eggs stored at 23 °C and 5 °C respectively, and were injected between every six or seven samples throughout the analytical sequences. As this was a targeted study, there was no need for large QC analyses. The CV%s of the choline chloride standards that were injected several times throughout the analyses were calculated, and Relative Error percentages (RE%s) and CV%s were calculated to monitor the accuracy and precision of the standards.

There was no pre-processing of data in this study; analysis took place using the raw data. The peak areas from the EICs of m/z 104.1 were used to carry out statistical analysis, and standard and drift calibration curves were produced based on the mean peak areas of the EIC peaks of the standards. Concentrations of choline in each of the yolk organic extracts were calculated, and a Levene's test was carried out to determine the variance of the data, followed by two ANOVA tests; one for eggs up to five weeks of age, and one for eggs up to the full twelve weeks of age, to test for statistical significance. Post-hoc Tukey tests were then carried out pairwise between eggs of different ages to see the significant difference between specific ages of eggs.

7.3.3 Non-targeted Liquid Egg Yolk Ageing Study

The experiment studying the ageing of liquid egg yolks had a different experimental set up. Eggs were collected from Oaklands Farm Eggs Ltd. (Shrewsbury, U.K.); laying hens were of the Hy-line brown breed, 39 weeks old at the point of lay, fed the same diet, and kept in enriched cages of 20 birds per cage. The yolks of several eggs were mixed together, in order to create one pooled sample, which was then separated into six different sample storage tubes, and stored in refrigerated conditions at 5 ℃. Although the legislation states that liquid egg must be stored at a temperature no greater than 4 ℃ [19], this was not possible with the resources available, so it was stored at just 1 ℃ greater, at 5 ℃. Prior to refrigeration, an organic liquid extraction was carried out on aliquots of all six samples in order to obtain fresh sample extracts. Extractions were then carried out at 24, 48, 60, 72, 84, and 96 hours after initial refrigeration, resulting in six sample extracts from each of the seven different time points.

Organic metabolite extraction of the yolk, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. The data analysis workflow for this study was as described in Chapter 3, Section 3.9, with an additional PCA using only the top statistically significant compounds following ANOVA/Welch tests.

7.4 Eggs Stored at 23 °C for 5 weeks: Yolk Organic Extracts

Yolk extracts from eggs that were stored at 23 °C for different lengths of time were compared to see whether there were any differences in their metabolite profiles.

7.4.1 Results and Discussion

7.4.1.1 Quality Control Analysis

Table 7.2 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 7.2: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 23 ℃*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	17660138	676558073	338452895	52414292	1709932400	1745631673
QC2	16977071	674815768	325103750	53416426	1735152487	1766907532
QC3	18286237	705286169	352518112	56070672	1807245238	1814532294
QC4	18224577	703380473	357815202	61281466	1759746130	1799166375
QC5	18552238	726328050	368542129	67119901	1775757372	1847255347
QC6	18621771	727037767	368372573	67221322	1742267941	1794506277
QC7	19072402	745908453	352160984	71550969	1855436871	1840364341
SD	690180	26621821	15719381	7555002	49013000	36845272
Mean	18199205	708473536	351852235	61296435	1769362634	1801194834
CV%	3.79	3.76	4.47	12.33	2.77	2.05

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 7.3 shows the retention times (RTs) for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 7.3: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 23 ℃*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	5.398	10.956	33.033	56.284	74.568	71.432
QC2	5.363	10.937	33.037	56.116	74.116	71.098
QC3	5.424	10.948	33.048	56.127	74.062	71.041
QC4	5.380	10.921	33.104	56.232	74.062	71.041
QC5	5.343	10.820	33.036	56.197	74.315	71.258
QC6	5.362	10.821	32.987	56.140	74.133	71.097
QC7	5.287	10.671	32.887	56.048	74.017	71.030
SD	0.044	0.104	0.070	0.079	0.210	0.160
Mean	5.365	10.868	33.024	56.163	74.237	71.186
CV%	0.81	0.96	0.21	0.14	0.28	0.22

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

7.4.1.2 Metabolite Profiling

The initial comparison between yolk extracts of eggs that were stored for different lengths of time was carried out by studying the Total Ion Chromatograms (TICs) of the extracts following HPLC-MS analysis. Figure 7.3 shows overlaid TICs for the average chromatograms of yolk extracts from fresh eggs and eggs that were five weeks old.

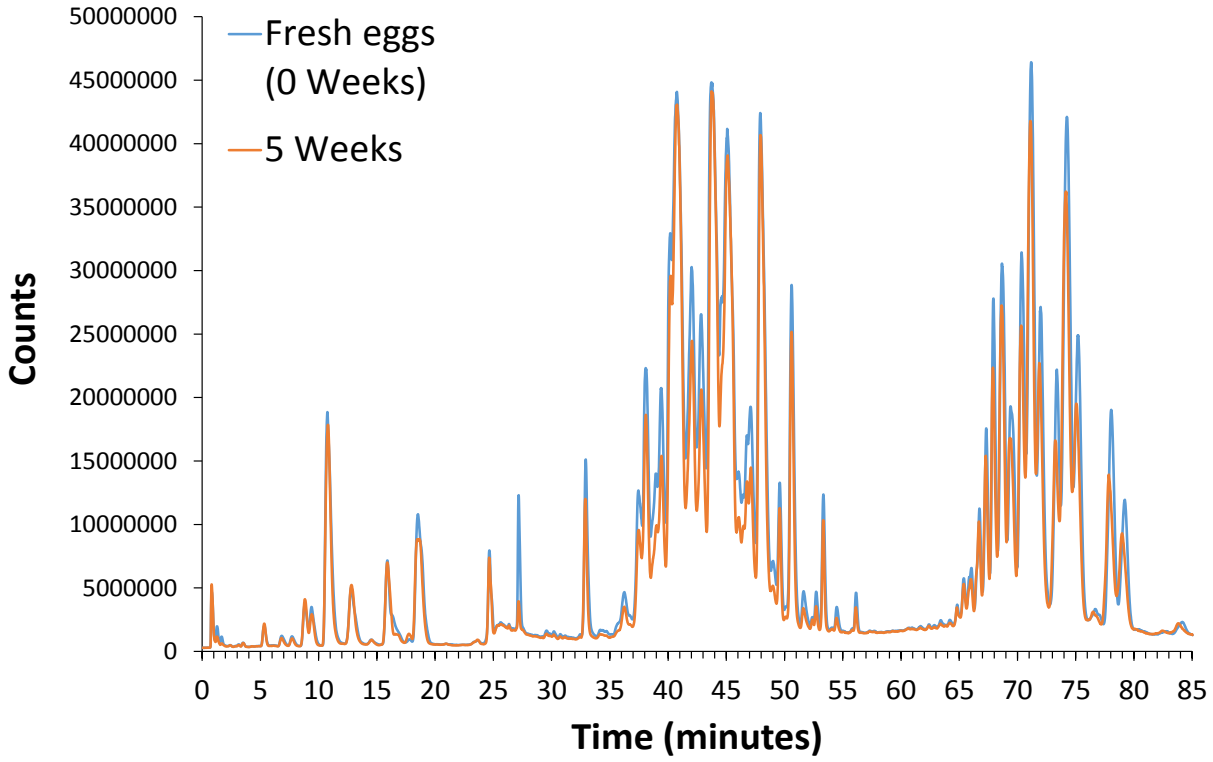


Figure 7.3: *Overlaid average TICs of yolk organic extracts from fresh eggs and eggs that were five weeks old, stored at 23 ℃*

As can be seen from these overlaid chromatograms, there is a visible difference in the metabolite profile of yolk between fresh eggs and eggs that were stored at 23 °C for five weeks. The peak intensity is higher for most peaks in the average TIC for fresh eggs compared to the older eggs. It may be that the more highly abundant compounds that contribute to the peaks observed in the TICs have all experienced a higher rate of catabolism or degradation during storage than the rate of production of new molecules, resulting in a lower abundance of these compounds in the yolks of eggs that were five weeks old, compared to the yolks of fresh eggs.

7.4.1.3 Multivariate Statistics

Following the first PCA, scores plots were produced to display the variation between the yolk extracts of eggs of different ages.

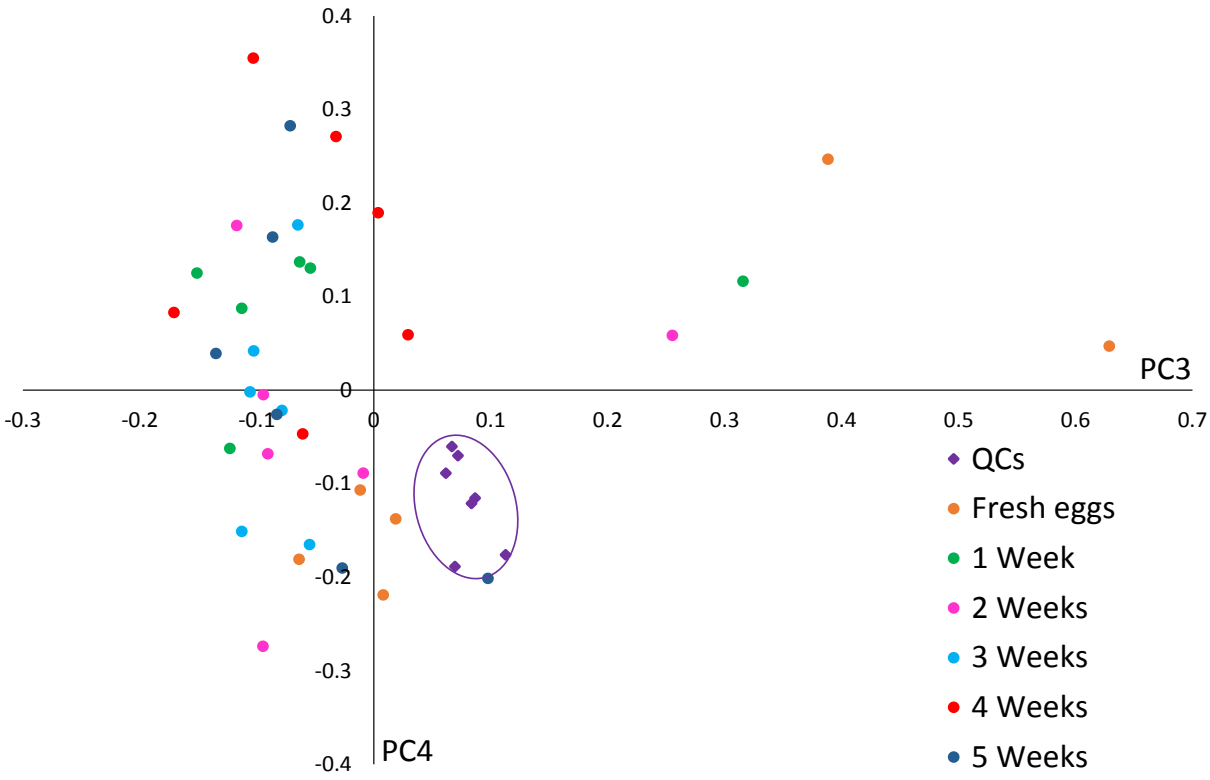


Figure 7.4: *PCA scores plot showing PC3 vs PC4 for eggs up to five weeks old that were stored at 23 ℃, including all compounds with CV%<30%. PC3 explains 0.41% of the variance, and PC4 explains 0.19% of the variance.*

Although the overlaid TICs in Figure 7.3 show a clear difference between fresh eggs and eggs that were five weeks post-lay, the scores plot in Figure 7.4 shows no obvious differences between the eggs of different ages; the most variation between sample extracts appears to be between random eggs, rather than specific sets of eggs, indicating that the differences between eggs of different ages are very subtle. This could be because the more highly abundant compounds that contribute to the visible difference between the TICs in Figure 7.3 may contribute less to the variance between samples that is explained by PC3 and PC4, compared to the less abundant compounds which may experience a greater difference between random samples, than between the yolks of eggs of different ages. The

QC samples are grouped more tightly together than the samples for the other sample sets, which shows that the analysis was robust and that there was little instrumental drift affecting the analysis.

After the initial ANOVA was carried out and compounds with $p>0.01$ were removed, a second PCA was carried out using the remaining compounds, and scores plots were again produced to display the variation between the yolk extracts from eggs of different ages.

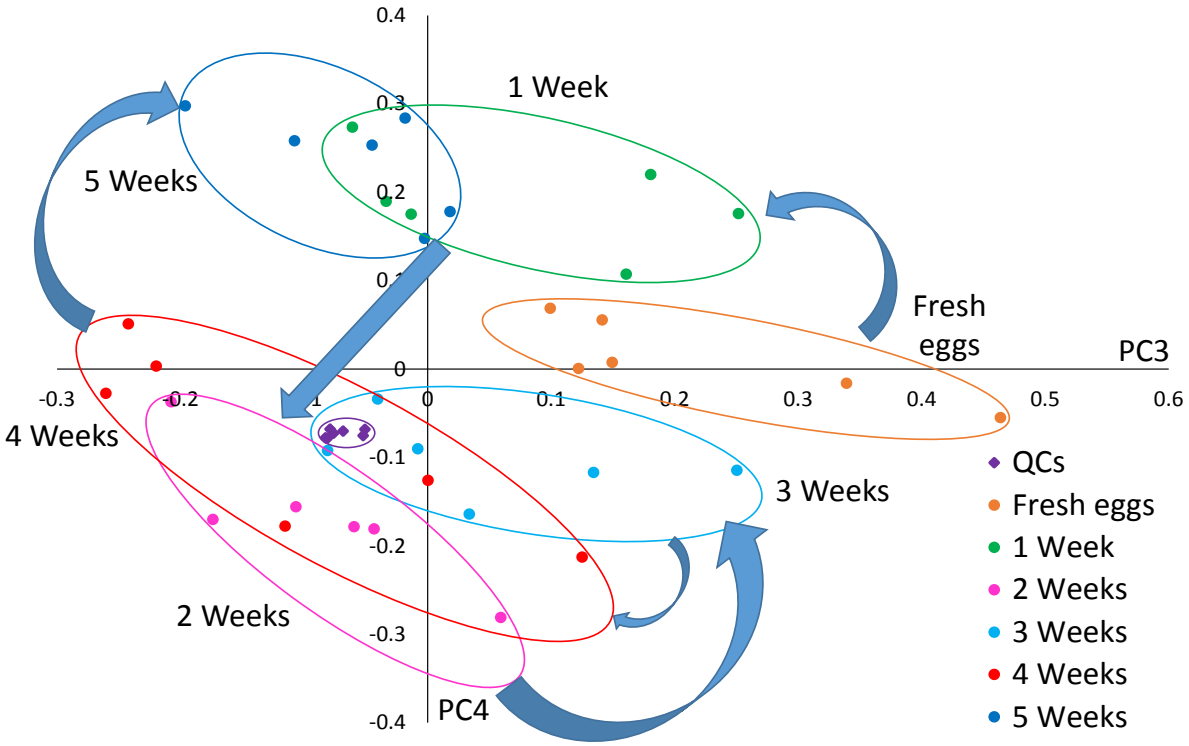


Figure 7.5: *PCA scores plot showing PC3 vs PC4 for eggs up to five weeks old that were stored at 23 ℃, including only compounds with CV%<30% and p<0.01. PC3 explains 0.22% of the variance, and PC4 explains 0.18% of the variance.*

Figure 7.5 shows that when only the statistically significant compounds ($p<0.01$) are included in the PCA, there is some visible separation between the sets of eggs that were stored for different lengths of time. Fresh eggs form their own group on the plot, showing that lots of metabolic change occurs in the yolk within the first week following the eggs being laid. There is also a large amount of separation between eggs that were one week old and eggs that were two weeks old, again showing that lots of changes occur during

this time period. Eggs that were two, three and four weeks old are grouped relatively closely together, indicating that not many metabolic changes occur within this period. However, eggs that were five weeks old are separate from these slightly younger eggs, grouping more closely to eggs that were just one week old. This is interesting, because it would be expected that the oldest eggs would be the most different, and thus show the most separation on a scores plot, from the younger eggs. This could be due to the possibility of several different metabolic pathways being involved in the production and breakdown of the compounds that contribute the most to the variation described by PC3 and PC4, resulting in an erratic change of abundance over an increasing storage time. In the yolks of eggs that were five weeks post-lay, these compounds that contribute to the variation described by PC3 and PC4 may have been present in a similar abundance to that in the yolks of eggs that were just one week old, resulting in them grouping together on the scores plot.

Although the separation between eggs of different ages is quite subtle on the scores plot in Figure 7.5, it is important to note that the differences between the sets of eggs themselves were also very subtle; the eggs were all laid by the same batch of birds, of the same age and breed, kept in the same conditions and fed the same diet, with the same number of birds per cage. The only difference was the length of time that the eggs were stored for prior to metabolite extraction. The subtlety of these differences is highlighted by the very small percentage variances explained by PCs 3 and 4, which are 0.22% and 0.18% respectively. The QC samples are tightly clustered together indicating that there was little instrumental drift occurring throughout the analysis.

7.4.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 44 compounds remaining, after the removal of any duplicates, adducts, or isotopes from the top 100 compounds based on PC1 loadings, 41 compounds were found to be statistically

significant between eggs of different ages. These compounds, and their CV%s and *p*-values, can be seen in Table D.1 in Appendix D; those in bold were also found to be statistically significant following confirmation using the raw data, and attempts were made to identify them.

Of these 38 compounds that were still found to be statistically significant when confirmed using the raw data, two were putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The observed *m/z* for both compounds was due to the $[M+H]^+$ adduct. The compound *m/z* 104.1070 was putatively identified as choline, and the compound *m/z* 331.2860 was putatively identified as 1-monopalmitin. The comparisons between mass spectra provided by METLIN and those resulting from the analysis in this study can be seen in Appendix D, Figures D.1 and D.2. Table 7.4, shows the putative identification of the compound *m/z* 104.1070, as well as the *p*-value resulting from ANOVA, and the *p*-values resulting from the Tukey tests.

Table 7.4: *Table showing the putative identification of compound m/z 104.1070 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
104.1070	$C_5H_{13}NO$	98.85	Choline	<0.001

Post-hoc test *p*-values						
	Fresh	**1 Week**	**2 Weeks**	**3 Weeks**	**4 Weeks**	**5 Weeks**
Fresh		0.877	0.583	0.139	0.001	<0.001
1 Week			0.994	0.692	0.023	<0.001
2 Weeks				0.937	0.078	<0.001
3 Weeks					0.412	0.001
4 Weeks						0.152
5 Weeks						

The probability score for the predicted formula $C_5H_{13}NO$ was 98.85 which is quite a high score, meaning that it is very likely that this formula corresponds to the compound

of interest. The p-value from ANOVA shows that the abundance of this compound is highly significantly different between the eggs of different ages, and the post-hoc tests corroborate this. These post-hoc test p-values show that there is a high statistical significance, particularly between the fresh/younger eggs and the older eggs, indicating that there may be a steady increase or decrease in the abundance of this compound with increasing egg age. This trend is confirmed by the graph in Figure 7.6, which shows a positive correlation between the abundance of this compound and increasing egg age.

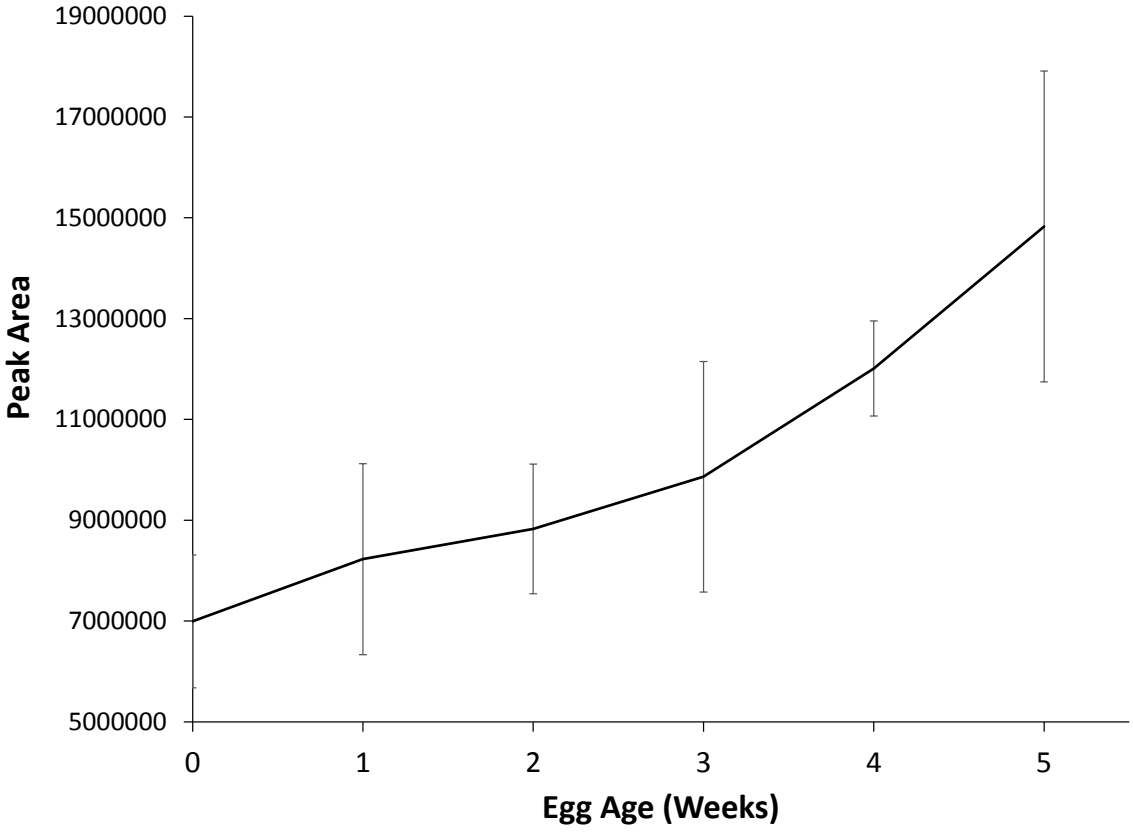


Figure 7.6: *Line graph showing the trend in changing abundance of compound m/z 104.1070 across eggs of six different ages. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

As this compound shows a correlation between abundance in egg yolk and increasing egg age, it has potential to act as a biomarker of egg age. If there was no correlation, even though there are significant differences in the abundance of this compound between eggs of different ages, it would not be suitable as an age related biomarker.

Table 7.5, shows the putative identification of the compound *m/z* 331.2860, as well as the *p*-value resulting from ANOVA, and the *p*-values resulting from the Tukey tests.

Table 7.5: *Table showing the putative identification of compound m/z 331.2860 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
331.2860	$C_{19}H_{38}O_4$	95.48	1-monopalmitin	<0.001

Post-hoc test *p*-values						
	Fresh	**1 Week**	**2 Weeks**	**3 Weeks**	**4 Weeks**	**5 Weeks**
Fresh		1	0.027	<0.001	0.731	0.613
1 Week			0.025	<0.001	0.694	0.827
2 Weeks				0.406	0.839	0.007
3 Weeks					0.231	<0.001
4 Weeks						0.370
5 Weeks						

This compound has been putatively identified as 1-monopalmitin; a monoglyceride with a palmitic acid chain attached to the first carbon of the glycerol backbone. This table shows that the probability score of the predicted formula, $C_{19}H_{38}O_4$, is 95.48 which, although it is only just within the 95 cut off, is still a high score, indicating a high likelihood that this formula corresponds to the compound of interest. The ANOVA *p*-value shows that the difference in abundance of this compound between eggs of different ages is highly significant, and this is confirmed by the *p*-values resulting from the post-hoc tests. However, these post-hoc test *p*-values seem to indicate that there is a more erratic change in abundance between different ages of eggs, rather than any correlation. The graph in Figure 7.7 confirms this; there is clearly no correlation between compound abundance and egg age, with the mean peak areas for this compound changing erratically over the six different time points. As there is such an erratic change in compound abundance between eggs of different ages, this compound is clearly not suited to being a biomarker of egg age, even though it is statistically significant.

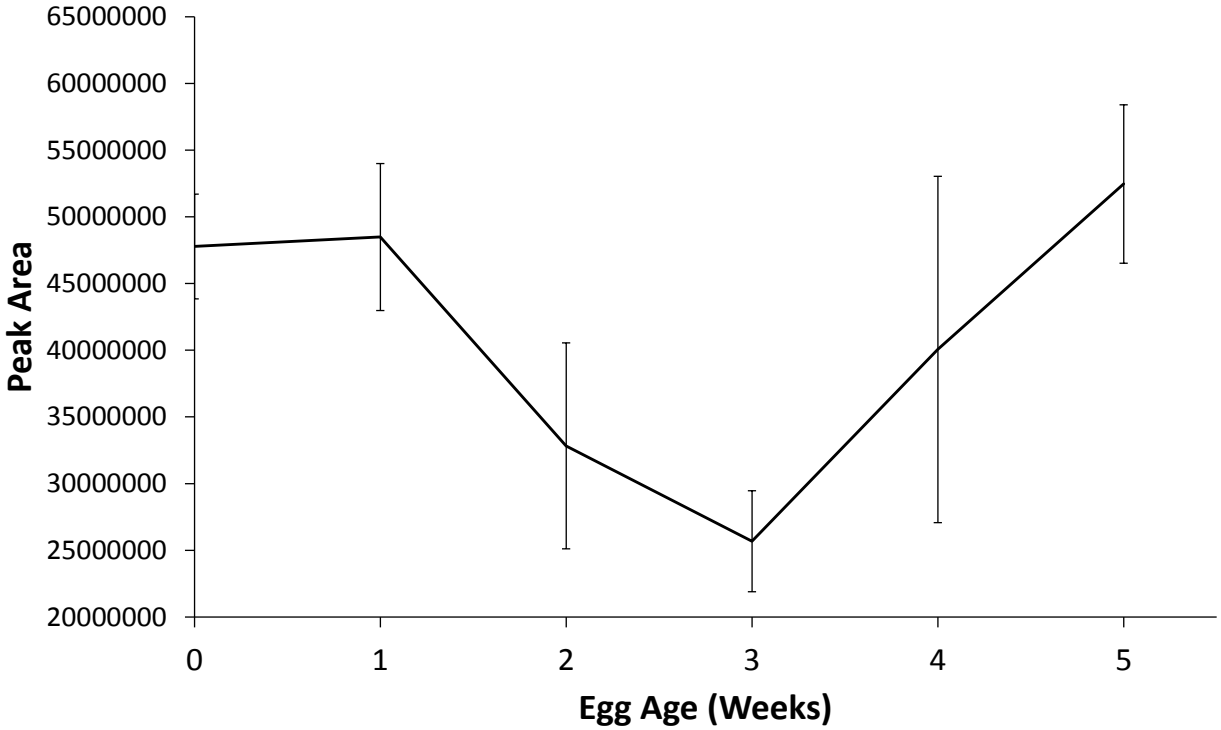


Figure 7.7: *Line graph showing the trend in changing abundance of compound m/z 331.2860 across eggs of six different ages. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Although the scaling of the y-axis makes the standard deviation error bars in Figures 7.6 and 7.7 appear larger, they are quite wide due to the large standard deviations resulting from the biological variation within the sets of different aged eggs, because the eggs all came from different birds. Although the variables regarding the birds were kept as controlled and uniform as possible, there was naturally some variation between them, and thus between the eggs that they laid. Larger sample sizes would likely reduce the standard deviations of the compounds within eggs of different ages.

7.4.1.5 Choline Identification

The compound m/z 104.1070 was putatively identified as choline, by comparing the mass spectrum provided by METLIN with the mass spectrum from the analysis. It has been shown to have potential as a biomarker of egg age, as its abundance in egg yolk is positively correlated with egg age, as can be seen in Figure 7.6, and so a confirmed identification was required.

Figure 7.8 shows overlaid EICs of the compound m/z 104.1000 (the m/z of the $[M+H]^+$ adduct of choline) for both the choline chloride standard and the QC sample. From this, it can be seen that the retention times of these compounds, in both samples, match (0.861 minutes in the choline chloride standard and 0.869 minutes in the QC sample) indicating that it is possible that the compound of interest is choline.

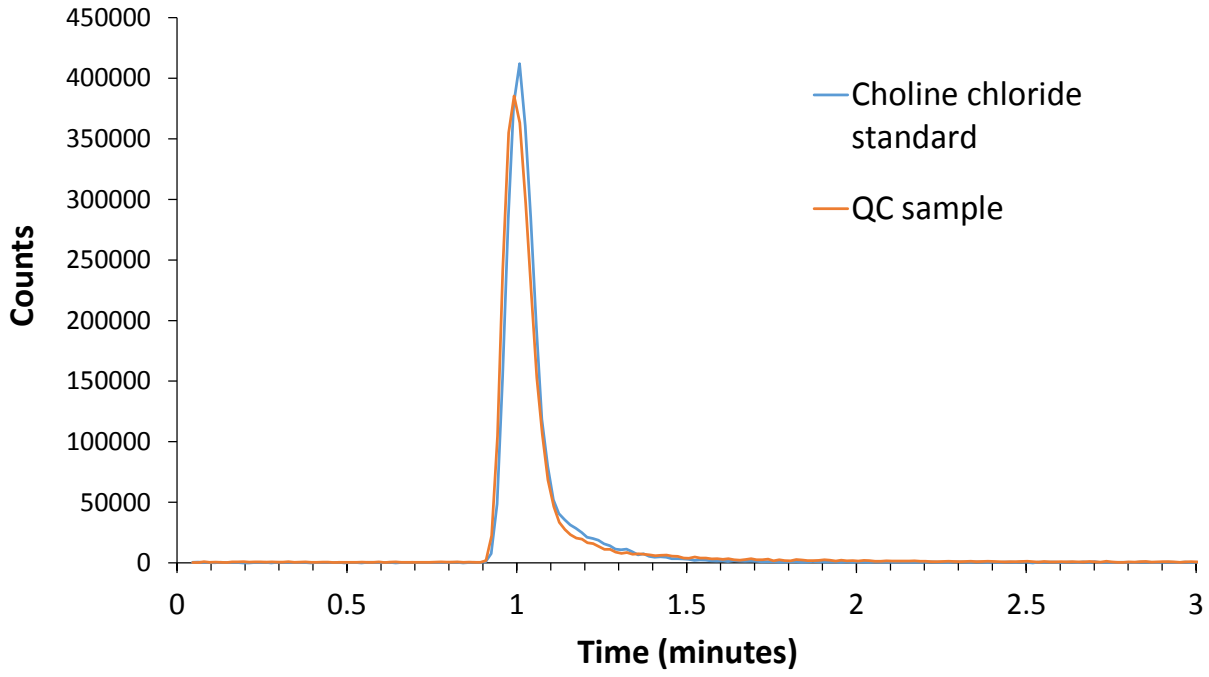


Figure 7.8: *Overlaid EICs of compounds with m/z 104.1000 for choline chloride standard and QC sample*

Figure 7.9 shows the mass spectra for the compounds in the QC sample and the choline chloride sample at three different collision energies.

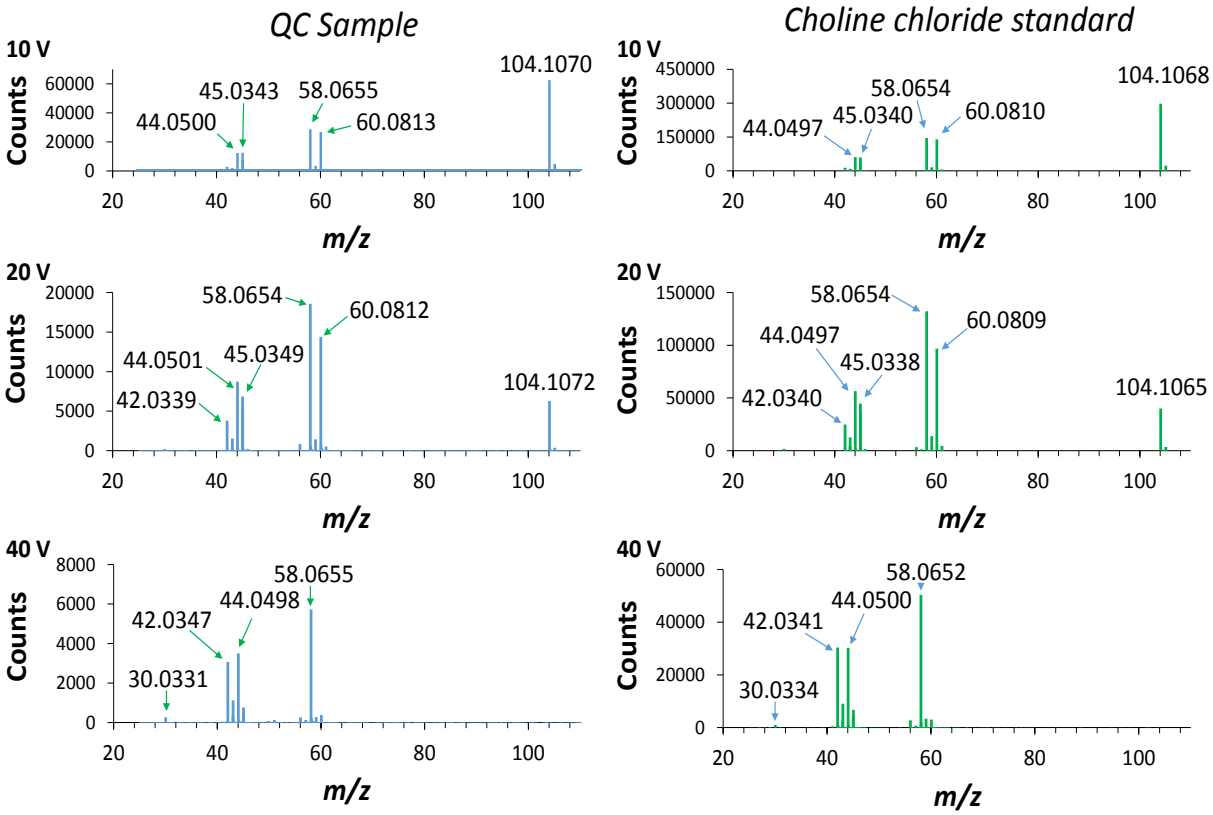


Figure 7.9: *Comparison of mass spectra at 10 V, 20 V, 40 V collision energy, between a QC sample and a choline chloride standard, for compound m/z 104.1000*

As can be seen in Figure 7.9, the mass spectra for the compound *m/z* 104.1000 in both samples are a very good match, with all fragment peaks matching at all three collision energies used. From this, it can be confidently confirmed that the compound *m/z* 104.1070 has been identified as choline.

7.4.1.6 Choline

Choline is a precursor to acetylcholine, a neurotransmitter [20][21], and is a component of various different phospholipids [21][22]. It exists in its free form mainly due to the catabolism of these phospholipids, particularly phosphatidylcholines [20][23], which are found in cell membranes [24] and are highly abundant in egg yolk [24][25]. Phosphatidylcholines (PCs) are catabolised by phospholipases (PLs), which catalyse the hydrolysis of certain bonds within the PC molecules. PLA_1 and PLA_2 catalyse the hydrolysis of either

the 1-acyl or 2-acyl groups respectively, breaking the ester bonds between the fatty acids and the glycerol backbone, then either PLB, or lysoPLA$_1$ or lysoPLA$_2$, catalyses the hydrolysis of the remaining acyl group, breaking the ester bond between the fatty acid and glycerol backbone of the resulting lysophosphatidylcholine molecule. PLC and PLD catalyse the hydrolysis of the phosphodiester bond; PLC acting on the glycerophosphate bond between the phosphate group and the glycerol backbone, and PLD acting on the choline phosphate ester bond between the phosphate group and choline, releasing a free choline molecule [20][26][27]. Figure 7.10 shows an example of this catabolic pathway.

It has been observed by Wang et al. that the PC content of egg yolk decreases over an increasing storage time of eggs as the PLs, which are endogenous to egg yolk, hydrolyse the PCs, breaking down the molecules [28]. This catabolism of the PC molecules results in the release of choline molecules, so as the PC content of an egg yolk decreases with increasing egg age, the choline content increases, which is what was observed in this study. Some of this increase in choline concentration could also be due to the degradation of PCs following egg lay. Similar results regarding the decreasing PC content of egg yolk were also observed in this study. Several compounds of the top 38 in bold in Table D.1 in Appendix D were recognised, through the use of Lipid Maps and METLIN, as lipids belonging to the phosphatidylcholine-choline metabolic pathway. Some of these 38 compounds were classified as potential PCs and they show a general trend of decreasing abundance with increasing egg age, as can be seen in Appendix D Figure D.3a, supporting the previous results. Some other of these 38 compounds were classified as potential lysophosphatidylcholines, which are intermediates in the catabolism of PCs to choline, and result from the hydrolysis of one of the fatty acyl groups in PC molecules by PLA$_1$ or PLA$_2$ enzymes. These were found to follow a general trend of increasing in abundance with increasing egg age, as can be seen in Appendix D Figure D.3b, which is what would be expected as the PC molecules are catabolised by PLs over time.

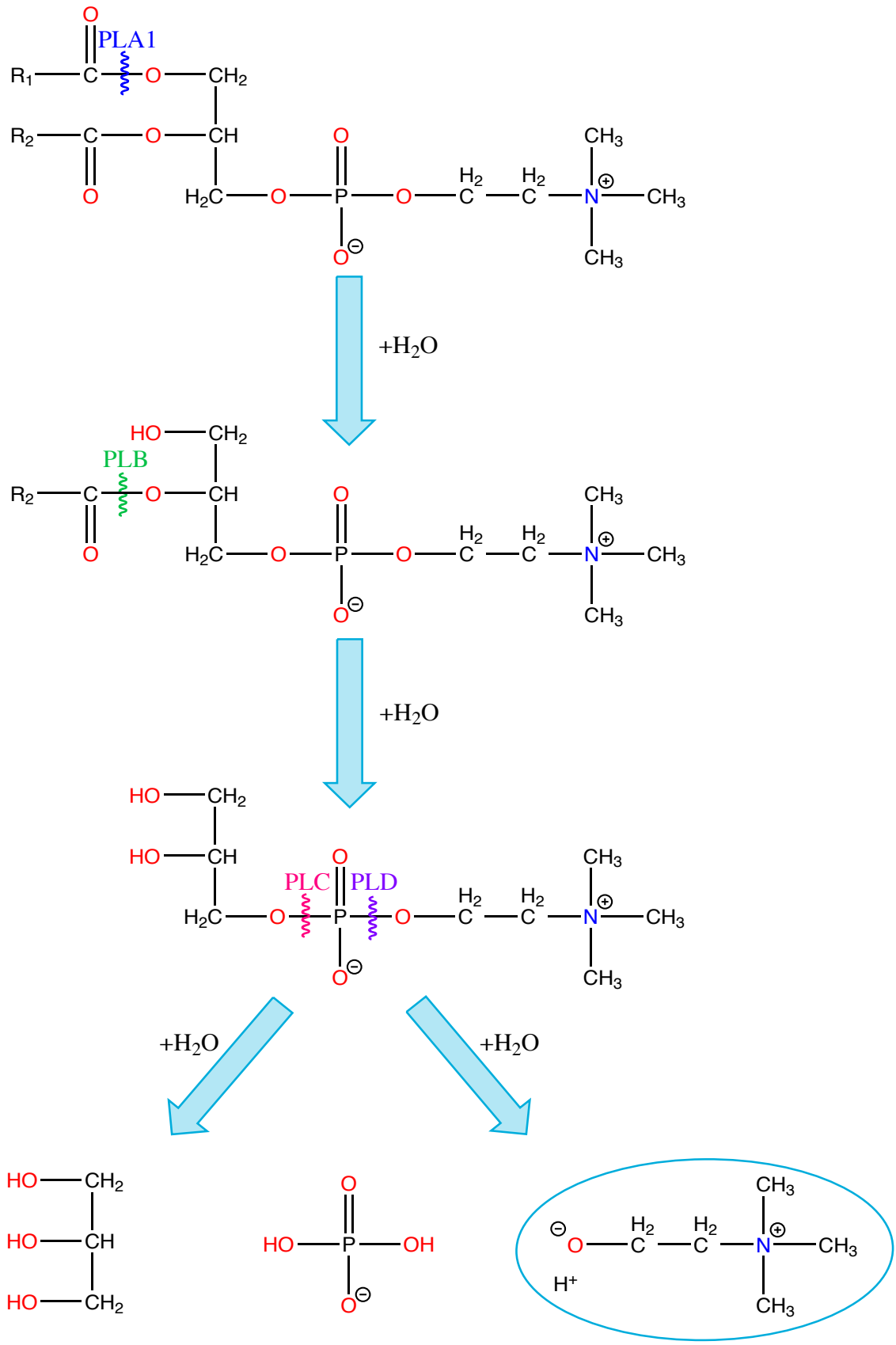


Figure 7.10: *Example catabolic pathway of the hydrolysis of phosphatidylcholine to choline (circled)*

7.4.2 Choline Targeted Study

As the compound identified as choline showed potential as a biomarker of egg age, due to the correlation between its abundance in egg yolk and egg age, a targeted study was carried out in order to obtain more quantitative information about the abundance of choline over a wider range of egg age.

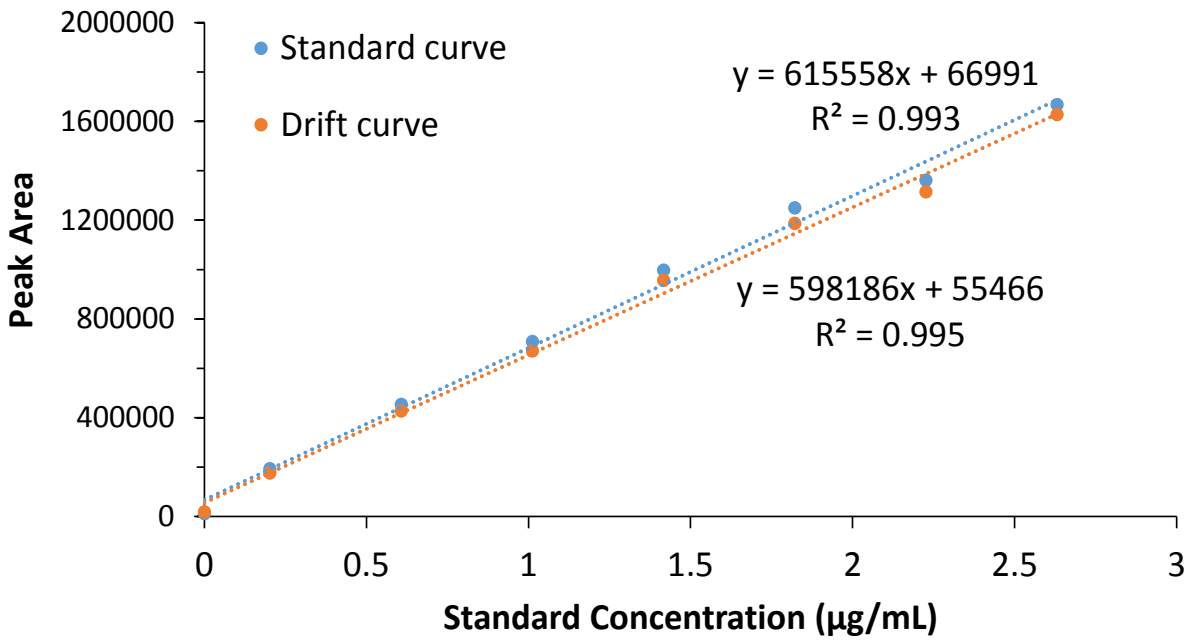


Figure 7.11: *Standard and drift calibration curves for choline chloride standards ranging from 0.2-2.6 µg/mL*

Figure 7.11 shows the standard and drift calibration curves that were produced based on the mean peak areas of the standards. The coefficient of determination value, R^2, for both calibration curves is greater than 0.99 (0.993 for the standard curve, and 0.995 for the drift curve), indicating that there is good linearity in the calibration, and that it is a good model for predicting choline concentration based on peak area. The Relative Error percentages (RE%s) for the standards were all found to be lower than 8% for the standard curve and lower than 14% for the drift curve, and the CV%s of the standards were all found to be lower than 6% for the standard curve and lower than 8% for the drift curve. These values are all within the acceptable limit of 15% according to the published

Food and Drug Administration (FDA) guidelines [29]. The CV% of the QC samples was also calculated to ensure that there was minimal instrumental drift affecting the analysis, and this was found to be 1.6%, which again is within accepted limits.

The concentrations of choline in the yolk organic extracts, and in the yolks themselves, can be seen in Appendix D, Table D.2. The concentration of choline in egg yolk ranges from an average of 6.8 µg/g in fresh egg yolk, up to an average of 28.7 µg/g in the yolks of eggs that were stored for twelve weeks. The p-values resulting from ANOVA were p=0.001 over the first five weeks of egg age, and p<0.001 for the full twelve weeks of age, confirming the results of the non-targeted study, that choline concentration is significantly different between eggs of different ages over five, as well as twelve weeks of age. Table D.3 in Appendix D shows the results of the post-hoc Tukey tests for pairwise comparisons of eggs of different ages.

There are significant differences in choline concentration between eggs that had been stored for different lengths of time, up to twelve weeks of storage, so if eggs are mislabelled regarding their sell-by and best before dates, it is likely that by observing their yolk choline concentrations it would be possible to uncover this fraudulence. The first point at which there is a significant difference between eggs of different ages is between fresh eggs and eggs that were three weeks old, with a p-value of 0.002, showing a strong significance. This could be useful because when eggs reach their sell-by dates at three weeks post-lay, supermarkets and other food businesses may be tempted to change the dates on the eggs or the packaging, making them appear to be fresh, in order to give them a longer shelf life and encourage consumers to purchase them. However, if it is possible to observe a significant difference in choline concentration between eggs of these two ages, then it could help to discourage and prevent this mislabelling fraud.

There was again a positive correlation between yolk choline concentration and increasing egg age, as can be seen in Figures 7.12 and 7.13.

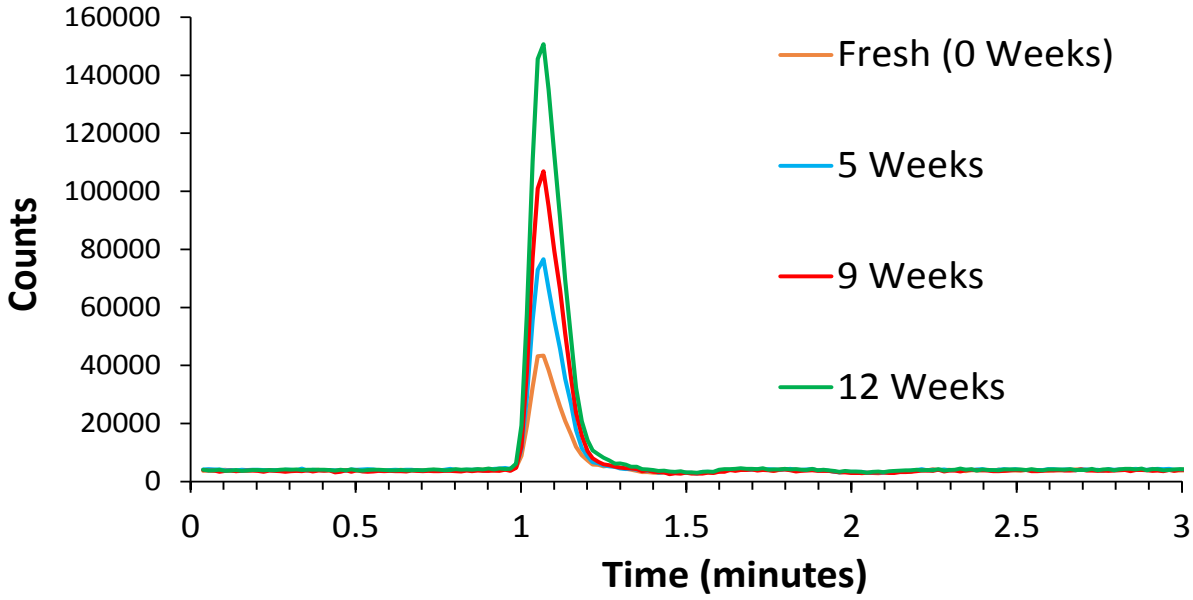


Figure 7.12: *Average EICs of choline $[M+H]^+$ at m/z 104.1 in yolk extracts for eggs stored for 0, 5, 9, and 12 weeks*

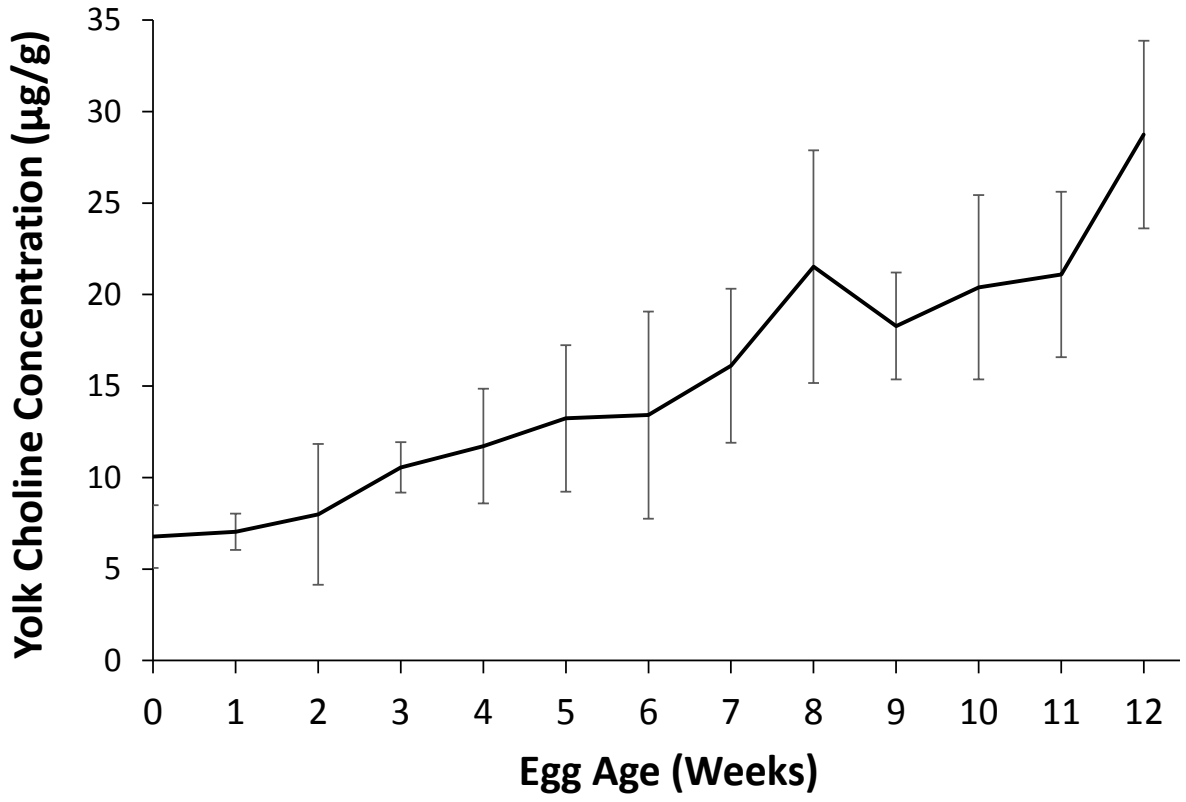


Figure 7.13: *Line graph showing the trend in changing abundance of choline in eggs ranging from fresh to twelve weeks old. n=6. Measured as mean values of concentration at each egg age, with error bars ±1 standard deviation.*

Figure 7.12 shows average EICs of the choline adduct $[M+H]^+$ at m/z 104.1 in yolk extracts from eggs of four different ages. The peak intensity steadily increases with egg age, indicating an increasing choline concentration, and therefore a positive correlation between yolk choline concentration and egg age. However, Figure 7.13 gives a more complete picture of the changing concentration of choline. It shows an increasing concentration in yolk choline content over a range of twelve weeks of egg age, however it also shows a high increase in concentration, and in concentration variance, at eight weeks of egg age before it decreases again at nine weeks and then continues its gradual increase up to twelve weeks. As can be seen in Table D.2 in Appendix D, two of the six eggs that underwent extraction at eight weeks post-lay show a much higher concentration of choline in the yolk than the other four, which has caused this anomaly in the trend. If larger sample sizes were used, then it would be interesting to determine whether or not these two eggs were outliers. Larger sample sizes would also reduce the standard deviations of choline concentrations at each egg age, and thus the error bars in Figure 7.13, which represent one standard deviation either side of the mean, would be reduced.

7.5 Eggs Stored at 5 °C for 5 weeks: Yolk Organic Extracts

Yolk extracts of eggs that were stored at 5 °C for different lengths of time were compared to see whether there were differences in their metabolite profiles.

7.5.1 Results and Discussion

7.5.1.1 Quality Control Analysis

Table 7.6 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 7.6: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic samples of eggs stored at 5 ℃*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	9638442	616287074	340689505	37992095	1687641898	480764040
QC2	9879820	628814511	352591958	42475407	1715105714	502104802
QC3	9459914	612521438	332773895	42095952	1711439108	478354196
QC4	11482568	621028757	334581359	47740219	1695721497	499163567
QC5	10666069	621413929	332171860	48679965	1633021165	502140203
QC6	12019996	630533320	335246585	48339357	1638561447	498680541
QC7	10409730	623398566	317973847	40543072	1510147168	454216582
SD	959147	6389817	10366421	4256638	72170108	17908390
Mean	10508077	621999656	335147001	43980866	1655948285	487917704
CV%	9.13	1.03	3.09	9.68	4.36	3.67

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 7.7 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 7.7: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 5 ℃*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.683	10.258	32.457	56.946	70.899	77.609
QC2	4.482	10.073	32.339	56.794	70.599	77.136
QC3	4.414	10.022	32.266	56.743	70.447	76.802
QC4	4.452	10.034	32.254	56.714	70.469	76.840
QC5	4.501	9.976	32.225	56.697	70.402	76.723
QC6	4.476	9.972	32.169	56.641	70.282	76.553
QC7	4.445	10.020	32.269	56.774	70.645	77.232
SD	0.088	0.098	0.092	0.097	0.201	0.361
Mean	4.493	10.051	32.283	56.758	70.535	76.985
CV%	1.97	0.97	0.29	0.17	0.29	0.47

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

7.5.1.2 Metabolite Profiling

The yolk extracts of eggs that were stored for different lengths of time were compared initially by observing the TICs of the extracts following HPLC-MS analysis. Figure 7.14 shows overlaid TICs for the average chromatograms of yolk extracts from fresh eggs and eggs that were five weeks old.

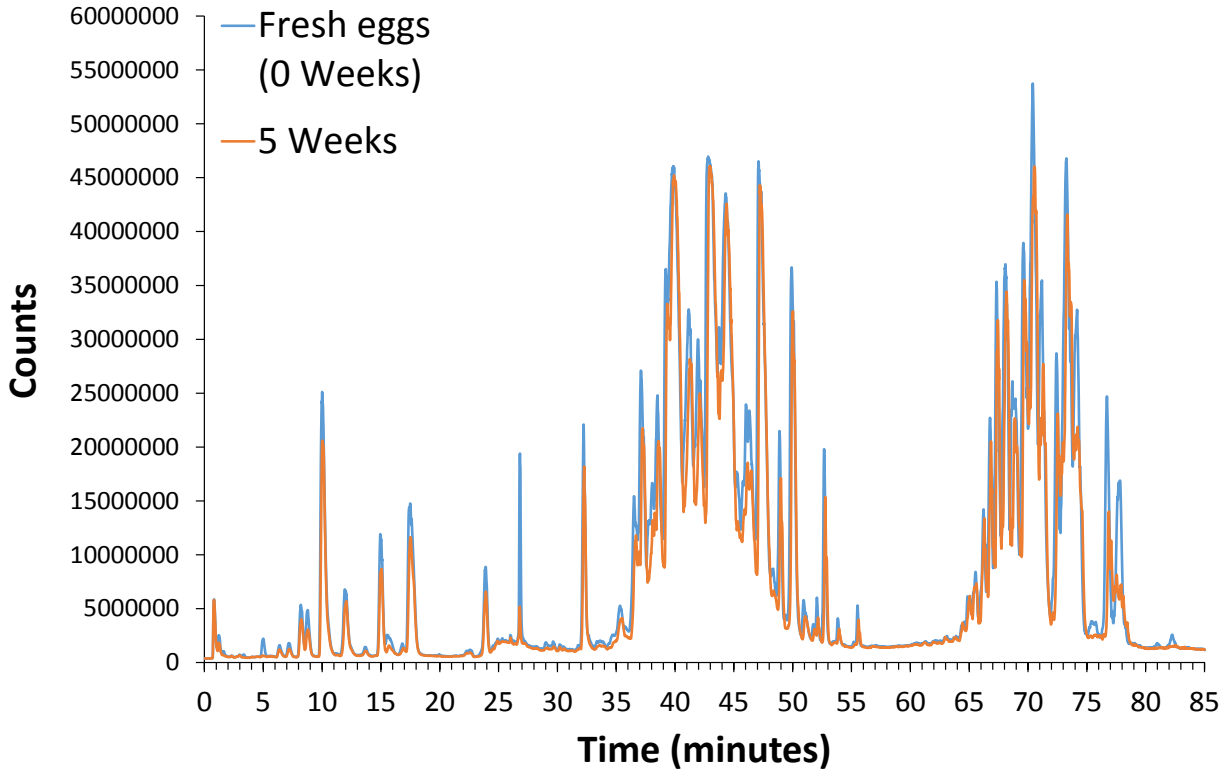


Figure 7.14: *Overlaid average TICs of yolk organic extracts from fresh eggs and eggs that were five weeks old, stored at 5 ℃*

There is a visible difference in the average TICs of yolk extracts from fresh eggs and eggs that were five weeks old; the peak intensities for most peaks are greater in the TIC for fresh eggs. This is comparable to what was observed when comparing average TICs for fresh and five week old egg yolk extracts, when eggs were stored at 23 ℃.

The peak shape towards the end of the TIC of the yolks of eggs that had been refrigerated for five weeks is different compared to other TICs throughout this research. This is because there was some retention time drift that occurred throughout the analytical sequence, that affected the later peaks in the chromatogram. This resulted in a poor

peak shape for these peaks in the average TIC that was produced. Although there was some retention time drift, the analysis was still robust as reported in Section 7.5.1.1.

7.5.1.3 Multivariate Statistics

Following PCA, scores plots were produced to display any variation between the yolk extracts of eggs of different ages.

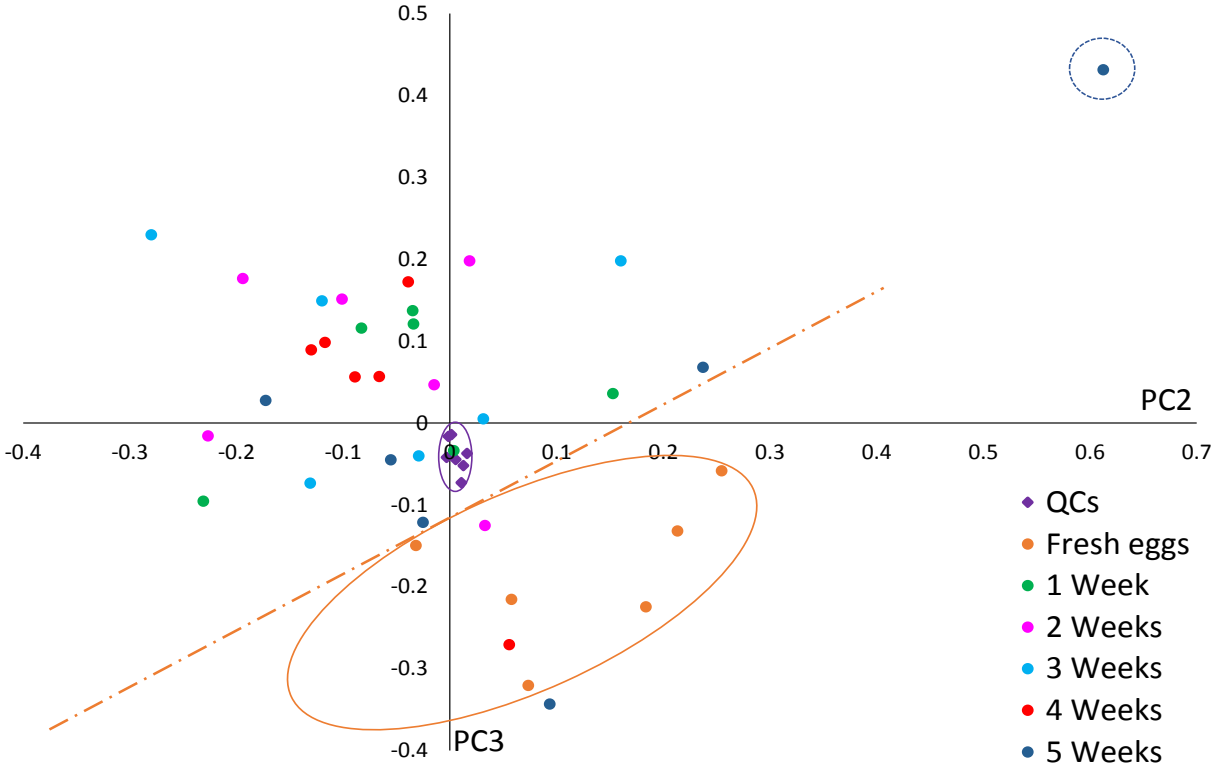


Figure 7.15: *PCA scores plot showing PC2 vs PC3 for eggs up to five weeks old that were stored at 5 ℃, including all compounds with CV%<30%. PC2 explains 0.37% of the variance, and PC3 explains 0.15% of the variance. Circled: sample separate to the rest of sample set.*

Although the TICs in Figure 7.14 show a clear difference between fresh eggs and eggs that were five weeks old, the scores plot in Figure 7.15 does not show the same observation; there appears to be little difference between the different ages of eggs. This could be because the more highly abundant compounds that contribute to the differences between the TICs of fresh egg yolk and the yolk of eggs that had been refrigerated for five weeks, may not contribute to as much of the variance explained by PC2 and PC3 as

the less abundant compounds, which may experience more variance between the yolks of random egg samples, than between those of eggs of different ages.

Although there is little variation between eggs of different ages on the scores plot in Figure 7.14, there is some variation between fresh eggs and older eggs. Fresh eggs form a group that is almost separate to any other samples on the plot; there is complete separation between fresh eggs and eggs that were one and two weeks old (indicated by the orange dashed line on the plot), and only one sample from eggs that were three, four and five weeks old that overlap into this group. This indicates that most of the metabolic change occurs during the first week post-lay. The QC samples are clustered tightly together in the centre of the plot, indicating that there was little instrumental drift occurring throughout the analysis, and thus that the analysis was robust.

There is one sample from the group of eggs that were five weeks old that is very separate to all other samples and is circled on the plot. It was observed prior to metabolite extraction that this egg was very small compared to the other eggs, had a white speckled shell, and had very little liquid egg inside the shell. This explains why this egg shows so much variation compared to the other eggs.

As the scores plot shows little variation between eggs of different ages, which is highlighted by the fact that PC2 describes just 0.37% of the variation and PC3 describes 0.15%, it proves that the differences between eggs that have been refrigerated for different lengths of time are extremely subtle. Although the age, breed and housing condition of the laying birds were kept the same for all eggs, there will have been a natural variation between the birds themselves, resulting in a variation between their eggs that could not be controlled. As the differences due to egg age are so small, most of the variation between samples on the scores plot is likely to be due to this natural variation.

7.5.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games Howell tests on the top 50 compounds remaining, following the removal of any duplicates, adducts and isotopes from the top 100 based on PC1 loadings, 48 compounds were found to show significant differences between eggs of different ages. These compounds can be seen in Table D.4, Appendix D; those in bold were also found to be statistically significant when confirmed using the raw data. Attempts were then made to identify these compounds.

Of these 41 compounds that were still found to be statistically significant following confirmation using the raw data, four were putatively identified through comparing mass spectra provided by METLIN, with spectra obtained from HPLC-MS analysis in this study. The observed *m/z* values for all four compounds were due to the $[M+H]^+$ adduct. The compounds *m/z* 331.2857, 338.3435, 454.2945, and 496.3420 were putatively identified as 1-monopalmitin, docosenamide (erucamide), 1-palmitoyl-glycero-3-phosphatidylethanolamine (PE(16:0/0:0)), and 1-palmitoyl-glycero-3-phosphatidylcholine (PC(16:0/0:0)), respectively. The comparisons between the mass spectra provided by METLIN and the mass spectra resulting from the analysis for these compounds can be seen in Appendix D, Figure D.2 for the compound *m/z* 331.2857 which was previously putatively identified in Section 7.4, and Appendix D, Figures D.4, D.5, and D.6 for the compounds *m/z* 338.3435, 454.2945, and 496.3420 respectively.

The 48 statistically significant compounds in Table D.4 were compared with the 44 compounds in Table D.1 in Appendix D, for eggs that were stored at 23 ℃, and several compounds were found to be statistically significant in both studies. One of these compounds was that which was putatively identified in this study as erucamide. This was not identified in the previous study, as the likelihood score for the molecular formula corresponding to erucamide was below the 95 cut off.

Tables 7.8-7.11, and Figures 7.16-7.19, show the putative identifications of the four

compounds, the p-values resulting from ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests, as well as the trends in the changing abundance of the compounds in the yolks of eggs ranging from fresh to five weeks old.

Table 7.8: *Table showing the putative identification of compound m/z 331.2857 and the p-values resulting from the Welch test and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**Welch *p*-value**
331.2857	$C_{19}H_{38}O_4$	95.77	1-monopalmitin	<0.001

Post-hoc test *p*-values

	Fresh	**1 Week**	**2 Weeks**	**3 Weeks**	**4 Weeks**	**5 Weeks**
Fresh		0.037	0.13	<0.001	0.818	0.004
1 Week			0.037	<0.001	0.002	0.001
2 Weeks				0.988	0.198	1
3 Weeks					<0.001	0.840
4 Weeks						0.009
5 Weeks						

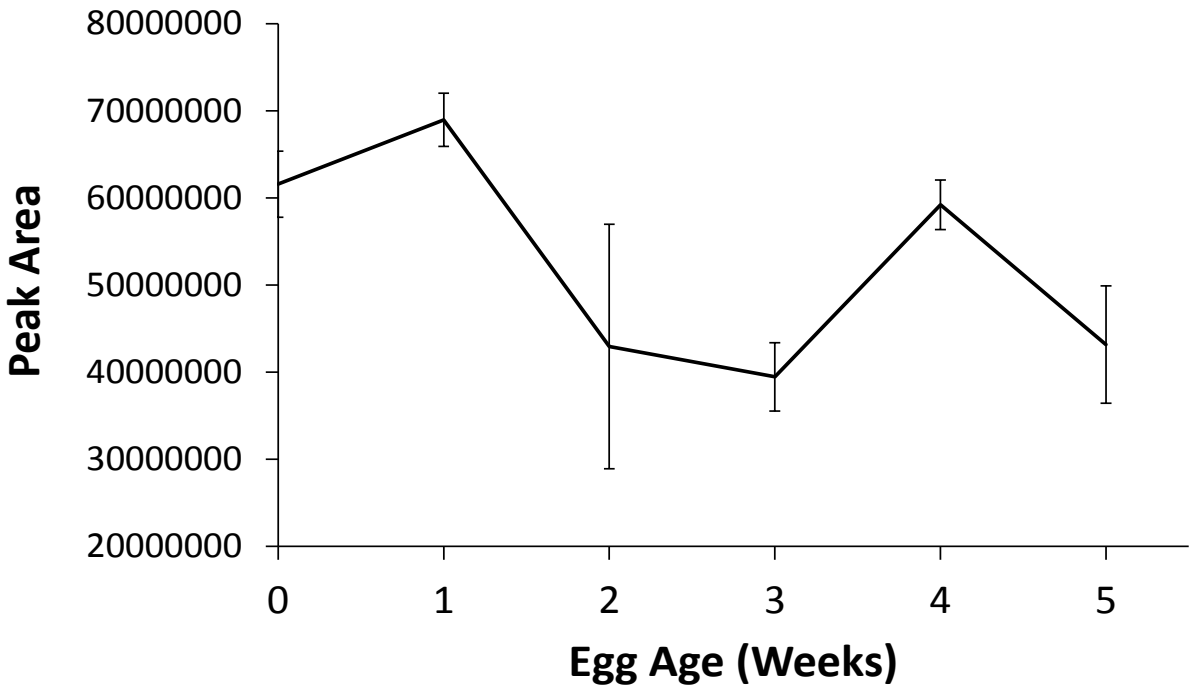


Figure 7.16: *Line graph showing the trend in changing abundance of compound m/z 331.2857 across eggs of six different ages. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Table 7.9: *Table showing the putative identification of compound m/z 338.3435 and the p-values resulting from the Welch test and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	Welch *p*-value
338.3435	$C_{22}H_{43}NO$	99.41	Erucamide	<0.001

Post-hoc test *p*-values						
	Fresh	**1 Week**	**2 Weeks**	**3 Weeks**	**4 Weeks**	**5 Weeks**
Fresh		0.017	<0.001	<0.001	<0.001	<0.001
1 Week			0.016	0.077	0.103	0.067
2 Weeks				0.02	0.005	0.061
3 Weeks					0.991	1
4 Weeks						0.969
5 Weeks						

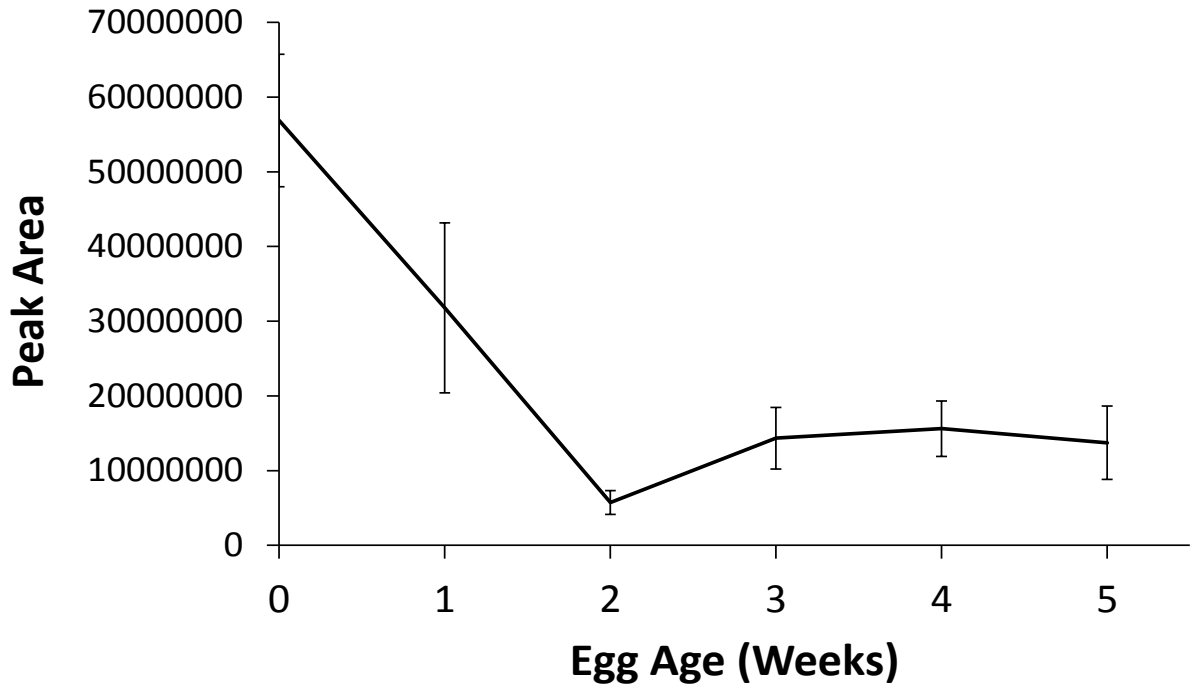


Figure 7.17: *Line graph showing the trend in changing abundance of compound m/z 338.3435 across eggs of six different ages. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Table 7.10: *Table showing the putative identification of compound m/z 454.2945 and the p-values resulting from the ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
454.2945	$C_{21}H_{44}NO_7P$	95.09	PE(16:0/0:0)	0.003

Post-hoc test *p*-values

	Fresh	1 Week	2 Weeks	3 Weeks	4 Weeks	5 Weeks
Fresh		0.002	0.022	0.037	0.251	0.022
1 Week			0.914	0.827	0.292	0.916
2 Weeks				1	0.858	1
3 Weeks					0.935	1
4 Weeks						0.855
5 Weeks						

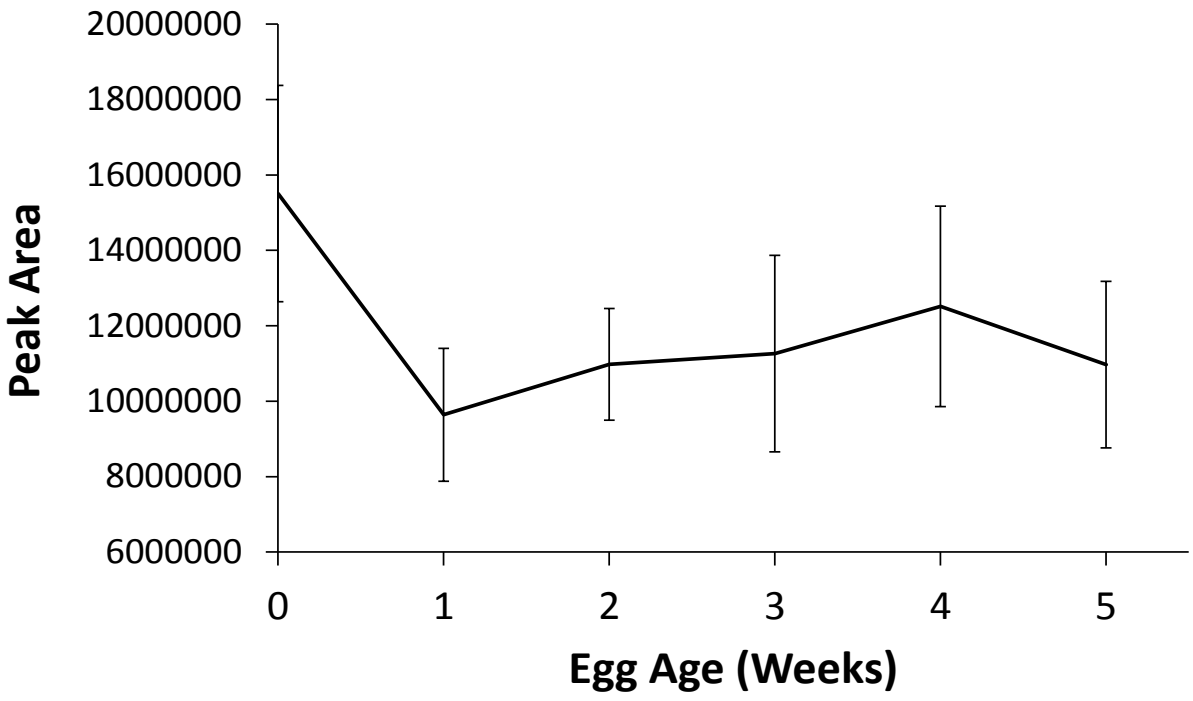


Figure 7.18: *Line graph showing the trend in changing abundance of compound m/z 454.2945 across eggs of six different ages. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Table 7.11: *Table showing the putative identification of compound m/z 496.3420 and the p-values resulting from the ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
496.3420	$C_{24}H_{50}NO_7P$	95.40	PC(16:0/0:0)	0.006

Post-hoc test *p*-values

	Fresh	1 Week	2 Weeks	3 Weeks	4 Weeks	5 Weeks
Fresh		0.004	0.019	0.043	0.133	0.372
1 Week			0.987	0.918	0.656	0.308
2 Weeks				0.999	0.949	0.685
3 Weeks					0.994	0.867
4 Weeks						0.991
5 Weeks						

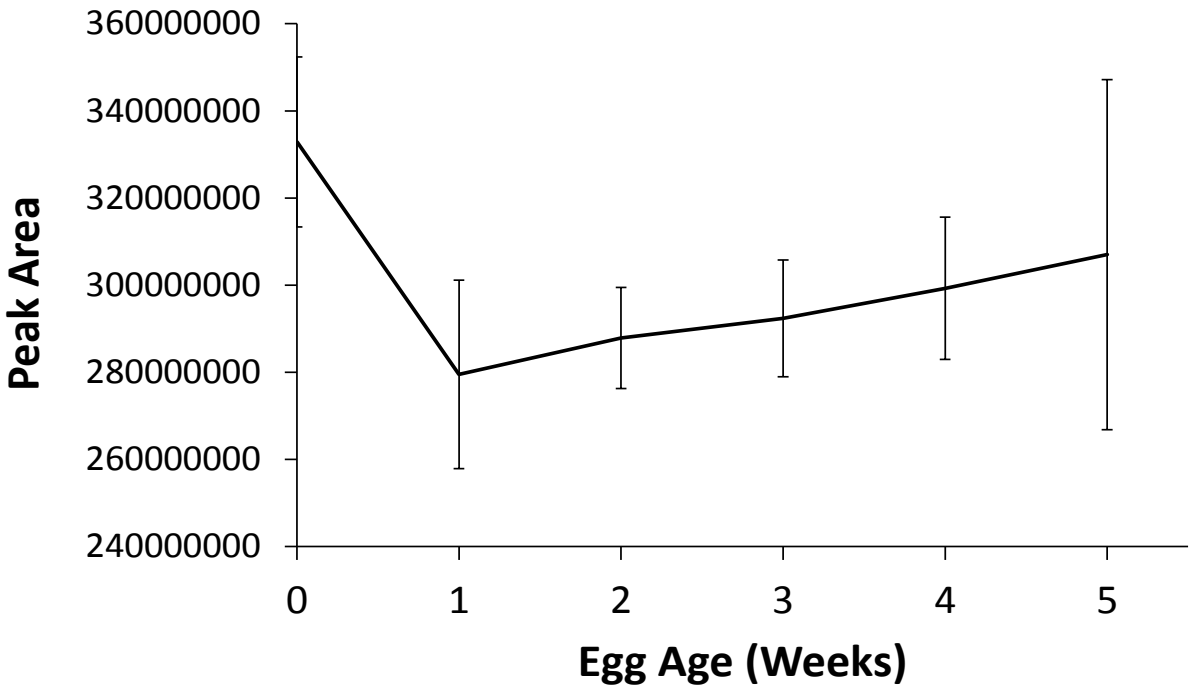


Figure 7.19: *Line graph showing the trend in changing abundance of compound m/z 496.3420 across eggs of six different ages. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

The probability scores for the predicted molecular formulae for all of these compounds are above 95, indicating that there is a strong likelihood that the predicted formulae correspond to the compounds of interest. The p-values from ANOVA and Welch tests show that the abundances of these compounds are statistically significant when comparing them between eggs of different ages, ranging from fresh to five weeks old, and the post-hoc test p-values mostly corroborate this.

The compound that was putatively identified as 1-monopalmitin shows several points of high statistical significance between eggs of different ages, as shown by the Games-Howell test p-values in Table 7.8. However, it is not suitable as a marker of egg age, as there is no correlation between its abundance in egg yolk and egg age, which would make it difficult to accurately predict the age of an egg. The reason for this erratic change of compound abundance is likely to be due to several different metabolic pathways resulting in the production and breakdown of this compound; different reactions involved in the various pathways may occur at varying rates, meaning that the abundance can change significantly from week to week, without following a continuous trend.

The compound that was putatively identified as docosenamide, or erucamide, an amide derivative of the fatty acid erucic acid, also shows several points of high statistical significance as shown by the post-hoc test p-values in Table 7.9. However, most of these points of significant difference are between fresh eggs and eggs that have been stored at 5 °C for any length of time, indicating that there may have been a large change in compound abundance during the first week of egg storage. This is corroborated by the line graph in Figure 7.17, which shows a large decrease in compound abundance over the first two weeks of egg storage, followed by little change thereafter. Although the changing abundance of this compound is statistically significant, it is not suitable as a biomarker of egg age, as there is little significant difference in abundance after the first two weeks post-lay. It may be possible to use this compound as a marker to discriminate between fresh eggs and eggs that are one week old or more, but nothing more than this.

The two compounds that were putatively identified as PE(16:0/0:0) and PC(16:0/0:0) both show similar points and values of statistical significance as shown by the post-hoc test p-values in Tables 7.10 and 7.11. This suggests that they may also show similar trends in changing abundance, which is confirmed by the line graphs in Figures 7.18 and 7.19. Both of these compounds experience a significant decrease in abundance over the first week post-lay, followed by very little change in abundance afterwards. This is likely to be due to the fact that these compounds are very similar to each other. PE(16:0/0:0) is a lysophosphatidylethanolamine with a palmitic acid group attached to the first carbon of the glycerol backbone, whilst PC(16:0/0:0) is a lysophosphatidylcholine with a palmitic acid group again attached to the first carbon of the glycerol backbone. These compounds both follow very similar metabolic pathways in their catabolism to glycerol-3-phosphate [30], and so the rate of catabolism of these compounds in egg yolk may be very similar, resulting in the changes in compound abundance with increasing egg age following very similar trends.

However, neither of these compounds are suitable as markers of egg age, as the one point of high statistical significance is between fresh eggs and eggs that were one week old, with little change and no correlation of abundance over a further increase of egg age. A slight increase in abundance of these compounds, following the initial decrease, can be observed in the graphs in Figures 7.18 and 7.19. This trend could be due to the initial catabolism and/or degradation of the PE(16:0/0:0) and PC(16:0/0:0) molecules reducing their abundance, followed by the catabolism and/or degradation of other compounds, e.g. phosphatidylethanolamines (PEs) and PCs, resulting in the production of more PE(16:0/0:0) and PC(16:0/0:0) molecules, thereby increasing their abundance over the following weeks. Although PE(16:0/0:0) experienced a decrease in abundance in egg yolk at five weeks post-lay, it would be interesting to observe the trend of these compounds over a wider range of egg age to see at which point they both stop increasing in abundance and begin to significantly decrease. This would be the point at which the rate of catabolism and/or degradation of the PE(16:0/0:0) and PC(16:0/0:0) molecules is greater than the

rate of catabolism and/or degradation of PEs and PCs, and therefore the production of the PE(16:0/0:0) and PC(16:0/0:0) molecules.

Some of the standard deviation error bars on these line graphs are quite wide. Whilst some only appear large due to the scaling of the y-axis, others are wide due to a large standard deviation, representing a high amount of variation in compound abundance at those time points. This is due to the natural variation between the eggs; although conditions were kept as uniform as possible, there will have been biological variation between the laying birds, and therefore also between their eggs.

7.5.1.5 Choline in refrigerated eggs: non-targeted

As choline was found to have potential as a biomarker of egg age when eggs are stored at 23 °C, the abundance of choline in the yolks of eggs stored for varying lengths of time at 5 °C was then studied, as choline did not appear in the statistically significant top 100 compounds based on PC1 loadings. Table 7.12, shows the p-value from ANOVA and the p-values from the post hoc Tukey tests, when comparing choline between the yolks of eggs that were stored at 5 °C for five weeks.

As can be seen from Table 7.12, choline showed no significant difference in abundance between eggs of different ages when stored at 5 °C, and thus is not suitable as a biomarker of egg age, over a five week age range, when eggs are refrigerated.

The graph in Figure 7.20 also shows that there is very little change in the abundance of choline in egg yolks when eggs are stored at 5 °C for five weeks. There is no correlation between abundance and egg age, which is very different to the increasing abundance of choline that was observed in the yolks of eggs that were stored at 23 °C for increasing lengths of time up to five weeks.

Table 7.12: *Table showing the p-values resulting from the ANOVA and Tukey tests for choline in the yolks of eggs stored at 5 °C for five weeks.*

ANOVA *p*-value	Post-hoc test *p*-values					
0.151	**Fresh**	**1 Week**	**2 Weeks**	**3 Weeks**	**4 Weeks**	**5 Weeks**
Fresh		1	0.999	0.936	0.756	0.775
1 Week			1	0.980	0.621	0.643
2 Weeks				0.994	0.514	0.536
3 Weeks					0.235	0.25
4 Weeks						1
5 Weeks						

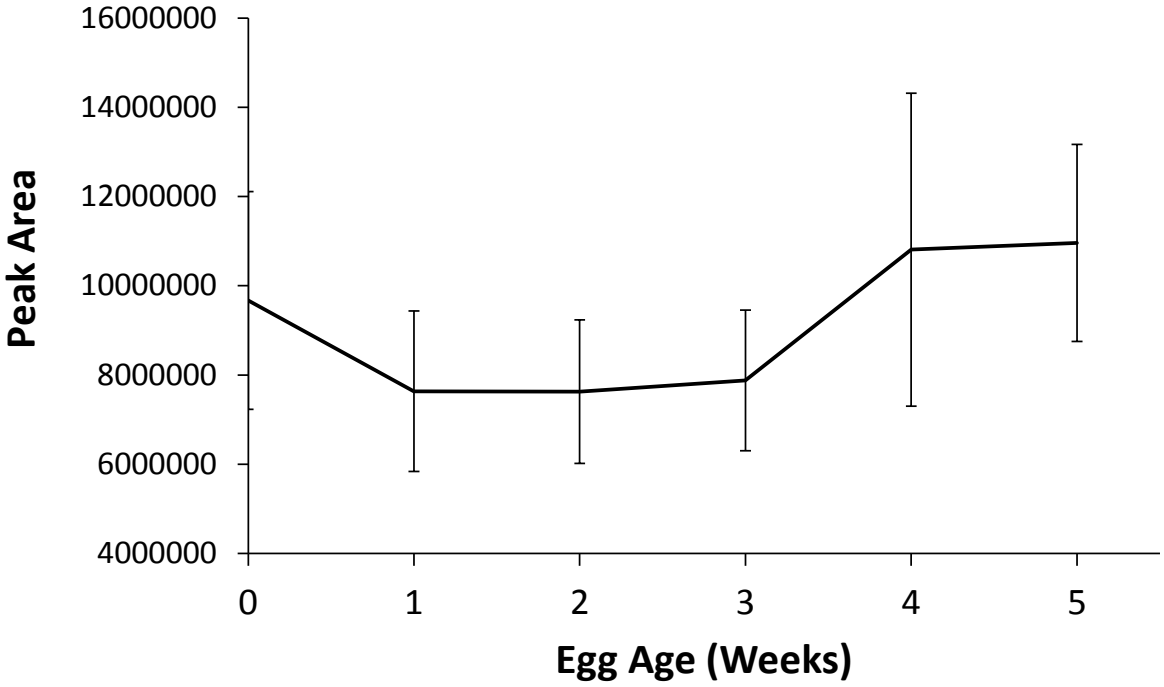


Figure 7.20: *Line graph showing the trend in changing abundance of choline across eggs of six different ages for eggs stored at 5 °C for five weeks. n=6. Measured as mean values of peak areas at each egg age, with error bars ±1 standard deviation.*

The lack of statistically significant increase in the abundance of choline in yolk between eggs of different ages indicates that storing eggs at lower temperatures slows down the metabolic processes, e.g. the catabolism of PCs, and any metabolite degradation, so that the production of choline is much slower and a trend in increasing choline abundance cannot be seen. This prevents the accurate prediction of egg age through yolk choline concentration.

7.5.2 Choline Targeted Study

A choline targeted study was carried out in order to gain more quantitative information about how the abundance of choline changes in the yolks of eggs over a longer period of time. It was previously observed in Section 7.4.2 that the change in choline abundance for eggs stored at 23 °C was statistically significant over twelve weeks of egg age, with the abundance increasing gradually over the whole twelve weeks. As there was no significant change in choline abundance in the yolks of eggs that were stored at 5 °C for five weeks, and no obvious trend in the minimally changing abundance, it was not expected that there would be much significance in the changing abundance of choline over twelve weeks. However, it was possible that a statistically significant trend may have developed when the abundance was measured over a longer period of time.

Figure 7.21 shows the standard and drift calibration curves that were produced based on the mean peak areas of the standards. The coefficient of determination value, R^2, for both calibration curves is greater than 0.99 (0.997 for the standard curve, and 0.999 for the drift curve), indicating that the calibration has good linearity and thus is a good model for calculating choline concentration based on the peak areas of the standard EICs. The Relative Error percentages (RE%s) for the standards were all found to be lower than 7% for the standard curve and lower than 15% for the drift curve, and the CV%s of the standards were all found to be lower than 13% for the standard curve and lower than 14% for the drift curve. These values are all within the acceptable limit of 15% according to the published Food and Drug Administration (FDA) guidelines [29]. The CV% of the QC samples was also calculated to ensure that there was minimal instrumental drift affecting the analysis, and this was found to be 3.6%, which again is within accepted limits.

The concentrations of choline in the yolk organic extracts, and in the yolks themselves, can be seen in Table D.5 in Appendix D. Over the first five weeks of egg storage, there is a statistical significance of p=0.017 for the changing concentration of choline. However, the Tukey tests show that the only point of significant difference in choline concentration

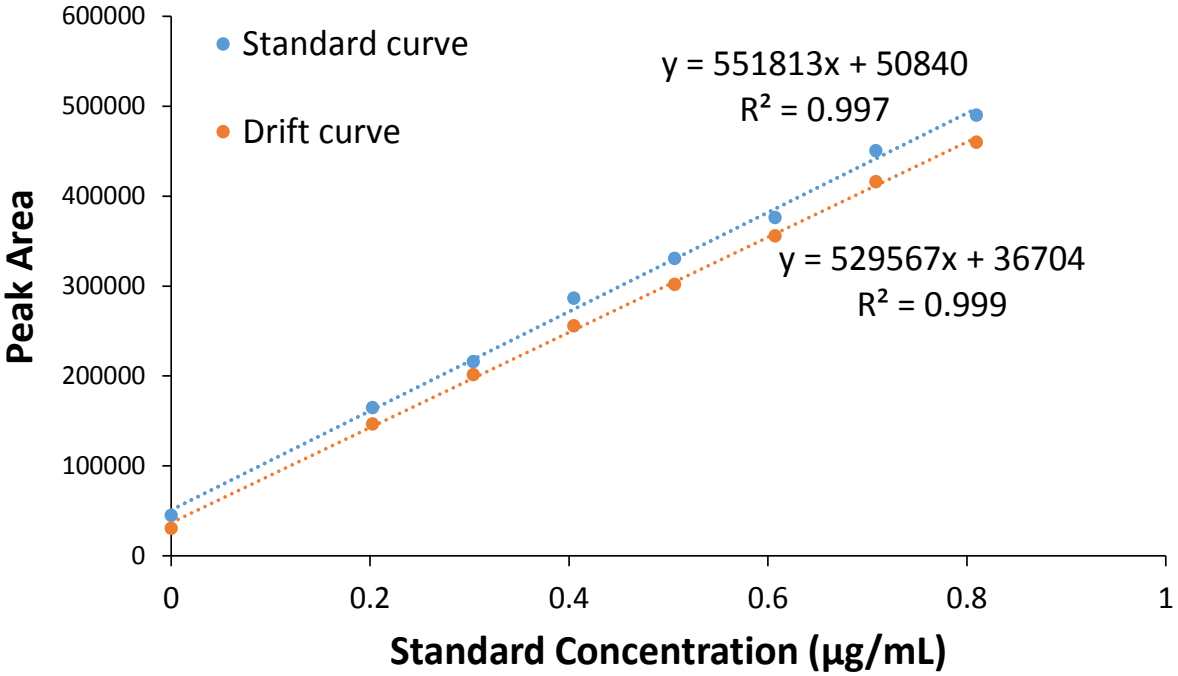


Figure 7.21: *Standard and drift calibration curves for choline chloride standards ranging from 0.2-0.8 µg/mL*

is between eggs of one and five weeks of age, with p=0.01.

The graph in Figure 7.22 shows the changing concentration of choline in the yolks of eggs up to five weeks old, which is quite different to the trend that was observed for choline abundance in the non-targeted study, which can be seen in Figure 7.20. This highlights the fact that there is no correlation between yolk choline concentration and egg age, when eggs have been stored at 5 °C. The choline concentration decreases over the first week post-lay, resulting in the lowest concentration of choline at this time point, before increasing over the following two weeks, decreasing slightly at four weeks, then increasing again at five weeks, resulting in the highest concentration of choline over the first five weeks post-lay. This explains why the only significant difference in choline concentration is between one week old eggs and eggs that were five weeks old. The standard deviations are again quite large, resulting in wide error bars, indicating a wide variance in the concentrations of choline at each time point.

Over the full twelve weeks of egg age that were studied, the ANOVA shows a high

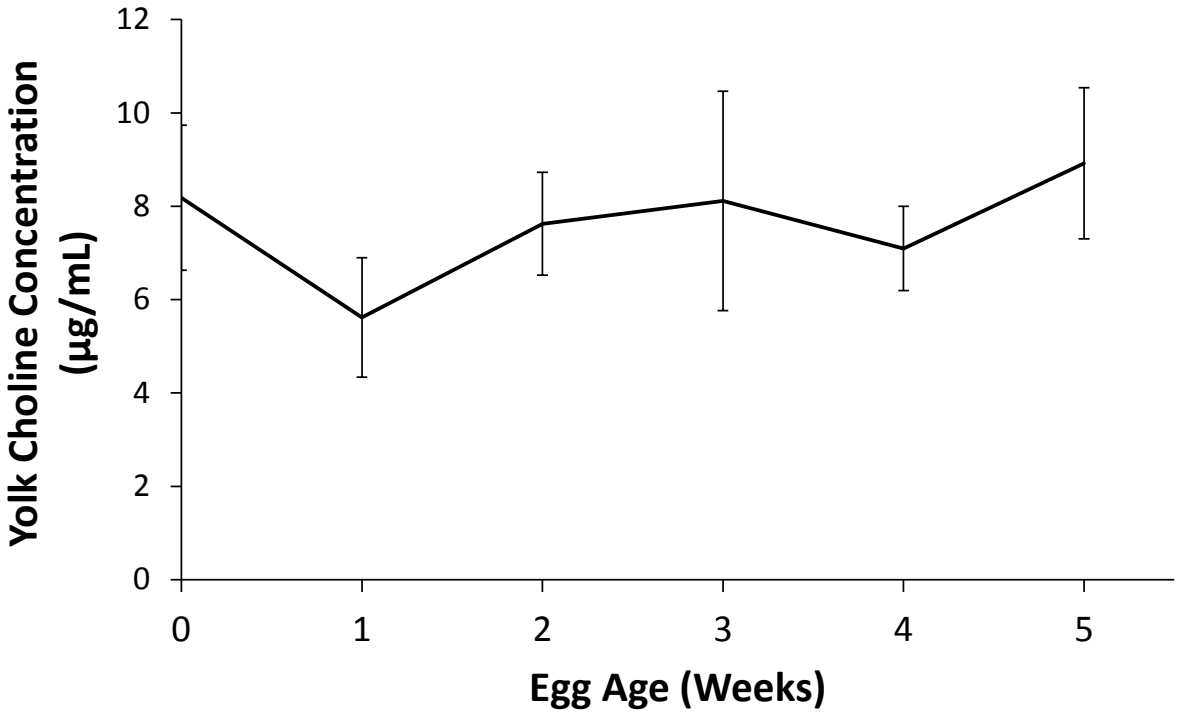


Figure 7.22: *Line graph showing the trend in changing yolk choline concentration across eggs of different ages, which were stored at 5 °C for up to five weeks. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

significance of p=0.001 for the changing concentration of choline in egg yolk. However, the post-hoc Tukey tests show that the only points of significant difference in choline concentration are between eggs that were one week old, and eggs that were six, eight, ten and twelve weeks old. This seems to indicate that even over a longer period of egg age, choline is not useful as a biomarker of age when eggs are stored at 5 °C.

The graph in Figure 7.23 shows the lack of correlation between yolk choline concentration and egg age, over twelve weeks of egg storage. The change in concentration is mainly insignificant, and is highly erratic, with the concentration changing randomly over the weeks. The concentration decreases during the first week post-lay, resulting in the lowest concentration observed in this study, then increases the following week. This explains why only one week old eggs show any significant differences in choline concentration compared to eggs of other ages. The error bars are again quite wide due to large standard

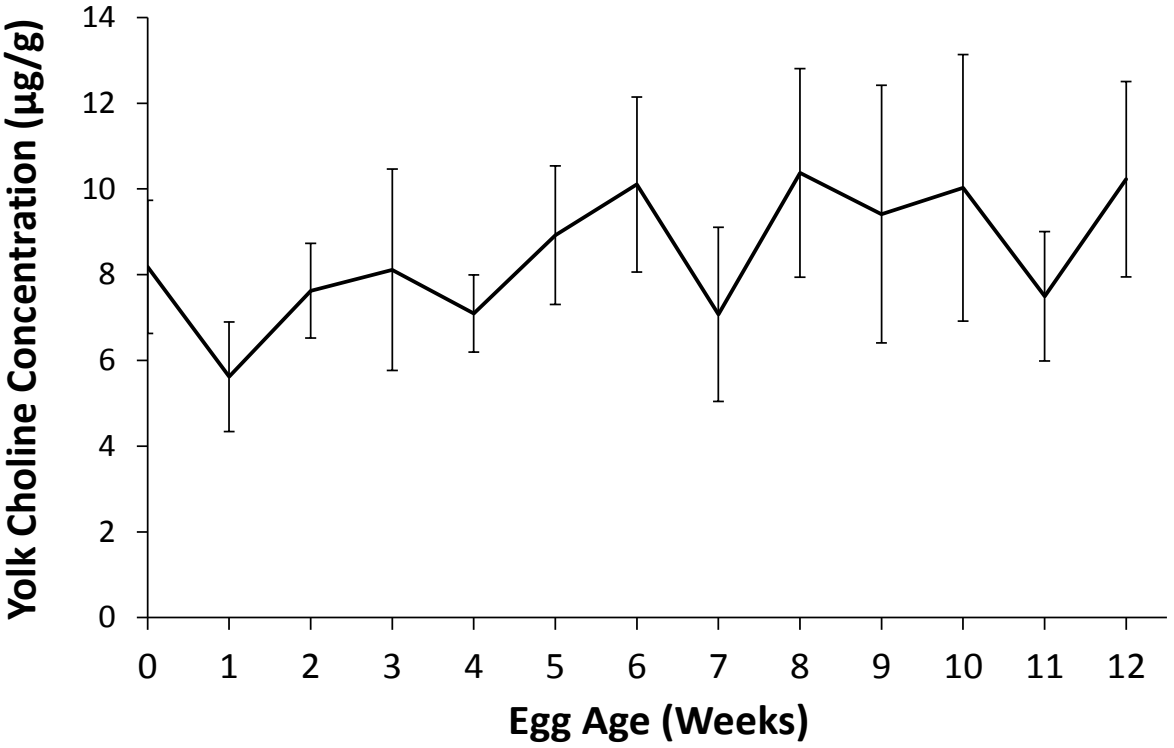


Figure 7.23: *Line graph showing the trend in changing yolk choline concentration across eggs of different ages, which were stored at 5 °C for up to twelve weeks. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

deviations, indicating a wide variation of yolk choline concentration in eggs of each age.

This targeted study has confirmed what was concluded from the non-targeted study for eggs that were stored at 5 °C for five weeks. Storing eggs at lower temperatures decreases the rate of the catabolism and potential degradation of PCs, and therefore the production of choline, resulting in no significant increase in choline abundance, preventing the accurate prediction of egg age from yolk choline concentration.

7.6 Comparison of Eggs Stored at 23 °C and 5 °C for 5 weeks: Yolk Organic Extracts

Yolk extracts from eggs that were stored at 23 °C and 5 °C were compared to see whether there were any differences in their metabolite profiles.

7.6.1 Results and Discussion

7.6.1.1 Quality Control Analysis

As there were only twelve samples to be analysed, the QC sample was injected between every three samples, resulting in just five QC analyses in total. Table 7.13 shows the peak areas for the six peaks of interest in the five QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 7.13: *Table showing peak areas, SDs, means, and CV%s for six peaks in five QC samples that were analysed throughout the analytical run for yolk organic samples of eggs stored at 5 °C and 23 °C for five weeks*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	14357668	561083847	238380106	41069769	1499532718	332415797
QC2	13746587	561480400	236612674	42211361	1457514100	338975053
QC3	15418169	551799453	225621898	44087771	1377970674	330990129
QC4	14383656	552734170	224197958	39658888	1387151410	313774448
QC5	14941709	569139146	234766715	39174967	1448479259	323989085
SD	635487	7143171	6541290	1990389	50968350	9583078
Mean	14569557	559257403	231915870	41240551	1434129632	328028902
CV%	4.36	1.28	2.82	4.83	3.55	2.92

The table shows that the CV%s for the peak areas of all six peaks across the five QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 7.14 shows the RTs for the six peaks of interest in the five QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 7.14: *Table showing RTs, SDs, means, and CV%s for six peaks in five QC samples that were analysed throughout the analytical run for yolk organic extracts of eggs stored at 5 °C and 23 °C for five weeks*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.947	10.057	32.19	55.467	70.563	78.563
QC2	4.93	9.907	32.045	55.235	70.184	77.774
QC3	4.914	9.892	31.969	55.139	70.019	77.518
QC4	4.932	10.026	32.201	55.436	70.468	78.366
QC5	4.932	10.091	32.232	55.485	70.567	78.514
SD	0.012	0.090	0.114	0.156	0.246	0.472
Mean	4.931	9.995	32.127	55.352	70.360	78.147
CV%	0.24	0.90	0.36	0.28	0.35	0.60

The table shows that the CV%s for the RTs of all six peaks throughout the five QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

7.6.1.2 Metabolite Profiling

The yolk extracts of eggs that were stored for five weeks at both 5 °C and 23 °C were compared by observing the TICs of the extracts following HPLC-MS analysis. Figure 7.24 shows overlaid TICs for the average chromatograms of yolk extracts from eggs that were stored at 5 °C and 23 °C for five weeks.

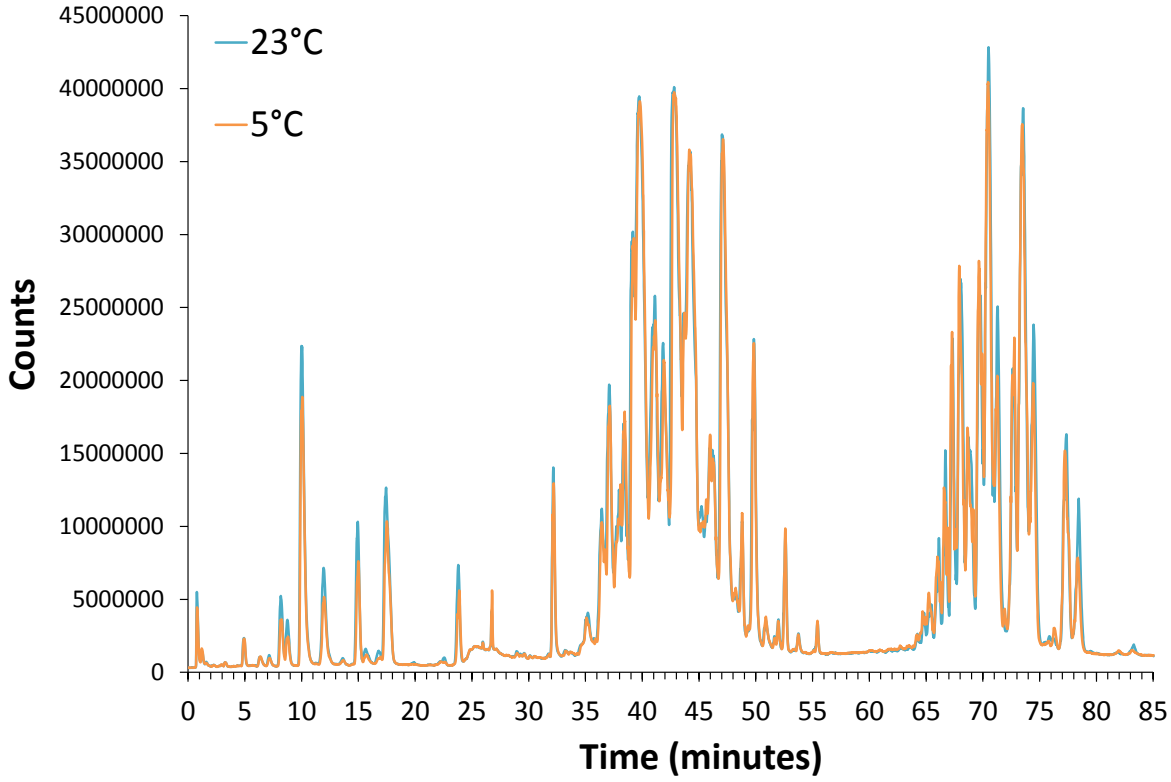


Figure 7.24: *Overlaid average TICs of yolk organic extracts from eggs that were stored at 5 °C and 23 °C for five weeks*

There is a visible difference in metabolite profile between eggs that were stored at the two different temperatures, as can be seen in Figure 7.24. Most chromatographic peaks appear to be of a higher intensity in the profiles of eggs that were stored at 23 °C, however there are also some peaks of a higher intensity in the profiles of eggs that were stored at 5 °C. This is likely to be due to some peaks corresponding to compounds that are products of catabolism and/or degradation, whilst others will correspond to compounds that undergo catabolism and/or degradation to produce these compounds. The compounds that are products of catabolism and/or degradation will be of a higher

intensity in the yolks of eggs that were stored at 23 °C, as the rate of catabolism due to enzyme-catalysed reactions, and degradation, is higher at higher temperatures, whereas the compounds that undergo catabolism and/or degradation will be of a higher intensity in the yolks of eggs stored at 5 °C, as the rate of catabolism and degradation is slower at lower temperatures.

It would be expected that the later peaks in the chromatogram would correspond to larger lipid or phospholipid compounds that undergo catabolism or break-down during storage. However, some of these peaks have a higher intensity in the TIC of yolks from eggs that were stored at 23 °C for five weeks, compared to that of yolks from eggs that were stored at 5 °C. It is not known why this might be the case, however it could be that just one yolk extract from one of the sets of eggs that were stored at either temperature had an anomalous metabolite profile, in terms of compound abundance, to the other yolk extracts from eggs that were stored at the same temperature, which affected the overall average TIC that was produced.

7.6.1.3 Multivariate Statistics

Following the first PCA, scores plots were produced to display any variation between the yolks of eggs that were stored at 23 °C and 5 °C for five weeks. Although there are some clear differences in metabolite profile between the eggs that were kept at 5 °C and 23 °C, that can be seen in Figure 7.24, the scores plot in Figure 7.25 does not appear to show any difference between the two sets of eggs. This indicates that the differences in yolk metabolite profile between the eggs that were stored at different temperatures are extremely subtle, and that the largest amount of variation is due to random differences between individual eggs, rather than any trend between eggs that were kept at different storage temperatures. The more highly abundant compounds that contribute to the observed differences between the TICs of the yolks of eggs that were stored at 23 °C and 5 °C for five weeks may not contribute as much to the variation that is explained

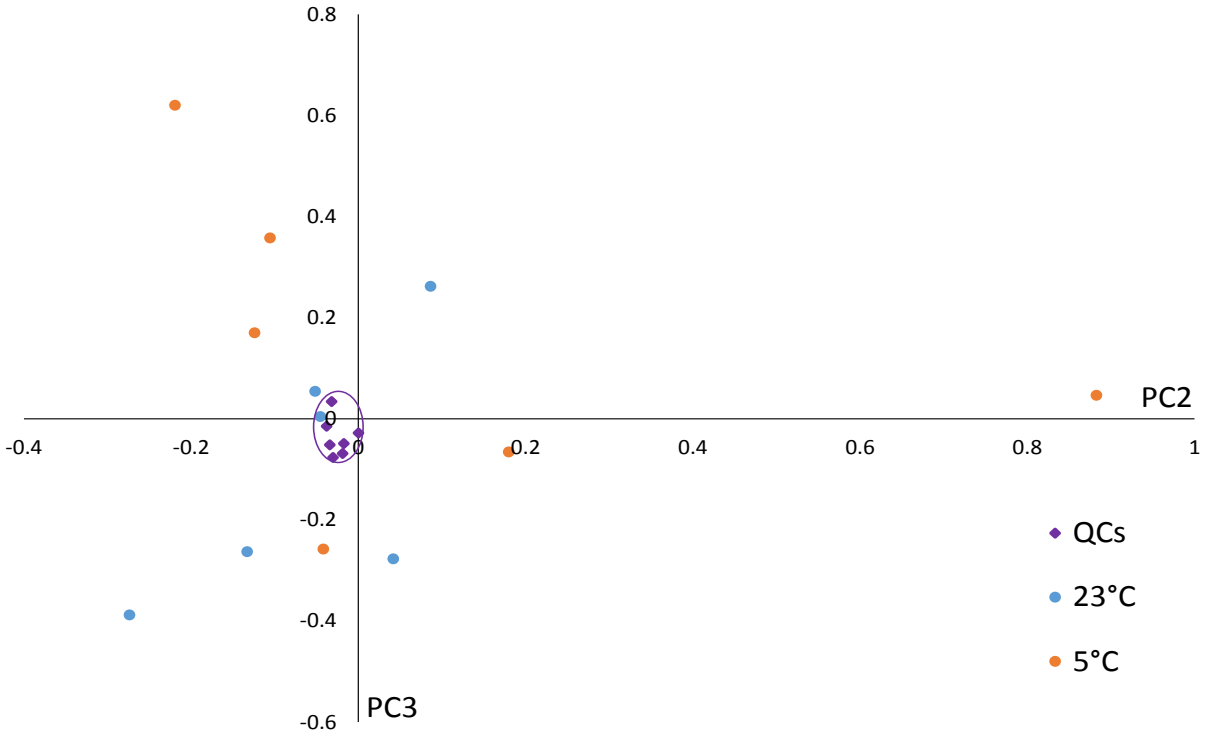


Figure 7.25: *PCA scores plot showing PC2 vs PC3 for eggs stored at 5 °C and 23 °C for five weeks, including all compounds with CV%<30%. PC2 explains 0.50% of the variance, and PC3 explains 0.15% of the variance.*

by PC2 and PC3 as the less abundant compounds, which may experience more variance between random eggs, rather than between eggs that were stored at the two different temperatures. This would explain why the scores plot in Figure 7.25 does not show the differences between the two sets of eggs that were observed from the overlaid TICs in Figure 7.24. There is a tight clustering of QC samples, which gives a good indication that the analysis was robust and that there was little instrumental drift throughout the analysis.

After the *t*-tests were carried out, and following the removal of any compounds with $p>0.01$, a second PCA was carried out using the remaining compounds. Scores plots were again produced to display the variation between the yolk extracts from eggs that were stored at the two different temperatures.

As can be seen in Figure 7.26, when only statistically significant compounds are in-

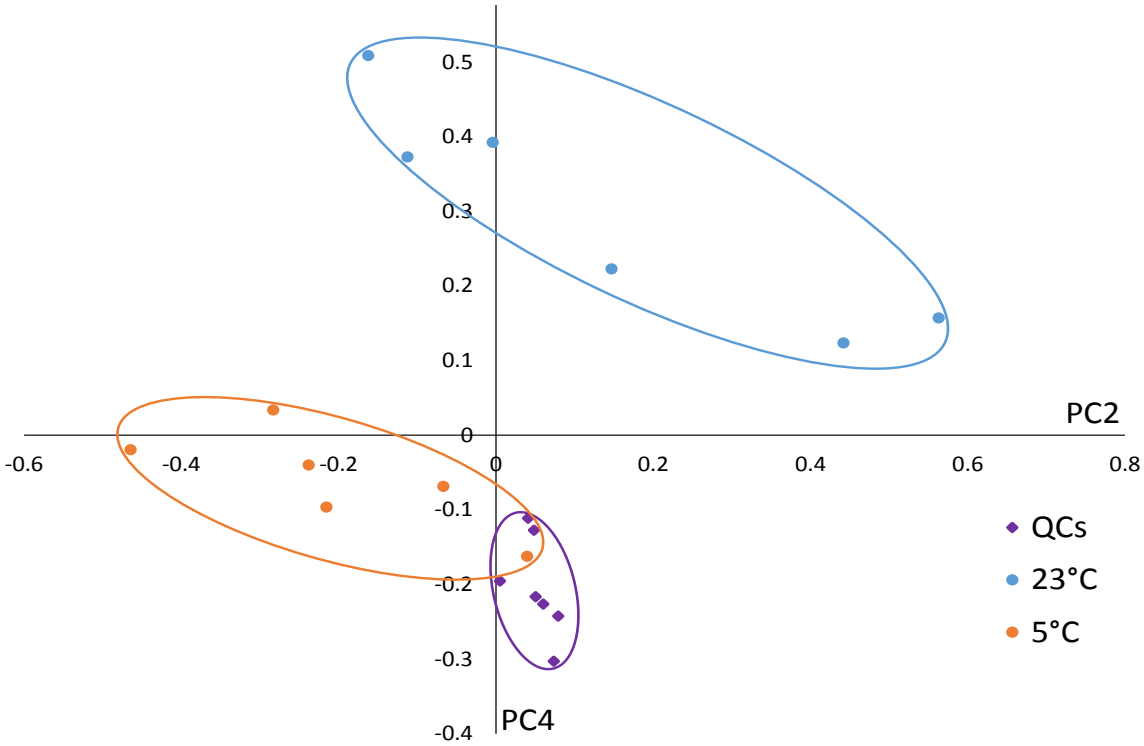


Figure 7.26: *PCA scores plot showing PC2 vs PC4 for eggs stored at 5 °C and 23 °C for five weeks, including all compounds with CV%<30% and p<0.01. PC2 explains 0.14% of the variance, and PC4 explains 0.0031% of the variance.*

cluded in the PCA, full separation can be observed between the two sets of eggs on the resulting scores plot. This separation between the eggs that were stored at the two different temperatures is across both PC2 and PC4, which explain only 0.14% and 0.0031% of the variance respectively. Although there is a complete separation between the two sets of eggs when plotting PC2 and PC4, the differences remain extremely subtle as evidenced by the very small percentage of variance explained by PC2 and PC4. PC1 explains 99.82% of the variance, which is between random eggs rather than between eggs that were stored at the two different storage temperatures, again showing how subtle the differences are between the yolks of eggs stored at two different temperatures for five weeks. The QC samples are grouped more closely together than the samples in the two sample sets, indicating that there was little instrumental drift affecting the analysis, making it robust.

7.6.1.4 Univariate Statistics

The CV%s and t-test p-values of the top 50 statistically significant compounds remaining, following the removal of any duplicates, adducts, or isotopes from the top 100 compounds based on PC1 loadings of the second PCA, including only compounds with $p<0.01$, can be seen in Table D.6 in Appendix D. Those in bold remained statistically significant when the same analyses were carried out using the raw data. The table also shows whether the compounds were of a higher or lower abundance in egg yolk when the eggs were stored at the two different temperatures. As can be seen in the table, most of the lower molecular weight compounds were of a lower abundance in the yolks of eggs that were stored at 5 °C, compared to those stored at 23 °C, whilst the reverse trend was seen with the higher molecular weight compounds. This is likely to be due to the lower storage temperature reducing the rate of enzyme-catalysed catabolic reactions and metabolite degradation in the egg yolks, resulting in the production of fewer molecules that are products of catabolism or degradation, which are of a lower molecular weight. As the higher molecular weight compounds undergo a lower rate of catabolism and degradation at 5 °C, these are of a higher abundance in the yolks of eggs kept at this temperature as they are not being broken down as quickly as they would be in the eggs stored at 23 °C.

Of the 29 compounds that still showed statistical significance when studying the raw data, two were putatively identified. These two putatively identified compounds were m/z 104.1072 and 331.2856 and were identified as choline and 1-monopalmitin respectively, which were previously identified in Section 7.4.

The abundance of the putatively identified 1-monopalmitin was significantly lower, with $p<0.001$, in the yolks of eggs that were stored at 5 °C for five weeks than in the yolks of eggs that were stored at 23 °C for the same length of time, as can be seen in Figure 7.27. This is what would be expected, as monoglycerides are catabolic and degradation products of other compounds, such as larger lipids and phospholipids, which

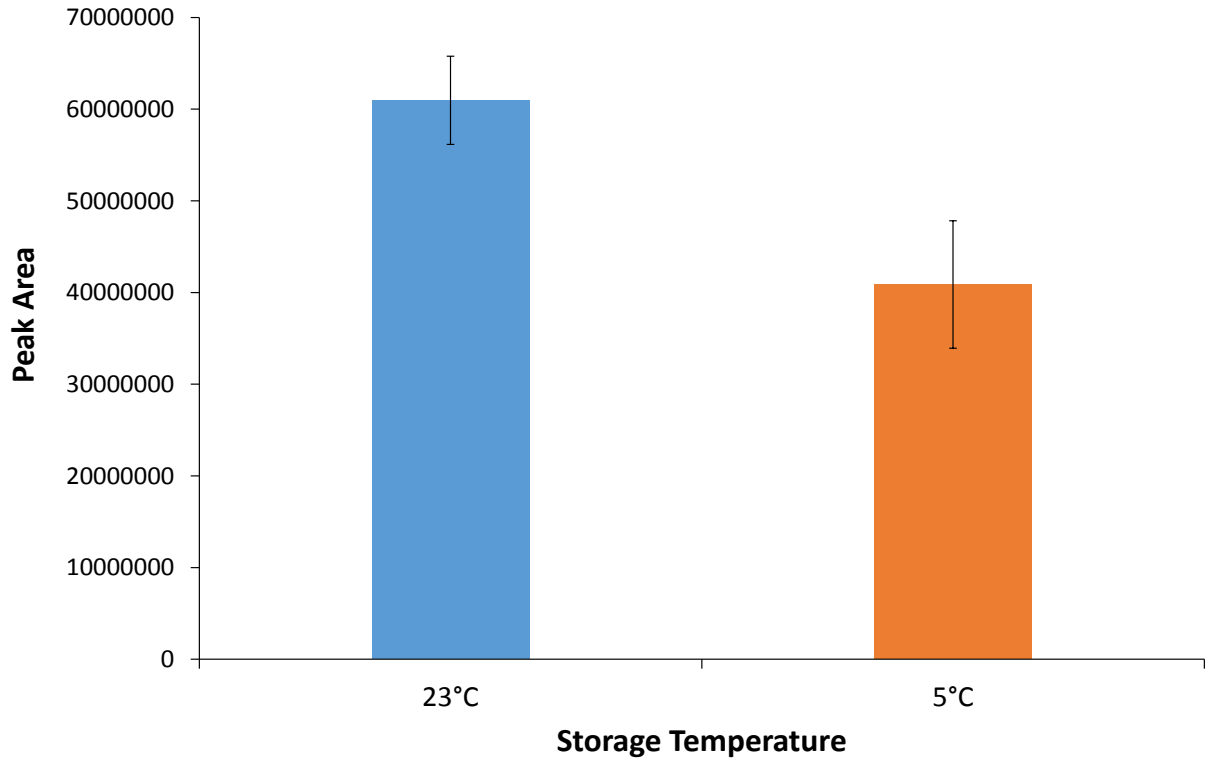


Figure 7.27: *Bar chart showing the difference in abundance of the compound putatively identified as 1-monopalmitin between eggs stored at 23 °C and 5 °C for five weeks. n=6. Measured as mean values of peak area at each storage temperature, with error bars ±1 standard deviation.*

will have been breaking down throughout the five weeks post-lay. Metabolic processes and metabolite degradation occur at slower rates in lower temperatures, which explains why the abundance of 1-monopalmitin is lower in the yolks of eggs that were stored at 5 °C; the enzyme-catalysed reactions and degradation causing the breakdown of the lipids and phospholipids that produce monoglycerides are slower at this temperature, resulting in fewer monoglyceride molecules being produced.

The statistical significance of choline was quite high, with p=0.004, and the abundance of choline was lower in the yolks of eggs that were stored at 5 °C, compared to those that were stored at 23 °C, as can be seen in Figure 7.28. Again, this is expected and is due to the lower storage temperature reducing the rate of phospholipase-catalysed reactions and metabolite degradation in the egg yolks, resulting in a reduced rate of PC catabolism and degradation. Therefore, fewer choline molecules were produced, compared to when

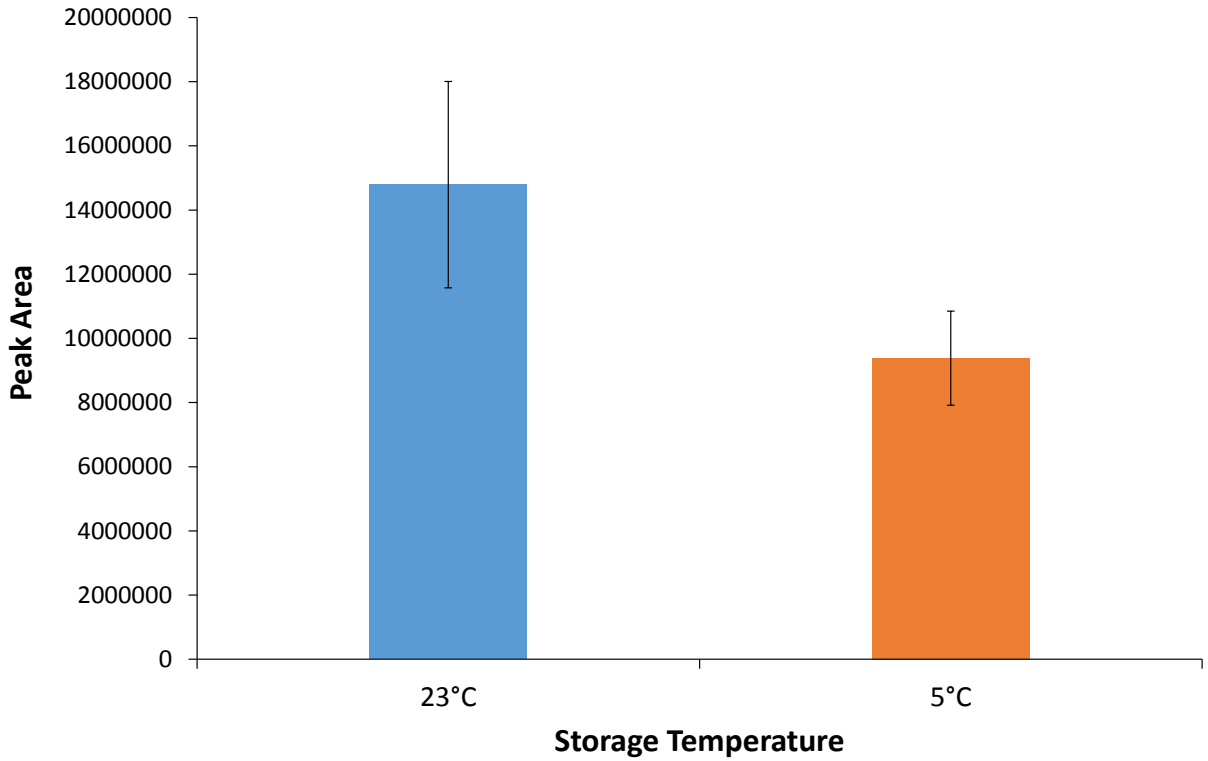


Figure 7.28: *Bar chart showing the difference in abundance of choline between eggs stored at 23 ˚C and 5 ˚C for five weeks. n=6. Measured as mean values of peak area at each storage temperature, with error bars ±1 standard deviation.*

eggs were stored at the higher temperature.

This clear difference of choline abundance in egg yolk between eggs that were stored at the two different temperatures confirms what was concluded previously; that storing eggs in a refrigerated environment can sufficiently reduce the rate of phosphatidylcholine catabolism and degradation, and therefore choline production, thereby inhibiting the increase of choline concentration with increasing storage time, preventing the accurate prediction of egg age based on yolk choline concentration.

7.7 Liquid Eggs Stored at 5 °C for 96 hours: Yolk Organic Extracts

Extracts of liquid egg yolk samples that had spent varying lengths of time in the refrigerator were compared to see whether any differences in their metabolite profiles could be observed.

7.7.1 Results and Discussion

7.7.1.1 Quality Control Analysis

Table 7.15 shows the peak areas for the six peaks of interest in the five QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 7.15: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for organic extracts of liquid egg yolk stored at 5 ℃*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	1605529	135851252	58016201	37514692	360943898	76295402
QC2	1482151	128098765	53849362	35417721	366632117	63091223
QC3	1459456	115277006	51075412	33926202	387771662	98715569
QC4	1605619	134660417	57289172	32151508	358167911	53980078
QC5	1465560	133838151	57688026	35086345	368768983	57742355
QC6	1539140	129582278	57332150	35265501	380836210	51265887
QC7	1566615	135187340	57015620	35842661	396332858	55346117
SD	63557	7271667	2590022	1660680	14362063	16948991
Mean	1532010	130356458	56037991	35029232	374207662	65205233
CV%	4.14	5.58	4.62	4.74	3.84	25.99

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift affecting the detector response occurring throughout the analysis.

Table 7.16 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 7.16: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples that were analysed throughout the analytical run for organic extracts of liquid egg yolk stored at 5 ℃*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.430	8.911	32.233	56.318	71.590	78.147
QC2	4.421	8.953	32.308	56.493	71.814	78.487
QC3	4.412	8.977	32.382	56.633	72.220	79.043
QC4	4.458	9.056	32.477	56.729	72.548	79.619
QC5	4.479	9.010	32.515	56.883	72.685	79.889
QC6	4.435	9.067	32.571	57.055	72.941	80.311
QC7	4.469	9.083	32.621	57.122	73.024	80.460
SD	0.025	0.064	0.142	0.293	0.550	0.891
Mean	4.443	9.008	32.444	56.748	72.403	79.422
CV%	0.57	0.71	0.44	0.52	0.76	1.12

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

7.7.1.2 Metabolite Profiling

The extracts of liquid egg yolk samples that were stored at 5 °C for different lengths of time were compared initially by observing the TICs of the extracts following HPLC-MS analysis. Figure 7.29 shows overlaid TICs for the average chromatograms of extracts from fresh egg yolk and liquid egg yolk that was stored for 96 hours.

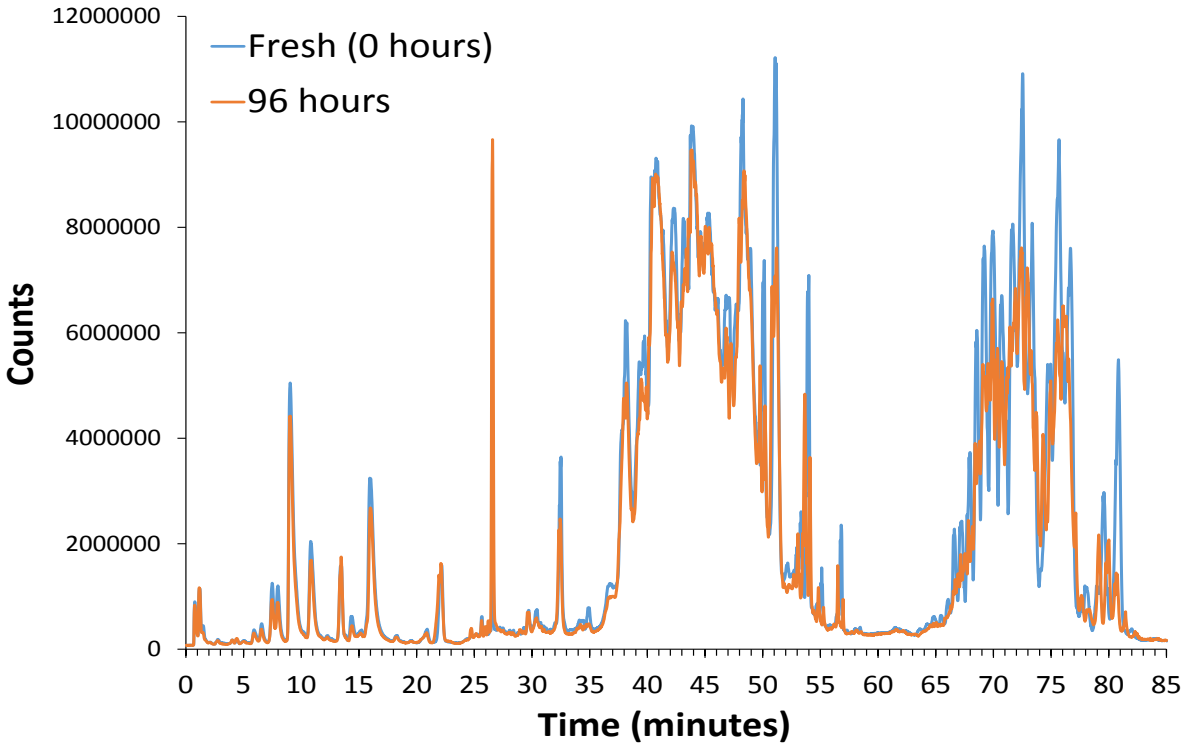


Figure 7.29: *Overlaid average TICs of organic extracts from fresh egg yolk and liquid egg yolk that was stored at 5 °C for 96 hours*

The overlaid TICs in Figure 7.29 show a clear difference in metabolite profile between fresh egg yolk and liquid egg yolk that was refrigerated for 96 hours. There is a higher peak intensity for most of the peaks in the TIC of fresh egg yolk extracts compared to the TIC of extracts of yolk that was stored for 96 hours, which is comparable to what was observed when comparing the TICs of the yolks of fresh eggs and eggs that had been stored for five weeks at both 23 °C and 5 °C. This is particularly noticeable in the latter half of the analysis, after 45 minutes.

After 45 minutes of analysis, the TIC of the extracts of liquid yolk that was stored at

5 °C for 96 hours starts to look different compared to other TICs produced throughout this research. This is due to retention time drift occurring throughout the analytical sequence, which affected the later eluting compounds. Because of this, the average TICs that were produced resulted in a poor peak shape for the later peaks in the chromatogram. Therefore, the observed lower peak intensity for the peaks after 45 minutes in the TIC of extracts of yolk that was refrigerated for 96 hours may be due to the poor peak shape, in addition to or rather than, true differences in metabolite profile.

Although retention time drift occurred throughout the analytical sequence, the analysis was still robust as can be seen in Section 7.7.1.1.

7.7.1.3 Multivariate Statistics

Following PCA, scores plots were produced to display the variation between the egg yolk samples that were stored for different lengths of time at 5 °C. Although the TICs in Figure 7.29 show a clear difference between fresh egg yolk and yolk that was refrigerated for 96 hours, the scores plot in Figure 7.30 does not reflect this. There is very little separation between the sets of egg yolk that were stored for different lengths of time, and in fact most of the variation appears to be between random samples. It may be that most of the differences between the TICs of fresh yolk and yolk that was refrigerated for 96 hours, that can be seen in Figure 7.29, are due to the poor peak shape of the later peaks in the average TIC of yolk that was refrigerated for 96 hours, rather than true differences in metabolite profile. It could also be that the more highly abundant compounds that contribute to the differences in peak intensity between the TICs of fresh yolk and refrigerated yolk contribute less to the variance that is explained by PC3 and PC4 than the less abundant compounds, which may experience more variation between random yolk samples, rather than between yolk samples that have been refrigerated for different lengths of time.

However, fresh egg yolk samples do show some very slight separation to the other yolk

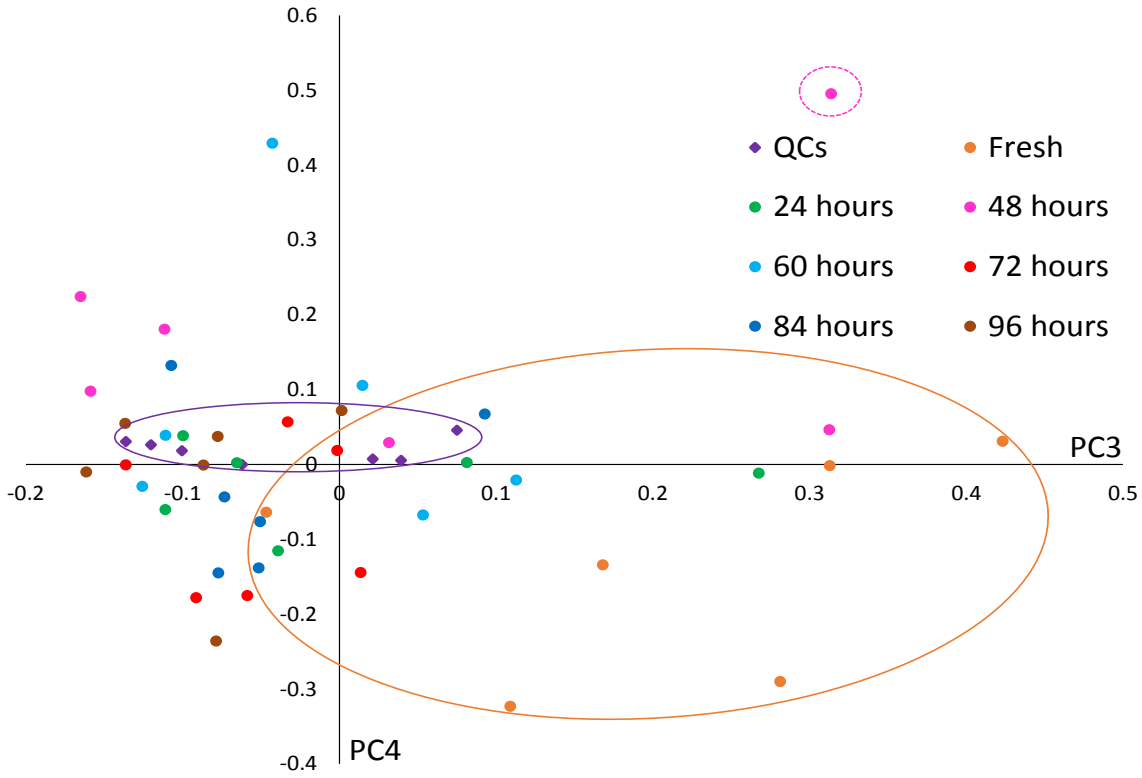


Figure 7.30: *PCA scores plot showing PC3 vs PC4 for egg yolk stored at 5 °C up to 96 hours, including all compounds with CV%<30%. PC3 explains 0.20% of the variance, and PC4 explains 0.16% of the variance. Circled: sample separate to the rest of sample set.*

samples, across PC3. This indicates that some of the differences between fresh yolk and yolk that was refrigerated for 96 hours, that were observed in the overlaid TICs in Figure 7.29, are represented by the variance that is explained by PC3. The variation between samples within the fresh egg yolk sample set is quite high, with the samples spreading out over most of the scores plot, across PC3. This is not what would be expected, as these extracts should all be very similar, as they were all produced from the same pooled sample of fresh egg yolk. The QC samples also appear to be spread out across PC3, however there is no run order effect in the spread and they are still the most clustered group of samples, indicating that the analysis was robust and that instrumental drift did not affect the results. This spread of QC samples, and fresh yolk samples, suggests that the differences between the egg yolk samples that were stored for different lengths of time were extremely subtle, as the PCA began exploiting differences between the QC samples,

which were all identical, and showed a high variance between the fresh yolk samples which should all have been very similar.

There is one sample which corresponds to an extract of yolk that was refrigerated for 48 hours (circled), that is very separate to the other yolk extracts that were stored for 48 hours. It is not known why this may be, as all yolk samples were created from one large pooled sample of egg yolk, and nothing unusual was noticed during the extraction procedure, or in the resulting chromatogram. It may be that the location of this sample was at a slightly different temperature to the rest of the refrigerator, or that the lid was not applied to the sample tube correctly, or that an error occurred during the extraction or analysis that was not observed at the time.

7.7.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 50 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, all 50 compounds were found to be significantly different between yolk samples that were stored at 5 °C for different lengths of time. These compounds, and their CV%s and p-values, can be seen in Table D.7 in Appendix D; those in bold were also found to be significantly different when the raw data was used to carry out the same tests.

Some of these top 50 compounds were also found to be statistically significant in the study observing the yolks of whole shell eggs that were stored at 5 °C for five weeks. Some compounds appear to follow similar trends between these two studies; they experience a significant increase in abundance over 96 hours of liquid yolk storage at 5 °C as well as over the first week of whole shell egg storage at the same temperature. However, others do not appear to follow similar trends. Some compounds show significant differences in abundance over 96 hours of liquid yolk storage, but not over one week of whole shell egg storage, and others do show significant differences across the similar time periods in

both studies, but with the abundance changing in opposite directions. These differences could be due to slightly different temperature conditions between the two studies; there may have been temperature fluctuations in the refrigerator, or the eggs may have been stored in different positions with slightly different temperatures. The differences could also be due to the eggs in the previous study being stored as whole shell eggs, and in the current study as liquid yolk. It would be interesting to repeat this study over several more days, to see whether the changes in compound abundance that appear to be opposite to those observed in the previous study, start to change direction due to further metabolic changes, emulating what was observed previously.

Of the 36 compounds that were found to remain statistically significant when analysing the raw data, four were putatively identified through comparing mass spectra provided by METLIN with mass spectra resulting from the chemical analysis in this study. The observed *m/z* values for all of these compounds are due to the $[M+H]^+$ adduct. The compounds *m/z* 310.3103, 331.2840, 338.3416, and 525.3743 were putatively identified as oleoyl ethyl amide, 1-monopalmitin, docosenamide (erucamide), and 1-palmitoyl-2-acetyl-glycero-3-phosphatidylcholine (Platelet Activating Factor (PAF) C-16), respectively. The compounds *m/z* 331.2840 and 338.3416 have been previously putatively identified, and the comparison of mass spectra provided by METLIN with mass spectra obtained from chemical analysis for these compounds can be seen in Appendix D, Figures D.2 and D.4 for *m/z* 331.2840 and 338.3416 respectively. The comparison of mass spectra provided by METLIN with mass spectra obtained from analysis in this study for compounds *m/z* 310.3103, and 525.3743 can be found in Appendix D, Figures D.7 and D.8 respectively.

Tables 7.17-7.20, and Figures 7.31-7.34, show the putative identifications of the four compounds, the *p*-values resulting from ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests, as well as the trends in the changing abundance of the compounds in the liquid yolk ranging from fresh to 96 hours of storage at 5 °C.

Table 7.17: *Table showing the putative identification of compound m/z 310.3103 and the p-values resulting from Welch test and Games-Howell tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	Welch *p*-value
310.3103	$C_{20}H_{39}NO$	97.31	Oleoyl ethyl amide	<0.001

Post-hoc test *p*-values						
	24 hours	**48 hours**	**60 hours**	**72 hours**	**84 hours**	**96 hours**
Fresh	0.045	0.026	0.014	0.004	0.004	0.004
24 hours		0.349	0.073	<0.001	0.004	<0.001
48 hours			0.665	0.001	0.016	0.001
60 hours				0.019	0.141	0.021
72 hours					0.989	1
84 hours						0.993
96 hours						

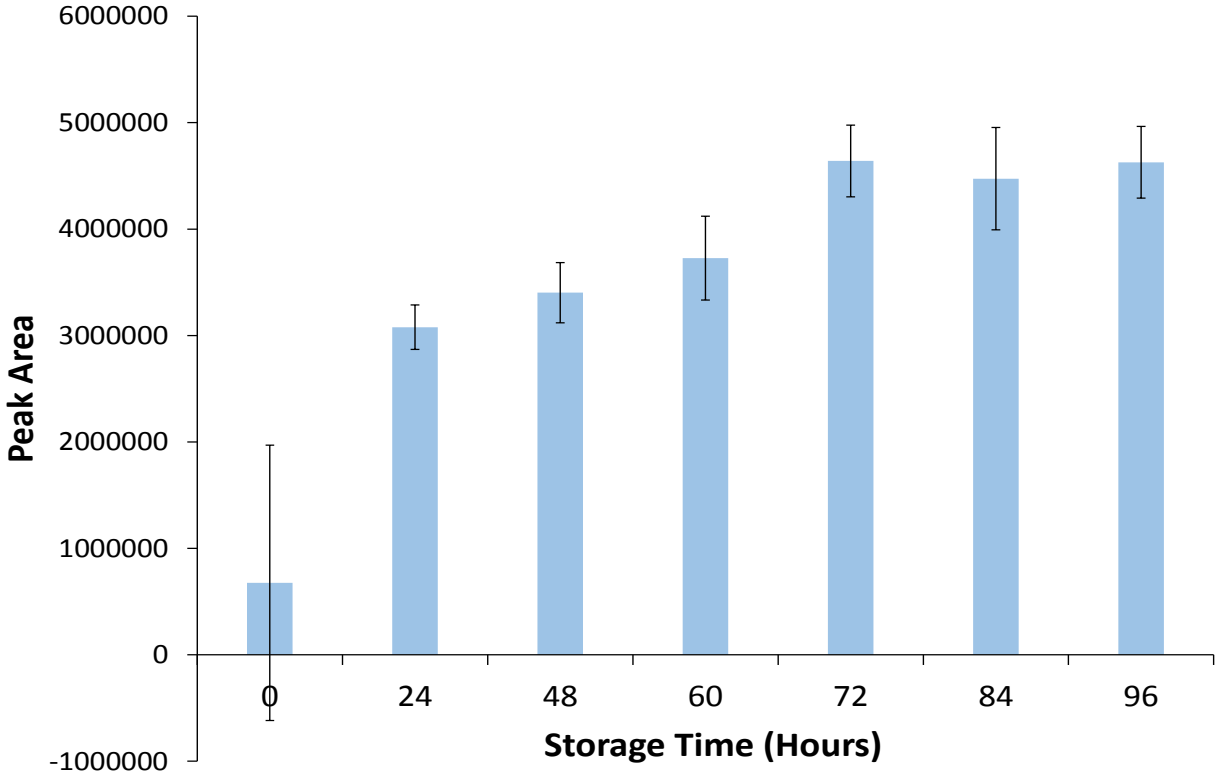


Figure 7.31: *Bar chart showing the trend in changing abundance of compound m/z 310.3103 across liquid egg yolk stored for up to 96 hours. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Table 7.18: *Table showing the putative identification of compound m/z 331.2840 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
331.2840	$C_{19}H_{38}O_4$	99.57	1-monopalmitin	<0.001

Post-hoc test *p*-values						
	24 hours	**48 hours**	**60 hours**	**72 hours**	**84 hours**	**96 hours**
Fresh	<0.001	0.033	<0.001	0.005	<0.001	0.002
24 hours		<0.001	0.47	0.002	0.978	0.005
48 hours			0.045	0.99	0.002	0.917
60 hours				0.21	0.925	0.402
72 hours					0.018	1
84 hours						0.048
96 hours						

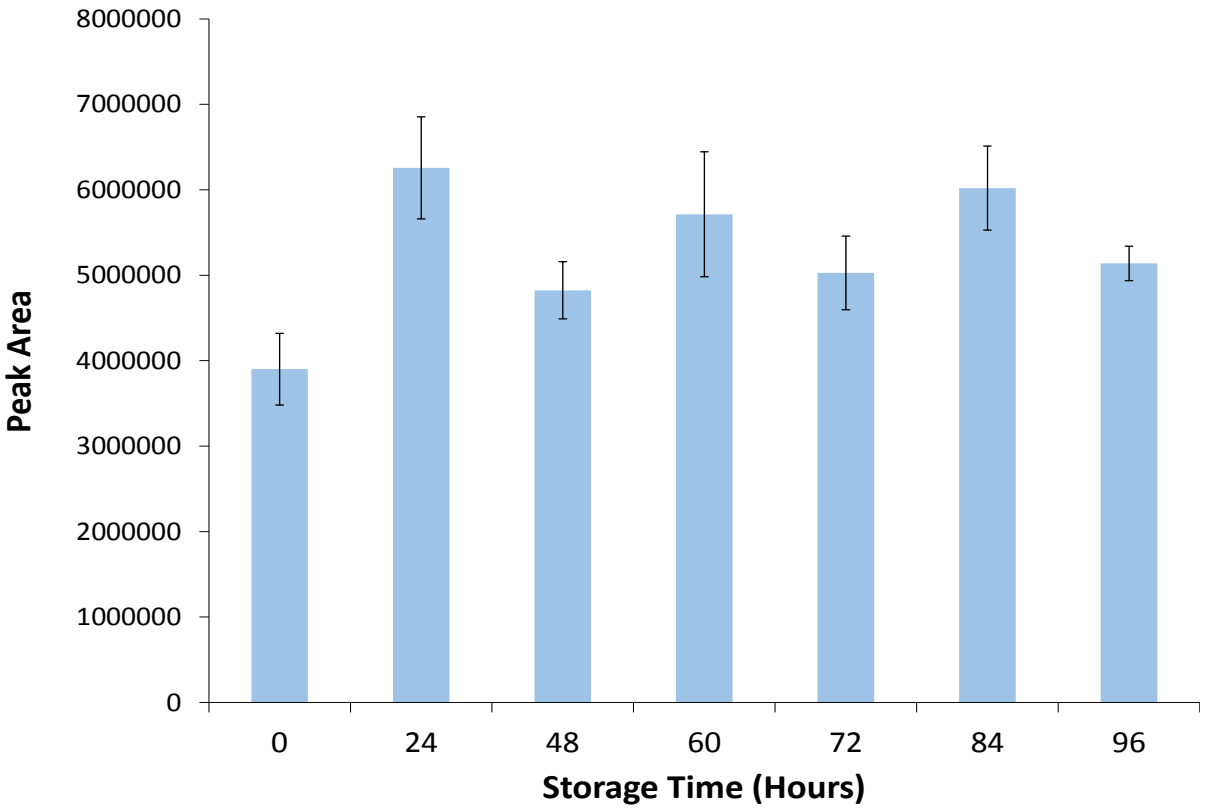


Figure 7.32: *Bar chart showing the trend in changing abundance of compound m/z 331.2840 across liquid egg yolk stored for up to 96 hours. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Table 7.19: *Table showing the putative identification of compound m/z 338.3416 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
338.3416	$C_{22}H_{43}NO_4$	99.92	Erucamide	<0.001

Post-hoc test *p*-values						
	24 hours	**48 hours**	**60 hours**	**72 hours**	**84 hours**	**96 hours**
Fresh	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
24 hours		0.999	0.908	0.002	0.005	0.004
48 hours			0.992	0.005	0.016	0.013
60 hours				0.033	0.087	0.072
72 hours					1	1
84 hours						1
96 hours						

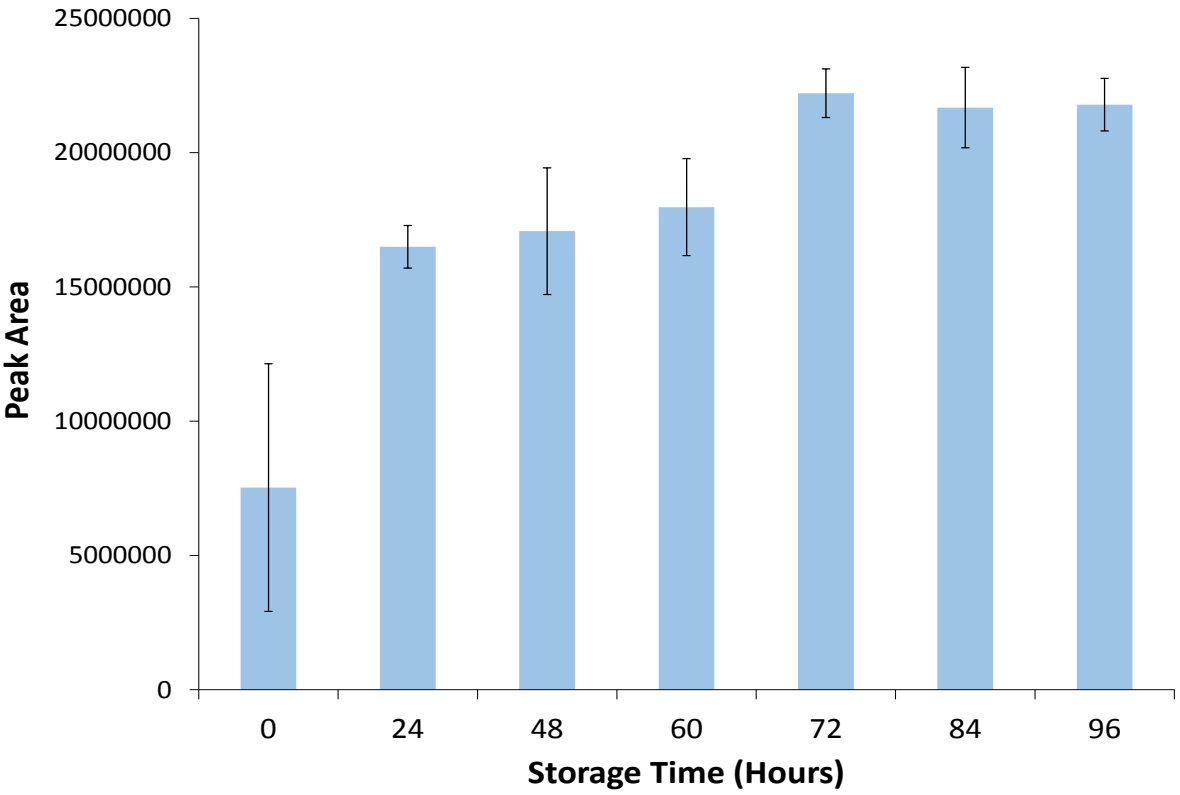


Figure 7.33: *Bar chart showing the trend in changing abundance of compound m/z 338.3416 across liquid egg yolk stored for up to 96 hours. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

Table 7.20: *Table showing the putative identification of compound m/z 525.3743 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
525.3743	$C_{26}H_{54}NO_7P$	99.71	PAF C-16	0.009

Post-hoc test *p*-values						
	24 hours	**48 hours**	**60 hours**	**72 hours**	**84 hours**	**96 hours**
Fresh	0.996	0.626	0.864	0.012	0.092	0.151
24 hours		0.932	0.995	0.057	0.301	0.432
48 hours			0.999	0.435	0.902	0.965
60 hours				0.219	0.686	0.82
72 hours					0.98	0.933
84 hours						1
96 hours						

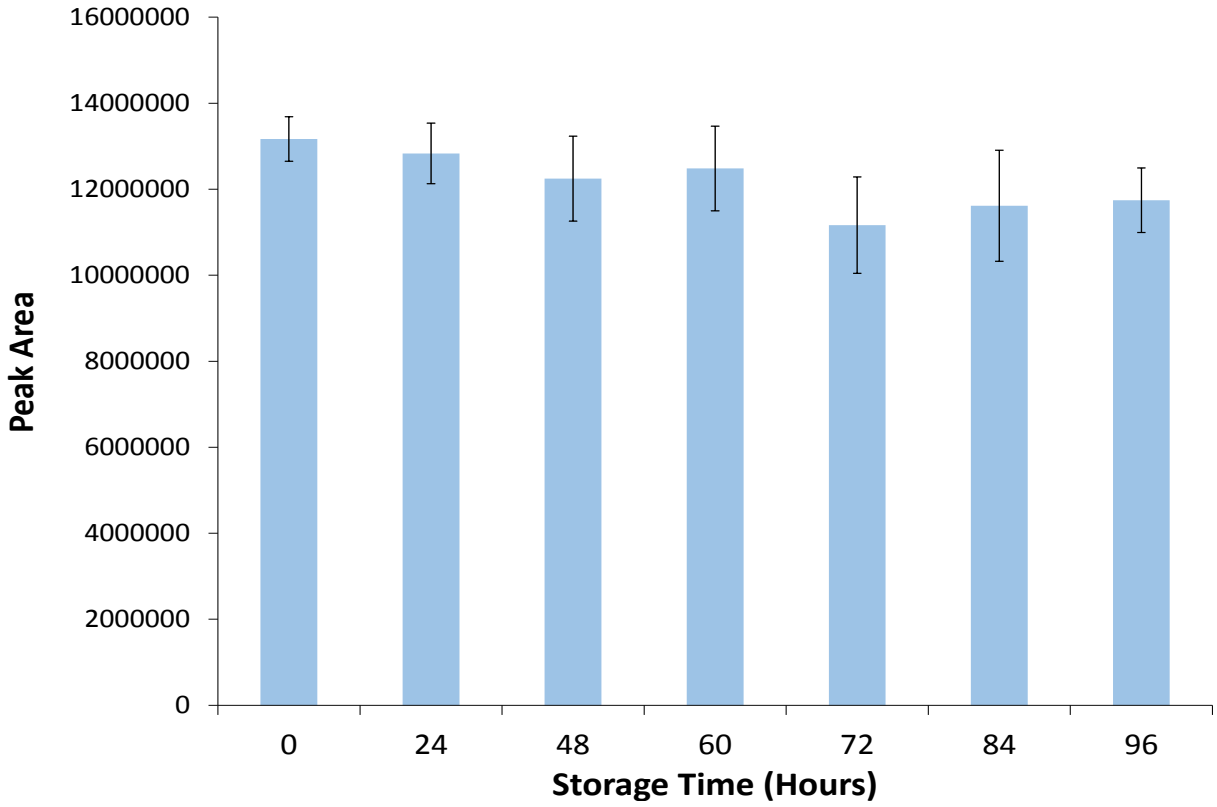


Figure 7.34: *Bar chart showing the trend in changing abundance of compound m/z 525.3743 across liquid egg yolk stored for up to 96 hours. n=6. Measured as mean values of peak area at each egg age, with error bars ±1 standard deviation.*

The probability scores for the predicted molecular formulae for all of these compounds are above 95, indicating that there is a strong likelihood that the predicted formulae correspond to the compounds of interest. The p-values from ANOVA and Welch tests show that the abundances of these compounds are statistically significant when comparing them between yolk that was stored for different lengths of time, ranging from fresh to 96 hours of storage, and the post-hoc test p-values mostly corroborate this.

The compound that was putatively identified as oleoyl ethyl amide, an amide derivative of the fatty acid oleic acid, with an ethyl group attached to the nitrogen of the amide group, has several points of high statistical significance between liquid yolk samples that were stored for different lengths of time. The post-hoc test p-values in Table 7.17 show a significant difference between fresh yolk and yolk that has been refrigerated for 24 hours, with an increasing significant difference between fresh yolk and yolk that has been stored for increasing lengths of time, up to 72 hours of storage, at which point the significance is high and remains the same over the remaining storage time. This suggests that there was a gradual change in compound abundance with increasing storage time of the yolk, up to 72 hours, at which point the metabolic changes affecting the abundance of this compound slowed down, resulting in no further increase in significant difference between fresh egg yolk and yolk that was stored for 84 and 96 hours. The bar chart in Figure 7.31 confirms this trend in compound abundance. It shows that there is an increase in compound abundance with increasing storage time of the yolk over the first 72 hours of storage, and then very little change, with no correlation, afterwards.

Both the post-hoc test p-values in Table 7.17 and the bar chart in Figure 7.31 show that there is a significant difference in compound abundance between yolks that were stored at 5 °C for 48 hours, and those that were stored for 72 hours. As EU legislation states that liquid egg can only be stored at refrigerated temperatures for 48 hours before it must be disposed of [19], this compound has potential as a biomarker for this type of fraud, as it can distinguish between liquid egg yolk which has been stored for 48 hours

and that which has been stored for 72 hours and over. As there is a significant difference in compound abundance over the first 24 hours of yolk storage, this compound could also be used to discriminate between fresh yolk, and yolk that has been refrigerated for at least 24 hours.

The compound that was putatively identified as 1-monopalmitin has several points of high statistical significance between yolks that were stored for various lengths of time. However, the bar chart in Figure 7.32 shows that the changing abundance of this compound with increasing storage time of the yolk is very erratic and does not experience any correlation. Therefore, this compound is not useful in predicting the age of liquid egg yolk that has been stored at 5 °C. However, both the post-hoc test p-values in Table 7.18, and the bar chart in Figure 7.32, show that there is a significant difference in the abundance of this compound between fresh egg yolk and liquid egg yolk that has been stored for any length of time up to 96 hours. So although this compound does not show much potential as a marker of the age of liquid egg yolk, it could be useful when determining whether egg yolk is fresh or not. This compound was also in the top 48 statistically significant compounds when whole shell eggs were stored at 5 °C for five weeks, where it was found to significantly increase in abundance in yolk over the first week of storage. In this study, although the changing abundance of this compound is quite erratic, it does experience an overall increase after the four days (96 hours) of liquid egg yolk storage, which corresponds to the trend seen in the previous study.

The compound that was putatively identified as erucamide has several points of high statistical significance, particularly between fresh yolk and yolk that was refrigerated for any length of time, as shown by the post-hoc test p-values in Table 7.19. This indicates that there was a large change in compound abundance that occurred during the first 24 hours of storage. There are also significant differences in compound abundance between eggs that were stored for 60 hours and those that were stored for 72 hours, and between eggs that were stored for 24 and 48 hours and those that were stored for 72 hours and

longer. This suggests that there may have been a further greater change in compound abundance between eggs that were stored for 60 hours and those that were stored for 72 hours. The bar chart in Figure 7.33 confirms this; there is a large increase in compound abundance over the first 24 hours of yolk storage, followed by another jump in compound abundance between 60 and 72 hours of storage, after which point there is very little change in abundance. This could be due to the metabolic processes that result in the production of this compound reaching the point at which most substrate has been catabolised or degraded, and thus most product has been produced, at 72 hours of refrigerated storage.

Although there was not a steady increase in compound abundance observed over the 96 hours of yolk storage, this compound does show some potential as a marker of yolk storage time, particularly between fresh yolk, and yolk that has been refrigerated for at least 24 hours. The post-hoc test p-values in Table 7.19, and the bar chart in Figure 7.33, both show that there is a significant difference in compound abundance between yolk that was stored for 48 hours, and yolk that was stored for 72 hours and over. As legislation states that liquid egg can only be stored for up to 48 hours before being used [19], this compound has potential use in uncovering fraud where liquid egg yolk may have been kept for 72 hours or longer.

The trend in changing abundance of the compound that was putatively identified as erucamide appears to be very similar to that of the compound that was putatively identified as oleoyl ethyl amide. Oleoyl ethyl amide is an amide derivative of oleic acid, and erucamide is an amide derivative of erucic acid, which has a hydrocarbon chain consisting of four more carbon atoms than oleic acid. These compounds are clearly very similar, and experience similar metabolic processes, therefore it is not surprising that they undergo such a similar trend in changing abundance in liquid egg yolk.

This compound was also in the top 48 statistically significant compounds when whole shell eggs were stored at 5 °C for five weeks, where it was found to significantly decrease in abundance in yolk over the first week of storage. This is contradictory to what has

been observed in this study, as the compound was found to experience an overall large, significant increase in abundance in yolk that was stored at 5 °C for four days. This difference in the changing abundance of this compound may be due to the eggs being stored as whole shell eggs in one study, and just as liquid yolk in the other. It could also be that if this current study was continued over a longer period of time, the abundance of this compound might be observed to decrease and follow the trend that was observed in the previous study.

The compound that was putatively identified as PAF C-16 is a glycerophosphatidylcholine with a palmitoyl group attached to the first carbon of the glycerol backbone by an ether linkage and an acetyl group attached to the second carbon by an ester bond. The m/z 525.3743 is actually the $[M+H]^+$ adduct of the C-13 isotope of the compound, as can be seen in the mass spectra comparison in Appendix D, Figure D.8. It is not known why the isotope was in the top 50 compounds in the statistical workflow and not the actual compound itself. Although the ANOVA p-value in Table 7.20 shows that this compound is highly statistically significant when comparing egg yolk that has been stored at 5 °C for various lengths of time, the post-hoc test p-values do not corroborate this. The only point of significant difference is between fresh yolk, and yolk that was stored for 72 hours. The bar chart in Figure 7.34 shows that the abundance of this compound decreases at 72 hours, explaining this point of statistical significance. As this compound shows little statistical significance, and there is no correlation between abundance and storage time of the yolk, this compound is not suitable as a marker of liquid yolk storage time.

Some of the standard deviation error bars on these bar charts, particularly for fresh egg yolk, are quite wide, indicating a wide amount of variation in compound abundance in yolk samples that were stored for different lengths of time. The large standard deviations observed for fresh yolk corroborate what was observed in the scores plot in Figure 7.30, which showed wide variation amongst the fresh egg yolk samples.

7.7.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 50 statistically significant compounds, and scores plots were again produced to display the variation between egg yolk samples that were stored for different lengths of time.

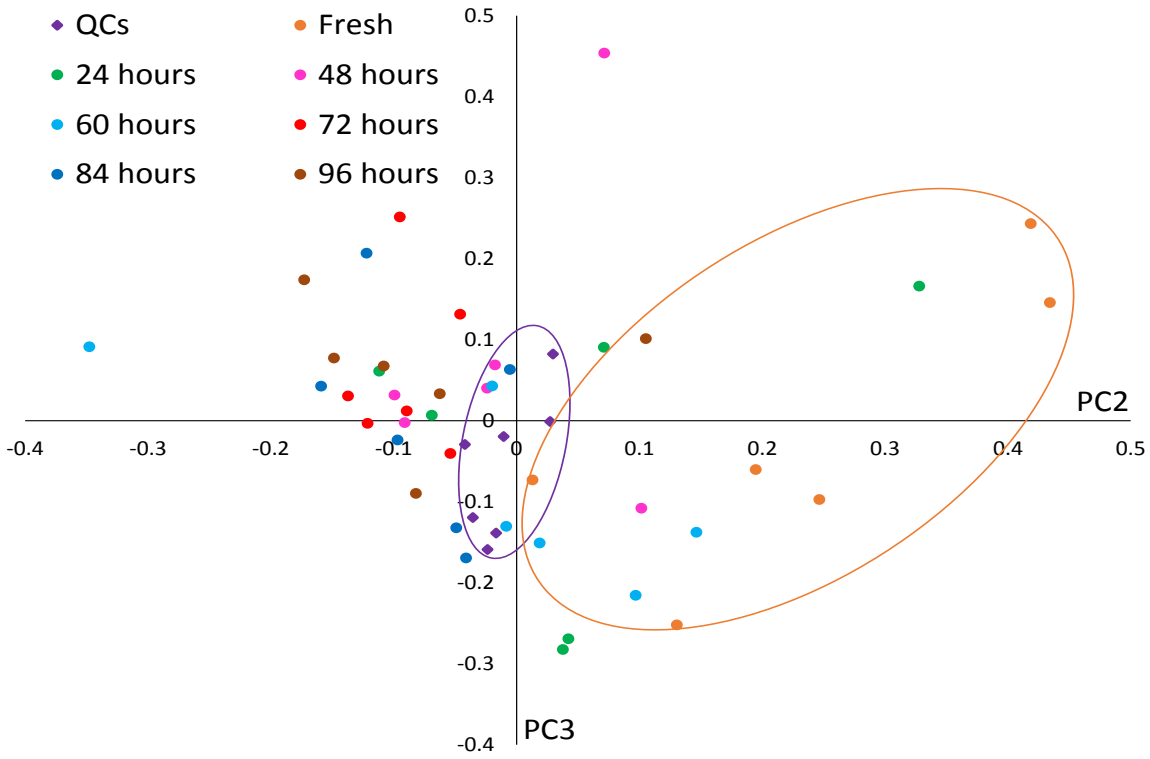


Figure 7.35: *PCA scores plot showing PC2 vs PC3 for egg yolk stored at 5 °C up to 96 hours, including the top 50 compounds. PC2 explains 0.23% of the variance, and PC3 explains 0.10% of the variance.*

The scores plot in Figure 7.35 still only shows very slight separation between fresh egg yolk and yolk that was refrigerated, and no separation between the sets of yolk that were refrigerated for different lengths of time, even though only statistically significant compounds have been included in the PCA. This highlights the subtlety of the differences between fresh egg yolk, and egg yolk samples that have been stored for different lengths of time at 5 °C. There is again a lot of variation within samples of fresh yolk, as well as yolk that was stored for 24, 48, and 60 hours, but from 72 hours of storage onwards, the variance appears to decrease and the samples all group more closely together. This

is likely to be due to the metabolic processes occurring at slightly different rates in the different yolk samples over the first few days of storage, possibly due to their different positions in the refrigerator resulting in slightly different storage temperatures, causing a wide variance of compound abundance between yolk samples that were stored for the same length of time. By 72 hours of storage, the initial metabolic processes will have slowed down and the metabolite profiles will have all reached a similar state in the different yolk samples, resulting in a smaller variance of compound abundance.

7.8 Summary & Conclusions

7.8.1 Summary

The various studies throughout this work have shown, through overlaid TICs and PCA scores plots, that there are differences, although sometimes very subtle, in the metabolite profiles of the yolks of eggs that have been stored for different lengths of time, both at 23 °C, and 5 °C, as well as between liquid yolk samples that have been refrigerated for different lengths of time at 5 °C.

A compound that shows statistical significance when comparing eggs that have been stored at 23 °C for different lengths of time up to five weeks, and that was found to increase gradually in abundance with increasing egg age, was identified as choline. This compound was also found to increase significantly in abundance with increasing egg age over a longer storage time of twelve weeks, when a follow-up targeted study was carried out. This targeted study also quantified the concentration of choline in egg yolk, and found it to increase from an average of 6.8 µg/g to 28.7 µg/g over the twelve week period.

Although several compounds were found to be statistically significant when comparing eggs that were stored for different lengths of time up to five weeks at 5 °C, none of these were determined to be suitable as markers of egg age during refrigerated storage. Choline was found to not be statistically significant when eggs were stored at 5 °C, in both the

non-targeted and follow-up targeted studies, and did not show a correlation between abundance and egg age. This indicates that storing eggs at a lower temperature can prevent the accurate prediction of egg age based on choline concentration in the egg yolk.

When comparing the metabolite profiles of the yolks of eggs that had been stored for five weeks at 23 °C and 5 °C, several compounds were found to be statistically significant. Only two of these were putatively identified, and one of these compounds was choline, which was found to be lower in abundance in the yolks of eggs that were stored for five weeks at 5 °C. This is due to the lower storage temperature reducing the rate of phospholipase-catalysed reactions and metabolite degradation, thereby slowing down the catabolism and break-down of phosphatidylcholines, and therefore the production of choline. This confirms that a lower storage temperature can prevent the accurate prediction of the age of an egg based on the choline concentration of the egg yolk.

The comparison of the metabolite profiles of liquid yolk that had been stored at 5 °C for various lengths of time resulted in the discovery of several compounds that showed a statistical significance between the yolk samples. Four of these were putatively identified, and two were determined to have potential as markers of the age of liquid yolk that has been stored at 5 °C. Oleoyl ethyl amide and erucamide were both found to increase in abundance with increasing storage time of the yolk, and both showed a high statistical significance between yolk that was stored for 48 hours, and yolk that was stored for 72 hours. This makes these compounds potentially useful as markers to detect fraud when liquid yolk is refrigerated for longer than the 48 hours that is legally allowed. These two compounds, as well as the compound that was putatively identified as 1-monopalmitin, also showed potential as markers to discriminate between fresh yolk, and yolk that has been refrigerated for at least 24 hours.

7.8.2 Conclusions

This work has resulted in the identification of choline, a compound that could be used to predict the age of an egg based on its concentration in the yolk. This could help to identify cases of fraud due to date mislabelling, and prevent further instances of fraud by deterring potential fraudsters due to the increased risk of being caught. However, this research has also shown that refrigerating the eggs can prevent the accurate prediction of egg age based on yolk choline concentration. Two compounds were found to show potential as markers capable of discriminating between samples of liquid yolk that have been refrigerated for different lengths of time. This could be useful in detecting fraud, when liquid yolk is refrigerated for longer than the 48 hours that is allowed, prior to use.

Future work would involve repeating these studies with larger sample sizes in order to obtain more robust results. It would be interesting to repeat the choline targeted study using birds of different breeds, to determine whether similar results are obtained. Although it is probable that the positive correlation between yolk choline concentration and egg age would be comparable between breeds, the starting concentration may differ, which would need to be accounted for. Further work would also involve the confirmation that choline abundance increases with increasing egg age due to the catabolism and/or degradation of phosphatidylcholines (PCs). This would be done by observing the differences in PC abundance between the yolks of eggs of different ages, by carrying out MS analysis in precursor ion scan mode for the m/z 184 fragment (phosphocholine), in order to detect the PC ions that fragment to produce this fragment ion [31].

It would also be interesting to repeat the liquid egg storage study using birds of different ages and breeds, and shell eggs of different ages prior to breaking, again to determine whether similar results are obtained. Carrying out similar studies with liquid albumen, and liquid yolk and albumen combined, both refrigerated and stored at room temperature, as well as with liquid yolk stored at room temperature, would also be another aspect of further work regarding this research.

7.9 References

[1] *Regulation of the European Parliament and Council No 853/2004, Laying down specific hygiene rules for on the hygiene of foodstuffs*, Official Journal of the European Union, L139/55, 2004.

[2] A. Fearne and D. Lavelle, "Segmenting the UK egg market: Results of a survey of consumer attitudes and perceptions," *British Food Journal*, vol. 98, pp. 7–12, 1996.

[3] M. Lucisano, A. Hidalgo, E. M. Comelli, and M. Rossi, "Evolution of chemical and physical albumen characteristics during the storage of shell eggs," *Journal of Agricultural and Food Chemistry*, vol. 44, pp. 1235–1240, 1996.

[4] *Commission Regulation No 2295/2003, Introducing detailed rules for implementing Council Regulation No 1907/90 on certain standards for eggs*, Official Journal of the European Union, L340/16-34, 2003.

[5] *Guidance on Legislation Covering the Marketing of Eggs*, Department for Environment, Food and Rural Affairs, Animal and Plant Health Agency, 2016.

[6] *The Eggs and Chicks (England) Regulations*, Food, England, 2009.

[7] *London egg packer and wholesaler found guilty of egg marketing fraud*, Animal and Plant Health Agency, 2016.

[8] *Egg marketing: Technical inspection report, gov.uk*, Animal and Plant Health Agency, https://www.gov.uk/government/publications/egg-marketing-technical-inspection-report, 2014.

[9] Y. C. Liu, T. H. Chen, Y. C. Wu, Y. C. Lee, and F. Tan, "Effects of egg washing and storage temperature on the quality of eggshell cuticle and eggs," *Food Chemistry*, vol. 211, pp. 687–693, 2016.

[10] C. Lapao, L. T. Gama, and M. C. Soares, "Effects of broiler breeder age and length of egg storage on albumen characteristics and hatchability," *Poultry Science*, vol. 78, pp. 640–645, 1999.

[11] R. Karoui, B. Kemps, F. Bamelis, B. deKetelaere, E. Decuypere, and J. deBaerdemaeker, "Methods to evaluate egg freshness in research and industry: A review," *European Food Research and Technology*, vol. 222, pp. 727–732, 2006.

[12] Q. Huang, N. Qiu, M. H. Ma, Y. G. Jin, H. Y. F. Geng, S. G. Sun, and S. H. Sun, "Estimation of egg freshness using s-ovalbumin as an indicator," *Poultry Science*, vol. 91, pp. 739–743, 2012.

[13] A. Al-Shami, G. Mikaylenko, and J. Tang, "Dielectric characterization of hen eggs during storage," *Journal of Food Engineering*, vol. 82, pp. 450–459, 2007.

[14] K. C. Williams, "Some factors affecting albumen quality with particular reference to haugh unit score," *World's Poultry Science Journal*, vol. 48, pp. 5–16, 1992.

[15] D. Shin, C. Narciso-Gaytán, J. M. Regenstein, and M. X. Sánchez-Plata, "Effect of various refrigeration temperatures on quality of shell eggs," *Journal of the Science of Food and Agriculture*, vol. 92, pp. 1341–1345, 2012.

[16] H. Akyurek and A. A. Okur, "Effect of storage time, temperature and hen age on egg quality in free- range layer hens," *Journal of Animal and Veterinary Advances*, vol. 8, pp. 1953–1957, 2009.

[17] A. Giunchi, A. Berardinelli, L. Ragni, A. Fabbri, and F. A. Silaghi, "Non-destructive freshness assessment of shell eggs using FT-NIR spectroscopy," *Journal of Food Engineering*, vol. 89, pp. 142–148, 2008.

[18] A. Hidalgo, M. Lucisano, E. M. Comelli, and C. Pompei, "Evolution of chemical and physical yolk characteristics during the storage of shell eggs," *Critical Reviews in Food Science and Nutrition*, vol. 44, pp. 1447–1452, 1996.

[19] *Regulation (EC) 853/2004 of 29 April 2004, Laying down specific hygiene rules for for food of animal origin*, Official Journal of the European Union, L226/22-82, 2004.

[20] P. Lockman and D. Allen, "The Transport of Choline," *Drug Development and Industrial Pharmacy*, vol. 28, pp. 749–771, 2002.

[21] P. Ueland, "Choline and betaine in health and disease," *Journal of Inherited Metabolic Disease*, vol. 34, pp. 3–15, 2011.

[22] J. Simon, "Choline, betaine and methionine interactions in chickens, pigs and fish (including crustaceans)," *World's Poultry Science Journal*, vol. 55, pp. 353–374, 1999.

[23] S. Zeisel and K. D. Costa, "Choline: An essential nutrient for public health," *Nutrition Reviews*, vol. 67, pp. 615–623, 2009.

[24] D. Küllenberg, L. Taylor, M. Schneider, and U. Massing, "Health effects of dietary phospholipids," *Lipids in Health and Disease*, vol. 11, pp. 1–16, 2012.

[25] M. Anton, F. Nau, and Y. Nys, "Bioactive egg components and their potential uses," *World's Poultry Science Journal*, vol. 62, pp. 429–438, 2006.

[26] M. I. Gurr and J. L. Harwood, *Lipid Biochemistry: An Introduction*, 4th ed. London: Chapman & Hall, 1991.

[27] G. S. Richmond and T. K. Smith, "Phospholipases A_1," *International Journal of Molecular Sciences*, vol. 12, pp. 588–612, 2011.

[28] Q. Wang, G. Jin, N. Wang, X. Guo, Y. Jin, and M. Ma, "Lipolysis and Oxidation of Lipids during Egg Storage at Different Temperatures," *Czech Journal of Food Science*, vol. 35, pp. 229–235, 2017.

[29] *Guidance for Industry: Bioanalytical Method Validation*, Food and Drug Administration, 2018.

[30] (2019). KEGG PATHWAY: Glycerophospholipid metabolism, [Online]. Available: `https://www.kegg.jp/kegg-bin/show_pathway?map00564` (visited on 07/2019).

[31] S. Esposito, S. Colicchia, X. de la Torre, M. Mazzarino, and F. Botrè, "Liposomes as potential masking agents in sport doping. Part 1: Analysis of phospholipids and sphingomyelins in drugs and biological fluids by aqueous normal-phase liquid chromatography-tandem mass spectrometry," *Drug Testing and Analysis*, vol. 9, pp. 75–86, 2016.

8. Differences in metabolite profile between the yolks of eggs from hens kept in cages of different stocking densities

8.1 Introduction

From January 2003, until January 2012, EU legislation stated that un-enriched, or battery cage, systems for laying hens must provide no less than 550 cm^2 of cage area per hen [1]. However, due to welfare considerations, as of January 2012 the use of these un-enriched cages has been prohibited, and all cages must now comply with the requirements of enriched cages, which allow a minimum of 750 cm^2 of cage area per hen, of which 600 cm^2 must be usable [1]. This change created the potential for fraud, as poultry farmers may have chosen not to abide by this new regulation, as it required rearranging of the cages and housing systems of the birds, as well as more space to house the same number of hens. Therefore, it is important to have a method of detecting when this regulation regarding the housing of laying birds in enriched cages is not being followed. There would clearly be a benefit to having a robust and reliable scientific method capable of predicting the stocking density of laying hens, based on the analysis of the egg, to check

for non-compliance.

There has been some research into how the stocking density of the cage of the laying bird affects both egg production, and some physical characteristics of the resulting eggs. Several studies have found that an increase in the stocking density of birds in cages results in lower egg production [2][3][4], which may be attributable to an observed decrease in the feed intake of the birds with the increase in stocking density [2][3]. A decrease in the mass of eggs produced by the birds was also observed in several studies [2][3][4], which could also be attributed to a decreased feed intake of the birds. Yolk colour has been found in one study to be affected by stocking density [2], however another study found that the stocking density of birds did not affect the yolk colour of the resulting eggs [3]. There has been some research into whether the Haugh Unit (HU) of an egg (based on the relationship between the weight of the intact egg, and the albumen height once the egg has been broken) is affected by the stocking density of the cage of the laying hen. However, no relationship has been observed between the stocking density of the cage of the laying hen, and the HU of the eggs [2][3].

Different stocking densities can either be due to different numbers of birds in cages of the same size, or an equal number of birds in cages of different sizes. Differences observed between eggs laid by birds kept in cages of different stocking densities, with different numbers of birds in cages of the same size, could be due to differences in the population size of the cage, rather than just the stocking density. When studying the effect of the cage stocking density of laying hens on eggs, it is important to ensure that any statistically significant potential markers are related to the differences in stocking density, rather than population size.

There has been little published research into the effect of the cage population size of the laying hen on the characteristics of the egg. This could be due to the fact that there is no legislation regarding the population size of a cage, or because there are very few effects on the egg. In fact, one study found that cage population size had no effect on

the rate of egg production, or the mass of the eggs [5].

Although there has been some research into the effects of the stocking density of the cage of the laying hen on the physical characteristics of eggs, there has been no research into how the stocking density affects the chemical composition of the eggs. A metabonomic approach would allow information to be obtained about how a wide range of compounds, and compound classes, are affected by the stocking density of the laying hen. This could result in the development of a method capable of predicting the stocking density of the laying hen, from the egg itself, helping to detect cases of fraud regarding battery cages.

8.2 Aims and Objectives

Having a robust scientific method that is capable of predicting the stocking density of the cage of a laying hen, based on the egg itself, would help to detect breaches of legislation, where birds are kept in un-enriched cages with greater stocking densities than enriched cages with lower stocking densities.

This work aimed to use HPLC-MS analysis as part of a non-targeted metabonomic workflow to determine whether a difference in metabolite profile can be observed between the yolks of eggs laid by birds from cages of different stocking densities. It aimed to identify any compounds that showed potential use as markers, capable of predicting the stocking density of the laying hen cage, which could help to identify cases of fraud. It also used a similar metabonomic workflow to compare the yolks of eggs from cages of different population size, in order to determine whether any observed differences in metabolite profile, between the yolks of eggs laid by birds from cages of different stocking densities, were separate to any differences observed due to cage population size.

8.3 Experimental

8.3.1 Cage Population Size

Eggs were collected on the day of lay from Oaklands Farm Eggs Ltd. (Shrewsbury, U.K.). Laying hens were of the Hy-line brown breed, 21 weeks old at the start of the study, fed the same diet, and kept in enriched cages of the same stocking density. Four sets of six eggs were collected from four different cages, each housing a different number of birds: 20, 40, 60, and 80 birds per cage, just three weeks after the birds were moved into these conditions. These eggs were treated as the control group. Following a further six weeks of the birds living in these cages, another four sets of six eggs were collected from the same cages, in order to observe whether the population size of the cage of the laying hen affects the metabolite profiles of the yolks of eggs. All eggs underwent metabolite extraction on the day of lay, and the resulting extracts were stored at -80 °C until chemical analysis took place. Figure 8.1 shows the experimental design for this study.

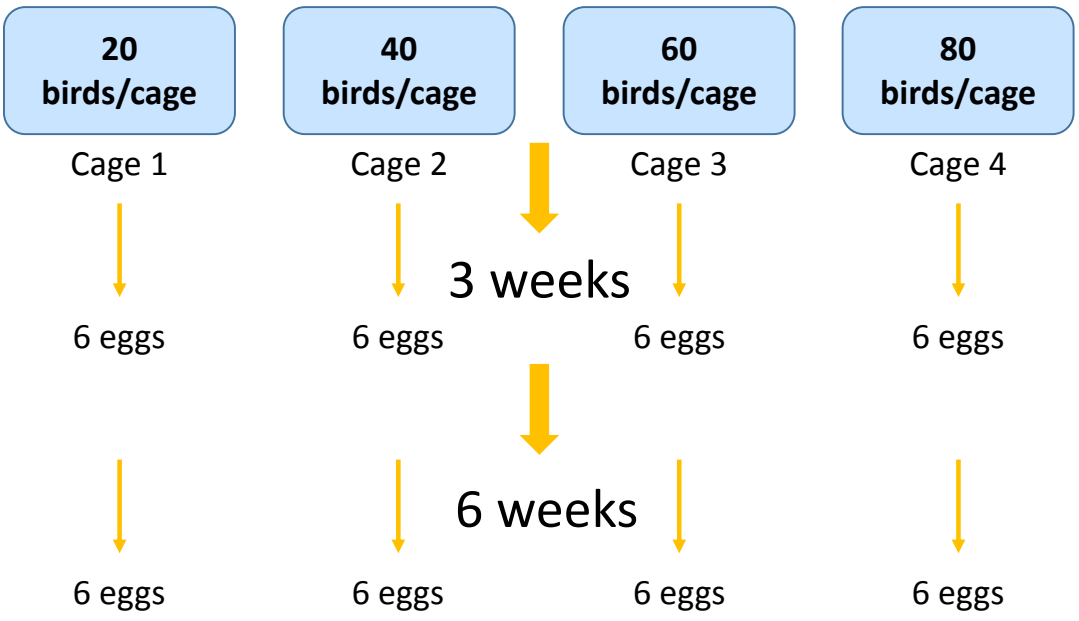


Figure 8.1: *Diagram showing the experimental design for the cage population size study*

Organic metabolite extraction of the yolk, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. Data analysis for the yolk samples from eggs that were laid nine weeks after the birds were moved into the cages was then carried out as described in Chapter 3, Section 3.9, with no attempt at compound identification. The p-values for the top significantly different compounds were then compared with the p-values from ANOVA/Welch tests for these compounds when comparing eggs laid by birds from cages of different population size after just three weeks of the birds living in the cages.

Unfortunately it was not possible to collect eggs any earlier than three weeks after the birds were moved into these cages. Therefore, any influence that the cage population size may have on the birds and their eggs may have started to affect the control sample set.

8.3.2 Stocking Density

Eggs were collected on the day of lay from the National Institute of Poultry Husbandry (Harper Adams University, Newport, U.K.). Laying hens were of the Hy-line brown breed, 52 weeks old at the start of the study, fed the same diet, and kept in enriched cages, with different stocking densities. The cages were all the same size, with different numbers of birds kept in each cage: 2, 4, 6, and 8 birds per cage. Four sets of six eggs were collected from the four different cages, just four days after the birds were moved into these conditions. These eggs were treated as the control group. Following a further ten days of the birds living in these conditions, another four sets of six eggs were collected from the same cages, in order to observe whether the stocking density of the cage of laying hens affects the metabolite profiles of the yolks of eggs. All eggs underwent metabolite extraction on the day of lay, and the resulting extracts were stored at -80 °C until chemical analysis took place. Figure 8.2 shows the experimental design for this study.

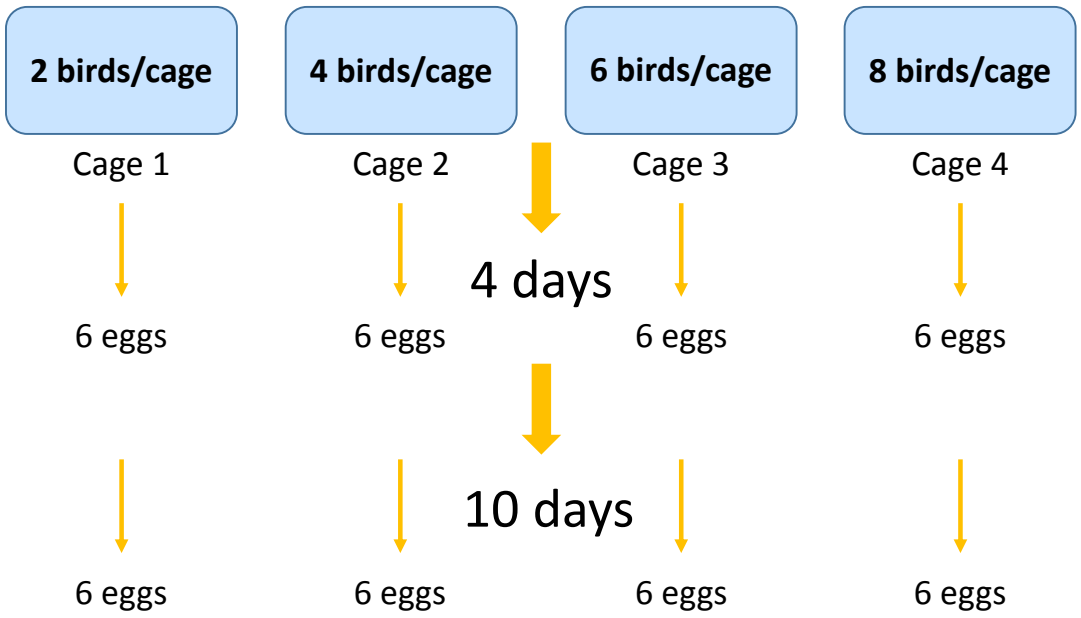


Figure 8.2: *Diagram showing the experimental design for the stocking density study*

Organic metabolite extraction of the yolk, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. Data analysis for the yolk samples from eggs laid two weeks after the birds were moved into the cages was then carried out as described in Chapter 3, Section 3.9, with the addition of a second PCA, using only the top statistically significant compounds following ANOVA/Welch tests. The p-values for the top significantly different compounds were then compared with the p-values from ANOVA/Welch tests for these compounds when comparing eggs laid by birds from cages of different stocking densities, after just four days of the birds living in the cages. These compounds were also compared to the top statistically significant compounds when comparing eggs laid by birds kept in cages with different population sizes. This was done in order to determine whether the statistical significance of these compounds was genuinely due to differences in stocking density, or whether it might be due to the difference in population size of the cages, or some other variable that may exist between the birds in the different cages, rather than the stocking density itself.

Unfortunately it was not possible to collect eggs any earlier than four days after the birds were moved into these cages, therefore the control sample set may have started to be affected by the stocking density of the laying birds. It was also not possible to continue this study for longer than two weeks, therefore the effect of the stocking density of the laying hens on the metabolite profile of egg yolk may not be fully represented by the eggs used in this study.

8.4 Cage Population Size (9 Weeks): Yolk Organic Extracts

Organic extracts of the yolks of eggs laid by hens that were kept in cages of different population size for nine weeks were compared to see whether the number of birds living together in a cage affects the metabolite profiles of the egg yolks.

8.4.1 Results and Discussion

8.4.1.1 Quality Control Analysis

Table 8.1 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 8.1: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different population size*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	1382062	129763553	50992569	37490454	345503620	103486405
QC2	1352582	124934116	54167801	35286834	336437753	145882880
QC3	1425724	124802495	50778089	34209741	323566408	76506007
QC4	1358155	124796201	52120718	34606960	327177489	72500495
QC5	1414995	125410743	52848218	33843394	320578161	74670385
QC6	1442398	126247115	49929978	33844492	319334818	71394737
QC7	1429934	116317274	46335284	33123420	330134859	110017280
SD	36278	4058459	2505209	1431568	9364436	27956772
Mean	1400836	124610214	51024665	34629328	328961873	93494027
CV%	2.59	3.26	4.91	4.13	2.85	29.90

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis. The CV% for the peak area of Peak F is only just within the 30% cut off, indicating that there was some instrumental drift throughout the analysis, but as it is still under 30%, the analysis can still be considered to be robust.

Table 8.2 shows the retention times (RTs) for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 8.2: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different population size*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.222	8.770	31.893	55.746	70.586	76.539
QC2	4.324	8.822	31.962	55.815	70.638	76.614
QC3	4.378	8.860	31.999	55.836	70.709	76.784
QC4	4.256	8.871	32.027	55.897	70.737	76.779
QC5	4.345	8.810	31.999	56.085	70.908	77.033
QC6	4.407	8.905	32.045	55.998	70.854	76.963
QC7	4.389	8.904	32.126	56.074	70.869	77.010
SD	0.070	0.050	0.072	0.133	0.123	0.194
Mean	4.332	8.849	32.007	55.922	70.757	76.817
CV%	1.61	0.57	0.23	0.24	0.17	0.25

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

8.4.1.2 Metabolite Profiling

Comparisons were made between the yolks of eggs laid by hens kept in cages of different population size, by studying the Total Ion Chromatograms (TICs) of the extracts following HPLC-MS analysis. Figure 8.3 shows overlaid TICs for the average chromatograms of organic extracts of yolk from eggs laid by hens that were kept in cages with 20, 40, 60, and 80 birds per cage.

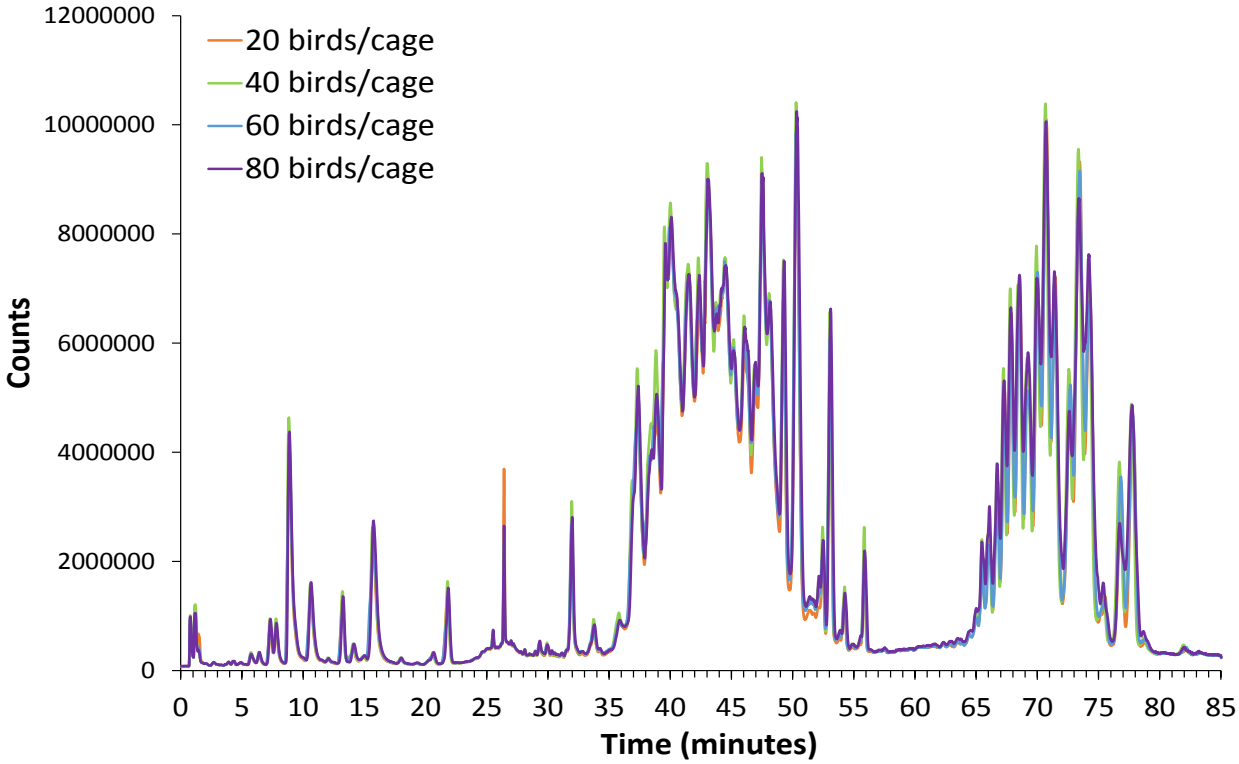


Figure 8.3: *Overlaid average TICs of organic extracts of yolks of eggs laid by hens that were kept in cages of different population size*

The overlaid TICs in Figure 8.3 show that there are some slight differences in metabolite profile between the yolks of eggs laid by birds kept in cages of different population size. However, there does not appear to be a specific trend between peak intensity and population size. There is a higher peak intensity for most peaks in the TICs of organic extracts of yolk from eggs laid by birds kept in cages of 40 and 80 birds per cage, whilst there is a lower peak intensity for most peaks in the TICs of organic extracts of yolk from eggs laid by birds kept in cages of 20 and 60 birds per cage.

8.4.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the yolks of eggs laid by hens that were kept in cages of four different population sizes.

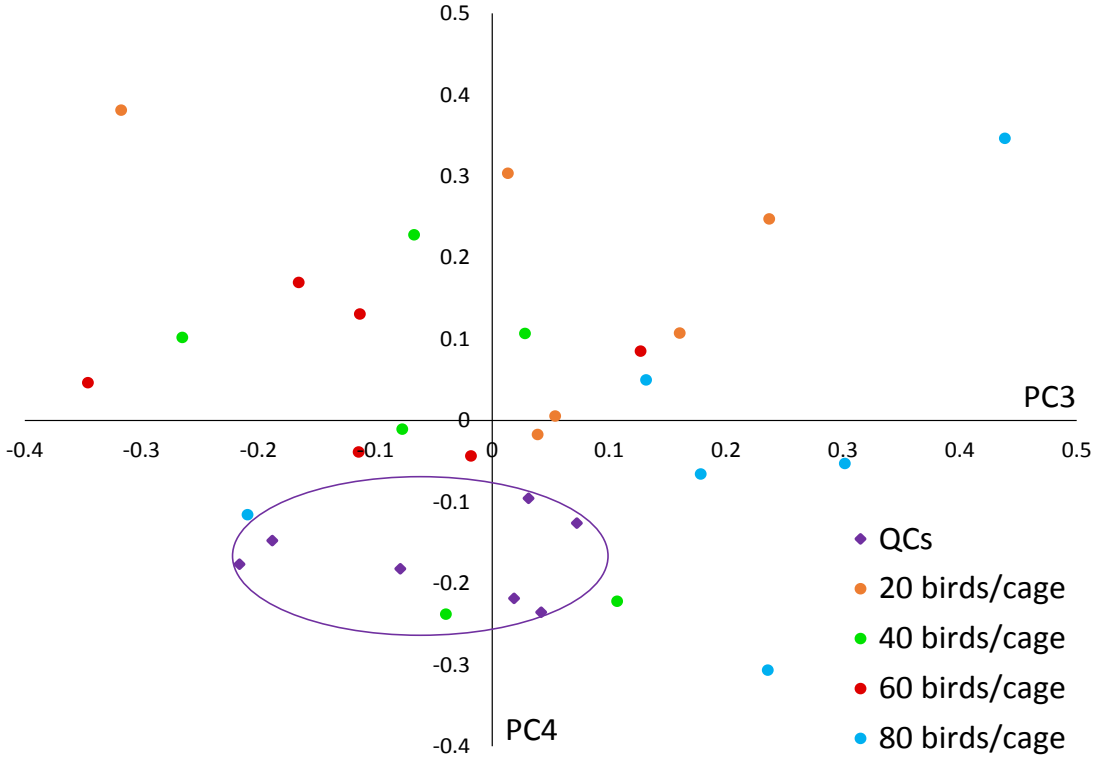


Figure 8.4: *PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds that were kept in cages of four different population sizes, including all compounds with CV%<30%. PC3 explains 0.23% of the variance, and PC4 explains 0.22% of the variance.*

The scores plot in Figure 8.4 shows no separation between the yolks of eggs laid by birds that were kept in cages of different population size. This suggests that the cage population size of the laying hen has very little effect on the metabolite profile of the egg yolk. The more highly abundant compounds that contribute to the observed higher peak intensity in the TICs of yolk from eggs laid by birds that were kept in cages of 40 and 80 birds may contribute less to the variation that is explained by PC3 and PC4 than the less abundant compounds which may experience more variation between random eggs, rather than eggs laid by birds that were kept in cages with different population sizes.

The QC samples, whilst not particularly tightly clustered, are grouped closer together than the other sample sets, showing that there was little instrumental drift affecting the analysis, and that the analysis was robust. The QC samples are likely to be more spread out than expected due to the very small amount of variation between samples. As there is such little variation between samples, the PCA has exploited any differences between the QC samples.

8.4.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 50 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, 47 were found to be statistically significant. These compounds, their CV%s, and p-values resulting from ANOVA/Welch tests, can be seen in Table E.1 in Appendix E. Those in bold were also found to be statistically significant following confirmation using the raw data.

8.4.1.4.1 Comparison of top compounds between eggs collected after 9 weeks and control eggs collected after 3 weeks of birds living in cages of different population size

Table E.2 in Appendix E shows the ANOVA/Welch test p-values of the top 34 compounds that were found to be significantly different, following confirmation using the raw data, between the yolks of eggs laid by birds that were kept in cages of different population size for nine weeks. It also shows the p-values for the abundances of these compounds when comparing the yolks of eggs laid by birds that were kept in cages of different population size after just three weeks of the birds living in these cages.

Of these 34 compounds, 7 were found to be significantly different between the yolks of eggs laid by birds that were kept in cages of different population size for just three weeks. This suggests that either the abundances of these compounds were starting to be affected

by the influence of cage population size, or that these compounds were not statistically significant due to the different population sizes of the cages, and were affected by some other variable that may have existed between the birds in the different cages.

8.5 Stocking Density (2 weeks): Yolk Organic Extracts

Organic extracts of the yolks of eggs laid by hens kept in cages of different stocking densities for two weeks were compared to see whether the stocking density of the cage of the laying hen affects the metabolite profiles of the eggs.

8.5.1 Results and Discussion

8.5.1.1 Quality Control Analysis

Table 8.3 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 8.3: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different stocking density*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	2170115	162312313	39338711	48651457	410707348	82819839
QC2	2705074	169213082	42536688	48733441	445343289	82562610
QC3	2869388	174598801	38187109	43697911	472921967	83377337
QC4	2812153	164155777	41888271	42804274	422475903	75739428
QC5	2983482	173606052	38561378	48774188	446895936	75339025
QC6	3178658	166308520	41092277	49767646	438909541	66218091
QC7	2948875	158726851	39486607	44423547	392218132	66785769
SD	318665	5849012.8	1687229	2916245	26597889	7359738
Mean	2809678	166988771	40155863	46693209	432781731	76120300
CV%	11.34	3.50	4.20	6.25	6.15	9.67

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all well below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis.

Table 8.4 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 8.4: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens kept in cages of different stocking density*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC1	4.836	9.998	33.752	57.804	72.677	79.848
QC2	4.874	10.070	33.840	57.859	72.649	79.754
QC3	4.905	10.100	33.871	57.973	72.746	80.067
QC4	4.887	10.082	33.786	57.922	72.695	79.966
QC5	4.902	10.114	33.851	57.943	72.744	80.014
QC6	4.905	10.134	33.888	58.040	72.946	80.250
QC7	4.855	10.117	33.920	58.022	72.962	80.315
SD	0.027	0.045	0.058	0.085	0.128	0.202
Mean	4.881	10.088	33.844	57.938	72.774	80.031
CV%	0.55	0.45	0.17	0.15	0.18	0.25

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well below the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

8.5.1.2 Metabolite Profiling

Comparisons were made between the yolks of eggs laid by hens kept in cages of different stocking densities, by studying the TICs of the extracts following HPLC-MS analysis. Figure 8.5 shows overlaid TICs for the average chromatograms of organic extracts of yolk from eggs laid by hens that were kept in cages with 2, 4, 6, and 8 birds per cage.

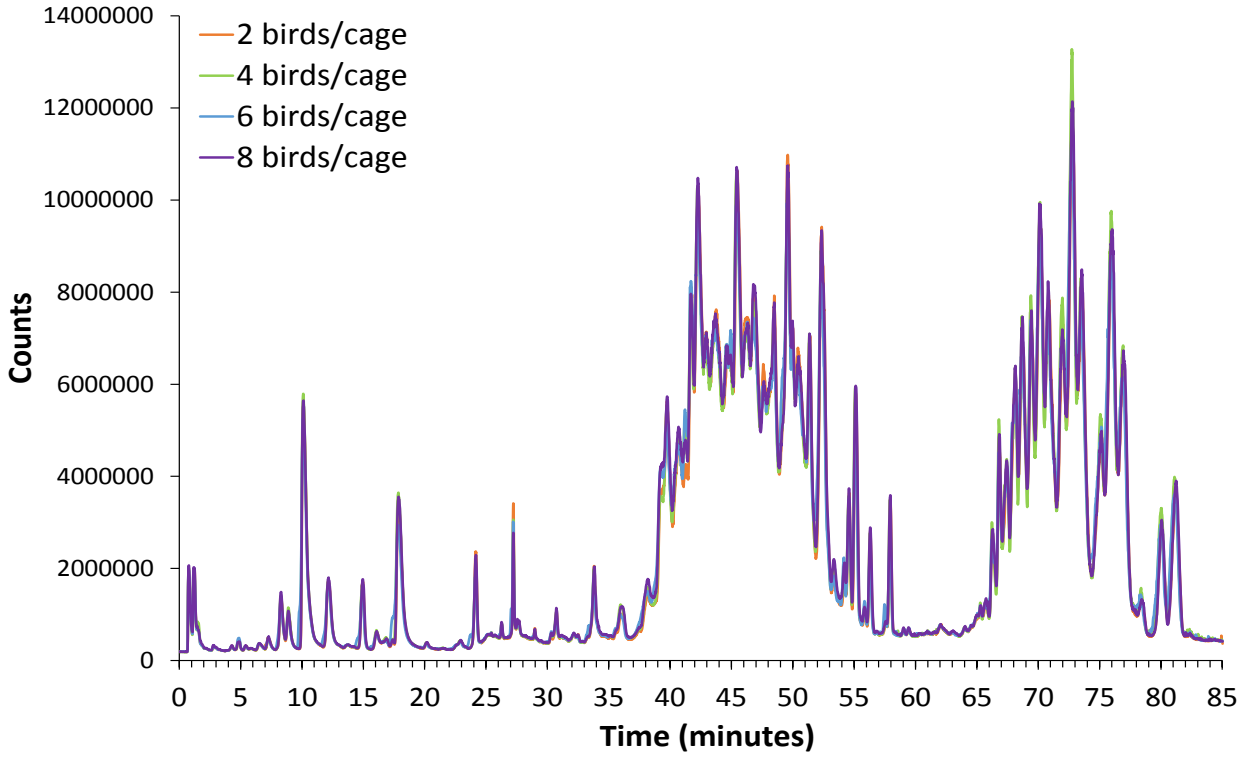


Figure 8.5: *Overlaid average TICs of organic extracts of yolks of eggs laid by hens that were kept in cages of different stocking densities*

The overlaid TICs in Figure 8.5 show very little difference in the metabolite profiles of the yolks of eggs laid by birds that were kept in cages of different stocking densities for two weeks. Most difference is seen after 65 minutes of analysis, when there is a greater peak intensity for most peaks in the TIC of the yolks of eggs laid by birds that were kept in a cage of 4 birds. However, the differences in peak intensity are still quite subtle. This suggests that the stocking density of laying birds does not have much of an impact on the metabolite profiles of the yolks of eggs laid by the birds, at least not after only two weeks of living in these conditions.

8.5.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the yolks of eggs laid by hens which were kept in cages of four different stocking densities.

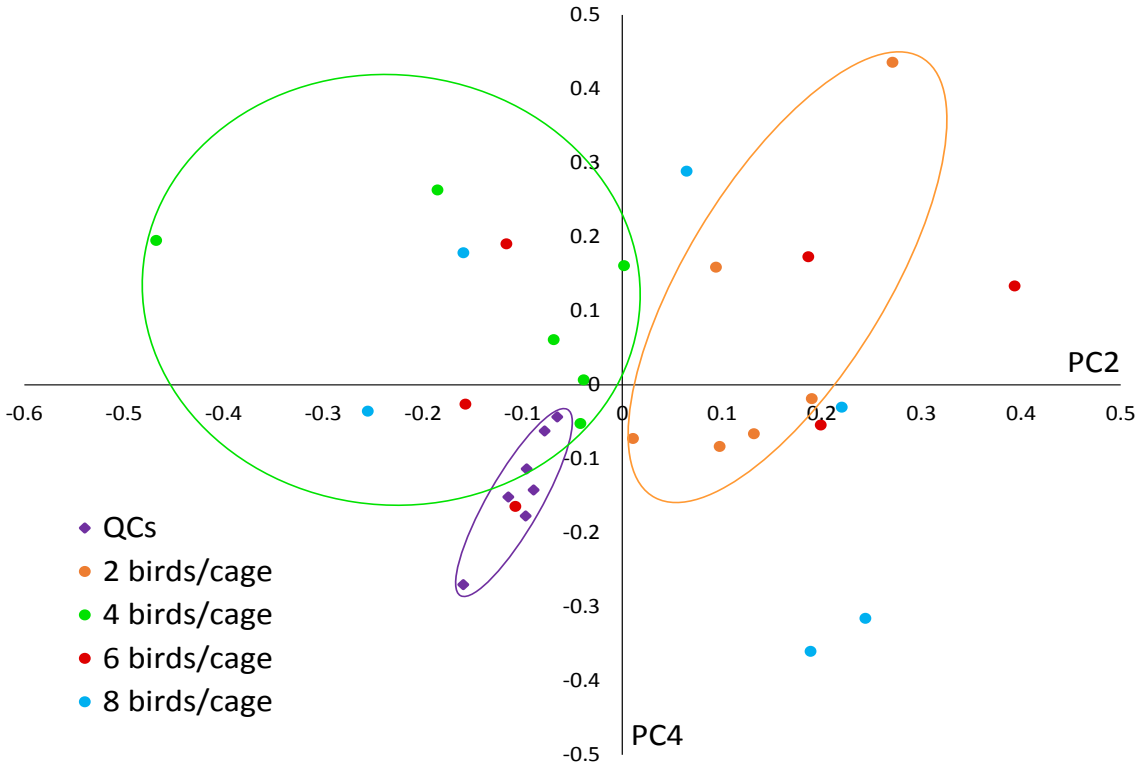


Figure 8.6: *PCA scores plot showing PC2 vs PC4 for yolk organic extracts of eggs laid by birds that were kept in cages of four different stocking densities, including all compounds with CV%<30%. PC2 explains 0.19% of the variance, and PC4 explains 0.10% of the variance.*

The scores plot in Figure 8.6 shows some slight difference between the yolks of eggs laid by birds that were kept in cages with stocking densities of 2 and 4 birds per cage, but no separation between these and the yolks of eggs laid by birds that were kept in cages with 6 and 8 birds per cage, which are spread throughout the plot. This confirms what was observed from the overlaid TICs in Figure 8.5, which showed very little difference in metabolite profile between the yolks of eggs laid by birds that were kept in cages with the four different stocking densities, with some difference in the TIC of the yolks of eggs laid by birds that were kept in a cage of 4 birds per cage, which showed a higher peak intensity towards the end of the chromatogram.

The QC samples are more tightly grouped together than the other samples, confirming that the differences between samples are due to true biological differences rather than instrumental drift.

8.5.1.4 Univariate Statistics

Following ANOVA/Welch tests and post-hoc Tukey/Games-Howell tests on the top 62 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, 59 were found to be statistically significant. These compounds, their CV%s, and p-values resulting from ANOVA/Welch tests, can be seen in Table E.3 in Appendix E. Those in bold were also found to be statistically significant following confirmation using the raw data.

Of the 40 compounds that were still found to be statistically significant when confirmed using the raw data, one was putatively identified through comparing mass spectra provided by METLIN with mass spectra from the analysis. The compound m/z 565.4040 was putatively identified as canthaxanthin, which is a type of carotenoid known as a xanthophyll. The comparison between the mass spectrum provided by METLIN and that resulting from the analysis in this study can be seen in Appendix E, Figure E.1. The observed m/z for this compound was due to the $[M+H]^+$ adduct.

Table 8.5 shows the putative identification of this compound, and the p-values resulting from ANOVA and post-hoc Tukey tests. As the probability score for the predicted formula is over 99, it indicates that there is a high probability that this formula corresponds to this compound.

The post-hoc test p-values in Table 8.5 show that there is only a significant difference in the abundance of this compound between the yolks of eggs laid by birds that were kept in a cage of 8 birds, and the yolks of eggs laid by birds that were kept in cages with lower stocking densities. This suggests that the abundance of this compound is more similar between the yolks of eggs laid by birds that were kept in cages with stocking densities of

Table 8.5: *Table showing the putative identification of compound m/z 565.4040 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
565.4040	$C_{40}H_{52}O_2$	99.55	Canthaxanthin	0.001
Post-hoc test *p*-values				
	2 birds/cage	**4 birds/cage**	**6 birds/cage**	**8 birds/cage**
2 birds/cage		0.779	0.861	0.005
4 birds/cage			0.334	0.001
6 birds/cage				0.026
8 birds/cage				

2, 4, and 6 birds per cage, and that there is a greater difference in compound abundance in the yolks of eggs laid by birds that were kept in a cage of 8 birds. This is confirmed by the bar chart in Figure 8.7, which shows that the abundance of this compound is similar between the yolks of eggs laid by birds that were kept in cages with 2 and 4 birds per cage. It also shows that the abundance of this compound in egg yolk decreases slightly when the laying hens are kept in cages with 6 birds per cage, and then decreases further when the laying hens are kept in cages with 8 birds per cage.

This bar chart shows a slight negative correlation between compound abundance and stocking density of the laying hen; the compound abundance decreases with an increasing stocking density. However, as there are only significant differences between the yolks of eggs laid by birds kept in cages of lower stocking densities and those laid by birds kept in a cage of 8 birds, this compound is not useful as a marker of stocking density.

As this study took place only two weeks after the birds were moved into these cages, repeating this study after a longer time frame may show more significant differences. It would be interesting to observe differences in the abundance of this compound between the yolks of eggs laid by birds kept in cages with larger population sizes and greater differences in stocking density, to see whether the observed trend is replicated or enhanced. This

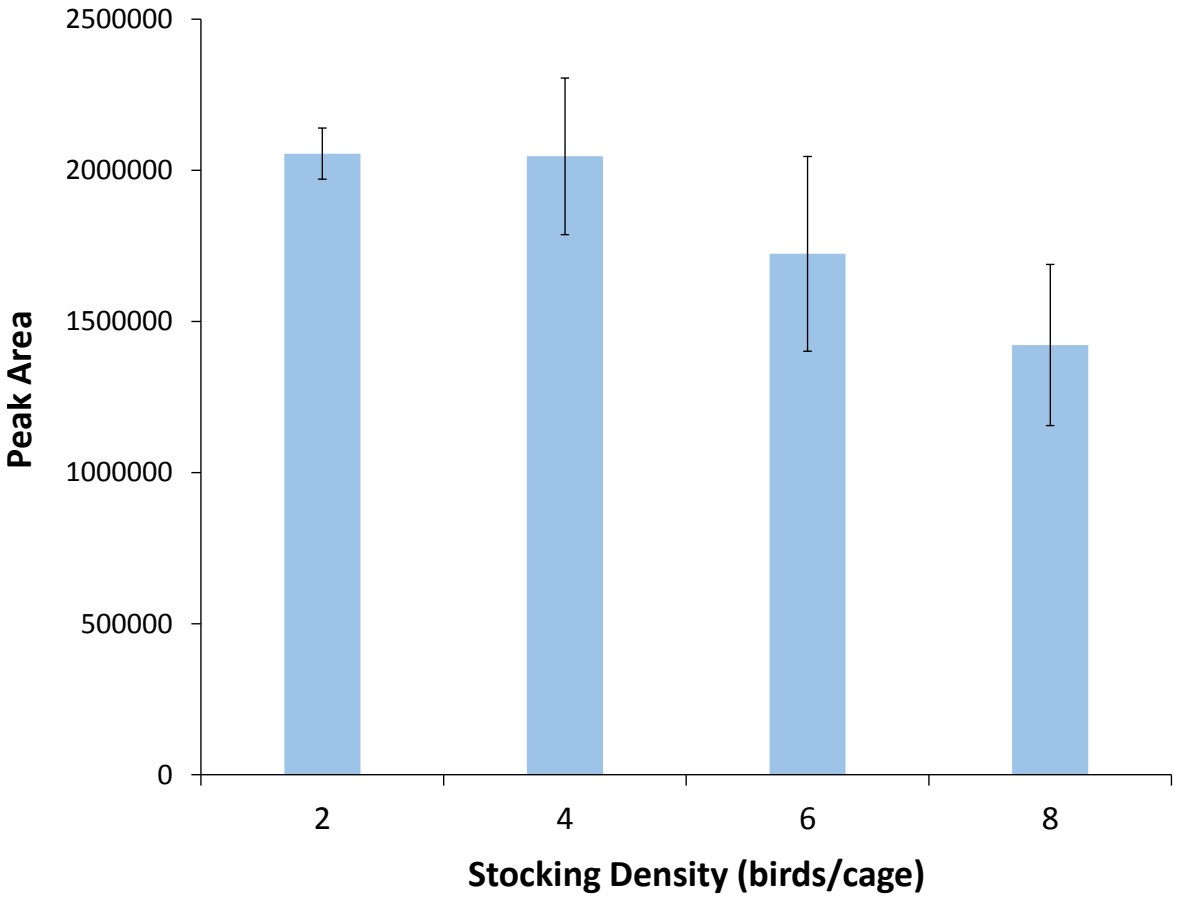


Figure 8.7: *Bar chart showing the trend in abundance of compound m/z 565.4040 between the yolks of eggs laid by hens kept in cages with different stocking densities of 2, 4, 6, and 8 birds per cage. n=6. Measured as mean values of peak area for each cage population size, with error bars ±1 standard deviation.*

could reveal the compound to be useful as a marker of stocking density, when cages with more realistic, larger population sizes and greater differences in stocking density between them, are studied.

Some of the error bars on the bar chart are quite wide, showing the large standard deviations of the compound abundances due to the natural biological variation between different birds and therefore between their eggs. The standard deviation for eggs laid by hens kept in a cage with just 2 birds is quite small, indicating that the amount of variation in the abundance of this compound is lower in eggs laid by birds kept in the cage with the lowest stocking density. This is probably because there are fewer birds for the eggs to originate from, and therefore less variation between eggs resulting from the

variation between laying birds.

Canthaxanthin is a carotenoid pigment that is commonly used as an additive in poultry feed, in order to produce a more intensely coloured egg yolk [6]. Carotenoids, and therefore canthaxanthin, are not endogenous to poultry, therefore all canthaxanthin found in the egg is due to canthaxanthin absorbed from the hen's feed [7]. There is a direct relationship between dietary levels of canthaxanthin and its deposition in egg yolk [8].

As the birds were all fed the same feed, the difference in the abundance of this compound, between the yolks of eggs laid by birds that were kept in cages of different stocking densities, could be due to different rates or different amounts of absorption of this compound from the feed, or different rates of metabolism of this compound once it has been absorbed, which may be affected by the stocking density of the birds. It could also be due to birds kept in cages with higher stocking densities perhaps eating less food, possibly due to more competition, or a higher stress level, therefore receiving less canthaxanthin from the feed, resulting in a lower abundance of this exogenous compound in their eggs.

8.5.1.4.1 Top compounds: Multivariate Statistics

A second PCA was carried out using only the top 59 statistically significant compounds, and scores plots were again produced to display the variation between the yolks of eggs that were laid by birds which were kept in cages with different stocking densities.

This scores plot in Figure 8.8 shows some more separation between the yolks of eggs laid by birds that were kept in cages of different stocking densities, than was observed in the previous scores plot in Figure 8.6. Although there are no distinct groups on the plot, the spread of the samples can be seen to be developing into some groupings based on the stocking densities of the laying hens. The plot shows a complete separation between the yolks of eggs laid by birds that were kept in cages of 2 and 6 birds per cage. The yolks of eggs laid by birds that were kept in cages with 6 and 8 birds per cage show a wider spread across the plot than the yolks of eggs laid by birds that were kept in cages with

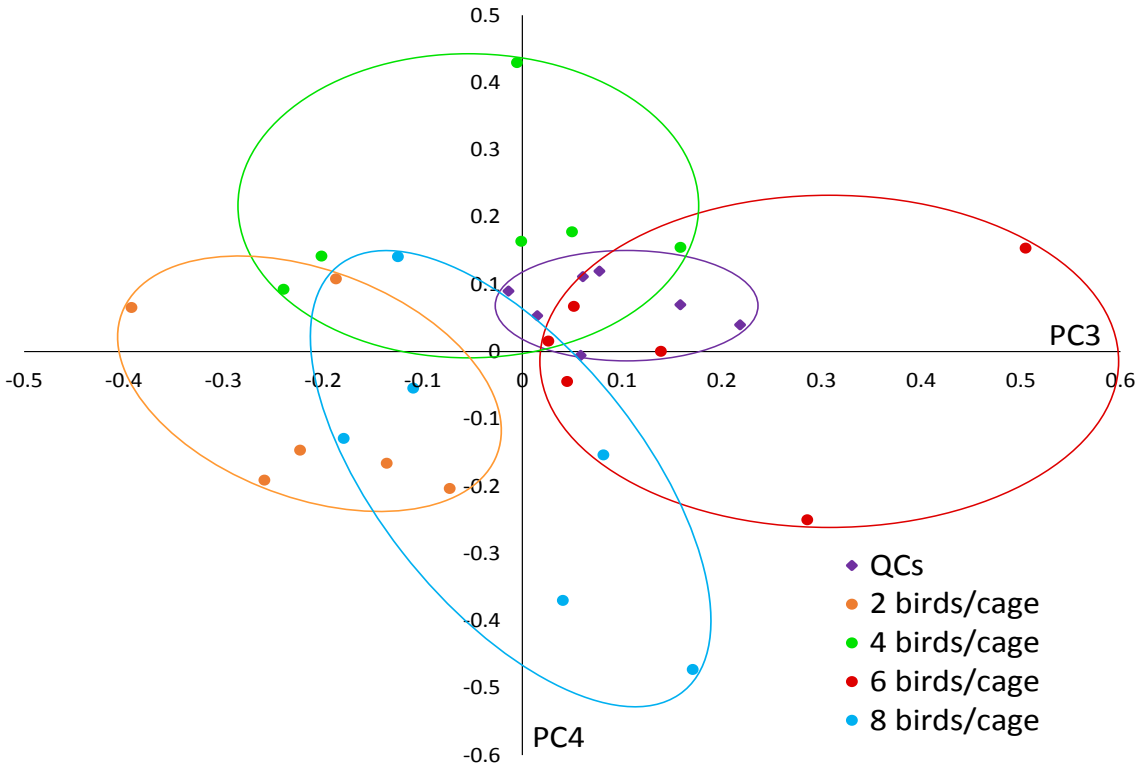


Figure 8.8: *PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds that were kept in cages of four different stocking densities, including the top 59 statistically significant compounds. PC3 explains 0.01% of the variance, and PC4 explains 0.005% of the variance.*

2 and 4 birds per cage. This suggests that there may be more variation in metabolite profile between the yolks of eggs laid by birds kept in cages of higher stocking densities. This is probably due to the fact that there is an increased likelihood that the eggs were laid by different birds, therefore the eggs show the variation due to the natural biological differences between the birds themselves.

The variation described by PC3 and PC4 is just 0.01% and 0.005% respectively, showing just how subtle the differences in the metabolite profile of egg yolk are, between different stocking densities of the laying hens. The QC samples are grouped more tightly than the other samples, confirming that the analysis was robust and that there was little instrumental drift affecting the analysis.

8.5.1.4.2 Comparison of top compounds in eggs collected after 2 weeks of birds living in cages of different stocking densities, with control eggs collected after 4 days, and eggs laid by birds from cages with different population size

Table E.4 in Appendix E shows the ANOVA/Welch test p-values of the top 40 compounds that were found to be significantly different, following confirmation using the raw data, between the yolks of eggs laid by birds that were kept in cages of different stocking densities for two weeks. It also shows the p-values for the abundances of these compounds when comparing the yolks of eggs laid by birds that were kept in cages of different stocking densities after just four days of the birds living in these cages.

Of these 40 compounds, 21 were found to be significantly different between the yolks of eggs laid by birds that were kept in cages of different stocking densities for just four days. This suggests that either the abundances of these compounds were affected very soon after the birds were moved into these cages of different stocking densities, or that these compounds were not statistically significant due to the different stocking densities of the cages, and were affected by some other variable that may have existed between the birds in the different cages.

Very few of the top 40 compounds, and only one of the 19 compounds that were not found to be statistically significant after just four days of the birds living in these cages, were found to also show significant differences in abundance between eggs laid by birds in cages of different population size. This suggests that most of the observed differences in metabolite profile between the yolks of eggs laid by birds in cages of different stocking densities were not due to differences in population size, and that the statistical significance of the remaining 18 compounds is likely to be due to differences in the stocking density of the cages, rather than population size or some other variable between the birds in the different cages.

The compound that was putatively identified as canthaxanthin was not statistically

significant after just four days of the birds living in cages with different stocking densities, nor between eggs laid by birds from cages of different population sizes, and so it is likely that the stocking density of the cage of the laying hen does affect the abundance of this compound in the egg yolk.

The bar chart in Figure 8.9 shows the differences in the abundance of the compound putatively identified as canthaxanthin, between the yolks of eggs laid by birds from cages of different stocking densities, after just four days of them living in these cages.

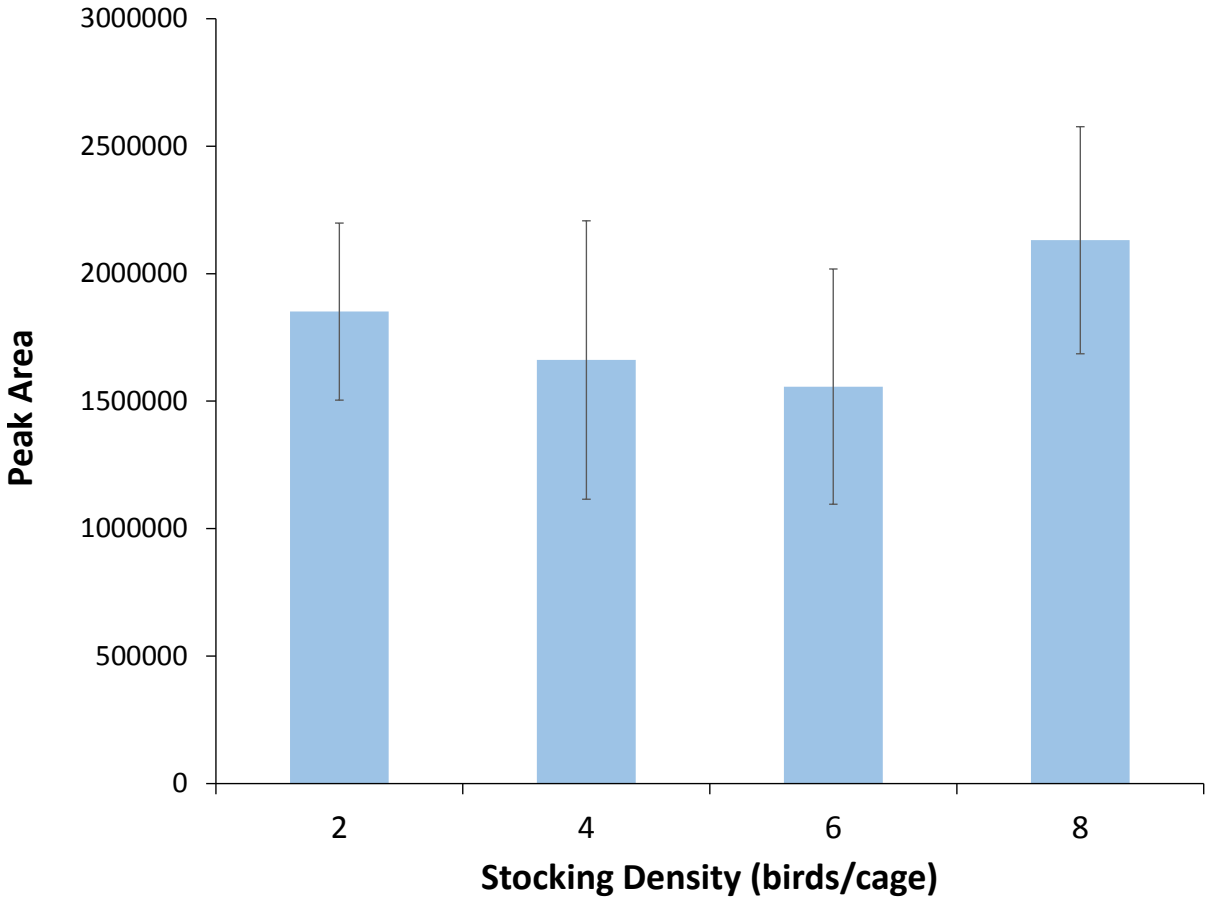


Figure 8.9: *Bar chart showing the trend in abundance of compound m/z 565.4040 between the yolks of eggs laid by hens kept in cages with different stocking densities of 2, 4, 6, and 8 birds per cage, after just four days of living in these cages. n=6. Measured as mean values of peak area for each cage population size, with error bars ±1 standard deviation.*

Although the error bars in Figure 8.9 are quite wide, showing the variation in the abundance of this compound within the yolks of eggs laid by birds from the same cage,

there is a slight trend which is similar to the trend shown in the bar chart in Figure 8.7, after two weeks of the birds living in these cages. This trend shows a negative correlation between compound abundance in egg yolk and stocking density of the laying birds; as the stocking density increases, the compound abundance decreases.

However, the average abundance of this compound in the yolks of eggs laid by birds from a cage containing 8 birds, after four days of the hens living in these cages, is higher than in the yolks of eggs laid by birds from a cage containing 2 birds, which breaks the trend. It could be that after only four days of the birds living in these cages there was no real trend between cage stocking density and compound abundance, as there had not been enough time for stocking density to have an effect on the birds and their eggs, and it is coincidence that the trend appears to be similar. However, it could also be that the stocking density was starting to have an effect on the abundance of this compound in the egg yolk, but not enough for there to be significant differences in the abundance, and not enough for the trend to follow through all of the stocking densities. Following a further ten days of the birds living in these cages, the influence of the stocking density on the abundance of this compound was strong enough to have affected the yolks of eggs laid by birds from cages of all different stocking densities, and therefore the negative correlation between compound abundance and stocking density was observed, as well as some significant difference in compound abundance.

8.6 Summary & Conclusions

8.6.1 Summary

The overlaid TICs in Section 8.4 showed a very small difference in the metabolite profiles of the yolks of eggs laid by birds that were kept in cages with different population sizes for nine weeks, and several compounds were found to show significant differences in abundance between these eggs. However, the scores plot did not show any difference

in metabolite profile between the eggs laid by birds from cages of different population size. Although the overlaid TICs showed a small difference in metabolite profile, and several compounds showed some statistical significance, it can be concluded that the cage population size of the laying hen has very little effect on the metabolite profile of the egg yolk.

In Section 8.5, the overlaid TICs and PCA scores plots showed only a small amount of difference in the metabolite profiles of the yolks of eggs laid by birds that were kept in cages with different stocking densities. However there was an observable difference, even after only two weeks of the birds living in these cages. Very few statistically significant compounds were also found to show statistical significance due to cage population size, which indicates that the observed differences in metabolite profile between the yolks of eggs laid by birds kept in cages of different stocking densities are not due to differences in population size between the cages.

One compound that was statistically significant when comparing the yolks of eggs laid by birds that were kept in cages of different stocking densities was putatively identified as canthaxanthin. This compound was not found to be statistically significant after just four days of the birds living in these cages, or between eggs laid by birds living in cages of different population size, indicating that the differences in abundance are likely to be due to differences in the stocking density of the laying birds. There was a slight negative correlation between the abundance of this compound in egg yolk and the stocking density of the cage of the laying hen, however the only significant differences in abundance were between the yolks of eggs laid by birds that were kept in a cage with 8 birds, and those laid by birds that were kept in cages with lower stocking densities. Therefore, although there is a slight trend between compound abundance and stocking density, as there are so few points of statistical significance between eggs laid by birds from cages of different stocking densities, this compound is not suitable as a marker of stocking density when such small cage population sizes and small differences in stocking density are studied.

8.6.2 Conclusions

This study has shown that the population size of the cage of the laying hen has very little effect on the metabolite profile of the yolk of eggs, and the effects that it does have, and the compounds that are affected, are mostly different to those affected by stocking density. Although the stocking density of the cage of the laying hen also has only a small effect on the metabolite profile of the yolk, there are several compounds that have significantly different abundances in the yolks of eggs laid by birds from cages of different stocking densities. One of these compounds was putatively identified as canthaxanthin. Further research, studying larger cage population sizes, and greater differences in stocking density between these cages, may show bigger differences in the abundance of this compound between the yolks of eggs laid by birds from cages with different stocking densities. This would reveal this compound to be useful as a marker of stocking density, in more realistic scenarios with greater cage population sizes of the laying birds.

Further studies, with larger samples sizes and eggs collected after a longer time frame, could then be carried out in order to confirm the statistical significance of the compound putatively identified as canthaxanthin, as well as the correlation between its abundance in egg yolk and the cage stocking density of the laying hen. If these results were confirmed, and shown to be repeatable with different ages and breeds of birds, and the identification of this compound was confirmed, and quantification carried out, then the concentration of canthaxanthin in egg yolk could be used to distinguish between eggs laid by birds from cages of different stocking densities. This could be used to determine whether a cage egg was laid by a bird in an enriched cage or an illegal battery cage, therefore helping to detect cases of fraud regarding poultry egg production.

8.7 References

[1] *Council Directive 1999/74/EC of 19 July 1999, Laying down minimum standards for the protection of laying hens*, Official Journal of the European Communities, L203/53-57, 1999.

[2] A. Saki, P. Zamani, M. Rahmati, and H. Mahmoudi, "The effect of cage density on laying hen performance, egg quality, and excreta minerals," *Journal of Applied Poultry Research*, vol. 21, pp. 467–475, 2012.

[3] H. K. Kang, S. B. Parks, S. H. Kim, and C. H. Kim, "Effects of stock density on the laying performance, blood parameter, corticosterone, litter quality, gas emission and bone mineral density of laying hens in floor pens," *Poultry Science*, vol. 95, pp. 2764–2770, 2016.

[4] M. Sarica, S. Boga, and U. S. Yamak, "The effects of space allowance on egg yield, egg quality and plumage condition of laying hens in battery cages," *Czech Journal of Animal Science*, vol. 8, pp. 346–353, 2008.

[5] J. Carey, F. Kuo, and K. Anderson, "Effects of cage population on the productive performance of layers," *Poultry Science*, vol. 74, pp. 633–637, 1995.

[6] Y. Nys, "Dietary carotenoids and egg yolk coloration," *Archiv für Geflügelkunde*, vol. 64, pp. 45–54, 2000.

[7] Y. Nys and N. Guyot, "Improving the safety and quality of eggs and egg products," in. Woodhead Publishing Limited, 2011, vol. 1: Egg Chemistry, Production and Consumption, ch. 4: Egg formation and chemistry, pp. 83–132.

[8] *Opinion of the scientific committee on animal nutrition on the use of canthaxanthin in feedingstuffs for salmon and trout, laying hens, and other poultry*, European Commission, 2002.

9. Differences in metabolite profile between the yolk and albumen of eggs laid by hens from different housing systems

9.1 Introduction

According to European Union (EU) legislation, shell eggs fall into four different categories regarding the farming method used in their production [1]. There are eggs produced from conventional farming methods: eggs from hens kept in enriched cages (cage eggs), eggs from barn hens (barn eggs), and eggs from free range hens (free range eggs), as well as eggs produced from organic farming systems (organic eggs) [1]. The minimum requirements for the production of eggs by conventional farming methods are specified in the Council Directive 1999/74/EC [2], and the requirements of organic production are set out in Council Regulation (EC) No 834/2007 [3], including the use of organic, non-genetically modified feed. The main differences between the four housing systems can be seen in Table 9.1.

All eggs must have the farming method clearly labelled on the packaging, which is usually in the form of a code: 0 refers to organic eggs, 1 to free range eggs, 2 to barn eggs,

Table 9.1: *Table showing the main requirements for the four different housing systems*

Cage	Barn	Free range	Organic
600cm^2 per bird	9 birds/m^2	9 birds/m^2	6 birds/m^2
no access to open runs	no access to open runs	access to open runs	access to open runs

and 3 to eggs from enriched cages [1]. According to the U.K. Egg Statistics released from the Department for Environment, Food, and Rural Affairs (DEFRA), the average U.K. farm-gate egg price is approximately 30p greater for free range eggs than for cage eggs [4]. The farm-gate prices of barn and organic eggs are kept confidential. By comparing the supermarket prices of eggs from different housing systems, it can be easily observed that cage eggs are the cheapest, whilst organic eggs fetch the highest price. This creates the temptation to mislabel eggs with false farming methods, in order to sell them at a higher price and gain a greater profit. As there is no way for the consumer to confirm whether the eggs that they purchase actually originate from the farming method of production that is stated on the label, it would be easy for eggs to be mislabelled with regards to their housing system, with a low likelihood of the fraud being uncovered. In 2018, an individual from the U.K. was fined a total of £505,381 for selling barn eggs as free range eggs in order to make a profit [5]. In the same year, another individual from the U.K. was given a suspended jail sentence for selling eggs that they falsely claimed to be free range [6].

There is currently no scientific testing of eggs, in order to determine whether they have been correctly labelled with regard to their method of production. Cases of fraud are only discovered through whistleblowing, or by inspectors noticing inconsistencies in the paperwork. Implementing a random testing of eggs at any point during the production process, from laying to selling, would deter farmers, or food business operators, from falsely labelling the eggs, as the risk of being caught would increase. Clearly, there would be a benefit to having a robust scientific method capable of predicting the housing system

of the laying birds, from the eggs themselves, which could help to identify cases of fraud.

There has been much research into what effect the housing system of the laying bird has on the resulting eggs, including studying both physical and chemical characteristics of the eggs. However, the results of studies investigating the effects of housing system on the physical characteristics of the egg are quite varied, and sometimes contradictory.

Some studies have found that the Haugh Unit (HU) of the egg (which is based on the relationship between the weight of the intact egg, and the albumen height once the egg has been broken [7]) is significantly different between eggs from different housing systems [8][9][10]. However, these studies produced contrasting results regarding whether the HU is higher or lower between eggs from certain housing systems. Another study found that the HU of eggs varies with the breed of the laying hen, and reported that one breed of bird did not show any statistical significance in HU between eggs from different housing systems [11]. This suggests that the HU of eggs laid by birds of different breeds may be affected to different extents by the housing system of the birds. Another study did not observe any significant difference in HU between eggs from different housing systems [12], which could be due to the breed of bird being used in the study.

The albumen height itself has also been compared between eggs from different housing systems. One study found that the albumen height was significantly greater in cage eggs than barn eggs [8], however other studies found no significant difference in albumen height between eggs from different housing systems [10][11][13]. The shape index (ratio between width and length) of the eggs has also been compared between eggs from different housing systems, with contrasting results reported. Some studies found the shape index to be greater for free range eggs than for cage eggs [9][13], whereas another found it to be greater for cage eggs than for free range and barn eggs [12]. Other studies observed no significant difference in shape index between eggs from different housing systems [8][10][11].

The colour of the yolk is another physical characteristic of the egg that has been compared between eggs from different housing systems. The yolk has been found to be

darker in free range eggs compared to cage and barn eggs, which may be due to the potential for the free range birds to forage for other foodstuffs, potentially consuming food that is more rich in xanthophylls, which influence yolk colour [13]. Organic eggs have been found to have a darker yolk than free range eggs, which again could be due to differences in diet [11]. One study reported no significant difference in yolk colour between cage and barn eggs [8], which is likely to be due to the fact that neither cage or barn systems allow outdoor access, so there is little potential for foraging. However, another study found that cage eggs had the darkest yolks, followed by free range, then barn, then organic egg yolks, which contradicts these other results [10]. Yet another study found no significant difference in yolk colour observed between barn, free range and organic eggs [12].

As well as the physical characteristics of eggs laid by hens from different housing systems, there have also been some mixed and contradictory results from studies comparing the chemical characteristics of eggs. Cholesterol content has been found to be significantly different between eggs from different housing systems. One study reported that cholesterol content was higher in organic eggs than cage eggs [14], whilst another found that it was higher in cage eggs than free range eggs [15]. Other studies have found that the housing system of the laying bird has no significant effect on the cholesterol content of the egg [16][17][18]. The lipid content of eggs was found to be greater in free range eggs than in cage eggs [16], however other studies reported no statistical significance in lipid content between eggs from different housing systems [10][15].

The fatty acid profile of egg yolks has been able to successfully predict the production method of eggs, between cage and organic farming systems [19]. However, other studies have found only a slight significant difference [10] or very little difference [18][20] in the fatty acid compositions of eggs from different housing systems.

The vitamin content of eggs has also been studied to see how it differs in eggs from different housing systems. Retinol (vitamin A), various tocopherols (vitamin E), and

cholecalciferol (vitamin D_3) have been found to be significantly different between the yolks of eggs from cage and organic farming systems, with most vitamins showing a higher abundance in cage eggs [14]. However, another study found that neither vitamin A or vitamin E showed a significant difference between eggs from different housing systems [17].

Significant differences in the presence of some carotenoids in egg yolk have been observed when comparing eggs laid by birds from different housing systems [16][17]. A method based on the carotenoid profiling of egg yolks has been shown to successfully classify barn, free range, and organic eggs, due to different compositions of carotenoids in the egg yolk [21].

The trace mineral content of eggs from different housing systems has been found to be significantly different [22], and the nitrogen isotope composition of eggs has been successfully used to discriminate free range eggs, from cage and barn eggs [23].

There has been much research into how the housing system of the laying hen affects the resulting eggs, and how the characteristics of an egg can be used to determine the farming method of production. However, a metabonomic approach would allow a wide range of compounds and compound classes to be studied, in order to gain as much information about the differences between these eggs as possible, creating the potential for the development of a new method to discriminate between eggs from different housing systems.

9.2 Aims and Objectives

Having a robust scientific method capable of predicting the housing system of the laying hen, from the egg itself, would clearly be of use in detecting cases of fraud regarding the mislabelling of eggs with false farming methods.

This work aimed to use a metabonomic workflow to uncover the differences in metabolite profile between eggs from different housing systems. It initially explored the differences between the yolks of eggs from cage and barn housing systems, and attempted to identify any compounds that showed potential as markers of these two housing systems[†]*. The same process was then carried out for the albumen of eggs from cage and barn housing systems. A similar workflow was then tentatively carried out to uncover potential differences in metabolite profile between eggs from cage, barn, free range, and organic housing systems, for both egg yolk and albumen.*

[†]This research has been published in *Food Control*[*].

[*]A. E. Johnson, K. L. Sidwick, V. R. Pirgozliev, A. Edge, and D. F. Thompson, "The use of metabonomics to uncover differences between the small molecule profiles of eggs from cage and barn housing systems", *Food Control*, vol. 100, pp. 165-170, 2019.

9.3 Experimental

Eggs were collected on the day of lay from Oaklands Farm Eggs Ltd. (Shrewsbury, U.K.). It was not possible to collect eggs laid by birds of the same age and breed from all four housing systems, therefore eggs from some housing systems were laid by birds of different ages and breeds. The diets were kept the same for laying hens in the cage and barn housing systems, but may have been different for the free range birds. The hens kept in the organic housing system were fed a diet consisting of 95% organic feed, with 5% non-organic protein feed.

Table 9.2: *Table showing the age and breed of birds from the four different housing systems*

	Cage	**Barn**	**Free range**	**Organic**
Breed	Novogen	Novogen	Lohmann Classic	Hy-line brown
Age	50 weeks	50 weeks	73 weeks	71 weeks

Four sets of six eggs were collected from each of the four different housing systems. All eggs underwent metabolite extraction on the day of lay, and the resulting extracts were stored at -80 °C until chemical analysis took place.

Initially, a metabonomic workflow was only applied to the comparison of the yolk and albumen of eggs laid by birds kept in cage and barn housing systems, as these birds were the only ones of the same age and breed, therefore only these eggs were directly comparable. Organic metabolite extraction of the yolk and albumen of cage and barn eggs, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing were carried out as described in Chapter 3. However, the data analysis workflow was slightly different and was carried out as described in Figure 9.1.

The removal of any compounds with a CV% >30%, followed by PCA and the production of scores plots, was as described in Chapter 3, Section 3.9. Following PCA, the compounds were ranked from the highest to lowest based on either the loadings from the

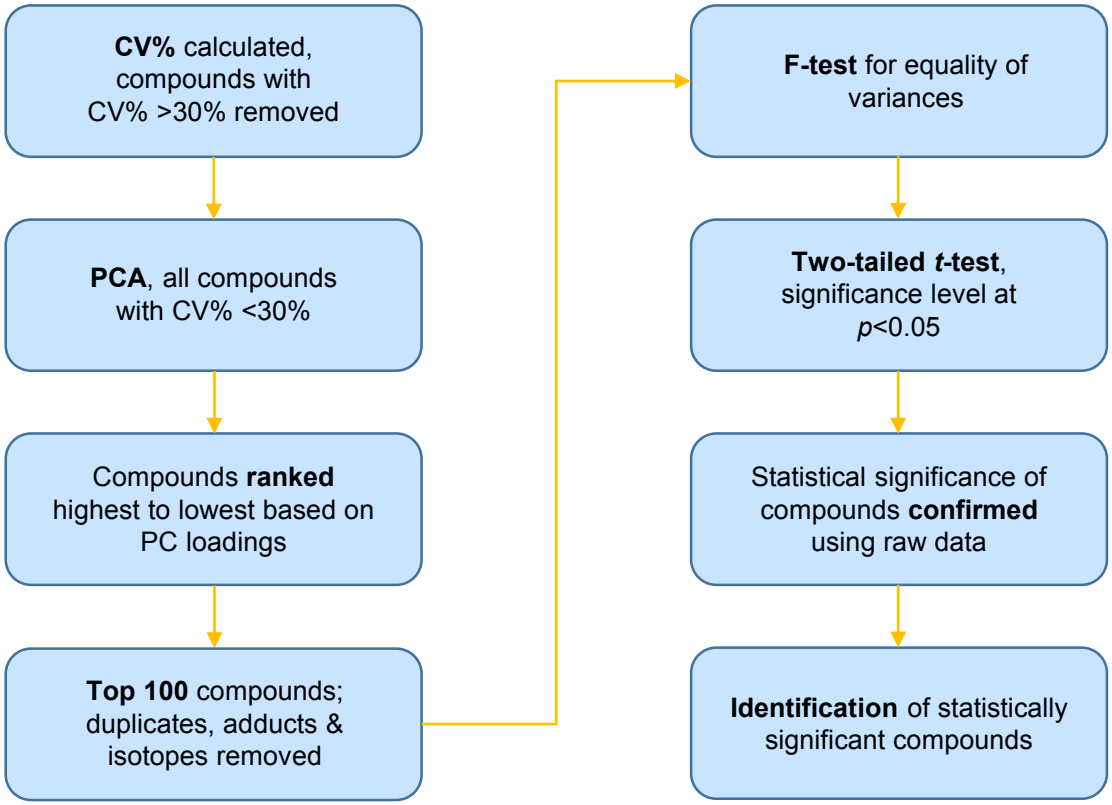


Figure 9.1: *Diagram showing the data analysis workflow that was implemented in the study comparing barn and cage eggs*

PC which showed the greatest separation between cage and barn eggs on the PCA scores plots, or if this was not obvious, PC1. The top 100 of these compounds were taken and any duplicates, adducts, and isotopes were removed. F-tests were carried out on the remaining compounds to test the equality of variances, and corresponding two tailed t-tests were then carried out. Compounds that were found to be statistically significant when comparing eggs laid by cage and barn hens then had their significance confirmed using the raw data. Identification of the compounds that were still found to be statistically significant then took place.

A second, tentative study was carried out, applying a metabonomic workflow to the observation of differences between both the yolk and albumen of eggs from all four housing systems, in order to determine whether differences in the metabolite profiles of the yolk and albumen can be detected between cage, barn, free range and organic eggs. Birds from

free range and organic production systems were of different breeds and ages to each other, and to the birds from cage and barn housing systems, and birds from free range systems may also have been fed a different diet. Therefore, any differences that were observed between eggs laid by these birds may not be due to the different housing systems, but due to the uncontrolled variables.

Organic metabolite extraction of the yolk and albumen of eggs from all four housing systems, chemical analysis of the resulting extracts using HPLC-MS, quality control analysis, and data pre-processing and analysis were carried out as described in Chapter 3.

9.4 Cage and Barn eggs: Yolk Organic Extracts

Organic extracts of the yolks of eggs laid by hens from cage and barn housing systems were compared in order to observe the differences in metabolite profile between eggs from these two housing systems.

9.4.1 Results and Discussion

9.4.1.1 Quality Control Analysis

Table 9.3 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the standard deviation (SD), the mean, and the CV% of the peak areas for each of the peaks.

Table 9.3: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens from cage and barn housing systems*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC 1	2464400	134201096	58993546	42417371	418368179	92921477
QC 2	2396322	137373301	55329453	42728536	412572398	96043702
QC 3	2349150	138831953	56034065	44263881	411815527	116765777
QC 4	2655438	141451706	58988882	46407224	439647546	157504670
QC 5	3127109	136429229	56822472	42746257	440913920	149375041
QC 6	3075148	132939364	57515538	43553547	424756538	97694791
QC 7	2767749	129184017	50562190	42814004	433675022	90268187
SD	316217	4051986	2891594	1403435	12307641	28126336
Mean	2690759	135772952	56320878	43561546	425964161	114367664
CV%	11.75	2.98	5.13	3.22	2.89	24.59

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that there was little instrumental drift occurring throughout the analysis. The CV% for the peak area of Peak F is quite high, indicating some instrumental drift, but as it is still within the 30% cut-off, the analysis can still be considered robust.

Table 9.4 shows the retention times (RTs) for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 9.4: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of yolk from eggs laid by hens from cage and barn housing systems*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC 1	4.545	9.292	32.555	56.708	72.984	80.454
QC 2	4.540	9.288	32.560	56.695	73.179	80.731
QC 3	4.566	9.247	32.486	56.619	72.574	79.712
QC 4	4.665	9.413	32.768	56.920	73.021	80.375
QC 5	4.547	9.712	32.663	56.682	72.883	80.419
QC 6	4.583	9.248	32.503	56.668	73.056	80.658
QC 7	4.596	9.310	32.616	56.784	73.168	80.820
SD	0.044	0.165	0.098	0.099	0.207	0.367
Mean	4.577	9.359	32.593	56.725	72.981	80.453
CV%	0.96	1.77	0.30	0.17	0.28	0.46

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all within the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis. The CV% for the retention time of Peak B is quite high, but as it is still under 2%, the retention time drift is still within acceptable limits.

9.4.1.2 Metabolite Profiling

Comparisons were made between the yolks of eggs laid by hens from cage and barn housing systems, by studying the Total Ion Chromatograms (TICs) of the extracts following HPLC-MS analysis. Figure 9.2 shows overlaid TICs for the average chromatograms of organic extracts of yolk from eggs laid by hens from these two housing systems.

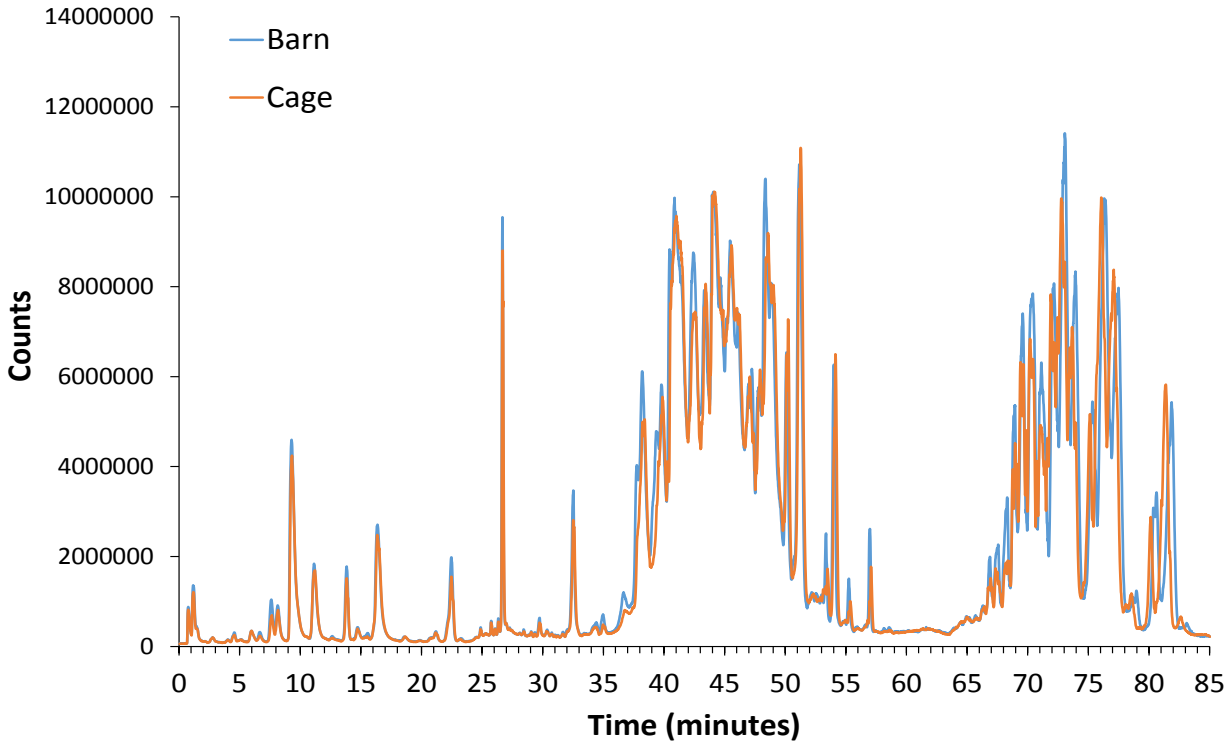


Figure 9.2: *Overlaid average TICs of organic extracts of yolks of eggs laid by hens from cage and barn housing systems*

The overlaid TICs in Figure 9.2 show some slight differences in the metabolite profiles of the yolks of eggs laid by birds from cage and barn housing systems. Up until approximately 35 minutes of analysis, there is a lower peak intensity for all peaks in the TIC of the yolks of cage eggs, compared to the TIC of the yolks of barn eggs, suggesting that cage egg yolks have a lower abundance of the less non-polar compounds than the yolks of barn eggs. Throughout the rest of the chromatogram, most peaks still have a higher intensity in the TIC of barn egg yolk, however some peaks do show a higher intensity in the TIC of cage egg yolk. This suggests that there may be less of a trend in the abundances of

the more non-polar compounds, between the yolks of cage and barn eggs. These slight differences in the TICs of cage and barn egg yolks suggest that there are differences in the metabolite profiles of the yolks of eggs from cage and barn housing systems.

9.4.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the yolks of eggs laid by hens from cage and barn housing systems.

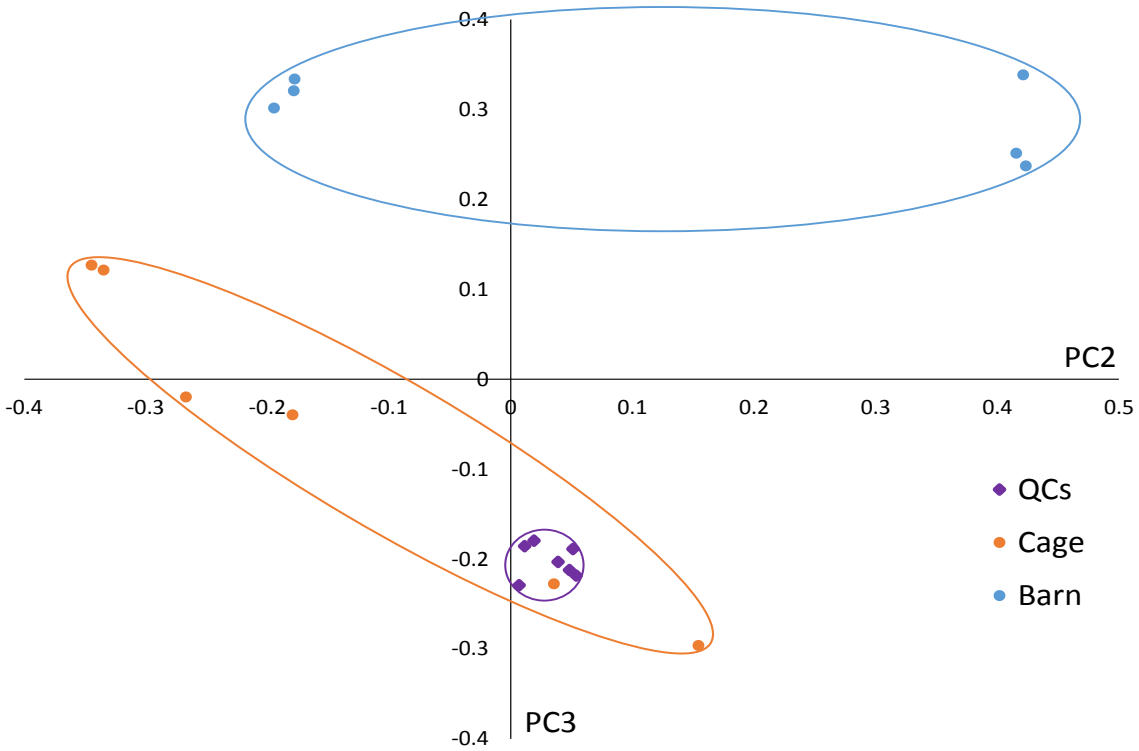


Figure 9.3: *PCA scores plot showing PC2 vs PC3 for yolk organic extracts of cage and barn eggs, including all compounds with CV%<30%. PC2 explains 1.08% of the variance, and PC3 explains 0.76% of the variance.*

The scores plot in Figure 9.3 shows complete separation between eggs laid by birds from cage and barn housing systems. This confirms what was deduced from the overlaid TICs in Figure 9.2; that there is a difference in metabolite profile between the yolks of cage and barn eggs. This could be due to the fact that the birds kept in cages have more limited space to move around compared to birds in barns, which would affect the

metabolite profiles of the birds, and of their eggs. PC2 only describes 1.08% of the total variance between samples, and PC3, which shows the most separation between the cage and barn eggs, only describes 0.76% of the variance. This shows just how subtle the differences in metabolite profile are between the yolks of eggs laid by birds from cage and barn housing systems. The QC samples form a small, tight group on the plot, confirming that there was little instrumental drift affecting the analysis.

9.4.1.4 Univariate Statistics

Following t-tests on the top 59 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC3 loadings, 29 were found to be statistically significant. These compounds, their CV%s, and p-values resulting from t-tests, can be seen in Table F.1 in Appendix F. Those in bold were also found to be statistically significant following confirmation using the raw data.

All 29 compounds in Table F.1 were present in a significantly higher abundance in the yolks of barn eggs compared to cage eggs, which supports what was observed in Figure 9.2, with most peaks showing a higher intensity in the TIC of barn egg yolk.

Of the 23 compounds that were still found to be statistically significant when confirmed using the raw data, 12 produced potential metabolite matches through METLIN, as seen in Table F.2 in Appendix F. The probability scores for the predicted molecular formulae for these compounds are above 95, indicating that there is a strong likelihood that the predicted formulae correspond to the compounds of interest. All 12 of these compounds were matched with various lipids: three potential diglycerides, four potential triglycerides, and five potential phospholipids. As all of these compounds were present in a higher abundance in barn eggs compared to cage eggs, this indicates that there is a higher lipid content in eggs from barn production systems compared to eggs from cage systems. These results are similar to those discovered by Pignoli et al., who found that there was a higher lipid content in free range eggs compared to cage eggs [16]. This could

be due to the fact that birds kept in cages experience a higher level of stress compared to birds in other housing systems, which leads to a greater lipid catabolism and, therefore, to a lower lipid content of egg yolk [16].

Of these 12 compounds in Table F.2, just one was putatively identified through comparing mass spectra provided by METLIN with the mass spectrum from the analysis. The compound *m/z* 734.5699 was putatively identified as 1,2-dipalmitoyl-glycero-3-phosphatidylcholine (PC(16:0/16:0)), a phosphatidylcholine with a palmitic acid group attached to the first two carbons of the glycerol backbone. The mass spectra comparison can be seen in Appendix F, Figure F.1. The observed *m/z* for this compound was due to the $[M+H]^+$ adduct. Figure 9.4 shows the difference in abundance of this compound in the yolks of cage and barn eggs, and Table 9.5 shows the putative identification of this compound and the *t*-test *p*-value.

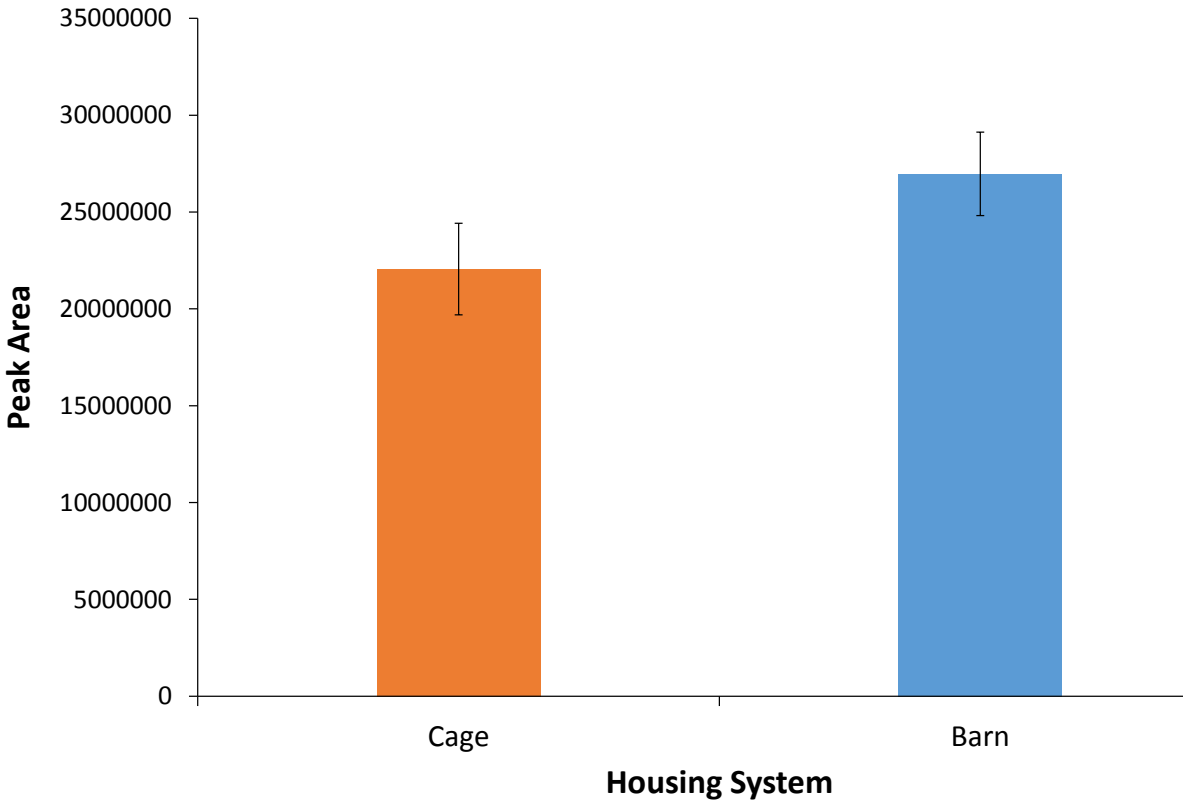


Figure 9.4: *Bar chart showing the difference in abundance of compound m/z 734.5699 between the yolks of cage and barn eggs. n=6. Measured as mean values of peak area for each housing system, with error bars ±1 standard deviation.*

Table 9.5: *Table showing the putative ID of compound m/z 734.5699 and the t-test p-value*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	***t*-test *p*-value**
734.5699	$C_{40}H_{80}NO_8P$	99.36	PC(16:0/16:0)	0.004

The *t*-test *p*-value in Table 9.5, and the bar chart in Figure 9.4, show that the abundance of this compound is significantly greater in the yolks of eggs laid by birds from barn housing systems, compared to the yolks of eggs laid by birds from cage systems. Therefore, this compound has potential as a marker capable of discriminating between cage and barn eggs.

As all of the 12 statistically significant compounds that produced potential matches through METLIN were matched with various lipids and phospholipids, and they were all of a significantly higher abundance in barn eggs compared to cage eggs, a lipid profile could also be used as a method of distinguishing between eggs from cage and barn housing systems. This would make the distinction between cage and barn eggs even stronger and more absolute than using just one compound, e.g. PC(16:0/16:0), as a marker, as several compounds would be involved in the discrimination between the two housing systems.

9.5 Cage and Barn eggs: Albumen Organic Extracts

Organic extracts of the albumen of eggs laid by hens from cage and barn housing systems were compared in order to observe the differences in metabolite profile between eggs from these two housing systems.

9.5.1 Results and Discussion

9.5.1.1 Quality Control Analysis

Table 9.6 shows the peak areas for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the peak areas for each of the peaks.

Table 9.6: *Table showing peak areas, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens from cage and barn housing systems*

	Peak Area					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC 1	14606949	48867111	95513889	1214770	1021874	11325776
QC 2	16530783	54242083	104034266	1395287	1068341	10353206
QC 3	15861207	52949716	102609987	1496926	970506	10142054
QC 4	11970273	40929579	87748859	1445819	824491	8090161
QC 5	11652672	41847098	86346936	1410197	659606	7736775
QC 6	11294999	41265024	83630770	1426971	710891	7422017
QC 7	11826816	43174120	88113021	1522392	829388	7631578
SD	2210536	5738467	8191882	99862	155887	1597817
Mean	13391957	46182104	92571104	1416052	869300	8957367
CV%	16.51	12.43	8.85	7.05	17.93	17.84

The table shows that the CV%s for the peak areas of all six peaks across the seven QC samples are all below the 30% cut off, indicating that the analysis was robust and that any instrumental drift occurring throughout the analysis was minimal.

Table 9.7 shows the RTs for the six peaks of interest in the seven QC samples that were analysed throughout the run, as well as the SD, the mean, and the CV% of the RTs for each of the peaks.

Table 9.7: *Table showing RTs, SDs, means, and CV%s for six peaks in seven QC samples analysed throughout the analytical run for organic extracts of albumen from eggs laid by hens from cage and barn housing systems*

	Retention Time (minutes)					
Sample	**Peak A**	**Peak B**	**Peak C**	**Peak D**	**Peak E**	**Peak F**
QC 1	7.606	11.677	14.601	16.343	18.848	22.381
QC 2	7.611	11.668	14.616	16.337	18.825	22.342
QC 3	7.611	11.667	14.622	16.354	18.843	22.359
QC 4	7.631	11.707	14.669	16.390	18.928	22.429
QC 5	7.671	11.754	14.705	16.430	18.919	22.436
QC 6	7.639	11.703	14.656	16.381	18.886	22.420
QC 7	7.697	11.794	14.745	16.473	18.978	22.511
SD	0.034	0.048	0.052	0.050	0.055	0.057
Mean	7.638	11.710	14.659	16.387	18.890	22.411
CV%	0.45	0.41	0.35	0.30	0.29	0.25

The table shows that the CV%s for the RTs of all six peaks throughout the seven QC samples are all well within the 2% cut off, indicating that the analysis was robust and there was minimal retention time drift throughout the analysis.

9.5.1.2 Metabolite Profiling

Comparisons were made between the albumen of eggs laid by hens from cage and barn housing systems, by studying the TICs of the extracts following HPLC-MS analysis. Figure 9.5 shows overlaid TICs for the average chromatograms of organic extracts of albumen from eggs laid by hens from these two housing systems.

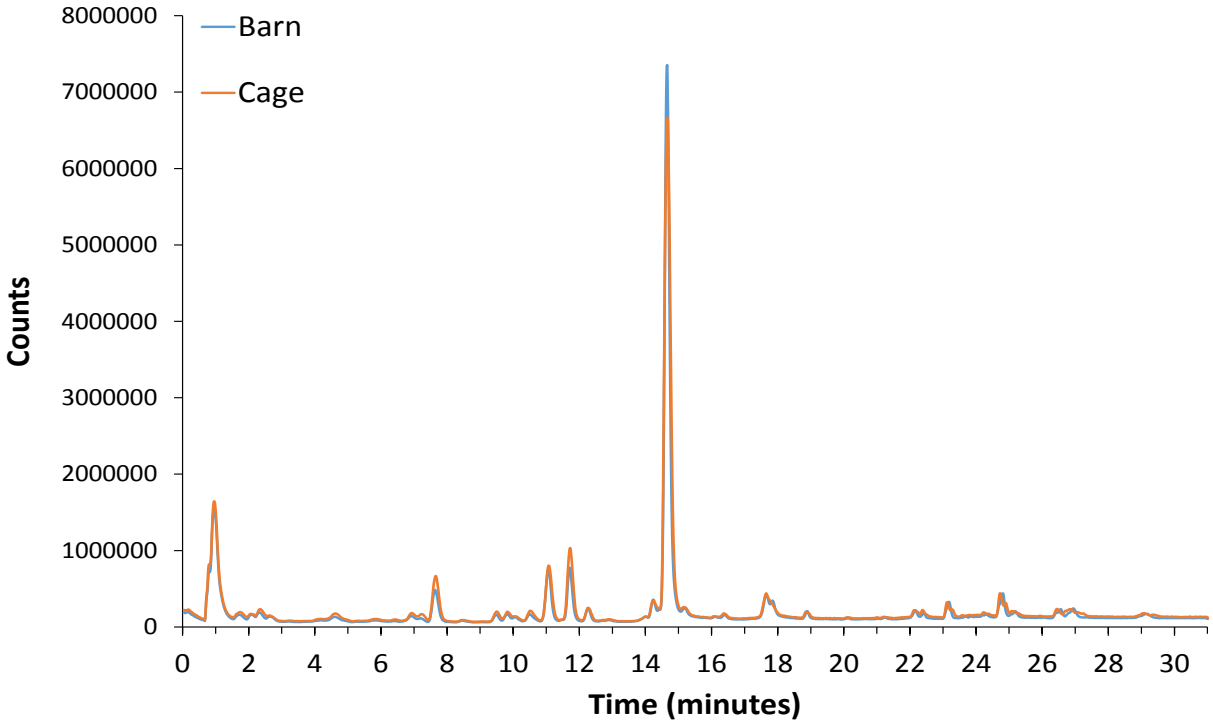


Figure 9.5: *Overlaid average TICs of organic extracts of albumen of eggs laid by hens from cage and barn housing systems*

The overlaid TICs in Figure 9.5 show only a very slight difference in the metabolite profiles of the albumen of eggs laid by birds from cage and barn housing systems. Up until 14 minutes of analysis, there is a very slightly higher peak intensity for all peaks in the TIC of the albumen of cage eggs, compared to the TIC of the albumen of barn eggs; however, the peak just before 15 minutes shows a higher intensity in the TIC of barn egg albumen compared to the TIC of cage egg albumen. After this point in the analysis there is no real difference in the peak intensity of the TICs of the albumen of eggs from the two different housing systems. This suggests that there is minimal difference in metabolite profile between the albumen of cage and barn eggs.

9.5.1.3 Multivariate Statistics

Following PCA, scores plots were produced in order to display the variation between the albumen of eggs laid by hens from cage and barn housing systems.

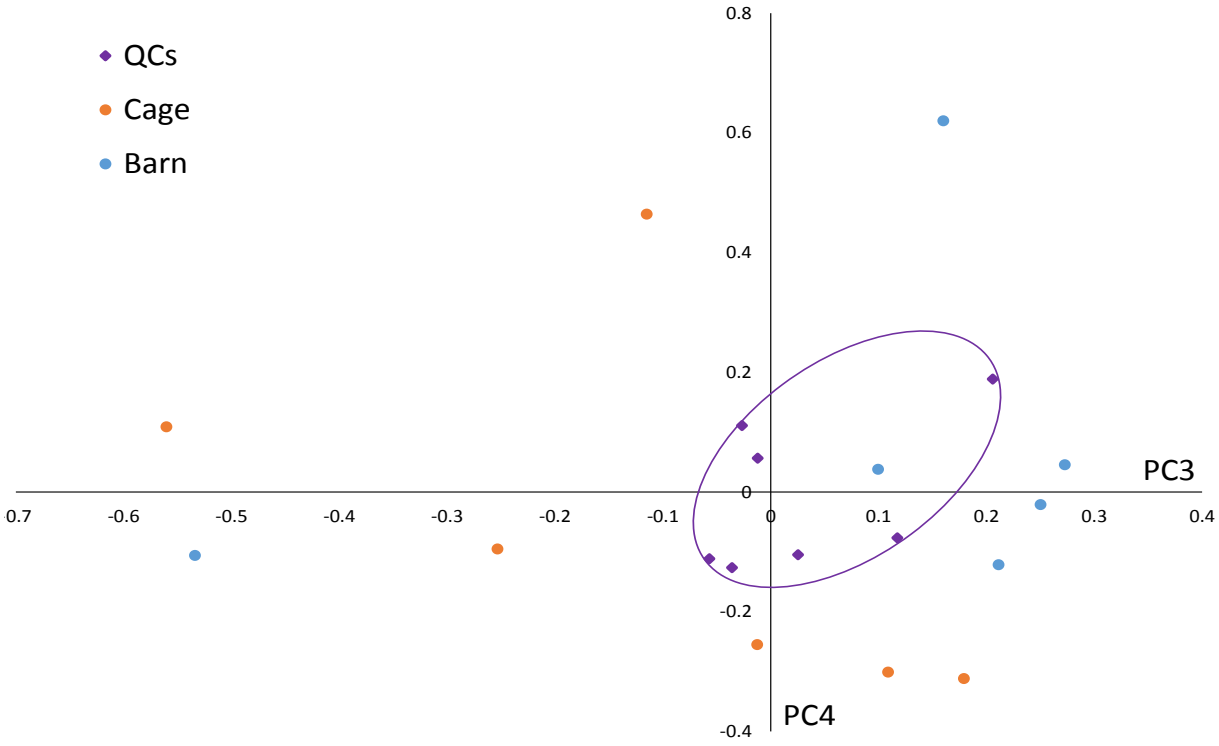


Figure 9.6: *PCA scores plot showing PC3 vs PC4 for albumen organic extracts of cage and barn eggs, including all compounds with CV%<30%. PC3 explains 0.14% of the variance, and PC4 explains 0.04% of the variance.*

The scores plot in Figure 9.6 shows no separation at all between eggs laid by birds from cage and barn housing systems. This confirms what was deduced from the overlaid TICs in Figure 9.5; that there is very little difference in metabolite profile between the albumen of eggs laid by birds from these two housing systems.

The QC samples are quite spread out, rather than forming a tight group on the plot. This, again, shows how little difference there is between the metabolite profiles of the albumen of eggs laid by birds from cage and barn housing systems; as there is such a small difference between the samples, the PCA is exploiting very small differences between the QC samples, resulting in the wider spread of QC samples that is observed on the plot.

Although the QC samples appear to be widely spread due to the lack of difference in metabolite profile between the albumen of eggs from cage and barn housing systems, the analysis was robust and there was minimal instrumental drift occurring throughout the analysis, as shown in the QC analysis in Section 9.5.1.1.

9.5.1.4 Univariate Statistics

Following *t*-tests on the top 37 compounds remaining, after the removal of any duplicates, adducts or isotopes from the top 100 compounds based on PC1 loadings, only six were found to be statistically significant. This confirms that there is very little difference in the metabolite profile of the albumen between cage and barn eggs. These six compounds, their CV%s, and *t*-test *p*-values, can be seen in Table F.3 in Appendix F. Those in bold were also found to be significant following confirmation using the raw data.

All six compounds in Table F.3 were again present in a significantly higher abundance in the albumen of barn eggs compared to cage eggs, similarly to what was observed when comparing the metabolite profiles of the yolks of eggs from cage and barn housing systems.

Of the two compounds that were still found to be statistically significant when confirmed using the raw data, only one of these produced potential metabolite matches through METLIN. Compound *m/z* 316.3209 was predicted the molecular formula $C_{19}H_{38}O_2$ with a likelihood score of 99.55, indicating a high probability that this predicted formula is correct for this compound. Several potential matches of fatty acids, and fatty acid esters, were found for this formula through METLIN; however, this compound was not successfully putatively identified as any of these fatty acids or fatty acid esters.

9.6 Cage, Barn, Free Range and Organic Eggs

A metabonomic workflow was also carried out to tentatively compare the yolk and albumen of eggs from four different housing systems: cage, barn, free range, and organic.

9.6.1 Yolk Organic Extracts

The overlaid TICs and PCA scores plot show some difference between eggs laid by birds from different housing systems, as seen in Figures 9.7 and 9.8. This suggests that there is some difference in yolk metabolite profile between these sets of eggs laid by the different groups of birds. However, as the age and breed, and potentially diet, of the birds was not the same between birds kept in all four housing systems, the observed differences could be due to these variables rather than housing system.

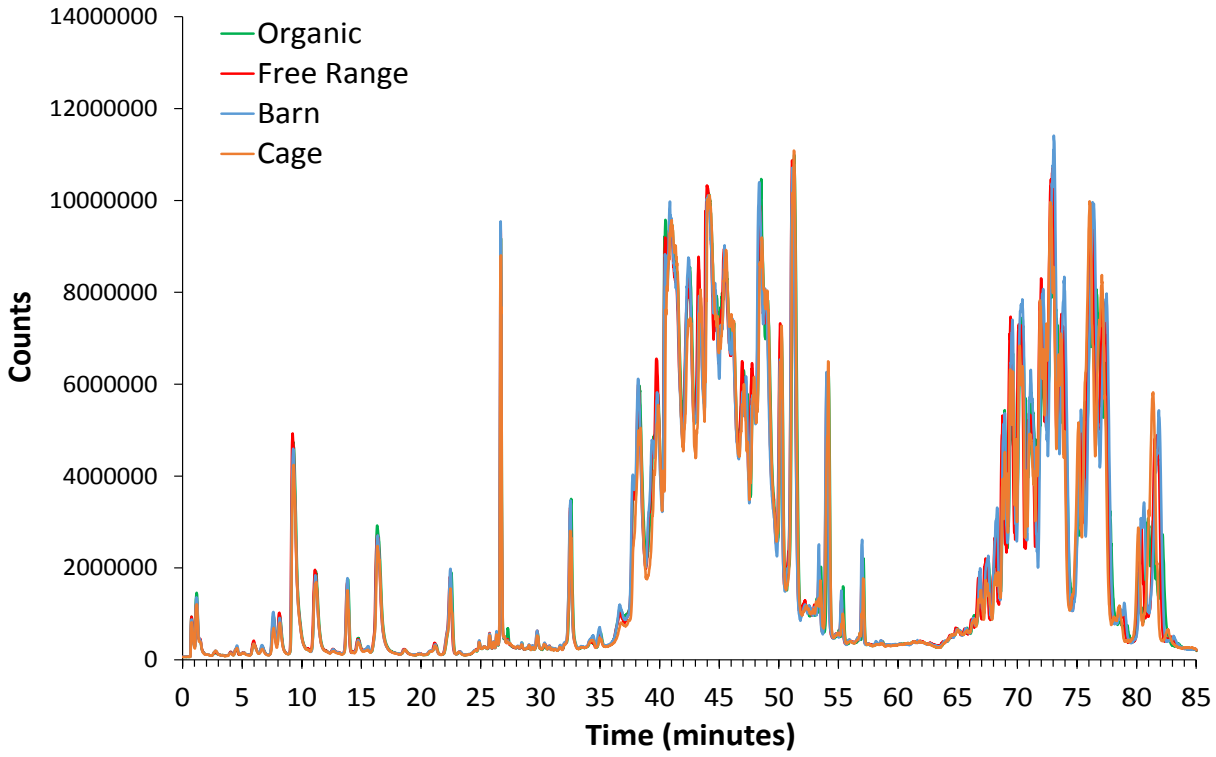


Figure 9.7: *Overlaid average TICs of organic extracts of yolks of eggs laid by hens from four different housing systems*

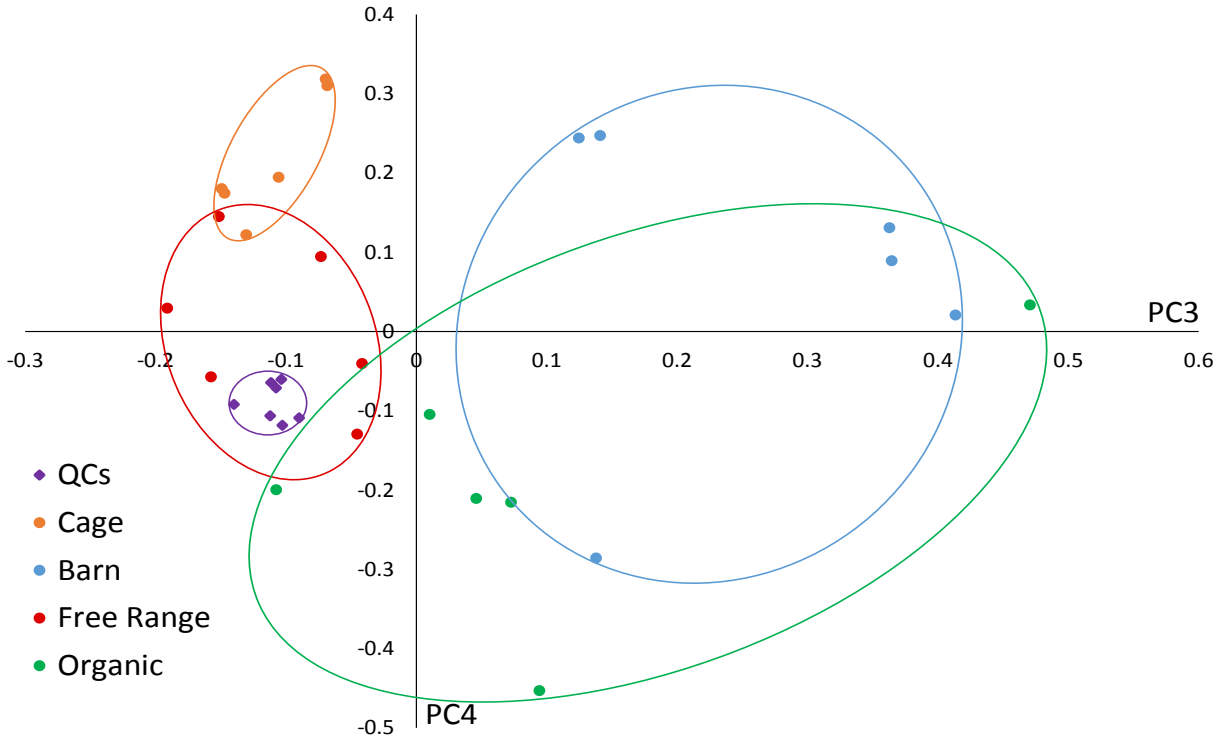


Figure 9.8: *PCA scores plot showing PC3 vs PC4 for yolk organic extracts of eggs laid by birds from four different housing systems, including all compounds with CV%<30%. PC3 explains 0.76% of the variance, and PC4 explains 0.27% of the variance.*

The greatest difference, shown in both Figure 9.7 and Figure 9.8, is between eggs laid by birds from cage and barn housing systems, which confirms what was observed previously; that cage and barn eggs can be distinguished from each other by the metabolite profile of the yolk.

Of the 48 compounds that were found to be statistically significant, following the removal of any duplicates, adducts, and isotopes from the top 100, 41 had their significance confirmed from the raw data analysis. Three of these compounds were putatively identified. Compounds *m/z* 716.5227, 725.5556, and 734.5699, were putatively identified as 1-palmitoyl-2-linoleoyl-glycero-3-phosphatidylethanolamine (PE(16:0/18:2)), 1-palmitoyl-sphingomyelin (SM(18:1/16:0)), and 1,2-dipalmitoyl-glycero-3-phosphatidylcholine (PC (16:0/16:0)). The observed *m/z* values for compounds *m/z* 716.5227 and 734.5699 were due to the $[M+H]^+$ adduct, and the observed *m/z* value for the compound *m/z* 725.5556

was due to the $[M+Na]^+$ adduct. The comparisons between mass spectra from METLIN, and those resulting from the chemical analysis, can be seen in Appendix F, Figures F.2 and F.3 for compounds *m/z* 716.5227 and 725.5556, and Figure F.1 for the compound *m/z* 734.5699, which was previously putatively identified and found to be significantly different in abundance between the yolks of cage and barn eggs.

This compound was again found to have a significantly different abundance between cage and barn egg yolks, as seen in Table 9.8.

Table 9.8: *Table showing the putative ID of compound m/z 734.5699 and ANOVA and Tukey test p-values. n=6. Orange indicates significance (p<0.05).*

Marker *m/z*	**Predicted Formula**	**Probability Score**	**Putative Identification**	**ANOVA *p*-value**
734.5699	$C_{40}H_{80}NO_8P$	99.36	PC(16:0/16:0)	0.001
Post-hoc test *p*-values				
	Cage	**Barn**	**Free range**	**Organic**
Cage		0.010	0.100	0.022
Barn			0.011	0.985
Free range				0.024
Organic				

This compound was also found to have statistical significance when comparing the yolks of eggs from the other housing systems. However, it was previously found to be statistically significant in the research in Chapter 5, when the effect of laying hen age on the metabolite profile of the egg yolk was studied. Therefore, the observed statistical significance could be due to the differences in the age of the laying birds, or possibly the breed, and potentially the diet of the laying birds, rather than the housing system. This compound could be used as a marker to discriminate between cage and barn eggs, but further work would need to be carried out, with more controlled variables, to determine whether it is useful as a marker between eggs from all four housing systems.

Compound *m/z* 716.5227, which was putatively identified as PE(16:0/18:2), a phosphatidylethanolamine with a palmitic acid group on the first carbon and a linoleic acid

Table 9.9: *Table showing the putative ID of compound m/z 716.5227 and the p-values resulting from ANOVA and Tukey tests. n=6. Orange indicates significance (p<0.05).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	ANOVA *p*-value
716.5227	$C_{39}H_{74}NO_8P$	98.5	PE(16:0/18:2)	0.009
Post-hoc test *p*-values				
	Cage	**Barn**	**Free range**	**Organic**
Cage		0.031	0.056	0.991
Barn			0.991	0.057
Free range				0.100
Organic				

group on the second carbon of the glycerol backbone, was also found to be significantly different in abundance between cage and barn eggs, as seen in Table 9.9. This compound was not found to be statistically significant when comparing the yolks of eggs laid by birds from the other housing systems, therefore the variables of hen age, breed, and diet are not of a concern. This compound could be useful as a marker of housing system, between cage and barn eggs.

Table 9.10: *Table showing the putative ID of compound m/z 725.5556 and the Welch test and Games-Howell test p-values. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

Marker *m/z*	Predicted Formula	Probability Score	Putative Identification	Welch *p*-value
725.5556	$C_{39}H_{79}N_2O_6P$	97.19	SM(18:1/16:0)	0.001
Post-hoc test *p*-values				
	Cage	**Barn**	**Free range**	**Organic**
Cage		0.219	0.016	0.001
Barn			0.921	0.180
Free range				0.040
Organic				

The compound *m/z* 725.5556, which was putatively identified as SM(18:1/16:0), a sphingomyelin with palmitic acid forming an ester bond to the amino group of the molecule,

was only found to show statistical significance when comparing the yolks of eggs laid by birds from cage, free range and organic systems, as seen in Table 9.10. However, this compound was previously found to be statistically significant in the research in Chapter 5, when the effect of laying hen age on the metabolite profile of the egg yolk was studied. As the age, breed, and potentially diet of the birds in these housing systems were all different, these observed differences in compound abundance could be due to the difference in age of the laying birds, or the breed or possibly diet, rather than housing system. Further studies, with more controlled variables, would be required in order to determine whether this compound is statistically significant due to the different housing systems, and therefore whether it is suitable as a marker of egg housing system.

9.6.2 Albumen Organic Extracts

The overlaid TICs in Figure 9.9 show only a very small difference in metabolite profile between the albumen of eggs from the four different housing systems.

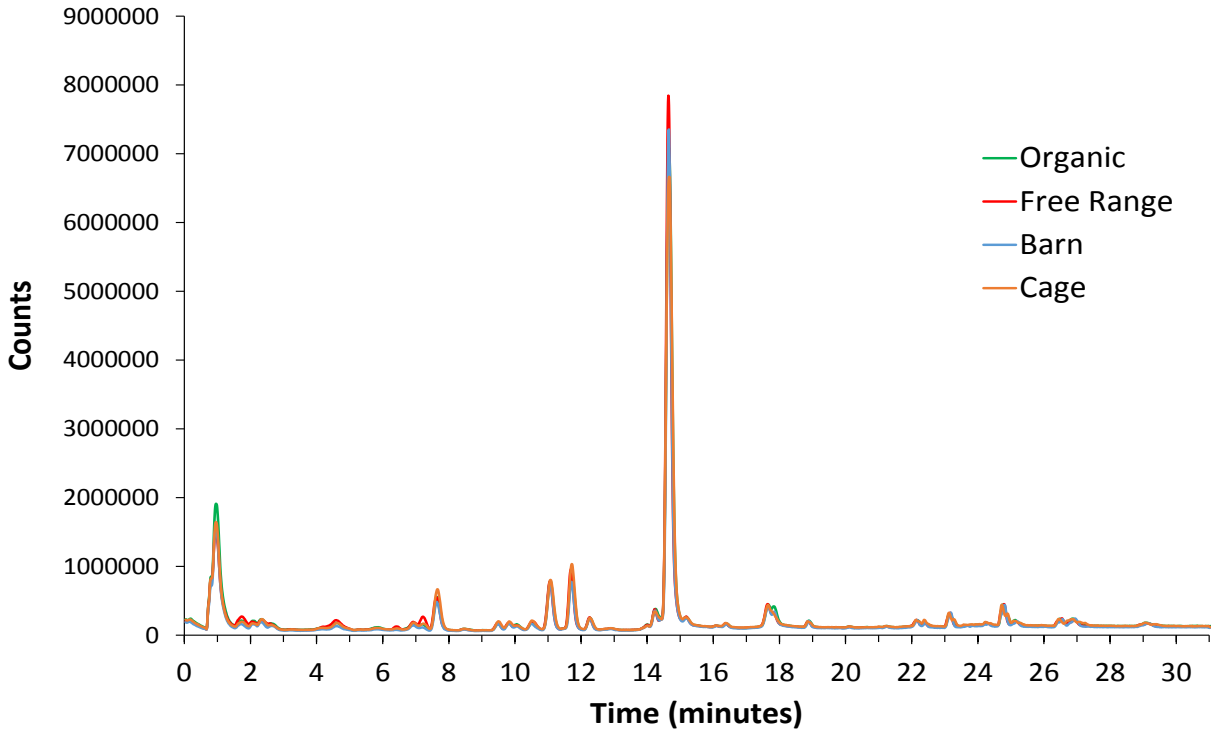


Figure 9.9: *Overlaid average TICs of organic extracts of albumen of eggs laid by hens from four different housing systems*

The PCA scores plot in Figure 9.10 shows even less difference between the albumen of eggs from the four different housing systems than the overlaid TICs in Figure 9.9, as there is no separation at all between the eggs from the different housing systems. This suggests that the housing system of the laying hen may not have much effect on the resulting metabolite profile of the albumen.

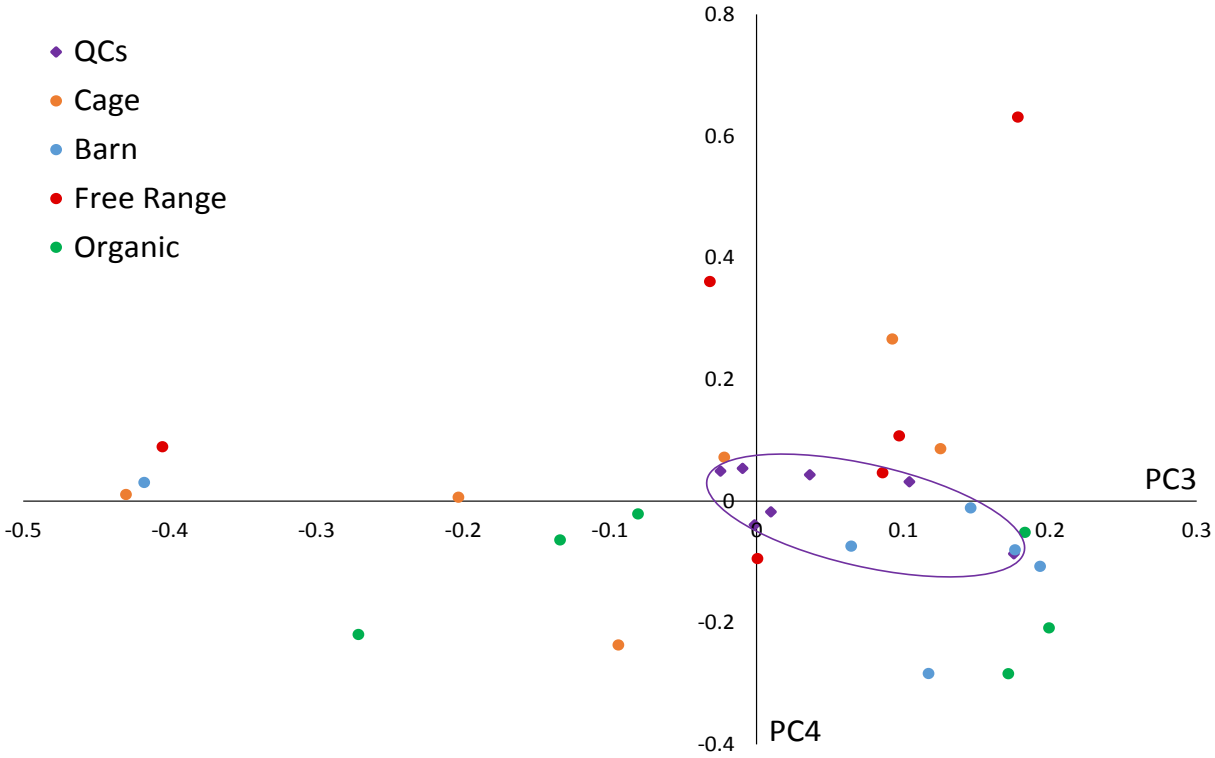


Figure 9.10: *PCA scores plot showing PC3 vs PC4 for albumen organic extracts of eggs laid by birds from four different housing systems, including all compounds with CV%<30%. PC3 explains 0.15% of the variance, and PC4 explains 0.06% of the variance.*

Although 26 of the 40 statistically significant compounds that remained, following the removal of any duplicates, adducts, or isotopes from the top 100, had their significance confirmed through the raw data analysis, none were successfully putatively identified.

The differences in abundance of these statistically significant compounds may have been influenced by the differences in breed, age, and potentially diet between the laying birds from the different housing systems, rather than the housing systems themselves, and therefore they may not be suitable as markers of egg production system. Further

work, including studies with more controlled variables, would be required in order to determine whether the statistical significance of these compounds is due to housing system, rather than the other variables, and therefore whether these compounds have potential as markers of egg housing system if they were to be identified.

9.7 Summary & Conclusions

9.7.1 Summary

The overlaid TICs and the PCA scores plot for the comparison of the yolks of eggs laid by birds from cage and barn housing systems showed that there was a difference in metabolite profile between the yolks of cage and barn eggs. All of the top 29 statistically significant compounds were found to be of a higher abundance in the yolk of barn eggs compared to cage eggs, and 12 of these produced potential metabolite matches to a variety of different lipids, including diglycerides, triglycerides, and phospholipids. One of these compounds was putatively identified as PC(16:0/16:0) and was determined to be suitable for discriminating between eggs from cage and barn housing systems.

The overlaid TICs and PCA scores plot for the comparison of the albumen of eggs laid by birds from cage and barn housing systems showed that there was very little difference in metabolite profile between the albumen of cage and barn eggs. Much fewer compounds in the albumen were found to be statistically significant compared to in the yolk, and none were successfully putatively identified, although the compound m/z 316.3209 produced potential matches with fatty acids and fatty acid esters.

When comparing the yolks of eggs laid by birds from all four housing systems, the overlaid TICs and PCA scores plot did show some difference in metabolite profile between the yolks of eggs laid by these birds. However, as the breed, age, and potentially diet of the birds kept in these systems were different, the observed differences in metabolite profile may have been due to these variables, rather than the differences in housing system.

Two putatively identified compounds, PC(16:0/16:0) and PE(16:0/18:2), were found to be significantly different in abundance between cage and barn egg yolk, making them potentially useful as markers to discriminate between eggs from these two housing systems. The compound PC(16:0/16:0) was previously discovered to be statistically significant when comparing just cage and barn egg yolk, therefore this confirmation of its significant difference increases its potential as a marker between cage and barn housing systems, as it is reproducibly statistically significant. This compound also showed statistical significance between eggs from free range and organic housing systems, as well as cage and barn systems, as did the compound that was putatively identified as SM(18:1/16:0). However, both of these compounds were previously found to be statistically significant between the yolks of eggs laid by birds of different ages. As the birds from the different housing systems, other than cage and barn systems, differed in age as well as breed, and potentially diet, the observed significant differences in compound abundance between eggs from these housing systems may be due to these factors, particularly the difference in laying hen age, rather than the differences in housing system. Therefore, these compounds may not be suitable as markers of egg production system for all four housing systems.

When comparing the albumen of eggs laid by birds from all four housing systems, the overlaid TICs and PCA scores plot again showed very little difference between the eggs laid by birds from the different systems. No statistically significant compounds found in the albumen were putatively identified.

9.7.2 Conclusions

This research has shown that there is a significant difference in metabolite profile between the yolk of cage and barn eggs. All compounds which were found to be statistically significant were significantly higher in the yolk of barn eggs compared to cage eggs, and 12 of these compounds were found to have potential matches to various lipids. Therefore, a lipid profile consisting of various lipids could be used as a method of discriminating

between cage and barn eggs. One of these 12 compounds was putatively identified as PC(16:0/16:0), and its statistical significance between cage and barn eggs was confirmed when eggs from all four housing systems were compared. Therefore, this compound shows potential as a marker to discriminate between cage and barn eggs.

The housing system of the laying hen appears to have very little effect on the resulting metabolite profile of egg albumen. Therefore, analysis of the albumen may not be as successful as the yolk, in discriminating between eggs from different housing systems.

Any compounds that were found to be statistically significant when comparing the eggs laid by birds from all four housing systems may have been affected by the difference in breed, age, or potentially diet between the birds in the different systems, rather than the housing systems themselves. Therefore further studies, with more controlled variables, would have to be carried out to confirm whether the statistical significance of these compounds is due to the difference in housing system of the laying bird, and therefore whether these compounds are suitable as markers of egg housing system.

Further work would involve the development of a yolk lipid profile capable of discriminating between cage and barn eggs. Following further studies comparing the yolks of free range and organic eggs, as well as cage and barn eggs, with more controlled variables, this profile could then be developed into a lipid profile capable of discriminating between eggs from all four housing systems. Further work would also involve comparing the abundance of the compound that was putatively identified as PC(16:0/16:0) between the yolks of eggs from all four housing systems, with birds of the same age, breed, and diet, to determine whether it is statistically significant, and therefore whether it has potential as a marker of egg housing system for all four systems. The identification of this compound would then need to be confirmed through the use of a chemical standard. Studies comparing the developed yolk lipid profile and the abundance of PC(16:0/16:0) between eggs from different housing systems would then be repeated for different ages and breeds of birds, to ensure that the results are replicable.

9.8 References

[1] *Commission Regulation (EC) 589/2008 of 23 June 2008, Laying down detailed rules for implementing Council Regulation (EC) No 1234/2007 as regards marketing standards for eggs*, Official Journal of the European Union, L163/6-23, 2008.

[2] *Council Directive 1999/74/EC of 19 July 1999, Laying down minimum standards for the protection of laying hens*, Official Journal of the European Communities, L203/53-57, 1999.

[3] *Council Regulation (EC) No 834/2007 of 28 June 2007, on organic production and labelling of organic products and repealing Regulation (EEC) No 2092/91*, Official Journal of the European Union, L189/1-23, 2007.

[4] *United Kingdom egg statistics - Quarter 4, 2018*, Department for Environment Food & Rural Affairs, 2018.

[5] (2018). Preston free range egg fraud farmer to pay back £500,000, [Online]. Available: `https://www.bbc.co.uk/news/uk-england-lancashire-44552810` (visited on 07/2019).

[6] (2018). Chicken farmer given suspended sentence for free-range egg fraud, [Online]. Available: `https://www.theguardian.com/uk-news/2018/jul/02/chicken-farmer-given-suspended-sentence-for-free-range-egg` (visited on 07/2019).

[7] R. Karoui, B. Kemps, F. Bamelis, B. deKetelaere, E. Decuypere, and J. deBaerdemaeker, "Methods to evaluate egg freshness in research and industry: A review," *European Food Research and Technology*, vol. 222, pp. 727–732, 2006.

[8] M. Ahammed, B. J. Chae, J. Lohakare, B. Keohavong, M. H. Lee, S. J. Lee, D. M. Kim, J. Y. Lee, and S. J. Ohh, "Comparison of aviary, barn and conventional cage raising of chickens on laying performance and egg quality," *Asian-Australasian Journal of Animal Sciences*, vol. 27, pp. 1196–1203, 2014.

[9] B. Y. Dikmen, A. Ipek, Ü. Şahan, A. Sözcü, and S. C. Baycan, "Impact of different housing systems and age of layers on egg quality characteristics," *Turkish Journal of Veterinary Animal Science*, vol. 41, pp. 77–84, 2017.

[10] A. Hidalgo, M. Rossi, F. Clerici, and S. Ratti, "A market study on the quality characteristics of eggs from different housing systems," *Food Chemistry*, vol. 106, pp. 1031–1038, 2008.

[11] Z. Sokołowicz, J. Krawczyk, and M. Dykiel, "Effect of alternative housing system and hen genotype on egg quality characteristics," *Emirates Journal of Food and Agriculture*, vol. 30, pp. 695–703, 2018.

[12] A. Sekeroglu, M. Sarica, E. Demir, Z. Ulutas, M. Tilki, M. Saatci, and H. Omed, "Effect of different housing systems on some performance traits and egg qualities of laying hens," *Journal of Animal and Veterinary Advances*, vol. 9, pp. 1739–1744, 2010.

[13] H. van den Brand, H. K. Parmentier, and B. Kemp, "Effects of housing system (outdoor *vs* cages) and age of laying hens on egg characteristics," *British Poultry Science*, vol. 45, pp. 745–752, 2004.

[14] D. Matt, E. Veromann, and A. Luik, "Effect of housing systems on biochemical composition of chicken eggs," *Agronomy Research*, vol. 7, pp. 662–667, 2009.

[15] R. M. Radu-Rusu, M. G. Usturoi, A. Leahu, S. Amariei, C. G. Radu-Rusu, and I. Vacaru-Opriş, "Chemical features, cholesterol and energy content of table hen eggs from conventional and alternative farming systems," *South African Journal of Animal Science*, vol. 44, pp. 33–42, 2014.

[16] G. Pignoli, M. T. Rodriguez-Estrada, M. Mandrioli, L. Barbanti, L. Rizzi, and G. Lercker, "Effects of rearing and feeding systems on lipid oxidation and antioxidant capacity of freeze-dried egg yolks," *Journal of Agricultural and Food Chemistry*, vol. 57, pp. 11 517–11 527, 2009.

[17] G. B. Anderson, W. Bolton, R. M. Jones, and M. H. Draper, "Effect of age of the laying hen on the composition of the egg," *British Food Journal*, vol. 19, pp. 741–745, 1978.

[18] K. Küçükyilmaz, M. Bozkurt, E. N. Herken, M. Çinar, A. U. Çath, E. Bintaş, and F. Çöven, "Effects of rearing systems on performance, egg characteristics and immune response in two layer hen genotype," *Asian-Australasian Journal of Animal Sciences*, vol. 25, pp. 559–568, 2012.

[19] A. Tres, R. O'Neill, and S. van Ruth, "Fingerprinting of fatty acid composition for the verification of the identity of organic eggs," *Lipid Technology*, vol. 23, pp. 40–42, 2011.

[20] A. Sekeroglu, M. Sarica, E. Demir, and Z. Ulutas, "Effect of housing systems on some fatty acid contents of hen's eggs," *Asian Journal of Chemistry*, vol. 20, pp. 4047–4050, 2008.

[21] S. van Ruth, M. Alewijn, K. Rogers, E. Newton-Smith, N. Tena, M. Bollen, and A. Koot, "Authentication of organic and conventional eggs by carotenoid profiling," *Food Chemistry*, vol. 126, pp. 1299–1305, 2011.

[22] I. Giannenas, P. Nisianakis, A. Gavriil, G. Kontopidis, and I. Kyriazakis, "Trace mineral content of conventional, organic and courtyard eggs analysed by inductively coupled plasma mass spectrometry (ICP-MS)," *Food Chemistry*, vol. 114, pp. 706–711, 2009.

[23] K. M. Rogers, "Stable isotopes as a tool to differentiate eggs laid by caged, barn, free range, and organic hens," *Journal of Agricultural and Food Chemistry*, vol. 57, pp. 4236–4242, 2009.

10. Conclusions and Future work

10.1 Conclusions

This research has resulted in the development of a non-targeted metabonomic workflow that is capable of discovering differences in the metabolite profiles of the yolk and albumen of eggs belonging to different experimental groups. It has shown how this workflow can be successfully applied to studies investigating the development of methods for the potential detection of fraud in the production and marketing of poultry eggs.

The first three studies that were carried out during this research were preliminary studies, the results of which influenced the experimental design and interpretation of the results of the remaining three studies. The work described in Chapter 4 investigated which condition and storage temperature were the best at maintaining metabolite stability of yolk and albumen organic metabolite extracts during storage. It was concluded that for yolk organic extracts, over both 6 and 12 weeks of storage, and for albumen organic extracts over a longer storage period, that storing the extracts dried at -80 °C provides the best metabolite stability. Therefore, throughout this research the organic metabolite extracts were stored dried at -80 °C prior to chemical analysis.

The work described in Chapters 5 and 6 investigated whether the age and diet of the laying bird respectively, affect the metabolite profiles of the yolk and albumen of the egg. These studies concluded that both the age and the diet of the laying bird have an effect on the metabolite profiles of the egg yolk and albumen, with hen age having

a greater effect on the metabolite profile of the albumen than on that of the yolk. The effects of both hen age and diet on the metabolite profiles of both the yolk and albumen were found to only be very subtle. However, as metabonomic studies are often used to detect subtle differences in metabolite profile between experimental groups, these very slight differences due to hen age and diet may have a relatively large impact on the metabolite profile compared to the effect of the variable that is under investigation, and thus a large influence on the results of the study. Therefore it was determined that, throughout this research, the age and diet of the birds should be kept the same between different experimental groups in the same study. Any compounds that were found to be statistically significant between eggs laid by birds of different ages should be treated with caution when found to be significant in studies when birds are not of the same age.

The research in Chapter 5 also investigated, from a fraud detection perspective, whether a metabonomic approach was capable of identifying potential markers of laying hen age in the yolk and albumen of eggs. However the top, putatively identified, statistically significant compounds in egg yolk showed no significant correlation between their abundance in yolk and the age of the laying hen, over the total age range that was studied, and thus were determined to not be useful as markers of hen age. One statistically significant compound found in the albumen, which was putatively identified as oleoyl ethyl amide, showed some potential as a marker as its abundance was positively correlated with hen age.

The work in Chapter 6 was a useful preliminary study investigating whether a difference in the diet of the laying hen can be observed from the metabolite profiles of the egg yolk and albumen, and if so, how quickly the metabolite profile of the yolk is affected by a change in diet of the laying hen. This research revealed that two weeks after a change in diet of laying hens, a difference can be seen in the metabolite profiles of the yolks, and five weeks after a diet change a difference can be seen in the metabolite profiles of the albumen. However, it also showed that different diets take different lengths of time

to have an effect on the metabolite profile of the yolk, and may take different lengths of time to affect the yolk compared to the albumen.

The latter three studies that were carried out as part of this research were all related to the application of metabonomic workflows to the development of methods for the detection of fraud during egg production and marketing.

The work in Chapter 7 showed how a metabonomic approach uncovered differences in the metabolite profiles of the yolks of shell eggs of different ages, when stored at both 23 °C and 5 °C for up to five weeks, as well as between samples of liquid yolk that had been refrigerated at 5 °C for different lengths of time, up to 96 hours. One of the compounds that was found to be statistically significant when comparing the yolks of eggs of different ages following storage at 23 °C, was identified as choline. Choline abundance was found to be positively correlated with egg age when eggs were stored at 23 °C, and increased with an increasing egg age up to twelve weeks, making it potentially useful as a marker of egg age. However, choline abundance was not found to be correlated with egg age when eggs were stored at 5 °C for up to either five weeks or twelve weeks, and when the abundance of choline in egg yolk was compared between eggs that were stored at 23 °C and those that were stored at 5 °C for five weeks, the abundance of choline in the yolks of the refrigerated eggs was significantly lower. Therefore, it was determined that refrigerating eggs can prevent the increase of choline abundance over time, and therefore prevent use of choline as a marker of egg age. Although several other compounds were found to be statistically significant when comparing eggs that were stored at 5 °C for different lengths of time, none of these were determined to be suitable as markers of egg age. Two statistically significant compounds that were putatively identified as oleoyl ethyl amide and erucamide, were found to have potential as markers of liquid yolk storage time when stored at 5 °C, particularly as markers capable of discriminating between 48 and 72 hours of storage, and therefore capable of detecting fraud, as they both showed a positive correlation with storage time. These two compounds, as well as that which

was putatively identified as 1-monopalmitin, were also determined to have potential as markers capable of discriminating between fresh yolk and yolk that has been refrigerated for 24 hours.

The research that was described in Chapter 8 showed that the population size of laying hen cages does not have a great effect on the metabolite profile of egg yolk, and the effects that it does have, and the compounds that are affected, are mostly different to those affected by the stocking density. Some differences in metabolite profile were observed between the yolks of eggs laid by birds from cages of different stocking densities. One compound that was found to show statistical significance when comparing the yolks of eggs laid by birds from cages of different stocking densities was putatively identified as canthaxanthin. This compound showed some slight potential as a marker capable of discriminating between eggs laid by birds from cages of different stocking densities.

The work in Chapter 9 showed that there is a clear difference in metabolite profile between the yolks of eggs laid by birds that have been kept in cage and barn housing systems. All of the statistically significant compounds with potential metabolite matches were potential lipids and phospholipids, and were all of a higher abundance in the yolk of barn eggs compared to cage eggs. Therefore, a lipid profile may be capable of discriminating between eggs from cage and barn housing systems. One of these compounds was putatively identified as PC(16:0/16:0) and has potential as a marker of egg housing system for cage and barn eggs. Although no real conclusions could be drawn from the comparison of the metabolite profiles of the yolk and albumen of eggs from free range and organic housing systems, with those of the yolk and albumen of eggs from cage and barn systems, there were some observable differences in the metabolite profiles of the yolk of eggs from the four different housing systems. However, the metabolite profile of the albumen was not very strongly affected by the housing system of the laying hen.

Table 10.1 summarises the putatively identified, potentially useful markers that were discovered in each study throughout this work.

Table 10.1: *Summary table showing putatively identified potential markers in each chapter*

Chapter	Potential Markers	
Chapter 4: Extract Storage	N/A	
Chapter 5: Hen Age	Albumen: oleoyl ethyl amide	
Chapter 6: Hen Diet	N/A	
Chapter 7: Egg Age	Yolk at 23 °C: choline	Liquid yolk at 5 °C: oleoyl ethyl amide & erucamide
Chapter 8: Hen Stocking Density	Yolk: canthaxanthin	
Chapter 9: Hen Housing System	Yolk: PC(16:0/16:0)	Yolk: Lipid profile

10.2 Future work

Although throughout this research there were several putatively identified compounds that showed potential as markers of various factors, which could help to detect cases of fraud in egg production and marketing, only one of these putative identifications was confirmed through the use of a chemical standard. Further work would involve the confirmation of all putative identifications of potentially useful compounds, followed by targeted, quantitative studies focussing on these compounds, using greater sample sizes to increase the power of the studies. Validation of the developed analytical methods for the quantitative evaluation of these potential markers would then be carried out.

Throughout this research the age and breed of the laying birds were kept the same, both within and between experimental groups, to ensure that any observed differences were due to the factor that was under investigation, rather than another variable. However, the factors that have been investigated may have different effects on birds of different ages and breeds, and therefore their eggs, or the age and breed of the birds may affect the concentrations of potential markers in the eggs. Therefore these studies, including any follow-up targeted studies, should be repeated across different breeds and ages of laying birds, to determine whether the metabolite profiles of the yolk and albumen of eggs are affected similarly, and whether the potential use of any identified compounds as markers is replicable, across the different ages and breeds of laying bird.

Only organic extracts of the yolk and albumen were analysed throughout this research, and only using positive ionisation in the MS analysis. Future work could involve further chromatographic development for the aqueous metabolite extracts of yolk and albumen, including investigating the use of columns with different stationary phases and a variety of different mobile phases and modifiers. Different mobile phase modifiers may also improve the detection of compounds during negative ionisation MS analysis, allowing further investigation of how the negatively ionised compounds in the metabolite profiles

of egg yolk and albumen are affected by various factors, potentially discovering further compounds that may have potential as markers to help detect cases of fraud. It may also be possible to combine the aqueous and organic extracts and carry out a more complex separation using 2D-LC, where two columns of different stationary phases are connected in sequence and used to separate a wider range of compounds within a sample. This would result in a greater number of compounds being separated and detected in just one analysis, increasing the likelihood of discovering compounds of interest, without carrying out extra analyses.

Only unsupervised multivariate analysis, in the form of PCA, was carried out as part of the data analysis workflow throughout this research. It would be interesting to use supervised MVA methods such as Partial Least Squares (PLS) on the data sets generated during this work, to build training set models capable of predicting the categories of unknown samples [1]. It would also be interesting to introduce the data sets to Artificial Neural Networks (ANNs), in order to detect patterns in the data and make predictions about unknown samples [2][3].

In addition to what has previously been described, further work following on from the research in Chapter 5 would involve carrying out similar metabonomic studies, both non-targeted and targeted, over a wider hen age range up to a higher age of bird, to observe how the metabolite profiles of the yolk and albumen of eggs change over the full laying period of a bird. The targeted studies would then be able to determine the age of laying hen at which a correlation between the abundance of potential markers, such as oleoyl ethyl amide in the albumen, and hen age, is no longer observed and the compounds are no longer useful as markers of hen age.

The research described in Chapter 6 determined that different diets affect the metabolite profiles of the yolk and albumen differently, and at different rates, to each other. Therefore further, more diet specific, studies would have to be carried out investigating how, to what extent, and how quickly, switching from certain laying hen diets, e.g. differ-

ent varieties of organic feed, 95% organic feed with 5% non-organic feed, and conventional feed, to different varieties of 100% organic feed, affects the metabolite profiles of the yolk and albumen.

In addition to repeating the targeted egg age study described in Chapter 7 with larger sample sizes, and validating the analytical method for the quantitative evaluation of choline in egg yolk, further work would also involve the confirmation that the choline concentration in egg yolk increases due to the catabolism and/or degradation of phosphatidylcholine over time. This would be done by carrying out MS analysis in precursor ion scan mode for the m/z 184 fragment, in order to detect the phosphatidylcholine ions that fragment to produce this ion, allowing the abundance of phosphatidylcholines to be monitored over an increasing egg age. If it is found that the abundance of phosphatidylcholine decreases with increasing egg age, then it can be concluded that the increase in choline concentration is due to the catabolism and/or degradation of phosphatidylcholines, releasing free choline molecules.

Further work relating to the liquid egg ageing study in Chapter 7 would involve carrying out similar studies observing the differences in metabolite profile between samples of liquid albumen, and liquid yolk and albumen combined, that are both refrigerated and stored at room temperature for different lengths of time, as well as between samples of liquid yolk that are stored for different lengths of time at room temperature. It would also be interesting to repeat this study using shell eggs of different ages prior to breaking, to see if similar results are obtained.

Following the research that was described in Chapter 8, further work would involve conducting similar studies on a larger scale, over a longer period of time, with laying hens kept in cages of greater population sizes, and more highly different stocking densities, to emulate a more realistic scenario. Greater differences in metabolite profile, and more potential markers, may be discovered when carrying out these studies on a larger scale. It may also be possible to observe larger differences in the abundance of canthaxanthin

in the yolks of eggs laid by birds from cages of different stocking densities, increasing its potential as a marker of cage stocking density. This could lead to a greater possibility of developing a method capable of detecting cases of fraud regarding battery cages.

In addition to the putative identification of PC(16:0/16:0), in the research described in Chapter 9, several compounds that were found to be of a significantly higher abundance in the yolk of barn eggs, compared to cage eggs, were tentatively classified as different types of lipids. Therefore it may be possible, with further work, to develop a lipid profile capable of discriminating between cage and barn eggs. Further research would also involve carrying out a similar study, comparing eggs from free range and organic housing systems, with eggs from cage and barn systems, where laying hens are all of the same age and breed, in order to observe differences in the metabolite profiles of yolk and albumen between eggs from all four housing systems, and to potentially discover and identify compounds that show potential as markers of egg housing system. The lipid profile could then be further developed in order to successfully discriminate between eggs from all four housing systems. It would also be interesting to observe the difference in abundance of the compound putatively identified as PC(16:0/16:0), when comparing the yolks of eggs from free range and organic housing systems with those from cage and barn systems, to see if this compound has potential as a marker capable of discriminating between all four housing systems, as well as between just cage and barn systems.

This research has shown that a metabonomic approach can be successful in observing differences between poultry eggs from different experimental groups, as well as in the detection and identification of compounds that show potential as markers capable of detecting instances of fraud in the production and marketing of poultry eggs. However, there is clearly the potential for much more work regarding further method development, marker identification, conducting larger and more powerful targeted studies, and the validation of analytical methods for the detection and quantification of potential markers, as well as the development of multivariate statistical and computational models for the

prediction of sample information, and therefore the potential detection of fraud.

10.3 References

[1] S. Ren, A. Hinzman, E. Kang, R. Szczesniak, and L. Lu, "Computational and statistical analysis of metabolomics data," *Metabolomics*, vol. 11, pp. 4192–1513, 2015.

[2] Y. Wu and J. Feng, "Development and application of artificial neural network," *Wireless Personal Communications*, vol. 102, pp. 1645–1656, 2018.

[3] J. Miller and J. Miller, *Statistics and chemometrics for analytical chemistry*, 6th ed. Essex: Pearson, 2010.

Appendices

Appendix A

Table A.1: *Table showing the significantly different features between yolk organic extracts that were stored either dried or re-dissolved at -25 °C, -46 °C, and -80 °C for six weeks. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance. 25, 46, and 80 refer to the storage of re-dissolved extracts at -25 °C, -46 °C, and -80 °C respectively. 25*, 46* and 80* refer to the storage of dried extracts at the same temperatures.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend 25 | 25* | 46 | 46* | 80 | 80*
369.3527	32.222	2.51	0.001	
496.3402	9.098	4.01	<0.001	
524.3717	16.058	3.54	<0.001	
603.5350	50.716	1.46	<0.001	
610.5415	46.577	15.80	<0.001	
612.5568	49.585	1.77	<0.001	
638.5725	50.615	1.72	<0.001	
640.5883	53.436	1.81	<0.001	
734.5704	42.568	3.47	0.001	
744.5544	45.369	1.70	<0.001	
758.5698	40.289	1.58	<0.001	
762.5922	43.344	1.33	<0.001	
768.5545	45.154	2.15	0.005	
782.5704	40.107	1.87	<0.001	
784.5855	43.127	2.86	<0.001	
784.5860	41.751	1.49	<0.001	

Table A.1 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend 25 | 25* | 46 | 46* | 80 | 80*
786.6017	45.543	3.69	<0.001	
806.5707	39.784	1.05	<0.001	
808.5863	42.591	3.25	<0.001	
810.6020	44.789	2.51	<0.001	
818.7253	65.562	1.34	<0.001	
836.6182	47.103	9.19	<0.001	
844.7413	66.163	2.02	<0.001	
848.7723	70.012	1.14	0.007	
850.7884	72.673	5.55	0.001	
860.7727	69.275	1.83	0.003	
862.7876	71.674	3.54	0.007	
870.7593	66.771	2.08	<0.001	
876.8048	73.438	2.95	0.004	
890.8184	75.496	5.01	<0.001	
896.7739	67.311	1.13	<0.001	
902.8187	74.266	2.57	0.002	
904.8346	77.771	3.31	<0.001	

Table A.2: *Table showing the significantly different features between yolk organic extracts that were stored either dried or re-dissolved at -25 °C, -46 °C, and -80 °C for twelve weeks. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance. 25, 46, and 80 refer to the storage of re-dissolved extracts at -25 °C, -46 °C, and -80 °C respectively. 25*, 46* and 80* refer to the storage of dried extracts at the same temperatures.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend 25 | 25* | 46 | 46* | 80 | 80*
339.2891	52.415	5.57	<0.001	
524.3718	17.357	14.58	0.003	
603.5349	52.531	6.19	<0.001	
608.5248	44.978	4.41	<0.001	
611.5439	48.233	4.80	<0.001	
617.5115	51.502	5.30	0.003	
638.5727	52.432	4.57	<0.001	
659.5010	52.332	22.68	<0.001	
703.5752	39.600	3.15	0.009	
718.5390	45.867	4.77	0.001	
728.5584	49.480	8.39	0.008	
734.5699	44.613	5.40	0.003	
744.5534	40.447	4.52	0.007	
744.5547	47.381	5.95	0.009	
746.5693	43.500	3.66	0.002	
762.6003	49.394	9.54	<0.001	
766.5392	44.730	4.14	0.004	
794.5710	49.096	4.64	0.010	
800.6168	49.761	3.70	0.010	
808.5869	43.480	3.78	0.004	
811.6040	45.109	3.56	0.010	
818.7253	67.074	2.53	<0.001	
820.7445	69.215	2.56	0.004	
822.7585	71.774	3.74	0.002	
836.6080	46.439	3.92	0.001	

Table A.2 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend 25 | 25* | 46 | 46* | 80 | 80*
844.7411	67.751	7.23	<0.001	
849.7760	72.482	3.71	0.044	
850.7888	75.885	10.49	<0.001	
860.7736	71.534	4.08	<0.001	
862.7930	74.575	10.82	0.047	
867.6846	69.835	19.73	0.009	
870.7590	69.348	4.32	<0.001	
871.7624	68.470	17.12	<0.001	
872.7721	71.835	6.78	0.002	
886.7871	72.255	4.94	<0.001	
894.7560	67.140	2.41	<0.001	
896.7717	70.064	5.67	<0.001	
898.7872	72.020	11.56	<0.001	
898.7908	71.332	4.80	<0.001	
905.8284	77.921	19.14	0.004	
922.7884	71.732	5.38	0.002	
926.8183	75.606	10.57	<0.001	
928.8331	78.984	21.26	0.003	

Table A.3: *Table showing the significantly different features between yolk organic extracts that were stored under different conditions and at different temperatures for six weeks, as well as twelve weeks of storage. The p-value was calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	6 weeks of storage *p*-value	12 weeks of storage *p*-value
369.3527	32.222	0.001	0.317
603.5350	50.716	<0.001	<0.001
610.5415	46.577	<0.001	0.268
612.5568	49.585	<0.001	0.075
636.5567	47.660	0.001	0.050
636.6297	54.646	0.006	0.123
638.5725	50.615	<0.001	<0.001
640.5883	53.436	<0.001	0.040
760.5855	44.266	0.001	0.079
764.5233	39.967	0.005	0.117
782.5704	40.107	<0.001	0.260
786.6017	45.543	<0.001	0.483
788.6173	48.324	0.001	0.497
792.5557	44.538	0.003	0.050
818.7253	65.562	<0.001	<0.001
844.7413	66.163	<0.001	<0.001
848.7723	70.012	0.007	0.050
850.7884	72.673	0.001	<0.001
860.7727	69.275	0.003	<0.001
862.7876	71.674	0.007	0.008
870.7593	66.771	<0.001	<0.001
876.8048	73.438	0.004	0.195
890.8184	75.496	<0.001	0.050
896.7739	67.311	<0.001	<0.001
898.7877	69.118	0.025	<0.001
902.8187	74.266	0.002	0.004
904.8346	77.771	<0.001	0.173

Table A.4: *Table showing the significantly different features between yolk organic extracts that were stored under different conditions and at different temperatures for twelve weeks, as well as six weeks of storage. The p-value was calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	12 weeks of storage *p*-value	6 weeks of storage *p*-value
608.5248	44.978	<0.001	<0.001
818.7253	67.074	<0.001	<0.001
844.7411	67.751	<0.001	<0.001
850.7888	75.885	<0.001	0.001
870.7590	69.348	<0.001	<0.001
871.7624	68.470	<0.001	<0.001
886.7871	72.255	<0.001	<0.001
894.7560	67.140	<0.001	<0.001
896.7717	70.064	<0.001	<0.001
898.7908	71.332	<0.001	0.025
905.8284	77.921	0.004	0.002
926.8183	75.606	<0.001	0.406

Table A.5: *Table showing the significantly different features between albumen organic extracts that were stored either dried or re-dissolved at -25 °C, -46 °C, and -80 °C for forty-four weeks. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance. 25, 46, and 80 refer to the storage of re-dissolved extracts at -25 °C, -46 °C, and -80 °C respectively. 25*, 46* and 80* refer to the storage of dried extracts at the same temperatures.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend 25 | 25* | 46 | 46* | 80 | 80*
169.1102	0.843	10.68	0.006	
250.1778	10.199	6.38	0.004	
266.1558	11.213	5.50	0.010	
271.2627	15.062	9.24	<0.001	
288.2895	15.047	10.69	<0.001	
294.1547	0.875	8.20	<0.001	
299.1209	0.843	11.58	<0.001	
299.2941	18.502	7.77	0.004	
310.2376	2.219	11.67	0.001	
316.3208	18.509	6.86	0.002	
328.1362	0.910	11.77	<0.001	
331.2842	8.241	4.89	0.011	
371.3144	12.505	6.84	0.008	
409.3110	11.215	5.23	0.002	
419.3151	22.903	22.41	0.007	
581.5404	27.584	6.31	0.008	
617.5126	25.801	8.17	0.001	
623.3403	2.865	7.98	0.002	
647.4600	26.418	7.23	<0.001	
674.4642	6.548	6.32	<0.001	
685.4206	26.410	10.82	0.006	
743.6370	12.423	2.27	0.001	
934.6395	16.654	4.01	0.001	

Appendix B

Table B.1: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (21-26 weeks old). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
331.2847	14.512	10.24	<0.001
359.3165	23.642	8.19	0.001
369.3520	33.319	10.53	<0.001
387.1752	1.211	2.46	<0.001
409.3443	33.319	5.68	<0.001
482.3252	17.433	4.49	<0.001
496.3412	8.548	3.61	<0.001
520.3413	7.975	2.22	<0.001
522.3564	10.246	6.24	0.001
610.5425	47.541	6.88	<0.001
612.5573	50.806	5.71	0.007
634.5425	45.863	2.72	0.009
638.5730	51.778	4.08	0.005
640.5887	54.517	6.88	0.004
642.5161	49.004	2.52	<0.001
648.6307	54.607	6.83	0.007
683.5448	14.512	3.29	<0.001
703.5761	38.996	3.40	<0.001
716.5243	42.043	3.89	<0.001
718.5398	45.190	3.63	<0.001
731.6074	43.777	4.47	<0.001
732.5551	40.457	13.30	<0.001
734.5709	43.913	4.30	<0.001
739.6071	23.642	3.89	<0.001
740.5239	41.831	2.95	<0.001
744.5551	39.705	4.55	0.007
746.5714	42.797	3.04	<0.001

Table B.1 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value
756.5555	40.185	3.34	<0.001
760.5932	45.474	8.65	<0.001
762.6029	48.549	4.11	<0.001
764.5392	42.863	24.48	<0.001
773.5895	44.290	6.24	0.008
774.6031	47.479	4.99	<0.001
780.5549	38.189	4.92	<0.001
782.5708	41.419	1.90	0.005
784.5864	43.665	2.99	0.007
795.7003	33.318	14.45	0.001
809.5900	42.777	3.09	0.005
810.6023	46.092	2.17	0.003
813.6853	53.126	8.07	<0.001
851.7120	68.925	1.92	<0.001
872.7720	69.604	6.59	0.004
896.7732	69.097	9.15	0.010
900.7946	70.244	8.11	<0.001
900.8045	72.860	9.52	0.001
922.7913	70.591	10.03	0.008

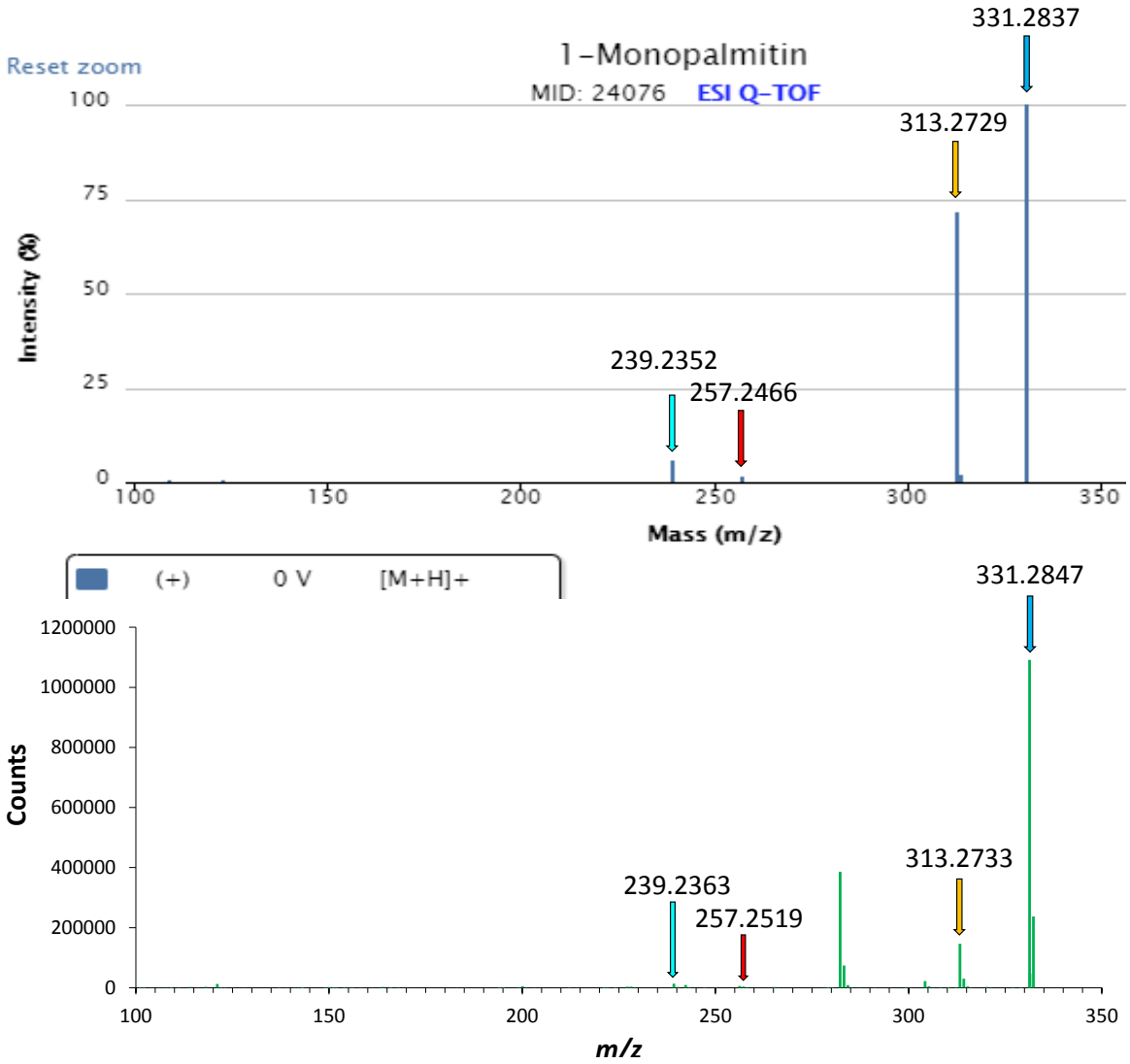


Figure B.1: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds 1-monopalmitin and m/z 331.2847 respectively*

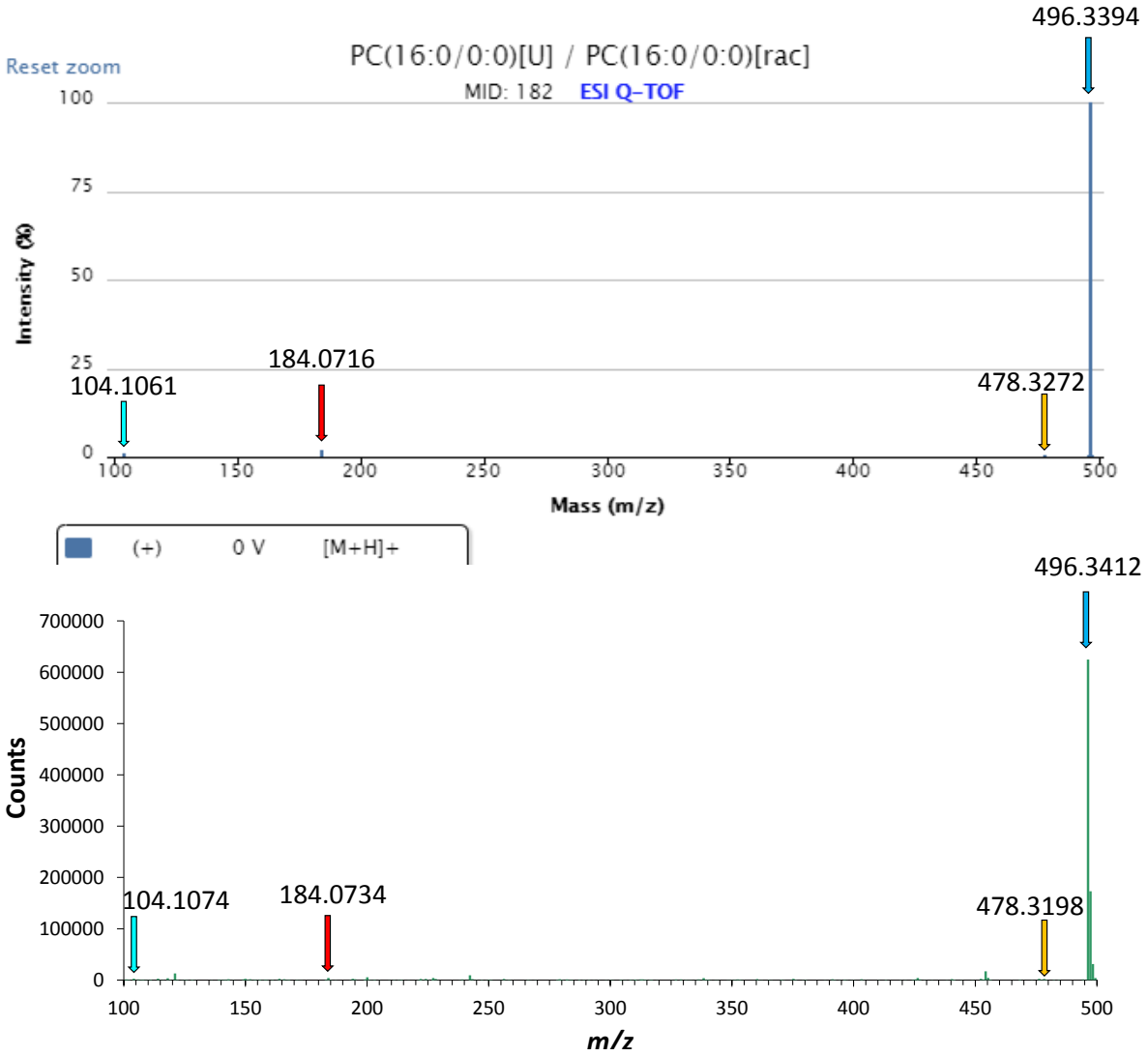


Figure B.2: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PC(16:0/0:0) and m/z 496.3412 respectively*

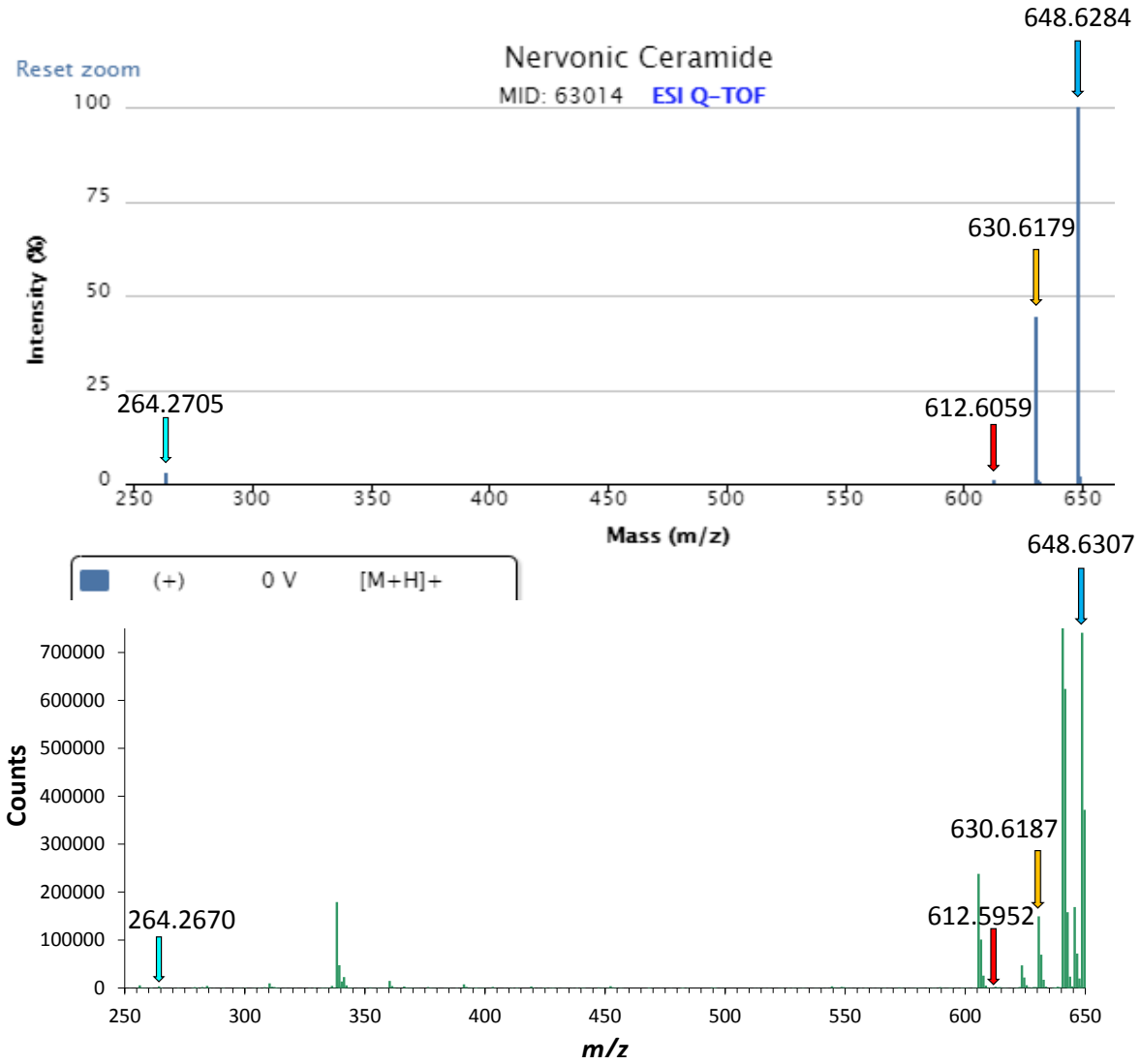


Figure B.3: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds nervonic ceramide and m/z 648.6307 respectively*

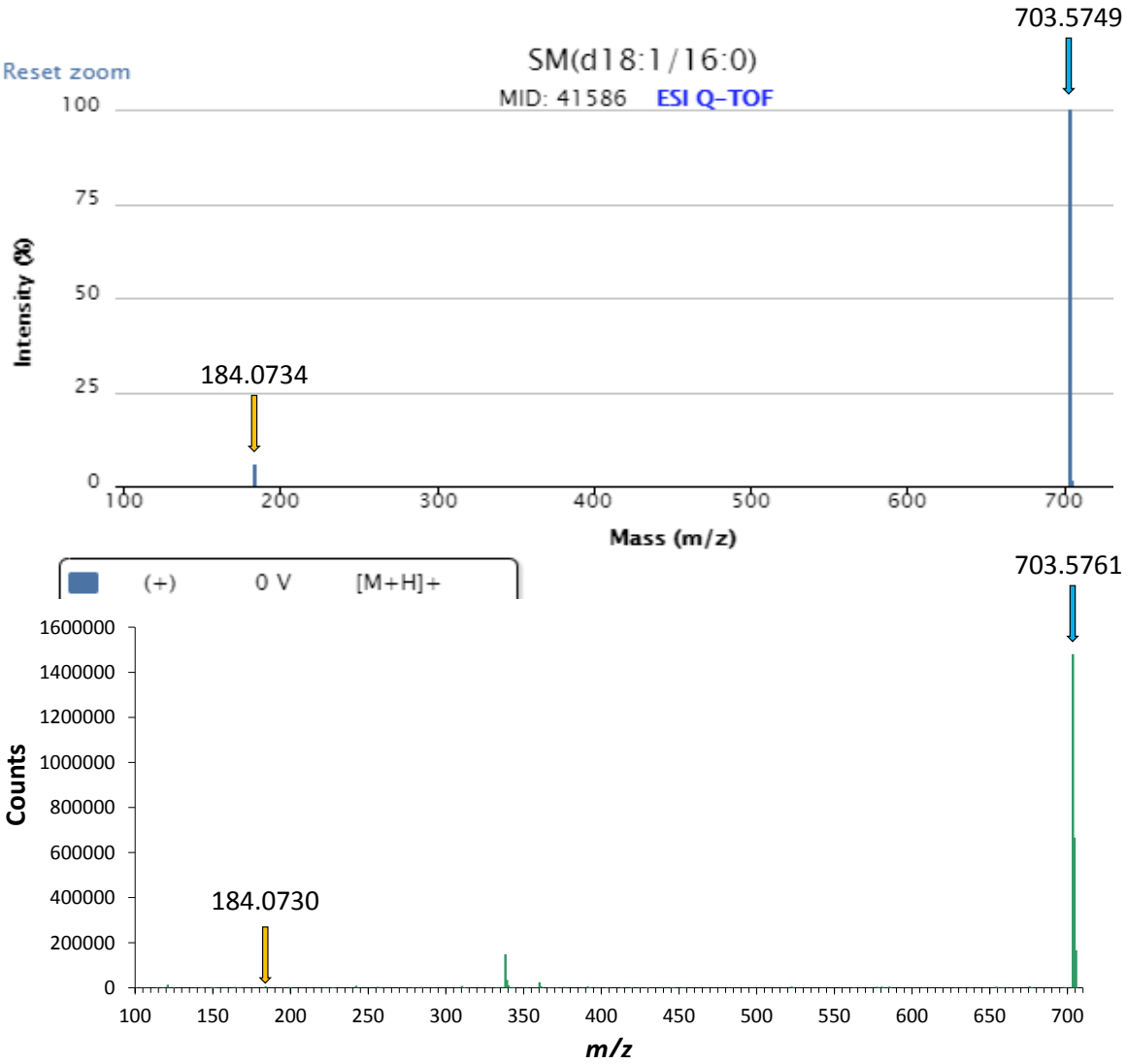


Figure B.4: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds SM(18:1/16:0) and m/z 703.5761 respectively*

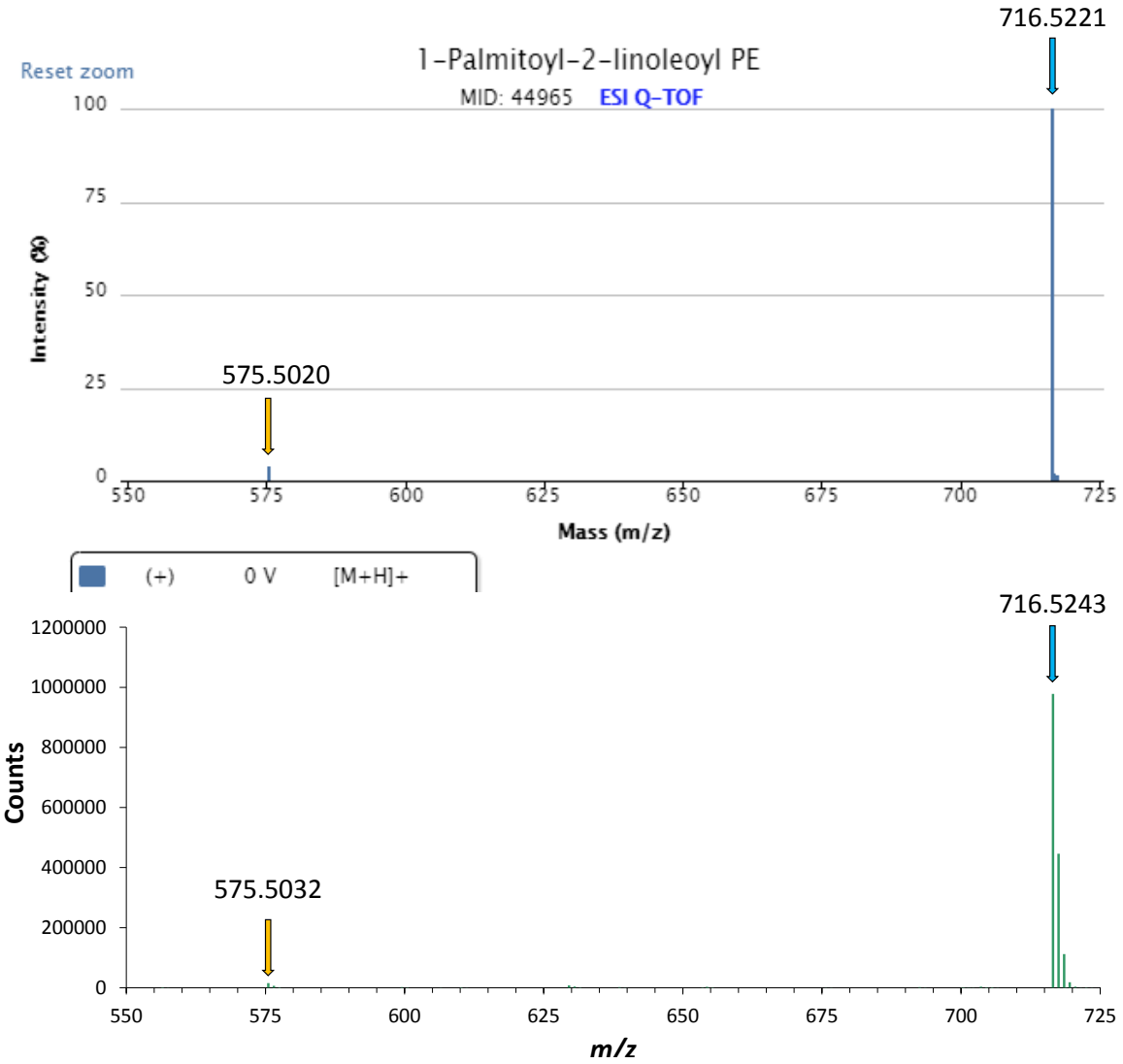


Figure B.5: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PE(16:0/18:2) and m/z 716.5243 respectively*

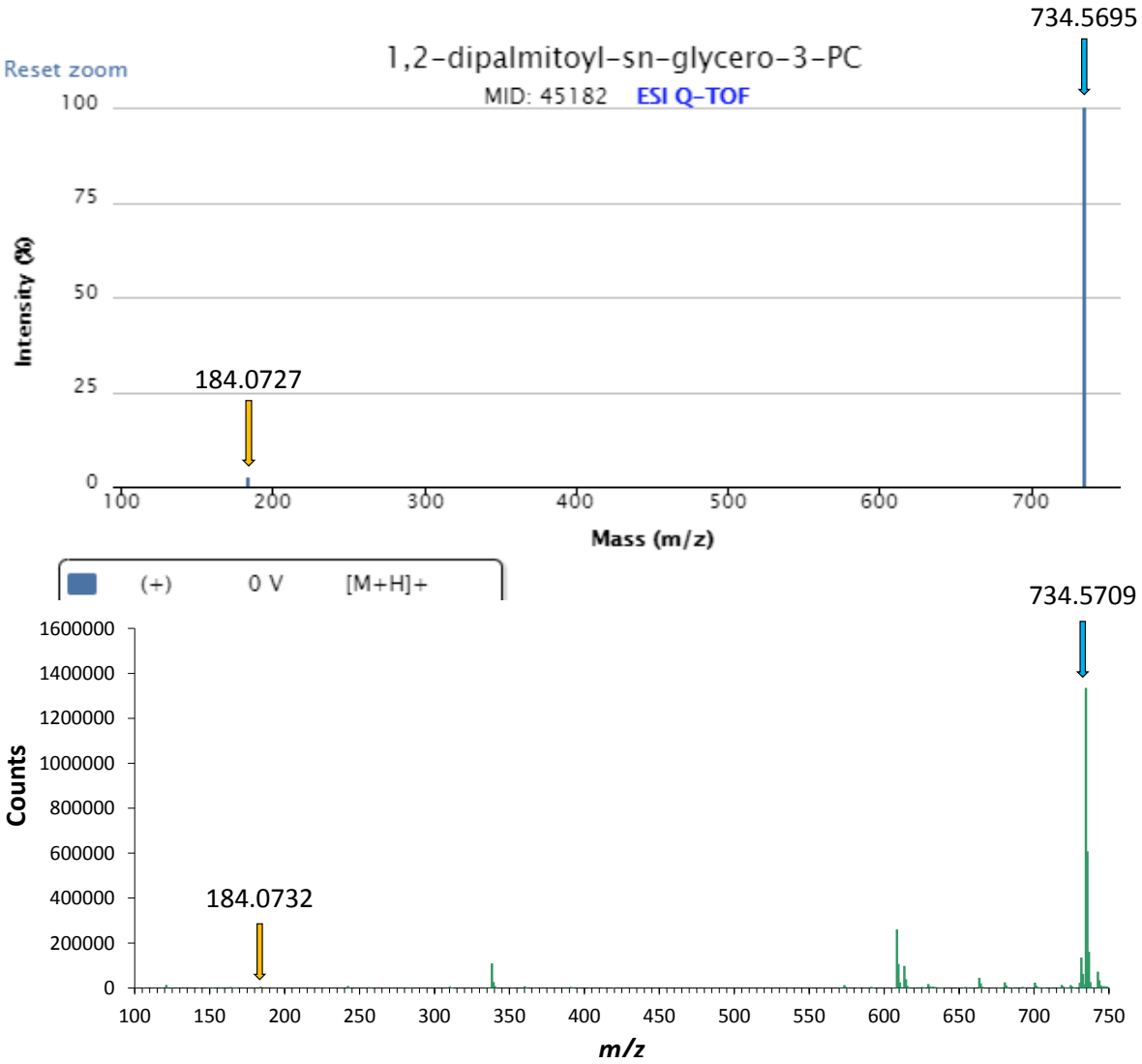


Figure B.6: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PC(16:0/16:0) and m/z 734.5709 respectively*

Table B.2: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (26-30 weeks old). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
282.2792	14.402	1.86	0.002
313.2733	14.435	3.81	<0.001
331.2847	14.435	2.48	<0.001
338.3420	26.985	0.71	<0.001
359.3162	23.513	1.04	<0.001
369.3520	33.209	1.19	0.001
387.1735	1.206	5.38	<0.001
409.3443	33.209	2.44	0.004
454.2932	9.871	1.05	0.001
482.3251	17.339	3.52	0.002
496.3405	9.705	2.30	0.001
520.3400	6.958	2.90	0.003
522.3557	10.204	3.29	0.010
548.5040	47.402	0.76	0.001
577.5181	71.894	3.11	0.007
613.4807	44.297	3.16	0.001
627.5347	39.700	4.52	0.005
630.6182	54.539	3.18	0.009
683.5443	14.436	1.63	<0.001
703.5758	38.969	1.64	0.003
732.5903	51.933	5.06	0.009
739.6069	23.515	2.54	<0.001
740.5238	41.774	2.61	0.002
744.5673	47.153	11.20	<0.001
767.5419	42.837	21.57	<0.001
768.5900	44.480	3.25	0.009
778.5756	48.547	2.14	0.001
782.5708	41.359	2.18	<0.001
782.6057	52.019	22.16	<0.001
786.6022	46.240	24.81	0.003
794.5750	47.119	14.23	0.006
816.6480	53.792	7.28	0.010
824.6508	52.654	8.29	0.004
846.7560	68.732	0.87	0.002
848.7718	71.122	1.81	0.005
874.7880	71.902	1.10	0.007
876.7148	67.379	26.22	0.002
877.7281	69.217	4.11	0.010

Table B.2 continued

Feature *m/z*	**RT median (minutes)**	**CV%**	***p*-value**
878.8031	71.885	1.46	0.010
878.8096	74.873	1.74	0.006
880.7453	71.580	3.62	0.009
880.8184	74.871	1.89	0.010
904.5917	39.700	4.13	0.003
909.7889	79.720	4.00	0.001
917.8290	74.874	1.72	0.007

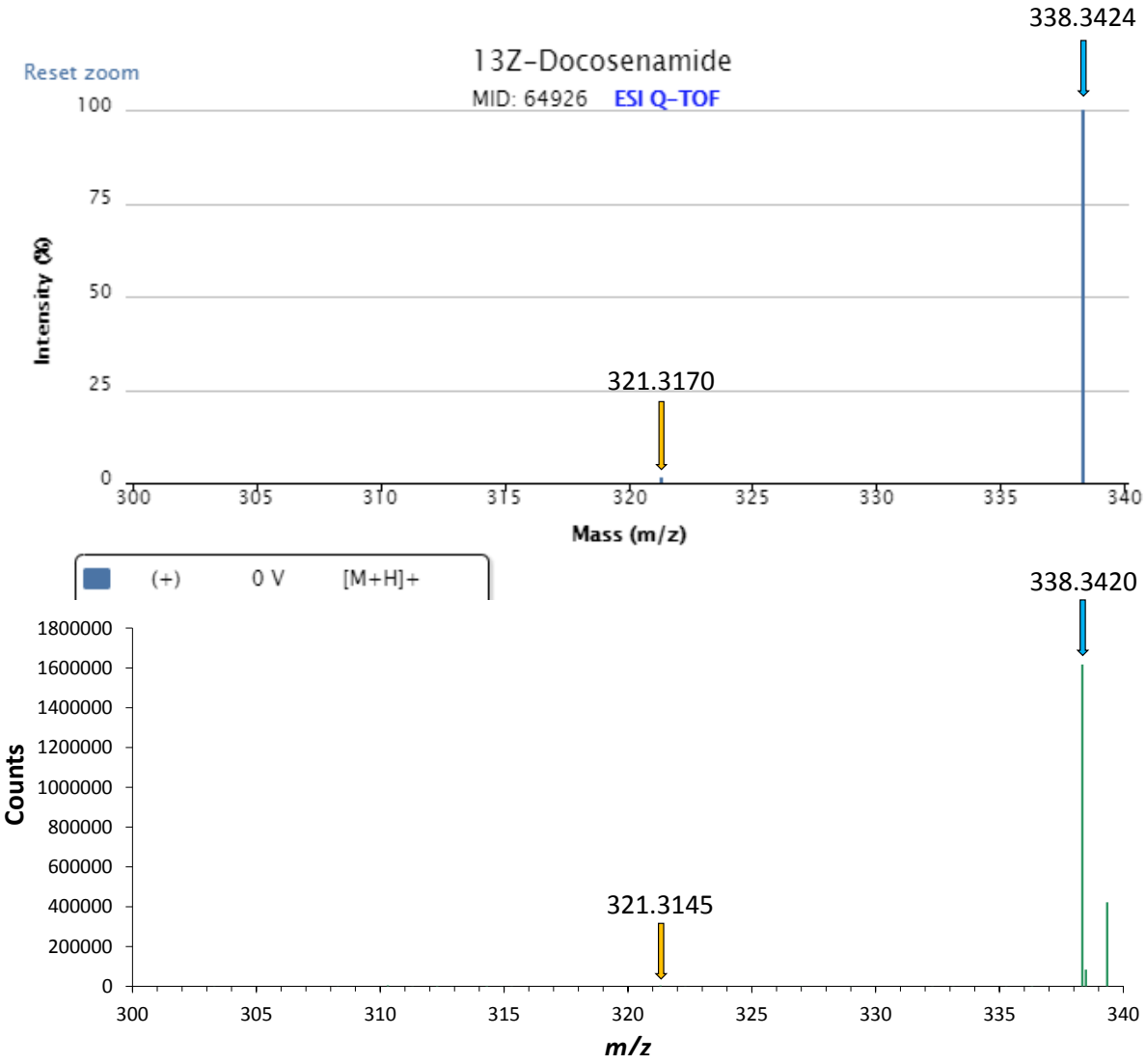


Figure B.7: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds docosenamide and m/z 338.3420 respectively*

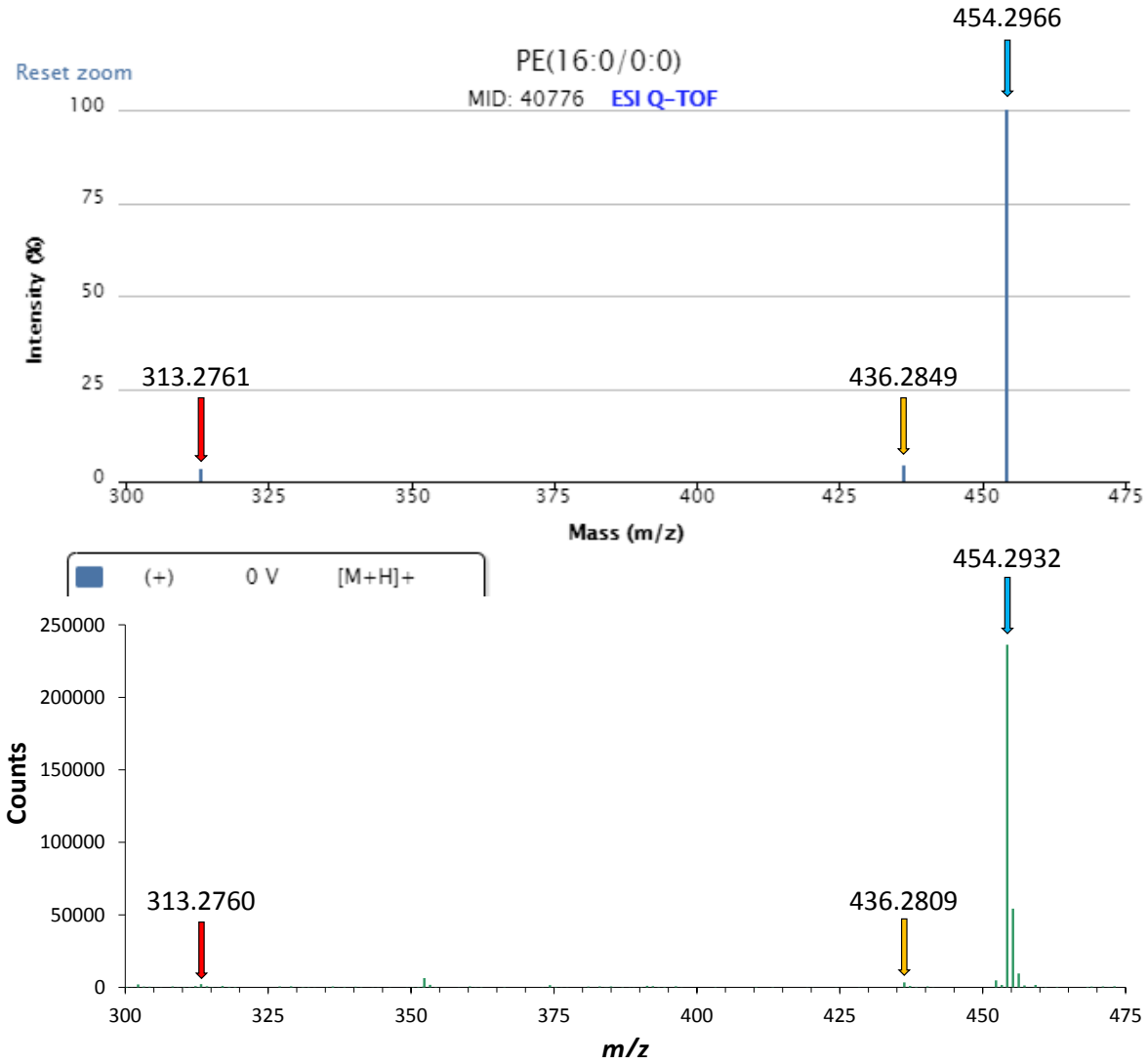


Figure B.8: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PE(16:0/0:0) and m/z 454.2932 respectively*

Table B.3: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (30-38 weeks old). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
331.2848	14.004	7.99	0.001
338.3412	43.652	5.46	0.003
338.3415	68.740	8.23	0.006
341.3044	22.865	6.99	<0.001
387.1751	1.206	3.69	<0.001
548.5038	46.762	3.07	0.004
575.5035	34.690	3.80	0.002
599.5035	34.560	3.61	<0.001
627.5349	39.400	4.98	0.001
650.4388	16.189	11.17	<0.001
654.3322	14.312	3.51	<0.001
675.5437	33.542	3.87	0.006
675.6763	27.135	6.25	0.008
706.5449	38.630	6.08	0.001
734.5610	39.709	1.17	0.009
740.5233	41.078	4.68	0.01
752.5228	34.524	22.00	0.005
754.5378	36.312	1.75	0.008
762.6020	47.804	1.17	0.004
764.5249	40.526	1.09	0.002
778.5383	35.942	2.92	0.004
780.5539	39.029	0.84	0.005
787.6061	46.129	4.33	0.004
790.6243	48.901	2.46	0.005
792.5566	45.117	1.47	0.007
806.5548	35.483	20.34	<0.001
808.5695	37.208	1.85	<0.001
824.7704	72.795	3.39	0.003
859.5331	34.561	1.18	<0.001
880.7454	71.719	3.22	0.008
884.7772	77.945	18.65	0.034
887.5644	39.397	5.27	0.002
888.8042	73.225	6.95	0.001
890.8184	76.537	9.42	0.008
898.7876	69.706	1.55	0.008
900.8040	72.180	2.80	0.006
901.7286	67.697	2.17	0.005
906.8299	75.223	4.17	0.005

Table B.3 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value
906.8482	83.802	11.17	0.001
926.8178	73.274	3.35	0.009
928.8326	76.582	8.41	<0.001

Table B.4: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (38-44 weeks old). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
105.0663	1.208	3.36	<0.001
147.1075	0.793	21.36	0.003
150.0547	0.826	5.60	0.008
237.1120	1.888	7.07	<0.001
247.2261	4.050	5.05	0.003
267.1187	1.224	6.34	0.033
267.2674	23.493	3.17	0.001
286.2286	12.530	7.37	0.006
310.2377	3.829	7.06	<0.001
352.3055	4.162	4.90	<0.001
359.3153	21.807	6.35	0.003
360.3189	23.484	3.16	0.002
381.2978	23.484	1.71	0.001
387.1770	1.208	5.44	0.003
396.3317	4.196	4.57	<0.001
409.1588	1.175	5.46	0.001
419.3154	27.019	10.34	0.004
429.3720	36.842	9.50	0.004
429.3723	32.962	4.09	<0.001
429.3724	30.803	5.36	<0.001
440.3578	4.229	4.78	<0.001
445.3673	29.175	4.09	<0.001
464.2820	1.490	3.24	0.001
489.2269	5.290	3.20	<0.001
580.3603	10.221	7.26	0.007
604.3844	4.710	4.35	0.009
612.5475	47.754	10.77	0.006
637.3052	14.653	2.64	<0.001
642.4364	28.729	3.73	0.008
650.4385	16.746	8.08	<0.001
666.4336	16.030	4.14	<0.001
688.5872	51.903	10.61	0.004
694.4652	24.574	5.30	0.007
734.5332	38.936	4.72	0.007
739.6059	23.492	4.29	0.009
764.5247	41.179	1.38	0.003
774.5630	32.962	1.72	<0.001
790.3784	1.208	8.57	0.003
792.5653	29.991	6.31	<0.001

Table B.4 continued

Feature *m/z*	**RT median (minutes)**	**CV%**	***p*-value**
792.5741	32.116	2.04	<0.001
806.5543	27.120	6.93	<0.001
806.5544	36.238	23.91	<0.001
808.5693	38.722	2.79	<0.001
808.7383	69.164	20.58	0.004
809.5724	37.958	2.45	<0.001
811.3086	1.175	4.61	0.002
814.5576	28.862	23.67	0.009
820.5958	33.458	9.71	0.007
822.6336	50.198	27.57	0.006
834.5838	40.255	2.91	<0.001
836.5986	42.462	0.93	0.006
868.7317	65.042	3.25	0.006
878.7442	59.190	16.23	0.001
904.7455	70.117	27.44	0.003
904.7595	59.906	22.39	0.001
906.7760	61.558	13.96	0.003
931.7764	73.444	25.28	<0.001

Table B.5: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens of different ages (21-44 weeks old). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
331.2851	13.676	5.56	<0.001
359.3166	22.350	3.49	<0.001
369.3526	32.344	5.53	<0.001
387.1780	1.203	4.16	<0.001
454.2937	9.353	2.34	0.001
494.3245	5.984	3.51	<0.001
524.3717	14.599	2.76	0.005
575.5037	46.834	18.38	0.026
577.5204	49.805	3.08	<0.001
603.5365	50.884	5.46	<0.001
605.5514	53.658	2.63	<0.001
610.5418	46.778	11.21	0.028
612.5573	49.723	1.80	<0.001
614.5729	52.674	2.10	<0.001
615.4965	46.796	1.51	0.007
617.5122	49.788	1.19	<0.001
636.5574	47.881	2.47	0.002
638.5730	50.752	1.58	<0.001
640.5887	53.591	1.07	<0.001
641.5125	47.881	4.15	0.001
703.5762	37.838	1.20	<0.001
718.5396	44.155	2.71	0.001
730.5752	50.520	5.07	0.001
732.5553	39.483	21.44	<0.001
748.5288	42.529	1.67	0.006
760.5856	46.051	9.50	0.01
761.5798	40.494	2.64	0.005
762.5925	44.437	6.38	0.005
762.6016	47.672	2.18	<0.001
764.5238	40.178	2.48	0.001
766.5387	41.732	1.63	0.002
772.5864	43.314	0.97	0.001
778.5754	47.632	2.49	<0.001
780.5550	37.257	1.91	<0.001
782.5703	39.695	6.29	0.004
782.6064	51.367	4.29	0.002
783.5753	38.999	0.39	0.002
784.5768	39.007	1.36	0.003
786.6021	47.136	3.07	<0.001

Table B.5 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value
788.6178	48.146	2.65	0.002
788.6541	51.883	12.68	0.001
806.5708	39.946	0.94	0.009
810.6011	43.348	3.11	<0.001
812.6177	47.666	6.75	<0.001
813.6855	52.605	14.08	<0.001
818.7256	65.684	6.72	0.004
820.7434	67.561	2.98	0.001
824.6458	52.240	2.55	0.003
846.7468	66.306	12.59	<0.001
868.7415	65.244	3.75	0.002
870.7576	66.919	2.88	0.006
886.7870	69.950	5.38	<0.001
888.8036	72.497	4.20	0.003
894.7568	65.807	3.89	<0.001
896.7727	67.447	10.22	<0.001
898.7875	69.251	2.91	<0.001
903.8131	71.549	3.90	0.008
926.8185	73.826	4.98	0.003

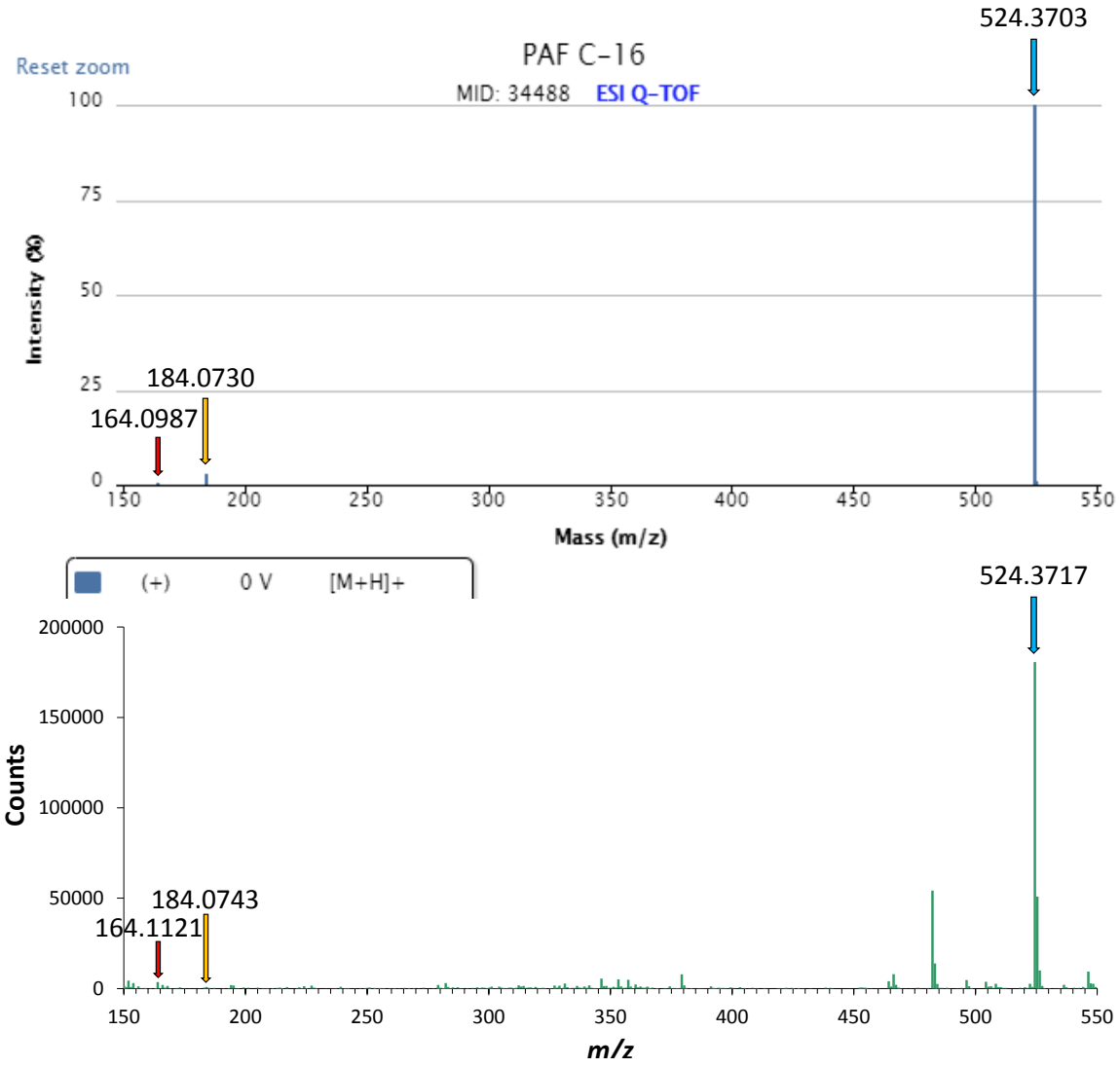


Figure B.9: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PAF C-16 and m/z 524.3717 respectively*

Table B.6: *Table showing the significantly different features in albumen organic extracts between eggs that were laid by hens of different ages (21-44 weeks old). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
203.0525	0.775	6.64	<0.001
267.2680	12.314	10.38	0.005
288.2894	14.908	14.38	<0.001
310.2373	2.185	13.89	<0.001
310.3104	11.649	12.27	<0.001
312.3256	15.043	13.01	<0.001
316.3208	18.410	14.27	<0.001
331.2840	8.158	13.19	<0.001
336.3253	12.863	12.47	<0.001
338.3418	15.241	9.40	<0.001
350.3408	14.828	11.74	<0.001
359.3153	12.305	10.67	0.007
366.3724	18.211	12.03	<0.001
387.1789	1.008	12.20	<0.001
390.3335	14.827	12.49	<0.001
396.3937	15.324	12.05	0.001
409.1605	0.975	15.48	<0.001
425.1338	0.958	12.59	<0.001
537.5342	23.619	10.33	<0.001
548.5027	24.697	11.06	0.001
551.5024	25.604	13.85	<0.001
565.5660	25.230	12.01	<0.001
579.5343	27.484	27.71	<0.001
586.5402	25.626	10.91	<0.001
593.5973	27.103	15.31	<0.001
604.3836	2.210	12.17	<0.001
614.5718	27.518	19.39	<0.001
619.5269	27.502	14.02	<0.001
647.5583	29.940	21.91	0.002
654.3314	6.217	11.97	<0.001
663.4535	22.491	11.70	<0.001
686.3575	7.228	11.97	<0.001
690.5997	11.610	12.83	0.001
739.6054	11.558	16.25	0.003
777.3290	22.797	10.81	0.009

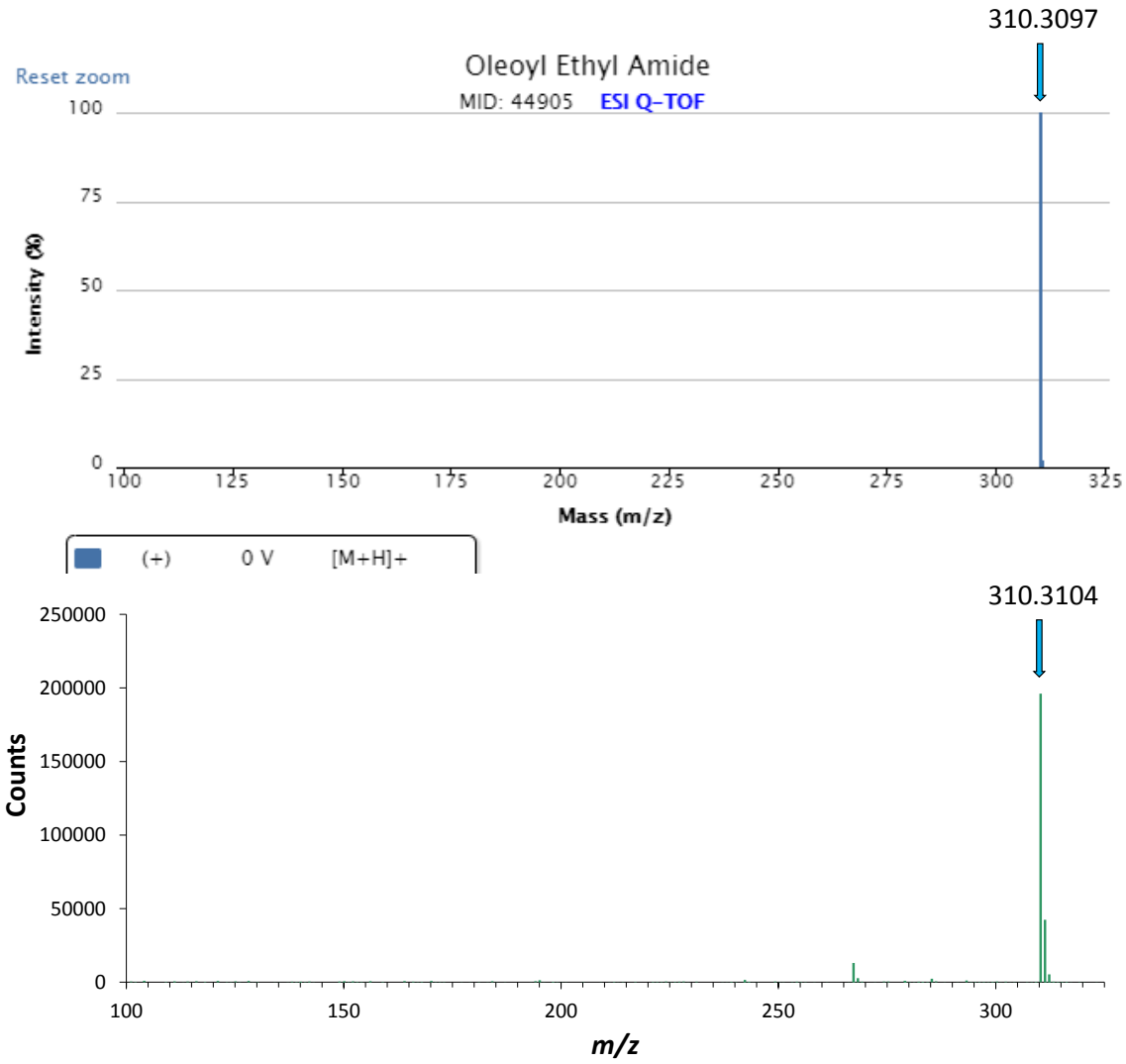


Figure B.10: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds oleoylethylamide and m/z 310.3104 respectively*

Appendix C

Table C.1: *Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed four different diets, two weeks post diet change. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend A B C D
185.1170	2.917	5.47	0.004	
191.0852	2.567	5.25	<0.001	
250.1773	18.904	2.10	0.002	
301.1417	2.734	4.40	0.005	
302.3054	25.817	7.40	<0.001	
303.2528	7.475	4.72	0.001	
308.2799	4.007	6.79	<0.001	
352.3057	4.046	4.46	<0.001	
359.3157	22.649	3.55	0.003	
381.2973	21.025	3.83	<0.001	
387.1735	1.190	4.87	0.001	
396.3314	4.088	5.25	<0.001	
425.1394	1.787	3.14	<0.001	
445.3672	28.936	5.75	0.008	
489.2267	5.142	3.70	0.001	
522.5965	38.157	10.77	0.010	
577.5186	70.547	3.02	0.003	
637.3050	14.368	3.50	0.001	

Table C.1 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend A B C D
642.4361	28.543	1.35	0.002	
670.4673	30.175	1.49	0.004	
684.5554	47.360	8.23	0.001	
684.6258	55.980	1.51	0.009	
716.5653	36.094	2.65	0.006	
739.6053	22.650	4.26	0.010	
743.5341	38.572	8.79	0.002	
759.5735	41.162	1.47	0.029	
771.5808	44.135	8.30	0.005	
780.7070	66.672	7.93	0.003	
822.6680	64.765	16.97	0.007	
839.5539	35.237	9.06	0.001	
843.5807	40.282	4.62	0.005	
846.7571	68.795	2.15	0.007	
868.7327	65.085	5.22	0.009	
905.5947	39.448	1.77	0.008	
933.6248	43.246	2.68	0.010	
933.8629	74.654	4.68	0.004	
933.8662	79.855	8.71	0.005	
944.7740	67.958	3.47	<0.001	
946.7404	72.292	4.70	0.004	
967.6956	74.699	5.93	0.009	
970.7905	68.503	17.99	0.001	
977.8578	68.812	3.87	0.008	
992.8751	72.206	3.22	0.008	

Table C.2: *Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed the four different diets two weeks post diet change, as well as one week prior to diet change. p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	2 weeks post diet change *p*-value	1 week pre diet change *p*-value
185.1170	2.917	0.004	0.348
191.0852	2.567	<0.001	0.002
250.1773	18.904	0.002	0.565
301.1417	2.734	0.005	0.008
302.3054	25.817	<0.001	0.007
303.2528	7.475	0.001	0.004
308.2799	4.007	<0.001	0.010
352.3057	4.046	<0.001	0.001
359.3157	22.649	0.003	0.002
381.2973	21.025	<0.001	0.003
387.1735	1.190	0.001	0.090
396.3314	4.088	<0.001	0.002
445.3672	28.936	0.008	0.101
489.2267	5.142	0.001	0.001
637.3050	14.368	0.001	0.005
642.4361	28.543	0.002	0.008
670.4673	30.175	0.004	0.020
684.6258	55.980	0.009	0.243
716.5653	36.094	0.006	0.787
739.6053	22.650	0.010	0.005
905.5947	39.448	0.008	0.532
933.6248	43.246	0.010	0.275
944.7740	67.958	<0.001	0.288

Table C.3: *Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed four different diets, five weeks post diet change. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend A B C D
310.3106	23.254	2.31	<0.001	
331.2843	14.441	3.51	0.001	
338.3423	27.022	1.18	<0.001	
359.3155	23.553	5.91	0.002	
366.3731	30.308	1.61	0.003	
387.1711	1.210	5.15	<0.001	
520.3398	7.025	3.21	0.002	
575.5033	35.886	7.55	<0.001	
603.5346	40.467	5.36	0.001	
604.3842	4.729	1.27	0.005	
604.6028	54.311	6.33	0.003	
618.6183	56.071	5.22	<0.001	
632.6344	57.731	7.20	0.007	
642.5154	49.417	23.51	0.019	
760.5857	45.730	2.03	0.006	
776.5632	48.274	1.92	0.01	
782.5709	41.746	5.10	0.008	
782.6058	52.469	2.55	0.005	
786.6021	46.653	2.96	<0.001	
799.6049	50.029	14.32	0.007	
808.5827	47.272	2.83	0.001	
837.6202	49.498	9.79	0.002	
850.6994	67.491	0.54	0.007	
877.7299	70.438	2.34	0.004	
880.7473	73.119	3.22	0.009	
893.7025	70.467	4.07	0.003	

Table C.3 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend A B C D
895.7160	72.532	4.05	0.008	
897.7325	75.716	3.23	0.005	
900.7529	80.225	3.72	0.002	
991.6720	9.758	2.18	0.007	

Table C.4: *Table showing the significantly different features in yolk organic extracts between eggs laid by birds that were fed the four different diets five weeks post diet change, as well as one week prior to diet change. The p-value was calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	5 weeks post diet change *p*-value	1 week pre diet change *p*-value
310.3106	23.25	<0.001	0.336
331.2843	14.44	0.001	0.021
338.3423	27.02	<0.001	0.362
359.3155	23.55	0.002	0.002
387.1711	1.21	<0.001	0.090
520.3398	7.03	0.002	0.197
575.5033	35.89	<0.001	0.744
603.5346	40.47	0.001	0.541
604.3842	4.73	0.005	0.247
604.6028	54.31	0.003	0.017
618.6183	56.07	<0.001	0.211
632.6344	57.73	0.007	0.007
642.5154	49.42	0.019	0.126
760.5857	45.73	0.006	0.195
776.5632	48.27	0.010	0.738
782.5709	41.75	0.008	0.167
782.6058	52.47	0.005	0.154
808.5827	47.27	0.001	0.191
837.6202	49.50	0.002	0.616
991.6720	9.76	0.007	0.019

Table C.5: *Table showing the significantly different features in albumen organic extracts between eggs laid by birds that were fed four different diets, five weeks post diet change. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend A B C D
105.0695	0.988	20.37	<0.001	
202.1799	0.756	7.50	<0.001	
209.1577	9.697	1.49	0.003	
226.1800	10.704	3.12	0.001	
267.1228	0.988	19.89	<0.001	
267.2679	11.922	2.11	0.008	
274.2736	10.246	2.91	<0.001	
288.2890	14.499	11.99	0.002	
302.3048	14.226	3.22	<0.001	
316.2108	0.772	9.88	<0.001	
316.3208	18.058	15.85	0.002	
321.3152	14.857	3.08	0.016	
338.3410	15.404	7.05	0.002	
341.3087	11.190	13.93	0.006	
362.3385	17.793	2.77	0.004	
368.3879	20.317	4.25	0.007	
383.3984	11.285	2.57	0.005	
387.1794	0.988	4.92	<0.001	
397.4105	14.939	21.01	0.008	
397.4131	11.287	4.15	0.003	
408.3081	9.693	4.58	<0.001	
409.3285	15.487	3.46	0.001	
418.2191	1.038	4.18	0.002	
425.1338	0.938	4.22	0.002	
489.2264	2.746	3.80	<0.001	
519.4356	9.687	3.08	<0.001	

Table C.5 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value	Trend A B C D
543.1256	0.839	6.20	0.003	
557.1529	0.822	10.94	0.003	
642.4358	16.605	1.78	<0.001	
649.6596	14.741	7.32	0.005	
654.3311	6.013	1.72	<0.001	
659.2870	6.003	2.03	<0.001	
686.3568	6.973	2.66	<0.001	
691.3127	6.960	4.09	<0.001	
795.3347	1.021	4.58	<0.001	
811.3092	0.955	4.82	<0.001	
955.5692	16.324	5.10	<0.001	

Table C.6: *Table showing the significantly different features in albumen organic extracts between eggs laid by birds that were fed the four different diets five weeks post diet change, as well as one week prior to diet change. The p-value was calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	5 weeks post diet change *p*-value	1 week pre diet change *p*-value
105.0695	0.99	<0.001	0.540
202.1799	0.76	<0.001	0.004
209.1577	9.70	0.003	0.253
267.1228	0.99	<0.001	0.848
267.2679	11.92	0.008	0.145
274.2736	10.25	<0.001	0.112
288.2890	14.50	0.002	0.990
302.3048	14.23	<0.001	0.012
316.2108	0.77	<0.001	0.331
316.3208	18.06	0.002	0.702
341.3087	11.19	0.006	0.186
383.3984	11.28	0.005	0.670
387.1794	0.99	<0.001	0.436
409.3285	15.49	0.001	0.599
418.2191	1.04	0.002	0.663
425.1338	0.94	0.002	0.294
489.2264	2.75	<0.001	0.028
519.4356	9.69	<0.001	0.371
642.4358	16.61	<0.001	0.016
654.3311	6.01	<0.001	0.112
659.2870	6.00	<0.001	0.102
686.3568	6.97	<0.001	<0.001
691.3127	6.96	<0.001	<0.001
795.3347	1.02	<0.001	0.900
811.3092	0.95	<0.001	0.244
955.5692	16.32	<0.001	0.511

Appendix D

Table D.1: *Table showing the significantly different features in yolk organic extracts between eggs of different ages stored at 23 ℃. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
104.1070	0.800	2.58	<0.001
331.2860	15.833	3.40	<0.001
338.3443	27.150	2.41	<0.001
341.3063	24.630	3.55	<0.001
360.3204	24.630	2.69	<0.001
369.3519	32.852	4.26	<0.001
387.1805	1.286	1.18	<0.001
518.3239	10.637	18.15	0.001
520.3424	8.761	2.95	0.001
522.3582	12.728	2.56	0.003
524.3737	16.547	4.40	0.010
615.4991	46.645	3.08	0.002
622.6168	52.657	8.10	<0.001
636.6320	54.401	9.67	<0.001
641.5152	47.604	12.03	0.002
648.6329	53.311	4.22	0.005
650.4412	17.988	15.38	<0.001
650.6489	56.081	10.46	<0.001
703.5791	37.935	0.66	0.001
731.6091	42.515	3.31	0.001
732.7675	39.280	25.56	0.008
734.5727	42.626	2.60	0.001
740.5284	40.785	3.45	<0.001
746.6093	46.617	2.40	<0.001
756.5583	31.317	1.70	<0.001
762.6026	46.961	1.77	<0.001
764.5277	40.241	12.29	0.003
772.5891	43.007	0.85	0.006
774.6034	45.977	3.91	<0.001

Table D.1 continued

Feature *m/z*	**RT median (minutes)**	**CV%**	***p*-value**
776.5631	46.593	1.26	0.003
780.5578	37.103	5.04	<0.001
784.5877	42.413	0.78	0.001
787.6721	50.876	4.77	<0.001
794.5735	47.036	2.73	0.005
802.6347	50.022	3.74	0.020
806.5702	42.012	25.86	0.015
808.5880	41.629	1.97	0.009
811.6696	48.765	5.39	<0.001
813.6884	51.540	4.74	<0.001
814.6355	48.973	14.95	<0.001
838.6235	46.945	2.58	0.006

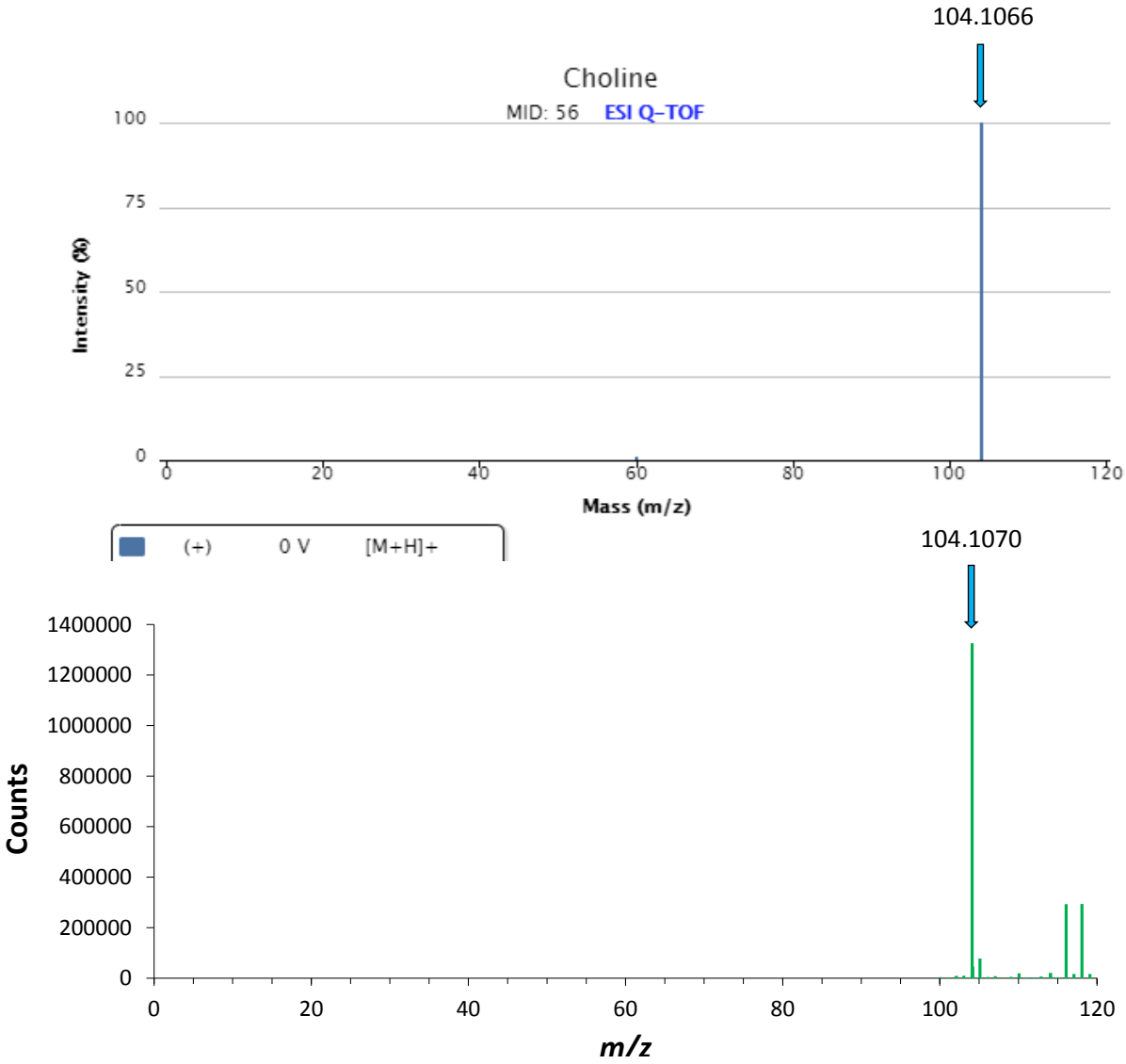


Figure D.1: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds choline and m/z 104.1070 respectively*

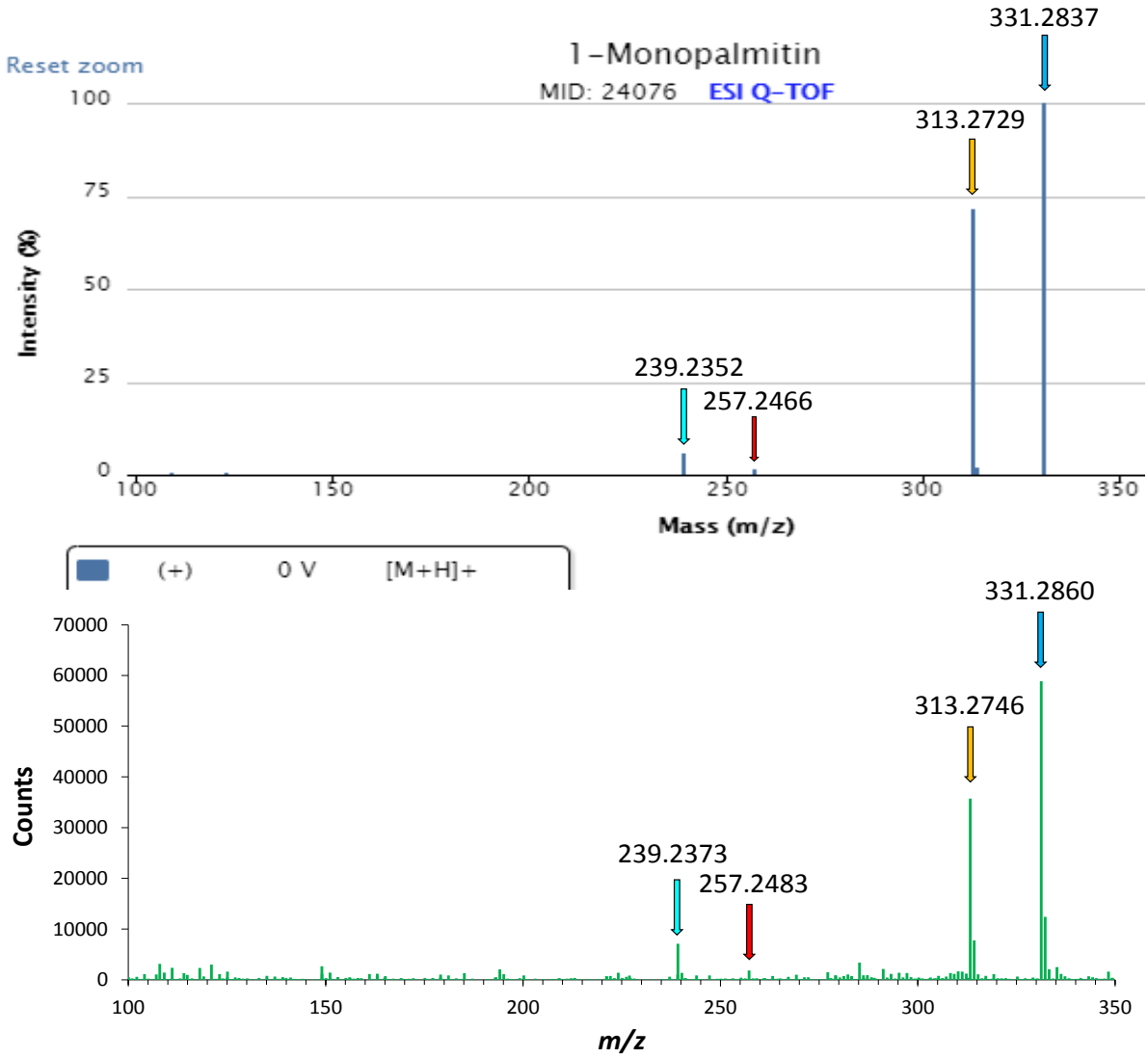
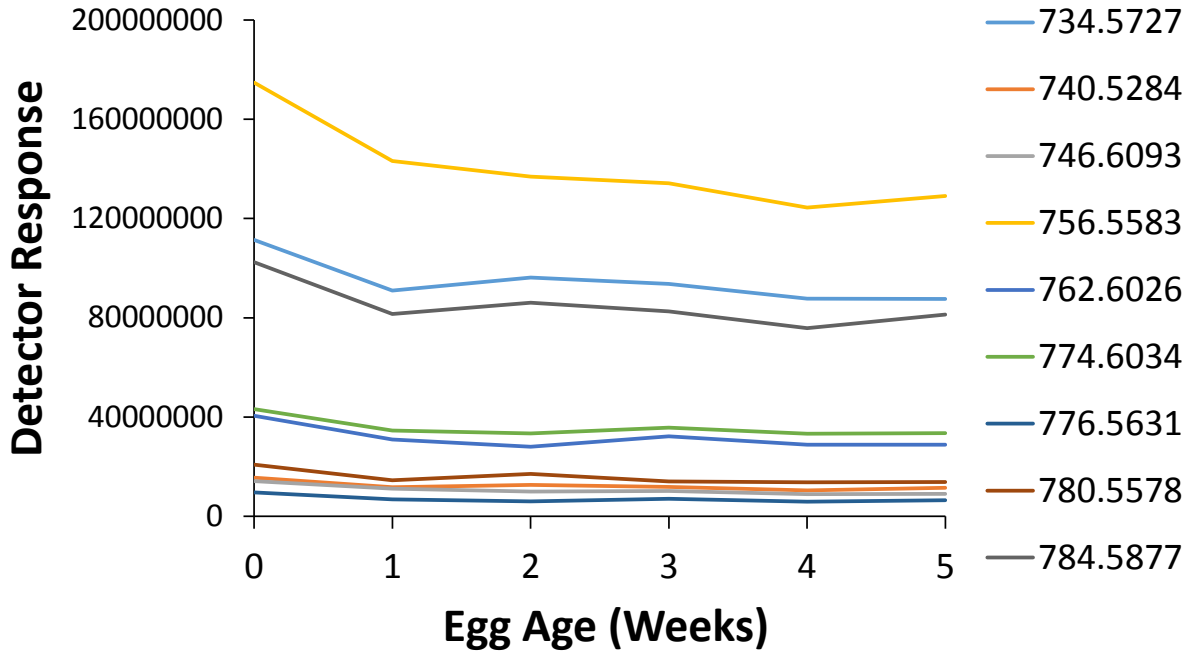
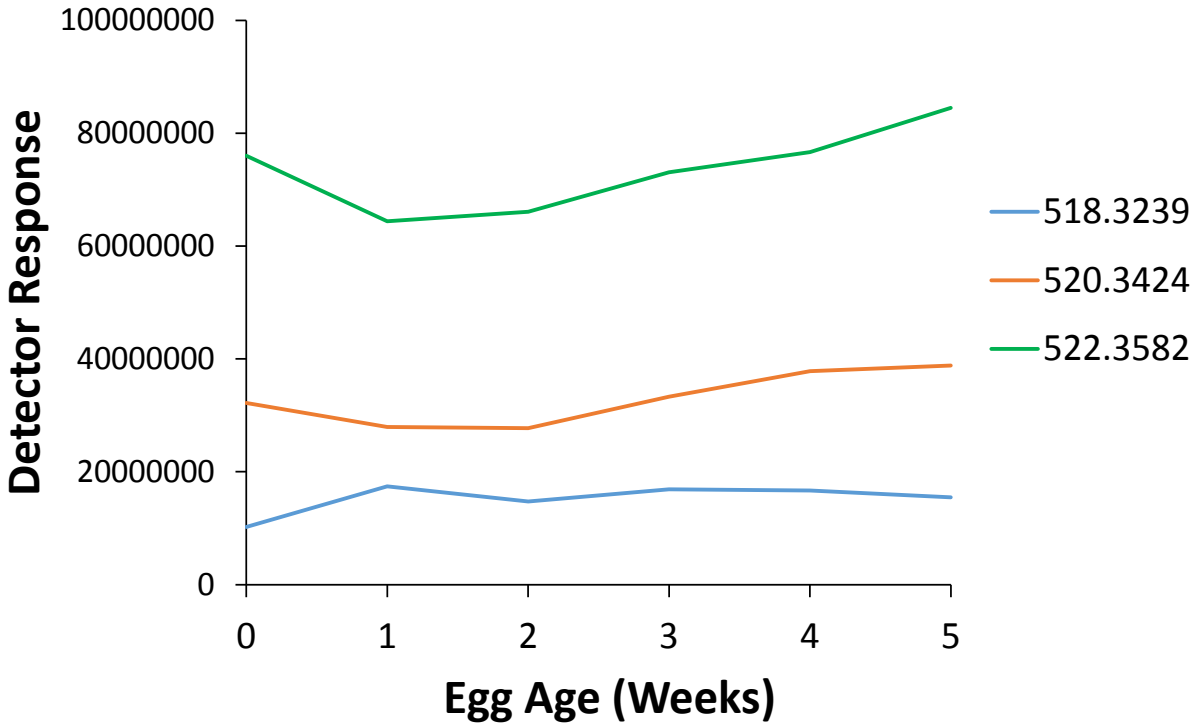


Figure D.2: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds 1-monopalmitin and m/z 331.2860 respectively*



(a) *Line graphs showing trend in changing abundance of compounds classified as potential phosphatidylcholines*



(b) *Line graphs showing trend in changing abundance of compounds classified as potential lysophosphatidylcholines*

Figure D.3

Table D.2: *Table showing concentration of choline in yolk organic extracts, the yolks themselves, and the mean choline concentration in egg yolks from eggs of different ages, stored at 23 °C*

Sample	Choline concentration in extract (µg/mL)	Choline concentration in yolk (µg/g)	Mean choline concentration in yolk (µg/g)
Fresh 1	0.4630	9.2605	6.7758
Fresh 2	0.4034	8.0680	
Fresh 3	0.3087	6.1747	
Fresh 4	0.2155	4.3100	
Fresh 5	0.3354	6.7085	
Fresh 6	0.3067	6.1331	
1 Week 1	0.3516	7.0319	7.0414
1 Week 2	0.4055	8.1092	
1 Week 3	0.3821	7.6429	
1 Week 4	0.2680	5.3601	
1 Week 5	0.3250	6.5004	
1 Week 6	0.3802	7.6036	
2 Weeks 1	0.4113	8.2256	7.9906
2 Weeks 2	0.3939	7.8772	
2 Weeks 3	0.1838	3.6757	
2 Weeks 4	0.7453	14.9057	
2 Weeks 5	0.4010	8.0201	
2 Weeks 6	0.2620	5.2396	
3 Weeks 1	0.6208	12.4155	10.5512
3 Weeks 2	0.5441	10.8828	
3 Weeks 3	0.4634	9.2689	
3 Weeks 4	0.5872	11.7448	
3 Weeks 5	0.5032	10.0637	
3 Weeks 6	0.4466	8.9318	
4 Weeks 1	0.8512	17.0243	11.7145
4 Weeks 2	0.5449	10.8989	
4 Weeks 3	0.3763	7.5257	
4 Weeks 4	0.5910	11.8190	
4 Weeks 5	0.6326	12.6527	
4 Weeks 6	0.5183	10.3663	
5 Weeks 1	0.5383	10.7658	13.2322
5 Weeks 2	0.8325	16.6508	
5 Weeks 3	0.9443	18.8865	
5 Weeks 4	0.4336	8.6715	
5 Weeks 5	0.7061	14.1225	
5 Weeks 6	0.5148	10.2959	

Table D.2 continued

Sample	Choline concentration in extract (µg/mL)	Choline concentration in yolk (µg/g)	Mean choline concentration in yolk (µg/g)
6 Weeks 1	1.1872	23.7446	13.4123
6 Weeks 2	0.4583	9.1654	
6 Weeks 3	0.5068	10.1368	
6 Weeks 4	0.4661	9.3216	
6 Weeks 5	0.6066	12.1329	
6 Weeks 6	0.7986	15.9727	
7 Weeks 1	1.1465	22.9302	16.1104
7 Weeks 2	0.8526	17.0513	
7 Weeks 3	0.8369	16.7390	
7 Weeks 4	0.8358	16.7166	
7 Weeks 5	0.5968	11.9368	
7 Weeks 6	0.5644	11.2884	
8 Weeks 1	0.8822	17.6435	21.5221
8 Weeks 2	0.9522	19.0445	
8 Weeks 3	1.4868	29.7366	
8 Weeks 4	1.4602	29.2040	
8 Weeks 5	0.7297	14.5942	
8 Weeks 6	0.9455	18.9100	
9 Weeks 1	1.0859	21.7185	18.2840
9 Weeks 2	0.9048	18.0967	
9 Weeks 3	0.7665	15.3309	
9 Weeks 4	0.7409	14.8187	
9 Weeks 5	0.9144	18.2874	
9 Weeks 6	1.0726	21.4516	
10 Weeks 1	1.4027	28.0535	20.3955
10 Weeks 2	0.6425	12.8503	
10 Weeks 3	0.8935	17.8706	
10 Weeks 4	1.0995	21.9892	
10 Weeks 5	1.0910	21.8207	
10 Weeks 6	0.9895	19.7890	
11 Weeks 1	0.8926	17.8526	21.0965
11 Weeks 2	0.9839	19.6771	
11 Weeks 3	1.4022	28.0438	
11 Weeks 4	0.7607	15.2140	
11 Weeks 5	1.1215	22.4306	
11 Weeks 6	1.1680	23.3607	

Table D.2 continued

Sample	Choline concentration in extract (µg/mL)	Choline concentration in yolk (µg/g)	Mean choline concentration in yolk (µg/g)
12 Weeks 1	1.6912	33.8238	28.7473
12 Weeks 2	1.7233	34.4655	
12 Weeks 3	1.4349	28.6971	
12 Weeks 4	1.0779	21.5584	
12 Weeks 5	1.4845	29.6903	
12 Weeks 6	1.2124	24.2487	

Table D.3: *Table showing p-values resulting from Tukey tests for yolk choline concentration between all pairwise egg ages. n=6. Orange indicates significance (p<0.05) and green indicates high significance (p<0.01).*

	Fresh	1 Week	2 Weeks	3 Weeks	4 Weeks	5 Weeks	6 Weeks	7 Weeks	8 Weeks	9 Weeks	10 Weeks	11 Weeks	12 Weeks
Fresh		0.750	0.496	0.002	0.007	0.009	0.034	0.001	0.002	<0.001	<0.001	<0.001	<0.001
1 Week			0.572	<0.001	0.006	0.012	0.040	<0.001	0.002	<0.001	<0.001	<0.001	<0.001
2 Weeks				0.156	0.096	0.043	0.081	0.006	0.001	<0.001	0.001	<0.001	<0.001
3 Weeks					0.425	0.152	0.278	0.012	0.008	<0.001	0.001	0.002	<0.001
4 Weeks						0.481	0.535	0.068	0.011	0.004	0.005	0.002	<0.001
5 Weeks							0.951	0.253	0.022	0.031	0.021	0.010	<0.001
6 Weeks								0.371	0.042	0.091	0.048	0.027	0.001
7 Weeks									0.113	0.324	0.141	0.076	0.001
8 Weeks										0.294	0.741	0.896	0.056
9 Weeks											0.395	0.230	0.001
10 Weeks												0.805	0.017
11 Weeks													0.021
12 Weeks													

Table D.4: *Table showing the significantly different features in yolk organic extracts between eggs of different ages stored at 5 ℃. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
331.2857	14.942	1.76	<0.001
338.3435	26.791	1.67	<0.001
359.3174	23.839	1.96	<0.001
369.3530	32.168	3.38	0.004
387.1822	1.258	3.28	<0.001
454.2945	10.234	3.15	0.003
482.3266	17.733	2.76	<0.001
494.3256	6.353	3.38	0.003
496.3420	9.987	2.11	0.006
497.3449	8.731	6.31	<0.001
518.3237	9.955	8.14	<0.001
520.3423	8.181	3.31	0.010
522.3573	10.500	6.38	<0.001
548.5041	48.361	10.35	<0.001
612.5589	48.844	0.97	0.006
622.6164	52.041	1.48	<0.001
639.7720	49.855	17.51	0.038
643.5305	50.053	19.66	0.019
648.6325	52.710	0.94	0.009
650.6481	55.512	1.86	0.003
730.5414	35.321	1.80	0.003
732.5577	38.469	1.38	0.001
734.5729	41.907	1.95	0.001
740.5290	40.080	3.20	0.009
746.6088	45.921	1.15	0.001
756.5580	37.717	1.25	<0.001
761.5912	43.919	2.77	0.009
764.5284	39.507	5.69	<0.001
772.5888	42.306	1.19	0.006
774.6029	45.264	2.65	<0.001
780.5575	36.268	6.61	<0.001
785.5916	41.692	2.43	0.010
791.6336	47.150	1.13	0.001
792.5579	43.955	3.21	0.001
804.5563	35.819	2.04	0.001
810.5947	40.912	1.63	0.008
820.7426	66.738	1.08	<0.001
822.7548	68.962	1.38	0.001
848.7763	69.597	1.25	<0.001
850.7920	72.379	1.62	0.001

Table D.4 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value
850.7920	72.379	1.62	0.001
851.7863	69.589	1.29	<0.001
876.8115	73.182	1.69	0.004
877.7281	67.874	2.65	<0.001
879.7482	70.672	3.35	<0.001
881.7652	73.514	3.04	<0.001
903.7424	68.653	3.09	0.002
905.7651	71.041	28.87	0.002
908.7800	74.192	27.29	0.007

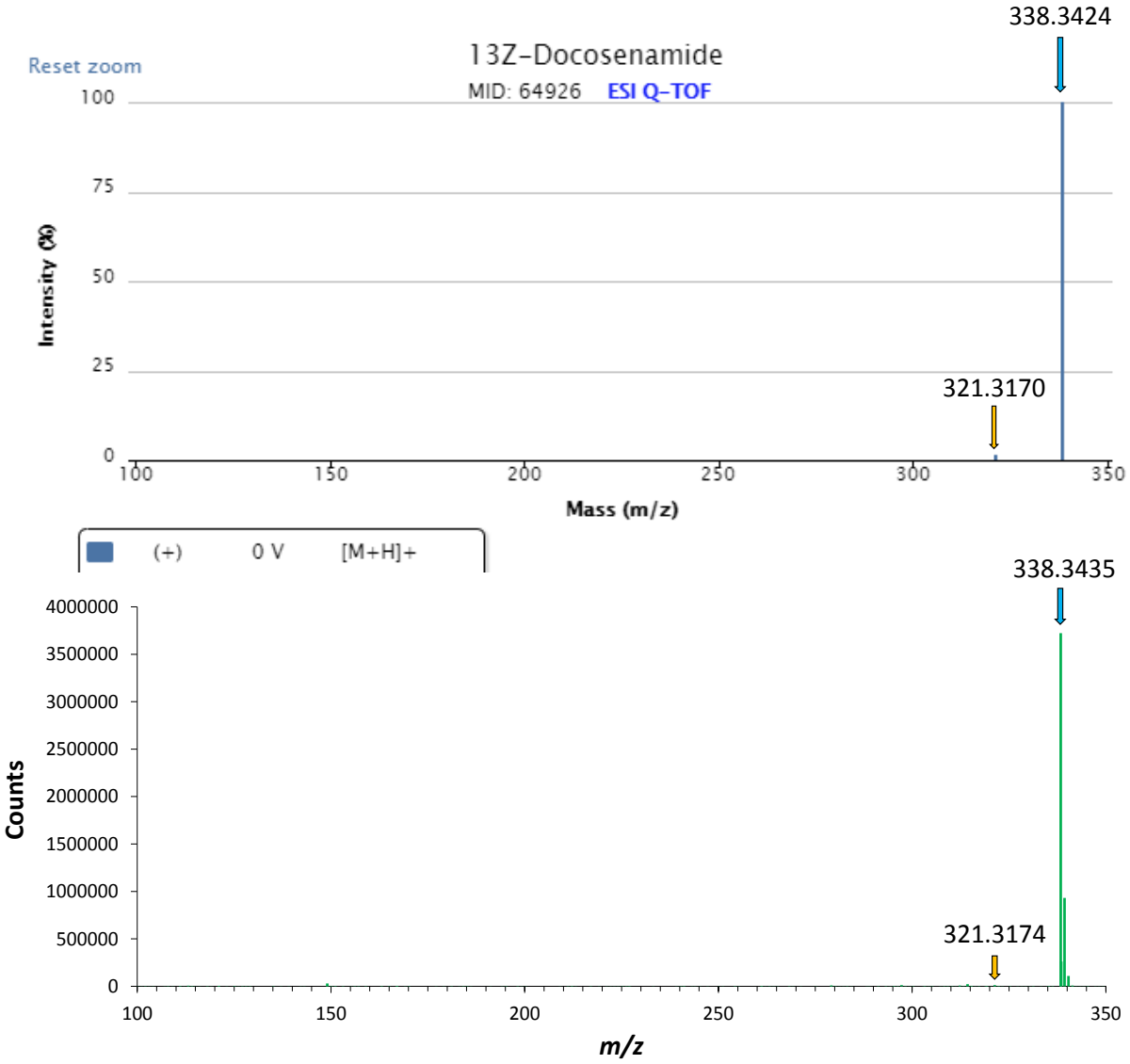


Figure D.4: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds docosenamide and m/z 338.3435 respectively*

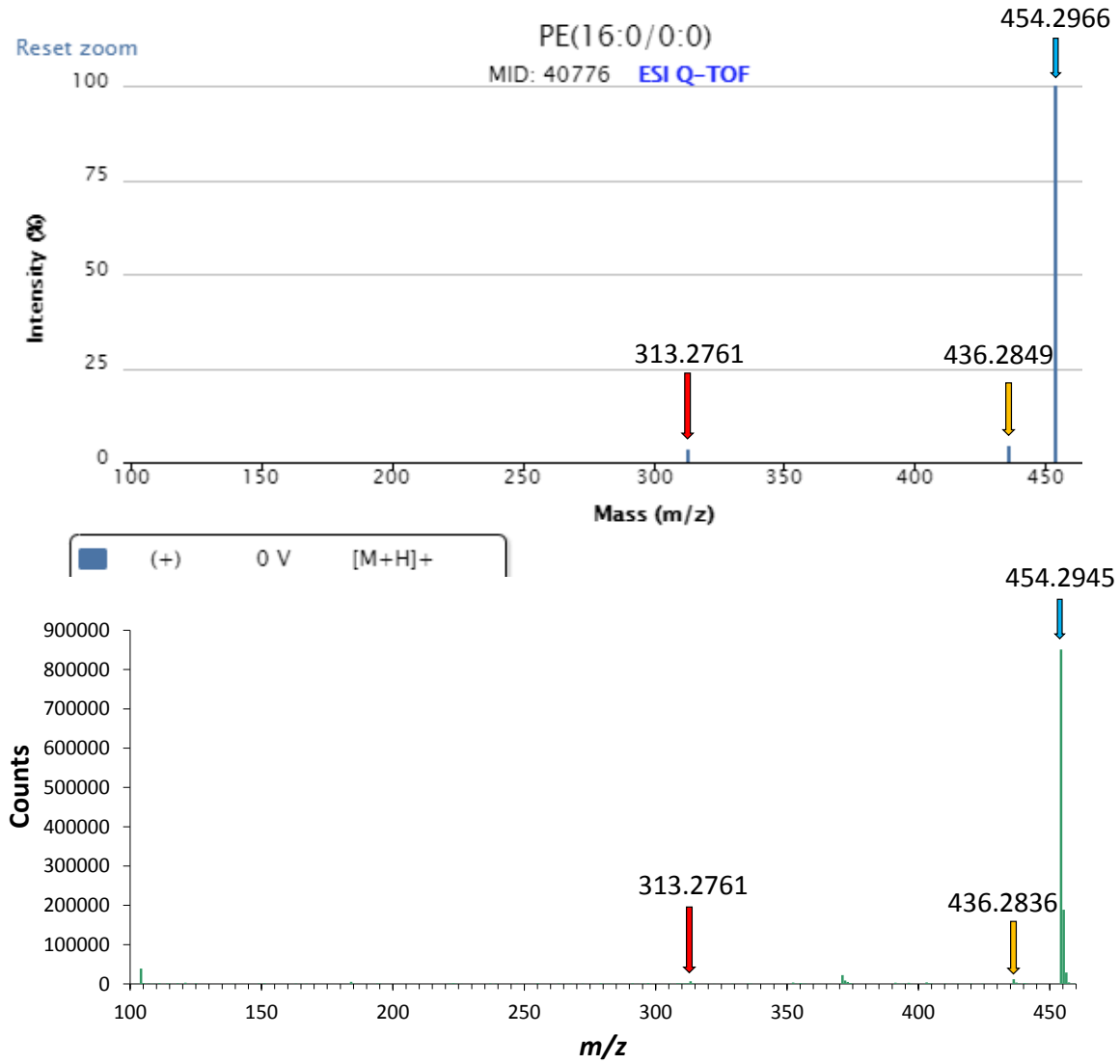


Figure D.5: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PE(16:0/0:0) and m/z 454.2945 respectively*

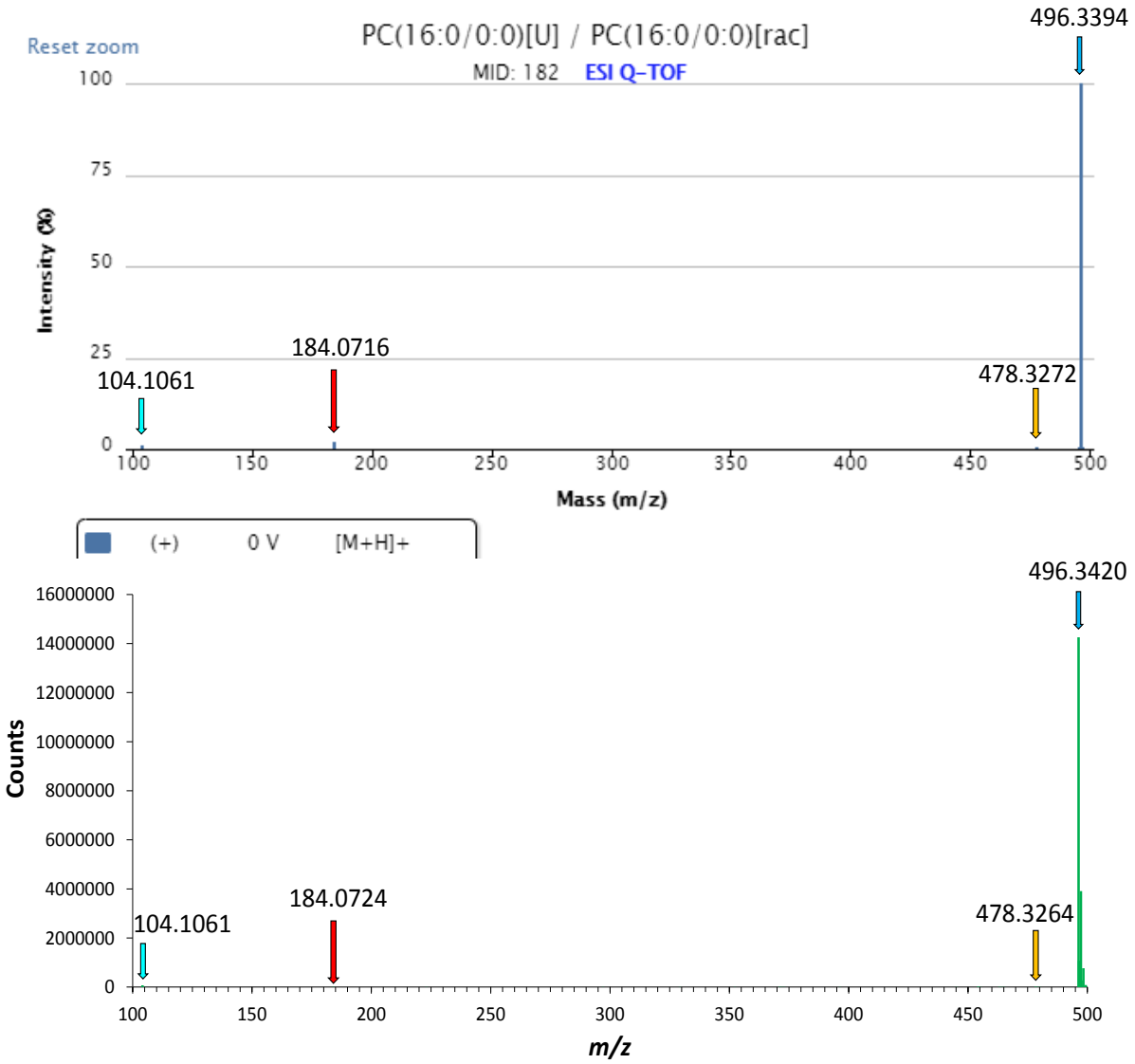


Figure D.6: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PC(16:0/0:0) and m/z 496.3420 respectively*

Table D.5: *Table showing concentration of choline in yolk organic extracts, the yolks themselves, and the mean choline concentration in egg yolks from eggs of different ages, stored at 5 °C*

Sample	Choline concentration in extract (µg/mL)	Choline concentration in yolk (µg/g)	Mean choline concentration in yolk (µg/g)
Fresh 1	0.3045	6.0898	8.1827
Fresh 2	0.5187	10.3741	
Fresh 3	0.4209	8.4178	
Fresh 4	0.4691	9.3820	
Fresh 5	0.3879	7.7572	
Fresh 6	0.3538	7.0754	
1 Week 1	0.2758	5.5160	5.6184
1 Week 2	0.2664	5.3288	
1 Week 3	0.2451	4.9014	
1 Week 4	0.2284	4.5680	
1 Week 5	0.2632	5.2644	
1 Week 6	0.4066	8.1320	
2 Weeks 1	0.4664	9.3288	7.6251
2 Weeks 2	0.3833	7.6667	
2 Weeks 3	0.3574	7.1474	
2 Weeks 4	0.3078	6.1558	
2 Weeks 5	0.4173	8.3452	
2 Weeks 6	0.3553	7.1066	
3 Weeks 1	0.4036	8.0720	8.1143
3 Weeks 2	0.2715	5.4292	
3 Weeks 3	0.4066	8.1318	
3 Weeks 4	0.6108	12.2170	
3 Weeks 5	0.3134	6.2678	
3 Weeks 6	0.4284	8.5678	
4 Weeks 1	0.3212	6.4244	7.0958
4 Weeks 2	0.3123	6.2465	
4 Weeks 3	0.3266	6.5316	
4 Weeks 4	0.4290	8.5801	
4 Weeks 5	0.3851	7.7014	
4 Weeks 6	0.3545	7.0908	
5 Weeks 1	0.4232	8.4640	8.9193
5 Weeks 2	0.4131	8.2616	
5 Weeks 3	0.5336	10.6730	
5 Weeks 4	0.5533	11.0659	
5 Weeks 5	0.4093	8.1851	
5 Weeks 6	0.3433	6.8664	

Table D.5 continued

Sample	Choline concentration in extract (µg/mL)	Choline concentration in yolk (µg/g)	Mean choline concentration in yolk (µg/g)
6 Weeks 1	0.4538	9.0754	10.1038
6 Weeks 2	0.5359	10.7179	
6 Weeks 3	0.4520	9.0404	
6 Weeks 4	0.4377	8.7543	
6 Weeks 5	0.7006	14.0112	
6 Weeks 6	0.4512	9.0236	
7 Weeks 1	0.3567	7.1333	7.0734
7 Weeks 2	0.2804	5.6081	
7 Weeks 3	0.2676	5.3522	
7 Weeks 4	0.5120	10.2406	
7 Weeks 5	0.2701	5.4022	
7 Weeks 6	0.4352	8.7040	
8 Weeks 1	0.5548	11.0958	10.3745
8 Weeks 2	0.6136	12.2719	
8 Weeks 3	0.2884	5.7677	
8 Weeks 4	0.4893	9.7855	
8 Weeks 5	0.6090	12.1802	
8 Weeks 6	0.5573	11.1456	
9 Weeks 1	0.2973	5.9457	9.4120
9 Weeks 2	0.4637	9.2742	
9 Weeks 3	0.4487	8.9738	
9 Weeks 4	0.3240	6.4802	
9 Weeks 5	0.6399	12.7973	
9 Weeks 6	0.6500	13.0009	
10 Weeks 1	0.7169	14.3384	10.0272
10 Weeks 2	0.5961	11.9220	
10 Weeks 3	0.2916	5.8320	
10 Weeks 4	0.3669	7.3374	
10 Weeks 5	0.4821	9.6420	
10 Weeks 6	0.5546	11.0911	
11 Weeks 1	0.3805	7.6095	7.4946
11 Weeks 2	0.3126	6.2522	
11 Weeks 3	0.4074	8.1479	
11 Weeks 4	0.5010	10.0193	
11 Weeks 5	0.2893	5.7857	
11 Weeks 6	0.3576	7.1527	

Table D.5 continued

Sample	Choline concentration in extract (µg/mL)	Choline concentration in yolk (µg/g)	Mean choline concentration in yolk (µg/g)
12 Weeks 1	0.4679	9.3584	10.2271
12 Weeks 2	0.4790	9.5793	
12 Weeks 3	0.5924	11.8484	
12 Weeks 4	0.3526	7.0520	
12 Weeks 5	0.6837	13.6740	
12 Weeks 6	0.4925	9.8501	

Table D.6: *Table showing the significantly different features in yolk organic extracts between eggs stored at 5 °C and 23 °C for five weeks. CV% calculated from peak areas of QC samples, p-value calculated from two-sample t-tests. The final two columns show whether the compound is of higher abundance (↑) or lower abundance (↓) in the yolks of eggs stored at that temperature, compared to the other temperature.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value	25 °C	5 °C
104.1072	0.790	1.85	0.004	↑	↓
165.0544	0.856	2.83	0.005	↑	↓
257.2471	14.893	5.47	<0.001	↑	↓
279.2275	19.740	5.69	0.003	↑	↓
297.2747	13.165	14.90	0.002	↑	↓
283.2980	26.382	14.29	0.003	↑	↓
303.2524	8.173	3.02	0.004	↑	↓
317.2408	24.813	23.93	0.008	↓	↑
331.2856	14.878	2.67	<0.001	↑	↓
348.2877	21.151	9.74	0.001	↑	↓
353.2543	21.143	4.05	0.003	↑	↓
359.3172	23.673	2.03	0.005	↑	↓
367.2789	19.098	6.18	0.009	↑	↓
377.2643	12.419	8.70	0.010	↑	↓
387.3465	26.842	9.16	0.002	↑	↓
404.2076	1.263	3.19	0.002	↓	↑
405.3092	29.677	7.32	0.008	↑	↓
427.3098	13.650	13.12	<0.001	↑	↓
438.3448	32.571	20.21	0.004	↓	↑
464.2827	1.637	14.19	0.001	↑	↓
485.3340	17.632	3.47	0.003	↑	↓
504.3074	17.682	2.57	0.003	↑	↓
517.3000	1.630	25.97	0.002	↑	↓
526.2922	8.373	21.70	0.003	↑	↓
549.5044	70.216	3.97	0.008	↓	↑
578.5236	69.717	2.44	0.009	↓	↑
626.5725	48.992	19.11	0.007	↓	↑
642.4310	28.270	12.75	<0.001	↓	↑
658.5427	45.229	0.59	0.004	↑	↓
670.4635	31.281	4.68	0.001	↓	↑
686.5865	52.916	11.43	0.003	↓	↑
721.5882	45.805	4.97	0.008	↓	↑
774.5572	29.219	10.37	0.005	↑	↓
841.6701	67.063	6.06	0.004	↓	↑
848.7762	69.808	1.84	0.009	↓	↑
851.7169	67.677	27.31	0.009	↓	↑
853.7331	69.592	6.89	0.005	↓	↑
859.6359	30.145	15.33	0.009	↓	↑
877.8112	73.508	2.08	0.005	↓	↑

Table D.6 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value	25 °C	5 °C
878.7347	68.499	3.26	0.003	↓	↑
880.8255	73.492	1.83	0.003	↓	↑
897.7243	70.332	3.38	0.007	↓	↑
924.2692	75.864	29.87	0.010	↓	↑
928.8320	68.014	1.28	0.005	↓	↑
932.8582	73.176	5.51	0.006	↓	↑
935.8432	72.662	21.57	0.009	↓	↑
950.8387	66.876	5.63	0.007	↓	↑
972.6113	10.199	2.02	0.009	↑	↓
985.6246	17.659	2.62	0.002	↑	↓
992.8881	73.673	2.35	0.005	↓	↑

Table D.7: *Table showing the significantly different features in organic extracts between egg yolks stored at 5 °C for different lengths of time. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
310.3103	21.919	2.03	<0.001
331.2840	13.489	4.95	<0.001
338.3416	26.614	1.57	<0.001
359.3152	22.198	2.01	<0.001
387.1770	1.200	6.31	<0.001
525.3743	16.040	4.10	0.009
603.5350	51.347	3.35	0.004
605.5500	54.103	3.93	0.002
608.5247	43.676	1.96	<0.001
614.5717	53.123	1.76	<0.001
615.4956	47.230	3.18	<0.001
617.5114	50.277	2.04	0.007
622.6141	53.486	2.13	<0.001
632.6338	56.991	3.07	<0.001
636.6296	55.282	1.71	<0.001
640.5882	54.069	1.65	0.003
641.5118	48.328	3.94	0.002
666.6215	69.330	11.20	0.001
720.5889	46.301	1.39	0.002
730.5743	48.370	3.00	0.001
744.5869	46.955	5.12	0.002
746.6053	47.721	5.37	<0.001
758.5701	41.928	5.28	0.003
761.5895	44.950	5.78	0.007
764.5244	40.574	2.82	0.005
769.5596	45.786	1.31	0.001
773.5889	43.701	2.31	0.010
774.6311	52.517	14.62	0.001
775.6047	46.968	8.58	0.003
778.5742	49.770	1.52	0.001
780.5909	50.435	1.69	<0.001
782.5714	40.708	2.39	0.004
782.6063	51.629	1.61	<0.001
784.5863	42.450	0.99	0.001
787.6685	52.238	1.96	0.003
788.6175	49.025	2.03	0.001
792.5556	45.193	2.23	0.001
794.5775	46.534	9.45	0.002
806.5710	40.408	1.56	0.003
808.5866	43.241	1.02	0.008

Table D.7 continued

Feature *m/z*	RT median (minutes)	CV%	*p*-value
812.6171	47.166	26.98	<0.001
813.6848	52.844	2.13	0.007
824.6467	52.372	1.86	<0.001
832.5855	41.836	1.50	0.005
834.6020	45.134	1.40	<0.001
836.6168	46.821	1.49	<0.001
838.6312	49.358	2.28	<0.001
874.7875	73.017	2.58	<0.001
901.8093	73.856	4.46	0.010
902.8189	77.328	4.58	0.001

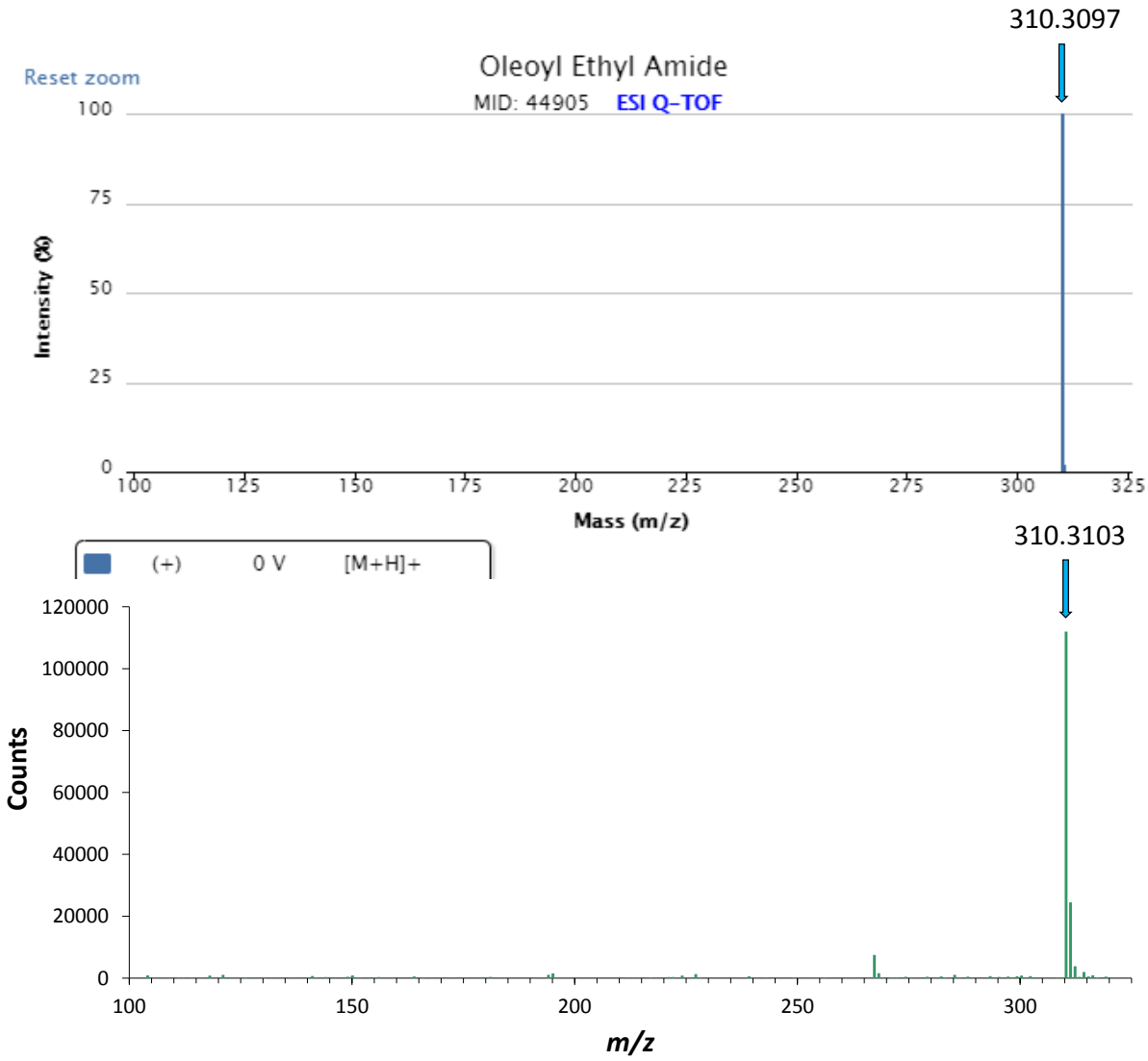


Figure D.7: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds oleoyl ethyl amide and m/z 310.3103 respectively*

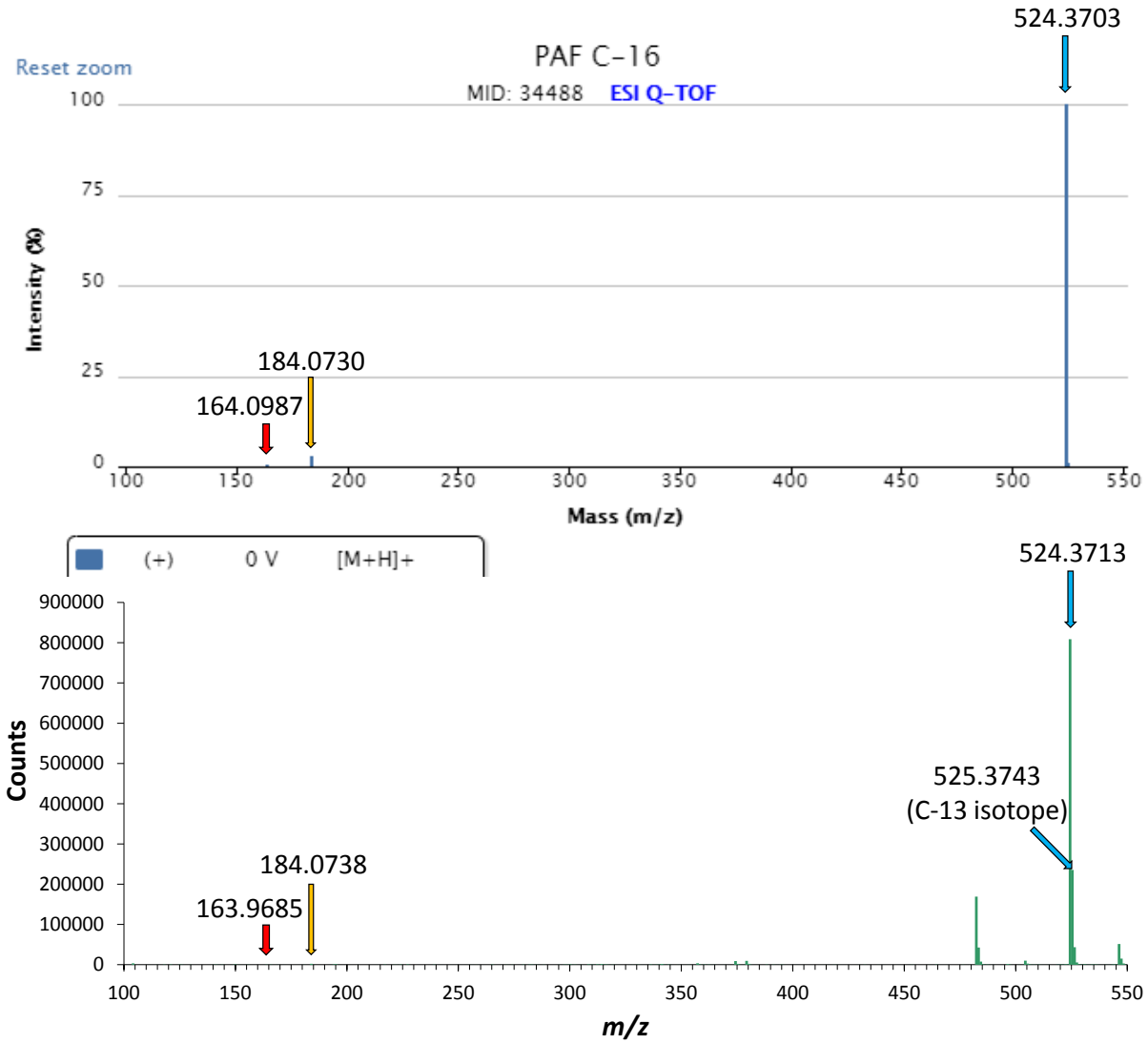


Figure D.8: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PAF C-16 and m/z 525.3743 respectively*

Appendix E

Table E.1: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages with different numbers of birds (20, 40, 60, 80). CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
341.3046	52.891	1.62	0.002
387.1733	1.184	3.09	0.005
522.3555	9.227	2.48	<0.001
549.5066	45.919	1.50	0.010
577.5196	49.179	2.60	0.001
579.5346	52.044	3.44	0.003
595.5295	49.138	1.75	0.008
604.5381	50.252	7.68	0.004
604.6027	52.359	2.54	<0.001
618.6183	54.086	2.57	<0.001
621.5452	50.203	1.79	0.005
632.6343	55.746	2.24	<0.001
636.5571	47.279	6.90	0.003
640.5885	52.891	1.26	0.001
641.5118	47.231	3.17	0.005
644.5950	52.360	1.34	0.003
650.4387	15.263	9.02	0.004
654.3316	13.547	1.15	<0.001
658.6106	54.086	0.73	0.001
666.6032	53.723	1.89	0.001
672.6264	55.762	1.02	0.005
700.5270	42.583	2.62	0.004
728.5582	47.065	4.80	0.001
732.5435	35.712	3.22	0.006
742.5554	42.565	23.90	0.005
744.5861	45.789	3.02	0.009
758.5699	39.861	1.06	0.003
759.5719	42.533	16.99	0.008

Table E.1 continued

Feature *m/z*	**RT median (minutes)**	**CV%**	***p*-value**
760.5860	51.293	8.46	0.005
806.5545	34.697	20.24	<0.001
808.5695	36.989	2.69	<0.001
816.6999	54.908	7.06	0.007
836.7707	70.846	28.35	0.003
842.7237	65.176	1.51	0.002
847.7599	67.714	1.97	0.009
868.7395	65.723	2.39	0.001
870.7607	67.382	0.83	<0.001
872.7681	69.442	4.64	<0.001
877.8078	73.242	10.10	0.001
882.7618	72.910	2.87	0.009
898.7868	69.516	3.40	0.007
916.8339	76.265	5.43	<0.001
922.7806	69.319	3.29	0.001
926.8182	72.300	11.56	<0.001
928.8330	75.811	9.09	0.003
930.8489	78.505	6.92	<0.001
941.8296	71.265	7.44	0.003

Table E.2: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages of different population size for nine weeks, as well as three weeks. The p-value was calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	9 weeks *p*-value	3 weeks *p*-value
341.3046	52.891	0.002	0.241
387.1733	1.184	0.005	<0.001
522.3555	9.227	<0.001	0.449
549.5066	45.919	0.010	0.378
577.5196	49.179	0.001	0.289
579.5346	52.044	0.003	0.246
595.5295	49.138	0.008	0.706
604.5381	50.252	0.004	0.048
604.6027	52.359	<0.001	0.150
618.6183	54.086	<0.001	0.017
621.5452	50.203	0.005	0.622
632.6343	55.746	<0.001	0.387
636.5571	47.279	0.003	0.005
640.5885	52.891	0.001	0.546
641.5118	47.231	0.005	0.039
644.5950	52.360	0.003	0.337
650.4387	15.263	0.004	0.269
654.3316	13.547	<0.001	0.037
658.6106	54.086	0.001	0.102
666.6032	53.723	0.001	0.686
672.6264	55.762	0.005	0.377
700.5270	42.583	0.004	0.873
728.5582	47.065	0.001	0.740
808.5695	36.989	<0.001	<0.001
836.7707	70.846	0.003	0.092
842.7237	65.176	0.002	0.200
847.7599	67.714	0.009	0.481
868.7395	65.723	0.001	0.254
870.7607	67.382	<0.001	0.092
916.8339	76.265	<0.001	0.556
922.7806	69.319	0.001	0.713

Table E.2 continued

Feature *m/z*	**RT median (minutes)**	**9 weeks *p*-value**	**3 weeks *p*-value**
926.8182	72.300	<0.001	0.650
928.8330	75.811	0.003	0.906
930.8489	78.505	<0.001	0.699

Table E.3: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages with different stocking densities. CV% calculated from peak areas of QC samples, p-value calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
162.1106	0.748	11.08	0.001
291.2525	4.251	3.26	0.007
310.2377	3.923	8.23	0.005
369.1660	1.234	4.53	0.006
387.1880	1.881	6.42	0.002
448.3574	31.178	24.16	0.002
489.2270	5.451	2.69	0.001
498.3457	8.920	5.80	0.008
548.5037	47.996	6.93	<0.001
565.4040	30.764	3.88	0.001
599.5033	35.893	5.74	0.004
601.5184	38.135	8.82	0.009
624.6273	55.856	3.82	0.003
627.5340	41.489	7.21	0.009
630.6188	55.232	2.88	0.015
637.3050	15.078	2.88	0.006
638.6438	57.508	5.00	<0.001
640.8873	52.327	18.48	0.001
641.5122	49.672	5.69	0.003
650.6425	56.459	4.88	0.010
652.6597	59.050	6.10	0.001
658.5402	46.134	4.10	0.020
659.5010	52.401	29.15	0.010
662.6441	56.875	3.92	0.009
667.6324	53.406	11.16	0.005
708.5446	40.285	11.51	0.008
720.5897	47.731	5.06	0.005
734.4781	35.303	8.62	0.009
745.5474	43.894	14.89	0.014
754.5322	36.340	4.13	0.004
758.5694	45.106	3.67	0.009
758.5817	55.811	9.97	0.009
760.5863	46.143	4.30	0.004
762.5973	51.252	24.62	0.001
764.5186	41.211	9.66	0.009
771.5713	46.135	10.20	0.025
775.6074	49.803	6.06	0.006
786.5074	41.919	7.57	0.041
786.6013	54.209	25.31	0.001

Table E.3 continued

Feature *m/z*	**RT median (minutes)**	**CV%**	***p*-value**
787.6080	52.891	5.59	0.002
788.6072	48.385	5.45	0.010
796.3372	1.250	6.62	0.008
796.7383	70.633	8.43	0.009
802.6805	56.295	4.44	0.007
807.5718	41.122	24.96	0.009
815.5437	43.762	8.03	0.002
842.5891	35.678	3.12	0.006
842.6917	56.227	17.23	0.003
860.5361	35.893	6.33	<0.001
878.5743	38.119	6.04	0.010
892.6918	68.113	4.42	0.004
896.7716	69.656	3.45	0.002
906.6050	45.189	5.81	<0.001
908.7751	69.093	14.63	<0.001
923.7255	66.468	11.66	<0.001
946.7851	69.721	4.10	0.003
952.8326	75.584	5.82	0.001
980.8631	80.775	13.97	0.006
989.5638	15.086	3.96	0.010

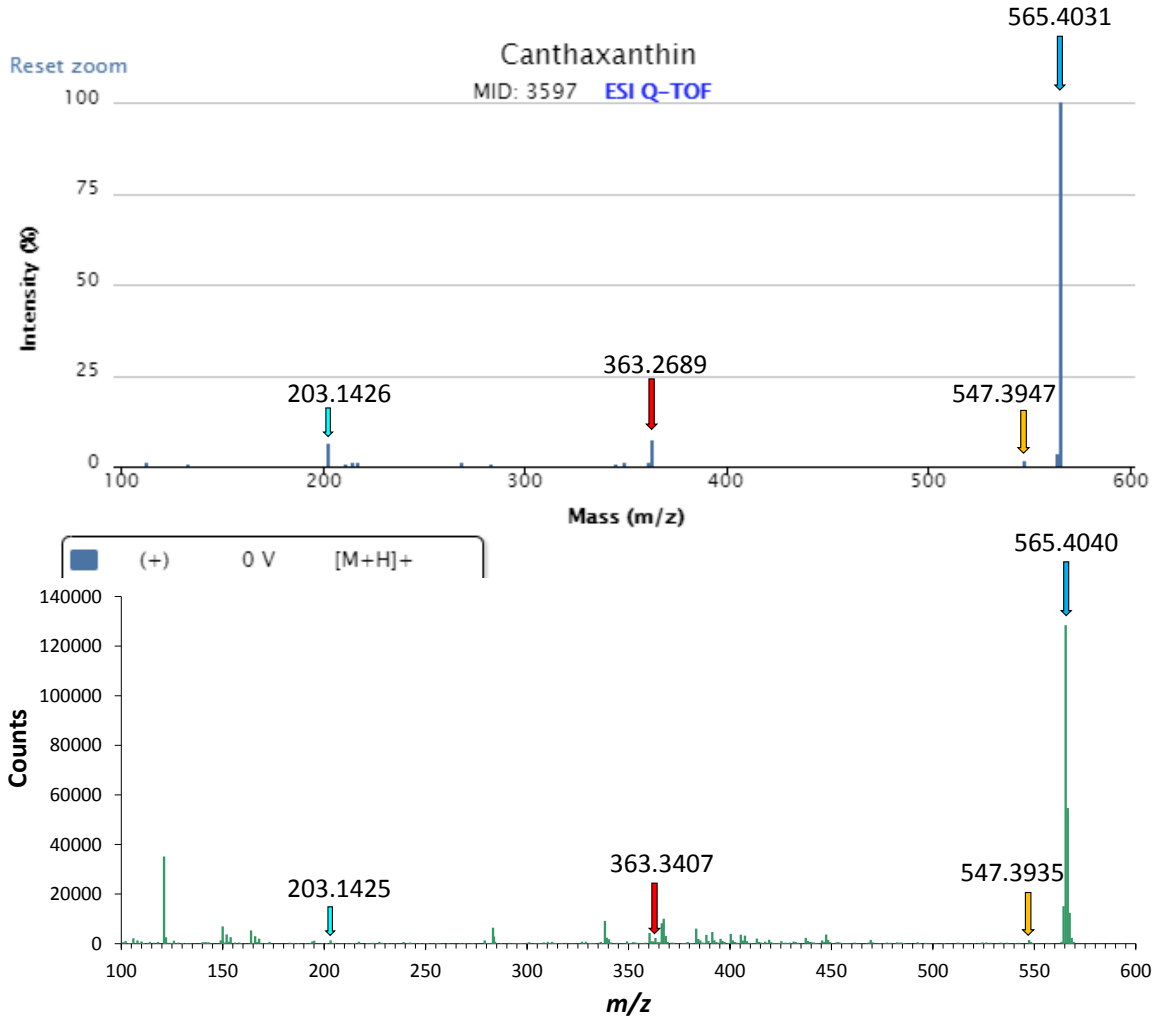


Figure E.1: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds canthaxanthin and m/z 565.4040 respectively*

Table E.4: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens kept in cages of different stocking densities for two weeks, as well as four days. The p-value was calculated from either ANOVA or Welch test depending on variance.*

Feature *m/z*	RT median (minutes)	2 weeks *p*-value	4 days *p*-value
291.2525	4.251	0.007	<0.001
310.2377	3.923	0.005	0.025
369.1660	1.234	0.006	<0.001
387.1880	1.881	0.002	<0.001
489.2270	5.451	0.001	0.001
548.5037	47.996	<0.001	<0.001
565.4040	30.764	0.001	0.172
599.5033	35.893	0.004	0.001
601.5184	38.135	0.009	0.004
624.6273	55.856	0.003	0.019
630.6188	55.232	0.015	0.079
637.3050	15.078	0.006	<0.001
638.6438	57.508	<0.001	0.007
641.5122	49.672	0.003	0.651
650.6425	56.459	0.010	0.028
652.6597	59.050	0.001	0.010
658.5402	46.134	0.020	0.308
667.6324	53.406	0.005	0.032
708.5446	40.285	0.008	0.270
720.5897	47.731	0.005	0.384
734.4781	35.303	0.009	0.163
754.5322	36.340	0.004	0.375
758.5694	45.106	0.009	0.001
760.5863	46.143	0.004	0.277
771.5713	46.135	0.025	0.423
786.5074	41.919	0.041	0.022
788.6072	48.385	0.010	0.052
796.3372	1.250	0.008	<0.001
796.7383	70.633	0.009	0.045
842.5891	35.678	0.006	0.118
842.6917	56.227	0.003	0.764

Table E.4 continued

Feature *m/z*	RT median (minutes)	2 weeks *p*-value	4 days *p*-value
860.5361	35.893	<0.001	0.005
878.5743	38.119	0.010	0.005
896.7716	69.656	0.002	0.252
906.6050	45.189	<0.001	0.020
923.7255	66.468	<0.001	0.119
946.7851	69.721	0.003	0.490
952.8326	75.584	0.001	0.139
980.8631	80.775	0.006	0.129
989.5638	15.086	0.010	0.007

Appendix F

Table F.1: *Table showing the significantly different features in yolk organic extracts between eggs that were laid by hens from cage and barn housing systems. CV% calculated from peak areas of QC samples, p-value calculated from t-test.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
520.3398	8.158	3.21	0.004
601.5182	48.420	5.03	0.001
608.5243	43.851	5.74	0.047
634.5395	45.408	7.66	0.001
636.5566	48.386	4.54	<0.000
700.5266	43.843	3.37	<0.001
728.5580	48.135	10.36	0.001
734.5699	43.389	3.11	0.004
744.5543	46.159	1.82	0.013
752.5215	34.904	3.05	0.001
754.5371	37.355	6.71	0.011
756.5546	37.946	18.63	0.043
762.6007	48.003	1.25	0.035
772.5851	43.834	3.40	0.007
780.5529	39.391	2.44	0.023
783.5737	39.557	2.26	0.011
787.6679	52.084	1.38	0.030
812.6620	49.804	3.31	0.002
870.7605	68.124	8.92	0.003
872.7705	70.111	1.63	0.015
886.7880	71.777	5.73	0.008
888.8098	74.217	4.39	0.033
896.7722	69.705	1.74	0.049
896.7766	68.779	1.97	0.000
898.7852	70.960	2.41	0.000
898.7907	72.334	3.94	0.008
901.8043	73.763	5.92	0.010
926.8194	76.619	12.83	0.006
928.8331	78.336	6.39	0.036

Table F.2: *Table showing the 12 compounds that produced potential metabolite matches through METLIN. PC=phosphatidylcholine, PE=phosphatidylethanolamine, PA=phosphatidate*

Feature m/z	RT median (minutes)	Potential Formulae	Species	Score	Potential ID
520.3398	8.158	$C_{26}H_{50}NO_7P$	$(M+H)^+$	99.22	PE, PC, LysoPC
608.5243	43.851	$C_{37}H_{66}O_5$	$(M+NH_4)^+$	99.39	Diglycerides
634.5395	45.408	$C_{39}H_{68}O_5$	$(M+NH_4)^+$	99.54	Diglycerides
636.5566	48.386	$C_{39}H_{70}O_5$	$(M+NH_4)^+$	97.79	Diglycerides
700.5266	43.843	$C_{39}H_{74}NO_7P$	$(M+H)^+$	97.37	PE
		$C_{39}H_{71}O_7P$	$(M+NH_4)^+$	97.37	PA
734.5699	43.389	$C_{40}H_{80}NO_8P$	$(M+H)^+$	99.36	PE, PC
		$C_{40}H_{77}O_8P$	$(M+NH_4)^+$	99.36	PA
744.5543	46.159	$C_{41}H_{78}NO_8P$	$(M+H)^+$	98.16	PE, PC
772.5851	43.834	$C_{43}H_{82}NO_8P$	$(M+H)^+$	95.00	PE, PC
		$C_{43}H_{79}O_8P$	$(M+NH_4)^+$	95.00	PA
872.7705	70.111	$C_{55}H_{98}O_6$	$(M+NH_4)^+$	99.53	Triglycerides
886.7880	71.777	$C_{56}H_{100}O_6$	$(M+NH_4)^+$	96.56	Triglycerides
898.7852	70.960	$C_{57}H_{100}O_6$	$(M+NH_4)^+$	98.76	Triglycerides
928.8331	78.336	$C_{59}H_{106}O_6$	$(M+NH_4)^+$	98.03	Triglycerides

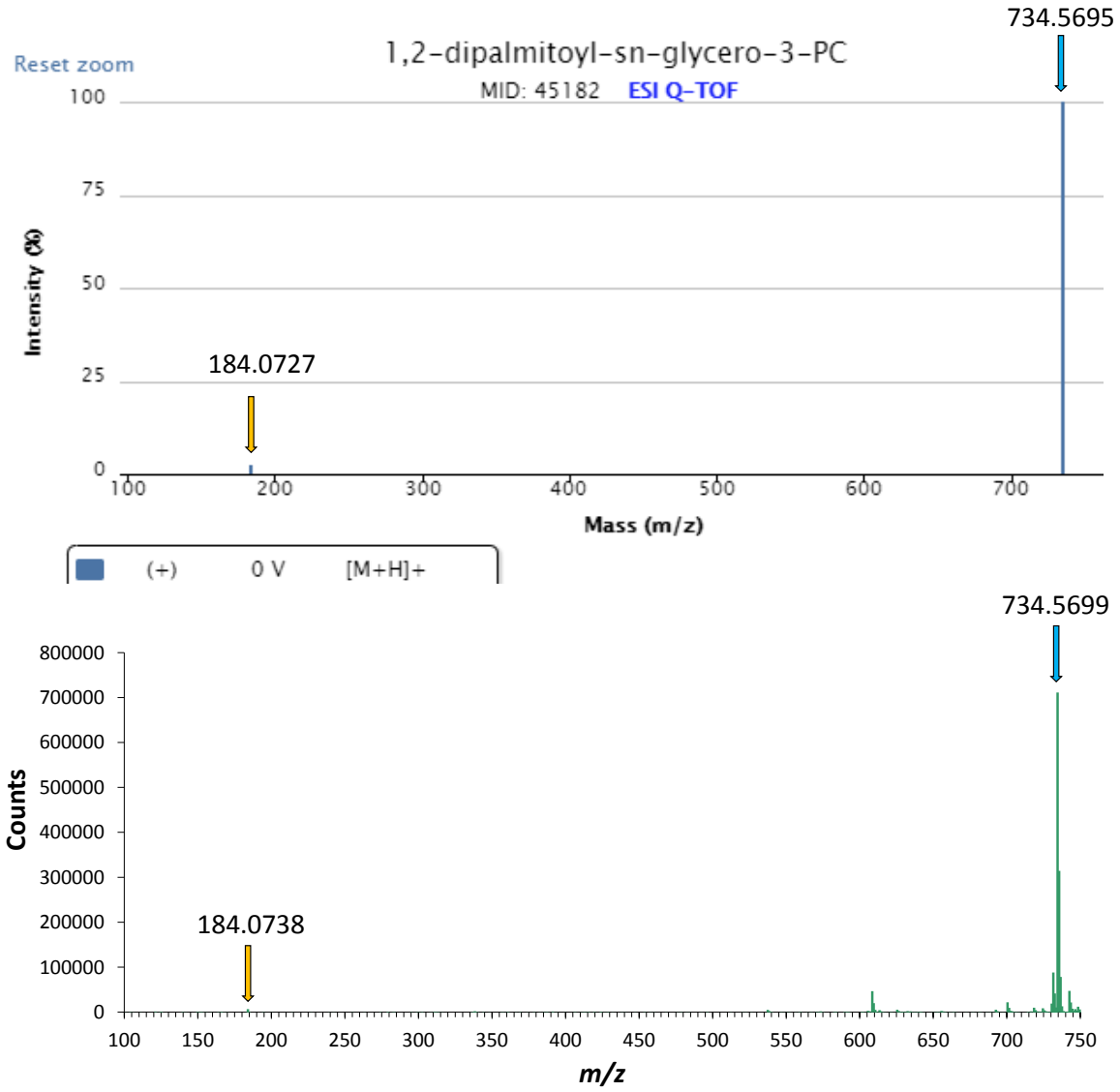


Figure F.1: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PC(16:0/16:0) and m/z 734.5699 respectively*

Table F.3: *Table showing the significantly different features in albumen organic extracts between eggs that were laid by birds from cage and barn housing systems. CV% calculated from peak areas of QC samples, p-value calculated from t-test.*

Feature *m/z*	RT median (minutes)	CV%	*p*-value
310.2369	2.102	6.30	0.031
316.3209	17.813	3.24	0.034
425.1295	0.942	3.11	0.004
506.2529	2.667	1.80	0.019
586.5397	25.068	3.34	0.041
811.3078	0.958	2.30	0.002

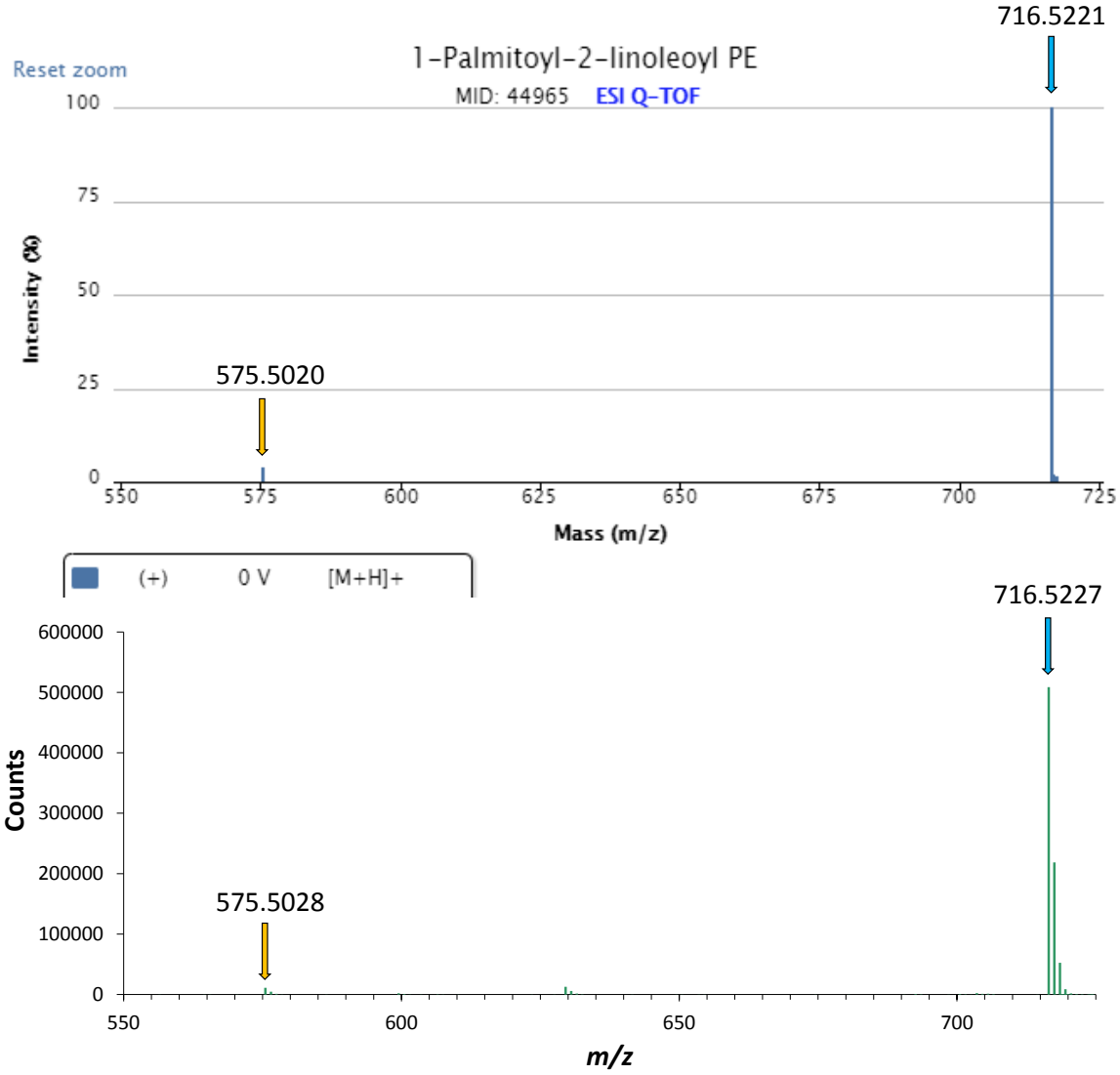


Figure F.2: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds PE(16:0/18:2) and m/z 716.5227 respectively*

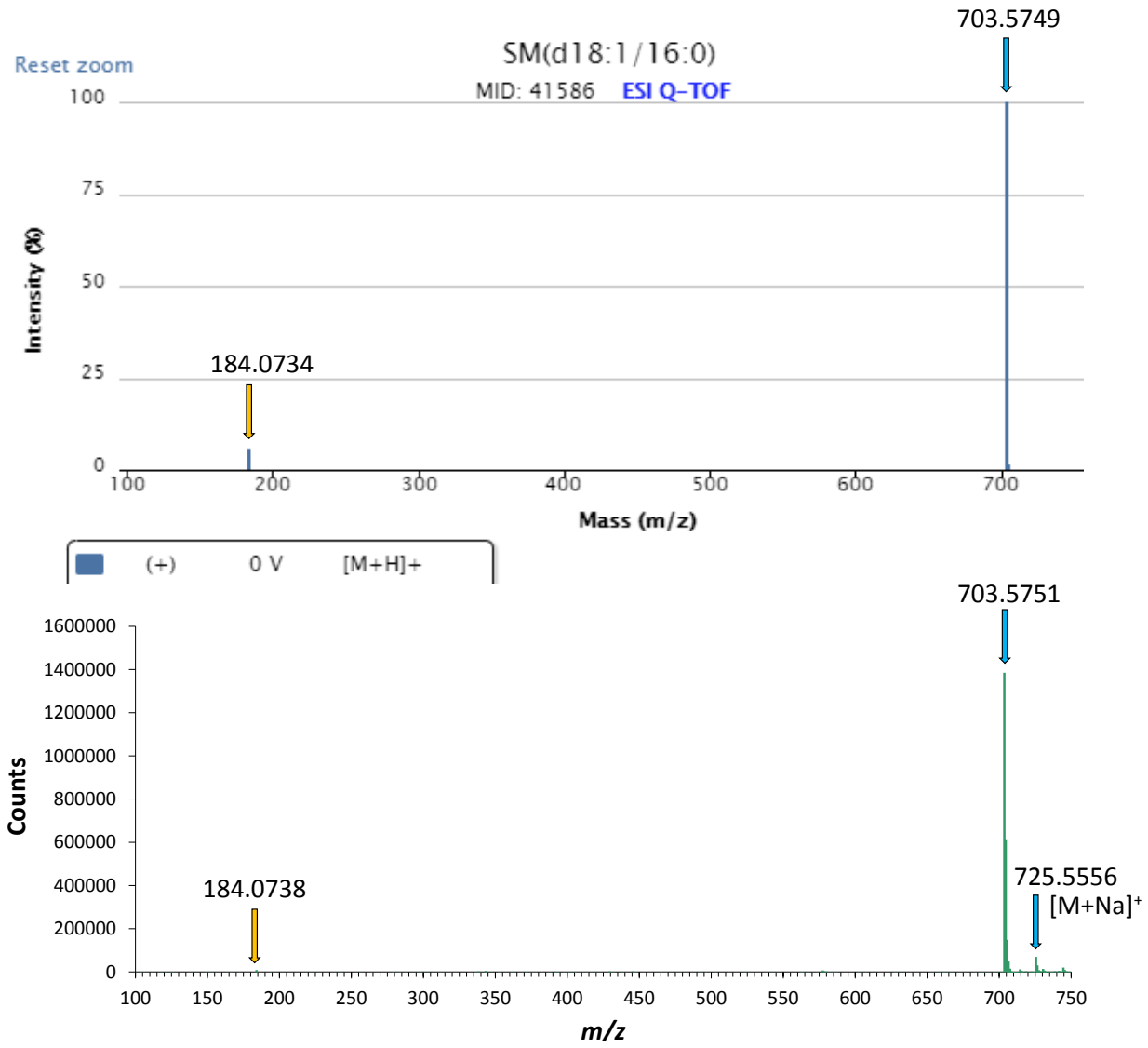


Figure F.3: *Comparison between the mass spectrum provided by METLIN (top), and the mass spectrum resulting from analysis (bottom) for the compounds SM(18:1/16:0) and m/z 725.5556 respectively*