

2010-10-01

# Metabolic Profiling and Fingerprinting for the Detection and Discrimination of Mechanical Damage in Mushrooms (Agaricus bisporus) during Storage

Aoife O'Gorman

*Technological University Dublin*

# Metabolic profiling and fingerprinting for the detection and discrimination of mechanical damage in mushrooms (*Agaricus bisporus*) during storage

Aoife O'Gorman, BSc

A thesis submitted to Dublin Institute of Technology

in fulfilment of the requirements for the degree of

**DOCTOR OF PHILOSOPHY**

School of Food Science and Environmental Health

College of Sciences and Health

Dublin Institute of Technology

Supervisors:
Dr. Jesus M. Frias Celayeta
Dr. Catherine Barry-Ryan

October 2010

## Abstract

Horticultural products such as mushrooms are exposed to external agents during their post-harvest life, which are going to affect product quality. Loss of whiteness during storage is particularly important in the mushroom industry. Rough handling and distribution of crops, fruiting body senescence, bacterial and viral infection are among the causes of mushroom discolouration. The aim of this work was to study the use of metabolic fingerprinting and metabolic profiling tools for the detection and discrimination of mechanical damage on mushrooms. This research involved:

1. Investigating whether the chemical changes induced by mechanical damage and ageing of mushrooms could be (a) detected in the mid-infrared absorption region using FTIR spectroscopy as a fingerprinting tool and (b) identified using chemometric data analysis.
2. Investigating metabolites in mushroom tissues using GC/MS as a metabolic profiling technique. The method was used to profile mushroom samples to identify metabolic markers for damage and to gain understanding of the many metabolic processes that occur.
3. Studying low levels of damage in mushrooms using NMR spectroscopy as a fingerprinting technique coupled with chemometrics to identify markers and determine metabolite structure.

The results from this study show the usefulness of FTIR spectroscopy coupled with chemometric data analysis for evaluating damage in mushrooms with specific wavenumbers identified. Metabolic profiling using GC/MS has led to a library of metabolites being built. Specific metabolites have been identified as markers for damage.

## Declaration

I certify that this thesis which I now submit for examination for the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for an award in any other institute or university.

The work reported in this thesis conforms to the principles and requirements of the institutes guidelines for ethics in research.

The institute has permission to keep, to lend or to copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature:_____ Date:_____

Aoife O'Gorman (candidate)

## Acknowledgements

I would like to sincerely thank my supervisors Dr. Jesus Frias and Dr. Catherine Barry-Ryan for their untiring enthusiasm, support, advice and guidance throughout this project, without which this would not have been possible, thank you both!!

I would like to thank Dr. Gerard Downey for facilitating collaborative work with Teagasc (Ashtown Food Research Centre). Thanks to Dr. Padraig McLaughlin for his NMR expertise.

Thanks to all the academic and technical staff of DIT for their assistance and practical advice throughout my research. I am also very grateful to Dr. Helen Grogan and Mr. Ted Cormican from Teagasc Research Station in Kinsealy for provision of mushrooms and technical advice.

The biggest thanks go to my fellow postgraduates (Laura Massini, Ziad, Patricia, Laura Thornton, Laura Alvarez, Alka, Lubna) who provided the laughs along the way and who were always there to answer questions, borrow/steal chemicals from, etc.

A special thanks to Dr. Aoife Smith who provided me with much entertainment over the years. For all the giggles, drama (always drama) and FAB Fridays… will never be forgotten!! Thanks also for all the help with those pesky NMRs.

A special mention must go to Dr. Edurne Gaston who started this journey with me, thanks for everything over the past four years, from debating what a billion means to camping in Achill (soooooo sorry about that one!!!), to fixing my MANY Endnote problems to Oktoberfest madness. Always remember no fungi no future my mushroom girl!!!!

To the bestest friends in the world, Kate and Sarah (Dr. Rowe), particularly thanks to Kate for having to listen to all the exciting science talk over the years….god love you!!! Now I know why you drag me to those “interesting” art exhibitions…..Revenge!! Best of luck Sarah in your new position in Boston, so proud of you and can’t wait to come visit!!!

I would like to say a heartfelt thank you to my family. Thanks to my parents for their endless support and faith in me over the years, especially to my mum for reading my thesis!!! Thanks also to my sisters and brother, Carol, Claire and Patrick.

A final mention goes to Eddie, thank you so much for all the love and support and for solving all my computer problems, which there were many!!!

## List of Symbols & Abbreviations

| | |
|---|---|
| AA | Amino acids |
| AMDIS | Automated mass spectral deconvolution and identification System |
| ANOVA | Analysis of variance |
| APCI | Atmospheric pressure chemical ionisation |
| ATP | Adenosine triphosphate |
| ATR | Attenuated total reflectance |
| BP | Base peak (ion) |
| $CaCl_2$ | Calcium chloride |
| Caret | Classification and regression training |
| CART | Classification and regression trees |
| $CDCl_3$ | Deuterated chloroform |
| CE/MS | Capillary electrophoresis mass spectrometry |
| $CHCl_3$ | Chloroform |
| COSY | Correlation spectroscopy |
| $CO_2$ | Carbon dioxide |
| CZE | Capillary zone electrophoresis |
| D | Damaged |
| DIMS | Direct injection mass spectrometry |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| EM | Electromagnetic |
| EOF | Electroosmotic flow |
| FAO | The Food and Agricultural Organisation of the United Nations |

| | |
|---|---|
| FAs | Fatty acids |
| EESI | Extractive electrospray ionisation |
| EI | Electron impact |
| ESI | Electrospray ionisation |
| EST | Expressed sequence tag |
| FDA | Factorial discriminant analysis |
| FTIR | Fourier transform infrared |
| GC/EI/TOF/MS | Gas chromatography/electrospray ionisation/time-of-flight/mass spectrometry |
| GC/MS | Gas chromatography mass spectrometry |
| GDH | Glutamate dehydrogenase |
| GM | Genetically modified |
| GOGAT | Glutamine oxoglutarate aminotransferase |
| GS | Glutamine synthetase |
| HK | Hexokinase |
| HPL | Hydroperoxide lyase |
| HPLC | High performance liquid chromatography |
| HPODs | Hydroperoxides |
| $H_2SO_4$ | Sulphuric acid |
| IMS | Ion mobility mass spectrometry |
| IR | Infrared |
| IS | Internal standard |
| KBr | Potassium bromide |
| LC/MS | Liquid chromatography mass spectrometry |
| LOX | Lipoxygenase |
| MAP | Modified atmosphere packaging |

| | |
|---|---|
| MDH | Mannitol dehydrogenase |
| MeOH | Methanol |
| MIR | Mid infrared |
| MPD | Mannitol-1-phosphate dehydrogenase |
| MPP | Mannitol-1-phosphate phosphatase |
| MS | Mass spectrum |
| MSTFA | N-methyl-N-trimethylsilylfluoracetamide |
| MW | Molecular weight |
| MVX | Mushroom virus X |
| NIH | National Institute of Health |
| NIR | Near infrared |
| NIST | National Institute of Standards and Technology |
| NMR | Nuclear magnetic resonance |
| NOESY | Nuclear overhauser enhancement spectroscopy |
| OOB | Out of bag classification error |
| PC | Principal components |
| PCA | Principal component analysis |
| PLS | Partial least squares regression |
| PLS-DA | Partial least squares-discriminant analysis |
| PLS-LDA | Partial least squares-linear discriminant analysis |
| ppm | Parts per million |
| PPO | Polyphenol oxidase |
| PPP | Pentose phosphate pathway |
| Rf | Radio frequency |
| RF | Random forest |

| | |
|---|---|
| RMSECV | Root mean square error of cross validation |
| RNA | Ribonucleic acid |
| ROS | Reactive oxygen species |
| rpm | rotations per minute |
| RT | Retention time |
| SD | Standard deviation |
| SIMS | Soft ionisation mass spectrometry |
| SNV | Standard normal variance |
| TCA cycle | Tricarboxylic acid cycle |
| TMS | Trimethylsilyl ester |
| TMS | Tetramethylsilane (internal standard added to $CDCl_3$ for NMR experiments) |
| TOCSY | Total correlation spectroscopy |
| UD | Undamaged |
| UDP-glucose | Uridine diphosphate glucose |
| UHPLC | Ultra-high pressure liquid chromatography |
| UV | Ultraviolet |
| VIP | Variable importance plot |
| v/v | Volume to volume |
| #L | Latent number of variables |
| 10-ODA | 10-oxodecenoic acid |

**TABLE OF CONTENTS**

**1. Introduction……..1**

1.1 General Introduction……1

*1.1.1 Mushroom History*……1

*1.1.2 Fungal Kingdom*.....2

*1.1.3 Mushroom Morphology*.....2

*1.1.4 Mushroom Physiology*……..3

*1.1.5 Nutritional Attributes*……...4

*1.1.5.1 Carbohydrate & Fibre*……4

*1.1.5.2 Protein & Amino Acids*……...5

*1.1.5.3 Lipids*……..5

*1.1.5.4 Vitamins & Minerals*……...5

*1.1.6 Mushroom Production*……..6

*1.1.7 Irish Mushroom Industry*……..8

1.2 Factors Effecting Mushroom Quality…….....9

*1.2.1 Shelf-life*……9

*1.2.2 Mushroom Quality*……..10

*1.2.2.1 Mushroom Discolouration*……11

*1.2.2.1.1 Phenolic Compounds*……11

*1.2.2.1.2 Senescence*……14

*1.2.2.1.3 Mechanical Damage*…….14

*1.2.2.1.4 Microbial & Viral Spoilage*……..16

*1.2.3 Mushroom Variability*…….17

1.3 Metabolomics…….18

*1.3.1 Mushroom Metabolism*……...18

*1.3.1.1 Carbohydrate Metabolism*……19

*1.3.1.1.1 Metabolism of Mannitol*……20

*1.3.1.1.2 Metabolism of Trehalose*……...21

*1.3.1.2 Nitrogen Metabolism*……22

*1.3.1.2.1 Urea Metabolism & The Ornithine Cycle (urea cycle)*……23

*1.3.2 Defining Metabolomics*……..…...24

*1.3.3 Development of Metabolomics*……....27

*1.3.4 Metabolomic Limitations*……29

*1.3.5 Metabolomic Technologies*…….30

*1.3.5.1 Gas Chromatography-Mass Spectrometry*…….……..…31

*1.3.5.2 Liquid Chromatography-Mass Spectrometry*……..……32

*1.3.5.3 Capillary Electrophoresis-Mass Spectrometry*……..……..34

*1.3.5.4 Emerging Mass Spectrometry Technologies*…….34

*1.3.5.5 Nuclear Magnetic Resonance Spectroscopy*…….....35

*1.3.5.6 Vibrational Spectroscopy*……...36

*1.3.5.6.1 Infrared Spectroscopy*……...37

*1.3.5.6.2 Raman Spectroscopy*……38

1.4 Chemometrics…….39

*1.4.1 Defining Chemometrics*……..39

*1.4.2 Chemometric Tools*……40

*1.4.2.1 Principal Component Analysis*……..……..40

*1.4.2.2 Partial Least Squares*……..…….42

*1.4.2.3 Random Forests*……43

1.5 Metabolomic Applications……..45

*1.5.1 Metabolomics in Food Quality*……...46

*1.5.1.1 Metabolomics in Food Safety*……...49

*1.5.1.2 Metabolomics in Food Component Analysis*……49

*1.5.1.3 Metabolomics & Mushrooms (A. bisporus)*……..50

**2. Aims & Objectives…….52**

**3. The use of fourier transform infrared spectroscopy and chemometric data analysis to evaluate damage and age of mushrooms (*Agaricus bisporus*)…….53**

3.1 Material and Methods……...53

*3.1.1. Mushrooms*…….53

*3.1.2. Mushroom Treatments*……54

*3.1.3. FTIR Spectroscopy*……..54

*3.1.4. Chemometric Data Analysis*……...55

3.2 Results and Discussion……..58

*3.2.1 Spectral Data*……..58

*3.2.2 Principal Component Analysis*……..……60

*3.2.3 Detection of Damage*……..63

*3.2.3.1 Random Forests*……63

*3.2.3.2 Partial Least Squares*……..……67

*3.2.4 Predicting Postharvest Age*…….……68

*3.2.4.1 Random Forests*……...68

*3.2.4.2 Partial Least Squares*……..……70

3.3 Conclusions……....72

**4. The use of nuclear magnetic resonance spectroscopy and chemometric data analysis to evaluate low levels of damage in mushrooms (*Agaricus bisporus*)……..……74**

4.1 NMR Lipid Profile of Mushrooms…….....74

4.2 Material and Methods……75

*4.2.1 Mushrooms*……..75

*4.2.2 NMR Profiling Protocol & Overview*…….....76

*4.2.2.1 Extraction*……..76

*4.2.3 NMR Measurements*……77

*4.2.4 Chemometric Data Analysis*……...78

4.3 Results and Discussion……..79

*4.3.1 Non-polar Phase Spectral Analysis*……79

*4.3.1.1 Assignment of Signals*……...82

*4.3.2 Polar Phase Spectral Analysis*……84

*4.3.3 Principal Component Analysis*……..……86

*4.3.4 Detection of Damage*……..88

*4.3.4.1 Random Forests*……88

*4.3.4.2 Partial Least Squares*……..…….91

4.4 Conclusions……92

**5. Metabolic profiling of mushrooms (*Agaricus bisporus*) using GC/MS & chemometrics to identify markers of damage and investigation of metabolic pathways……93**

5.1 Materials and Methods……...93

*5.1.1 Mushroom Treatments*……93

*5.1.2 Metabolic Profiling Protocol and Overview*……..…….94

*5.1.2.1 Extraction of Polar Phase*……95

*5.1.2.2 Extraction of Lipid Phase*…….96

*5.1.2.3 Fractionation*……96

*5.1.2.4 Transmethylation & Derivatisation of Lipid Phase (non-polar)*…..96

*5.1.2.5 Derivatisation of Polar Phase*……..97

*5.1.3 Analysis of Metabolites by GC/MS*…….97

*5.1.4 Data Analysis*……..98

*5.1.5 Chemometric Data Analysis*…….100

5.2 Results and Discussion……102

*5.2.1 Non-polar Metabolites*……..……102

*5.2.1.1 Fatty Acids (FAs)*……102

*5.2.1.1.1 Saturated Fatty Acids*……..104

*5.2.1.1.2 Unsaturated Fatty Acids*…….107

*5.2.1.2 Phenolic Compounds*……..109

*5.2.2 Polar Metabolites*…….110

*5.2.2.1 Amino acids*…….……110

*5.2.2.2 Sugars & Polyols*……112

*5.2.3 Principal Component Analysis*……..……..114

*5.2.4 Detection of Damage*……115

*5.2.4.1 Random Forests*……..……115

*5.2.4.1.1 Random Forests (models 2-6)*…….……120

*5.2.4.2 Partial Least Squares*……...……...122

*5.2.4.3 Metabolic Pathways*……125

*5.2.4.3.1 Lipoxygenase Pathway (LOX)*……126

*5.2.4.3.2 β-Oxidation Pathway (fatty acid metabolism)*……130

*5.2.4.3.3 Isomerisation of Glucose (myo-inositol)*…….132

*5.2.4.3.4 Glycolytic Pathway (D-mannose)*……...133

*5.2.4.3.5 TCA Cycle (Krebs cycle)*…….134

*5.2.4.3.6 Polyol Metabolism (glucitol/mannitol & hexitol)*……..…..134

*5.2.4.4 Correlation of Metabolites (correlation matrices)*…….136

*5.2.4.4.1 Correlation Matrices (polar metabolites)*……...136

*5.2.4.4.2 Correlation Matrices (non-polar metabolites)*……...143

5.3 Conclusions……..145

**6. Evaluating metabolic technologies for identifying markers of damage in mushrooms (*Agaricus bisporus*)……..147**

6.1 Sample Preparation……..147

*6.1.1 FTIR Spectroscopy*…..……..147

*6.1.2 NMR Spectroscopy*……..……..147

*6.1.3 GC/MS*…….……..148

6.2 Analysis……....148

*6.2.1 FTIR Spectroscopy*…..……..148

*6.2.2 NMR Spectroscopy*..……..148

*6.2.3 GC/MS*……...149

6.3 Results……..151

*6.3.1 FTIR Spectroscopy*……..……..151

*6.3.2 NMR Spectroscopy*…..……..151

*6.3.3 GC/MS*……..152

6.4 Conclusions……..153

**7. Conclusions……..155**

7.1 Overall Conclusions……155

7.2 Future Investigations……...156

**References……158**

**Publications……..193**

**LIST OF FIGURES**

**Figure 1.1** A schematic representation of a mushroom fruiting body……..…...3

**Figure 1.2** *A. bisporus* growing in a polyethylene bag.....….....7

**Figure 1.3** Structure of phenol……..……..…….....11

**Figure 1.4** Natural melangenous phenolics present in *A. bisporus*….....12

**Figure 1.5** Mechanisms by which L-tyrosine converts L-tyrosinase firstly to L-DOPA and then to *o*-dopaquinone and the following steps that lead to melanin formation……...13

**Figure 1.6** Pathway of carbohydrate routing in *Agaricus bisporus*……...19

**Figure 1.7** The mannitol cycle as originally postulated by Hult and Gatenbeck…….21

**Figure 1.8** Scheme representing a typical trehalose synthesis and degradation of fungi……..22

**Figure 1.9** Possible pathways of ammonia incorporation: (1) glutamate dehydrogenase (2) glutamine synthetase (3) glutamine: (GOGAT)…….23

**Figure 1.10** Ornithine/urea cycle with relevant biochemical pathways……..24

**Figure 1.11** General classification of metabolomics……...25

**Figure 1.12** Schematic representation of the process of metabolomic analysis…..27

**Figure 3.1** FTIR transmittance spectra of all the mushroom tissues in (a) 400-1800 $cm^{-1}$ (b) 2800-3050 $cm^{-1}$ and (c) 3050-4000 $cm^{-1}$ ranges……..…..…59

**Figure 3.2** Average FTIR spectrum of undamaged cap tissue 4000-400 $cm^{-1}$…..60

**Figure 3.3** PC1 *versus* PC2 score plots of undamaged mushroom tissue (a) caps (b) gills (c) stipes and damaged tissue (d) cap (e) gills and (f) tissues 0-7: Sample ages from 0-7 days……62

**Figure 3.4** Relative importance plots of variables that are important in the random forest models for predicting damaged/undamaged samples…..……..…….64

**Figure 3.5** Box plot showing the absorbance values for each tissue at (a) 1560 $cm^{-1}$ and (b) 1868 $cm^{-1}$……..……..……65

**Figure 3.6** Relative importance plot of variables that are important in the random forest model for predicting age……..……...70

**Figure 3.7** Frequency of generation of PLS regression modes for mushroom post-harvest age on the basis of the number of latent variables selected (a) undamaged caps (b) undamaged gills (c) undamaged stipes (d) damaged caps (e) damaged gills (f) damaged stipes…….71

**Figure 4.1** Overview of metabolite fingerprinting NMR spectroscopy protocol……...……..76

**Figure 4.2** (a) Bruker Avance III 500 MHz spectrometer and (b) NMR spectroscopy tubes containing mushroom samples stacked in the autosampler…….……..….78

**Figure 4.3** Illustration of the splitting into training & test datasets to assess the prediction of damage in mushrooms…….79

**Figure 4.4** Representative $^{1}H$ NMR spectrum of undamaged day zero non-polar phase cap tissue…….80

**Figure 4.5** $^{1}H$ NMR spectra of (a) undamaged cap *versus* (b) damaged cap day zero (non-polar phase )…… ……81

**Figure 4.6** $^{1}H$ NMR spectra of non-polar phase day zero undamaged (red) and damaged (blue) (a) gills and (b) stipes…….…….82

**Figure 4.7** $^{1}$H NMR spectrum (expanded) of lipidic fraction of undamaged cap tissue day zero containing possible signals for 1: C*H*$_2$COO, 2: C*H*$_2$CH=CHCH$_2$, 3: C*H*$_2$CH$_2$COO, 4: C*H*$_2$, 5: C*H*$_3$……..84

**Figure 4.8** Representative $^{1}$H NMR spectrum of undamaged day zero polar phase cap tissue……...……85

**Figure 4.9** $^{1}$H NMR spectra of (a) undamaged cap day zero *versus* (b) damaged cap day zero (polar phase)…….……...85

**Figure 4.10** PC1 *versus* PC2 score plots of non-polar phase gill tissue for (a) day zero samples and (b) day one samples……..87

**Figure 4.11** PC1 *versus* PC2 score plots of polar phase gill tissue for (a) day zero samples and (b) day one samples……..88

**Figure 4.12** Relative importance plot of variables that were identified as important by RF models for predicting damage (a) non-polar phase stipe tissue and (b) non-polar phase gill tissue……...…….90

**Figure 5.1** Overview of metabolic profiling (GC/MS) protocol……...95

**Figure 5.2** Structures of internal standards injected during the metabolic profiling study……..95

**Figure 5.3** Structure of the silylating agent N-methyl-N-trimethylsilyl-fluoroacetamide (MSTFA)……97

**Figure 5.4** Initial AMDIS window of a typical non-polar phase mushroom showing the chromatogram total ion count (TIC) in the top window and a raw MS spectrum in the bottom window….……...99

**Figure 5.5** AMDIS component window of a typical non-polar phase mushroom chromatogram showing the chromatogram (top) with all the deconvoluted components identified, the purity window and component information window

(middle) and a deconvoluted and noise removed mass spectrum belonging to the selected component (bottom)……..……...…….100

**Figure 5.6** AMDIS component window of a typical polar phase mushroom chromatogram showing the chromatogram (top) with all the deconvoluted components identified, the purity window and component information window (middle) and a deconvoluted and noise removed mass spectrum belonging to the selected component (bottom)……...……..100

**Figure 5.7** Non-polar phase GC/MS total ion chromatogram (TIC) of a mushroom extract……..……104

**Figure 5.8** McLafferty rearrangement mechanism of fatty acids……105

**Figure 5.9** Mass spectrum of octadecanoic acid methyl ester…….106

**Figure 5.10** Cleavage process of octadecanoic acid methyl ester…….106

**Figure 5.11** Mass spectrum of dodecanoic acid methyl ester……107

**Figure 5.12** Mass spectrum of linoleic acid methyl ester……..108

**Figure 5.13** Mass spectrum of benzoic acid methyl ester…….109

**Figure 5.14** (a) Polar phase GC/MS total ion chromatogram (TIC) of a mushroom extract (b) enlarged version……111

**Figure 5.15** Mass spectrum of the amino acid valine……112

**Figure 5.16** Mass spectrum of methyl-silyated (a) D-mannose and (b) myo-inositol……...…….113

**Figure 5.17** PC1 *versus* PC2 score plots of cap tissue for (a) day zero samples and (b) day one samples……..……..114

**Figure 5.18** VIP plot of metabolites that are important variables in the RF model for predicting damage…….116

**Figure 5.19** Box plot showing the quantity of myo-inositol at each damage level…….....117

**Figure 5.20** Bar plots with the semi-quantitative concentration of benzoic acid in mushroom tissues at different damage levels in day one samples……..118

**Figure 5.21** Tukey multiple comparison test plots comparing differences in mean levels of damage in (a) day zero samples and (b) day one samples (myo-inositol)……...…120

**Figure 5.22** Evolution of bootstrap resampling accuracy as a function of latent variables …….123

**Figure 5.23** A general overview of metabolism adapted from KEGG……..126

**Figure 5.24** Bar plots with the semi-quantitative concentration of linoleic acid in mushroom tissues at different damage levels in (a) day zero samples and (b) day one samples…….127

**Figure 5.25** Structures of the main eight-carbon volatiles……128

**Figure 5.26** Formation of 1-octen-3-ol……..130

**Figure 5.27** Formation and structures of myo-, neo- and scyllo-inositols……132

**Figure 5.28** Structure of D-mannose…….133

**Figure 5.29** Bar plot with the semi-quantitative concentration of D-mannose in mushroom cap tissues at different damage levels in day zero samples and day one samples……133

**Figure 5.30** Hexitol synthesis……135

**Figure 5.31** Correlation matrices of polar metabolites in day zero cap tissues at 0 min damage……...……..137

**Figure 5.32** Correlation matrices of polar metabolites in day zero cap tissues at 20 min damage…….137

**Figure 5.33** Correlation matrices of polar metabolites day zero cap tissues at 40 min damage……138

**Figure 5.34** Selected plots of response ratios of highly correlated metabolites seen in undamaged cap tissue……...139

**Figure 5.35** Correlation matrices of polar metabolites in day one cap tissues at 0 min damage……141

**Figure 5.36** Correlation matrices of polar metabolites in day one cap tissues at 20 min damage……..141

**Figure 5.37** Correlation matrices of polar metabolites in day one cap tissues at 40 min damage……142

**Figure 5.38** Correlation matrices of non-polar metabolites day zero cap tissues at 40 min damage……143

**Figure 5.39** Plots of response ratios for linoleic acid and pentadecanoic acid (gill tissue) at each damage level and day……144

**LIST OF TABLES**

**Table 1.1** Some recent metabolomic applications in food analysis…….48

**Table 1.2** A summary of metabolites that have been identified in mushrooms by metabolomic techniques……..……..51

**Table 3.1** Summary of samples analysed by FTIR spectroscopy……..55

**Table 3.2** ANOVA table of the effect of damage, tissue and age on the absorption at specific wavenumbers identified as important variables for predicting damage by RFs…….…….66

**Table 3.3** Summary of results for mushroom discrimination using PLS-DA on the basis of damage……...67

**Table 3.4** Confusion matrix and the error rate for the prediction of mushroom age…….....69

**Table 3.5** Summary of PLS regression results for the prediction of post-harvest age (day 0-7 inclusive) in undamaged and damaged mushrooms……71

**Table 4.1** Summary of samples analysed by NMR spectroscopy…….……..…….…….75

**Table 4.2** Summary of RF models produced to discriminate between damaged and undamaged mushrooms using NMR data……..89

**Table 4.3** ANOVA table of the effect of damage on each important variable as indicated by the RF model for all non-polar phase samples……..…...91

**Table 4.4** Performance statistics of PLS-DA models built using NMR data…...91

**Table 5.1** Summary of samples analysed by GC/MS……..94

**Table 5.2** Metabolites identified by GC/MS as components of *A. bisporus*…..103

**Table 5.3** Mass spectrometric characteristic ions and gas chromatographic retention times of fatty acid methyl esters……..108

**Table 5.4** Mass spectrometric characteristic ions and gas chromatographic retention times of abundant phenolic compounds……..110

**Table 5.5** Mass spectrometric characteristic ions and gas chromatographic retention times of amino acids……112

**Table 5.6** Mass spectrometric characteristic ions and gas chromatographic retention times of abundant sugars and polyols……..113

**Table 5.7** ANOVA table of the effect of damage on individual metabolites….119

**Table 5.8** Summary of all RF models for predicting damage, including OOB error rates, variables of importance & associated metabolic pathways……..122

**Table 5.9** Summary of results for mushroom discrimination on the basis of damage (all data)…….123

**Table 5.10** Performance statistics of PLS-DA models built using GC/MS data.124

**Table 6.1** Comparison of analytical platforms used for detecting damage in mushrooms……..150

**Table 6.2** Summary of metabolomic techniques and the best models used to evaluate damage in mushrooms……..153

# 1. INTRODUCTION

# 1. Introduction

## 1.1. General Introduction

### *1.1.1. Mushroom History*

Mushrooms have been part of the fungal diversity for around 300 million years, probably being collected in the wild as food and possibly for medicinal purposes by prehistoric humans. With the widespread cultivation of plants for food, mushrooms were eventually cultivated and not simply picked in the wild.

The first mushroom to be cultivated was probably *Auricularia auricular* around 600 A.D. The historical record for *Lentinula edodes* is much more considerable and indicates that *L. edodes* was first cultivated circa 1000 A.D. according to the famous Chinese Book of Agriculture published in 1313 (Chang and Miles, 1978). The most significant advance in the field occurred in the seventeenth century when *Agaricus bisporus* was cultivated in France (Skovgaard, 2002) using horse manure as a substrate. The horse manure was initially colonised by mushroom spores from wild sources and when a mushroom bed had been harvested the substrate was broken up, dried and used as 'spawn' to inoculate new beds. A side effect of this method was that each site had to be abandoned after a year or two due to the build up of disease and insects.

Later developments in mushroom cultivation included the growing of mushrooms indoors using a pure culture spawn containing living mycelium of the desired mushroom species. Pure culture mushroom spawn for *A. bisporus* was first achieved in 1886 in the United Kingdom, in 1894 in France, and in 1902 in the United States.

In Ireland mushrooms were first grown commercially in the mid 1930's with exports to Great Britain beginning in 1947 (Chang and Miles, 2004).

### *1.1.2. Fungal Kingdom*

Historically mushrooms were classified with plants. However, modern studies established that mushrooms along with other fungi, have features of their own that are sufficiently and significantly distinct to place them in a separate fungal kingdom, the Kingdom Myceteae. Fungi differ from the plant kingdom by their possession of a cell wall that is different in composition. The mode of nutrition of fungi is heterotrophic but, unlike animals is absorptive as opposed to digestive.

A mushroom is defined as a higher fungus belonging to the sub-kingdom Dikarya with a distinctive fruiting body that can be above or below ground. Higher fungi may be ascomycota or basidomycota, edible or non-edible, poisonous and also medicinal species (Courtecuise and Duhem, 1995; Chang and Miles, 2004). The phyla basidiomycota include some of the most familiar and conspicuous of all fungi including the common button mushroom *A. bisporus*.

### *1.1.3. Mushroom Morphology*

*A. bisporus* mushrooms consist of three different tissues cap, gills and stipe (Figure 1.1).

1. The cap is fleshy and hemispherical and as the cap expands it becomes flattened in order to protect the gills. The cap colour ranges from white to cream at first, becoming brownish with age and damage.
2. The gills are the reproductive tissues of the mushroom and produce millions of spores, which are located underneath the cap. In many mushrooms the gills are covered early in development by a veil and in the mature mushroom the remains of this veil can be seen as a ring around the stipe. Over time the colour of the gills change from a pinkish colour to a brown black colour as the spores mature.

3. The stipe is white in colour and it is connected at its base to the mycelium in the compost. Its function is to lift the cap above the compost in order for the spores to be released (Flegg *et al.*, 1985; Courtecuise and Duhem, 1995).

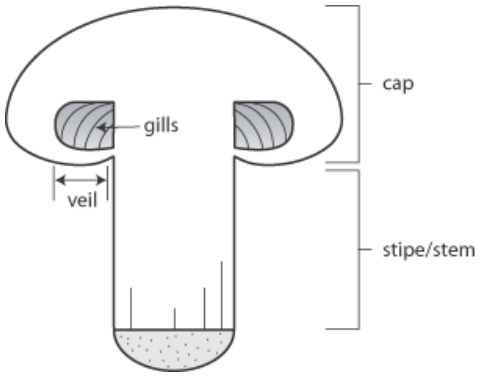



**Figure 1.1 A schematic representation of a mushroom fruiting body (Mohacek-Grosev *et al.*, 2001)**

The vegetative part of the mushroom is called the mycelium, which is made up of filaments called hyphae. These filaments grow only at the tip or at specialised regions and form a system of branching threads and cordlike strands that branch out throughout the soil, compost or other material on which the mushroom is growing. After a period of growth and under favourable conditions the established (matured) mycelium produces the fruiting body (Chang and Miles, 2004).

### *1.1.4. Mushroom Physiology*

Mushrooms are heterotrophs and acquire their nutrients by absorbing soluble inorganic and organic materials from substances like wood logs, manure composts or other organic synthetic composts. It is important for them to find organic carbon in their environment i.e. in their substrate. This carbon source provides the skeletal carbon for organic compounds and the energy for the anabolic processes. Other elements necessary for fungal life include; oxygen, hydrogen, phosphorus, potassium, copper, iron, zinc and vitamins. Water is also

essential to fungi for its role during the growth cycle. Heat and light also play a part in their physiology (Beelman *et al.*, 2003; Chang and Miles, 2004).

### *1.1.5. Nutritional Attributes*

The nutritional value of the mushroom originates from its chemical composition. It should be noted that mushroom composition varies greatly due to their strains, cultivation techniques (including different substrates), maturity at harvest and methods of analysis (Beelman and Edwards, 1989). Mushrooms are considered health foods as they are low in calories, fats and essential fatty acids, high in protein, vitamins and minerals. The high levels of potassium, phosphorus and selenium along with very low levels of sodium add to the beneficial attributes of mushrooms. Mushrooms have also been reported as therapeutic foods, useful in the prevention of diseases such as hypertension, hypercholesterolemia, atherosclerosis and cancer (Crisan and Sands, 1978; Kuraswa *et al.*, 1982; Manzi *et al.*, 2001).

#### *1.1.5.1. Carbohydrate & Fibre*

Carbohydrates are generally the main components of the mushroom (ranging from 30-75% of dry weight) (Manzi *et al.*, 2004; Colak *et al.*, 2007). Glucose, mannitol and α-trehalose are the main representatives of monosaccharides, their derivatives and oligosaccharide groups, respectively. Usual contents of glucose and trehalose are low. The content of mannitol, which participates in volume growth and firmness of fruiting bodies, differs widely (Barros *et al.*, 2007).

The reserve polysaccharide of mushrooms is glycogen with usual content at about 5-10% of dry matter. Chitin is a water-insoluble structural polysaccharide, accounting for up to 80-90% of dry matter in mushroom cell walls (Manzi *et al.*, 2004).

The content of fibre varies greatly between species with values for *A. bisporus* reported as 10.4% (dry weight) (Chang and Miles, 2004).

*1.1.5.2. Protein & Amino Acids*

The proteins of cultivated mushroom contain all nine essential amino acids i.e. those which the body cannot synthesise (lysine, methionine, tryptophan, threonine, valine, leucine, isoleucine, histidine and phenylalanine). Mushrooms generally contain 19-35% protein (dry weight) (Chang and Miles, 2004).

*1.1.5.3. Lipids*

Lipids in the cultivated mushroom *A. bisporus* have been investigated extensively (Weete, 1980). The lipid content in different species of mushrooms ranges from 1.1 to 8.3% (dry weight), with an average content of 4.0% (Chang and Miles, 2004).

The acids include $C_{12}$-$C_{20}$ even-numbered fatty acids (Holtz and Schisler, 1971; Prostenik *et al.*, 1978; Weete *et al.*, 1985) and $C_{16}$-$C_{24}$ hydroxy fatty acids (Prostenik *et al.*, 1978), with oleic, linoleic and palmitic acids predominating. These may exist in their free form or be conjugated to other lipid constituents. Esterification of the glycerol with fatty acids may lead to the formation of mono-, di- or triglycerides. Stancher and colleagues expanded the observed range of free and bound fatty acids to include $C_8$ and $C_{13}$-$C_{17}$ odd-numbered acids (Byrne and Brennan, 1975; Stancher *et al.*, 1992).

The chief unsaturated fatty acid of mushrooms lipids, linoleic acid, is the precursor of the mushroom alcohol (1-octen-3-ol) (Tressl *et al.*, 1982). This alcohol together with the two associated $C_8$ ketones (1-octen-3-one, 3-octanone), constitute the main volatiles and are considered the major contributors to the characteristic mushroom flavour.

*1.1.5.4. Vitamins & Minerals*

It has been reported that edible mushrooms are a good source of several vitamins including thiamine ($B_1$), riboflavin ($B_2$), niacin, biotin and ascorbic acid (Vitamin C) (Crisan and Sands, 1978; Chang and Miles, 2004). Mushrooms are deficient in vitamin $D_2$. However, they are found to be a rich source of ergosterol, the precursor of vitamin $D_2$. Ergosterol in

mushrooms can be converted into vitamin $D_2$ by UV irradiation (Mau *et al.*, 1998; Jasinghe and Perera, 2006).

They are also a good source of minerals. The major mineral constituents are potassium (K), phosphorus (P), sodium (Na), calcium (Ca), magnesium (Mg) and selenium (Se). Copper (Cu), zinc (Zn), iron (Fe), manganese (Mn), molybdenum (Mo) and cadmium (Cd) make up the minor mineral constituents (Chang and Miles, 2004).

### *1.1.6. Mushroom Production*

The use of plastic bags for mushroom growing was developed in Denmark in 1959 and spread to France and Germany (MacCanna, 1984). The system of growing in plastic bags (Figure 1.2) and tunnels was the basis of the expansion of the Irish mushroom industry in the 1980s.

The two basic requirements for mushroom growing are good substrate (i.e. compost) and the right environmental conditions. As compost quality is largely outside the control of mushroom growers, their main contribution to final product quality is crop management. This involves controlling temperature, relative humidity, watering, ventilation and $CO_2$ levels. Modern mushroom houses are equipped with computerised environmental control systems for this purpose (Teagasc, 1994).

During the crop cycle, mushrooms are harvested in a rhythmic pattern of breaks or flushes that occur at approximately seven day intervals. After two flushes, production declines rapidly and a grower must decide to terminate the crop and start anew or face dwindling harvest of mushrooms from each successive flush.

The steps involved in mushroom production are:

Phase 1: Preparation of compost

Phase 2: Pasteurizing and final composting

Phase 3: Spawning

Phase 4: Casing

Phase 5: Harvest (MacCanna, 1984; Teagasc, 1994; Chang and Miles, 2004)

The Food and Agricultural Organisation of the United Nations (FAO) issued a small booklet called 'International Standards for Edible Fungi' (Codex Alimentarius Commission No. 38) in 1970. The standards are as follows:

1. Buttons – Mushrooms with membranes closed, only just forming. Stem length not to exceed 2 cm (¾ inch), cap diameter 2.5 to 6 cm (1 to 2½ inches).
2. Caps – Mushrooms with membranes well developed or just opening, with cap retaining a pronounced cap shape. Stem length not to exceed 2.5 cm (1 inch) from the apex. Cap diameter 2.5 to 7 cm (1 to 2¾ inch).
3. Flats or Opens – Mushrooms that have advanced beyond the cap stage, the cap forming the letter 'T' with the stipe. Cap diameter 2.5 to 7 cm (1 to 3½ inch) and stem length not to exceed 2.5 or 3 cm, according to the class.

These definitions are widely accepted but the measurements may vary on a wider range.

**Figure 1.2 *A. bisporus* growing in a polyethylene bag**

### *1.1.7. Irish Mushroom Industry*

Mushrooms are the single most important horticultural crop grown in Ireland (Teagasc, 2007). From small beginnings the industry has grown to become a major producer of fresh mushrooms and the third largest fresh *A. bisporus* exporter in Europe, after Poland and the Netherlands (Van Horen, 2008). Over 60,000 tonnes are produced annually with 80% (>45,000 tonnes) being exported to UK markets (Teagasc, 2007).

The mushroom industry expanded dramatically during the 1980s and 1990s with the introduction of a new concept of growing called the 'satellite' system. The satellite system was invented in Ireland and is quite simple. Compost companies would sell compost to an associated group of growers, and then buy the mushroom crop back from the growers. Further marketing of the mushrooms was handled by the sales organisation of the compost company. This resulted in a very efficient production and marketing system, with growers having a secure source of compost and a guaranteed market for their mushrooms (http://www.mushroombusiness.com).

Ireland was ranked third in the world fresh-mushroom exports behind China and The Netherlands between the years 2000-2004 (Grogan, 2008). However, in recent years the production of mushrooms has decreased in Ireland by 20%, while in Poland it has increased by 100%. Poland is now the biggest producer and exporter of mushrooms in Europe (Van Horen, 2008).

Competition between The Netherlands and Poland has had a negative effect on established mushroom industries throughout Europe, with the numbers of farms declining in most mushroom producing countries. Ireland has seen a steady decline in the number of mushroom growers from 504 in 2000 to 100 in 2006 (Grogan, 2008).

The only way to move forward and remain profitable is to maximise efficiencies in all links of the production and supply chain. The Irish mushroom industry still controls a

significant share of the UK mushroom market, however, in order to continue to consolidate this position a number of challenges must be addressed:

- Top quality compost needs to be consistently supplied to growers.
- Growers need to maximise efficiencies at production level by increasing quality, yields and throughput of compost.
- Logistical costs need to be kept to a minimum (Neary, 2003).

Despite the difficulties faced over the past few years, growers are investing in order to improve the growing system, which will allow the Irish industry to maintain a strong position in the European export market place (Grogan, 2008).

## 1.2. FACTORS EFFECTING MUSHROOM QUALITY

### *1.2.1. Shelf-life*

Mushrooms have a short shelf-life compared to most vegetables. This is due to their thin and porous epidermal structure which results in high respiration rates which give rise to deterioration immediately after harvest (Brennan *et al.*, 2000). Maturation and senescence are phenomena which together with microbial activity contribute to postharvest deterioration of fresh mushrooms (Beelman, 1987).

Mushrooms are very perishable products with a usual shelf-life of less than 3 days at ambient temperature (Lee, 1999), and from 8 to 10 days under refrigeration (Burton, 1989). Extending the shelf-life of mushrooms is desirable as an extra few days would compensate for time in transit, thereby giving more flexibility to processors, retailers and consumers alike (Brennan *et al.*, 2000). A longer shelf-life would be particularly important for any exporting country (i.e. Ireland) for which access to the food markets in larger neighbouring countries within Europe is vital.

A number of recent studies have been carried out into extending the shelf-life of mushrooms. This research has focused on developing postharvest technologies and improving packaging designs to increase shelf-life (Kim *et al.*, 2006; Aguirre *et al.*, 2008; Gowen *et al.*, 2008b; Taghizadeh *et al.*, 2009). Although postharvest treatments delay mushroom senescence and extend shelf-life the product quality during storage is mainly dependent on the quality of the harvested mushroom (Flegg *et al.*, 1985).

### *1.2.2. Mushroom Quality*

Quality is a difficult term to define as it is not a single recognisable characteristic. One definition of quality is the 'totality of characteristics of an entity (product, service, process, activity, system, organisation, person) that bear on its ability to satisfy stated and implied needs'(Will and Guenther, 2007). This means that the product must meet the consumer's expectations. Consumers judge quality of fresh produce on the basis of appearance including 'freshness' at the time of initial purchase. They are also concerned about the nutritional quality and safety of the products.

Quality in mushrooms is judged by the consumer and can be assessed as a combination of visual appearance, freshness, colour, size, maturity stage, development stage, firmness, turgor, microbial growth, blemish-free, weight loss and blotching (Burton, 1989; Carey and O'Connor, 1991; Vizhanyo and Felfoldi, 2000). High quality mushrooms are defined as white, unblemished, firm, clean and with a closed veil (Eastwood and Burton, 2002).

Mushroom quality is affected by pre-harvest factors such as strain, composting, casing, watering and temperature regime, atmosphere composition and flush number. The post-harvest factors include picking, transportation, chilling and packaging (Gormley, 1986). Another factor that affects mushroom quality after harvest includes disease (Beelman *et al.*, 1989), together with improper handling during harvest, packaging and transport which results in bruising.

The single most important consideration of quality in fresh mushrooms is colour as it is the first characteristic that consumers see. Therefore any discolouration i.e. brown marks are easily seen and viewed by the consumer as an indictor of low quality (Burton, 2004).

*1.2.2.1. Mushroom Discolouration*

The loss of whiteness during storage results in large economic losses every year. Cap browning, softening, moisture loss and development of some off-flavours cause the mushroom to loose its marketability (Jolivet *et al.*, 1998). The mushroom is susceptible to discolouration induced by senescence, damage and physiological disorders (Vizhanyo and Felfoldi, 2000).

Browning of mushrooms is a result of the formation of pigments called melanins. These melanins are formed by enzymatic reactions containing phenolic or polyphenolic molecules (Falguera *et al.*, 2010).

*1.2.2.1.1. Phenolic Compounds*

The structure of phenol ($C_6H_5OH$) (Figure 1.3) is that of a hydroxyl group (-OH) bonded to a phenyl ring ($-C_6H_5$).

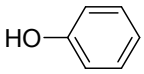



**Figure 1.3 Structure of phenol**

Polyphenolics are a multiplicity of different phenolic compounds i.e. compounds composed of aromatic benzene ring(s) substituted with hydroxyl groups, including all functional derivatives. The main natural phenolics present in the mushroom *A. bisporus* are glutaminyl-4-hydroxybenzene, p-aminophenol, phenylalanine and tyrosine (Figure 1.4).

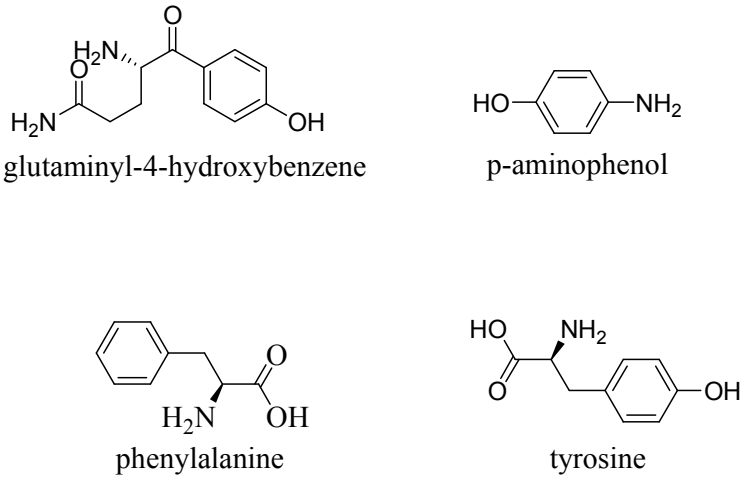



**Figure 1.4 Natural melangenous phenolics present in *A. bisporus***

Two distinct mechanisms of phenol oxidation are responsible for mushroom browning: (a) activation of tyrosinase, an enzyme belonging to the polyphenoloxidase (PPO) family and/or (b) spontaneous oxidation (Jolivet *et al.*, 1998).

(a) Activation of tyrosinase: Tyrosinase (PPO) is a copper-containing enzyme that catalyses two distinct reactions involving molecular oxygen with various phenolic substrates: (1) the *o*-hydroxylation of monophenols to *o*-diphenols (monooxygenase or cresolase activity) and (2) the subsequent oxidation of *o*-diphenols to *o*-quinones (diphenolase or cathecholase activity). Later polymerisation of these compounds leads to the formation of a heterogenous group of melanins (Figure 1.5) (Duckworth and Coleman, 1970; Aydemir, 2004).

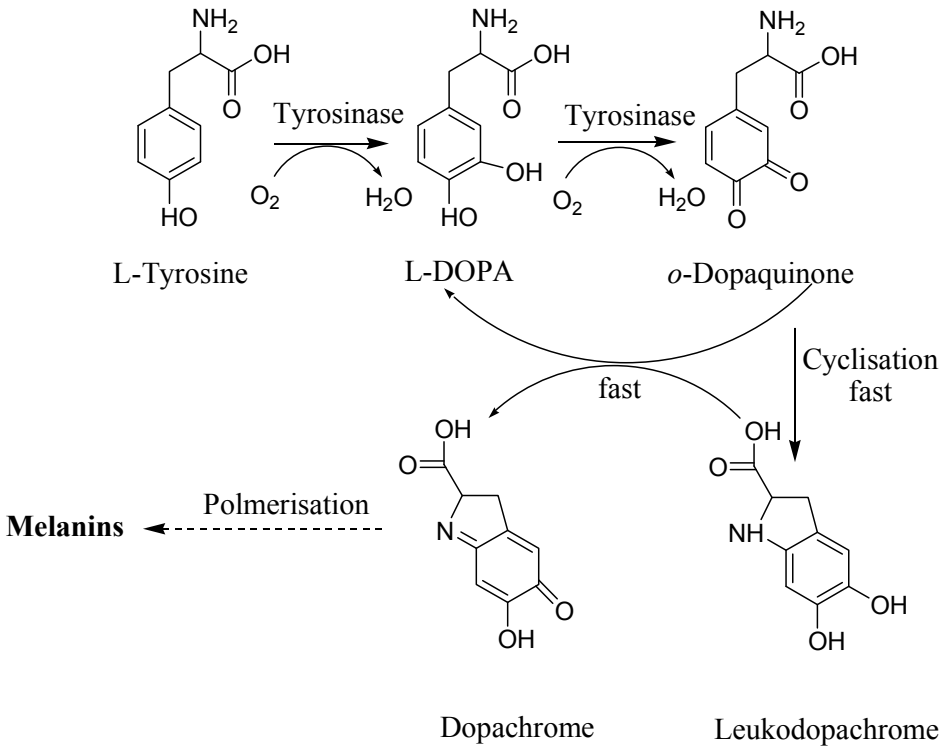



**Figure 1.5 Mechanism by which L-tyrosine converts L-tyrosinase firstly to L-DOPA and then to *o*-dopaquinone, and the following steps that lead to melanin formation (Falguera *et al.*, 2010)**

Colour is the main quality parameter for fresh mushrooms, with initial studies of tyrosinases being motivated by a desire to understand and prevent enzymatic browning that occurs in the presence of air when mushrooms are cut or bruised. This problem of postharvest browning has been tackled from several aspects (Kuyper *et al.*, 1993; Martinez and Whitaker, 1995; Brennan *et al.*, 2000). However, there has been a recent interest in tyrosinases due to discoveries of beneficial properties on health, such as antioxidative, anti-inflammatory, immune and anti-tumour properties (de Faria *et al.*, 2007).

The three major stress factors that induce mushroom discolouration are; senescence and damage affecting healthy mushrooms and microbial spoilage e.g. bacterial/viral infections affecting diseased mushrooms (Vizhanyo and Felfoldi, 2000).

*1.2.2.1.2. Senescence*

Senescence is a natural deteriorative process that takes place postharvest as soon as the mushroom is picked i.e. removed from the soil. Mushrooms after harvest are cut off from nutrients and water and begin to mobilise and redistribute nutrients (Burton, 2004). It is an oxidative process that involves degradation of the cellular and sub-cellular structures and macromolecules, and the mobilisation of the products of degradation to other parts of the mushroom. Susceptibility to oxidative stress depends on the overall balance between production of oxidants and antioxidant capability of the cell (del Rio *et al.*, 1998). Permeability changes during senescence have been linked with a simultaneous decline in membrane lipid (Ferguson and Simon, 1973). Therefore, as a result of senescence or natural ageing the cell membranes become disrupted and compartmentalisation is lost, allowing enzymes and substances to mix, thereby accelerating browning (Jolivet *et al.*, 1998).

*1.2.2.1.3. Mechanical Damage*

Mechanical damage or bruising has been defined as damage to tissue by external forces causing physical change in texture and/or eventual chemical alterations of colour, flavour and texture (Mohsenin, 1986). The brown discolouration in mushrooms caused by mechanical damage is largely confined to the skin tissue, the high levels of phenols and polyphenol oxidase in the skin tissue being one of the main reasons for this. Another reason is that the skin tissue is only loosely attached to the main flesh of the mushroom sporophore, which means that these surface cells absorb most of the energy of mechanical damage (Burton *et al.*, 2002).

Mushroom bruisability can vary from crop to crop and even within a crop. Sensitivity to mechanical damage is not only determined by genetics but also depends on a number of environmental and agronomic factors. Flush number and strain were highlighted by Burton

and colleagues as having a large influence on bruisability. The other major factors identified in the study were water potential of casing and humidity in the growing room (Burton *et al.*, 2002).

Mechanical damage is a consequence of inappropriate harvest, manipulation and transportation. Mushroom cultivation may be the first exposure to mechanical damage as mushrooms growing in close proximity of each other may brush off one another in turn causing tissue damage and bruising. The presence of small flies in a growing tunnel can interact with mushroom surfaces also resulting in damage. The problem of bruising can be controlled somewhat during production, although it is harder to control during transport and packaging.

Mushroom picking and handling is another area where mechanical damage occurs. A small mechanical force such as hands slipping over a mushroom during picking can cause damage and brown discolouration (Burton *et al.*, 2002; Burton, 2004). The importance of careful handling must be stressed. Mushrooms should be picked gently, and as dry as possible, into the tray from which they are going to be marketed (Gormley, 1986; Gormley, 1987).

Transportation is an area where mechanical damage also occurs. Damage caused by transport vibration was assessed on different species of fruit and vegetables such as peaches (O'Brien *et al.*, 1995), apricots (O'Brien and Guillou, 1969) and tomatoes (Singh and Singh, 1992). The loss of fresh fruit and vegetables during transport and distribution has been estimated to be above 30% in China, given their sensitivity to mechanical damage (Zhou *et al.*, 2007). An efficient transport system in terms of trucks with good suspension to ensure a smooth journey together with controlled refrigeration for the mushroom is essential (Gormley, 1987).

Much research has been carried out to investigate ways to reduce bruising in order for mushroom growers to optimise growing procedures for less bruised, high quality mushrooms. Treatments including refrigeration, overwrapped packages, modified atmosphere packages, irradiation, wash solutions and stipe trimming have been introduced to improve postharvest quality (Simon *et al.*, 2005; Kim *et al.*, 2006; Roy *et al.*, 2006; Eissa, 2007). These treatments can only act to preserve the quality of the product generated in cultivation. Different irrigation water treatments applied throughout the growth of a mushroom crop can begin to improve mushroom quality during the cultivation process (Miklus and Beelman, 1996).

Recent studies have shown the potential use of spectroscopic techniques coupled with chemometrics to detect damage in fresh mushrooms. These findings reveal the possibility of developing a tool that could detect damage before browning becomes visible thereby reducing economic losses for the industry (Esquerre *et al.*, 2009).

*1.2.2.1.4. Microbial & Viral Spoilage*

The most important disease on mushroom farms is brown blotch, which accounts for an estimated crop loss of 5-10% in the UK, with further losses occurring after mushrooms have been harvested (Fermor *et al.*, 1991).

Discolouration of *A. bisporus* caused by pathogenic pseudomonads, the so-called blotch diseases are well documented. These bacteria grow and break down mushroom fibres which soften the mushroom leading to enzymatic browning. The major species responsible for this is *Pseudomonas tolaasii* which produces a toxin that lyses mushroom cells (Gormley, 1975), resulting in sunken dark brown lesions (Tolaas, 1915; Paine, 1919). *Pseudomonas reactans* causes mild dark purple to light brown discolouration and a slight surface depression that becomes deeper and darker with age (Wells *et al.*, 1996), while the pale yellowish red discolouration that develops into a reddish ginger-coloured

discolouration (ginger blotch disease) is characteristic of *Pseudomonas gingeri* (Wong *et al.*, 1982).

Much research has been carried out in order to find an adequate method to prevent or control the disease. Manipulation of the environmental conditions (i.e. relative humidity, temperature etc) has proved to be essential but not sufficient.

A pathogen called La France virus impacted severely on the mushroom industry in the USA and the UK in the 1960s (Schisler *et al.*, 1967). In recent years a novel disease has become prevalent in the British mushroom industry. Mushroom virus X (MVX) is an enigmatic disease which causes crop delay, pinning disruption, poor quality and occasionally brown off-coloured mushrooms (Gaze *et al.*, 2000). In Ireland the brown mushroom symptom tends to occur in isolation and is not associated with yield reductions, but the symptom causes a significant loss of quality.

### *1.2.3. Mushroom Variability*

A batch has been defined by Schoten and Van Kooten as all individuals sharing the same harvest date, grower and cultivar (Schouten and Van Kooten, 1998), implying a common growth history. If all individuals in a batch were identical and stored under the same conditions, they would all reach the quality limit at the same time. However, this is not the case because of the presence of biological variation (Schouten *et al.*, 2004).

Biological variation is described as the composite of biological properties that differentiate individual units of a batch. This means that due to differences between individuals, every batch will be different and will behave differently, depending on the extent of variation in the product and the batch (Tijskens and Konopacki, 2003).

One of the main problems in mushroom technology (as in many fresh products) is the uncontrollable effect that product variability has on the management of the product. Agricultural products (including mushrooms) are managed by sorting and grading of

produce based on appearance, texture, colour shapes and sizes. The manual sorting and grading systems are based on traditional visual quality inspection performed by human operators which is tedious, time consuming, slow and inconsistent. Cost effective, consistent, high speed and accurate sorting systems can be achieved with machine vision-assisted grading and imaging techniques (hyperspectral imaging systems). Computer application and imaging systems in agriculture and food industries have been applied in the areas of sorting fresh products, detection of defects such as bruises and dark spots on fresh fruits, mushrooms and vegetables (Brosnan and Sun, 2004; ElMasry *et al.*, 2009; Gowen *et al.*, 2009).

## 1.3. METABOLOMICS

### *1.3.1. Mushroom Metabolism*

Metabolism is described as the sum of the chemical activities of a cell. It can be divided into two parts – catabolism and anabolism. Catabolism refers to the breakdown of a substance into simpler forms with the liberation of energy. An example of catabolism is the breakdown of the polysaccharide cellulose to glucose, then to pyruvic acid and eventually $CO_2$. Anabolism is the synthesis of cell materials and this process requires energy. An example of this process is the synthesis of cell wall polysaccharides such as glucans from component monosaccharides.

In mushroom biology metabolism is illustrated as follows:

1. The breakdown of substrate materials yielding energy and providing smaller more soluble compounds.
2. Those energy controlled processes that are needed for the transport of certain materials across the cytoplasmic membrane, and

3. Synthesis of cell materials, including the cell wall, from the compounds that have entered the cell (Chang and Miles, 2004).

*1.3.1.1. Carbohydrate Metabolism*

Although studied in a wide variety of organisms carbohydrate metabolism in fungi has received little attention and therefore data on acquisition and conversion of carbon compounds by fungi is scarce. Research has mainly focused on the yeast *Saccharomyces cerevisiae* and the filamentous fungi *Aspergillus nidulans*, *A. niger* and *Neurospora crassa* (Wannet *et al.*, 2000).

Simple saccharides (mainly glucose) generally enter glycolysis after being transported into the *Agaricus bisporus* cell. Depending on the metabolic status of the cell, the glucose molecule can then be diverted to several biochemical pathways (Figure 1.6).

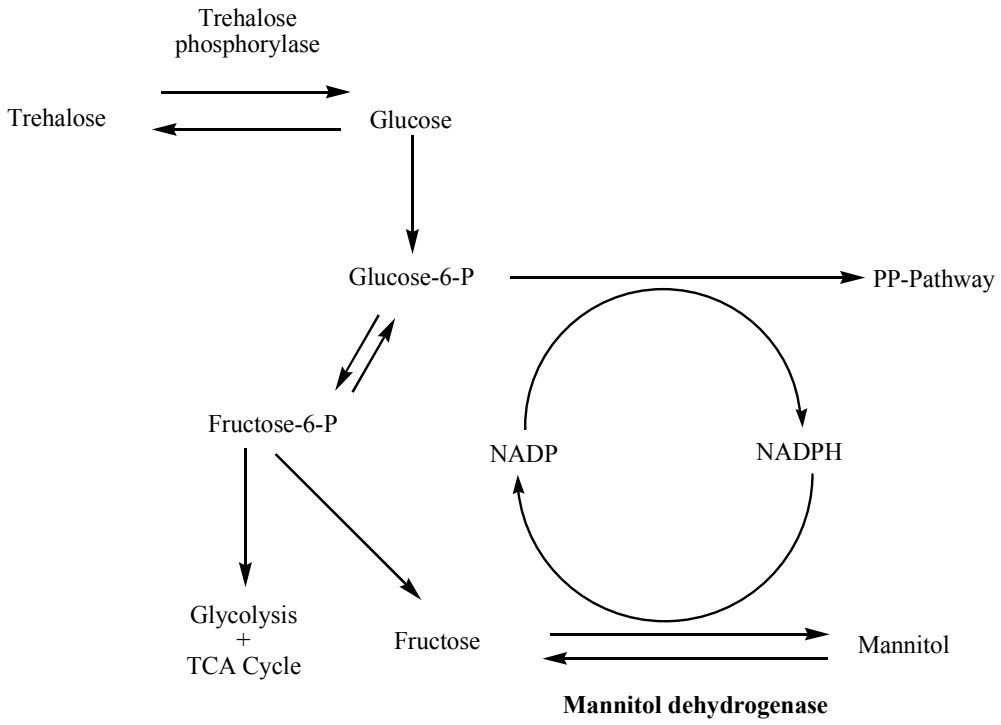



**Figure 1.6 Pathway of carbohydrate routing in *Agaricus bisporus* (Wannet *et al.*, 2000*)***
PP-Pathway: Pentose phosphate pathway

It should be noted that in *A. bisporus* the Krebs cycle (or TCA cycle) is blocked at the 2-oxoglutarate dehydrogenase complex (Rast *et al.*, 1976). This complex can be functionally replaced by the enzymes of the 4-aminobutyrate shunt: glutamate decarboxylase, 4-aminobutyrate aminotransferase and succinate-semialdehyde dehydrogenase.

In the case of *Agaricus bisporus* studies on the carbon metabolism have mainly focused on abundantly synthesised mannitol and trehalose (Wannet *et al.*, 2000).

*1.3.1.1.1. Metabolism of Mannitol*

Mannitol is the main storage carbohydrate where it can contribute up to 20% of the mycelium dry weight and up to 50% of the fruit body dry weight (Stoop and Mooibroek, 1998; Ruijter *et al.*, 2003). Mannitol is an abundant sugar in nature occurring in bacteria, algae, lichens, fungi and other vascular plants.

In fungi mannitol is reputed to play different roles including osmoregulation, serving as a storage or translocation carbohydrate, regulating co-enzymes and regulating cytoplamsic pH by acting as a sink or source for protons (Lewis and Smith, 1967; Jennings, 1984). Mannitol has also been proposed to have a role in oxidative stress protection. The ability of mannitol to quench reactive oxygen species (ROS) *in vitro* is proven (Smirnoff and Cumbes, 1989) and there is a growing pool of evidence that mannitol may have a similar role in pathogenic fungi during infection.

The mannitol cycle was postulated by Hult and Gatenbeck in 1978 from studies of cell-free extracts of the fungus *A. alternate* (Figure 1.7). The cycle consists of four enzymes: mannitol-1-phosphate dehydrogenase (MPD), $NADP^+$-mannitol-2-dehydrogenase (MDH), mannitol-1-phosphate phosphatase (MPP) and hexokinase (HK). The mannitol cycle pathway branches off from glycolysis at fructose-6-phosphate. Mannitol biosynthesis is traditionally thought to occur through the dephosphorylation of mannitol 1-phosphate. The

formed mannitol is then consumed through oxidation to fructose thereby completing the cycle (Hult and Gatenbeck, 1978).

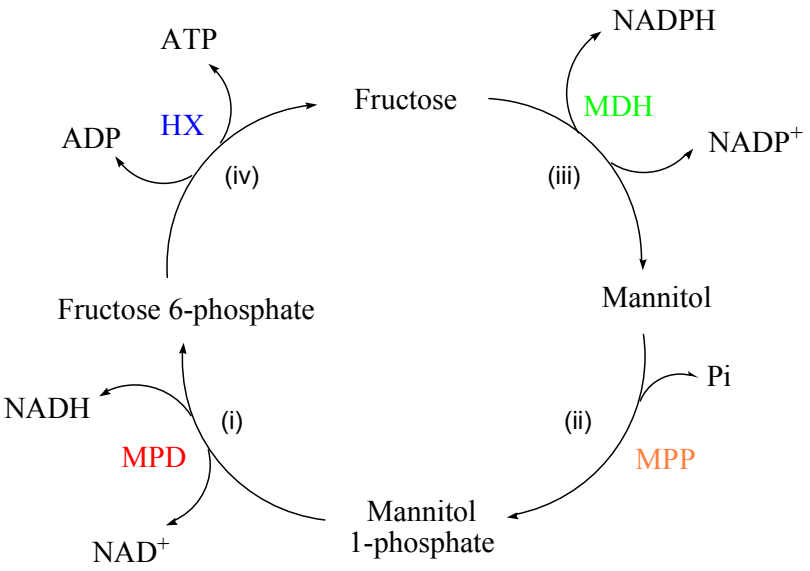



**Figure 1.7 The mannitol cycle as originally postulated by Hult and Gatenbeck (Hult and Gatenbeck, 1978)**
(i) The mannitol pathway branches off from glycolysis at fructose-6-phosphate, which is converted to mannitol-1-phosphate by MPD. (ii) Mannitol biosynthesis is traditionally thought to occur through the dephosphorylation of mannitol-1-phosphate to mannitol MPP, (iii) Mannitol is then converted by MDH through oxidation to fructose, which is subsequently phosphorylated by HX (iv) to fructose-6-phosphate, thus completing the cycle. HX: hexokinase: MDH, $NADP^+$ -mannitol-2-dehydrogenase; MPP: mannitol-1-phosphate phosphatise; MPD: mannitol-1-phosphate dehydrogenase

*1.3.1.1.2. Metabolism of Trehalose*

Trehalose is a non-reducing disaccharide (α-D-glucopyranosyl-α-D-glucopyranoside) and consists of two α-linked glucose sugars and in *A. bisporus* it is thought to serve as a reserve carbohydrate which is degraded under specific conditions (e.g. fructification) and has been suggested as a possible sugar translocated from the mycelium to the sporopore (Wells *et al.*, 1987; Wannet *et al.*, 1999). For a long time trehalose in fungi was thought to function as a storage compound, however, it has been found to function as a stress protectant. In *Saccharomyces cerevisiae* a link was found between the induction of genes responsible for trehalose synthesis and stress conditions such as heat, dehydration and radiation (Parrou *et*

*al.*, 1997). Synthesis of trehalose in bacteria and yeasts occur predominantly via the trehalose synthase complex (trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatases) and the hydrolytic cleavage of trehalose occurs via acid and neutral trehalases (Figure 1.8) (Wannet *et al.*, 1999).

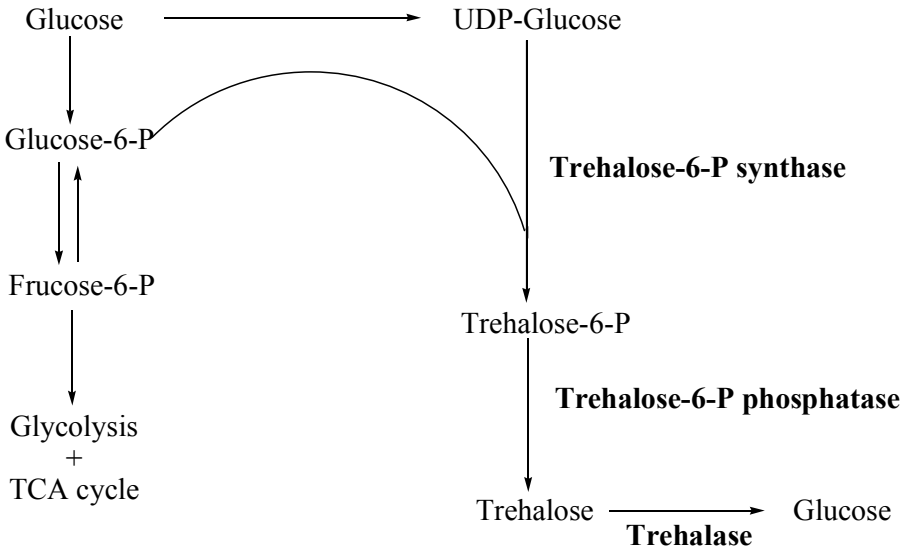



**Figure 1.8 Scheme representing a typical trehalose synthesis and degradation of fungi (Wannet *et al.*, 1999)**

*1.3.1.2. Nitrogen Metabolism*

Nitrogen is a major component of nearly all of the complex macromolecules central to the structure and function of all living organisms. Fungi can use a surprisingly diverse array of compounds as nitrogen sources and are capable of expressing upon demand the catabolic enzymes of many different pathways. Nitrogen metabolism and its regulation have been extensively studied in a large variety of organisms. However basidiomycetes have received relatively little attention (Marzluf, 1997).

Fungi assimilate simple nitrogenous sources for example ammonium ions for the biosynthesis of amino acids and proteins. Ammonium ions are readily translocated and can be directly assimilated into the amino acids glutamate and glutamine by glutamate

dehydrogenase (GDH) and glutamine synthetase (GS), respectively. Transfer of one amino acid from glutamine to 2-oxoglutarate yielding two molecules of glutamate is catalysed by glutamate synthase (glutamine:2-oxoglutarate aminotransferase, GOGAT) (Figure 1.9). *Agaricus bisporus* has two distinct glutamate dehydrogenases using NADH (NAD-GDH) or NADPH (NADP-GDH) as a cofactor (Moore and Al-Gharawi, 1976; Baars *et al.*, 1995). It should be noted that glutamate is also an intermediate in the 4-aminobutyrate shunt (section 1.3.1.1).

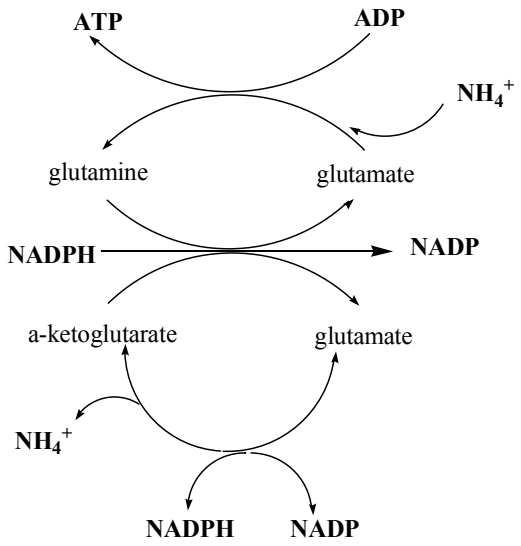



**Figure 1.9 Possible pathways of ammonia incorporation: (1) glutamate dehydrogenase (2) glutamine synthetase (3) glutamine: (GOGAT)**

*1.3.1.2.1. Urea Metabolism & The Ornithine Cycle (urea cycle)*

Higher fungi including cultivated mushrooms accumulate substantial amounts of urea in their fruit bodies (Hammond, 1979). Despite its abundance very little is known about its physiological role. Urea is chemically inert and highly soluble and therefore may serve as an osmotically favourable form of fungal nitrogen reserve. Accumulation of urea facilitates the translocation of water and metabolites in fruit bodies (Donker and Van As, 1999) which is needed for the production of spores.

Two major mechanisms for urea formation can be envisaged: the ornithine cycle and nucleic acid degradation. The ornithine cycle was demonstrated to be the major route for urea synthesis in the fruit body of *A. bisporus* (Reinbothe *et al.*, 1967).

Arginase is an ornithine cycle enzyme that catalyses the hydrolysis of arginine to urea and ornithine, fulfilling a prominent role in nitrogen metabolism of many organisms. Its activity controls the cellular levels of arginine and ornithine which are needed for essential metabolic processes such as protein synthesis and production of polyamines and prolines. In addition the ornithine cycle (Figure 1.10) enables the organism to detoxify $NH_4^+$ and to excrete nitrogen from the system (Wagemaker *et al.*, 2005).

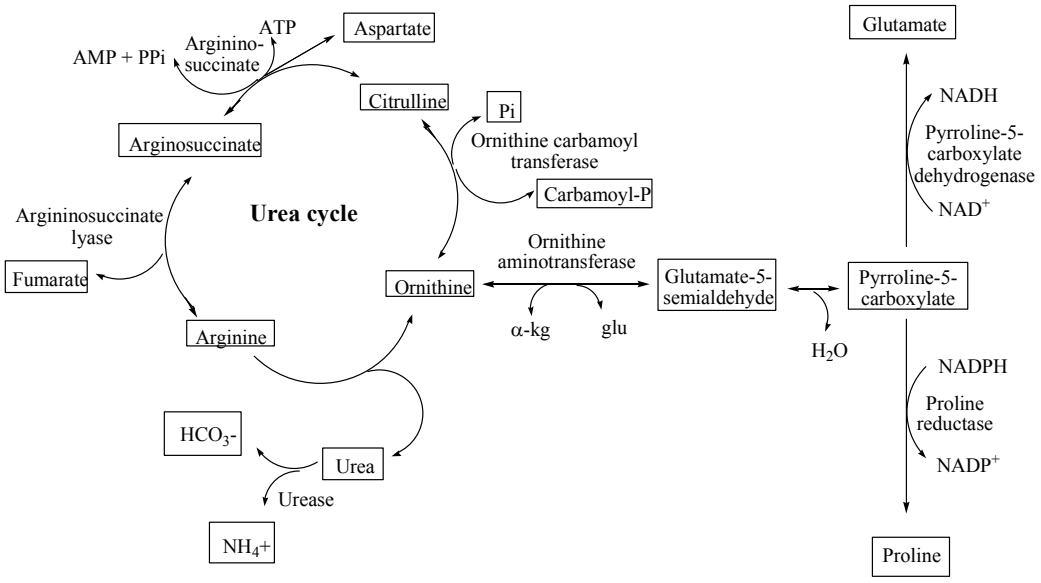



**Figure 1.10 Ornithine/urea cycle with relevant biochemical pathways (Wagemaker *et al.*, 2007)**

### *1.3.2. Defining Metabolomics*

Metabolomics is one of the more recently introduced 'omnics' technologies, joining genomics, transcriptomics and proteomics as tools in global systems biology (Sumner *et al.*, 2003). 'Omnics' technologies are based on the comprehensive biochemical and molecular characterisation of an organism, tissue or cell type. Metabolomics is a

technology that aims to identify and quantify the metabolome. It has been defined as the total quantitative collection of small molecular weight compounds (metabolites) present in a cell or organism which participate in metabolic reactions required for growth, maintenance and normal function (Oliver *et al.*, 1998; Harrigan and Goodacre, 2003). Generally, these include organic species such as amino and fatty acids, carbohydrates, vitamins and lipids, although inorganic and elemental species can also be studied (Lahner *et al.*, 2003).

In general metabolomic analyses have been classified as targeted or untargeted (Figure 1.11). Targeted analyses focus on a specific group of intended metabolites with most cases requiring identification and quantification of as many metabolites as possible within the group (Ramautar *et al.*, 2006). In comparison untargeted metabolomics focuses on the detection of as many groups of metabolites as possible to obtain patterns or fingerprints without necessarily identifying nor quantifying a specific compound/compounds (Monton and Soga, 2007).

Metabolomic studies may also be classified as discriminative, informative and/or predictive depending on the objective of the analysis (Figure 1.11).

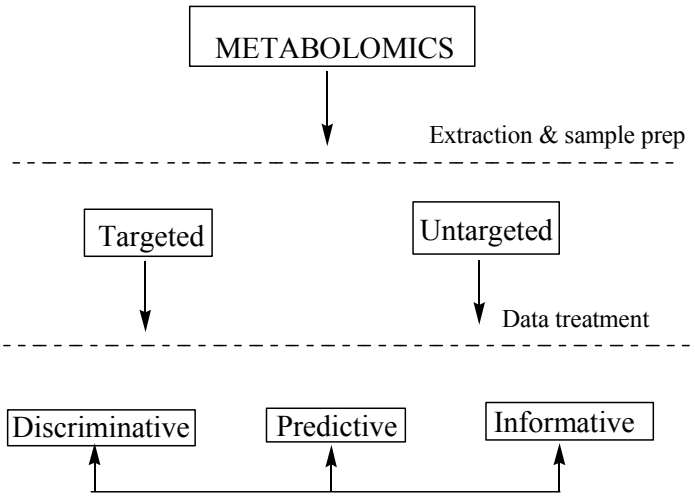



**Figure 1.11 General classification of metabolomics (Cevallos-Cevallos *et al.*, 2009)**

Discriminative analyses have been aimed at finding differences between sample populations without necessarily creating statistical models or evaluating possible pathways that may elucidate such differences. This type of metabolomics was used on wine in order to classify it by grape variety and production area (Son *et al.*, 2008). In contrast informative metabolomics have focused on the identification and quantification of targeted or untargeted metabolites to obtain sample intrinsic information. This type of metabolomics has been used in the development and continuous update of metabolite databases such as the human genome database (Wishart *et al.*, 2007).

Some metabolomic studies have been predictive. In this instance, statistical models based on metabolite profile and abundance were created to predict a variable that was difficult to quantify by other means. These types of models have been created for prediction of green tea sensory quality (Ikeda *et al.*, 2007).

Metabolomic analyses consist of a sequence of steps including sample preparation, metabolite extraction, derivatisation, metabolite separation, detection and data treatment (chemometrics) (Figure 1.12). In some cases not every step is required. Only detection and data analysis have been essential steps in all reported metabolomic studies. The selection of the steps depends on the type of instrumentation to be used for separation (e.g. gas chromatography (GC) *vs* liquid chromatography (LC)), the detection method (e.g. mass spectrometry (MS) *vs* nuclear magnetic resonance (NMR) spectroscopy) and the kind of sample being analysed for example whether it is liquid or solid (Cevallos-Cevallos *et al.*, 2009).

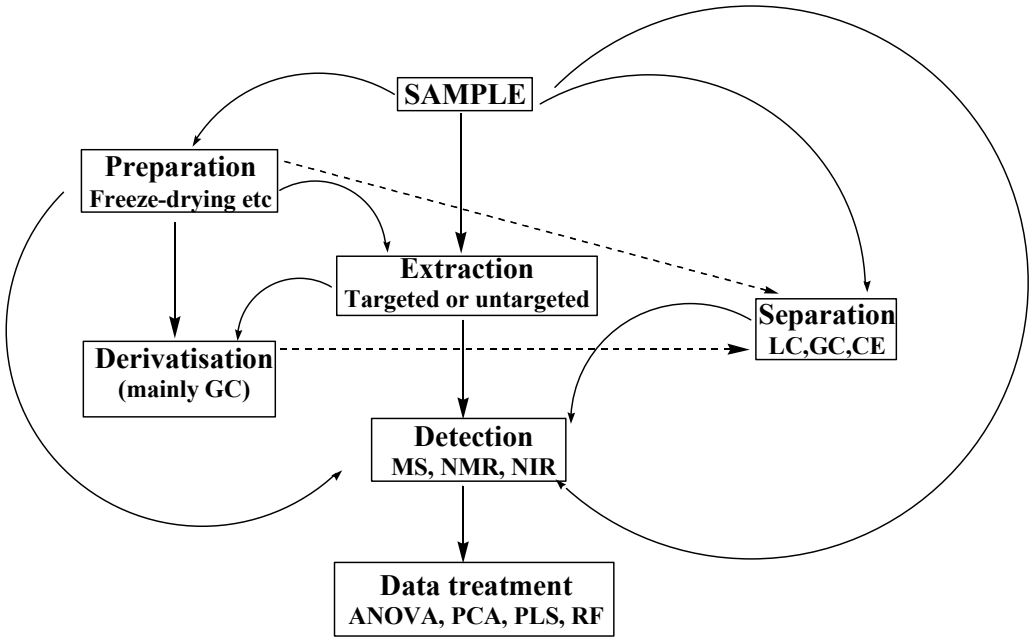



**Figure 1.12 Schematic representation of the process of metabolomic analysis (Cevallos-Cevallos *et al.*, 2009)**

### *1.3.3. Development of Metabolomics*

Metabolomics originated from metabolite profiling, which has been a part of medical practice for thousands of years. As far back as the fifth century BC, Hippocrates and Hermongenes both described the diagnosis and detection of diseases through sensory analysis of urine i.e. colour, taste, smell. The analysis of biofluids became more quantitative with the development of clinical chemistry in the mid-19$^{th}$ century. However, it was not until the early 20$^{th}$ century that clinical chemistry and metabolic profiling became a part of routine medical practice with the development of colorimetric tests and early instrumentation used to quantify metabolites in blood and urine (Rosenfeld, 2002).

A new generation of analytical instrumentation appeared in the 1970s which permitted the identification of not just a single compound but a whole class of compounds. Gas chromatographic (GC) columns started to be coupled to mass spectrometers to create GC/MS systems, which could detect organic acids from blood and urine. The earliest metabolite profiling publications originated in the early 1970s from the Baylor College of

Medicine (Devaux *et al.*, 1971; Horning and Horning, 1971). These authors illustrated their concept through the multicomponent analyses of steroids, acids and neutral and acidic urinary metabolites using GC/MS. Following this type of research, the concept of using metabolite profiles to screen, diagnose and assess health began to spread.

At the turn of the century, multiple genome and Expressed Sequence Tag (EST) sequencing projects were underway, fuelling the 'genomics' era (Genome, 2000; Goff *et al.*, 2002; Yu *et al.*, 2002). These high throughput sequencing projects revealed a large number of predicted genes. The genes could not be assigned a function based on sequence information alone, which led to the proposal to assess gene function using large-scale analyses at the transcriptome level, initiating the 'functional genomics' era. Following on from this it became apparent that proteomics might be more insightful in terms of monitoring results of gene expression.

This way of thinking eventually led to consideration of the metabolome. It is believed that Oliver was the first to make the connection based on the perceived need for quantitative and qualitative measurement of phenotype to assess genetic function and redundancy in yeast (Oliver, 1997). His group estimated the number of yeast metabolites to be approximately 600 and proposed the concept of metabolomics. This approach was then pioneered for plants by researchers at the Max Planck Institute (Trethewey *et al.*, 1999; Fiehn *et al.*, 2000; Trethewey, 2001; Fernie *et al.*, 2004; Fiehn, 2008).

Plant-science papers still form the majority of published papers on GC/MS metabolite profiling compared to biomedical research or microbiology. Metabolite profiling in plants is regarded as a standard tool in plant research and is routinely applied in many laboratories. Applications range from environment studies, genetic studies of complex traits to agricultural and food-quality investigations for example the substantial equivalence of genetically modified food to classic bred cultivars (Fiehn, 2008).

### *1.3.4. Metabolomic Limitations*

The major limitation of metabolomics is its current inability to comprehensively profile all of the metabolome. This failure is directly related to the chemical complexity of the metabolome, the biological variance inherent in most living organisms and the limitations of most instrumental approaches. The chemical components of metabolites range from ionic inorganic species to hydrophilic carbohydrates, hydrophobic lipids and complex natural products. This chemical diversity and complexity makes it extremely difficult to profile all of the metabolome simultaneously. No single analytical method has the ability to profile all of the metabolome. Despite this several developments aim to analyse many metabolites, even of different chemical classes by a single method. The use of selective extraction and parallel analyses using comprehensive visualisation of the metabolome is being employed (Sumner *et al.*, 2003).

Variability in metabolomics can present itself in the form of (a) analytical variance and (b) biological variance.

(a) Analytical variance can be defined as the coefficient of variance or relative standard deviation that is directly related to the experimental approach. The variance will differ depending on the analytical platform being used and its indeterminate origin.

(b) Biological variance arises from quantitative variations in metabolite levels between for example mushrooms of the same species grown under identical conditions and is indeterminate in origin. Biological variations typically exceed analytical variations (Roessner *et al.*, 2000; Sumner *et al.*, 2003).

Ways to minimise biological variance include pooling samples either by analysing different tissues of the plant within a single sample or by pooling multiple replicate plants. This helps to minimise random variations through statistical averaging. However, many

variations in metabolite levels often have biological significance and result from functional differentiation of tissues. Another option would be to start with homogenous tissue such as cell cultures. There is a need to incorporate strategies to minimise variability. Parameters including the growth stage, environmental conditions and in particular sampling are important factors that can also reduce variability (Sumner *et al.*, 2003).

Sampling is one of the most underestimated parts of metabolomic analyses. The composition and quantity of metabolites detected depend to a large extent on the sample preparation chosen. The time and method of sampling can greatly influence the reproducibility of the analytical sample. The storage of samples is also important and needs to be considered as the continued freeze/thawing of samples can be detrimental to stability and composition (Roessner *et al.*, 2000).

Another challenge encountered in metabolomics is dynamic range. This is when the presence of excessive amounts of some metabolites can cause significant or severe chemical interferences that limit the range in which other metabolites may be successfully profiled. For example high levels of primary metabolites such as sugars often interfere with the ability to profile secondary metabolites such as flavonoids. A positive aspect of this is that highly expressed metabolites are often unique and can provide exclusive bases for the differentiation of cell states, organs, tissues, varieties and organisms. They are often referred to as biomarkers. A number of different analytical approaches have been developed to improve dynamic range and to reduce complications (Sumner *et al.*, 2003).

### *1.3.5. Metabolomic Technologies*

Metabolomic approaches generally try to measure metabolite profiles in a non-targeted way. This means they try to separate and detect as many metabolites as possible in a single analysis and therefore the analytical technique must be suitable for a diverse range of small

endogenous metabolites in various concentrations. It must also separate compounds with different physical and chemical properties and be reproducible (Ramautar *et al.*, 2006).

The field of metabolomics requires profiling and fingerprinting methods. Metabolic profiling uses hyphenated techniques such as gas chromatography-mass spectrometry (GC/MS), liquid chromatography-mass spectrometry (LC/MS) or capillary electrophoresis-mass spectrometry (CE/MS). These techniques provide a detailed chromatographic profile of the sample and consequently measurements of the relative or absolute amounts of the components.

Nuclear magnetic resonance (NMR) spectroscopy, Raman spectroscopy and Fourier-transform infrared spectroscopy are referred to as 'fingerprinting' methods. They are more rapid, general screening methods that can be configured as 'high-throughput' and are suitable for determining differences and classifying samples (Halket *et al.*, 2005).

*1.3.5.1. Gas Chromatography-Mass Spectrometry*

Gas Chromatography-Mass Spectrometry (GC/MS) is one of the most widely used analytical techniques in metabolomics and has had a fairly long history in metabolic profiling (Horning and Horning, 1971). It combines the high separation efficiency and resolution of capillary GC that is essential for complex metabolic profiling with the high sensitivity of mass-selection detection. It is used to analyse qualitatively and quantitatively a wide range of volatile and/or derivatised non-volatile metabolites with high analytical reproducibility, although it is biased against non-volatile high molecular (MW) metabolites (Bedair and Sumner, 2008).

The majority of metabolites analysed require chemical derivatisation at room temperature or elevated temperatures to provide volatility and thermal stability prior to analysis. The most commonly utilised derivatising procedure for GC/MS metabolite profiling includes a two-step derivatisation scheme (Roessner *et al.*, 2000). The first step involves the use of

alkoxyamines to convert carbonyl groups to oximes in order to stabilise the reducing sugars and the second step involves replacing the active hydrogen in polar functional groups such as carboxylic acids, alcohols and amines with a trimethylsilyl group (typically *N*-methyl-*N*-(trimethylsilyl)trifluoracetamide) (MSTFA). Derivatisation increases the thermal stability and volatility of a broad range of metabolites (Bedair and Sumner, 2008). As a result of this limitation of GC/MS, most polyphosphorylated and activated intermediates are presently not accessible to GC/MS analyses (Kopka *et al.*, 2004).

Metabolite identification is provided by matching the retention time or retention index and mass spectrum of the sample peak with those of a pure compound previously analysed under the same experimental conditions. Many commercial and public domain databases (e.g. NIST, Golm metabolome consortium, NIH) exist to help metabolite identification. However, the available mass spectral databases do not contain all metabolites that would be expected from studying metabolic reaction networks. Many efforts are being made to create metabolomics-specific mass spectral libraries. Structural identification can also be performed via interpretation of fragment ions and fragmentation patterns (Dunn and Ellis, 2005).

In recent times applications in plant metabolomics are becoming widespread. Plants including potatoes (Roessner *et al.*, 2000) and tomatoes (Roessner-Tunali *et al.*, 2003) have been studied to measure effects of genetic or environmental modifications and stressors.

*1.3.5.2. Liquid Chromatography-Mass Spectrometry*

Liquid Chromatography-Mass Spectrometry (LC/MS) is another combined system that provides metabolite separation by LC followed by electrospray ionisation (ESI) or less typically atmospheric pressure chemical ionisation (APCI). This technique is simplified compared to GC/MS as lower analysis temperatures are used and sample volatility or

thermal stability are not required (Dunn and Ellis, 2005). It is a more universal separation technique for the targeted analysis of specific metabolite groups or utilised in a broader non-targeted manner. Additionally LC/MS also has the benefit of analyte recovery by fraction collection and/or concentration, which is more challenging when using GC separations (Bedair and Sumner, 2008).

The identification of metabolites in LC/MS is achieved through accurate mass determination, tandem MS analysis and/or coupling to nuclear magnetic resonance (NMR). However, a major disadvantage of LC/MS relative to GC/MS in metabolomic profiling is the lack of transferrable LC/MS libraries for metabolite identification. The mass spectral variability between LC/MS systems in terms of relative ion abundances, in-source fragmentation, tandem mass spectra fragment ions and lack of LC retention indices that compensate for instrument and experiment variability hinders comparison of LC/MS data between laboratories (Halket *et al.*, 2005). Currently there are a few limited mass spectral libraries available though none the size of GC/MS libraries.

A recent advancement in LC metabolic profiling has been the introduction of the commercially available ultra-high pressure liquid chromatography (UHPLC) systems, which operate at relatively higher pressures and use more tightly packed columns, which greatly enhance chromatographic resolution and efficiency (Wilson *et al.*, 2005). This provides enhanced opportunities for resolving complex biological mixtures in non-targeted metabolite profiling.

Applications of LC/MS are mainly focused on clinical applications in the discovery of biomarkers for a number of diseases. Applications in plant and microbial metabolomics are small in number. One example from the plant area is the determination of apple polyphenols and glucosides (Alonso-Salces *et al.*, 2004).

*1.3.5.3. Capillary Electrophoresis-Mass Spectrometry*

Capillary Electrophoresis (CE) is an analytical separation technique capable of high-resolution separation of a diverse range of metabolites and is particularly suitable for the separation of polar and charged compounds (Soga and Imaizumi, 2001). With respect to separation and efficiency it is a more powerful technique compared to LC due to the plug-flow profile generated by the electroosmotic flow (EOF) as compared to the parabolic flow in LC.

Capillary zone electrophoresis (CZE) has been the major CE technique used for CE/MS analysis of metabolites, due to the simplicity of the running buffer and the lack of surfactant or other additives necessary in other modes of separation. The charged molecules are separated in CZE based on their differential mobility, whilst neutral molecules migrate through the capillary using the electroosmotic flow without separation (Bedair and Sumner, 2008). The main advantage of this technique is the ability to separate almost any charged metabolite with very high resolution without prior derivatisation. It has been employed to identify primary metabolites in rice (Sato *et al.*, 2004) and over 1,600 compounds in *Bacillus subtilus* (Soga *et al.*, 2003).

The major drawback of CE is poor concentration sensitivity due to the limited amount of sample volume that can be introduced into the capillary and the low absorption path-length if UV detection is used (Ramautar *et al.*, 2006).

*1.3.5.4. Emerging Mass Spectrometry Technologies*

Soft ionisation techniques involving lasers to provide sample volatilisation and ionisation are currently being developed with the potential to provide rapid, high through-put global analyses (Dunn and Ellis, 2005). A review of such techniques has been provided by Bedair and Sumner (Bedair and Sumner, 2008).

The application of orthogonal multi-dimensional separations will impact the sensitivity and number of metabolites detected in the future through improved chromatographic resolution and increased signal to noise (S/N) ratios. Great potential is offered by the high throughput capabilities of GCxGC/TOF/MS. These instruments are designed for high scanning speeds (up to 500 scans $s^{-1}$) or high mass resolution. Higher scanning speeds are currently being employed for metabolic profiling with mass spectral deconvolution and have great potential to solve the high through-put problem in GC. The application of HPLC-HPLC-MS and HPLC-CE-MS (Evans and Jorgenson, 2004) is also in practice. These methods together with complex informatics technologies are employed to improve resolution of highly complex biological samples.

*1.3.5.5. Nuclear Magnetic Resonance Spectroscopy*

Nuclear Magnetic Resonance (NMR) Spectroscopy is a technique that takes advantage of the spin properties of the nucleus of atoms. $^{1}H$ is the most used nuclei for NMR measurements because of its very high natural abundance (99.98-99.97%) (de Laeter *et al.*, 2003) and good NMR properties. It has been used extensively for metabolomic research over the last twenty years (Lindon *et al.*, 2001; Lindon *et al.*, 2004a; Lindon *et al.*, 2004b) and benefits from the fact that it is a specific but non-selective technique. Hence, each separate resonance observed in an NMR spectrum is specific to a particular compound which provides a wealth of structural information regarding the components of a sample. A combination of chemical shift (indicating the nature of the chemical environment in which a particular nucleus is located), spin-spin coupling (indicating the number and nature of nearby nuclei and thus connectivity information) and relaxation or diffusion (which gives an indication of the size of a molecule and also the large scale environment in which a molecule is located) all allow the rapid identification of any components regarded as interesting in the analysis. NMR analysis is non-destructive and does not require pre-

selection of the analysis conditions that is required for mass spectrometry or chromatographic operating conditions (stationary phase, mobile phase, temperature) (Dunn *et al.*, 2005). However, MS does provide significant improvements in sensitivity in comparison to NMR spectroscopy.

Different metabolomic approaches may be applied when using NMR spectroscopy (Ratcliffe and Shachar-Hill, 2005). One of the approaches is directly related to the usage of NMR structure elucidation. Generally the compounds of interest are isolated from their tissues and solubilised for the acquisition of one-dimensional $^{1}$H NMR and when required, additional (2D)-NMR spectra (e.g. COSY, TOCSY and NOESY).

The spectra are complex resulting in thousands of signals relating to metabolites. For data processing the spectrum is usually divided into groups of chemical shifts with widths of 0.02-0.04 ppm. The chemical shifts can be assigned to specific metabolites and pure metabolites can be added for further clarification. Another alternative is the use of the spectrum pattern in classification of samples similar to other metabolic fingerprinting techniques (i.e. FTIR or Raman spectroscopy) (Choi *et al.*, 2004; Pereira *et al.*, 2006).

The technique is used in clinical and pharmaceutical applications for the analysis of biofluids or tissues, especially with $^{1}$H NMR spectroscopy. Studies are based on cells response to stress, including disease or therapeutic interventions by adjustment of their intra and extra-cellular environments to ensure homeostasis (Dunn and Ellis, 2005). NMR spectroscopy has also been employed in other fields for the analysis of plant-cell extracts such as *Arabidopsis* (Ward *et al.*, 2003) and tobacco (Choi *et al.*, 2004) and to determine the mode of action of biochemicals (Aranìbar *et al.*, 2001).

*1.3.5.6. Vibrational Spectroscopy*

Techniques such as Infrared (IR) and Raman spectroscopy are valuable analytical techniques. Although these two vibrational spectroscopic techniques are not as sensitive as

mass spectrometric methods and do not allow the collection of a list of metabolite identities in complex samples, they do provide a relevant metabolomics tool. Fourier transform infrared (FTIR) spectroscopy is the more widely used method to obtain spectral fingerprints of biological samples representing a snap-shot of the biochemistry at a given time. The two methods have the benefit of allowing the rapid (particularly with respect to FTIR), reagentless, non-destructive analysis of complex biological samples therefore facilitating high throughput screening and providing unbiased measurements of the whole system (Kell, 2004).

Optical spectroscopy mainly measures the vibrations and rotations of molecular functional groups that result from the energy exchange when radiation interacts with a sample. This interaction results in an increase of molecular energy which can produce three different transitions; electronic excitation, vibrational change and rotational change. IR spectroscopy utilises the IR region of the spectrum (12,000 $cm^{-1}$ to 10 $cm^{-1}$) and is divided into three sub-regions, near-IR (NIR), mid-IR and far-IR. The boundaries between these are not clearly defined but MIR is usually considered to range from 4,000 to 400 $cm^{-1}$, with NIR being at wavenumbers above 4,000 $cm^{-1}$ and far-IR at wavenumbers below 400 $cm^{-1}$ and into the microwave region. Raman spectroscopy utilises a monochromatic beam usually having a wavelength within the visible or UV regions of the spectrum. Both Raman and IR spectroscopy give information about molecular vibrations (Dunn and Ellis, 2005).

*1.3.5.6.1. Infrared Spectroscopy*

IR spectra are typically shown as percentage transmittance plotted against wavenumber. Absorbance is favoured over transmittance as the absorbance is proportional to concentration at a given wavelength (Beer's Law). An IR spectrum consists of many bands originating from the vibrational motion within the molecule due to the absorption of incident radiation. Bands due to rotational motion are absent from the spectra of biological

samples as the samples tend to be in the condensed form as solids, liquids or solutions, so only vibrational motion is observed. The features of the spectra, the number of bands, frequency, intensity and half-widths are characteristic so giving a fingerprint unique for the sample (Nelson, 1991).

In general spectral fingerprints are collected spanning either the MIR or the NIR regions. Despite the close proximity of these regions different attributes are observed. MIR absorption arises from fundamental molecular vibrations providing data containing chemical and structural information about the sample, which is amendable for direct interpretation (Griffiths and de Haseth, 2007). In contrast to this NIR absorption arises from overtones and combination-band absorption characteristics of CH, NH and OH groups, giving spectra containing broad overlapping features which are not directly interpretable at a chemical level (Belton *et al.*, 1987). Metabolomic applications of IR spectroscopy currently favour the use of the MIR region as this provides greater chemical and structural information about a sample (Johnson *et al.*, 2003).

The primary applications of FTIR spectroscopy to study complex biological systems are in the field of microbiology, typically with respect to biomedical and industrial applications (Nelson, 1991). Its potential as a diagnostic tool has been recognised for the identification of possible biomarkers for certain diseases. It has also been used as a diagnostic tool for quality assurance within the food industry.

*1.3.5.6.2. Raman Spectroscopy*

Raman spectroscopy operates by the detection of scattered energy after irradiation of a sample by monochromatic visible or UV radiation. The majority of the scattered energy has the same frequency as the incident radiation and is termed Rayleigh scattering. A small proportion of the scattered radiation consists of discrete frequencies above and below that of the incident radiation and this is called Raman scattering which is the result of an

inelastic collision in which there is the exchange of energy between the photon and the molecule (Everall *et al.*, 2000).

For a molecule to be Raman active i.e. that it is susceptible to Raman scattering, there needs to be a change in the molecular polarisability caused by internal vibration. An incident electric field induces an electric dipole moment which is a separation of charge within the molecule and under these conditions the molecule is said to be polarised. Electrons within the molecule are more easily displaced along a specific axis producing a polarizability ellipsoid. Raman scattering is a measure of the changes in the magnitude or direction of this ellipsoid. For IR the molecular vibration must produce a change in the electric dipole of the molecule (Banwell, 1983).

During the 1980s Raman spectroscopy was overlooked in the field of biological sciences. However over the past decade there has been an increasing number of publications demonstrating its potential use for the identification and characterisation of microorganisms. The use of Raman spectroscopy for the study of complex biological systems outside the field of microbiology is still very much in its infancy. Studies have demonstrated its potential in the biochemical analysis of honey (De Oliveira *et al.*, 2002) and for the analysis of plant pigments and essential oils (Schrader *et al.*, 2000).

## 1.4. CHEMOMETRICS

### *1.4.1. Defining Chemometrics*

The term chemometrics was introduced by Svante Wold and Bruce R. Kowalski in the early 1970s. Since then chemometrics has been developing and is now widely applied to different fields of chemistry, especially analytical chemistry (Wold, 1995). Chemometrics is a discipline using mathematical and statistical methods for the selection of the optimal experimental procedure and data treatment for data analysis (Massart *et al.*, 1997).

Metabolomic analysis generates large and complex datasets and because of this chemometrics has become an important part of metabolic profiling and fingerprinting due to its ability to provide interpretable models for complex inter-correlated data. It provides tools to make good use of measured data, enabling practitioners to make good use of measurements and to model quantitatively and produce visual representations of information. The use of multivariate statistical methods such as principal component analysis (PCA), partial least squares (PLS) and random forests (RF) is of great importance as these include efficient, validated and robust methods for modelling information-rich chemical and biological data. Therefore chemometrics can be defined as the tools used to extract the information of complex biological and chemical systems.

Chemometrics has grown into a well-established data analysis tool in areas such as multivariate calibration, quantitative structure-activity modelling, pattern recognition and multivariate statistical process monitoring and control. Although seemingly diverse the commonality in these areas is that high complexity data tables are generated and that these can be analysed and interpreted by chemometric methods (Trygg and Lundstedt, 2007).

### *1.4.2. Chemometric Tools*

#### *1.4.2.1. Principal Component Analysis*

Principal Component Analysis (PCA) is one of the oldest and most widely used multivariate techniques (Hotellin, 1933). The concept behind PCA is to describe the variance in a set of multivariate data in terms of a set of underlying orthogonal variables (principal components) (Sumner *et al.*, 2003). The definition of principal component analysis therefore, is the analysis of data that has been transformed from the original axes to the principal axes.

It is a useful technique to reduce the dimensionality of large data sets and is described as the powerhouse of chemometric tools. PCA is also useful for identifying significant signals

in noisy data. The mathematical technique used in PCA is called eigen analysis. The eigenvalues and eigenvectors of a square symmetric matrix with sums of squares and cross products can be solved from data matrix obtained from metabolite analysis. In many cases the data matrix for PCA should be prepared from data obtained by GC/MS, LC/MS etc. The eigenvector associated with the largest eigenvalue has the same direction as the first principal component. The eigenvector associated with the second largest eigenvalue determines the direction of the second principal component. The sum of the eigenvalues equals the trace of the square matrix and the maximum number of eigenvectors equals the number of rows (or columns) of this matrix. PCA has the ability to identify and indicate useful information from the metabolome using a few principal components. In fact the application of PCA to a metabolome data set provides two quantities: the score (related to the eigenvalue) and the loading (related to the eigenvector).

The loading allows the evaluation of the contribution that each metabolite makes to the information associated to a particular principal component. The loading is useful for understanding differences among samples in each metabolite level. The PCA score is defined as the coordinate of data vectors in the base of the principal component analysis. The score plot, limited to the significant principal components, gives a visual image of the differences of samples from an all around view point. The first principal axis is the direction in which the data are primarily distributed in n-dimensional space (Fukusaki and Kobayashi, 2005). Therefore PCA can indicate relationships among groups of variables/metabolites in a data set and show relationships that might exist between objects (Shin *et al.*, 2010).

The capability of PCA to manage and interpret large data sets has seen it being effectively employed to define relationships that exist for example in fatty acid characterisation

studies of food lipids (Kadegowda *et al.*, 2008). Applications of metabolomics and PCA are highlighted in section 1.5 (Table 1.1)

*1.4.2.2. Partial Least Squares*

Partial Least Squares (PLS) regression or discrimination is a wide class of methods for modelling relations between sets of observed variables by means of latent variables. It consists of regression and classification tasks as well as dimension reduction techniques and modelling tools. The underlying assumption of the PLS method is that the observed data is generated by a system or process which are driven by a small number of latent variables. Projections of the observed data to its latent structure by means of PLS was developed by Wold and co-workers (Wold, 1995).

PLS has received a lot of attention in the field of chemometrics. The algorithm has become a standard tool for processing a wide spectrum of chemical data problems. The success of PLS in chemometrics resulted in many applications in other scientific areas including the field of metabolomics (Yamamoto *et al.*, 2009; Kim *et al.*, 2010).

In its general form PLS creates orthogonal score vectors (also referred to as latent vectors or components) by maximising the covariance between different sets of variables. There are different PLS techniques to extract these latent vectors, with each one giving rise to a variant of PLS methods (Wangen and Kowalski, 1989; Westerhuis *et al.*, 1998).

PLS can be naturally extended to regression problems by linking the latent variables to the dependent variable through a Gaussian linear model. The predictor and predicted (response) variables are each considered as a block of variables. PLS then extracts the score vectors which serve as a new predictor representation and regresses the response variables on these new predictors. The natural asymmetry between predictor and response variables is reflected in the way in which score plots are computed. This variant is known

under the names of PLS1 (one response variable) and PLS2 (at least two response variables) (Helland, 2001).

In the same way PLS can be applied to classification problems by the correct linkage of the response and is referred to as partial least squares discriminant analysis (PLS-DA) (Barker and Rayens, 2003) which is more commonly used in metabolomics studies. It is particularly suited to dealing with problems where numbers of predictors are large and collinear. The reduced dimensions from PCA and PLS can facilitate the visualisation of high dimensional spectral data (Davis *et al.*, 2006).

*1.4.2.3. Random Forests*

Random Forests (RF) (similar to Support Vector Machine) belong to a family of statistical methods associated with analysis of large datasets in contrast to PLS and PCA which were originally designed with smaller data sets in mind. As such a random forest is a collection of decision trees which is created following a particularly efficient strategy aimed at increasing the diversity between the trees. Decision trees are unstable in nature i.e. they are unstable methods for which a small change in the dataset can result in large changes in the developed method (Breiman, 1996). To increase the diversity among the members of the ensemble, RF fits each tree on a bootstrap replicate of the full set of samples. A bootstrap replicate (Efron and Tibshirani, 1983) is a random subset of the available dataset of the same length taken with replacement (i.e. each sample is picked at random from the full original dataset irrespective of whether it has been picked before). Another source of diversity is introduced during the growing of each tree. For each node the method selects a small random subset of $m$ attributes (from the total $M$ attributes available) and uses only this subset to search for the best split. The combination of these two sources of diversity (i.e. bootstrapping plus selecting at each node only from a subset of attributes) produces easy-to-build ensemble models, where predictions on regression or classification are

performed through consensus (i.e. majority vote between all trees for classification and average prediction from all trees for regression). These features make random forests very efficient statistical methods for prediction and generalisation.

One of the most important properties of RF is the control of overfitting even when the ensemble contains thousands of individual trees. The error rate of a RF model on unseen samples shows a smooth convergence to a limiting value when the number of trees goes to infinity. The RF algorithm has only one free parameter in practice, the number $m$ of attributes made available at each node during the growing of trees. Breiman demonstrated that its results are not strongly dependent on this parameter and the default value of $m$ (the square root of the total number of attributes $M$) usually gives near optimal results, unlike PLS methods which are very sensitive to the number of latent variables used in the model (Breiman, 2001).

Each tree in a RF is typically grown on a bootstrap of the full dataset. On average 37% of the samples will not be present in a given bootstrap and these sets of unseen samples are referred to as out-of-bag (OOB) sets. These OOB sets are used in particular to give an unbiased estimate in the prediction error on unseen cases. This estimation can be used to give a measure of relevance of the attributes included in the RF model (once it has been grown) in the following way: one at a time, each attribute is shuffled (i.e. its values are randomly permuted between all samples in the dataset). An OOB estimation of the prediction error is made on this 'shuffled' dataset. Intuitively, an attribute that is irrelevant to the model will not change the prediction performance when altered in this way. However, if the model made strong use of a given attribute altering its values will lead to an important decrease in performance between the 'original' dataset and the 'shuffled' dataset which is therefore related to the relevance of the attribute affected by the process (Breiman, 2001).

A RF approach can be seen in the recent study by Zheng and colleagues (Zheng *et al.*, 2009). Other applications include the use of data mining tools in descriptive sensory analysis (Granitto *et al.*, 2007) and in the field of biological sciences e.g. RF has been used in microarray studies and other types of DNA and protein analyses (Shi *et al.*, 2005).

## 1.5. METABOLOMIC APPLICATIONS

Much work in the field of metabolomics has focused primarily on clinical or pharmaceutical applications such as drug discovery (Watkins and German, 2002; Kell, 2006), drug assessment (Lindon *et al.*, 2004a), clinical toxicology (Griffin and Bollard, 2004), clinical chemistry (Wishart, 2008a) and human diseases (Kaddurah-Daouk and Krishnan, 2009). Applications of metabolomics can be seen in many clinical areas, for example comparing metabolite profiles of urine of a healthy individual with that from a sick patient (Stenlund *et al.*, 2009), bacterial metabolomics (Ramautar *et al.*, 2006) and disease diagnostics (Madsen *et al.*, 2010).

Over the past few years metabolomics has emerged as a field of increasing interest to food scientists (Gibney *et al.*, 2005; García-Cañas *et al.*, 2010). The fact that metabolomics allows the simultaneous characterisation of large numbers of chemicals in biological matrices makes it an attractive tool to acquire a far more detailed and comprehensive molecular picture of food. Foods are now being analysed with more chemical detail leading to hundreds or even thousands of distinct chemical identities being detected and/or identified (Moco *et al.*, 2006; Ninonuevo *et al.*, 2006). Metabolomic applications within the food industry are diverse ranging from profiling of plant species, to discriminating between food spoilage bacteria, to studying the effects of stresses on plants and so forth.

### *1.5.1. Metabolomics in Food Quality*

Targeted metabolomics focused on volatiles has shown great potential to assess pre-harvest issues that affect quality. Pre-harvest fungal diseases in mango (Moalemiyan *et al.*, 2007), post-harvest bacterial contamination of onions (Vikram *et al.*, 2005) and McIntosh apples (Vikram *et al.*, 2004), as well as diseases of stored carrots (Vikram *et al.*, 2006) have been assessed by sampling headspace metabolites followed by GC/MS analysis. For each of the above examples the volatile profile was found to be disease-specific, with many compounds identified by GC/MS databases. In addition changes in polyphenolic compounds during berry breeding (Stewart *et al.*, 2007) have been characterised by informative metabolomics. Post-harvest metabolomic analysis has the potential for detecting and understanding food spoilage as reported in 2008 by Kushalappa and colleagues (Kushalappa *et al.*, 2008). Table 1.1 summarises some recent metabolomics studies used for food analysis.

Future trends will involve the use of discriminative and predictive metabolomics as the ultimate tool for quality control. The metabolite profile of products meeting minimum standards can be used as a baseline for quality acceptance. Individual samples can be analysed and compared to determine acceptability of the batch produced. Food adulteration may also be assessed in this way as adulteration can be detected by appearance of uncommon peaks in the samples metabolic profile. The use of predictive models coupled with profiling techniques provides a cost alternative to quality analyses. Predictive models have been developed to estimate sensory attributes of green tea (Ikeda *et al.*, 2007), watermelon (Tarachiwin *et al.*, 2008) and mushrooms (Cho *et al.*, 2007). The potential of metabolomic techniques to rapidly identify adulterated food and beverages as well as their potential to accurately monitor undetectable quality control issues suggests that

metabolomics could soon play a major role in many aspects of food quality assessment and quality control (Wishart, 2008a).

**Table 1.1 Some recent metabolomic applications in food analysis**

| Sample: Purpose of analysis | Separation-detection | Data treatment | Reference |
|---|---|---|---|
| Apples: light induced changes in peel | GC/MS<br>LC/MS | PCA | (Rudell *et al.*, 2008) |
| Berries: polyphenol composition | LC/MS<br>DIMS | Compound identification | (McDougall *et al.*, 2008) |
| Pine mushrooms: quality differentiation | NMR | PCA | (Cho *et al.*, 2007) |
| Potato: GM differentiation | GC/MS<br>DIMS | PCA | (Catchpole *et al.*, 2005) |
| Meat: quality/safety | EESI/MS | PCA | (Chen *et al.*, 2007) |
| Broccoli: variety differentiation | LC/UV/MS<br>DIMS<br>IMS | PCA<br>ANOVA | (Luthria *et al.*, 2008) |
| Tomato: variety differentiation | LC/TOF/MS<br>NMR | PCA | (Moco *et al.*, 2008) |
| Watermelon: quality evaluation | NMR | PLS-LDA | (Tarachiwin *et al.*, 2008) |

PCA: principal component analysis
ANOVA: Analysis of variance
PLS-LDA: Partial least squares-linear discriminant analysis
IMS: Ion mobility mass spectrometry
EESI: Extractive electrospray ionisation
DIMS: Direct injection mass spectrometry

*1.5.1.1. Metabolomics in Food Safety*

Metabolomics has the potential to assess safety of pre- and post- harvest technologies. Unintended effects of genetic modification of foods can be assessed by untargeted discriminative analyses. Metabolomics has been used to differentiate genetically modified (GM) foods from non-treated ones, an example of this was seen for potatoes (Catchpole *et al.*, 2005). Sample differentiation occurred based on the intended variations of fructans in GM samples. After removal of fructan derivatives from the model no discrimination was observed. This suggests that GM potatoes are similar in composition to original ones, concluding that no major unintended changes occurred after genetic modifications. Metabolomics could therefore be used to provide new information regarding new or controversial processing technologies such as irradiation.

There are a number of metabolomic approaches used for the detection of microbiologically spoiled or contaminated foods (Ellis and Goodacre, 2001). In a recent study Ammor and colleagues investigated the use of Fourier transform infrared spectroscopy in tandem with chemometrics to explore its potential as a rapid and accurate method for monitoring the spoilage of minced beef samples under different storage conditions. It showed that the comparison of FTIR spectra could highlight certain biochemical changes during meat spoilage (Ammor *et al.*, 2009).

*1.5.1.2. Metabolomics in Food Component Analysis*

Food component analysis traditionally involved identifying and classifying food components into broad categories such as carbohydrates, proteins, fats, vitamins, fibre, trace elements, solids and/or ash. With the introduction of metabolomics came the ability to analyse with considerably more chemical detail, allowing the identification of hundreds of distinct molecules being detected and/or identified in certain foods (Moco *et al.*, 2006; Ninonuevo *et al.*, 2006; Wishart *et al.*, 2007). The potential to "unravel" foods and beverages into their

chemical constituents offers food chemists a unique opportunity to understand the molecular details of what gives certain foods and drinks their unique taste, colour, texture or aroma. It also offers the nutritional scientist the opportunity to precisely identify the bioactive ingredients in foods and better understand their potentially beneficial (or harmful) consequences (Wishart, 2008b).

A large number of natural foods, spices and beverages have already been the subject of detailed metabolomic-based analysis including milk (Ninonuevo *et al.*, 2006), tomatoes and tomato juice (Moco *et al.*, 2006), and many others. These analyses used a combination of techniques such as NMR spectroscopy, GC/MS and LC/MS.

In the coming years it is likely that food consumption studies will become much more common with far more comprehensive metabolomic analyses being performed on many of the economically or pharmaceutically more important fruits (e.g. bananas, pomegranates, pineapples, blueberries), vegetables (e.g. avocadoes, corn, spinach, cauliflower), grains (e.g. wheat, barley, rye), meats (e.g. beef, chicken, fish), processed foods (e.g. cheese, yoghurt, vegetable oils), nuts (e.g. almonds, cashews) and the many nutraceutical foods or beverages (e.g. ginseng, garlic, coffee, green tea). Metabolomic-based food component studies will allow food scientists to more precisely follow the consequences of different preparation (frying *versus* baking; steaming *versus* boiling) and preservation (freezing, drying, smoking, refrigerating) processes on key food components/metabolites. This process will help in the breeding, selection or modification of better crops, the breeding and feeding of livestock/fishstock as well as the preparation of many processed foods (Wishart, 2008b).

*1.5.1.3. Metabolomics & Mushrooms (A. bisporus)*

A review of some metabolites that have been identified in mushrooms using different metabolomic techniques are outlined in Table 1.2.

**Table 1.2 A summary of metabolites that have been identified in mushrooms by metabolomic techniques**

| Metabolomic Technique | Metabolites identified | Analysis purpose | Reference |
|---|---|---|---|
| NMR | Diacylglycerophospholipids (DAGP)<br>Neutral acylglycerols<br>Ether lipids<br>Sphingolipids<br>Sterols | NMR lipid profile | (Bonzom *et al.*, 1999) |
| GC/MS | Fatty acids:<br>Caprylic acid<br>Capric acid<br>Lauric acid<br>Linoleic acid<br>Nonadecanoic acid<br>Tricosanoic acid<br>Oleic acid | Fatty acid composition | (Yilmaz *et al.*, 2006) |
| HPLC | Fructose<br>Glucose<br>Myo-inositol<br>Mannose<br>Ribose<br>Sucrose<br>Trehalose<br>Arabinose | Non-volatile taste components | (Chiang *et al.*, 2006) |
| HPLC | Free amino acids:<br>Alanine<br>Leucine<br>Serine<br>Tyrosine<br>Valine<br>Aspartic acid<br>Lysine | Non-volatile taste components | (Tsai *et al.*, 2007) |

# 2. AIMS AND OBJECTIVES

## Aims and Objectives

The work presented in this thesis aims to investigate the effects that mechanical damage have on the metabolism of mushrooms using novel metabolic fingerprinting and metabolic profiling tools. Those tools will be employed to detect and discriminate mechanical damage on mushrooms.

To achieve this aim, the following objectives were set:

- Experiments studying the metabolic effect of mechanical damage in mushrooms were set up, studying both the immediate response to damage as well as the metabolic changes after storage of mechanically damaged mushrooms.
- Fourier transform infrared (FTIR) spectroscopy and chemometric methods were used to investigate whether the chemical changes induced by mechanical damage and ageing could be (a) detected in the mid-infrared absorption region and (b) identified using chemometric data analysis.
- Nuclear magnetic resonance (NMR) spectroscopy and chemometric methods were used to determine if low levels of damage could be differentiated, in order to evaluate the potential of this technology to detect damage.
- Gas chromatography-mass spectrometry was used to profile metabolites in damaged and undamaged mushrooms. GC/MS coupled with chemometric data analysis was used to detect metabolic markers of damage and to assess the metabolic processes that occur.

# 3. THE USE OF FOURIER TRANSFROM INFRA-RED SPECTROSCOPY AND CHEMOMETRIC DATA ANALYSIS TO EVALUATE DAMAGE AND AGE OF MUSHROOMS (*AGARICUS BISPORUS)*

# 3. The use of fourier transform infra-red spectroscopy and chemometric data analysis to evaluate damage and age of mushrooms (*Agaricus bisporus*)

The aim of this study was to investigate whether the chemical changes induced by mechanical damage and ageing of mushrooms can be (a) detected in the mid-infrared absorption region and (b) identified using chemometric data analysis. Further, the ability to develop a rapid tool that could detect physical damage and age before browning becomes visible would be of importance to the mushroom industry and could reduce economic losses.

## 3.1. Material and Methods

### *3.1.1. Mushrooms*

Second flush mushrooms were grown at the Teagasc Research Centre Kinsealy (Dublin, Ireland) and harvested damage-free. A set of 160 ($n$ = 160) closed cap, defect-free *A. bisporus* strain Sylvan A15 (Sylvan Spawn Ltd., Peterborough, United Kingdom) mushrooms (3-5 cm cap diameter) were selected for this study and immediately transported by road to the testing laboratory. Special trays were designed to hold mushrooms by the stem using a metal grid to avoid contact between (a) the mushrooms, (b) the tops of mushroom caps and (c) the tray lid during transportation. Mushrooms arrived at the laboratory premises within 1 hour after harvesting and were either damaged for the specified time length or remained damage-free and then stored at 4 ºC until required for analysis.

### *3.1.2. Mushroom Treatments*

Mushrooms ($n$ = 160) were harvested in the conventional manner on a single occasion. On the day of harvest a random subset ($n$ = 80) was subjected to physical damage using a mechanical shaker (Gyrotory G2, New Brunswick Scientific Co., USA) set at 300 rpm (rotations per minute) for 20 min giving these mushrooms an L-value of 86 which according to Gormley and O'Sullivan places them in a category of good quality and acceptability at the beginning of the storage experiment (Gormley and O'Sullivan, 1975). These samples were labelled as damaged (D). The remaining 80 mushrooms were untreated and labelled undamaged (UD). 10 damaged and 10 undamaged mushrooms were selected at random from their respective subsets on the day of harvesting and prepared for spectroscopic analysis (see below). These are referred to as day 0 samples. The remainder of the mushrooms (70 each of damaged and undamaged) were placed in plastic punnets and stored as separate batches at 4 ºC in a controlled temperature facility (Labcold sparkfree refrigerator, UK). On each of the seven consecutive days of such storage, a set of 10 damaged and 10 undamaged mushrooms were randomly selected, removed from storage and prepared for FTIR spectroscopic analysis.

### *3.1.3. FTIR Spectroscopy*

Sample preparation involved the manual dissection of each mushroom into its three main tissue types (cap, gills and stipe) before freezing overnight at -70ºC in a cryogenic refrigerator (Polar 340V, Angelantoni Industrie spA, Massa Martana, Italy) followed by freeze-drying (Micromodulyo, EC Apparatus Inc, New York, USA) for 24 h. Freeze-dried samples were manually ground into fine particles using a pestle and mortar. Then 9 mg (3% w/w) of each sample was mixed with 291 mg (97%w/w) of KBr (Sigma Aldrich, Dublin, Ireland). KBr pellets were prepared by exerting pressure of 100 kg/cm² (1200 psi) for approximately 2 min in a pellet press (Specac, United Kingdom). To eliminate any

interference which might be caused by variation in pellet thickness different pellets were prepared from the same sample and their infrared spectra were compared. These samples were identical with their average spectra used for analysis (Garip *et al.*, 2009).

Spectra were collected using a Nicolet Avatar 360 FTIR E.S.P (Thermo Scientific, Waltham, MA, USA) over the frequency range 4000-400 $cm^{-1}$. One hundred scans of each pellet were collected at 4 $cm^{-1}$ resolution at room temperature using OMNIC software (version ESP 5.1). The average of the 100 scans was used for further data analysis. FTIR spectral data were discretized resulting in spectra containing 1868 individual points (discretized every 2 $cm^{-1}$) for chemometric analysis.

The table (Table 3.1) below contains a summary of the sample numbers, sample ages and number of spectra taken during the study.

**Table 3.1 Summary of samples analysed by FTIR spectroscopy**

| Age (Day) | Damage (Min) | Tissue | Number of Spectra |
|---|---|---|---|
| 0 | 0 | C,G,S | 30 |
| 0 | 20 | C,G,S | 30 |
| 1 | 0 | C,G,S | 30 |
| 1 | 20 | C,G,S | 30 |
| 2 | 0 | C,G,S | 30 |
| 2 | 20 | C,G,S | 30 |
| 3 | 0 | C,G,S | 30 |
| 3 | 20 | C,G,S | 30 |
| 4 | 0 | C,G,S | 30 |
| 4 | 20 | C,G,S | 30 |
| 5 | 0 | C,G,S | 30 |
| 5 | 20 | C,G,S | 30 |
| 6 | 0 | C,G,S | 30 |
| 6 | 20 | C,G,S | 30 |
| 7 | 0 | C,G,S | 30 |
| 7 | 20 | C,G,S | 30 |

C,G,S: Cap, gills, stipe

### *3.1.4. Chemometric Data Analysis*

Multivariate models for damage and age prediction in mushrooms using both raw (i.e. unmodified) and pre-treated spectral data were developed. The pre-treatment used was

standard normal variate (SNV) and was intended to reduce scatter-induced effects in the spectra (Preisner *et al.*, 2008). The frequency region studied was 2000-400 $cm^{-1}$ (fingerprint region). This spectral region encompasses absorptions from most of the chemical species present and attenuation of the dataset in this way avoids spectral regions which have low information content and may therefore interfere with effective model development.

Principal component analysis (PCA) was used to identify patterns in data in a way which emphasises differences and similarities. It is used to indicate relationships among groups of variables in a data set and show relationships that might exist between objects (Zheng *et al.*, 2009).

Random forest (RF) models were built to (a) discriminate between damaged and undamaged mushrooms and (b) to predict mushroom ages. The number of trees fitted to build the random forest was 1000. The number of random wavenumbers tried at every node of the tree was set at 500 after optimization and the RF model trained was made using a stratified random sampling strategy of the sample spectra that would take the same number of samples from each of the tissues.

Partial least squares (PLS) regression was applied to the spectral data sets to develop a quantitative model for prediction of the age of damaged mushrooms. A common problem in development of multivariate prediction models is selection of the optimum number of PLS loadings. Often this selection is based on an examination of the root-mean-square error of cross-validation (RMSECV). But identification of a minimum is not always possible or unambiguous and sub-optimal models incur a significant risk of overfitting. Experience has shown that this can be a problem when parameters which are of practical relevance, such as post-harvest age or damage, but have unclear molecular basis are being modelled. To avoid overfitting, model cross validation was employed as follows:

1. Samples were randomly-designated from each tissue/damage status/time grouping as calibration (60%) or validation (40%) samples. The validation subset was left out during the optimization of model based on the calibration set.
2. The model optimization step was carried out in order to estimate optimal dimensionality of the PLS model built on the calibration set. The method employed for this was based on the observation that an indication of overfitting is the appearance of noise in regression vectors. This takes the form of a reduction in apparent structure and the presence of sharp peaks with a high degree of directional oscillation. A simple method (Gowen *et al.*, 2010) for objectivity quantifying the shape of a regression vector, combined with the root mean square error of cross-validation (RMSECV) for the calibration set was applied in this study.
3. The random sample designation, model development and evaluation were performed 100 times. At the end of this cycle, models were initially examined on the basis of the number of latent variables selected. The most common number was then chosen as the optimum.

Mushroom discrimination (damaged *versus* undamaged) was performed using partial least squares discrimination analysis (PLS-DA). For PLS-DA, analytical contrasts were used to specify the damaged and undamaged factor. PLS-DA calibration models were developed and assessed using 100 randomly-populated calibration and validation sample sets.

Principal component analysis (PCA) and partial least squares (PLS) regression were performed using MATLAB and The Unscrambler software (v.9.7; Camo A/S, Oslo, Norway). The routine for selection of the optimum number of PLS loadings was also performed in MATLAB. Random forest (RF) modelling was performed using R 2.8.0 (R Development Core Team, 2007).

Univariate statistical analysis was carried out on specific wavenumbers highlighted by RF models as being important variables for detecting damage in mushrooms. The significance of damage and tissue factors and their interactions was tested by analysis of variance (ANOVA). R 2.8.0 was used to analyse the data.

## 3.2. RESULTS AND DISCUSSION

### *3.2.1. Spectral Data*

Average raw spectra of each of the three tissue types collected from all the damaged and undamaged samples (day 0-7) are shown in Figure 3.1(a-c). A number of observations may be made on these spectra:

First, the major feature is a vertical offset from one average plot to another. This offset originates in light scatter effects and may be a complication in further data analysis. Average spectra of the three tissue types also bear a close resemblance to each other with little visible difference in peak minima locations in Figure 3.1.

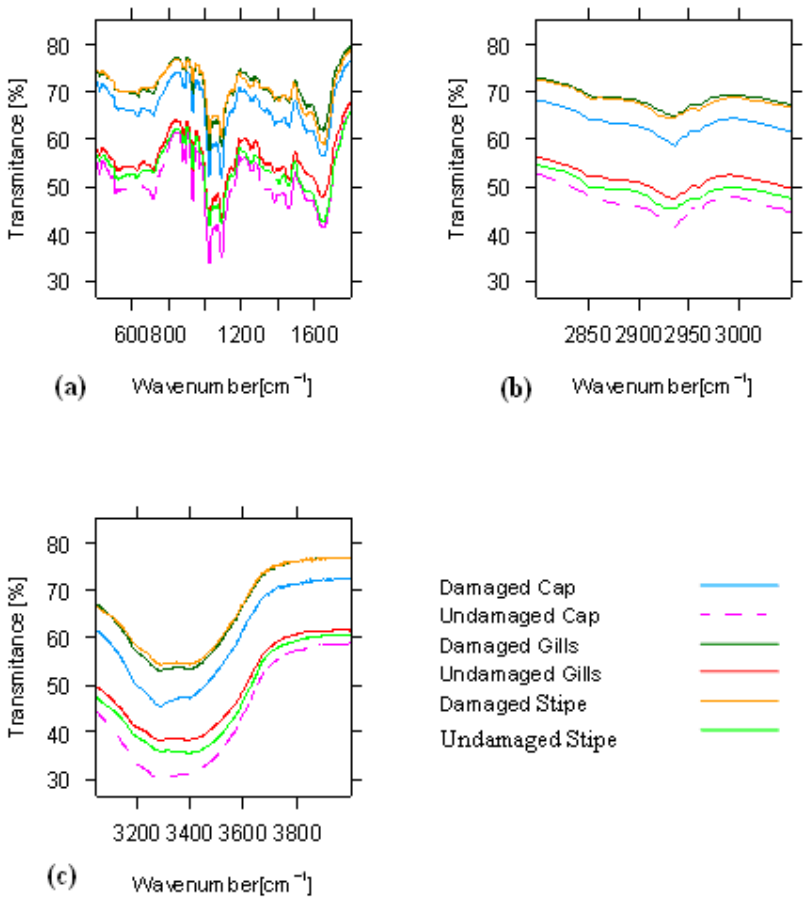



**Figure 3.1 FTIR transmittance spectra of all mushroom tissues in (a) 400-1800 $cm^{-1}$ (b) 2800-3050- $cm^{-1}$ and (c) 3050-4000 $cm^{-1}$ ranges**

In terms of minima locations there are major bands at 1650, 1090, 1020 and 935 $cm^{-1}$ and minor minima may be seen at 1560, 1150 and 1050 $cm^{-1}$ (Figure 3.2). Unambiguous identification of the molecular source of features in mid-infrared spectra of biological material is difficult but the peak at 1650 may be attributed to an amide I group vibration while the peak at 1560 $cm^{-1}$ may be identified as resulting from vibrations of amide II groups (Belton *et al.*, 1995; Di Mario *et al.*, 2008).

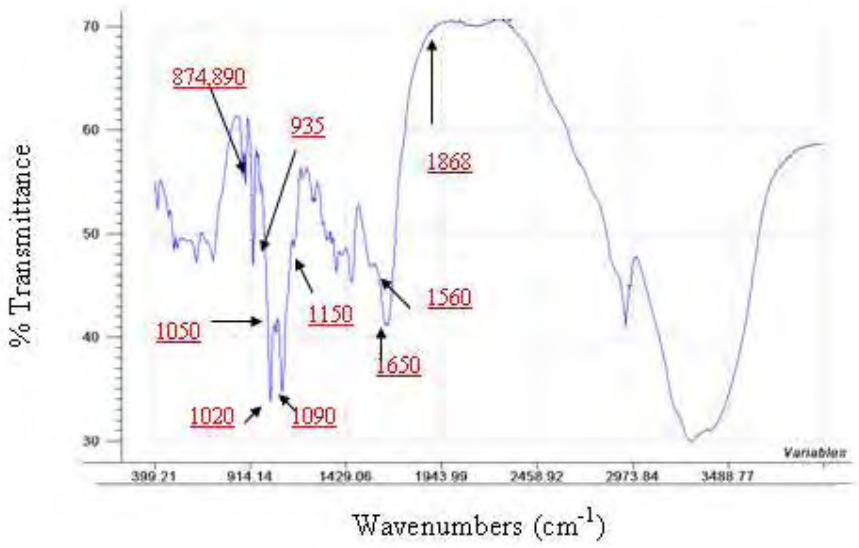



**Figure 3.2 Average FTIR spectrum of undamaged cap tissue 4000-400 $cm^{-1}$ (raw data)**

Both major absorption peaks at 1090 and 1020 $cm^{-1}$ have been attributed to vibrations of chitin, a major structural polysaccharide in mushrooms. Absorption at 1090 $cm^{-1}$ may also arise from vibrations of secondary alcohols. Smaller features at 1150 and 1050 $cm^{-1}$ have been attributed to vibrations of tertiary and primary alcohol structures (Workman, 2001). Minima at 935, 890 and 874 $cm^{-1}$ correspond to α- and ß-anomer $C_1$-H deformations. The vibrations at 935 and 890 $cm^{-1}$ are attributed to glucan vibrations, while the vibration at 874 $cm^{-1}$ is assigned to a mannan vibration (Pierce and Rast, 1995; Sandula *et al.*, 1999; Mohacek-Grosev *et al.*, 2001). An inability to attribute all spectral features is a common feature of spectroscopy of complex biological matrices, but the presence of such spectral detail implies the detection of a significant quantity of information which may be usefully interrogated by multivariate mathematical methods.

### *3.2.2. Principal Component Analysis*

Undamaged samples were studied separately on the basis of their tissue type i.e. caps, gills and stipes. The initial PCA of the mushroom caps data revealed a single sample (day 7)

that lay anomalously at some significant distance from the others. This was deleted and the resulting score plot is shown in Figure 3.3 for PC1 *versus* PC2. These first two principal components accounted for 97 and 2% respectively of the total variance in the spectral dataset and some sample clustering on the basis of storage time is readily apparent. As a general observation, it may be stated that the majority of the day 0 mushroom caps have a score value on PC1 greater than zero and are therefore located on the right-hand-side of Figure 3.3(a). While there are indications in the plots that samples of different storage time cluster together, the spread of these clusters is quite large and it is not possible to readily discern any trend relating plot position and storage time. There is a suggestion that the dispersion of the samples decreases as the length of storage time increases. With regard to undamaged gill tissue, observations similar to those made above in relation to undamaged caps may be made although the distribution patterns are somewhat different.

In the case of damaged mushroom tissues, a different pattern was found. It is clear from Figure 3.3(d-f) that day 0 samples clustered together but separately from those of day 1 to day 7 samples, irrespective of tissue type. This strongly suggested that physical damage had a significant effect on tissue structure and the subsequent ageing process. Some implications regarding the rate of change of mushroom tissue composition with ageing may be garnered from the observation that separation of day 0 from all other subsequent days accounts for the most variation in the spectral collection of damaged mushroom caps, gills and stipes.

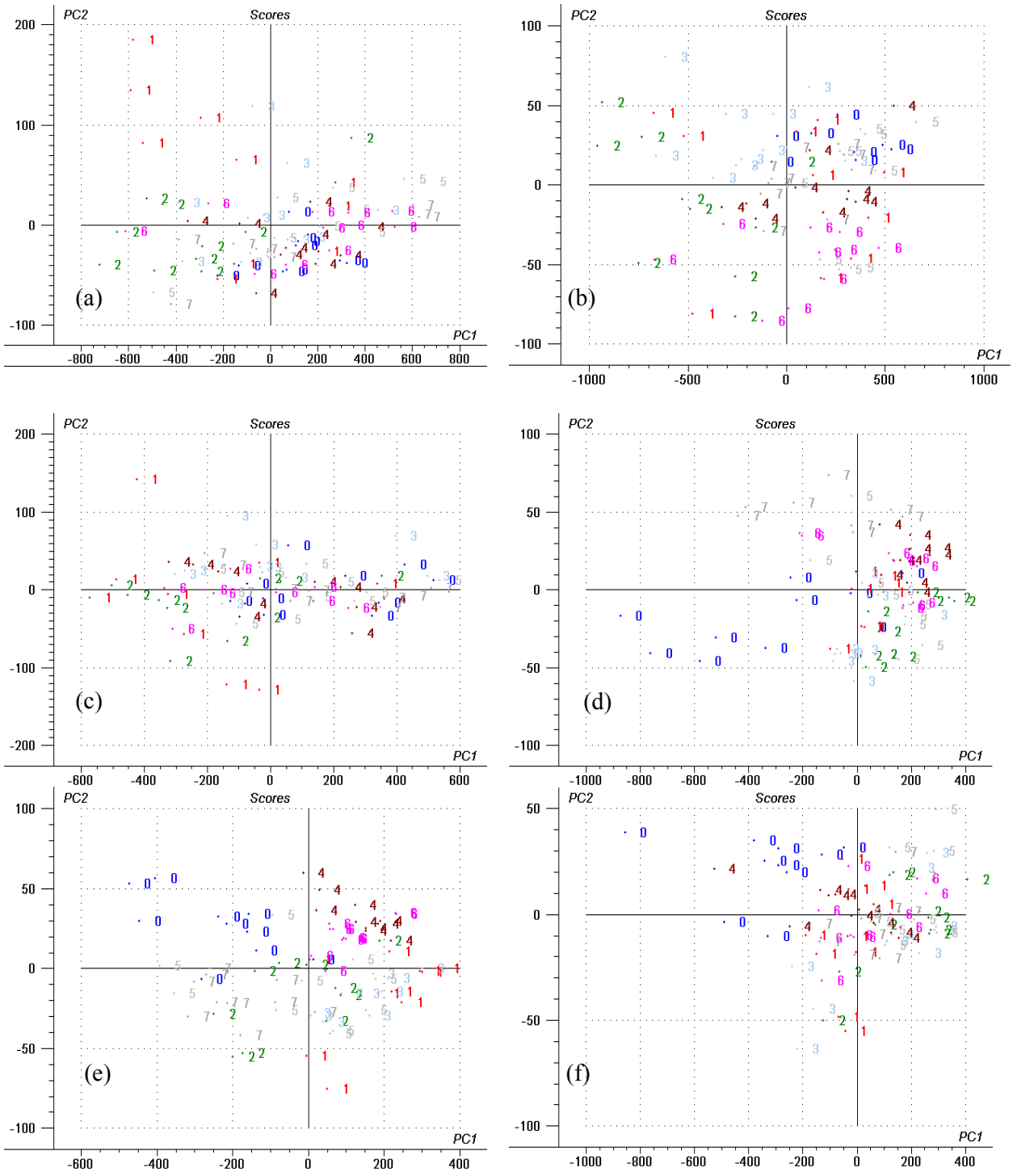



**Figure 3.3 PC1*versus* PC2 score plots of undamaged mushroom tissue (a) caps (b) gills (c) stipes and damaged tissue (d) caps (e) gills and (f) stipes; 0-7: Sample ages from zero to seven days**

### *3.2.3. Detection of Damage*

#### *3.2.3.1. Random Forests*

The first random forest model developed attempted to identify which wavenumbers could be used to predict damage specifically. The model tried to predict damage in mushrooms using the IR spectra, a variable indicating the tissue from which the spectra originated (cap, gill or stipe) and the age of the mushroom (in days from 0-7) as explanatory variables. This resulted in good classification between damaged and undamaged samples with an out-of-bag error rate (OOB) of 5.9%, sensitivity 93.3% and specificity 95%.

In RFs there are two measures of importance to indicate how informative a particular variable (a wavenumber in our case) is, the mean decrease in accuracy and the Gini index. The decrease in Gini index is not as reliable as the marginal decrease in accuracy (Breiman, 2001; Pang *et al.*, 2006) and for that reason the latter was analysed. The most important variables for predicting damage in the model are shown in Figure 3.4(a). The most important variable was the age of the mushroom samples followed by the wavenumbers 1868, 1870 and 1845 $cm^{-1}$.

Induced damage in mushrooms leads to an enzymatic response which is followed by brown discolouration. The enzymes involved in this response, tyrosinase or polyphenol oxidases, catalyse the oxidation of phenols, which in turn promote the formation of melanin-like compounds. This reaction is found not only in damaged mushrooms, but is also part of the natural ageing process, with mushrooms becoming darker and less firm during storage (Eastwood and Burton, 2002). The three wavenumbers identified can be used to differentiate between the chemical changes that are induced by the mechanical damage and are independent of those that take place due solely to ageing. The three wavenumbers are unassigned peaks.

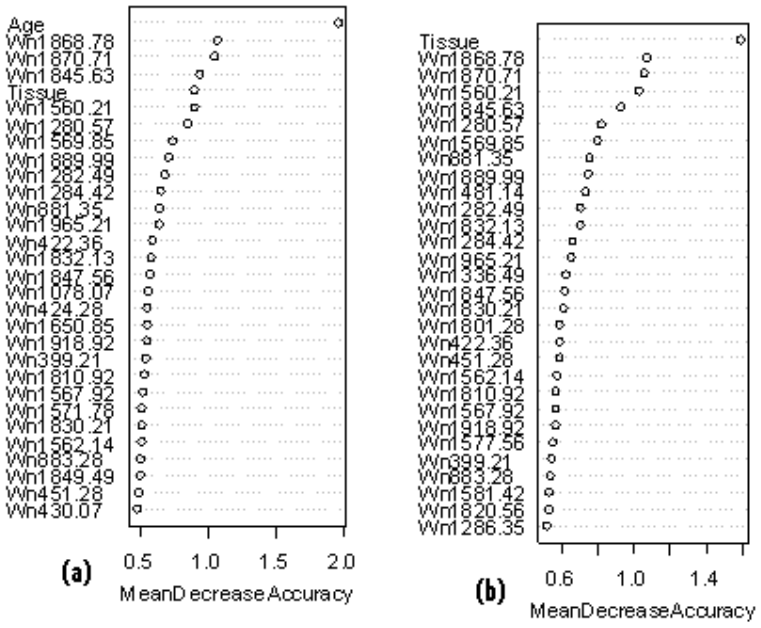



**Figure 3.4 Relative importance plot of variables that are important in the random forest model for predicting damaged/undamaged samples. (a)The most important variables are age followed by the wavenumbers 1868, 1870 and 1845 $cm^{-1}$ (b) When age is not an explanatory variable, the most important variables are tissue type followed by the wavenumbers 1868, 1870 and 1560 $cm^{-1}$**

By removing the variable age from the model a second model was built and used to predict whether there was damage or not. This random forest could be used as a classifier of mushroom damage and gave a very good prediction model with an OOB error rate of 9.8%, sensitivity of 89.2% and specificity of 91.2%. Even receiving mushrooms whose storage time after harvest was unknown, the model could still classify damaged and undamaged mushroom samples with a very good classification rate. The variables of importance involved in this classification model are shown in Figure 3.4(b).

The most important variable for predicting damage according to the mean decrease accuracy plot is the tissue used in the analysis followed by the wavenumbers 1868, 1870 and 1560 $cm^{-1}$. The peak located at 1560 $cm^{-1}$ is attributed to amide II vibrations of proteins (Mohacek-Grosev *et al.*, 2001). Amide II bands are associated with an out-of-phase combination of in-plane C-N stretching and N-H bending of amide groups (Militello *et al.*, 2004). Absorption of this band was found to be higher in damaged samples and

therefore an important variable for detecting damage in mushroom samples. The wavenumbers 1868 and 1870 $cm^{-1}$ are unassigned.

Appropriate univariate statistical treatment was applied to the wavenumbers identified by RFs as being important variables for detecting damage in mushrooms. These wavenumbers were 1868, 1870, 1845 and 1560 $cm^{-1}$. The significance of damage along with other factors such as tissue type and their interactions were tested by analyses of variance (ANOVA). Box-plots for the wavenumbers at 1560 and 1868 $cm^{-1}$ (Figure 3.5) show an increase in absorbance within damaged samples, a trend which was seen for all important wavenumbers.

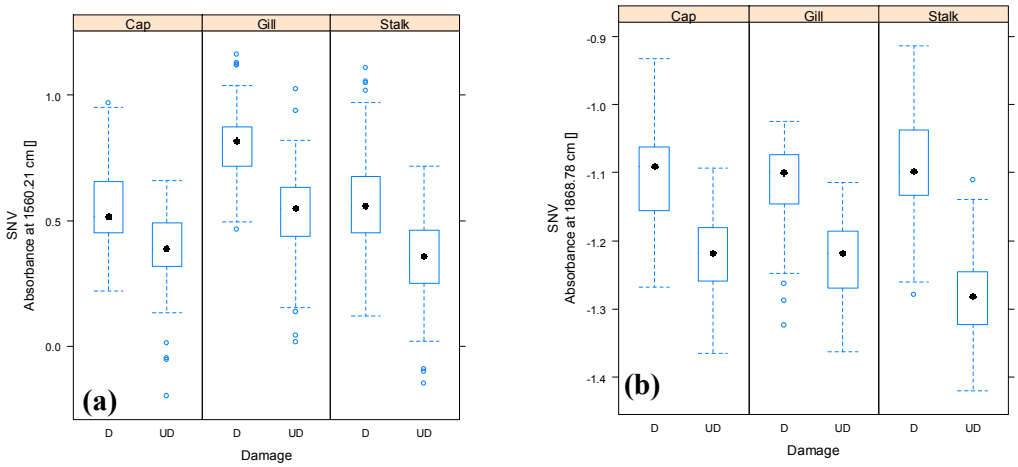



**Figure 3.5 Box plot showing the absorbance values for each tissue at (a) 1560 $cm^{-1}$ and (b) 1868 $cm^{-1}$**
D=Damaged samples
UD=Undamaged samples
SNV=Standard normal variate (pre-treatment)

The significance of the Damage factor ($p<0.001$) for each wavenumber (Table 3.2) indicates that the difference in absorbing species at these wavenumbers was significant between damaged and undamaged samples. For all the important variables identified by RFs, they each had higher absorption levels within damaged samples, an example of this was seen for 1560 and 1868 $cm^{-1}$ in the box-plots (Figure 3.5).

The Tissue factor also had a significant effect ($p<0.001$) on the absorption at these wavenumbers. This meant that each tissue type may react differently to physical damage. The significance of the Day*Tissue interaction means that damage and tissue type i.e. cap, gills or stipe had an effect on the absorption of the specific wavenumbers. It should be noted that tissue type was identified as an important variable for predicting damage (Figure 3.4b) in RF models. The box-plot (Figure 3.5) clearly shows that damaged samples had higher absorption than undamaged samples for all tissue types.

**Table 3.2 ANOVA table of the effect of damage, tissue and age on the absorption at specific wavenumbers identified as important variables for predicting damage by RFs**

| Wavenumber ($cm^{-1}$) | Factor | P-Value | Significance Level |
|---|---|---|---|
| 1868 | Damage | $2.2 \times 10^{-16}$ | *** |
| | Tissue | $5.1 \times 10^{-05}$ | *** |
| | Damage*Tissue | $3.5 \times 10^{-12}$ | *** |
| 1870 | Damage | $2.2 \times 10^{-16}$ | *** |
| | Tissue | $2.2 \times 10^{-16}$ | *** |
| | Damage*Tissue | $0.7 \times 10^{-02}$ | ** |
| 1845 | Damage | $2.2 \times 10^{-16}$ | *** |
| | Tissue | $4.0 \times 10^{-05}$ | *** |
| | Damage*Tissue | $1.8 \times 10^{-12}$ | *** |
| 1560 | Damage | $2.2 \times 10^{-16}$ | *** |
| | Tissue | $2.2 \times 10^{-16}$ | *** |
| | Damage*Tissue | $0.7 \times 10^{-02}$ | ** |

Significance levels at 95% (*), 99%(**), 99.9% (***)

In conclusion damaged samples had higher absorption at the wavenumbers 1868, 1870, 1845 and 1560 $cm^{-1}$ compared to undamaged samples regardless of tissue type. The ANOVAs are therefore complimentary to RF models for damage as the variables of importance for predicting damage were significantly different between damaged and undamaged samples.

*3.2.3.2. Partial Least Squares*

PLS-DA models were developed to discriminate between undamaged and damaged mushrooms of all tissue types separately. A summary of the average and dispersion of the results obtained on a percentage basis for each tissue is shown in Table 3.3.

**Table 3.3 Summary of results for mushroom discrimination using PLS-DA on the basis of damage**

| Tissue type | #Samples | #Loadings | % undamaged misclassified mean (std. deviation) | % damaged misclassified mean (std. deviation) |
|---|---|---|---|---|
| Caps | 160 | 7 | 4.1 (4.3) | 7.6 (4.0) |
| Gills | 160 | 9 | 2.1 (3.0) | 0.8 (1.7) |
| Stipes | 160 | 12 | 1.7 (2.1) | 0.6 (1.5) |

It is apparent that misclassification errors associated with all models are low, especially so in the case of gills and stipes. In terms of numbers of samples misclassified, these percentages translate to only 1 or 2 samples in each case. These results indicate that FTIR of freeze-dried mushroom tissues (especially gills and stipes) may be used to discriminate between damaged and undamaged mushrooms aged post-harvest from 0-7 days with almost complete confidence.

Modelling damage in mushrooms has been reported in literature in 2008 by Gowen and colleagues and Esquerre *et al.,* in 2009 (Gowen *et al.*, 2008a; Esquerre *et al.*, 2009). Gowen and colleagues investigated the use of hyperspectral imaging and principal component analysis (PCA) to develop models to predict damage on mushroom caps with correct classification ranging from 70-100%. Using near infra-red spectroscopy and partial least squares (PLS) regression, Esquerre and colleagues were able to correctly classify undamaged mushrooms from damaged ones with an overall correct classification model with 99% accuracy. The models for predicting damage using FTIR and random forests correctly classified 94 and 90% of samples respectively, whilst the PLS predictive models correctly classified 92-99% of undamaged samples from damaged ones.

These results highlight the usefulness of FTIR and chemometrics for detecting damage in mushrooms. This could lead to the development of a monitoring and classification system to detect physical damage before browning becomes visible, using specific wavenumbers identified as important variables in the 'fingerprint region' of mushroom spectra.

### *3.2.4. Predicting Postharvest Age*

#### *3.2.4.1. Random Forests*

Initial random forest models were built to predict the mushroom age from day zero to seven (0-7) using the IR spectra from the tissues and knowing whether they had been subjected to damage or not with the aim to identify specific wavenumbers associated with ageing. The random forest model produced an OOB error rate of 32% i.e. 68% of samples were correctly classified. The results of the model fit are shown in Table 3.4. Misclassification of samples was seen for all mushroom ages particularly days 4, 5 and 7. Classification of day zero samples performed quite well in the model with 82% of samples correctly classified, which leads to the possibility of using IR spectroscopy as a tool to discriminate fresh mushrooms (D0) from mushrooms that have been subjected to refrigeration. This type of tool could enable packers and producers to avoid fraud and 'recycling' of product, supporting the evidence from visual inspection.

**Table 3.4 Confusion matrix and the error rate for the prediction of mushroom age. The OOB error rate: 32%. The numbers in bold are correctly classified samples**

| | | Predicted Age | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **Error rate** |
| **Actual Age** | **0** | **49** | 3 | 0 | 3 | 2 | 0 | 3 | 2 | 0.18 |
| | **1** | 1 | **42** | 2 | 4 | 0 | 1 | 4 | 6 | 0.30 |
| | **2** | 4 | 5 | **43** | 2 | 3 | 0 | 0 | 3 | 0.28 |
| | **3** | 1 | 3 | 5 | **47** | 2 | 1 | 0 | 1 | 0.22 |
| | **4** | 3 | 0 | 3 | 3 | **32** | 2 | 8 | 9 | 0.47 |
| | **5** | 0 | 0 | 3 | 12 | 3 | **29** | 4 | 8 | 0.51 |
| | **6** | 1 | 0 | 6 | 0 | 2 | 0 | **48** | 3 | 0.20 |
| | **7** | 2 | 1 | 5 | 2 | 2 | 6 | 8 | **34** | 0.43 |

0-7: Sample age in days from day zero to day seven
Error rate: The % misclassification for each sample age

The variables of importance identified by the mean decrease accuracy plot were damage, tissue type and the wavenumbers 399, 952 and 1508 $cm^{-1}$ (Figure 3.6a).

A second model was developed to predict age using the same approach as above but removing the damage variable from the model. The model performed similarly to above with an OOB error rate of 33%. Again misclassification within all samples ages was seen. The model correctly classified 79% of day zero samples. The important variables identified to predict age were tissue type and the wavenumbers 399, 954, 952 and 1508 $cm^{-1}$ (Figure 3.6b). The peak located at 952 $cm^{-1}$ is assigned to a vibration of glucan (ß-anomer C-H deformation) (Mohacek-Grosev *et al.*, 2001). Glucans play many different roles in the physiology of fungi, with some accumulating in the cytoplasm as storage. However most are present in the cell wall structure (Ruiz-Herrera, 1992). This suggests that the ability to model ageing in mushrooms may depend on the effect of glucan levels changing in the cell wall due to natural senescence. The bands at 399 and 1508 $cm^{-1}$ are unassigned. The OOB errors produced to model ageing were quite large >33% which may be due to the low sample numbers.

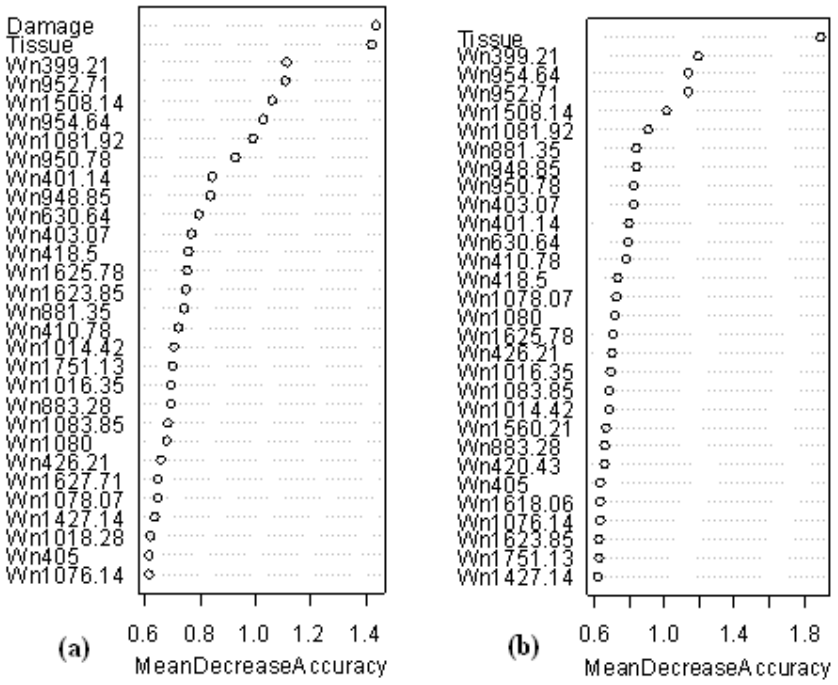



**Figure 3.6 Relative importance plot of variables that are important in the random forest model for predicting age. (a) The most important variables are damage followed by tissue type and the wavenumbers 399, 952 and 1508 $cm^{-1}$ (b) When damage is not a variable, the most important variables are tissue types and the wavenumbers 399, 954, 952 and 1508 $cm^{-1}$.**

### *3.2.4.2. Partial Least Squares*

PLS regression was applied separately to the caps, gills and stipes datasets in an attempt to develop separate quantitative models for prediction of the age of mushrooms, both damaged and undamaged. Selection of the appropriate number of latent variables for each model was assessed on the basis of the frequency of their occurrence. As shown in Figure 3.7, this was a clear and unambiguous choice. A summary of the results obtained using mushrooms from day 0 to 7 inclusive is shown in Table 3.5. In the case of undamaged mushrooms, root mean squared error of cross validation (RMSECV) values achieved were relatively high, only permitting the prediction of post-harvest age of damaged mushrooms to within ±2 to 3 days approximately (95% confidence limit) depending on tissue type.

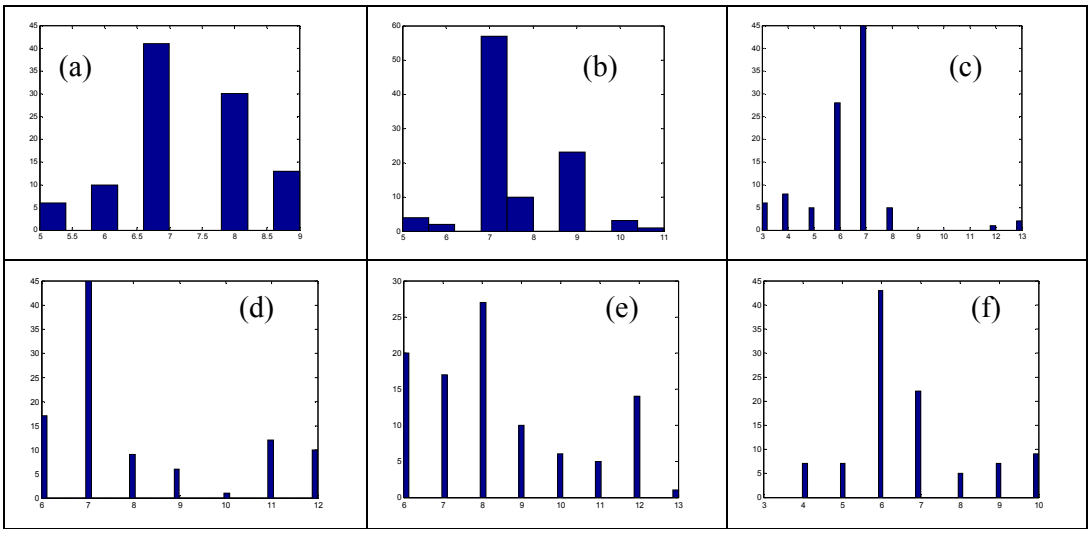



**Figure 3.7 Frequency of generation of PLS regression models for mushroom post-harvest age on the basis of the number of latent variables selected. (a) undamaged caps, (b) undamaged gills, (c) undamaged stipes, (d) damaged caps, (e) damaged gills and (f) damaged stipes. Abiscissa – no. of latent variables in model; ordinate – no. of occurrences**

**Table 3.5 Summary of PLS regression results for the prediction of post-harvest age (day 0-7 inclusive) in undamaged and damaged mushrooms**

| Treatment | Tissue | #Samples | #Loadings | RMSECV* | RER** |
|---|---|---|---|---|---|
| Undamaged | Caps | 80 | 7 | 1.2 | 2.0 |
| | Gills | 80 | 7 | 1.5 | 1.6 |
| | Stipes | 80 | 7 | 1.2 | 1.9 |
| Damaged | Caps | 80 | 7 | 1.3 | 1.9 |
| | Gills | 80 | 8 | 0.8 | 3.1 |
| | Stipes | 80 | 6 | 1.2 | 2.2 |

*RMSECV= root mean square error of cross-validation (mean of 100 runs); **RER = SD/RMSECV

The practical utility of such accuracy levels may be gauged by examination of the SD/RMSECV ratio, all but one of which are below 3.0, the generally accepted minimum value for a model to be of practical utility. With regard to damaged mushrooms, model predictive accuracies were similar for caps and stipes with RMSECV (and RER) values of 1.3 (1.9) and 1.2 (2.0) respectively. In the case of gill tissue, better predictive accuracy was achieved with RMSECV and RER values equal to 0.8 and 3.1 respectively. The number of latent variables associated with these models was low and similar in all cases, with a

variation between 6 and 8 only. The application of an objective indicator of the optimum number of PLS loadings to include in any model contributed to their stable performance.

The results presented for modelling age in mushroom using FTIR and chemometrics had misclassification errors of over 30% (RFs) yielding relatively unsuccessful results. However, random forest models were able to classify day zero samples reasonably well with correct classifications of 82 and 79% which leads to the possibility of using IR spectroscopy in detecting fresh mushrooms from old mushrooms and could be used within the sector for detecting fraud and 'recycling' of product.

The time required for freeze-dried sample preparation and measurement in this protocol is of the order of hours. Thus this approach would be applicable for research and quality control purposes. However, this may be reduced to the order of minutes by the use of specific wavenumbers, possibly raw mushroom tissue and alternative IR sample handling (i.e. attenuated total reflectance). This study highlights the usefulness of FTIR coupled with chemometric data analysis in particular for evaluating damage in mushrooms.

## 3.3. CONCLUSIONS

FTIR and chemometric data analysis was applied to evaluate damage and age in mushrooms. RF models were produced with the ability to predict damage in mushrooms with low error rates (<10%). The first model developed used the IR spectra, a variable indicating the tissue type and the age of the sample as explanatory variables. This model produced a very low OOB error rate of 5.9%. A second model was produced with the age variable removed and this model performed well with an OOB error rate of 9.8%.

PLS-DA models were developed to discriminate undamaged and damaged mushrooms of all tissue types separately. Misclassification errors were low in all models, particularly in the case of gills and stipes. The use of FTIR coupled with PLS-DA produced strong

models with the ability to discriminate between damage and undamaged samples with almost complete confidence.

Random forest (RF) models were developed to try to predict mushroom age from day zero to day seven (0-7). Unfortunately models produced high error rates (>30%). However the model could correctly classify 82% of day zero samples, which could be used to discriminate fresh mushrooms from mushrooms that are one day or older and have been subjected to refrigeration.

PLS regression was applied to tissue types separately in an attempt to develop separate quantitative models for age prediction. Undamaged and damaged samples were treated separately. The models were able to predict postharvest age to within 2-3 days depending on tissue type for undamaged mushrooms. Damaged models resulted with predictive accuracies similar for caps and stipes (2 days), with better predictive accuracy achieved for gills.

Results presented in this work show that FTIR spectroscopy and chemometrics could be used to classify mushrooms according to their damage class (i.e. undamaged or damaged). This study demonstrates the potential use of FTIR as a tool for discriminating damage in mushrooms with the potential for developing a classification system for the industry.

# 4. THE USE OF NUCLEAR MAGENETIC RESONANCE SPECTROSCOPY AND CHEMOMETRIC DATA ANALYSIS TO EVALUATE LOW LEVELS OF DAMAGE IN MUSHROOMS (*AGARICUS BISPORUS)*

# 4. The use of nuclear magnetic resonance spectroscopy and chemometric data analysis to evaluate low damage levels in mushrooms (*Agaricus bisporus*)

The aim of this study was to investigate the chemical changes induced by low levels of mechanical damage in mushrooms using $^{1}$H NMR spectroscopy and to identify the extent of damage using NMR together with chemometric data analysis.

## 4.1. NMR Lipid Profile of Mushrooms

Lipid extracts from *Agaricus bisporus* have been analysed by 1D-proton and 2D-proton COSY NMR spectroscopy. Bonzom and colleagues studied a series of lipids extracted from freeze-dried mushrooms and performed qualitative and quantitative analysis. The data obtained was both accurate and detailed and obtained without chemical modification of the samples (Bonzom *et al.*, 1999).

In addition, similar success with this method was observed for metabolic profiling of lettuce leaves by Sobolev *et al*.. In this study a large number of water soluble metabolites and complex spectra of metabolites extracted in organic solvents were fully assigned (Sobolev *et al.*, 2005).

The results of both these studies demonstrate the potential of NMR spectroscopy as a method for the study of plant metabolism. It could also provide a useful tool for the studies of plant diseases, toxicity and the monitoring of metabolic changes.

## 4.2. MATERIAL AND METHODS

### *4.2.1. Mushrooms*

Second flush mushrooms were grown at the Teagasc Research Centre Kinsealy (Dublin, Ireland) and harvested damage free. A set of 120 closed cap mushrooms ($n$ = 120) were selected each week for three weeks (i.e. 3 repetitions) for this study. A subset ($n$ = 60) were labelled day zero mushrooms and the remaining sixty samples were labelled as day one samples. From each subset of sixty, thirty (n = 30) were labelled as damaged (D) and subjected to physical damage using a mechanical shaker set at 300 rpm (rotations per minute) for 30 seconds, giving these mushrooms an L-value of 98, which according to Gormley and O'Sullivan places them in a category of excellent quality (Gormley and O'Sullivan, 1975). The remaining samples ($n$ = 30) were labelled as undamaged (UD).

Day zero samples were prepared for analysis while day one samples were stored overnight at 4 ºC before being prepared for NMR analysis. A total of 1200 (600 polar and 600 non-polar) samples were analysed by NMR spectroscopy, a summary of which is given in Table 4.1.

**Table 4.1 Summary of samples analysed by NMR spectroscopy**

| Age (Day) | Damage (Seconds) | Tissue | Number of Samples | Number of Spectra (polar & non-polar) |
|---|---|---|---|---|
| 0 | 0 | Cap | 30 | 60 |
| | | Gills | 10 | 20 |
| | | Stipes | 10 | 20 |
| 0 | 30 | Cap | 30 | 60 |
| | | Gills | 10 | 20 |
| | | Stipes | 10 | 20 |
| 1 | 0 | Cap | 30 | 60 |
| | | Gills | 10 | 20 |
| | | Stipes | 10 | 20 |
| 1 | 30 | Cap | 30 | 60 |
| | | Gills | 10 | 20 |
| | | Stipes | 10 | 20 |

### *4.2.2. NMR Profiling Protocol & Overview*

Sample preparation involved the manual dissection of each mushroom into its three main tissue types (cap, gills and stipes). The samples were then frozen overnight at -70 ºC in a cryogenic fridge (Polar 340V Cryogenic fridge, Angeelantoni Industrie spA, Massa Martana, Italy), followed by freeze-drying (Micro-modulyo, EC Apparatus Inc, New York, USA). Dried sample tissues were then ground into a fine powder using a pestle and mortar and an extraction of polar and non-polar phases was performed as described by Wu *et al.,* with minor modifications (Wu *et al.*, 2008). An overview of the experimental procedure is described in Figure 4.1.

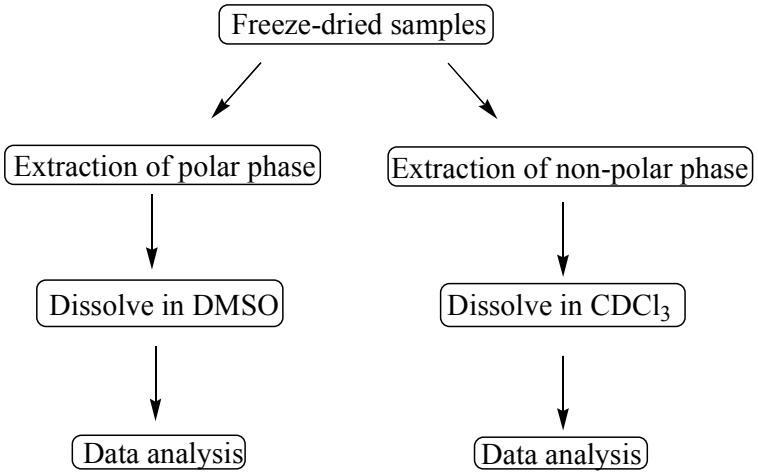



**Figure 4.1 Overview of metabolite fingerprinting NMR spectroscopy protocol**

#### *4.2.2.1. Extraction*

A tissue sample (400 mg) was homogenised in methanol (4 ml/g) and cold water (0.85 ml/g) in a Teflon tube. The homogenate was placed into a glass vial and the following solvents were added; chloroform (4 ml/g) and water (2 ml/g). Samples were then vortexed for 60 seconds, left on ice for 10 minutes to partition and centrifuged for 10 minutes at 2000 g at 4 ºC. The polar phase samples were dissolved in dimethyl sulfoxide-$d_6$ (DMSO-

$d_6$, 99.99 atom % D) (Sigma Aldrich, Dublin, Ireland). The non-polar phase samples were dissolved in deuterated chloroform ($CDCl_3$, 99.98 atom % D, containing 0.1% (v/v) tetramethylsilane (TMS)) (Sigma Aldrich, Dublin, Ireland). All solutions were filtered through glass wool and transferred to standard 7 mm NMR tubes (Sigma Aldrich, Dublin, Ireland).

### *4.2.3. NMR Measurements*

$^1$H NMR experiments were carried out on a Bruker Avance III 500 MHz spectrometer (Bruker Avance III UltraShield 500 MHz NMR, Germany) with a transmitter frequency of 500.13 MHz for protons (Figure 4.2). The pulse programme used was zg30 using a 30 degree flip, dwell time of 48.4 μs, acquisition time of 3.17 seconds and acquisition mode was DQD. 16 scans and 2 dummy scans were used and the sweep width was 20.7 ppm and receiver gain was 456. For processing a line broadening of 0.3 Hz was applied and the baseline correction used a 5 degree polynomial.

Spectra were referenced using the residual chloroform for DMSO signals. All spectra were obtained at 198 K. The NMR preprocessing software used was TOPSPIN 2.1 (version 2.1.4, Bruker BioSpin, Germany).


**Figure 4.2 (a) Bruker Avance III 500 MHz spectrometer and (b) NMR spectroscopy tubes containing mushroom samples stacked in the autosampler**

### *4.2.4. Chemometric Data Analysis*

The analysis of NMR data involved the following steps:

1. A preliminary observation of the data using principal component analysis (PCA) to identify clusters of data and outliers as appropriate.
2. Modelling the data using random forests (RF) was performed in order to confirm the ability of multivariate statistics to predict damage with the purpose of identifying important frequencies (signals) in the discrimination which may be used for the identification of markers of low level damage.
3. Univariate statistics (ANOVA) were used to assess the ability of the selected frequencies (signals) in the NMR spectra to discriminate low levels of damage.
4. A further step of modelling the data was performed in which the data was split into a training and test database using a random sampling procedure and resulting in 75% of the data being used for training and 25% for testing (Figure 4.3).

5. Partial least squares discriminant analysis (PLS-DA) using the indications of the PCA and optimising the hyperparameter of the number of components in the PLS regression step was performed. Confusion matrices for the training and test sets were used to identify the ability of PLS-DA models to discriminate damage.
6. Principal component analysis (PCA), random forests (RF), partial least squares (PLS) and univariate statistical analysis were performed using R 2.10.0 (R Development Core Team, 2009).

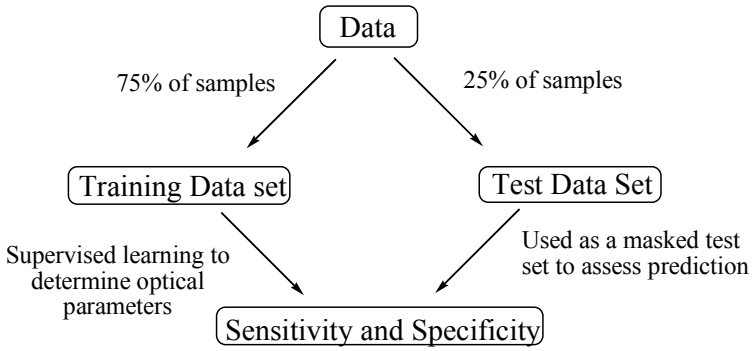



**Figure 4.3 Illustration of the splitting into training & test datasets to assess the prediction of damage in mushrooms (Pers *et al.*, 2008)**

## 4.3. RESULTS AND DISCUSSION

### *4.3.1. Non-polar Phase Spectral Analysis*

Figure 4.4 represents a typical $^{1}$H NMR spectra for a day zero undamaged non-polar phase fraction of a mushroom sample. Three distinct regions are apparent.

1. Aromatic groups are represented by signals between 6-10 ppm. A number of phenolic compounds are present in mushrooms. The main natural phenolics present in *Agaricus bisporus* are glutaminyl-4-hydroxybenzene, *p*-aminophenol, phenylalanine and tyrosine (Jolivet *et al.*, 1998). A number of other phenolics have

been reported in literature as being indentified in mushrooms such as *p*-hydroxybenzoic acid, *p*-coumaric acid and cinnamic acid (Gasowska *et al.*, 2004; Tsai *et al.*, 2009; Vaz *et al.*, 2011).

2. Carbohydrate groups are typically represented by signals between 3-6 ppm. Numerous carbohydrates have been reported in literature as present in mushrooms such as ribose, xylose, mannose, glucose, sucrose and trehalose (Beecher *et al.*, 2001; Heleno *et al.*, 2009; Kalač, 2009).
3. Finally lipid groups (aliphatic) are typically represented by signals between 0-3 ppm. A number of lipids have been reported in literature as present in mushrooms including fatty acids (Holtz and Schisler, 1971; Byrne and Brennan, 1975; Yilmaz *et al.*, 2006), sterols (Yokokawa and Mitsuhashi, 1981; Weete *et al.*, 1985; Bonzom *et al.*, 1999; Teichmann *et al.*, 2007), acylglycerols (Bonzom *et al.*, 1999) and phospholipids (Bonzom *et al.*, 1999).

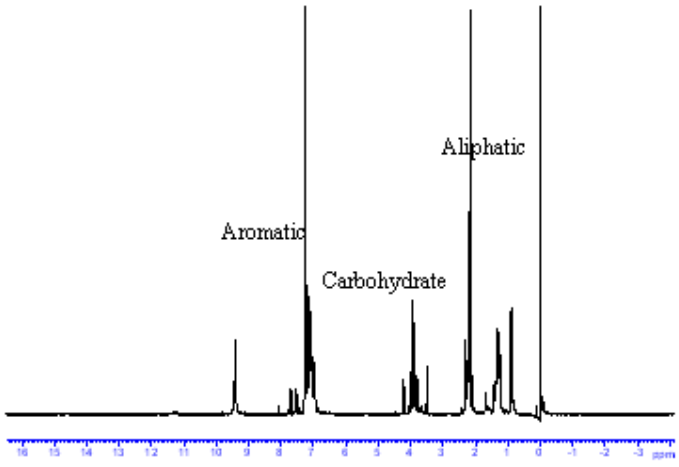



**Figure 4.4 Representative $^{1}$H NMR spectrum of undamaged day zero non-polar phase cap tissue**

Unfortunately poor resolution of the peaks hinders identification and assignment of individual groups for qualitative purposes. However, distinct variations can be observed between damaged and undamaged samples (Figure 4.5).

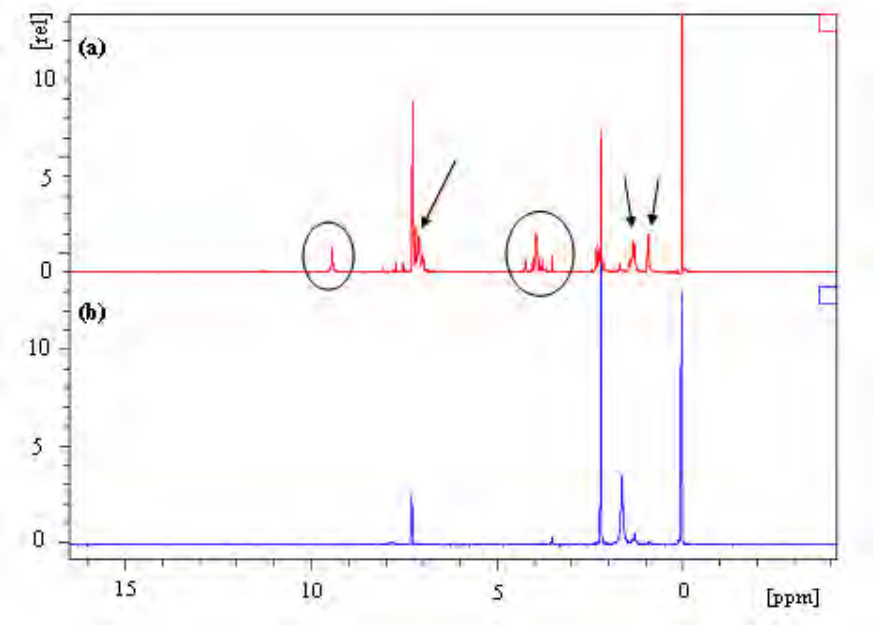



**Figure 4.5 $^{1}$H NMR spectrum of (a) undamaged cap *versus* (b) damaged cap day zero (non-polar phase)**
**Arrows and circles are used to highlight differences between the spectra**

Figure 4.5 depicts the typical spectra of (a) undamaged day zero cap and (b) damaged day zero cap. Visible differences in the spectra are highlighted. Peaks in the same areas appear less resolved and in some instances, signals are completely lost in the damaged samples. This is highlighted in the carbohydrate and aromatic regions (arrows and circles).

Although the identification of phenolic compounds was inconclusive, there were noted differences between this region in the damaged and undamaged mushroom samples. A definite decrease can be seen for the peak located at 7.3-7.4 ppm. There was also a loss of a number of signals beside this peak suggesting that phenolic compounds in mushrooms may be affected by low levels of damage indicating that this region of the spectra could possibly be used to discriminate between damaged and undamaged mushrooms.

The carbohydrate region also showed variation between the spectra of damaged and undamaged mushroom samples. The spectra for damaged mushroom caps contains less peaks in this region (as for all regions) again indicating that this type of damage may have an effect on carbohydrates present in mushroom cap samples which might be measureable by NMR spectroscopy.

Fatty acid chains are represented in the aliphatic region, with differences in the spectra also suggesting possible effects of damage to the mushrooms on these lipid groups. Differences in the spectra of both gills and stipes are also evident (Figure 4.6). The ability to detect these levels of damage i.e. 30 seconds, illustrates the potential of NMR analysis as a tool for profiling damage in mushroom samples.

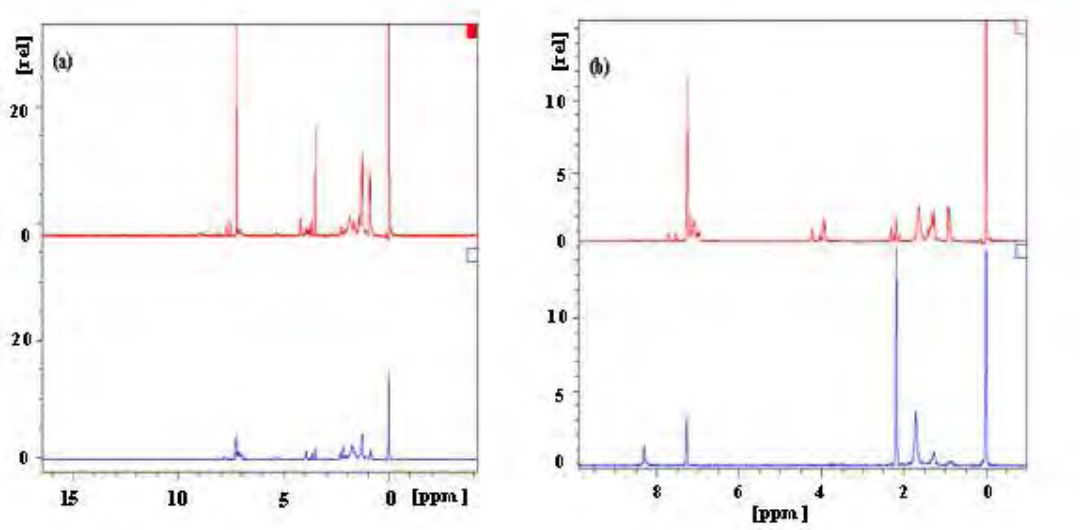



**Figure 4.6 $^{1}$H NMR spectra of non-polar phase day zero undamaged (red) and damaged (blue) (a) gills and (b) stipes**

*4.3.1.1. Assignment of Signals*

The assignment of individual metabolites by NMR analysis has been shown by Bonzom and colleagues. In this study lipid mixtures were separated on solid phase by ion exchange chromatography into four separate fractions corresponding to neutral lipids, free fatty

acids, neutral phospholipids and acidic lipids. The total lipid content was found to be represented by resonances at 2.31 ppm. This peak corresponds to the carboxylate protons ($CH_2COO$), which are present in all free fatty acids or conjugates. Hence the total lipid fraction was measurable (Bonzom *et al.*, 1999). Similar results were noted in the study of lettuce leaves (Sobolev *et al.*, 2005). NMR analysis also showed a similar peak at 2.3-2.4 ppm representing the $CH_2COO$ protons of the lipid fraction of the sample.

In addition further lipid protons were also identified in these two studies, allowing unequivocal assignment of individual fatty acid groups. For example, Bonzom *et al.* reported the following signals: A resonance at 2.0-2.1 ppm was assigned to protons next to the double bonds in the lipid chain ($CH_2CH=CHCH_2$). $^1H$ resonance at 1.6-1.7 ppm was assigned to protons next to the carboxylate groups ($CH_2CH_2COO$). The $^1H$ resonance peaks between 1.2-1.4 ppm were attributed to $CH_2$ protons in the lipid chain, with signals around 0.86 ppm assigned to methyl groups of the lipids ($CH_3$) (Bonzom *et al.*, 1999). These findings are also supported by lipids identified in lettuce leaves (Sobolev *et al.*, 2005).

Figure 4.7 shows the expanded aliphatic region of the spectrum of undamaged day zero cap tissue. Although the peaks are not resolved, comparisons can be made between results obtained in this study and the aforementioned studies. A peak at around 2.3 ppm (Figure 4.7) may represent the carboxylate protons ($CH_2COO$) of the lipids present in the sample, potentially identifying the total lipid content. The remaining groups identified by Bonzom *et al* $CH_2CH=CHCH_2$, $CH_2CH_2COO$, $CH_2$ and $CH_3$ may also be represented by peaks 2 to 5 respectively in Figure 4.7. Further resolution is needed to fully assign these signals.

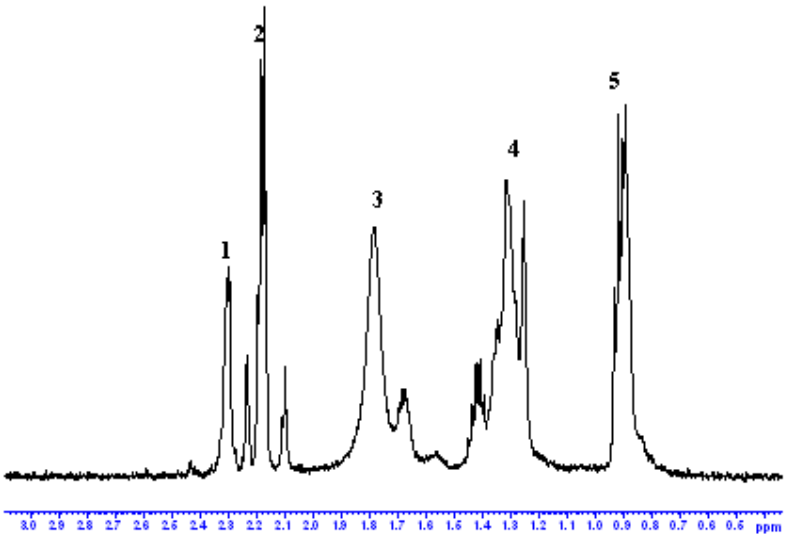



**Figure 4.7 $^{1}$H NMR spectrum (expanded) of lipidic fraction of undamaged cap tissue day zero containing possible signals for 1: C*H*$_2$COO, 2: C*H*$_2$CH=CHCH$_2$, 3: C*H*$_2$CH$_2$COO, 4: C*H*$_2$, 5: C*H*$_3$**

Identification of individual peaks was not attempted for signals in the aromatic or carbohydrate regions. Interpretation of the $^{1}$H NMR spectra was difficult in these regions without better resolution of the signals.

### *4.3.2. Polar Phase Spectral Analysis*

Figure 4.8 shows a typical $^{1}$H NMR spectrum of undamaged polar phase cap tissue in DMSO. The polar phase spectra are not as information rich compared to the non-polar phase spectra, nonetheless peaks are present in the aliphatic (amino acids) and midfield (carbohydrate) regions of the spectra.

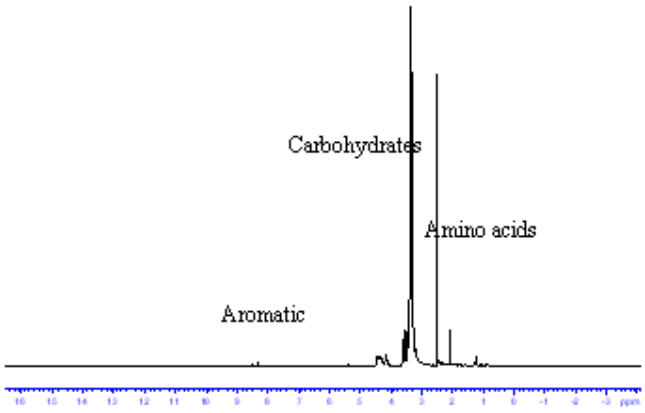



**Figure 4.8 Representative $^{1}$H NMR spectrum of undamaged day zero polar phase cap tissue**

Although a number of differences were noted between non-polar phase spectra for damaged and undamaged mushroom samples, this trend was not as clearly observed for polar phase samples (Figure 4.9).

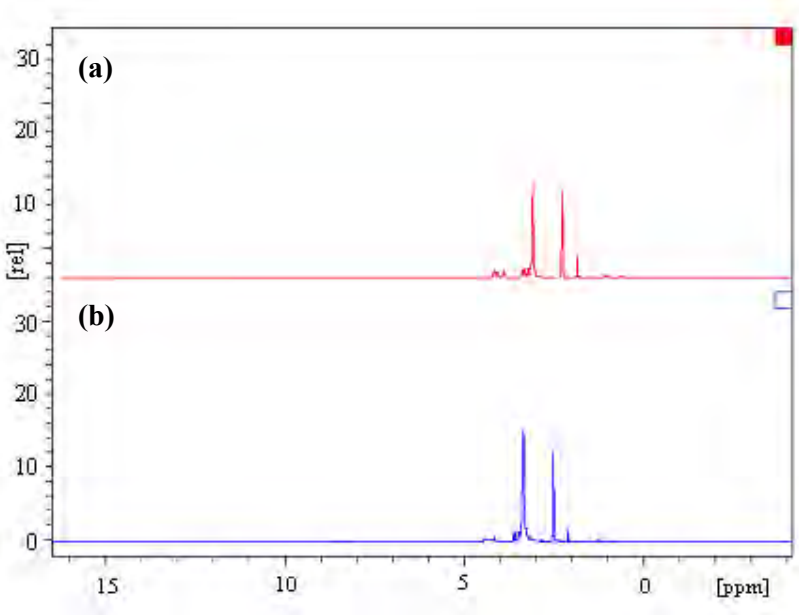



**Figure 4.9 $^{1}$H NMR spectrum of (a) undamaged cap day zero *versus* (b) damaged cap day zero (polar phase)**

Further extraction and purification procedures would be recommended in an attempt to identify individual peaks qualitatively and quantitatively. Suitable extractions may be

performed which specifically target individual groups. For example, extraction of phenolics using an acetone and water mixture described by Cheung *et al.* may be used to isolate these groups. The purpose of this study was to analyse the antioxidant activity and total phenolics of edible mushroom extracts. This extraction procedure could also be used for NMR analysis. The procedure involved freeze-drying the mushroom samples, followed by extraction using a Soxhlet extractor under reflux conditions. The residues were then extracted by boiling water with the organic solvent extracts removed using rotary evaporation and the water extracts dried in a freeze-dryer. Analysis was then carried out (Cheung *et al.*, 2003).

The Bligh and Dyer method was used by Bonzom *et al.* to extract lipids from freeze-dried and powdered cultivated *Agaricus bisporus* prior to NMR analysis (Bligh and Dyer, 1959). This was followed by solid phase chromatographic extraction, as it provides a rapid and effective way of isolating compounds of interest from complex matrices. Bond Elut ion-exchange chromatography was successfully used to separate lipids into four fractions, which were then analysed separately by NMR analysis (Bonzom *et al.*, 1999).

The successful extraction of individual groups of metabolites may optimise this method as a possible technique for profiling the effect of damage on mushroom metabolites. As seen by Bonzom *et al.,* integration of the signals aid in identification of individual metabolites, which may then be measured as means of assessing the effects of damage and time on the type and amounts of metabolites in these mushrooms.

### *4.3.3. Principal Component Analysis*

Samples were studied separately based on their tissue type i.e. cap, gills and stipes and on their age i.e. day zero and day one. Polar and non-polar phase groups were examined individually. PCA is an unsupervised method which converts high-throughput instrumental data (i.e. NMR) into a qualitative visual presentation (score plot) (Lindon *et al.*, 2001),

resulting in sample clustering into either similar or different groupings. The purpose of PCA in this study was to discriminate between classes i.e. damaged and undamaged mushrooms.

Score plots were analysed for a number of principal components (PCs). However, there was no separation between clusters i.e. no clear differentiation between damaged and undamaged samples. This trend was seen for all tissue types, days and phases (polar/non-polar). Examples are given in Figure 4.10 for non-polar gill tissue and in Figure 4.11 for polar gill tissue.

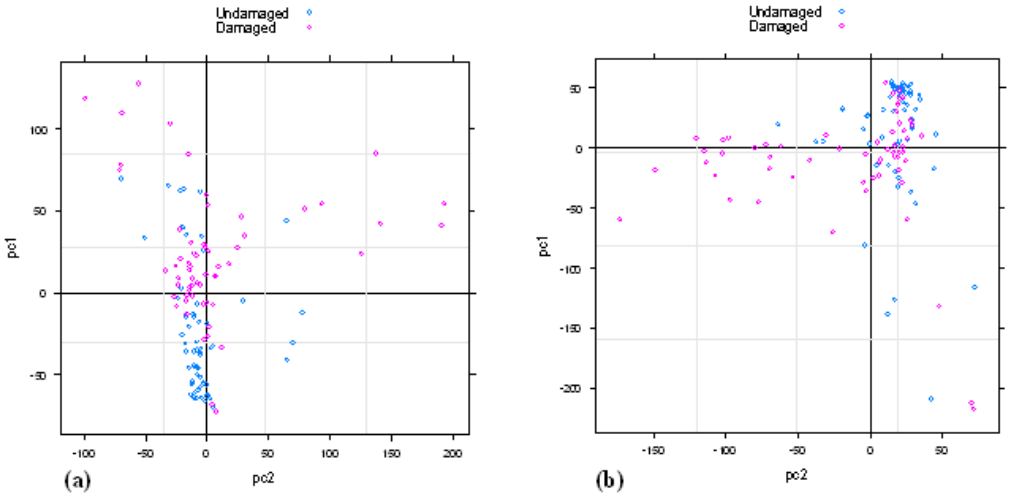



**Figure 4.10 PC1 *versus* PC2 score plots of non-polar phase gill tissue for (a) day zero samples and (b) day one samples**

The score plots for non-polar phase gill tissues are shown in Figure 4.10 for PC1 *versus* PC2. These first two principal components accounted for 46 and 43% respectively for day zero samples and 47 and 34% for day one samples. No clear separate clusters were found for non-polar phase samples. A similar trend was found in polar phase samples (Figure 4.11).

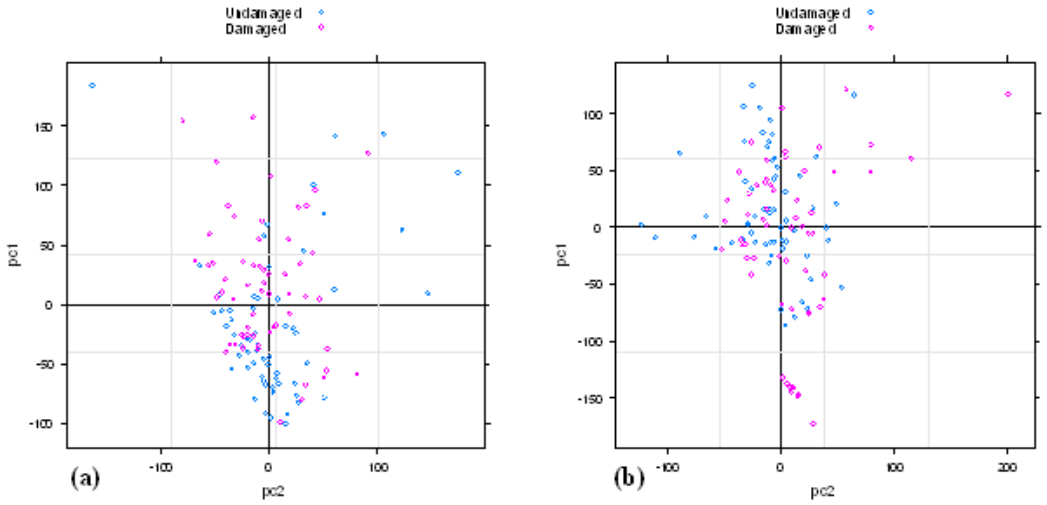



**Figure 4.11 PC1 *versus* PC2 score plots of polar phase gill tissue for (a) day zero samples and (b) day one samples**

### *4.3.4. Detection of Damage*

#### *4.3.4.1. Random Forests*

Overall random forest models were produced for all non-polar phase spectra (cap, gills, stipes) and for all polar phase spectra (cap, gills, stipes). RF models were developed for each day (day one/day zero) and for the individual tissues separately (Table 4.2). Non-polar phase samples produced the best models for discriminating between damaged and undamaged mushrooms, with non-polar phase stipes having the lowest OOB error rate (10%), followed by non-polar phase gill tissue (>15%). Visible differences could be clearly seen between damaged and undamaged non-polar phase $^{1}$H NMR spectra for all tissue types prior to chemometric analysis (Figure 4.5).

Differences between damaged and undamaged spectra were not clearly seen for polar phase samples. RF models produced high error rates when polar phase samples were used with gill and stipe tissue having error rates of >25% and therefore only non-polar phase samples may be used to predict damage in mushrooms.

**Table 4.2 Summary of RF models produced to discriminate between damaged and undamaged mushrooms using NMR data**

| Mushroom phase | RF Model | OOB (%) | Important Variables | Corresponding Signals (ppm) | Spectral region |
|---|---|---|---|---|---|
| NON-POLAR | All Samples | 17.33 | 14579, 14582, 13240, 14505, 22721 | 7.299, 7.297, 8.144, 7.346, 2.167 | Aromatic<br>Lipid |
| | Cap tissue | 25.0 | 14642, 14638, 14631, 14630, 26155 | 7.24, 7.262, 7.266, 7.267, 0.002 | Aromatic |
| | Gill tissue | 14.17 | 23297, 24641, 24645, 24635, 23289 | 1.804, 0.957, 0.954, 0.960, 1.809 | Lipid |
| | Stipe tissue | 10.83 | 14577, 14579, 14536, 14584, 14578 | 7.301, 7.299, 7.327, 7.296, 7.300 | Aromatic |
| | Day zero samples | 11.42 | 23289, 24644, 24626, 24042, 24635 | 1.809, 0.955, 0.966, 1.334, 0.960 | Lipid |
| | Day one samples | 9.97 | 22126, 22154, 23092, 23113, 14639 | 2.542, 2.524, 1.933, 1.920, 7.262 | Lipid<br>Aromatic |
| POLAR | All Samples | 45.67 | 22231, 22811, 22934, 22198, 22237 | 2.476, 2.110, 2.293, 2.497, 2.472 | Amino |
| | Cap tissue | 13.33 | 22233, 22241, 22193, 20944, 22196 | 2.475, 2.469, 2.499, 3.287, 2.498 | Amino |
| | Gill tissue | 25.83 | 20912, 23667, 22846, 14158, 13701 | 3.307, 1.564, 2.088, 7.565, 7.853 | Carb<br>Amino<br>Aromatic |
| | Stipe tissue | 35.0 | 17022, 22817, 17087, 12995, 26048 | 5.760, 1.571, 2.088, 7.565, 7.853 | Amino<br>Aromatic |
| | Day zero samples | 25.72 | 17095, 16040, 14712, 17331, 22826 | 5.713, 6.378, 7.216, 5.565, 2.101 | Aromatic<br>Amino |
| | Day one samples | 19.0 | 23667, 23117, 14027, 14042, 23651 | 1.571, 1.917, 7.647, 7.638, 1.581 | Amino<br>Aromatic |

RF: Random forest
OOB: Out of bag error rate

Variable importance plots (VIP) were produced for all of the RF models with non-polar phase samples indicating signals in the lipid and aromatic regions of the spectra as being important for modelling damage in mushrooms. RF models produced using the polar phase spectra identified signals in the amino acid and aromatic regions of the $^{1}$H NMR spectra as important for differentiating damaged and undamaged samples. Figure 4.12 shows examples of VIP plots produced for non-polar phase stipe tissue and non-polar phase gill tissue.

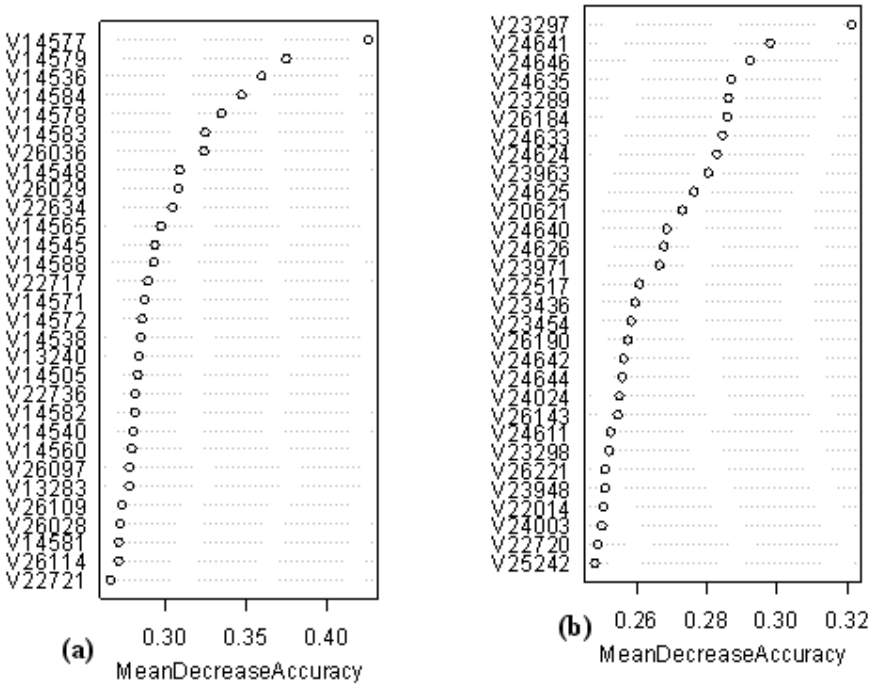



**Figure 4.12 Relative importance plot of variables that were identified as important by RF models for predicting damage (a) non-polar phase stipe tissue and (b) non-polar phase gills tissue**

The signals of importance, as indicated by the RF models, were examined by univariate statistical methods (ANOVA) to determine the significance of damage in mushroom samples. The significance of the Damage factor for each important variable (Table 4.3) indicated that the difference in NMR signals between damaged and undamaged samples were significant and therefore important as damage markers in mushrooms.

**Table 4.3 ANOVA table of the effect of damage on each important variable as indicated by the RF model for all non-polar phase samples**

| NMR Signal | Factor | P-Value | Significance Level |
|---|---|---|---|
| 14579 | Damage | $4.4 \times 10^{-4}$ | *** |
| 14582 | Damage | $0.2 \times 10^{-2}$ | ** |
| 13240 | Damage | $0.6 \times 10^{-2}$ | ** |
| 14505 | Damage | $0.9 \times 10^{-2}$ | ** |
| 22721 | Damage | $1.5 \times 10^{-9}$ | *** |

Significance levels at 95% (*), 99%(**), 99.9% (***)

#### *4.3.4.2. Partial Least Squares*

PLS-DA models were developed to discriminate between undamaged and damaged mushrooms of all tissue types and sample phases (i.e. polar/non-polar) separately. Table 4.4 shows the performance statistics of the models for each tissue.

**Table 4.4 Performance statistics of PLS-DA models built using NMR data**

| Model | #LV | Sensitivity[a] (%) | Specificity[a] (%) | Sensitivity[b] (%) | Specificity[b] (%) |
|---|---|---|---|---|---|
| NP caps | 4 | 98 | 97 | 95 | 96 |
| NP gills | 7 | 95 | 94 | 83 | 92 |
| NP stipes | 5 | 89 | 89 | 73 | 80 |
| Polar caps | 3 | 89 | 84 | 75 | 64 |
| Polar gills | 3 | 78 | 80 | 68 | 73 |
| Polar stipes | 4 | 85 | 66 | 79 | 44 |

#LV: Number of latent variables used in the model
[a]: Training Set
[b]: Testing Set

The PLS-DA model produced for non-polar phase caps achieved the best classification for damage in mushrooms with high sensitivity (i.e. percentage of samples correctly classified) and high specificity (i.e. percentage of samples from the other classes that are well classified by the model). Overall non-polar phase data models achieved better classification than polar phase data, with polar phase gills and stipes achieving the highest

misclassification of samples. These results showed that a high percentage of the spectra of the mushrooms that had been damaged were correctly classified.

When the models were applied to the test set of spectra, sensitivity and specificity were still high with a low misclassification error rate. This trend was not seen for the other tissue groups, with the sensitivities decreasing in the test set models, particularly for polar phase groups (caps and stipes).

## 4.4. CONCLUSIONS

NMR spectroscopy coupled with chemometric tools had the ability to predict low levels of damage in mushrooms. Non-polar phase spectra revealed visible differences between damaged and undamaged. RF models were able to predict damage with OOB error rates of 10% (stipe tissue). VIP plots indicated signals in the lipid and aromatic region of the spectra as being an important area for detecting damage. However, an inability to identify specific metabolite peaks indicated that further work in the extraction process would be required. NMR spectroscopy coupled with PLS-DA yielded models with very low error rates and could therefore be used for modelling damage in mushrooms.

Visible differences were not seen as clearly between damaged and undamaged polar phase spectra. However, when coupled with chemometric multivariate data analysis, RF models were produced. The majority of models produced high error rates and could not be used to successfully model damage in mushrooms. PLS-DA models were able to predict damage with low error rates.

# 5. METABOLIC PROFILING OF MUSHROOMS (*AGARICUS BISPORUS)* USING GC/MS & CHEMOMETRICS TO IDENTIFY MARKERS OF DAMAGE AND INVESTIGATION OF METABOLIC PATHWAYS

# 5. Metabolic profiling of mushrooms (*Agaricus bisporus*) using GC/MS & chemometrics to identify markers for damage and investigation of metabolic pathways

The aim of this study was to use a metabolic profiling approach using gas chromatography mass-spectrometry coupled with chemometric methods to profile damaged and undamaged mushrooms and to identify specific metabolites that could be used as markers of damage in mushrooms.

## 5.1. Material and Methods

### *5.1.1. Mushroom Treatments*

Second flush mushrooms were grown at the Teagasc Research Centre Kinsealy (Dublin, Ireland) and harvested damage-free. A set of 120 closed cap mushrooms ($n$ = 120) were selected for this study. A subset ($n$ = 80) was subjected to physical damage using a mechanical shaker set at 300 rpm. Two damage levels were studied, damage after 20 minutes (D20) which gives these mushrooms an L-value of 86 and places them in a good quality category and damage after 40 minutes (D40) which gives these mushrooms an L-value of 73 and places them in a poor quality category (Gormley and O'Sullivan, 1975). For each damage level twenty mushrooms were analysed. Mushroom samples were analysed on day zero and after 24 hours. The remaining samples (40) were labelled as undamaged (UD), a set of twenty for day zero and twenty for day one. All tissues i.e. cap, gills and stipes were analysed separately in this study. All sample handling was carried out with the utmost care to avoid damage to the mushroom samples. A summary of the experimental design is shown in Table 5.1.

**Table 5.1 Summary of samples analysed by GC/MS**

| Age (Day) | Damage (Minutes) | Tissue | Number of Samples | Number of Spectra (polar & non-polar) |
|---|---|---|---|---|
| 0 | 0 | Cap | 20 | 40 |
| | | Gills | 20 | 40 |
| | | Stipes | 20 | 40 |
| 0 | 20 | Cap | 20 | 40 |
| | | Gills | 20 | 40 |
| | | Stipes | 20 | 40 |
| 0 | 40 | Cap | 20 | 40 |
| | | Gills | 20 | 40 |
| | | Stipes | 20 | 40 |
| 1 | 0 | Cap | 20 | 40 |
| | | Gills | 20 | 40 |
| | | Stipes | 20 | 40 |
| 1 | 20 | Cap | 20 | 40 |
| | | Gills | 20 | 40 |
| | | Stipes | 20 | 40 |
| 1 | 40 | Cap | 20 | 40 |
| | | Gills | 20 | 40 |
| | | Stipes | 20 | 40 |

### *5.1.2. Metabolic Profiling Protocol and Overview*

Sample preparation involved the manual dissection of each mushroom into its three main tissue types (cap, gills and stipes) before freezing overnight at -70 ºC in a cryogenic fridge (Polar 340V Cryogenic fridge, Angeelantoni Industrie spA, Massa Martana, Italy). Once frozen, extraction and fractionation was carried out. Methoxyamination of carbonyl moieties followed by derivatisation of acidic protons with N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) prior to GC/MS analysis was performed as described by Fernie and Lisec with minor modifications (Lisec *et al.*, 2006; Fernie, 2007), following private communication with the authors. An overview of the protocol is described in Figure 5.1.

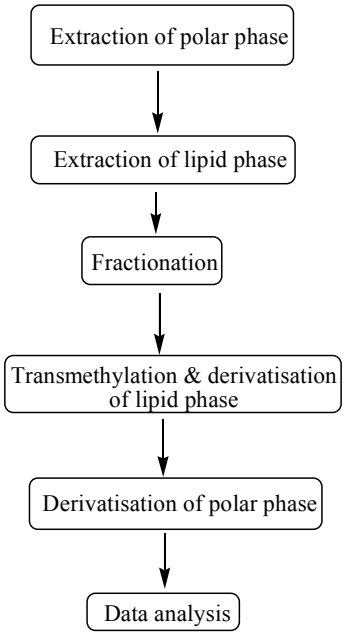



**Figure 5.1 Overview of metabolic profiling (GC/MS) protocol**

### *5.1.2.1. Extraction of Polar Phase*

200 mg of frozen mushroom tissue and 1 ml of methanol (Sigma Aldrich, Dublin, Ireland) were added to an eppendorf tube. The sample was then mixed using a vortex mixer for 10 sec. The methanol was used to inhibit enzymatic processes in the mushroom sample. Two internal standard (IS) compounds were added to the eppendorf tube, one polar (ribitol) and one non-polar (nonadecanoic acid) (Figure 5.2). The internal standard solutions were 50 µl of 0.2 mg / ml distilled water solution of ribitol (Sigma Aldrich, Dublin, Ireland) and 50 µl of 0.2 mg / ml of $CHCl_3$ solution of nonadecanoic acid (Sigma Aldrich, Dublin, Ireland). The contents of the eppendorf tube were mixed again using a vortex mixer for 10 sec.

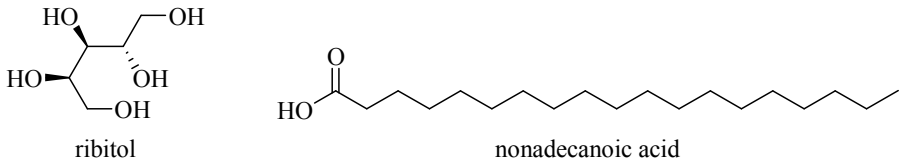



**Figure 5.2 Structures of internal standards injected during the metabolic profiling study**

The sample was then placed in a shaking bath for 15 min at 70 ºC. The different phases were then separated using a Micro Centrifugette 4212 (Medical Supply Co. LTD., Dublin, Ireland) at 14,000 rpm (rotations per minute) for 5 min. The supernatant (polar phase) was transferred from the eppendorf tube using a 200 µl micropipette and placed in a Teflon tube, and 1 ml of distilled water was added and mixed using the vortex.

#### *5.1.2.2. Extraction of Lipid Phase*

An aliquot of 750 µl of chloroform (Sigma Aldrich, Dublin, Ireland) was added to the pellet (lipid phase) and the eppendorf was shaken in a water bath at 37 ºC for 5 min. After centrifugation at 14,000 rpm for 5 min the supernatant was transferred to the same Teflon tube as the polar phase and homogenised with methanol, distilled water and chloroform using the vortex. The remaining phase was discarded.

#### *5.1.2.3. Fractionation*

The polar phase was separated from the lipid phase into a new eppendorf tube and dried in a freeze-dryer (Micro-modulyo, EC Apparatus Inc, New York, USA) for 24 h.

#### *5.1.2.4. Transmethylation & Derivatisation of Lipid Phase (non-polar)*

To extract the lipid phase 900 µl $CHCl_3$ and 1 ml MeOH solution containing 3 % v/v $H_2SO_4$ (Sigma Aldrich, Dublin, Ireland) were added to the Teflon tubes. The lipids and free fatty acids were transmethylated for 4 h in an oil bath at 100 ºC.

The next step involved removing the remaining polar phase. 4 ml of distilled water was added to the Teflon tube. After homogenisation using a vortex mixer and centrifugation at 4,000 rpm for 15 min, the water phase was removed using a pipette. This procedure was conducted twice.

The lipid phase was transferred to a glass vial. The vial was left unscrewed for 24 h or more to allow the chloroform to evaporate. After the evaporation, 10 µl of methoxyamine hydrochloride solution (20 mg / ml pyridine) and 10 µl of the silylation agent N-methyl-N-trimethylsilyltrifluoracetamide (MSTFA, Sigma Aldrich, Dublin, Ireland) (Figure 5.3) were added to the vials. After silylation for 30 min at 37 ºC, 1 µl was injected into the GC/MS. A total of 360 non-polar injections were taken.

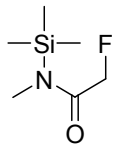



**Figure 5.3 Structure of the silylating agent N-methyl-N-trimethylsilyl-fluoroacetamide (MSTFA)**

*5.1.2.5. Derivatisation of Polar Phase*

When the polar phase was dried, 50 µl of methoxyamine hydrochloride (20 mg/ml pyridine) was added and the solution was mixed using a vortex mixer. The sample was placed for 90 min in a shaking water bath at 30 ºC. Then 80 µl of MSTFA was added and the sample placed in a water bath at 37 ºC for 30 min. The sample was stored at room temperature for 120 min and 1 µl was injected into the GC/MS. A total of 360 polar injections were taken.

*5.1.3. Analysis of Metabolites by GC/MS*

The polar and non-polar samples were analysed similarly using a Varian CP-3800 gas chromatograph coupled to a Varian Saturn 2200 quadrupole MS (JVA Analytical Ltd., Dublin, Ireland). Chromatography was performed on a Cp-sil 24 CB low bleed/MS capillary column (length 30 m, diameter 0.25 mm and film thickness 0.25 µl) using helium at 1.0 ml/min. Samples (1 µl) were injected into a programmed temperature ramp with a

split of 5:1. The GC temperature was initially 70 ºC for 5 min, increased at 3 ºC/min to 300 ºC and then isothermal for 5 min. The GC/MS interface temperature was 220 ºC. MS acquisition conditions were electron impact (EI) ionisation at 70 eV, solvent delay of 1 min, source temperature of 200 ºC, and mass range of 45-650 amu (atomic mass unit) at 2 scans/sec. Data were acquired using the Saturn software (Saturn GC/MS WS Ver 5.5, Varian Inc, USA).

### *5.1.4. Data Analysis*

A number of raw GC/MS data files were selected as representative examples for both polar and non-polar metabolites. These files were used with the Automated Mass Spectral Deconvolution and Identification System (AMDIS, V2.1, NIST, USA) software package to verify the presence of individual analytes and to deconvolute co-eluting peaks. Specific ion characteristics of each metabolite were selected to be used for compound detection in processing methods. Compounds were identified by analysis of standards, comparison with MS libraries and literature data.

Typical chromatograms obtained in this study are shown in Figures 5.4, 5.5 and 5.6. During the analysis of the chromatograms the following steps were taken.

- The chromatogram components were deconvoluted and the baseline noise subtracted with representative MS spectra selected (Figures 5.4, 5.5 and 5.6)
- In a preliminary search with a reduced number of sample chromatograms, the representative spectrum of every component was compared with the NIST library of MS spectra (NIST Mass Spectral Search Programme Version 1.7a, USA, 2001).
- Compounds identified which yielded weighted probabilities of over 70% were compiled in a library for automated batch search, which is an acceptable level to avoid false positives as reported by Norli and colleagues (Norli *et al.*, 2010).

- An automated analysis to report the presence and quantification of all the compounds in the built library was performed by analysis of MS spectra and retention index. The quantification of the compound concentration was done through the use of internal standard area and the known concentration of this internal standard.
- A matrix table with the concentration of each selected library metabolite in each of the samples was produced in the batch job, complete with sample information (i.e. flush, phase, tissue, storage age and damage level) and used for multivariate analysis.

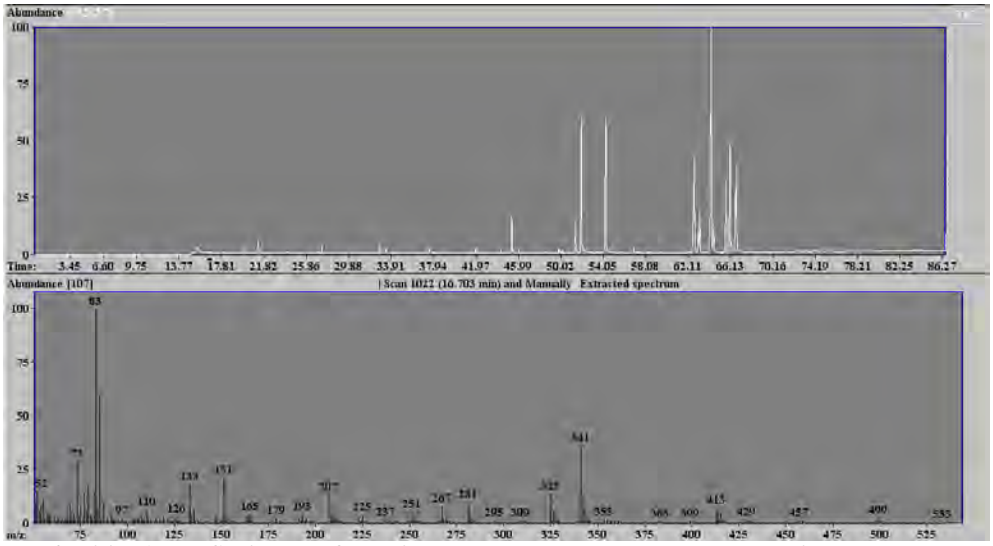



**Figure 5.4 Initial AMDIS window of a typical non-polar phase mushroom showing the chromatogram total ion count (TIC) in the top window and a raw MS spectrum in the bottom window**

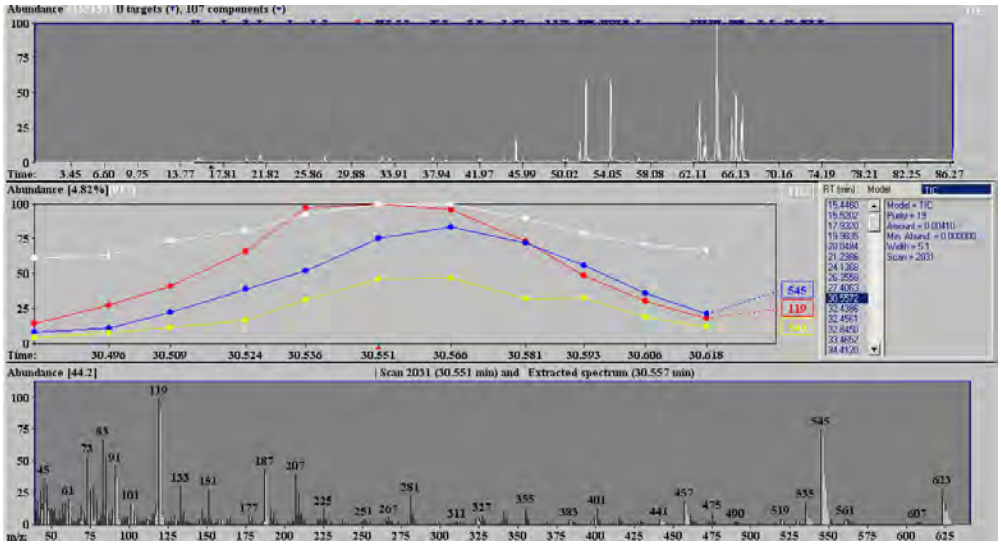


**Figure 5.5 AMDIS component window of a typical non-polar phase mushroom chromatogram showing the chromatogram (top) with all the deconvoluted components identified, the purity window and component information window (middle) and a deconvoluted and noise removed mass spectrum belonging to the selected component (bottom)**

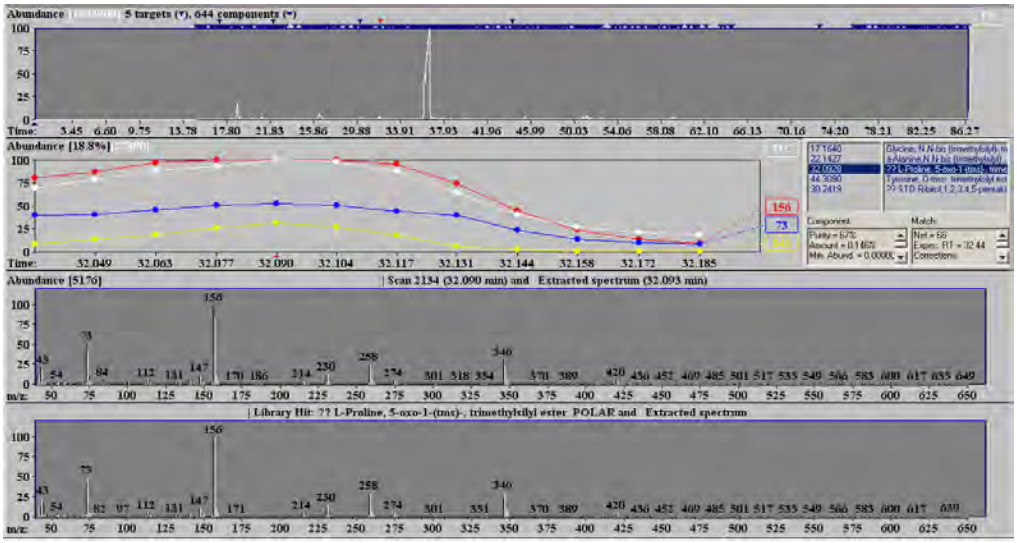


**Figure 5.6 AMDIS component window of a typical polar phase mushroom chromatogram showing the chromatogram (top) with all the deconvoluted components identified, the purity window and component information window (middle) and a deconvoluted and noise removed mass spectrum belonging to the selected component (bottom)**

### *5.1.5. Chemometric Data Analyses*

Principal component analysis (PCA), random forests modelling (RF), Partial least square discriminant analysis (PLS-DA), correlation matrices and univariate statistical analysis were performed using R 2.10.0 (R Development Core Team, 2009).

PCA was used to provide a preliminary observation of the data in order to identify clusters of data and outliers as appropriate. Samples were studied separately on the basis of tissue type and age (day zero/day one).

Random forest (RF) models were built to (a) discriminate between damaged and undamaged mushrooms and (b) to identify specific metabolites as markers for damaged mushrooms. The number of trees fitted to build the random forest was 1000. The number of random metabolites was set at 100 after optimization and the RF model trained was made using a stratified random sampling strategy of the targeted/identified metabolites that would take the same number of metabolites from each of the tissues.

Univariate statistical analysis was carried out on metabolites identified by RF models as being important markers of damage, which included analysis of variance (ANOVA) and Tukey tests.

PLS-DA was applied to the GC/MS data to develop models for the prediction of damage in mushrooms using the Caret (classification and regression training) package in R (R Development Core Team, 2009). The data was split into training and test sets, with 75% of the data used for model training and the remainder used for evaluating model performance i.e. the test set. For more information on the PLS-DA models refer to section 4.2.4.

Correlation matrices were used to determine patterns of correlation between (a) metabolites from the same group of compounds i.e. fatty acids, phenolic compounds, sugars and polyols or amino acids and (b) inter-correlations between metabolites in different groups of compounds. This approach can identify both biosynthetically related and co-ordinately regulated metabolites (Steuer *et al.*, 2003; Dobson *et al.*, 2008)

## 5.2. RESULTS AND DISCUSSION

During the analysis of the 720 chromatograms a library with 105 metabolites was built. Table 5.2 contains 62 metabolites from both polar and non-polar phases. A number (44) of metabolites were not included in the table as they were only found in a very small percentage of chromatograms (<10%). In the non-polar phases (360 chromatograms) fatty acids and phenolics were identified and the internal standard (IS) nonadecanoic acid was found with an average retention time of 56.42 min.

Sugars, polyols and amino acids were identified in polar phase chromatograms (360 chromatograms) and ribitol the polar IS was detected as ribitol, 1,2,3,4,5-pentakis-O-(trimethylsilyl), indicating that derivatisation was successful. This compound was found in the samples at an average retention time of 30.49 min.

### *5.2.1. Non-polar Metabolites*

#### *5.2.1.1. Fatty Acids (FAs)*

Fatty acids are chain-like structures with a carboxylic acid (HO-C=O) at one end and a methyl group ($CH_3$) at the other. The remainder of the compound consists of a hydrocarbon ($CH_2$) chain varying in length from 2-20 or more carbons. Fatty acids have the general formula $C_nH_{2n+1}COOH$. A representative non-polar phase total ion chromatogram containing the fatty acids and phenolics that were present in abundance in *Agaricus bisporus* are identified in Figure 5.7. A number of fatty acids have been previously reported in mushrooms (Byrne and Brennan, 1975; Hiroi and Tsuyuki, 1988; Hong *et al.*, 1988; Senatore *et al.*, 1988; Bonzom *et al.*, 1999; Yilmaz *et al.*, 2006).

**Table 5.2 Metabolites identified by GC/MS as components of *Agaricus bisporus***

| Fatty Acids[1] | Phenolics[2] | Amino Acids[3] | Sugars & Polyols[4] | Organic Acids[5] | Others |
|---|---|---|---|---|---|
| Dodecanoic acid | Benzoic acid[b] | Alanine[b] | D-mannose | Acetic acid[a] | Pyrimidine* |
| Tridecanoic acid | 2-(4-methoxyphenyl)ethanol* | Asparagine | **D-fructose**[ab] | Gluconic acid* | Urea[a6] |
| **Tetradecanoic acid** | **Diphenyl ether**[b] | Glycine | D-ribose[a] | Saccharic acid* | Silanamine* |
| **Pentadecanoic acid** | 2,6-bis(1,1-dimethylethyl)-4-chloro-phenol* | **Aspartic acid** | Erythrose[a] | Succinic[a] | 3-octanol[a7] |
| Hexadecanoic acid[b] | 8-phenyl-6-thio-theophylline* | **Proline**[b] | Sucrose | **Citric acid**[b] | |
| trans-9-Hexadecenoic acid | Tyrosine[b] | Threonine | **D-glucitol/mannitol** | **Phthalic acid*** | |
| **Heptadecanoic acid**[b] | 3,4-dihydroxybenzyl alcohol* | Tryptophan | D-ribo-hexitol | Propanedioic acid* | |
| Octadecanoic acid[b] | 1,3,8-trihydroxy-6-methylanthraquinone[a]* | Valine[b] | **Inositol**[b] | Quinaldic acid[a]* | |
| **9,12-Octadecadienoic acid**[b] | **phenol 2,4-bis(1,1-dimethylethyl)**[a] | Serine[b] | **Glycerol**[b] | 3-acetoxy-3-hydroxypropionic acid* | |
| **Eicosanoic acid**[b] | 4-phenyl-2-hydroxystilbene[a]* | Glutamine[a] | | Pentanedioic acid[a] | |
| Heneicosanoic acid[a] | | | | | |
| Tricosanoic acid[a] | | | | | |
| 11-Eicosanoic acid[a] | | | | | |
| 9,15-Octadecadienoic acid[a] | | | | | |
| Ricinoleic acid[a] | | | | | |
| Erucic acid[a] | | | | | |
| Docosanoic acid[a] | | | | | |
| **Hexanoic acid**[a] | | | | | |

[a] Metabolites found in a low percentage of sample (>10% less than 15%)
[b] Metabolites verified by authentic reference compounds
Compounds identified with very high probabilities (similarity coefficient or reverse similarity coefficient >85%) are indicated in **bold**
[1] Fatty acids identified and reported in literature (Byrne and Brennan, 1975; Yilmaz *et al.*, 2006; Öztürk *et al.*, 2011)
[2] Phenolic compounds identified and reported in literature (Rajarathnam *et al.*, 2003; Kim *et al.*, 2008; Yu *et al.*, 2011)
[3] Amino acids identified and reported in literature (Maggioni *et al.*, 1968; Tseng and Mau, 1999; Tsai *et al.*, 2007; Wagemaker *et al.*, 2007)
[4] Sugars and polyols identified and reported in literature (Tseng and Mau, 1999; Beecher *et al.*, 2001; Kim *et al.*, 2009)
[5] Organic acids identified and reported in literature (Fujita *et al.*, 1991; Valentão *et al.*, 2005)
[6] Urea reported in literature (Wagemaker *et al.*, 2005; Eastwood *et al.*, 2008)
[7] 3-octanol reported in literature (Flegg *et al.*, 1985; Combet *et al.*, 2006)
* Metabolites not previously reported in mushrooms

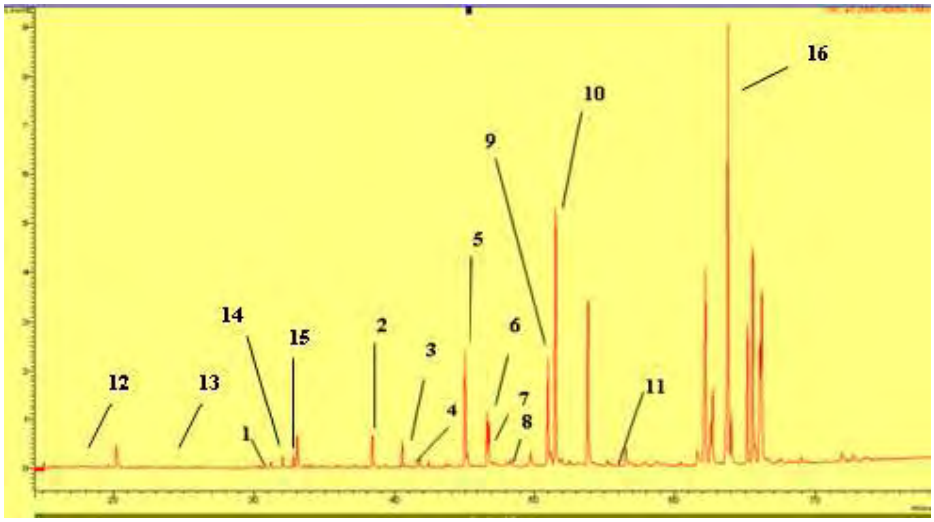



**Figure 5.7 Non-polar phase GC/MS total ion chromatogram (TIC) of a mushroom extract. Numbers in the figure correspond to compounds detailed in Tables 5.3 and 5.4**
Peaks 1-11 fatty acids
Peaks 12-16 phenolics

It should be noted that some fatty acid metabolites (e.g. dodecanoic acid and octadecanoic acid) produced two peaks corresponding to methyl esters and trimethylsily derivatives. Of these two peaks only one was included in the study. However, peaks 5 and 6 represent hexadecanoic acid methyl ester and hexadecanoic acid trimethylsilyl ester respectively (Figure 5.7) and are included to highlight this occurrence. A total of 18 fatty acids were separated and identified.

*5.2.1.1.1. Saturated Fatty Acids*

Saturated fatty acids contain no double bond in their hydrocarbon chain (Hui, 2006). Saturated fatty acids were largely predominant in non-polar phase chromatograms with 14 compounds identified. A number of other fatty acids were identified by mass spectrometry but only occurred in a small percentage of chromatograms (Table 5.2).

Fatty acid metabolites (and all other metabolites) were identified on the basis of their molecular weight, molecular ions, retention times, literature data and comparison to library databases and standards.

The electron impact (EI) mass spectra of saturated fatty acid methyl esters are very similar, having a reasonably abundant molecular ion and a characteristic base peak at *m/z* 74, which is a product of the well-recognised McLafferty rearrangement process (McLafferty, 1959; McCloskey, 1970). Other characteristic ions include an ion at *m/z* 87, the first member of the carbomethoxy ion series ($[(CH_2)nCO_2CH_3]^+$, where $n$ = 1, 2, 3, 4….. *m/z* 73, 87, 101, 115…), a hydrocarbon series of ions that are abundant in the low mass region and an acylium ion $[M-31]^+$ (Zirrolli and Murphy, 1993).

The McLafferty rearrangement ion is central in the identification of most fatty acid ester derivatives and its mechanistic aspects are shown in Figure 5.8. The resulting ion is important for identification purposes.

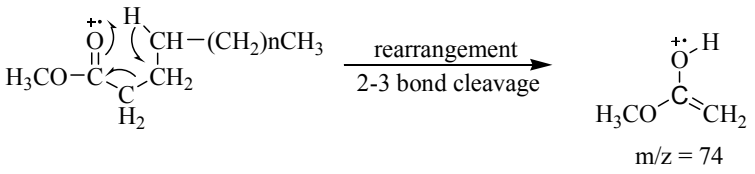



**Figure 5.8 McLafferty rearrangement mechanism of fatty acids**

A site-specific rearrangement is involved where a hydrogen atom from position 4 of the aliphatic chain migrates to the carbomethoxy group, presumably through a six-membered transition state, which is sterically favoured. If one of the hydrogen atoms on carbon 4 is substituted, the McLafferty ion will be noticeable lower than expected. This may explain why it is less evident in the mass spectra of unsaturated fatty acid derivatives with increasing numbers of double bonds, which can readily migrate to position 4 under

electron bombardment (Christie, 2010). The mass spectrum of octadecanoic acid methyl ester is presented in Figure 5.9.

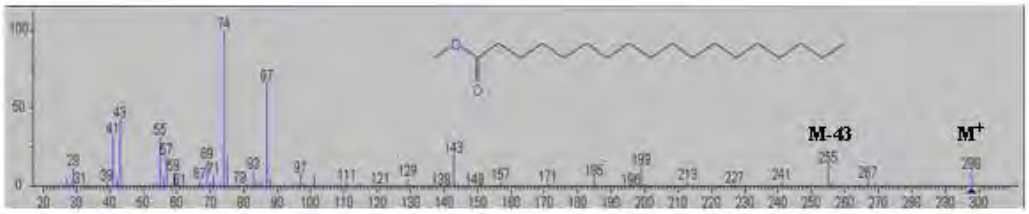



**Figure 5.9 Mass spectrum of octadecanoic acid methyl ester**

Its fragmentation mechanism is characterised as follows. The molecular ion peak of octadecanoic acid methyl ester is found at *m/z* 298 $[M]^+$. After the McLafferty rearrangement and α cleavage *m/z* 74 is produced, which was the base peak ion of the C6-C26 saturated fatty acid methyl ester. The peak at *m/z* 255 $[M-43]^+$ was the result of C-C bond cleavage and the loss of $C_3H_7^+$. These cleavage processes are shown in Figure 5.10.

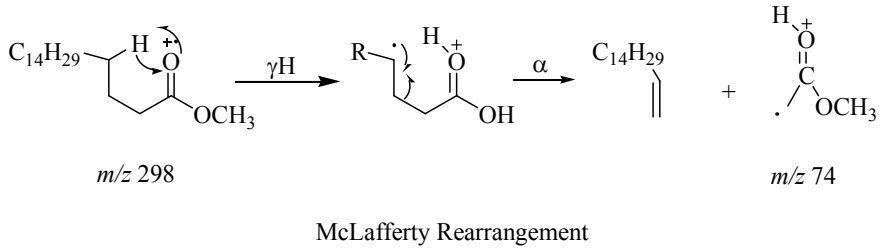



**Figure 5.10 Cleavage process of octadecanoic acid methyl ester**

The series of related ions giving rise to peaks at *m/z* 43, 55, 74 and 87 is formed by loss of neutral aliphatic radicals i.e. $CH_2$ fragmentation. The characteristic ions of saturated fatty acid esters are therefore *m/z* 74 (McLafferty rearrangement), $[M-43]^+$ the molecular ion $[M]^+$ and the series of ions resulting from $CH_2$ fragmentation (Zirrolli and Murphy, 1993; Cheung *et al.*, 1994; Wu *et al.*, 2007).

The mass spectrum of dodecanoic acid (Figure 5.11), shows a molecular peak at *m/z* 214 which corresponds to the molecular weight of the compound.

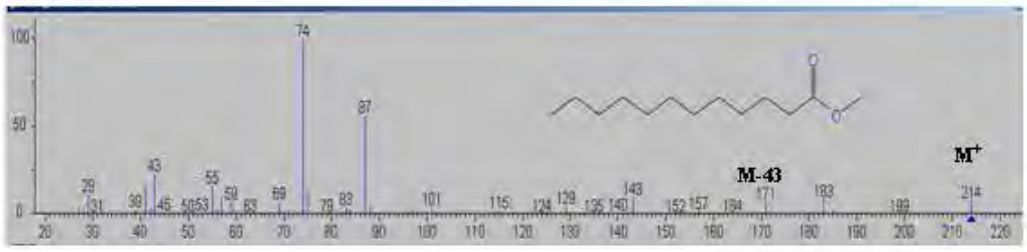



**Figure 5.11 Mass spectrum of dodecanoic acid methyl ester**

The molecular ion peak at *m/z* 214, the peak at *m/z* 171 $[M-43]^+$, the presence of the peaks at *m/z* 74 (McLafferty ion) and the series of related peaks at *m/z* 87, 101, 55 and 41 ($CH_2$ fragmentation) confirm this metabolite to be dodecanoic acid methyl ester (Wu *et al.*, 2007).

*5.2.1.1.2. Unsaturated Fatty Acids*

An unsaturated fatty acid is a fatty acid that contains at least one double bond within the fatty acid chain (Hui, 2006). A total of 4 unsaturated fatty acids were separated and detected by GC/MS; Linoleic acid, ricinoleic acid, erucic acid and palmitelaidic acid.

Linoleic acid (9, 12 octadecadienoic acid) was detected and identified in non-polar mushroom extract chromatograms with an average retention time of 52.56 min. It is the main unsaturated FA of mushroom lipids and is the precursor of the mushroom alcohol (1-octen-3-ol) (Mau *et al.*, 1992). The molecular ion peak at *m/z* 294, the peak at *m/z* 263 $[M-31]^+$, the peak at *m/z* 67 (the result of double bond transfer) and the series of related peaks at *m/z* 41, 55, 67, 81, 95, 109 and 123 ($CH_2$ fragmentation) confirm this metabolite to be linoleic acid methyl ester (Wu *et al.*, 2007; Christie, 2010).

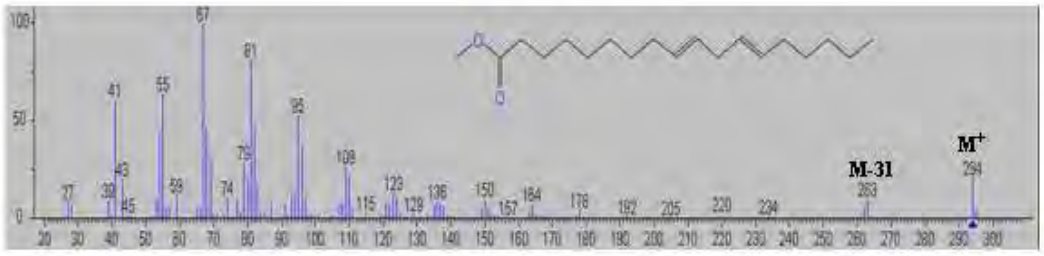



**Figure 5.12 Mass spectrum of linoleic acid methyl ester**

The gas chromatographic retention times and mass spectrometric characteristic ions of fatty acid methyl esters detected abundantly in *Agaricus bisporus* are shown in Table 5.3.

**Table 5.3 Mass spectrometric characteristic ions and gas chromatographic retention times of fatty acid methyl esters**

| No | Fatty Acid | Common name | MW | RT (min) | BP | Characteristic ions |
|---|---|---|---|---|---|---|
| **1** | Dodecanoic | Lauric acid (12:0) | 214 | 32.28 | 74 | 74, 171, 214 |
| **2** | Tridecanoic | (13:0) | 242 | 38.03 | 74 | 74, 199, 242 |
| **3** | Tetradecanoic | Myristic acid (14:0) | 300 | 41.04 | 74 | 74, 257, 300 |
| **4** | Pentadecanoic | (15:0) | 256 | 41.15 | 74 | 74, 213, 259 |
| **5** | Hexadecanoic | Palmitic acid (16:0) | 270 | 45.01 | 74 | 74, 227, 270 |
| **6** | Hexadecanoic acid TMS | As above | 328 | 46.66 | 73 | 73, 285, 328 |
| **7** | trans-9-hexadecenoic acid | Palmitelaidic acid (16:1) | 268 | 46.82 | 55 | 55, 236, 268 |
| **8** | Heptadecanoic | Margaric acid (17:0) | 284 | 47.77 | 74 | 74, 241, 284 |
| **9** | Octadecanoic | Stearic acid (18:0) | 298 | 50.74 | 74 | 74, 255, 298 |
| **10** | 9,12-Octadecadienoic acid | Linoleic acid (18:2) | 294 | 52.56 | 67 | 67, 263, 294 |
| **11** | Eicosanoic | Arachidic acid (20:0) | 326 | 55.82 | 74 | 74, 283, 326 |

Peak numbers 1-11 correspond to Figure 5.7
TMS: trimethylsilyl ester; MW: molecular weight; RT: retention time; BP: base peak (ion)

*5.2.1.2. Phenolic Compounds*

Phenolics are compounds possessing one or more aromatic rings with one or more hydroxyl groups and can range from simple molecules (phenolic acids, phenylpropanoids, flavonoids) to highly polymerised compounds (lignins, melanins, tannins) (Bravo, 1998). A number of phenolic compounds found in mushrooms have been previously reported in literature (Ribeiro *et al.*, 2008; Barros *et al.*, 2009; Vaz *et al.*, 2011).

Ten phenolic compounds were identified by mass spectrometry on the basis of their molecular weight, molecular ions, retention times and comparison to library databases and standard compounds. Out of these phenolic compounds seven were present in abundance in non-polar phase chromatograms (Table 5.4); 2,6-bis(1,1-dimethyl ethyl)-4-chloro-phenol, 2-(4-methoxyphenyl)ethanol, 3,4-dihydrobenzyl alcohol, 8-phenyl-6-thio-theophylline, diphenyl ether, tyrosine trimethylsilyl ester and benzoic acid methyl ester. The following three compounds were only present in a low percentage of chromatograms; 1,3,8-trihydroxy-6-methylanthraquinone, phenol 2,4-bis(1,1-dimethylethyl) and 4-phenyl-2-hydroxystilbene. The fragmentation pattern of benzoic acid methyl ester is described below.

Benzoic acid methyl ester had an average retention time of 20.56 minutes. The mass spectrum of benzoic acid methyl ester (Figure 5.13) on EI ionisation shows the molecular ion at *m/z* 136.

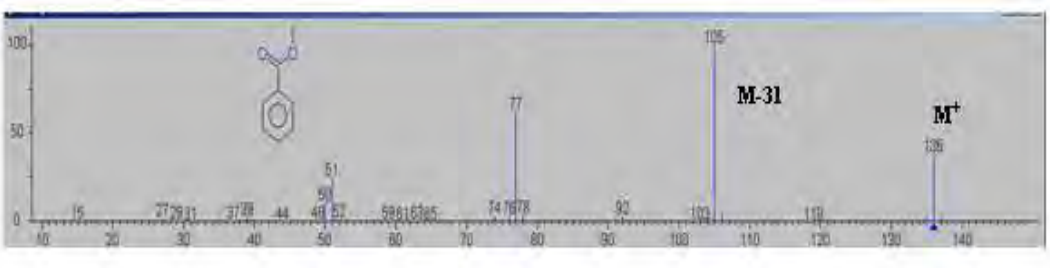



**Figure 5.13 Mass spectrum of benzoic acid methyl ester**

Its fragmentation mechanism is as follows; the molecular ion of benzoic acid methyl ester was *m/z* 136. The base peak is characterised by $C_6H_5CO^+$ benzoyl ions at *m/z* 105. Other prominent fragments are $C_6H_5^+$ ions at *m/z* 77 and $C_4H_3^+$ ions at *m/z* 5 (Opitz, 2007).

**Table 5.4 Mass spectrometric characteristic ions and gas chromatographic retention times of abundant phenolic compounds**

| Peak Number | Phenolic | Molecular Weight | Retention Time (min) | Base Peak | Characteristic ions |
|---|---|---|---|---|---|
| **12** | Benzoic acid | 136 | 17.04 | 105 | 136, 105, 77 |
| **13** | 2-(4-methoxyphenyl)ethanol | 152 | 25.50 | 121 | 152, 121, 77 |
| **14** | Diphenyl ether | 170 | 31.75 | 170 | 170, 141, 77 |
| **15** | 2,6-bis(1,1-dimethyl ethyl)-4-chloro-phenol | 240 | 33.00 | 225 | 240, 225 |
| **16** | 8-phenyl-6-thio-theophylline | 272 | 63.45 | 272 | 272, 243, 211 |
| **17*** | Tyrosine O-trimethylsilyl-,trimethyl ester | 325 | 44.53 | 179 | 73, 179 |
| **18 *** | 3,4-dihydroxybenzyl alcohol | 356 | 35.37 | 73 | 356, 267, 179, 73 |

Peak numbers 12-16 correspond to Figure 5.7
Peak numbers 17-18* correspond to Figure 5.14b
* Metabolites found in polar phase chromatograms

### *5.2.2. Polar Metabolites*

#### *5.2.2.1. Amino acids*

A total of ten amino acids (AA) were separated and identified by GC/MS (Table 5.2). The following AA were found abundantly in polar phase chromatograms (Figure 5.14a & 5.14b): alanine, asparagine, glycine, aspartic acid, proline, threonine, tryptophan, valine and serine. Glutamine was detected in a small percentage of polar phase chromatograms. Amino acids contribute to the flavour of mushrooms with a number being reported in literature as being present in *Agaricus bisporus* (Tseng and Mau, 1999; Tsai *et al.*, 2007).

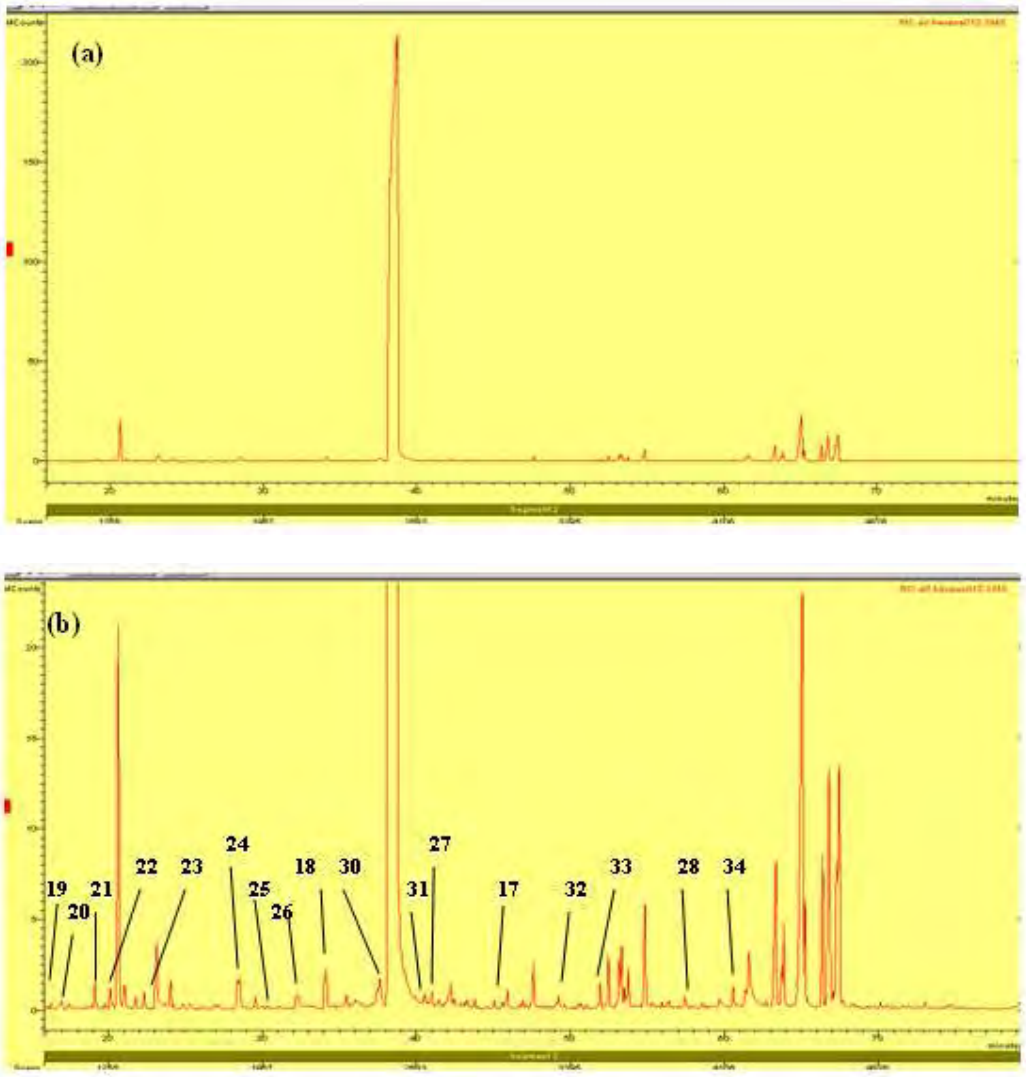



**Figure 5.14 (a) Polar phase GC/MS total ion chromatogram (TIC) of a mushroom extract (b) enlarged version. Numbers in the figure correspond to compounds detailed in Tables 5.5 and 5.6**

The fragmentation mechanism of valine is shown in Figure 5.15, with a table of all the amino acids and their characteristic ions given in Table 5.5. The common characteristic ions for identifying trimethylsilyated amino acids by EI ionisation are [M-$CH_3$], [M-$COCH_3$], [M-$CO_2$TMS] and [M-sidechain] (El-Khoury, 1999).

The molecular ion (*m/z* 261) was not detected in the mass spectrum; [M-15] ion *m/z* 246 was found in small proportion.

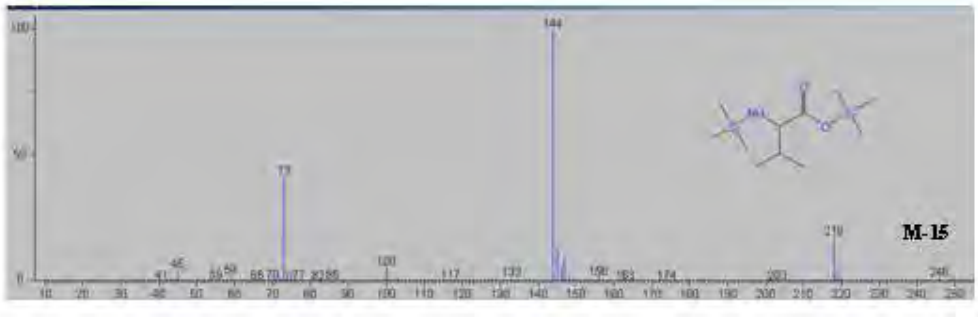



**Figure 5.15 Mass spectrum of the amino acid valine**

**Table 5.5 Mass spectrometric characteristic ions and gas chromatographic retention times of amino acids**

| Peak Number | Amino Acid | Molecular Weight | Retention Time (min) | Base Peak | Characteristic ions |
|---|---|---|---|---|---|
| **19** | L-Valine,N-(TMS)-,TMS ester | 261 | 15.85 | 144 | 246, 144, 73 |
| **20** | Glycine, N,N-bis(TMS)-TMS ester | 291 | 17.58 | 174 | 248, 174, 73 |
| **21** | Serine, bis(TMS)- | 259 | 18.46 | 116 | 132, 116, 73 |
| **22** | L-Threonine,O-(TMS)-,TMS ester | 263 | 19.82 | 73 | 219, 117, 73 |
| **23** | Alanine, N,N-bis(TMS)-TMS ester | 305 | 22.86 | 248 | 248, 174, 73 |
| **24** | L-Aspartic acid, N-(TMS)-,bis(TMS) ester | 349 | 28.53 | 232 | 232, 73 |
| **25** | Glutamine,tris(TMS) | 363 | 31.87 | 246 | 246, 73 |
| **26** | L-Proline (TMS)-,TMS ester | 273 | 32.04 | 156 | 156, 73 |
| **27** | Aspragin,N,N,N-tris(TMS)-,TMS ester | 420 | 41.70 | 188 | 405, 188, 73 |
| **28** | L-Tryptophan,N-1-bis(TMS)-,TMS ester | 420 | 57.26 | 202 | 202, 73 |

TMS trimethylsilyl
Peak numbers 19-28 correspond to Figure 5.14b

*5.2.2.2. Sugars & Polyols*

Nine sugars and sugar alcohols were identified by MS. Table 5.6 shows the major ions appearing in the mass spectra of sugars and sugar alcohols. Those found abundantly in polar phase chromatograms were: D-mannose, sucrose, D-glucitol/mannitol, D-ribo-

hexitol, myo-inositol and glycerol. A number of sugars and polyols are present in mushrooms and have been reported in literature. Mannitol is the most abundant polyol found in mushrooms (Hammond and Nichols, 1975; Beecher *et al.*, 2001; Yang *et al.*, 2001; Kim *et al.*, 2009). Glucitol was identified by MS with a very high probability. It should be noted that glucitol and mannitol are isomers.

D-mannose produced ions at *m/z* 73, 147 and 217 that are characteristic of aldohexases. A molecular ion was not observed but the [M-15] ion was present in small proportions (Figure 5.16a). The sugar alcohol myo-inositol produced characteristic ions at *m/z* 318, 147 and 73 and the parent ion $[M]^+$ was detected at *m/z* 612 (Figure 5.16b) (Reineccius *et al.*, 1970; Schoots and Leclercq, 1979; Roessner *et al.*, 2000; Mederios and Simoneit, 2007).

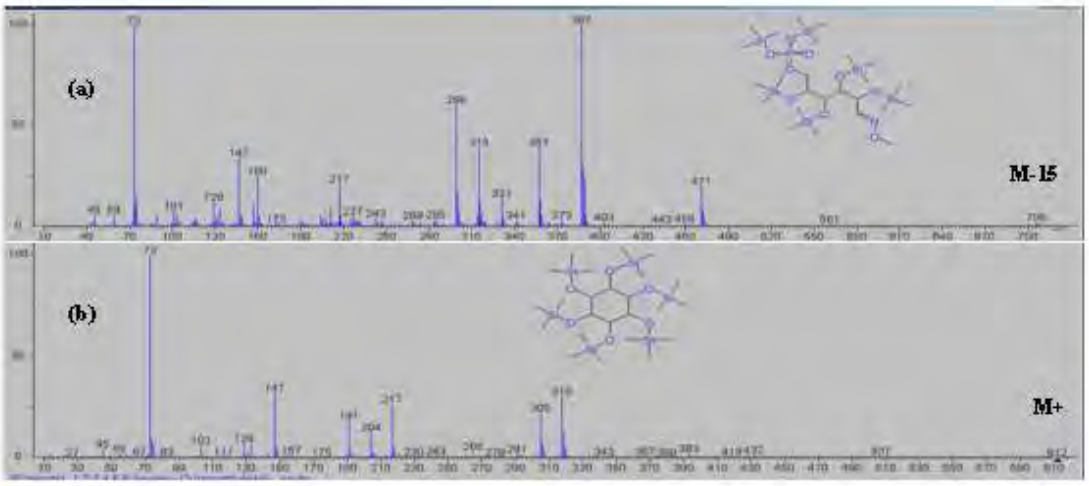



**Figure 5.16 Mass spectrum of methyl-silyated (a) D-Mannose and (b) inositol (Myo-inositol)**

**Table 5.6 Mass spectrometric characteristic ions and gas chromatographic retention times of abundant sugar and polyols**

| Peak Number | Sugar & Sugar Alcohols | Molecular Weight | Retention Time (min) | Base Peak | Characteristic ions |
|---|---|---|---|---|---|
| **29*** | Glycerol | 308 | 15.59 | 73 | 293, 147, 73 |
| **30** | D-ribo-hexitol | 526 | 38.50 | 73 | 333, 231, 73 |
| **31** | Myo-inositol | 612 | 40.75 | 73 | 318, 147, 73 |
| **32** | D-glucitol/mannitol | 766 | 49.45 | 73 | 751, 215, 73 |
| **33** | D-mannose | 721 | 52.34 | 73 | 706, 387, 73 |
| **34** | Sucrose | 918 | 60.16 | 361 | 361, 217, 73 |

Peak numbers 29-34 correspond to Figure 5.14b
*Not visible on chromatogram as peak too small

### *5.2.3. Principal Component Analysis*

Samples were studied separately on the basis of their tissue type i.e. caps, gills and stipes and also on their age i.e. day zero and day one. The score plot for day zero caps is shown in Figure 5.17a for PC1 *versus* PC2. These first two principal components accounted for 50 and 40%, respectively of the total variance in the GC/MS data set, and some samples clustering on the basis of damage is readily apparent. The majority of undamaged caps (blue) formed a cluster on the left hand side of the centre of the plot, with D20 (pink) caps forming a cluster to the right hand side of the centre. D40 samples were spread randomly throughout the score plot with some samples found in UD and D20 clusters.

In the case of day one caps (Figure 5.17b) a pattern can be seen in the score plot for PC1 *versus* PC2 (accounting for 64 and 25% of the total variance) and again clusters have formed for UD and D20 samples indicating that metabolite levels are affected by damage.

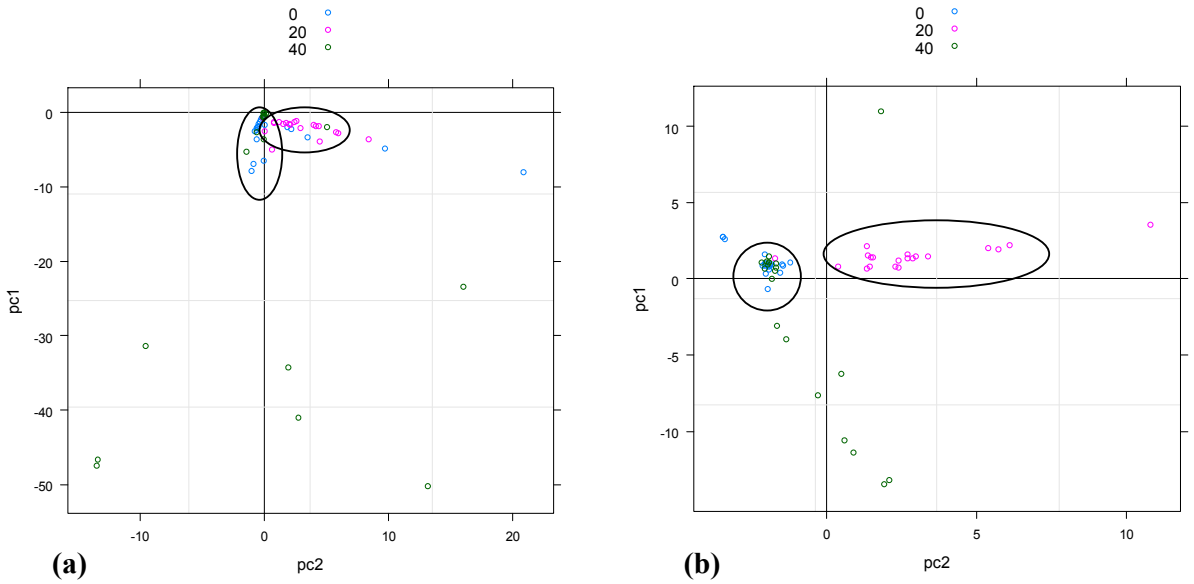



**Figure 5.17 PC1 *versus* PC2 score plots of cap tissue for (a) Day zero samples and (b) Day one samples**
0: Undamaged
20: 20 min damage
40: 40 min damage

In the case of gill and stipe tissue clusters for different damage levels were not clearly evident, with overlapping of damage levels seen for both day zero and day one.

### *5.2.4. Detection of Damage*

#### *5.2.4.1. Random Forests (Model 1: all data used)*

The first model developed attempted to identify specific metabolites that could be used as possible markers for damage in mushrooms. The model tried to predict damage in mushrooms using all metabolites identified by GC/MS, a variable indicating the tissue from which the metabolite originated (cap, gill, or stipe) and the age of the mushroom (day zero/day one) as explanatory variables. This resulted with an out-of-bag (OOB) error rate of 11.11%, sensitivity of 88.9% and specificity of 92%. The variable of importance (VIP) plot for predicting damage (Figure 5.18) identified pentadecanoic acid, linoleic acid, myo-inositol, benzoic acid and hexadecanoic acid as the five most important metabolites as damage markers. The variables tissue and day were identified in the VIP plot as the 13$^{th}$ and 22$^{nd}$ most important variables respectively. By removing the variable age and tissue another model was built that took the identified metabolites in mushrooms and tried to predict whether there was damage or not. This RF could be used as a classifier of mushroom damage and gave a good prediction model with an OOB error rate of 11.39%, sensitivity of 88.6% and specificity of 92% (similar to the previous model). The important variables (top five) identified were the same as Figure 5.18.

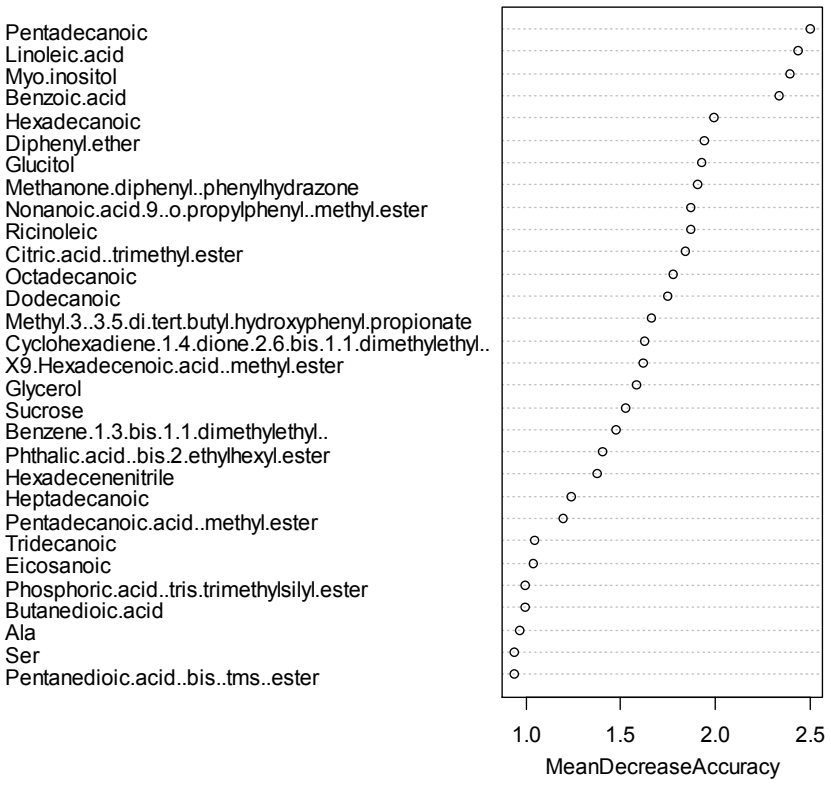



**Figure 5.18 VIP plot of metabolites that are important variables in the RF model for predicting damage**

The metabolites of importance as indicated by the RF model were examined by univariate statistical methods (ANOVA and Tukey tests) to determine the significance of damage for each of the three damage levels. A box plot (Figure 5.19) indicated that there was an increase in the quantity of myo-inositol from UD to D40 samples (undamaged samples to samples that had been exposed to high levels of damage).

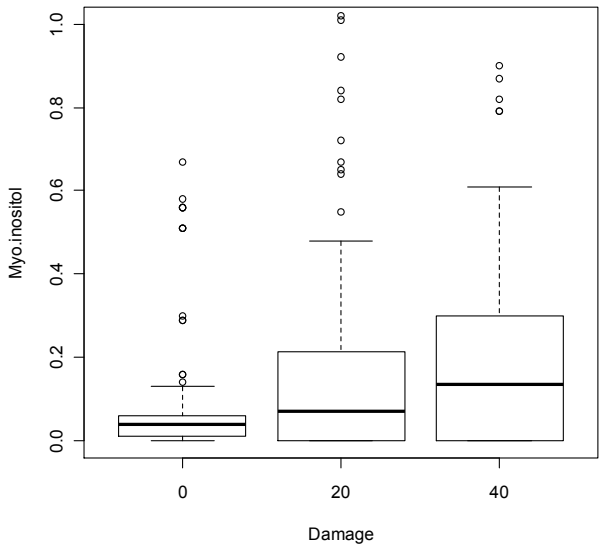



**Figure 5.19 Box plot showing the quantity of myo-inositol at each damage level**
0: Undamaged
20: 20 min damage
40: 40 min damage

Myo-inositol is found in *Agaricus bisporus* in low quantities compared to other sugar alcohols, such as mannitol. Recent studies showed that levels of myo-inositol (like mannitol) increased with maturation (Tsai *et al.*, 2007). The increased levels in damaged samples also suggest that myo-inositol is affected by mechanical damage and may be a metabolic marker of damage. Sugars play important roles in all aspects of mushroom life. They provide the main respiratory substrates for the generation of energy and metabolic intermediates that are then used for the synthesis of macromolecules and other cell constituents. The sugars ribose and deoxy-ribose form part of the structure of DNA and RNA. Polysaccharides such as chitin and glucan are the major structural elements of mushrooms cell walls (Ruiz-Herrera, 1992). A linkage to sugar is required for proper functioning of many lipids and proteins and therefore as a consequence, the abundance and depletion of sugars, polyols and their derivatives initiate various responses in mushrooms and have profound effects on mushroom metabolism, growth and depletion (Yu, 1999).

The phenolic compound benzoic acid was identified as an important metabolite by the RF model for differentiating between damaged and undamaged mushrooms. A number of phenolic compounds including benzoic acid have been identified in *Agaricus bisporus* and other mushrooms (Rajarathnam *et al.*, 2003; Kim *et al.*, 2008). Browning is a reaction that occurs when polyphenol oxidase (PPO) acts on a phenolic compound in the presence of oxygen to produce a dark colour (Martinez and Whitaker, 1995). Reports have found that the majority of phenolic compounds are present in mushroom skin rather than the flesh (Rajarathnam *et al.*, 2003). In this study the benzoic acid content was higher in stipe and gill tissue compared to cap tissue. Levels of benzoic acid were at their highest at 20 min damage for each of the tissue types. The levels decreased substantially in gills and stipe at 40 min damage (Figure 5.20) and to a lesser extent in the cap tissue suggesting that the browning reaction had come to completion i.e. the mushroom was completely damaged with levels of benzoic acid becoming depleted.

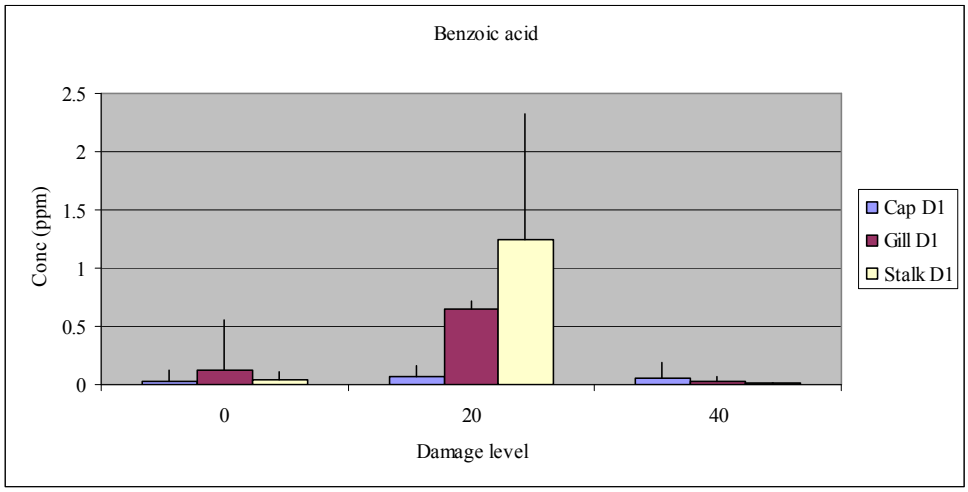



**Figure 5.20 Bar plots with the semi-quantitative concentration of benzoic acid in mushroom tissues at different damage levels in day one samples. A similar trend was found in day zero samples**

The fatty acids pentadecanoic acid, linoleic acid and hexadecanoic acid were identified as important variables by the model for detecting damage. Linoleic acid is the most abundant fatty acid found in *Agaricus bisporus* (Yilmaz *et al.*, 2006) and is the precursor of the

mushroom alcohol 1-octen-3-ol, which is considered to be responsible for the characteristic flavour obtained by the physical disruption of the edible mushroom. The biosynthesis of 1-octen-3-ol is due to aerobic oxidation (Holtz and Schisler, 1971) by lipoxygenase (LOX) of linoleic acid into regio- and stereo-specific hydroperoxides (HPODs), followed by an enzymatic cleavage by hydroperoxide lyase (HPL) of the corresponding HPODs to produce 1-octen-3-ol (Tressl *et al.*, 1982; Chen and Wu, 1984; Mau *et al.*, 1992; Assaf *et al.*, 1995). The enzymatic pathway in which linoleic acid is converted into 1-octen-3-ol and 10-oxo-*trans*-8-decenoic acid by *Agaricus bisporus* has been reported (Wurzenberger and Grosch, 1984; Combet *et al.*, 2006).

Lipids afford the means for fundamental metabolic processes and provide the basic composition for cell membranes. The primary function of fatty acids is as an energy reserve (Karlinski *et al.*, 2007) and they also play an important role in storage since most lipids are in the cell envelope which undergoes change as the mushroom deteriorates, either by damage or over time.

The significance of the Damage factor for each metabolite (Table 5.7) indicated that the difference in metabolite concentration between damaged and undamaged samples was significant and therefore these important metabolites, as indicated by RF models, may be markers of damage in mushrooms.

**Table 5.7 ANOVA table of the effect of damage on individual metabolites**

| Metabolite | Factor | P-Value | Significance Level |
|---|---|---|---|
| Linoleic acid | Damage | $2.0 \times 10^{-6}$ | *** |
| Myo-inositol | Damage | $5.2 \times 10^{-4}$ | ** |
| Pentadecanoic acid | Damage | $1.7 \times 10^{-9}$ | *** |
| Benzoic acid | Damage | $1.5 \times 10^{-9}$ | *** |
| Hexadecanoic acid | Damage | $9.7 \times 10^{-3}$ | ** |

Significance levels at 95% (*), 99% (**), 99.9% (***)

Tukey tests were carried out on the five important variables to determine which means among a set of means differ from the rest. The different samples ages (day zero and day one) were examined separately and the Tukey plots (Figure 5.21) showed that there were

significant differences in mean levels of damage for myo-inositol after day one. This trend was seen for all identified metabolites.

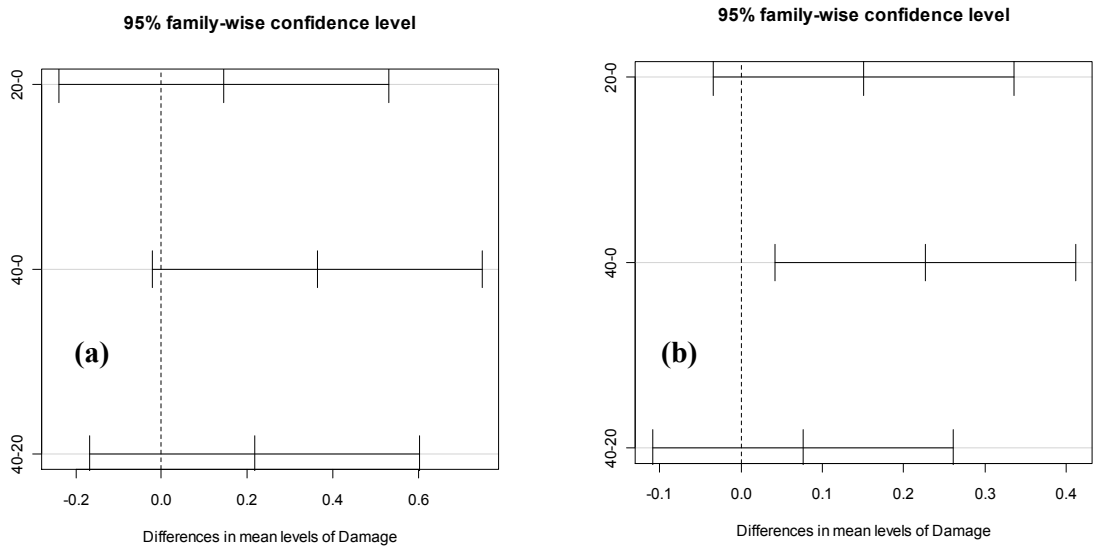



**Figure 5.21 Tukey multiple comparison test plots comparing differences in mean levels of damage in (a) Day zero samples and (b) Day one samples (myo-inositol)**
0: Undamaged
20: 20 min damage
40: 40 min damage

*5.2.4.1.1. Random Forests (models 2-6)*

RF models were produced separately for both day zero and day one samples. As before the model tried to predict damage in samples using the metabolites identified and separated by GC/MS and a variable indicating the tissue from which each sample came from as explanatory variables. The RF model for day zero samples (model 2) produced an OOB error rate of 10% and the variables identified by the model as being the most important for detecting damage were the following metabolites; linoleic acid, nonanoic acid, diphenyl ether, hexadecanoic acid and pentadecanoic acid.

The RF model for day one samples (model 3) produced an OOB error of 6.67% with the VIP plot indicating the following metabolites as the most important variables in the model

to discriminate between damage and undamaged mushrooms; phthalic acid, myo-inositol, pentadecanoic acid, glucitol and linoleic acid.

It has been previously reported that the metabolic response in the form of enzyme expression in mushrooms to both age (Mohapatra *et al.*, 2008) and damage (O'Gorman *et al.*, 2010) is delayed in time and it takes at least one day to develop. Therefore metabolite identification with day one samples is important in the sense of examining indicators of damage/aged metabolism, whereas the analysis of day zero samples will be useful for finding early indicators of damage (before it is perceived by the consumer).

Individual tissues were then subjected to RF modelling. Each tissue (model 4-6) was examined once (day zero and day one) to determine which metabolites were important in the model. A summary table of the OOB error rates for the tissues and VIPs are shown in Table 5.8 including previous models for damage.

**Table 5.8 Summary of all RF models for predicting damage, including OOB error rates, variables of importance and associated metabolic pathways**

| RF Model | OOB (%) | Important variables (top five) | Metabolic pathways involved |
|---|---|---|---|
| Model 1 (all samples) | 11.11 | Pentadecanoic acid | β-oxidation |
| | | Linoleic acid | Lipoxygenase pathway (LOX) |
| | | Myo-inositol | Isomerisation of glucose |
| | | Benzoic acid | |
| | | Hexadecanoic acid | β-oxidation |
| Model 2 (day zero samples ) | 10 | Linoleic acid | LOX |
| | | Nonanoic acid | β-oxidation |
| | | Diphenyl ether | |
| | | Hexadecanoic acid | β-oxidation |
| | | Pentadecanoic acid | β-oxidation |
| Model 3 (day one samples) | 6.67 | Phthalic acid | |
| | | Myo-inositol | Isomerisation of glucose |
| | | Pentadecanoic acid | β-oxidation |
| | | Glucitol/Mannitol | Polyol metabolism |
| | | Linoleic acid | LOX |
| Model 4 (Cap) | 8.33 | Linoleic acid | LOX |
| | | Hexadecanoic acid | β-oxidation |
| | | Heptadecanoic acid | β-oxidation |
| | | D-mannose | Glycolysis |
| | | Glycerol | β-oxidation |
| Model 5 (Gills) | 13.33 | Myo-inositol | Isomerisation of glucose |
| | | Glucitol/Mannitol | Polyol metabolism |
| | | Citric acid | Product of TCA cycle |
| | | Benzoic acid | |
| | | D-Ribo hexitol | Polyol metabolism |
| Model 6 (Stipe) | 9.17 | Benzoic acid | |
| | | Propanedioic acid | |
| | | Serine | Proteolysis |
| | | D-Ribo hexitol | Polyol metabolism |
| | | Ricinoleic acid | β-oxidation |

OOB: Out of bag error rate

*5.2.4.2. Patrial Least Squares (PLS)*

A PLS-DA model was developed to discriminate between the three levels of damage including all data i.e. all tissue types and mushroom ages (day zero & day one). Accuracy was used to select the optimal model i.e. appropriate number of latent variables to be used in the model and the number used was 4 (Figure 5.22).

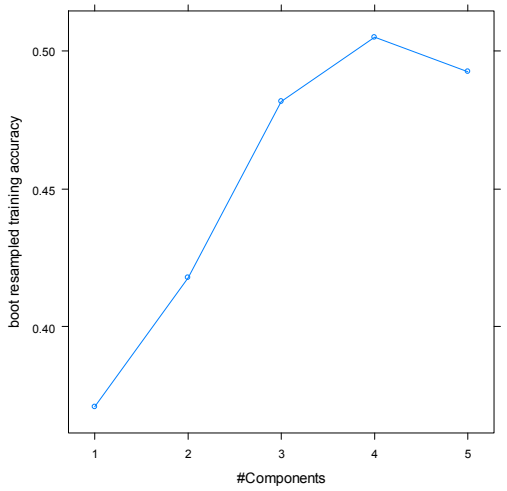



**Figure 5.22 Evolution of bootstrap resampling accuracy as a function of latent variables**

The initial PLS model built used all the data i.e. all tissues, damage levels and days to evaluate if it could differentiate between damaged and undamaged samples with high sensitivity and specificity. The results are presented in the Table 5.9 which shows the values of sensitivity (i.e. percentage of samples correctly classified as such) and specificity (i.e. percentage of samples from the other classes that are well classified by the model). The overall model performed with an accuracy of 55%. The training set performed with an accuracy of 53% and the test set with an accuracy of 63%.

**Table 5.9 Summary of results for mushroom discrimination on the basis of damage (all data)**

| Damage Level (Minutes) | Sensitivity (%) | Specificity (%) |
|---|---|---|
| 0 | 86[a] | 46[a] |
| | 93[b] | 53[b] |
| 20 | 44[a] | 89[a] |
| | 48[b] | 96[b] |
| 40 | 29[a] | 93[a] |
| | 48[b] | 95[b] |

[a] Training set, [b] Testing set

The overall models for both the training and testing models performed well for predicting undamaged samples (0 min damage) with high sensitivities (specificities were average). However, the models did not perform as well in classifying D20 and D40 samples. Models were then developed to differentiate between the different damage levels for each tissue

type. A summary of results is shown in Table 5.10. The models were able to detect undamaged samples quite well, particularly for cap tissue with high sensitivity and specificity. The RF model produced for cap tissue performed very well with an OOB error rate of 8.3%, which suggests that cap tissue alone could be used to predict damage in mushroom samples using PLS-DA and RF modelling.

There was misclassification of samples between D20 and D40 seen for all tissues. However, lower error rates were seen for cap tissue (training and testing models). Although the models did not perform as well for differentiating between the damage levels (D20/D40) they did perform well for differentiating undamaged samples from damaged ones, making PLS-DA an important tool for detecting damage in mushrooms.

**Table 5.10 Performance statistics of PLS-DA models built using GC/MS data**

| Damage Level (Minutes) | #LV | Tissue | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|
| 0 | 4 | Cap | 92[a], 97[b] | 76[a], 81[b] |
| | | Gills | 74[a], 69[b] | 78[a], 60[b] |
| | | Stipes | 76[a], 81[b] | 85[a], 82[b] |
| 20 | 4 | Cap | 69[a], 71[b] | 66[a], 72[b] |
| | | Gills | 56[a], 61[b] | 79[a], 74[b] |
| | | Stipes | 62[a], 54[b] | 71[a], 67[b] |
| 40 | 3 | Cap | 65[a], 75[b] | 80[a], 85[b] |
| | | Gills | 36[a], 54[b] | 89[a], 81[b] |
| | | Stipes | 45[a], 51[b] | 75[a], 80[b] |

[a] Training set, [b] Testing set
#LV: Number of latent variables

Modelling damage in mushrooms has been reported in literature in recent time (Gowen *et al.*, 2008a; Esquerre *et al.*, 2009; O'Gorman *et al.*, 2010; Taghizadeh *et al.*, 2010) using different techniques including fourier transform infrared spectroscopy, hyperspectral imaging and near infrared spectroscopy coupled with chemometrics.

These studies yielded models with low error rates for predicting damage in mushrooms highlighting the usefulness of imaging and spectroscopy for detecting physical damage in mushrooms, with the possibility of using these tools to develop classification systems for the industry.

The use of GC/MS and chemometrics also produced models with low error rates for detecting damage. RF models indicated the important variables for discriminating damage i.e. specific metabolites that could be used as metabolic markers for damage in mushroom samples. The ability to detect specific metabolites for damage allows the ability to gain understanding into metabolic pathways associated with the specific metabolites identified.

Metabolomics (GC/MS) coupled with chemometrics has not to the authors knowledge been used to detect damage in mushrooms. However, it has been used as a tool in the food industry for similar use e.g. identification of volatile quality markers for ready to use lettuce and cabbage (Lonchamp *et al.*, 2009). Metabolic profiling using GC/MS to profile metabolic changes in sound and brown pears was investigated using a PLS-DA multivariate statistical approach (Pedreschi *et al.*, 2009). GC/MS profiling has also found a function in determining phytochemical diversity in tubers of potatoes (Dobson *et al.*, 2008).

These examples highlight the usefulness of GC/MS profiling and when coupled with chemometrics the ability to develop models to predict damage with low error rates, making it an invaluable tool for the mushroom industry.

*5.2.4.3. Metabolic Pathways*

A general overview of metabolism is given in Figure 5.23. Metabolites that were found in GC/MS analysis of mushroom samples are highlighted in red in the diagram.

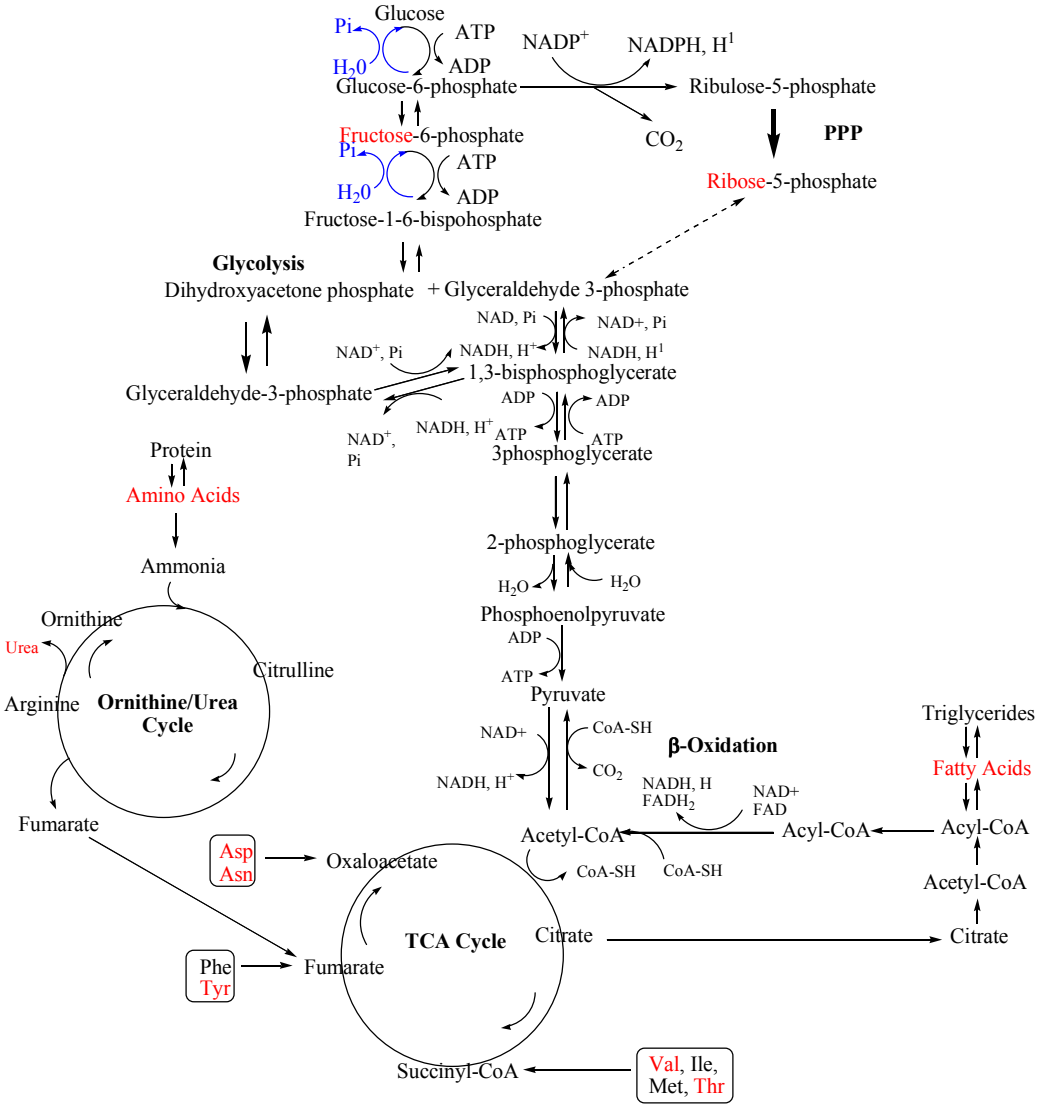



**Figure 5.23 A general overview of metabolism adapted from KEGG (KEGG, 1995)**
ATP: Adenosine triphosphate, ADP: Adensoine diphosphate, NADPH: nicotinamide adenine dinucleotide phosphate, PPP: pentose phosphate pathway, Acetyl-Co-A: Acetyl co-enzyme A, Pi: phosphatidylinositol, TCA: Tricarboxylic acid cycle, Asp: aspartic acid, Asn: asparagine, Phe: phenylalanine, Tyr: tyrosine, Val: valine, Ile: isoleucine, Met: methionine, Thr: threonine

### *5.2.4.3.1. Lipoxygenase Pathway (LOX)*

Linoleic acid was identified by RF models (1-4) as being an important variable for predicting damage in mushrooms. Day zero samples found levels of linoleic acid to be at their highest for cap and gill tissue in D40 samples. Stipe samples also had high levels of the compound (although the trend in stipe tissues was unusual). This shows that levels of

linoleic acid increased with damage. Day one samples showed levels that were found to be highest for gill and stipe tissue (D20). Levels in cap tissue were very low across all damage levels. It has been reported that the compound is found in higher concentrations in the gill tissue, particularly when damaged (Holtz and Schisler, 1971; Mau *et al.*, 1992; Cruz *et al.*, 1997). The trend for day one samples (Figure 5.24b) indicated that after 20 min damage linoleic acid levels decreased (gills and stipes). This suggests that over time and after a certain level of damage linoleic acid underwent oxidation to produce volatile components and therefore levels found in day one samples became depleted after D20.

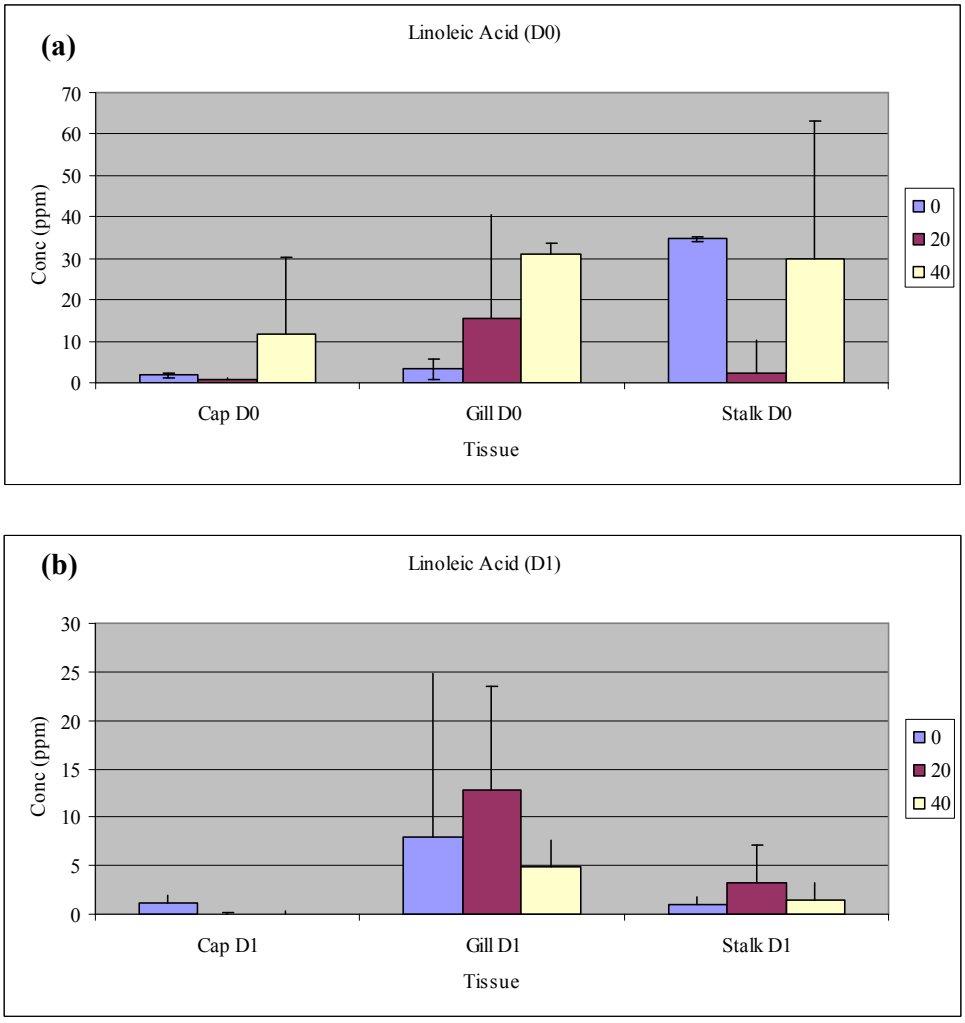



**Figure 5.24 Bar plots with the semi-quantitative concentration of linoleic acid in mushroom tissues at different damage levels in (a) Day zero samples and (b) Day one samples**

Linoleic acid levels may also be higher in damaged samples due to the degree of cellular disruption, as this could induce the release of fatty acids from membrane lipids.

As stated in section 5.2.1.1.2 and discussed in section 5.2.4.1 linoleic acid is the precursor of the mushroom alcohol 1-octen-3-ol. This aliphatic alcohol is the principal compound that contributes to the unique fungal aroma and flavour (Murahashi, 1938). Together with 1-octen-3-one, 3-octanol, 3-octanone and octanol it makes up the main eight-carbon volatile compounds present in mushrooms (Flegg *et al.*, 1985) (Figure 5.25).

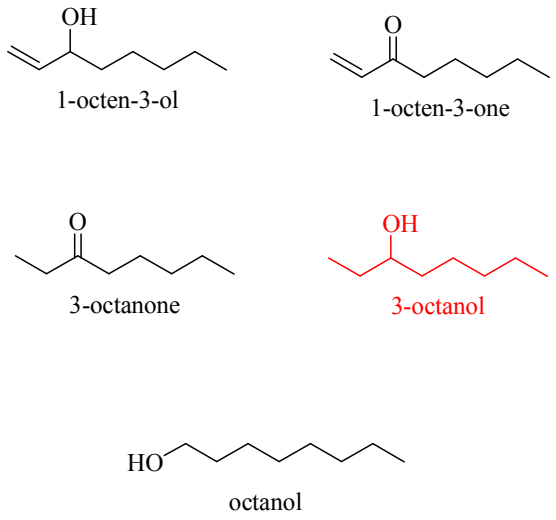



**Figure 5.25 Structures of the main eight-carbon volatiles. Structure in red was identified by GC/MS**

Eight-carbon volatile formation is unique to fungi and is likely to involve a fungal specific pathway. It is evident in literature that lipid and fatty acid metabolism has been under investigated in the fungal kingdom, with few genes and enzymes yet identified. This lack of knowledge adds to the difficulty in understanding unique systems such as eight-carbon compound production, having to depend on animal and plant resources, for sequence information to model biochemical pathways, complicating the characterisation of such unique systems (Combet *et al.*, 2006).

The lipoxygenase pathway is described in Figure 5.26. Linoleic acid is oxidised to form the eight carbon volatile 1-octen-3-ol, which is then cleaved to form a ten-carbon oxoacid (10-ODA) (Wurzenberger and Grosch, 1984).

It has been recently reported that 1-octen-3-ol may only be produced due to cellular damage, and the small amount detected from whole sporophores could be result of damage caused by the separation from mycelial cells upon harvest. The study indicated that tissue damage had a major effect on volatile formation. Mushrooms that had been sliced resulted in 10 times more volatiles being produced in comparison to whole sporophores indicating that the enzymic machinery was not operating maximally in whole mushrooms. This could possibly due to substrate limitation e.g. oxygen, or substrates in different intracellular locations. Damage also increased the range of volatiles (Combet *et al.*, 2009).

3-octanol was the only volatile identified in this study. There are two possible reasons for this. The first being that the mushroom is sliced into its three tissue types prior to extraction, causing damage (wounding). The second is the mushroom samples were subjected to mechanical damage (D20 and D40). These sources of damage could lead to the decline in volatile levels due to the reaction substrate(s) available, fatty acid and/or molecular oxygen, may rapidly become expended or that compartmentalisation reoccurs after wounding, preventing the access of enzyme to the substrate (Combet *et al.*, 2009).

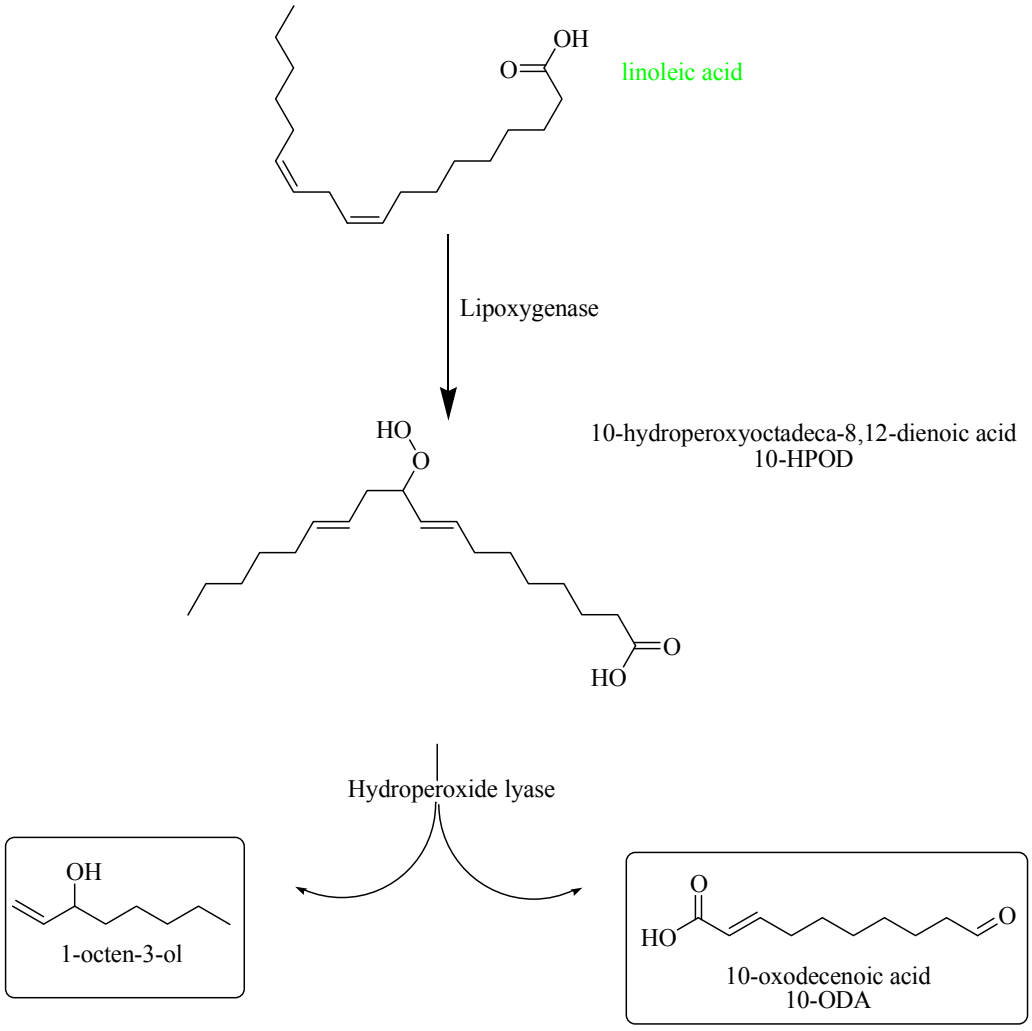



**Figure 5.26 Formation of 1-octen-3-ol (Wurzenberger and Grosch, 1984)**
Metabolite in green was identified as an important variable by RF models

*5.2.4.3.2. β-oxidation Pathway (fatty acid metabolism)*

The saturated fatty acids identified as important variables by RF models were pentadecanoic acid (model 1, 2 & 3), hexadecanoic acid (model 1, 2 & 4), heptadecanoic acid (model 4) and nonanoic acid (model 2). The unsaturated fatty acid ricinoleic acid was an important variable in model 6. The compound glycerol was present in mushroom samples and identified in RF model 4 as an important variable for detecting damage in mushroom caps. Lipases hydrolyse triglycerides releasing glycerol and fatty acids. The subsequent breakdown of glycerol (after phosphorylation) through glycolysis releases

energy (Chang and Miles, 2004). β-oxidation is the principal means by which fatty acids are metabolised by cells. The mechanism (Figure 5.23) involves a set of four consecutive reactions catalysed by four major enzymes in the process of fatty acid oxidation: acyl-CoA oxidase, 2-enoyl-CoA hydratase, 3-hydroxacyl-CoA dehydrogenase and 3-ketoacyl-CoA thiolase. Through this four-step pathway, a two-carbon unit is split from each fatty acid in the form of an acetyl-CoA unit, which can then be fed into the glyoxylate cycle or be degraded in the TCA cycle to produce $CO_2$ and $H_2O$. There are many other enzymatic activities such as *cis*-enoyl-CoA isomerise, which are necessary for the degradation of unsaturated fatty acids (Wang *et al.*, 2007).

A trend was found in damaged samples for day zero and day one samples. In day one samples levels of fatty acid were low in undamaged (UD) samples, increasing in D20 samples and even higher in D40 samples. The common trend seen in day one samples was as follows: low levels in UD samples, increasing in D20, with a decrease seen in D40 samples. A recent study reported that pentadecanoic acid, hexadecanoic acid and nonanoic acid levels decreased postharvest (Combet *et al.*, 2009). This suggests that these metabolites are affected by damage and could be used as markers of damage.

The cellular disruption caused by mechanical damage could induce the release of free fatty acid from membrane lipids and lipid globules, as well as breaking down the cellular compartments, increasing the levels of fatty acids. This suggests that when damaged or under stress the mushroom produces higher levels of fatty acids, releasing its reserves in a possible protective capacity or to do with a possible repair function. However, after certain levels of stress (i.e. D40) and after a certain length of time the levels of fatty acids begin to decrease.

*5.2.4.3.3. Isomerisation of Glucose (myo-inositol)*

The polyol myo-inositol was identified in RF models (1, 3 & 5), making it an important marker for damage particularly in the overall model, modelling damage in day one samples and in gill tissue samples. Myo-inositol can occur in the free form and in compounds such as inositol phospholipids (Ikawa *et al.*, 1968). Myo-inositol was synthesised by isomerisation of glucose-6-phosphate and dephosphorylation (Figure 5.27) (Loewus and Murthy, 2000) and is utilised by several pathways including phosphatidylinositol phosphate pathways.

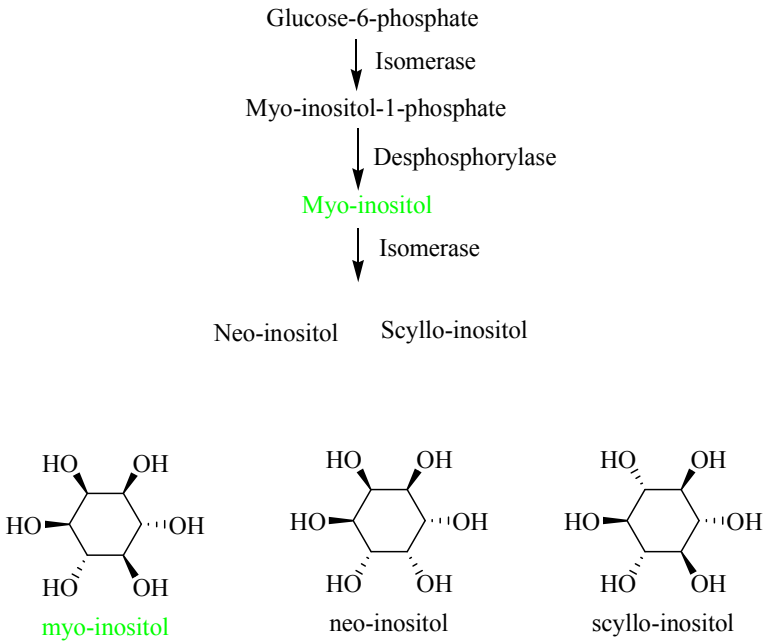
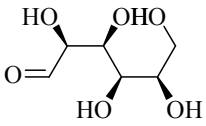



**Figure 5.27 Formation and structures of myo-, neo- and scyllo-inositols. Metabolite highlighted in green was identified as an important variable by RF models**

Levels of myo-inositol were slightly higher in damaged samples (D40) for day zero samples and levels were also higher in damaged day one samples (D20) indicating that this metabolite increased in damaged samples and may be a useful marker for damage.

*5.2.4.3.4. Glycolytic Pathway (D-mannose)*

D-mannose was identified as an important metabolite for modelling damage in mushrooms (RF model 4). This particular model was built using cap data only and it gave a very good OOB error rate of 8.33%. Mannose (Figure 5.28) undergoes glycolysis i.e. converts glucose-6-phosphate or fructose-6-phosphate to pyruvate and is phosphorlyated by hexokinases. Glycolysis occurs in the cytosol and is the ubiquitous means to convert glucose into pyruvate, providing the cell with energy, precursors and NADH (Van Laere, 1995; Arraes *et al.*, 2005).




D-mannose

**Figure 5.28 Structure of D-mannose**

The bar plot below (Figure 5.29) shows that the highest levels of D-mannose were in samples after D20. This suggests that after a certain amount of damage, D-mannose enters glycolysis and therefore levels become depleted and is an indicator of damage.

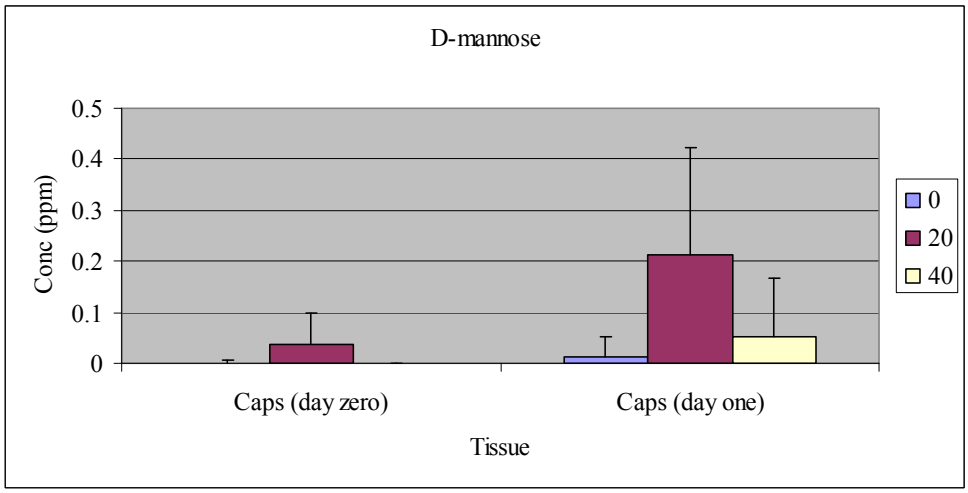



**Figure 5.29 Bar plot with the semi-quantitative concentration of D-mannose in mushroom cap tissues at different damage levels in day zero samples and day one samples**

#### *5.2.4.3.5. TCA Cycle (Krebs cycle)*

Citric acid was indicated in RF model 5 (gill tissue) as an important variable for discriminating damage in the gill tissue of mushrooms. The biochemical pathways related to citric acid accumulation and the role of the tricarboxylic acid cycle (TCA) (Figure 5.23) in fungi has been well established. Citric acid accumulation can be divided into three processes:

1. The breakdown of hexoses to pyruvate and acetyl-CoA by glycolysis
2. Formation of oxaloacetate
3. Condensation of acetyl-CoA and oxaloacetate to citric acid (Kubicek, 1988)

Succinic acid was also detected in mushroom samples and is a product of the TCA cycle, although it was not identified as an important metabolite by RF models.

Levels of citric acid were at their highest in day zero samples after 40 minutes of damage.

#### *5.2.4.3.6. Polyol Metabolism (glucitol/mannitol & hexitol)*

The sugar alcohol glucitol was identified as an important metabolite for modelling damage in models 3 & 5 (day one samples and gill samples) and D-ribo-hexitol was identified in models 5 & 6 (gills and stipe tissue). As mentioned previously in section 5.2.2.2 glucitol and mannitol are isomers. A limitation of mass spectral libraries is the inability to differentiate between structural isomers. Because mannitol is the main polyol found in *Agaricus bisporus* (Beecher et al., 2001; Tsai et al., 2007), it is recommended that mannitol was detected by mass spectrometry. Mannitol and hexitol together with myo-inositol and glycerol were detected by mass spectrometry, with mannitol and hexitol identified as important metabolites by RF models. Hexitols are synthesised from glucose by the routes shown in Figure 5.30.

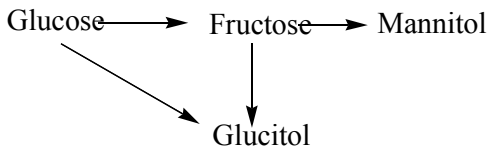



**Figure 5.30 Hexitol synthesis**

Overall levels of mannitol were found to be higher in day one samples, with concentrations increasing with damage levels i.e. D40 samples had the largest concentrations of mannitol, and this trend was also found in gill tissues. Hexitol levels decreased over time in gill and stipe tissues indicating the use of this metabolite as a useful marker for damage in these tissues.

Harvested mushrooms continue to have high rates of respiration linked to a switch from nutrient import to the breakdown of storage compounds (Hammond and Nichols, 1975). It has been reported that mannitol levels decrease postharvest while levels of cell wall and urea increase (Hammond, 1979; Eastwood *et al.*, 2001). In contrast mannitol levels were found to be higher in day one samples and samples that had been damaged. This indicates that mannitol could be used as a marker of damage in mushrooms as its levels increased with damage. Trehalose and mannitol are thought to act as storage carbohydrates for sporophores production and reserves, under conditions of water stress (Wells *et al.*, 1987; Burton *et al.*, 1994). Trehalose may be converted to glucose and then mannitol in the upper stipe and/or in the cap and transported to the upper stipe. This conversion of trehalose could be the reason mannitol levels increase over time and damage and also the reason trehalose is not identified by mass spectrometry.

Research has revealed over-expressed sugar transporter genes in *A. bisporus* cell membranes between stage 2 and 4 mushroom sporophores (Beecher *et al.*, 2001). This over expression may also occur when a mushroom becomes damaged, however further studies would be required.

*5.2.4.4. Correlation of Metabolites (correlation matrices)*

Polar and non-polar metabolite groups were examined. Polar extract metabolites included the amino acids and sugars/polyols whilst non-polar metabolites were the fatty acids and phenolic compounds. Pair-wise correlation analysis was performed on the response ratios of all metabolites. This approach can identify both synthetically related (Steuer *et al.*, 2003) and co-ordinately regulated metabolites. Correlations between metabolites were examined at for each damage level (0, 20, and 40), each tissue type and each day (day zero/day one) respectively. Two metabolites were considered to be highly correlated if the coefficient had a value of ≥0.9, and on this basis there were 132 highly positively correlated pairs for day zero samples and 121 medium correlated metabolites (≥0.7-0.9). Of the highly correlated metabolites 93 were between fatty acids, 21 between amino acids, 13 between sugars and 5 between phenolics. Day one samples gave 39 highly correlated and 74 medium correlated metabolites. Of the highly correlated metabolites 13 were between fatty acids, 13 between amino acids, 11 between phenolics and 2 between sugars.

*5.2.4.4.1. Correlation Matrices (polar metabolites)*

In the following sections only highly correlated metabolites will be discussed (i.e. coefficient values ≥0.9). Correlation matrices for polar metabolites (day zero samples) of cap tissue are shown in Figures 5.31, 5.32 and 5.33. A striking feature of the data was the extent of correlation within amino acids in D20 cap tissue (Figure 5.32). There were a number of high correlations between amino acids and sugars/polyols such as glucitol with alanine and myo-inositol with glycine (Figure 5.31). Examples can also be seen in D40 samples: glycerol with alanine, glucitol/mannitol with tyrosine and myo-inositol with proline etc (Figure 5.33). Amino acids play important roles as basic substrates and as regulators in many metabolic pathways (Brosnan, 2003).

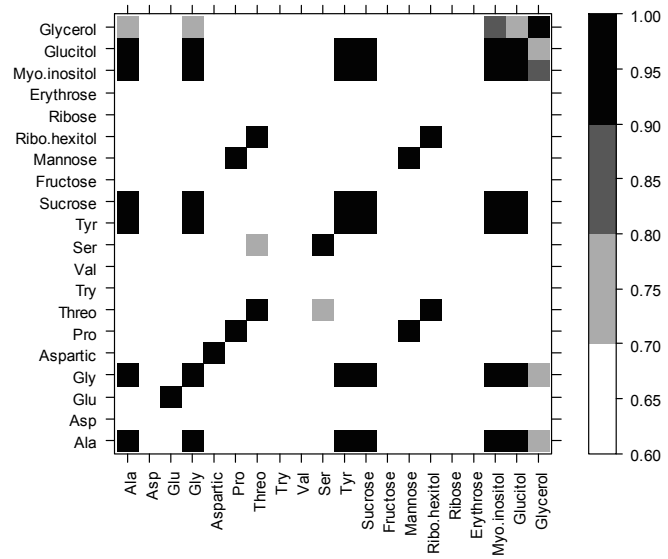



**Figure 5.31 Correlation matrices of polar metabolites in day zero cap tissues at 0 min damage**
Tyr: tyrosine, Ser: serine, Val: valine, Try: tryptophan, Threo: threonine, Pro: Proline, Aspartic: aspartic acid, Gly: glycine, Glu: glutamine, Asp: asparagine, Ala: alanine

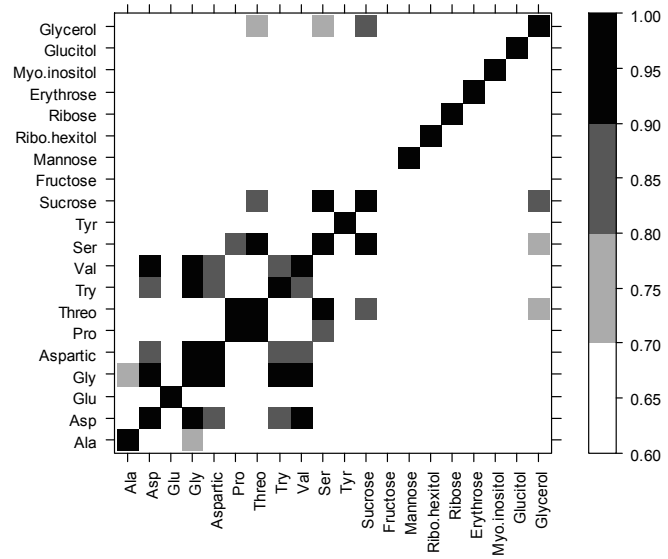



**Figure 5.32 Correlation matrices of polar metabolites in day zero cap tissues at 20 min damage**
Tyr: tyrosine, Ser: serine, Val: valine, Try: tryptophan, Threo: threonine, Pro: Proline, Aspartic: aspartic acid, Gly: glycine, Glu: glutamine, Asp: asparagine, Ala: alanine

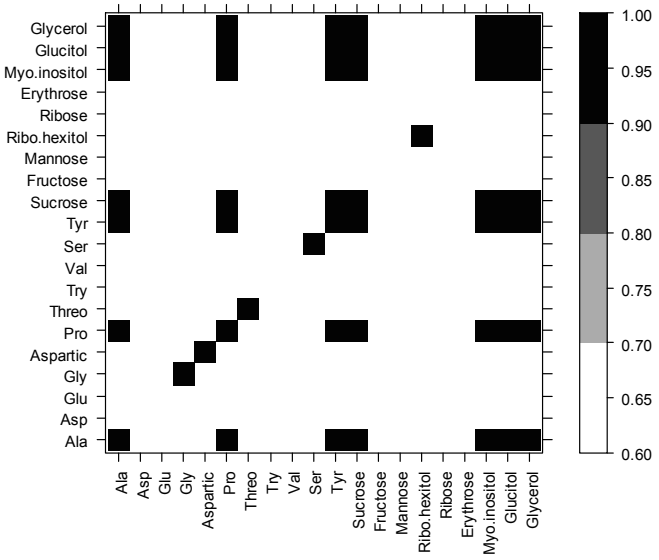



**Figure 5.33 Correlation matrices of polar metabolites in day zero cap tissues at 40 min damage**
Tyr: tyrosine, Ser: serine, Val: valine, Try: tryptophan, Threo: threonine, Pro: Proline, Aspartic: aspartic acid, Gly: glycine, Glu: glutamine, Asp: asparagine, Ala: alanine

The correlation matrix (Figure 5.31) for undamaged day zero caps showed a number of correlations between the same groups of metabolites and also a number of inter-correlations for e.g. alanine with myo-inositol. Metabolites that were highly correlated produced plots with linear relationships (Figure 5.34). The highly correlated metabolites between amino acids and sugars/polyols were not seen in D20 cap tissues. A number of amino acid metabolites with high correlations were seen. The matrix for D40 samples had fewer amino acid metabolite correlations with an increase of inter-correlations found, similar to the correlation matrix for undamaged cap tissue.

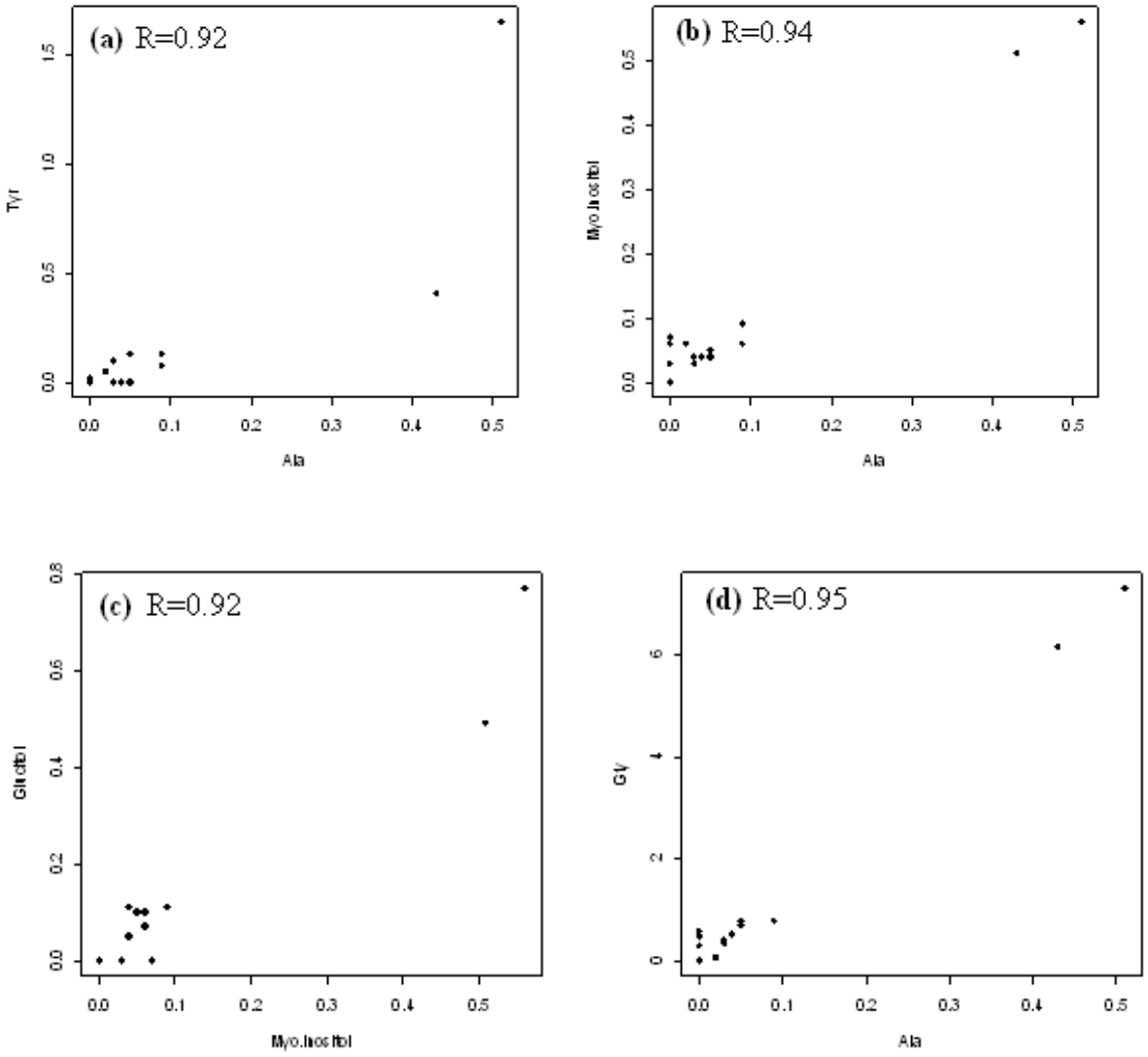



**Figure 5.34 Selected plots of response ratios of highly correlated metabolites seen in undamaged cap tissue**
Ala: alanine, Gly: glycine, Tyr: tyrosine
R=correlation value

Correlation matrices were also examined for gill and stipe tissues (Day zero samples). The following observations were made: No pairs of metabolites were highly correlated in UD and D20 matrices for gill tissue, with 5 pairs found in D40 samples: glycerol with sucrose, glycerol with myo-inositol, mannose with aspartic acid, mannose with tyrosine and sucrose with myo-inositol. Stipe tissues had the following trends: UD samples had 3 highly

correlated pairs of metabolites. D20 contained no correlated pairs, D40 samples had a number of correlation pairs and followed the same trend as cap D20 samples i.e. highly correlated pairs were found amongst amino acids only (Figure 5.32).

The observation of correlations shows that the metabolite concentrations are dependent on each other and therefore must be strongly connected to the underlying biophysical system. Cell metabolism constitutes a complex dynamical system, which is continuously subject to fluctuations. These fluctuations arise from a continuously changing environment and also from complex patterns of regulation, generated by the network itself. These fluctuations induce variability in certain metabolites, propagate through the network and generate an emergent pattern of correlations (Steuer *et al.*, 2003). The strong correlations between amino acids and sugars/polyols particularly in UD and D40 cap samples suggests the possibility that amino acid synthesis might be controlled, at least partly, by carbohydrates or associated factors.

Correlation matrices (polar metabolites) for day one cap samples (Figures 5.35, 5.36 and 5.37) gave 3 highly correlated pairs of metabolites in UD samples; glucitol/mannitol with alanine, glucitol/mannitol with tyrosine and tyrosine with alanine. D20 gave 1 pair of highly correlated metabolites; glucitol/mannitol with fructose and finally D40 samples had the following pairs; myo-inositol and glycine, mannose with alanine, tyrosine and alanine and glycine and alanine. The appearance of new correlations in the damaged samples in comparison to undamaged samples indicate the activation of new metabolic pathways through the effect of damage, affecting the ratios/relationships between the different metabolites and imply *de novo* enzyme production. Correlations of cap tissue (day one) show the adaption of mushroom metabolism to the mechanical damage (Figures 5.35, 5.36 and 5.37).

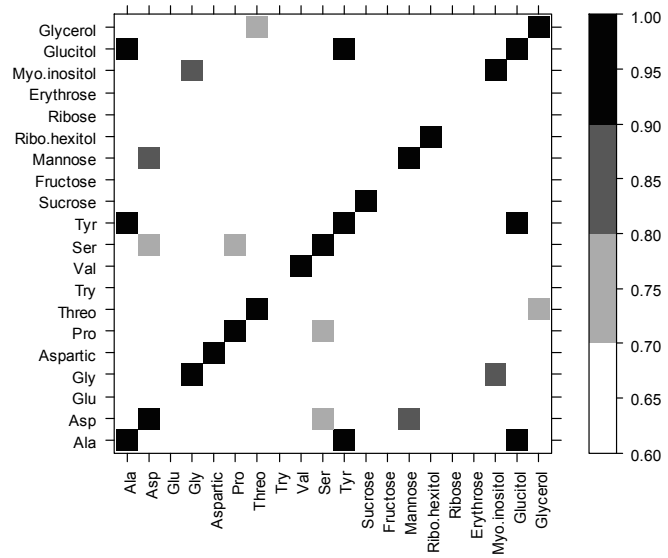



**Figure 5.35 Correlation matrices of polar metabolites in day one cap tissues at 0 min damage**
Tyr: tyrosine, Ser: serine, Val: valine, Try: tryptophan, Threo: threonine, Pro: Proline, Aspartic: aspartic acid, Gly: glycine, Glu: glutamine, Asp: asparagine, Ala: alanine

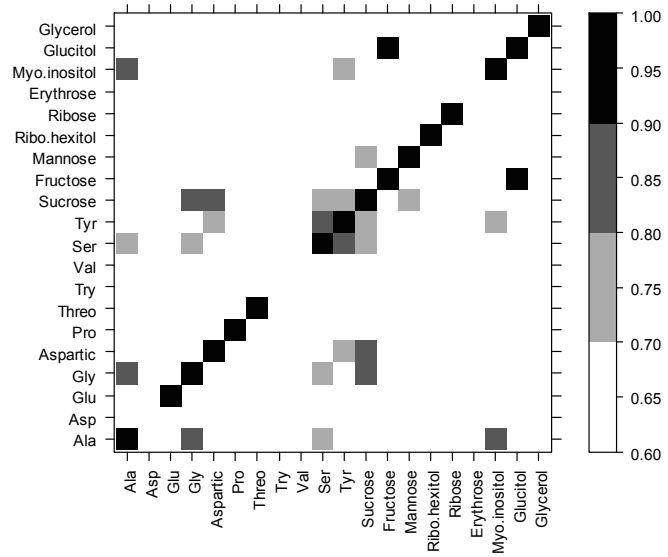



**Figure 5.36 Correlation matrices of polar metabolites in day one cap tissues at 20 min damage**
Tyr: tyrosine, Ser: serine, Val: valine, Try: tryptophan, Threo: threonine, Pro: Proline, Aspartic: aspartic acid, Gly: glycine, Glu: glutamine, Asp: asparagine, Ala: alanine

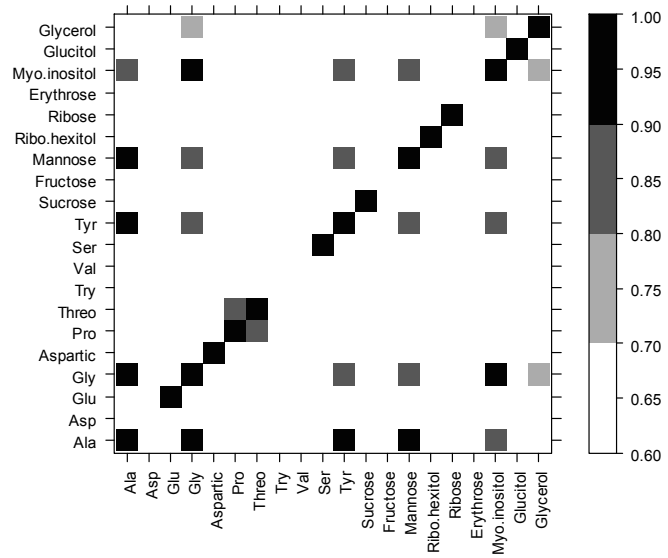



**Figure 5.37 Correlation matrices of polar metabolites in day one cap tissues at 40 min damage**
Tyr: tyrosine, Ser: serine, Val: valine, Try: tryptophan, Threo: threonine, Pro: Proline, Aspartic: aspartic acid, Gly: glycine, Glu: glutamine, Asp: asparagine, Ala: alanine

Day one gill samples had no highly correlated pairs for UD samples. The following pairs of metabolites were highly correlated in D20 samples; tyrosine with glycine and threonine with aspartic acid. The following were seen in D40 samples; glycerol with myo-inositol, threonine with aspartic acid, glycine with alanine.

Day one stipe samples had no highly correlated pairs for UD samples. For D20 there were a number of highly correlated metabolites seen, however, they were between amino acids only. D40 samples contained no highly correlated metabolites.

There was no trend seen that could be used to differentiate between damaged samples and undamaged ones. However, correlation matrices are useful in understanding metabolic pathway interactions between metabolites. It seems that pathways controlling carbon and amino acid metabolism should cross-link, since amino acids are based on carbon skeletons (Morcuende *et al.*, 1998) and therefore correlations between these groups can be seen in the correlation matrices.

*5.2.4.4.2. Correlation Matrices (non-polar metabolites)*

Non-polar metabolites included the fatty acid and phenolic compounds. The highest number of highly correlated metabolites was found for cap tissue after 40 min damage. There were high correlations seen between saturated fatty acids with expected correlations between fatty acids with even carbon numbers (e.g. octadecanoic acid and eicosanoic acid) and between those with odd carbon numbers (e.g. pentadecanoic acid and heptadecanoic acid), the members of each series being biosynthesised sequentially from the same starting unit by addition of a $C_2$ unit from malonyl-CoA (O'Hara *et al.*, 2002) (Figure 5.38).

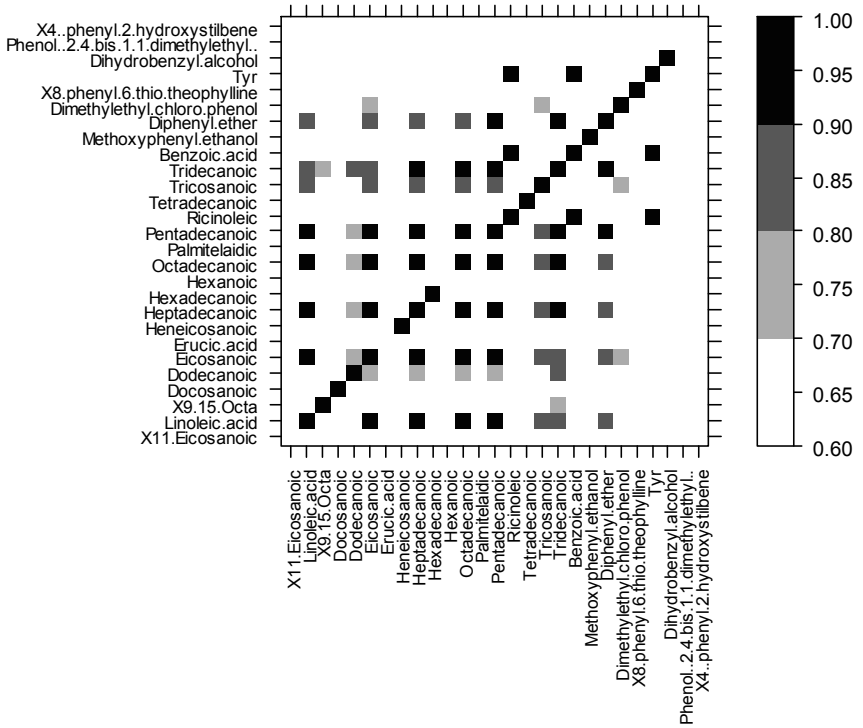



**Figure 5.38 Correlation matrices of non-polar metabolites in day zero cap tissues at 40 min damage**

Only a few correlations were seen between fatty acids and phenolic compounds in day zero cap samples. However, a number were found in day one samples that had been extensively damaged (D40). There were a total of 8 fatty acid and phenolic pair-wise correlations.

Interestingly high correlations were seen between pentadecanoic acid and linoleic acid in day zero caps (D20 & D40), day zero gills (D20, D40), day one gills (D20 & D40) and day zero stipes (D20, D40), which were identified in the first RF model as variables of importance for modelling damage in mushrooms. A correlation between these two metabolites was not significant in cap, gill or stipe undamaged tissue, for both day zero and day one samples. This suggests that a metabolic pathway (related to fatty acids and possibly membrane regeneration) becomes switched on when a mushroom becomes damaged.

These metabolites were also identified in a number of RF models as being important variables for predicting damage. Figure 5.39 shows the response ratios of linoleic acid and pentadecanoic acid for gill tissue at each day and damage level.

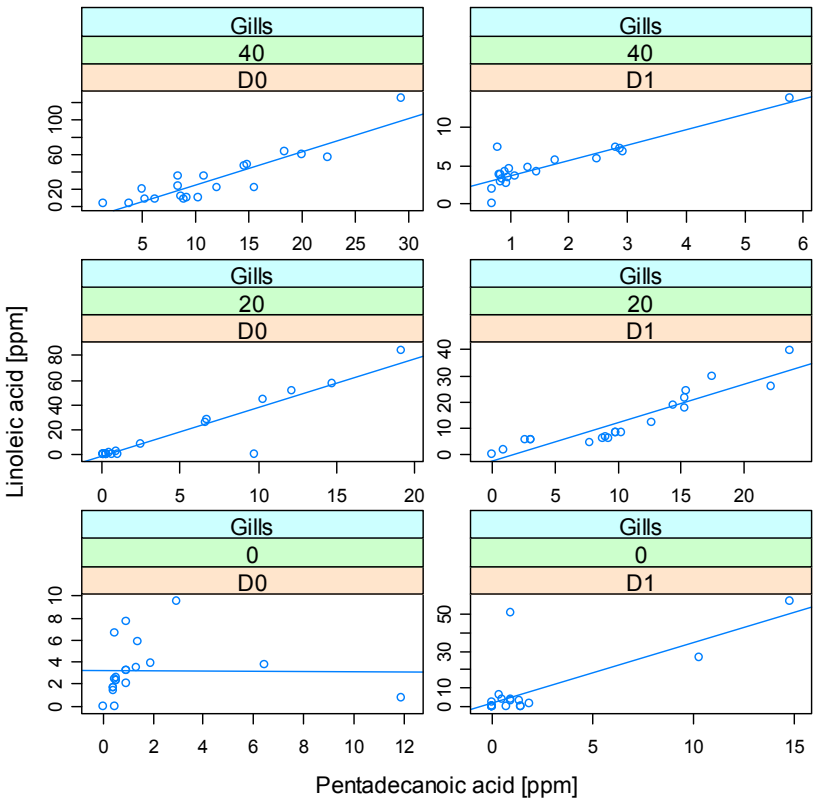



**Figure 5.39 Plots of response ratios for linoleic and pentadecanoic acid (gill tissue) at each damage level and day**
D0: Day zero; D1: Day one

A high correlation for these two metabolites was seen at damage levels D20 and D40 for both days. These metabolites were not highly correlated for undamaged gills and therefore may be used as indicators of damage. This pattern was also seen for cap and stipe tissue.

## 5.3. CONCLUSIONS

One hundred and five metabolites were identified by GC/MS analysis. These metabolites included fatty acids, phenolics, sugars and polyols.

PCA found patterns of clusters in UD and D20 samples for both day zero and day one cap tissues, indicating that damage had an effect on metabolite levels. Patterns were not so clear in gill and stipe tissues.

An overall RF model was developed using all the samples, a variable indicating the tissue from which the metabolite originated (cap, gill or stipe) and the age of the mushroom (day zero/day one) as explanatory variables. This model gave a good OOB error rate of 11.11%. A second model was produced removing the explanatory variables resulting with an OOB error rate of 11.39%. The two models identified the following five metabolites as important variables for predicting damage: pentadecanoic acid, linoleic acid, myo-inositol, benzoic acid and hexadecanoic acid. Univariate analysis confirmed that the difference in concentrations of these metabolites was significant between damaged and undamaged samples.

RF models were then built for the different days (day zero/day one) and the different tissue types (cap, gills or stipes). OOB error ranges were between 8-14% with model 3 (day one samples only) having the best prediction with an OOB error rate of 6.67%. Out of the total RF models produced (i.e. 6 models Table 5.8) 17 metabolites were identified (top 5 for each) as important variables for detecting damage, with a number of the same metabolites identified in different models. The RF models produced could be used as classifiers of

mushroom damage as they all resulted with good prediction models (OOB >14%) and can identify specific metabolic markers of damage.

The overall PLS-DA model did not perform as well in comparison to the RF models. However, the model was able to classify correctly undamaged samples from damaged samples very well in both training and testing sets. Models were then developed separately for each tissue type. All the tissue types produced models with low error rates for discriminating between undamaged and damaged samples. Gill and stipe tissues had the highest levels of misclassification in their ability to discriminate D20 and D40 samples. Cap tissue produced the best models (training and testing) for classifying correctly each damage level.

Correlation matrices were produced for non-polar and polar metabolites, with each tissue and age of sample examined separately. Correlation matrices yielded 171 highly correlated metabolites (≥0.9), which were mainly within each metabolite group. However, a number of inter-correlated correlations were also identified.

Non-polar correlation matrices indicated that linoleic acid and pentadecanoic acid were highly correlated within damaged samples. These metabolites were also highlighted by RF models as important indicators of damage.

# 6. EVALUATING METABOLOMIC TECHNOLOGIES FOR IDENTIFYING MARKERS OF DAMAGE IN MUSHROOMS (*AGARICUS BISPORUS*)

# 6. Evaluating metabolomic technologies for identifying markers of damage in mushrooms (*Agaricus bisporus*)

This study evaluated damage in mushrooms using a metabolomic approach. Three different metabolomic techniques (FTIR spectroscopy, NMR spectroscopy and GC/MS) were used coupled with chemometric methods (PCA, RF and PLS). The following chapter will aim to describe and examine the stages of sample preparation, analysis and results in each of the metabolomic trials.

## 6.1. SAMPLE PREPARATION

### *6.1.1. FTIR Spectroscopy*

FTIR spectroscopy is a technique used to obtain a spectral fingerprint of biological samples which represents a snap-shot of the biochemistry at a given time. Sample preparation was minimal, no extraction procedure was involved and samples did not require chemical derivatisation. FTIR spectroscopy enables reagentless analysis and is comparatively inexpensive. The technique also facilitated high-throughput analysis in terms of both sample preparation and analysis time (less than 1 min per sample).

### *6.1.2. NMR Spectroscopy*

In NMR analyses sample preparation was a straightforward procedure, which involved a simple extraction (chloroform/methanol/water) in order to separate the polar and non-polar mushroom extracts. A suitable deuterated solvent was needed for each phase to allow for locking of the signal. Derivatisation of analytes was not required. Like FTIR spectroscopy, NMR spectroscopy is also as a high-throughput technique (2-3 min per sample) and non-destructive, permitting subsequent analysis by other methods (Fan, 1996).

### *6.1.3. GC/MS*

In GC/MS analyses sample preparation was more extensive in comparison to FTIR and NMR spectroscopy. It required sample drying, which can result in loss of volatile metabolites. Subsequent two-stage chemical derivatisation was required to induce volatility and thermal stability (Roessner *et al.*, 2000). Oxime/silylation derivatisations which were used in the protocol are time consuming (1-2 h) and the stability of derivatised samples is an issue.

## 6.2. ANALYSIS

### *6.2.1. FTIR Spectroscopy*

FTIR spectroscopy results in an absorption spectrum that provides a characteristic fingerprint of the sample. Five major regions have been highlighted within the 4000 to 600 $cm^{-1}$ (MIR) region. These are broadly termed as the fatty acid region (3100-2800 $cm^{-1}$), the amide region (1700-1500 $cm^{-1}$), which can be divided into the amide I and amide II bands, the polysaccharide region (1200-1250 $cm^{-1}$) and a mixed region containing a variety of weak features.

The fingerprint region of the mushroom spectra was examined (2000-400 $cm^{-1}$). This region was information rich and a number of peaks were attributed to functional groups. Spectra were highly reproducible.

### *6.2.2. NMR Spectroscopy*

$^{1}H$ NMR analyses of samples was carried out on polar and non-polar extracts to discriminate low levels of damage in mushrooms. NMR spectroscopy is a specific and yet non-selective technique. This meant that each separate resonance observed in an NMR spectrum was specific to a particular compound, providing a wealth of structural

information regarding the components of a sample. NMR spectroscopy did not require pre-selection of the analysis conditions, such as ion source conditions for mass spectrometry or chromatographic operating conditions (stationary phase, mobile phase, temperature).

Manipulation of spectra was difficult as they contained thousands of signals relating to metabolites which made identification very difficult. However, it was possible to see visible differences between the spectra of damaged and undamaged mushrooms, prior to chemometric analysis. Non-polar extracts gave information rich spectra with more visible differences seen between damaged and undamaged samples compared to polar extracts. Polar phase spectra were not as information rich as expected compared to other NMR studies of mushrooms (Cho *et al.*, 2007). A reason for this could be that there are a range of polar compounds extracted, but individually they are present in very small amounts and are below the detection limit. Another reason could be that polar compounds are not too soluble in methanol and therefore were not extracted efficiently.

### *6.2.3. GC/MS*

In GC/MS analyses samples were analysed with small sample injection volumes (1 μl) on a high resolution capillary column, allowing sensitive analyses, which is one of the most important requirements for metabolomics (Sumner *et al.*, 2003). An electron impact spectrometer was used to provide molecular ion fragmentation to produce a mass spectrum indicative of the metabolites structure. Metabolites were detected with good sensitivity and spectra were highly reproducible.

The analysis of raw data was carried out using deconvolution software which verified individual analytes and deconvoluted co-eluting peaks. Identification of metabolites was achieved using commercially available libraries (NIST), standards and literature data. There are a number of commercially available MS libraries e.g. NIST, EPA and NIH which are extensive. However, they do not contain a large number of metabolites possibly

perceived when studying metabolic pathway networks. There is a need for metabolite specific libraries which are being produced within the community but are limited to the metabolites commercially available or those that can be identified from mass spectral interpretation (Dunn *et al.*, 2005).

GC/MS sample preparation was a lengthy procedure and together with the run time for each sample (90 min per sample) this makes the technique a much longer process in comparison to FTIR and NMR spectroscopy.

GC/MS analysis of mushrooms resulted in a library of 105 metabolites being built. A range of metabolites including amino acids, fatty acids, carbohydrates, polyols and organic acids were detected in high probabilities.

An overview of the three techniques is shown in Table 6.1.

**Table 6.1 Comparison of analytical platforms used for detecting damage in mushrooms**

| Method | Advantages | Disadvantages |
|---|---|---|
| FTIR | • High-throughput<br>• No sample preparation<br>• Inexpensive<br>• Reproducible<br>• Requires no derivatisation | • Low sensitivity<br>• Requires larger samples<br>• Qualitative rather than quantitative |
| NMR | • Non-destructive<br>• High throughput<br>• Little or no sample preparation requirement<br>• Robust, mature technology<br>• Requires no derivatisation<br>• Quantitative | • Relative low sensitivity<br>• Expensive instrumentation<br>• Requires larger samples (0.5 ml)<br>• Limited metabolites coverage<br>• Relative poor selectivity (signal overlap) |
| GC/MS | • Robust, mature technology<br>• Relatively inexpensive<br>• Quantitative (with calibration)<br>• Modest sample size<br>• Good sensitivity<br>• Excellent separation reproducibility | • Sample not recoverable<br>• Requires derivatisation<br>• Sample bias (volatile & stability)<br>• Standards or data base dependence |

## 6.3. RESULTS

Principal component analysis (PCA), random forests (RF) and partial least squares (PLS) were used to extract information from each metabolomic technique in order to develop models for predicting damage in mushrooms.

### *6.3.1. FTIR Spectroscopy*

In FTIR spectroscopy RF models had the ability to discriminate between damaged and undamaged samples with low error rates (5.9% and 9.8%). The models produced VIP plots (variables of importance) which identified specific wavenumbers that were important for detecting damage. The wavenumbers identified were 1868, 1870, 1845 and 1560 $cm^{-1}$. The wavenumbers 1868, 1870 and 1845 $cm^{-1}$ are unassigned. However, they are all located along the shoulder for the peak located at 1650 $cm^{-1}$, which is attributed to amide II. The wavenumber 1560 $cm^{-1}$ was identified as an important variable and is attributed to an amide I group. Univariate statistical analysis of these important variables showed that damaged samples have higher absorbancies at these wavenumbers, indicating that amide peaks are important for detecting damage using FTIR spectroscopy.

PLS-DA produced models with low error rates. Misclassification errors associated with all models were low, particularly in the case of gills and stipes. Correct classification ranged from 92-99% (different tissues), highlighting the usefulness of this technique to identify mushrooms that had been physically damaged.

### *6.3.2. NMR Spectroscopy*

In NMR spectroscopy RF models were produced for both non-polar and polar samples, and for each tissue type and day i.e. day zero and day one. Non-polar samples produced RF models with low error rates (10%) and indicated signals in the lipid and aromatic regions

as being important variables for determining whether samples were damaged or not. Polar samples resulted with models having large error rates and therefore it may be a useful tool for modelling damage in non-polar samples only.

PLS-DA models resulted with low error rates for predicting damaged samples using both non-polar and polar samples. However, non-polar samples were slightly more successful in their discrimination ability.

NMR spectroscopy coupled with chemometric tools was successful for modelling damage in mushroom samples that had been subjected to low levels of damage. The method did not prove successful for profiling metabolites, although further extraction procedures may yield more conclusive results.

### *6.3.3. GC/MS*

In GC/MS analyses an overall model to predict damage using all data produced an RF model with an OOB error rate of 11.1%. Further models were produced for the samples ages (day zero and day one) and for each tissue respectively. Model errors ranged from 6.67 (day one samples) to 13.33% (gill tissue). VIP plots identified important metabolites as being the most informative for discriminating between damage and undamaged samples. Fatty acids were the most important variables for detecting damage, with linoleic acid being identified in four out of six models.

This study identified a number of fatty acids as being important metabolic markers of damage in mushrooms, suggesting that lipid membranes were affected by physical damage, thereby changing the mushrooms physiology in some way. GC/MS allowed the identification of specific metabolites, therefore enabling a more extensive understanding into the metabolic pathways affected by damage.

GC/MS metabolic profiling also allowed a library of metabolites to be built. A total of 105 metabolites were separated and identified by this analysis, including a number of fatty acids, carbohydrates, amino acids, polyols and organic acids.

PLS-DA was applied to the GC/MS data, producing models with acceptable error rates. Misclassification was seen between the different levels of damage (D20 & D40). However, the models were ultimately able to differentiate between damage and undamaged samples (low error rates) for both days.

An overview of chemometric results for the three methods used is given in Table 6.2

**Table 6.2 Summary of metabolomic techniques and the best models used to evaluate damage in mushrooms**

| Technique | RF (%) | PLS-DA |
|---|---|---|
| **FTIR** | 5.9 | Low misclassification |
| **NMR** | 9.97 | Low misclassification |
| **GC/MS** | 6.67 | Low misclassification |

## 6.4. CONCLUSIONS

FTIR spectroscopy produced the model with the lowest error rate for predicting damage in mushrooms in comparison to the other metabolic techniques used, with amides identified as important variables for predicting damage using the random forest method. PLS-DA also resulted with low errors of misclassification. The method was rapid with many samples being analysed per day. The prospect of using ATR-FTIR would allow non-destructive analysis with even more samples being analysed per day.

However, in order to profile the metabolites in mushrooms, GC/MS was the most useful in terms of its ability to separate and identify specific metabolites with over 100 metabolites identified. Models for discriminating damage were also low (6.67%) with specific individual metabolites being identified as possible markers for damage. This allowed an

insight into which metabolites were affected by mechanical damage and which metabolic pathways were involved.

NMR spectroscopy also proved a useful tool for modelling low levels of damage in mushrooms (9.97%) with signals identified in the lipid and aromatic regions highlighted as important variables for discriminating damage. Further work would need to be carried out in this area in terms of extraction in order to successfully profile individual metabolites.

In conclusion there is no single analytical technique that has the ability to profile all of the metabolome and therefore a combination of techniques is useful and complimentary. This was found for NMR and GC/MS analyses with lipids being identified by RF models as important metabolites for detecting damage in mushrooms.

# 7. CONCLUSIONS

# 7. Conclusions

## 7.1. Overall Conclusions

This study employed metabolomic profiling and fingerprinting techniques (FTIR and NMR spectroscopy and GC/MS) for the detection and discrimination of mechanical damage in mushrooms. Mushrooms subjected to various levels of mechanical damage (from mushrooms with low damage levels and general good acceptable colour to mushrooms unacceptable for sale) were studied.

In summary, this study draws the following conclusions:

- The metabolic fingerprinting techniques (FTIR and NMR spectroscopy) proved to be the more efficient methods in terms of sample preparation and sample analysis.
- Employment of FTIR spectroscopy and chemometric tools was successfully applied to investigate the chemical changes induced by mechanical damage to mushrooms. RF and PLS-DA had the ability to model damage in mushrooms with low error rates. RF models identified specific wavenumbers (for amide vibrations) as being important variables for indicating damage. The absorbance values for these wavenumbers increased in the spectra of damaged samples and could be therefore used as markers for damage.
- FTIR spectroscopy was also used with chemometric tools to investigate whether the age of the mushrooms could be predicted. Models did not predict age in mushroom samples very well.
- NMR spectroscopy and chemometric tools were successfully employed to detect low levels of damage in mushrooms. Non-polar phase spectra yielded predictive models of damage with low error rates. RF models identified signals in the lipid

and aromatic region of the $^{1}H$ NMR spectra as being important variables for detecting damage. Polar phase spectra were not as successful for modelling damage in mushrooms with high error rates produced using RF models. PLS-DA models were however, successful for predicting low levels of damage.

- Identification and assignment of individual metabolites for qualitative purposes proved extremely difficult in NMR analysis.
- The metabolic profiling tool GC/MS was employed in order to build a library of mushroom metabolites. Over 100 metabolites were separated and identified including carbohydrates, fatty acids, phenolic compounds, amino acids, polyols and organic acids.
- Chemometric tools were successfully applied to GC/MS data to predict damage in mushrooms. RF models identified specific fatty acids as important markers of damage. PLS-DA models were also able to predict damage in an acceptable manner.

Overall the results from this study showed different metabolomics techniques had the potential to differentiate between mushrooms that had been mechanically damaged and those that were undamaged. These techniques could be used in the mushroom industry replacing older slower methods, reducing time and labour costs.

## 7.2. FUTURE INVESTIGATIONS

The work presented in this thesis could be extended and improved taking the following recommendations into consideration:

- A study using ATR-FTIR spectroscopy as opposed to pellet sampling to evaluate its ability to model damage using the same chemometric methods. If successful this

would enable non-destructive analysis of fresh mushrooms. This would reduce sample preparation time and allow analysis of fresh samples as opposed to freeze-dried samples.

- Further analysis needs to be carried out using NMR spectroscopy in terms of extractions and solvents used. A targeted metabolomics approach may be required in order to identify and assign individual metabolites. This would involve further extraction and purification procedures coupled with 2D NMR analysis for a more detailed profile of the metabolome.
- Increasing the number of mushroom metabolites indentified in the library and a further study of the relationship of these metabolites variation with different traits (i.e. variety, diseases, agricultural practices) will provide a much more precise knowledge of the mushroom metabolome and contribute to the general improvement of its cultivation and production.
- Investigating the use of these metabolomic techniques for the detection and discrimination of other types of damage to mushrooms such as microbial damage e.g. brown blotch disease and viral damage e.g. mushroom virus X.

# REFERENCES

Aguirre, L., Frias, J.M., Barry-Ryan, C. and Grogan, H. (2008). Assessing the effect of product variability on the management of the quality of mushrooms (*Agaricus bisporus*). Postharvest Biology and Technology **49**(2): 247-254.

Alonso-Salces, R.M., Ndjoko, K., Queiroz, E.F., Ioset, J.R., Hostettmann, K., Berrueta, L.A., Gallo, B. and Vicente, F. (2004). On-line characterisation of apple polyphenols by liquid chromatography coupled with mass spectrometry and ultraviolet absorbance detection. Journal of Chromatography A **1046**(1-2): 89-100.

Ammor, M.S., Argyri, A. and Nychas, G.E. (2009). Rapid monitoring of the spoilage of minced beef stored under conventionally and active packaging conditions using Fourier transform infrared spectroscopy in tandem with chemometrics. Meat Science **81**(3): 507-514.

Aranìbar, N., Singh, B.K., Stockton, G.W. and Ott, K.H. (2001). Automated mode-of-action detection by metabolic profiling. Biochemical and Biophysical Research Communications **286**(1): 150-155.

Arraes, F.B.M., Benoliel, B., Burtet, R.T., Costa, P.L.N., Galdino, A.S., Lima, L.H.A., Marinho-Silva, C., Oliveira-Pereira, L., Pfrimer, P., Procopio-Silva, L., Castelo-Branco Reis, V. and Felipe, M.S.S. (2005). General metabolism of the dimorphic and pathogenic fungus *Paracoccidiodes brasiliensis*. Genetics and Molecular Research **4**(2): 290-308.

Assaf, S., Hadar, Y. and Dosoretz, C.G. (1995). Biosynthesis of 13-hydroperoxylinoleate, 10-oxo-8-decenoic acids and 1-octen-3-ol from linoleic acid by a mycelial pellet homogenate of *Pleurotus pulmonarius*. Journal of Agricultural and Food Chemistry **43**(8): 2173-2178.

Aydemir, T. (2004). Partial purification and characterization of polyphenol oxidase from artichoke (Cynara scolymus L.) heads. Food Chemistry **87**(1): 59-67.

Baars, J.J., Op den Camp, H.J., Van Hoek, A.H., Van der Drift, C., Van Griensven, L.J., Visser, J. and Vogels, G.D. (1995). Purification and characterisation of NADP-dependent glutamate dehydrogenase from the commercial mushroom *Agaricus bisporus*. Current Microbiology **30**(4): 211-217.

Banwell, C.N. (1983). Fundamentals of molecular spectroscopy. London, McGraw-Hill Book Company.

Barker, M. and Rayens, W. (2003). Partial least squares for discrimination. Journal of Chemometrics **17**(3): 166-173.

Barros, L., Baptista, P., Correia, D.M., Casal, S., Oliveira, B. and Ferreira, I.C.F.R. (2007). Fatty acid and sugar compositions, and nutritional value of five wild edible mushrooms from Northeast Portugal. Food Chemistry **105**(1): 140-145.

Barros, L., Duenas, M., Ferreira, I.C.F.R., Baptista, P. and Santos-Buelga, C. (2009). Phenolic acids determination by HPLC-DAD-ESI/MS in sixteen different Portuguese wild mushrooms species. Food and Chemical Toxicology **47**(6): 1076-1079.

Bedair, M. and Sumner, L.W. (2008). Current and emerging mass-spectrometry technologies for metabolomics. TrAC Trends in Analytical Chemistry **27**(3): 238-250.

Beecher, T.M., Magan, N. and Burton, K.S. (2001). Water potentials and soluble carbohydrate concentrations in tissues of freshly harvested and stored mushrooms (*Agaricus bisporus*). Postharvest Biology and Technology **22**: 121-131.

Beelman, R.B. (1987). Factors influencing postharvest quality and shelf-life of fresh mushrooms. Mushroom News **35**(7): 12-18.

Beelman, R.B. and Edwards, C.G. (1989). Variability in the composition and nutritional value of the cultivated mushroom *Agaricus bisporus*. Mushroom News **37**: 20-26.

Beelman, R.B., Guthrie, B.D. and Royse, D.J. (1989). Influence of bacterial populations on postharvest deterioration of fresh mushrooms. Mushroom Science **12**(2): 655-665.

Beelman, R.B., Royse, D.J. and Chikthimmah, N. (2003). Bioactive components in button mushroom *Agaricus bisporus* (J. Lge) imbach (Agaricomycetideae) of nutritional, medicinal and biological importance (review). International Journal for Medicinal Mushrooms **5**(4): 1-18.

Belton, P.S., Colquhoun, I.J., Grant, A., Wellner, N., Field, J.M., Shewry, P.R. and Tatham, A.S. (1995). FTIR and NMR studies on the hydration of a high-Mr subunit of glutenin. International Journal of Biological Macromolecules **17**(2): 74-80.

Belton, P.S., Saffa, A.M. and Wilson, R.H. (1987). Use of Fourier transform infrared spectroscopy for quantitative analysis: a comparative study of different detection methods. Analyst **8**: 1117-1120.

Bligh, E.G. and Dyer, W.J. (1959). A rapid method of total lipid extraction and purification. Canadian Journal of Biochemical Physiology **37**: 911-917.

Bonzom, P.M.A., Nicolaou, A., Zloh, M., Baldeo, W. and Gibbons, W.A. (1999). NMR lipid profile of *Agaricus bisporus*. Phytochemistry **50**(8): 1311-1321.

Bravo, L. (1998). Polyphenols: chemistry, dietary sources, metabolism and nutritional significance. Nutrition Reviews **56**(11): 317-333.

Breiman, L. (1996). Bagging predictors. Machine Learning **24**(2): 123-140.

Breiman, L. (2001). Random Forests. Machine Learning **45**: 5-32.

Brennan, M., Le Port, G. and Gormley, R. (2000). Post-harvest treatment with citric acid or hydrogen peroxide to extend the shelf life of fresh sliced mushrooms. Lebensmittel-Wissenschaft und-Technologie **33**(4): 285-289.

Brosnan, J.T. (2003). Interorgan amino acid transport and its regulation. The Journal of Nutrition **133**: 2086-2072.

Brosnan, T. and Sun, D.W. (2004). Improving quality inspection of food products by computer vision--a review. Journal of Food Engineering **61**(1): 3-16.

Burton, K.S. (1989). The quality and storage life of *Agaricus bisporus*. Mushroom Science **12**: 287-293.

Burton, K.S. (2004). Cultural factors affecting mushroom quality - cause and control of bruising. Mushroom Science **16**: 397-402.

Burton, K.S., Hammond, J.B.W. and Mwamide, T. (1994). Protease activity in *Agaricus bisporus* during periodic fruiting (flushing) and sporophore development. Current Microbiology **28**: 275-278.

Burton, K.S., Molloy, S. and Willoughby, N. (2002). Water is the key to bruising. Horticultural Development News **88**: 29-30.

Byrne, P.F.S. and Brennan, P.J. (1975). The lipids of *Agaricus bisporus*. Journal of General Microbiology **89**: 245-255.

Carey, A.T. and O'Connor, T.P. (1991). Influence of husbandry factors on the quality of fresh mushrooms (*Agaricus bisporus*). Mushroom Science **13**: 673-682.

Catchpole, G.S., Beckmann, M., Enot, D.P., Mondhe, M., Zywicki, B., Taylor, J., Hardy, N., Smith, A., King, R.D., Kell, D.B., Fiehn, O. and Draper, J. (2005). Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. Proceedings of the national academy of sciences of the United States of America **102**(40): 14458-14462.

Cevallos-Cevallos, J.M., Reyes-De-Corcuera, J.I., Etxeberria, E., Danyluk, M.D. and Rodrick, G.E. (2009). Metabolomic analysis in food science: a review. Trends in Food Science & Technology **20**: 557-566.

Chang, S.T. and Miles, P.G. (1978). Historical record of the early cultivation of Lentinus in China. Mushroom Journal for the Tropics **7**: 31-37.

Chang, S.T. and Miles, P.G., Eds. (2004). Mushrooms: Cultivation, Nutritional Value, Medicinal Effect and Environmental Impact, CRC Press.

Chen, C.C. and Wu, C.M. (1984). Volatile components of mushroom (*Agaricus subrufecens*). Journal of Food Science **49**(4): 1208-1209.

Chen, H., Wortmann, A. and Zenobi, R. (2007). Neutral desorption sampling coupled to extractive electrospray ionization mass spectrometry for rapid differentiation of biosamples by metabolomic fingerprinting. Journal of Mass Spectrometry **42**(9): 1123-1135.

Cheung, L.M., Cheung, P.C.K. and Ooi, V.E.C. (2003). Antioxidant activity and total phenolics of edible mushroom extracts. Food Chemistry **81**: 249-255.

Cheung, M., Young, A.B. and Harrison, A.G. (1994). $O^{-\sqrt{}}$ and $OH^-$ chemical ionization of some fatty acid methyl esters and triacylglycerols. Journal of the American Society for Mass Spectrometry **5**(6): 553-557.

Chiang, P.D., Yen, C.T. and Mau, J.-L. (2006). Non-volatile taste components of canned mushrooms. Food Chemistry **97**(3): 431-437.

Cho, I.H., Kim, Y.-S. and Choi, H.-K. (2007). Metabolomic discrimination of different grades of pine-mushroom (*Tricholoma matsutake* Sing.) using $^1$H NMR spectrometry and multivariate data analysis. Journal of Pharmaceutical and Biomedical Analysis **43**(3): 900-904.

Choi, H.-K., Choi, Y.H., Verberne, M., Lefeber, A.W.M., Erkelens, C. and Verpoorte, R. (2004). Metabolic fingerprinting of wild type and transgenic tobacco plants by $^1$H NMR and multivariate analysis technique. Phytochemistry **65**(7): 857-864.

Christie, W.W. (2010). "The AOCS lipid library." Retrieved June, 2010, from http://www.aocs.org/index.cfm?navItemNumber=711.

Colak, A., Kolcuoglu, Y., Sesli, E. and Dalman, O. (2007). Biochemical composition of some Turkish fungi. Asian Journal of Chemistry **19**: 2193-2199.

Combet, E., Henderson, J., Eastwood, D. and Burton, K.S. (2006). Eight-carbon volatiles in mushrooms and fungi: properties, analysis , and biosynthesis. Mycoscience **47**: 317-326.

Combet, E., Henderson, J., Eastwood, D.C. and Burton, K.S. (2009). Influence of sporophore development, damage, storage, and tissue specificity on the enzymic formation of volatiles in mushrooms (*Agaricus bisporus*). Journal of Agricultural and Food Chemistry **57**(9): 3709-3717.

Courtecuise, R. and Duhem, B., Eds. (1995). Mushrooms and Toadstools. London, Harper Collins.

Crisan, E.V. and Sands, A., Eds. (1978). Nutritional value, In The Biology and cultivation of edible fungi. New York, Academic Press.

Cruz, C., Noel-Suberville, C. and Montury, M. (1997). Fatty acid content and some flavour compounds release in two strains of *Agaricus bisporus*, according to three stages of development. Journal of Agricultural and Food Chemistry **45**: 64-67.

Davis, R.A., Charlton, A.J., Oehlschlager, S. and Wilson, J.C. (2006). Novel feature selection method for genetic programming using metabolomic $^{1}$H NMR data. Chemometrics and Intelligent Laboratory Systems **81**(1): 50-59.

de Faria, R.O., Moure, V.R., Lopes de Almeida Amazonas, M., Krieger, N. and Mitchell, D.A. (2007). The biotechnological potential of mushroom tyrosinases. Food Technology and Biotechnology **45**(3): 287-294.

de Laeter, J.R., Bohlke, J.K., De Bievre, P., Hidaka, H., Peiser, H.S., Rosman, K.J.R. and Taylor, P.D.P. (2003). Atomic weights of the elements. Review 2000 (IUPAC Technical Report). Pure and Applied Chemistry **75**(6): 683-800.

De Oliveira, L.F.C., Colombra, R. and Edwards, H.G.M. (2002). Fourier transform Raman spectroscopy of honey. Applied Spectroscopy **56**(3): 306-311.

del Rio, L.A., Pastori, G.M., Palma, J.M., Sandalio, L.M., Sevilla, F., Corpas, F.J., Jimenez, A., Lopez-Huertas, E. and Hernandez, J.A. (1998). The activated oxygen role of peroxisomes in senescence. Plant Physiology **116**(4): 1195-1200.

Devaux, P.G., Horning, M.G. and Horning, E.C. (1971). Benzyl-oxime derivative of steroids; a new metabolic profile procedure for human urinary steroids. Analytical Letters **4**(151-160).

Di Mario, F., Rapanà, P., Tomati, U. and Galli, E. (2008). Chitin and chitosan from Basidiomycetes. International Journal of Biological Macromolecules **43**(1): 8-12.

Dobson, G., Shepherd, T., Verrall, S.R., Conner, S., McNicol, J.W., Ramsay, G. and Stewart, D. (2008). Phytochemical diversity in tubers of potato cultivars and landraces using GC/MS metabolomics approach. Journal of Agricultural and Food Chemistry **56**: 10280-10291.

Donker, H.C.W. and Van As, H. (1999). Cell water balance of white button mushrooms (*Agaricus bisporus*) during its post-harvest lifetime studied by quantitative magnetic resonance imaging. Biochimica et Biophysica Acta (BBA) - General Subjects **1427**(2): 287-297.

Duckworth, H.W. and Coleman, J.E. (1970). Physiochemical and kinetic properties of mushroom tyrosinase. Journal of Biological Chemistry **245**(7): 1613-1625.

Dunn, W.B., Bailey, N.J.C. and Johnson, H.E. (2005). Measuring the metabolome: current analytical technologies. The Analyst **130**: 606-625.

Dunn, W.B. and Ellis, D.I. (2005). Metabolomics: Current analytical platforms and methodologies. TrAC Trends in Analytical Chemistry **24**(4): 285-294.

Eastwood, D. and Burton, K.S. (2002). Mushrooms - a matter of choice and spoiling oneself. Microbiology Today **29**: 18-19.

Eastwood, D.C., Challen, M.P., Zhang, C., Jenkins, H., Henderson, J. and Burton, K.S. (2008). Hairpin-mediated down-regulation of the urea cycle enzyme argininosuccinate lyase in *Agaricus bisporus*. Mycological Research **112**: 708-716.

Eastwood, D.C., Kingsnorth, C.S., Jones, H.E. and Burton, K.S. (2001). Genes with increased transcript levels following harvest of the sporophore of *Agaricus bisporus* have multiple physiological roles. Mycological Research **105**(10): 1223-1230.

Efron, B. and Tibshirani, R.J. (1983). An introduction to the boostrap. New York, Chapman & Hall.

Eissa, H.A.A. (2007). Effect of chitosan coating on shelf life and quality of fresh-cut mushroom. Journal of Food Quality **30**(5): 623-645.

El-Khoury, A.E., Ed. (1999). Methods for investigation of amino acid and protein metabolism. Determination of amino acid isotopic enrichment: Methods, difficulties, and calculations. Florida, USA, CRC Press.

Ellis, D.I. and Goodacre, R. (2001). Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends. Trends in Food Science & Technology **12**(11): 414-424.

ElMasry, G., Wang, N. and Vigneault, C. (2009). Detecting chilling injury in red delicious apple using hyperspectral imaging and neural networks. Postharvest Biology and Technology **52**(1): 1-8.

Esquerre, C., Gowen, A.A., O'Donnell, C. and Downey, G. (2009). Initial studies on the quantitation of bruise damage and freshness in mushrooms using visible-near infrared spectroscopy. Journal of Agricultural and Food Chemistry **57**: 1903-1907.

Evans, C. and Jorgenson, J. (2004). Multidimensional LC-LC and LC-CE for high-resolution separations of biological molecules. Analytical and Bioanalytical Chemistry **378**(8): 1952-1961.

Everall, N., King, B. and Clegg, I. (2000). The raman effect. Chem Britain: 40-43.

Falguera, V., Pagán, J. and Ibarz, A. (2010). A kinetic model describing melanin formation by means of mushroom tyrosinase. Food Research International **43**(1): 66-69.

Fan, T.W.M. (1996). Metabolite profiling by one- and two- dimensional NMR analysis of complex mixtures. Progress in Nuclear Magnetic Resonance Spectroscopy(28): 161-219.

Ferguson, C.H.R. and Simon, E.W. (1973). Membrane lipids in senescing green tissues. Journal of Experimental Botany **24**(2): 307-316.

Fermor, T.R., Henry, M.B., Fenlon, J.S., Glenister, M.J., Lincoln, S.P. and Lynch, J.M. (1991). Development and application of a biocontrol system for bacterial blotch of the cultivated mushroom. Crop Protection **10**(4): 271-278.

Fernie, A.R. (2007). Metbolic profiling GC/MS protocol for quadrupole. A. O'Gorman. Dublin.

Fernie, A.R., Trethewey, R.N., Krotzky, A.J. and Willmitzer, L. (2004). Metabolite profiling: from diagnostics to systems biology. Nature Reviews Molecular Cell Biology **5**: 763-769.

Fiehn, O. (2008). Extending the breadth of metabolite profiling by gas chromatography coupled to mass spectrometry. TrAC Trends in Analytical Chemistry **27**(3): 261-269.

Fiehn, O., Kopka, J., Dormann, T., Altmann, T., Trethewey, R.N. and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. Nature Biotechnology **18**: 1157-1161.

Flegg, P.B., Spencer, D.M. and Wood, D.A., Eds. (1985). The biology and technology of the cultivated mushroom. Chichester, UK, John Wiley & Sons.

Fujita, T., Komemushi, S. and Yamagata, K. (1991). Contents of amino acids, organic acids and 5'-nucleotides in *Tricholoma giganteum*. Journal of the Science of Food and Agriculture **55**(1): 159-162.

Fukusaki, E. and Kobayashi, A. (2005). Plant metabolomics: potential for practical operation. Journal of Bioscience and Bioengineering **100**(4): 347-354.

García-Cañas, V., Simó, C., León, C. and Cifuentes, A. (2010). Advances in Nutrigenomics research: Novel and future analytical approaches to investigate the biological activity of natural compounds and food functions. Journal of Pharmaceutical and Biomedical Analysis **51**(2): 290-304.

Garip, S., Gozen, A.C. and Severcan, F. (2009). Use of Fourier transform infrared spectroscopy for rapid comparative analysis of Bacillus and Micrococcus isolates. Food Chemistry **113**(4): 1301-1307.

Gasowska, B., Kafarski, P. and Wojtasek, H. (2004). Interaction of mushroom tyrosinase with aromatic amines, *o*-diamines and *o*-aminophenols. Biochimica et Biophysica Acta 1 **673**: 170-177.

Gaze, R.H., Calvo-Bado, L., Challen, M.P., Adie, B. and Romaine, C.P. (2000). A new virus disease of *Agaricus bisporus*? Mushroom Science **15**: 701-705.

Genome, I.*A.* (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**(6814): 796-815.

Gibney, M.J., Walsh, M., Brennan, L., Roche, H.M., German, B. and Ommen, B. (2005). Metabolomics in human nutrition: opportunities and challenges. The American Journal of Clinical Nutrition **82**: 497-503.

Goff, S.A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.-l., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A. and Briggs, S. (2002). A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. japonica). Science **296**(5565): 92-100.

Gormley, R. (1986). Post-harvest quality of mushrooms. 4th National Mushroom Conference, Dublin, Ireland.

Gormley, T.R. (1975). Chill storage of mushrooms. Journal of the Science of Food and Agriculture **26**: 401-411.

Gormley, T.R. (1987). Handling, Packaging and Transportation of Fresh Mushrooms. Proceedings of the 5th National Mushroom Conference, Kinsealy, Dublin.

Gormley, T.R. and O'Sullivan, L. (1975). Use of a simple reflectometer to test mushroom quality. The Mushroom Journal **34**: 344-346.

Gowen, A.A., Downey, G., Esquerre, C. and O'Donnell, C.P. (2010). A simple method for estimation of PLS model dimensionality using regression vectors. Journal of Chemometrics: in press.

Gowen, A.A., O' Donnell, C.P., Taghizadeh, M., Cullen, P.J., Frias, J.M. and Downey, G. (2008a). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (*Agaricus bisporus*). Journal of Chemometrics **22**(3-4): 259-267.

Gowen, A.A., O'Donnell, C.P., Taghizadeh, M., Gaston, E., O'Gorman, A., Cullen, P.J., Frias, J.M., Esquerre, C. and Downey, G. (2008b). Hyperspectral imaging for the investigation of quality deterioration in sliced mushrooms (*Agaricus bisporus*) during storage. Sensing and Instrumentation for Food Quality and Safety **2**(3): 133-143.

Gowen, A.A., Taghizadeh, M. and O'Donnell, C.P. (2009). Identification of mushrooms subjected to freeze damage using hyperspectral imaging. Journal of Food Engineering **93**(1): 7-12.

Granitto, P.M., Gasperi, F., Biasioli, F., Trainotti, E. and Furlanello, C. (2007). Modern data mining tools in descriptive sensory analysis: A case study with a Random forest approach. Food Quality and Preference **18**(4): 681-689.

Griffin, J.L. and Bollard, M.E. (2004). Metabonomics: its potential as a tool in toxicology for safety assessment and data integration. Current Drug Metabolism **5**: 389-398.

Griffiths, P.R. and de Haseth, J.A. (2007). Fourier transform infrared spectrometry, 2nd edition. New York, John Wiley & sons.

Grogan, H. (2008). Changes and challenges for the Irish mushroom industry. Mushroom Business **28**: 22-23.

Halket, J.M., Waterman, D., Przyborowska, A.M., Patel, R.K.P., Fraser, P.D. and Bramley, P.M. (2005). Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. Journal of Experimental Botany **56**(410): 219-243.

Hammond, J.B.W. (1979). Changes in composition of harvested mushrooms (*Agaricus bisporus*). Phytochemistry **18**: 415-418.

Hammond, J.B.W. and Nichols, R. (1975). Change in respiration and soluble carbohydrate during the post-harvest storage of mushrooms *Agaricus bisporus*. Journal of the Science of Food and Agriculture **26**: 835-842.

Harrigan, G.C. and Goodacre, R., Eds. (2003). Metabolic profiling: its role in biomarker discovery and gene function analysis. Boston, Kluwer Academic Publishers.

Heleno, S.A., Barros, L., Sousa, M.J., Martins, A. and Ferreira, I.C.F.R. (2009). Study and characterization of selected nutrients in wild mushrooms from Portugal by gas chromatography and high performance liquid chromatography. Microchemical Journal **93**: 195-199.

Helland, I.S. (2001). Some theoretical aspects of partial least squares regression. Chemometrics and Intelligent Laboratory Systems **58**(2): 97-107.

Hiroi, M. and Tsuyuki, H. (1988). Comparison of fatty acids in fruit body and spore mushrooms. Bulletin of the College of Agricultural and Veterinary Medicine **45**: 100-109.

Holtz, R. and Schisler, L. (1971). Lipid metabolism of *Agaricus bisporus* (Lange) sing.: I. Analysis of sporophore and mycelial lipids. Lipids **6**(3): 176-180.

Hong, J.S., Kim, Y.H., Lee, K.R., Kim, M.K., Cho, C.I. and Part, K.K.H. (1988). Composition of organic and fatty acid in *Pleurotus ostreatus*, *Lentinus edodes* and *Agaricus bisporus*. Korean Journal of Food Science and Technology **20**(1): 100-105.

Horning, E.C. and Horning, M.G. (1971). Metabolic profiles: gas-phase methods for analysis of metabolites. Clinical Chemistry **17**(8): 802-809.

Hotellin, H. (1933). Analysis of a complex of statistical variables into principal components Journal of Educational Psychology **24**: 417-441.

http://www.mushroombusiness.com. "Irish mushroom sector recognises need to change." Retrieved 02/06/2010.

Hui, Y.H. (2006). Handbook of food science, technology and engineering. Florida, CRC Press.

Hult, K. and Gatenbeck, S. (1978). Production of NADPH in the mannitol cycle and its relation to polyketide formation of *Alternaria alternata*. European Journal of Biochemistry **88**(2): 607-612.

Ikawa, M., Borowski, P.T. and Chakravarti, A. (1968). Choline and inositol distribution in algae and fungi. Applied Microbiology **16**(4): 620-623.

Ikeda, T., Kanaya, S., Yonetani, T., Kobayashi, A. and Fukusaki, E. (2007). Prediction of Japanese green tea ranking by fourier transform near-infrared reflectance spectroscopy. Journal of Agricultural and Food Chemistry **55**(24): 9908-9912.

Jasinghe, V.J. and Perera, C.O. (2006). Ultraviolet irradiation: The generator of vitamin $D_2$ in edible mushrooms. Food Chemistry **95**(4): 638-643.

Jennings, D.H. (1984). Polyol metabolism in fungi. Advances in Microbial Physiology **25**: 149-193.

Johnson, H.E., Broadhurst, D., Goodacre, R. and Smith, A.R. (2003). Metabolic fingerprinting of salt-stressed tomatoes. Phytochemistry **62**(6): 919-928.

Jolivet, S., Arpin, N., Wichers, H.J. and Pellon, G. (1998). *Agaricus bisporus* browning: a review. Mycological Research **102**(12): 1459-1483.

Kaddurah-Daouk, R. and Krishnan, K.R.R. (2009). Metabolomics: a global biochemical approach to the study of central nervous system diseases. Neuropsychopharmacology **34**(1): 173-186.

Kadegowda, A.K.G., Piperova, L.S. and Erdman, R.A. (2008). Principal component and multivariate analysis of milk long-chain fatty acid composition during diet-induced milk fat depression. Journal of Dairy Science **91**(2): 749-759.

Kalač, P. (2009). Chemical composition and nutritional value of European species of wild growing mushrooms: A review. Food Chemistry **113**(1): 9-16.

Karlinski, L., Ravnskov, S., Kieliszewska-Rokicka, B. and Larsen, J. (2007). Fatty acid composition of various ectomycorrhizal fungi and ectomycorrhizas of Norway spruce. Soil Biology and Biochemistry **39**(4): 854-866.

KEGG. (1995, July 2010). "Pathway maps and pathway modules." KEGG, from http://www.genome.jp/kegg/.

Kell, D.B. (2004). Metabolomics and systems biology: making sense of the soup. Current Opinion in Microbiology **7**(3): 296-307.

Kell, D.B. (2006). Systems biology, metabolic modelling and metabolomics in drug discovery and development. Drug Discovery Today **11**: 1085-1092.

Kim, H.K., Khan, S., Wilson, E.G., Kricun, S.D.P., Meissner, A., Goraler, S., Deelder, A.M., Choi, Y.H. and Verpoorte, R. (2010). Metabolic classification of South American *Ilex* species by NMR-based metabolomics. Phytochemistry **71**(7): 773-784.

Kim, K.M., Ko, J.A., Lee, J.A., Park, H.J. and Hanna, M.A. (2006). Effect of modified atmosphere packaging on the shelf-life of coated, whole and sliced mushrooms. LWT - Food Science and Technology **39**(4): 365-372.

Kim, M.Y., Chung, I.M., Lee, S., J., Ahn, J.K., Kim, E.H., Kim, M.J., Kim, S.L., Moon, H.I., Ro, H.M., Kang, E.Y., Seo, S.H. and Song, H.K. (2009). Comparison of free amino, carbohydrates concentrations in Korean edible and medicinal mushrooms. Food Chemistry **113**(2): 386-393.

Kim, M.Y., Sequin, P., Ahn, J.K., Kim, J.J., Chun, S.C., Kim, E.H., Seo, S.Y., Kang, E.Y., Kim, S.L., Park, Y.J., Ro, H.M. and Chung, I.M. (2008). Phenolic compound concentration and antioxidant activities of edible and medicinal mushrooms from Korea. Journal of Agricultural and Food Chemistry **56**(16): 7265-7270.

Kopka, J., Fernie, A.R., Weckwerth, W., Gibon, Y. and Stitt, M. (2004). Metabolite profiling in plant biology: platforms and destinations. Genome Biology **5**(6): 109.

Kubicek, C.P. (1988). Regulatory aspects of the tricarboxylic acid cycle in filamentous fungi - A review. Transactions of the British Mycological Society **90**(3): 339-349.

Kuraswa, S., Sugahara, T. and Hayashi, J. (1982). Studies on dietary fiber of mushrooms and edible wild plants. Nutrition Reports International **26**: 167-173.

Kushalappa, A.C., Vikram, A. and Raghavan, G.S.V. (2008). Metabolomics of headspace gas for diagnosing diseases of fruits and vegetables after harvest. Stewart Postharvest Review **4**(2): 1-7.

Kuyper, L., Weinert, I.A.G. and McGill, A.E.J. (1993). The Effect of Modified Atmosphere Packaging and Addition of Calcium Hypochlorite on the Atmosphere Composition, Colour and Microbial Quality of Mushrooms. Lebensmittel-Wissenschaft und-Technologie **26**(1): 14-20.

Lahner, B., Gong, J., Mahmoudian, M., Smith, E.L., Abid, K.B., Rogers, E.E., Guerinot, M.L., Harper, J.F., Ward, J.M., McIntyre, L., Schroeder, J.I. and Salt, D.E. (2003). Genomic scaling of nutrient and trace elements in *Arabidopsis thaliana*. Nature Biotechnology **21**: 1215-1221.

Lee, J.S. (1999). Effects of modified atmosphere packaging on the quality of chitosan and $CaCl_2$ coated mushroom (*Agaricus bisporus*). Korean Journal of Food Science and Technology **31**(5): 1308-1314.

Lewis, D.H. and Smith, D.C. (1967). Sugar alcohols (polyols) in fungi and green plants. New Phytologist **66**(2): 143-184.

Lindon, J.C., Holmes, E., Bollard, M.E., Stanley, E.G. and Nicholson, J.K. (2004a). Metabonomics technologies and their application in physiological monitoring, drug safety assessment and disease diagnosis. Biomarkers **9**(1): 1-31.

Lindon, J.C., Holmes, E. and Nicholson, J.K. (2001). Pattern recognition methods and applications in biomedical magnetic resonance. Progress in Nuclear Magnetic Resonance Spectroscopy **39**: 1-40.

Lindon, J.C., Holmes, E. and Nicholson, J.K. (2004b). Toxicological applications of magnetic resonance. Progress in Nuclear Magnetic Resonance Spectroscopy **45**: 109-143.

Lisec, J., Schauer, N., Kopka, J., Willmitzer, L. and Fernie, A.R. (2006). Gas chromatography mass spectrometry-based metabolite profiling in plants. Nature Protocols **1**: 387-396.

Loewus, F.A. and Murthy, P.P.N. (2000). Review: *myo*-inositol metabolism in plants. Plant Science **150**: 1-19.

Lonchamp, L., Barry-Ryan, C. and Devereux, M. (2009). Identification of volatile quality markers of ready-to-use lettuce and cabbage. Food Research International **42**: 1077-1086.

Luthria, D.L., Lin, L.Z., Robbins, R.J., Finley, J.W., Banuelos, G.S. and Harnly, J.M. (2008). Discriminating between cultivars and treatments of broccoli using mass spectral fingerprinting and analysis of variance-principal component analysis. Journal of Agricultural and Food Chemistry **56**(21): 9819-9827.

MacCanna, C. (1984). Commercial Mushroom Production. Dublin, An Foras Taluntais - The Agricultural Institute.

Madsen, R., Lundstedt, T. and Trygg, J. (2010). Chemometrics in metabolomics--A review in human disease diagnosis. Analytica Chimica Acta **659**(1-2): 23-33.

Maggioni, A., Passera, C., Renosto, F. and Benetti, E. (1968). Composition of cultivated mushrooms (Agaricus bisporus) during the growing cycle as affected by the nitrogen source introduced in composting. Journal of Agricultural and Food Chemistry **16**(3): 517-519.

Manzi, P., Aguzzi, A. and Pizzoferrato, L. (2001). Nutritional value of mushrooms widely consumed in Italy. Food Chemistry **73**(3): 321-325.

Manzi, P., Marconi, S., Aguzzi, A. and Pizzoferrato, L. (2004). Commercial mushrooms: nutritional quality and effect of cooking. Food Chemistry **84**(2): 201-206.

Martinez, M.V. and Whitaker, J.R. (1995). The biochemistry and control of enzymatic browning. Trends in Food Science & Technology **6**(6): 195-200.

Marzluf, G.A. (1997). Genetic regulation of nitrogen metabolism in the fungi. Microbiology and Molecular Biology Reviews **61**(1): 17-32.

Massart, D.L., Vandeginste, B.G.M., Lewi, L.M.C. and Smeyers-Verbeke, J. (1997). Handbook of Chemometrics and Qualimetrics: Part A. Amsterdam, The Netherlands, Elsevier.

Mau, J.E.L., Beelman, R.B. and Ziegler, G.R. (1992). 1-Octen-3-ol in the Cultivated Mushroom, *Agaricus bisporus*. Journal of Food Science **57**(3): 704-706.

Mau, J.L., Chen, P.R. and Yang, J.H. (1998). Ultraviolet irradiation increased vitamin $D_2$ content in edible mushrooms. Journal of Agricultural and Food Chemistry **46**: 5269-5272.

McCloskey, J.A., Ed. (1970). Topics in Lipid Chemistry. London, Logos Press.

McDougall, G., Martinussen, I. and Stewart, D. (2008). Towards fruitful metabolomics: High throughput analyses of polyphenol composition in berries using direct infusion mass spectrometry. Journal of Chromatography B **871**(2): 362-369.

McLafferty, F.W. (1959). Mass spectrometric analysis. Molecular rearrangements. Analytical Chemistry **31**(1): 82-87.

Mederios, P.M. and Simoneit, B.R.T. (2007). Analysis of sugars in environmental samples by gas chromatography-mass spectrometry. Journal of Chromatography A **1141**: 271-278.

Miklus, M.B. and Beelman, R. (1996). $CaCl_2$ treated irrigation water applied to mushroom crops (*Agaricus bisporus*) increases Ca concentration and improves postharvest quality and shelf life. Mycologia **88**(3): 403-409.

Militello, V., Casarino, C., Emanuele, A., Giostra, A., Pullara, F. and Leone, M. (2004). Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering. Biophysical Chemistry **107**(2): 175-187.

Moalemiyan, M., Vikram, A. and Kushalappa, A.C. (2007). Detection and discrimination of two fungal diseases of mango (cv. Keitt) fruits based on volatile metabolite profiles using GC/MS. Postharvest Biology and Technology **45**(1): 117-125.

Moco, S., Bino, R.J., Vorst, O., Verhoeven, H.A., de Groot, J., van Beek, T.A., Vervoort, J. and Ric de Vos, C.H. (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiology **141**(4): 1205-1218.

Moco, S., Forshed, J., De Vos, R.C.H., Bino, R.J. and Vervoort, J. (2008). Intra- and inter-metabolite correlation spectroscopy of tomato metabolomics data obtained by liquid chromatography-mass spectrometry and nuclear magnetic resonance. Metabolomics **4**(3): 202-215.

Mohacek-Grosev, V., Bozac, R. and Puppels, G.J. (2001). Vibrational spectroscopic characterization of wild growing mushrooms and toadstools. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **57**(14): 2815-2829.

Mohapatra, D., Frias, J.M., Oliveira, F.A.R., Bira, Z.M. and Kerry, J. (2008). Development and validation of a model to predict enzymatic activity during storage of cultivated mushrooms (*Agaricus bisporus* spp.). Journal of Food Engineering **86**: 39-48.

Mohsenin, N.N. (1986). Physical Properties of Plant and Animal Materials. New York, Gordon and Breach.

Monton, M.R.N. and Soga, T. (2007). Metabolome analysis by capillary electrophoresis-mass spectrometry. Journal of Chromatography A **1168**(1-2): 237-246.

Moore, D. and Al-Gharawi, A. (1976). An elevated level of NADP-linked glutamate dehydrogenase is not a general feature of the caps of agaric sporophores. Transactions of the British Mycological Society **66**: 149-150.

Morcuende, R., Krapp, A., Hurry, V. and Stitt, M. (1998). Sucrose-feeding leads to increased rates of nitrate assimilation, increased rates of alpha-oxoglutarate synthesis, and increased synthesis of a wide spectrum of amino acids in tobacco leaves. Planta **206**: 394-409.

Murahashi, S. (1938). Uber die riechstoffe des matutake. Scienitific Papers of the Institute of Physical and Chemical Research **34**: 155-172.

Neary, N. (2003). Challenging times for Irish mushrooms. Farmers Journal.

Nelson, W.H., Ed. (1991). Modern techniques for rapid microbiological analysis. New York, VCH publishers.

Ninonuevo, M.R., Park, Y., Yin, H., Zhang, J., Ward, R.E., Clowers, B.H., German, J.B., Freeman, S.L., Killeen, K., Grimm, R. and Lebrilla, C.B. (2006). A strategy for

annotating the human milk glycome. Journal of Agricultural and Food Chemistry **54**(20): 7471-7480.

Norli, H.R., Christiansen, A. and Holen, B. (2010). Independent evaluation of a commercial deconvolution reporting software for gas chromatogrpahy mass spectrometry analysis of pesticide residues in fruits and vegetables. Journal of Chromatography A **1217**: 2056-2064.

O'Brien, M., Gentry, J.P. and Gibson, R.C. (1995). Vibrating characteristics of fruit as related to in-transit injury. Transactions of the American Scoiety of Agricultural Engineers **8**(2): 241-243.

O'Brien, M. and Guillou, R. (1969). An in-transit simulator for fruit handling studies. Transactions of the American Scoiety of Agricultural Engineers **12**(1): 94-97.

O'Gorman, A., Downey, G., Gowen, A.A., Barry-Ryan, C. and Frias, J.M. (2010). Use of fourier transform infrared spectroscopy and chemometric data analysis to evaluate damage and age in mushrooms (*Agaricus bisporus*) grown in Ireland. Journal of Agricultural and Food Chemistry **58**(13): 7770-7776.

O'Hara, P., Slabas, A.R. and Fawcett, T. (2002). Fatty acid and lipid biosynthetic genes are expressed at constant molar ratios but different absolute levels during embryogenesis. Plant Physiology **129**: 310-320.

Oliver, S.G. (1997). Yeast as a navigational aid in genome analysis. Microbiology **143**: 1483-1487.

Oliver, S.G., Winson, M.K., Kell, D.B. and Baganz, F. (1998). Systematic functional analysis of the yeast genome. Trends in Biotechnology **16**(10): 373-378.

Opitz, J. (2007). Electron-impact ionization of benzoic acid, nicotinic acid and their n-butyl esters: An approach to regioselective proton affinities derived from ionization

and appearance energy data. International Journal of Mass Spectrometry **265**(1): 1-14.

Öztürk, M., Emin Duru, M., Kivrak, S., Mercan-Dogan, N., Türkoglu, A. and Özler, M.A. (2011). *In vitro* antioxidant, anticholinesterase and antimicrobial activity studies on three *Agaricus* species with fatty acid compositions and iron contents: A comparative study on the three most edible mushrooms. Food and Chemical Toxicology **In press**.

Paine, S.G. (1919). Studies in bacteriosis II: a brown blotch disease of cultivated mushrooms. Annals of Applied Biology **5**: 206-219.

Pang, H., Lin, A., Holford, M., Enerson, B.E., Lu, B., Lawton, M.P., Floyd, E. and Zhao, H. (2006). Pathway analysis using random forests classification and regression. Bioinformatics **22**(16): 2028-2036.

Parrou, J.L., Teste, M.A. and Francois, J. (1997). Effects of various types of stress on the metabolism of reserve carbohydrates in *Saccharomyces cerevisiae*: genetic evidence for a stress-induced recycling of glycogen and trehalose. Microbiology **143**(6): 1891-1900.

Pedreschi, R., Franck, C., Lammertyn, J., Erban, A., Kopka, J., Hertog, M.L.A.T.M., Verlinden, B. and Nicolai, B. (2009). Metabolic profiling of 'Conference' pears under low oxygen stress. Postharvest Biology and Technology **51**: 123-130.

Pereira, G.E., Gaudillere, J.P., van Leeuwen, C., Hilbert, G., Maucourt, M., Deborde, C., Moing, A. and Rolin, D. (2006). $^{1}$H NMR metabolite fingerprints of grape berry: Comparison of vintage and soil effects in Bordeaux grapevine growing areas. Analytica Chimica Acta **563**: 346-352.

Pers, T.H., Martin, F.P., Verdich, C., Holst, C., Johansen, J.V., Astrup, A., Polak, J., Martinez, J.A., Rezzi, S., Blaak, E.E., Saris, W.H.M., Kochhar, S., Macdonald,

I.A., Sørensen, T.I.A. and Ramadan, Z. (2008). Prediction of fat oxidation capacity using $^{1}$H NMR and LC-MS lipid metabolomic data combined with phenotypic data. Chemometrics and Intelligent Laboratory Systems **93**: 34-42.

Pierce, J.A. and Rast, D.M. (1995). A comparison of native and synthetic mushroom melanins by fourier-transform infrared spectroscopy. Phytochemistry **39**(1): 49-55.

Preisner, O., Lopes, J.A. and Menezes, J.C. (2008). Uncertainty assessment in FT-IR spectroscopy based bacteria classification models. Chemometrics and Intelligent Laboratory Systems **94**(1): 33-42.

Prostenik, M., Burcar, I., Castek, A., Cosovic, C., Golem, J., Jandric, Z., Kljaic, K. and Ondrusek, V. (1978). Lipids of higher fungi. III. The fatty acids and 2-hydroxy fatty acids in some species of Basidiomycetes. Chemistry and Physics of Lipids **22**(2): 97-103.

R Development Core Team (2007). R: a language and environment for statistical computing. V. R Foundation for Statistical Computing, Austria, 2007.

R Development Core Team (2009). R: a language and environment for statistical computing. V. R Foundation for Statistical Computing, Austria, 2007.

Rajarathnam, S., Shashirekha, M.N. and Rashmi, S. (2003). Biochemical changes associated with mushroom browning in *Agaricus bisporus* (Lange) Imbach and *Pleurotus florida* (Block & Tsao): commercial implications. Journal of the Science of Food and Agriculture **83**(14): 1531-1537.

Ramautar, R., Demirci, A. and de. Jong, G.J. (2006). Capillary electrophoresis in metabolomics. TrAC Trends in Analytical Chemistry **25**(5): 455-466.

Rast, D., Stauble, E.J. and Zobrist, P. (1976). The krebs cycle and control of dormancy in the *Agaricus bisporus* spore. New Phytologist **76**(3): 469-477.

Ratcliffe, R.G. and Shachar-Hill, Y. (2005). Revealing metabolic phenotypes in plants: Inputs from NMR analysis. Biological reviews of the Cambridge philosophical society **80**(1): 27-43.

Reinbothe, H., Wasternack, C. and Miersch, J. (1967). Harnstoff-metabolismus by Basidiomyceten - IV. Untersuchungen zur physiologie des harnstoffs. Flora **158**: 27-57.

Reineccius, G.A., Kavanagh, T.E. and Keeney, P.G. (1970). Identification and quantitation of free neutral carbohydrates in milk products by gas-liquid chromatography and mass spectrometry. Journal of Dairy Science **53**(8): 1018-1022.

Ribeiro, B., Lopes, R., Andrade, P.B., Seabra, R.M., Gonçalves, R.F., Baptista, P., Quelhas, I. and cia Valentão, P. (2008). Comparative study of phytochemicals and antioxidant potential of wild edible mushroom caps and stipes. Food Chemistry **110**(11): 47-56.

Roessner-Tunali, U., Hegemann, B., Lytovchenko, A., Carrari, F., Bruedigam, C., Granot, D. and Fernie, A.R. (2003). Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. Plant Physiology **133**(1): 84-99.

Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N. and Willmitzer, L. (2000). Simultaneous analysis of metabolite in potato tuber by gas chromatography-mass spectrometry. The Plant Journal **23**(1): 131-142.

Rosenfeld, L. (2002). Clinical Chemistry Since 1800: Growth and Development. Clinical Chemistry **48**(1): 186-197.

Roy, S., Anantheswaran, R.C. and Beelman, R.B. (2006). Fresh mushroom quality as affected by modified atmosphere packaging. Journal of Food Science **60**(2): 334-340.

Rudell, D.R., Mattheis, J.P. and Curry, E.A. (2008). Prestorage ultraviolet white light irradiation alters apple peel metabolome. Journal of Agricultural and Food Chemistry **56**(3): 1138-1147.

Ruijter, G.J.G., Bax, M., Patel, H., Flitter, S.J., Van De Vondervoort, P.J.I., De Vries, R.P., Van Kuyk, P.A. and Visser, J. (2003). Mannitol is required for stress tolerance in *Aspergillus niger* conidiospores. Eukaryotic Cell **2**(4): 690-698.

Ruiz-Herrera, J., Ed. (1992). Fungal cell wall, structure, synthesis and assembly. Florida, CRC Press.

Sandula, J., Kogan, G., Kacurakova, M. and Machova, E. (1999). Microbial (1->3)-beta-d-glucans, their preparation, physico-chemical characterization and immunomodulatory activity. Carbohydrate Polymers **38**(3): 247-253.

Sato, S., Soga, T., Nishioka, T. and Tomita, M. (2004). Simultaneous determination of the main metabolites in rice leaves using capillary electrophoresis mass spectrometry and capillary electrophoresis diode array detection. Plant Journal **40**: 151-163.

Schisler, L.C., Sinden, J.W. and Sigel, E.M. (1967). Etiology, symptomatology and epidemiology of a virus disease of cultivated mushrooms. Phytopathology **57**: 519-526.

Schoots, A.C. and Leclercq, P.A. (1979). Chemical ionisation mass spectrometry of trimethylsilylated carbohydrates and organic acids retained in uremic serum. Biomedical Mass Spectrometry **6**(11): 502-507.

Schouten, R.E., Jongbloed, G., Tijskens, L.M.M. and van Kooten, O. (2004). Batch variability and cultivar keeping quality of cucumber. Postharvest Biology and Technology **32**(3): 299-310.

Schouten, R.E. and Van Kooten, O. (1998). Keeping quality of cucumber batches: Is it predictable? Acta Horticulturae **476**: 349-355.

Schrader, B., Schulz, H., Andreev, G.N., Klump, H.H. and Sawatzki, J. (2000). Non-destructive NIR-FT-Raman spectroscopy of plant and animal tissues, of food and works of art. Talanta **53**(1): 35-45.

Senatore, F., Dini, A. and Marino, M. (1988). Chemical constituents of some Basidiomycetes. Journal of Science Food Agriculture **45**: 337-345.

Shi, T., Seligson, D., Belldegrun, A.S., Palotie, A. and Horvath, S. (2005). Tumor classification by tissue microarray profiling: Random forest clustering applied to renal cell carcinoma. Modern Pathology **18**(4): 547-557.

Shin, E.C., Craft, B.D., Pegg, R.B., Phillips, D.R. and Eitenmiller, R.R. (2010). Chemometric approach to fatty acid profiles in Runner-type peanut cultivars by principal component analysis (PCA). Food Chemistry **119**(3): 1262-1270.

Simon, A., Gonzalez-Fandos, E. and Tobar, V. (2005). The sensory and microbiological quality of fresh sliced mushroom (*Agaricus bisporus* L.) packaged in modified atmospheres. International Journal of Food Science and Technology **40**(9): 943-952.

Singh, A. and Singh, Y. (1992). Effects of vibration during transportation on the quality of tomatoes. Agricultural Mechanization in Asia, Africa, and Latin America **23**(2): 70-72.

Skovgaard, N. (2002). The Mycota. A comprehensive treatise on fungi as experimental systems for basic and applied research: Volume X Industrial Applications. International Journal of Food Microbiology **77**(3): 233-234.

Smirnoff, N. and Cumbes, Q.J. (1989). Hydroxyl radical scavenging activity of compatible solutes. Phytochemistry **28**(4): 1057-1060.

Sobolev, A.P., Brosio, E., Gianferri, R. and Segre, A.L. (2005). Metabolic profile of lettuce leaves by high-field NMR spectra. Magnetic Resonance in Chemistry **43**: 625-638.

Soga, T. and Imaizumi, M. (2001). Capillary electrophoresis method for the analysis of inorganic anions, organic acids, amino acids, nucleotides, carbohydrates and other anionic compounds. Electrophoresis **22**(16): 3418-3425.

Soga, T., Ohashi, Y., Ueno, Y., Naraoka, H., Tomita, M. and Nishioka, T. (2003). Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. Journal of Proteome Research **2**(5): 488-494.

Son, H.S., Ki, M.K., Van Den Berg, F., Hwang, G.S., Park, W.M., Lee, C.H. and Hong, Y.S. (2008). $^{1}$H nuclear magnetic resonance-based metabolomic characterisation of wines by grape varieties and production areas Journal of Agricultural and Food Chemistry **56**(17): 8007-8016.

Stancher, B., Procida, G. and Calabrese, M. (1992). Characteristics of the most common mushrooms cultivated in Italy. IV. Lipds: Determination of the content of free and bound fatty acids. Industrie Alimentari **31**: 744-750.

Stenlund, H., Madsen, R., Vivi, A., Calderisi, M., Lundstedt, T., Tassini, M., Carmellini, M. and Trygg, J. (2009). Monitoring kidney-transplant patients using metabolomics and dynamic modeling. Chemometrics and Intelligent Laboratory Systems **98**(1): 45-50.

Steuer, R., Kurths, J., Fiehn, O. and Weckwerth, W. (2003). Observing and interpreting correlations in metabolomic networks. Bioinformatics **19**(8): 1019-1026.

Stewart, D., McDougall, G.J., Sungurtas, J., Verrall, S., Graham, J. and Martinussen, I. (2007). Metabolomic approach to identifying bioactive compounds in berries: Advances toward fruit nutritional enhancement. Molecular Nutrition & Food Research **51**(6): 645-651.

Stoop, J.M. and Mooibroek, H. (1998). Cloning and characterisation of NADP-mannitol dehydrogenase cDNA from the button mushroom, *Agaricus bisporus*, and its

expression in response to NaCl stress. Applied and Environmental Microbiology **64**: 4689-4696.

Sumner, L.W., Mendes, P. and Dixon, R.A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry **62**(6): 817-836.

Taghizadeh, M., Gowen, A. and O'Donnell, C.P. (2009). Prediction of white button mushroom (*Agaricus bisporus*) moisture content using hyperspectral imaging. Sensing and Instrumentation for Food Quality and Safety **3**(4): 219-226.

Taghizadeh, M., Gowen, A., Ward, P. and O' Donnell, C.P. (2010). Use of hyperspectral imaging for evaluation of the shelf-life of fresh white button mushrooms (*Agaricus bisporus*) stored in different packaging films. Innovative Food Science & Emerging Technologies **11**(3): 423-431.

Tarachiwin, L., Masako, O. and Fukusaki, E. (2008). Quality evaluation and prediction of *Citrullus lanatus* by $^{1}$H NMR-based metabolomics and multivariate analysis. Journal of Agricultural and Food Chemistry **56**(14): 5827-5835.

Teagasc (1994). Introduction to Mushroom Growing in Bags. Dublin, Teagasc.

Teagasc (2007). The Irish agricultue and food development authority. Teagasc Mushroom Newsletter **29**.

Teichmann, A., Dutta, P.C., Staffas, A. and Jägerstad, M. (2007). Sterol and vitamin $D_2$ concentrations in cultivated and wild grown mushrooms: Effects of UV irradiation. LWT - Food Science and Technology **40**: 815-822.

Tijskens, L.M.M. and Konopacki, P. (2003). Biological variance in agricultural products. Theoretical considerations. Acta Horticulturae **600**: 661-670.

Tolaas, A.G. (1915). A bacterial disease of cultivated mushroom. Phytopathology **5**: 51-54.

Tressl, R., Bahri, D. and Engel, K.H. (1982). Formation of eight carbon and ten carbon components in mushrooms (*Agaricus compestris*). Journal of Agricultural and Food Chemistry **30**: 89-93.

Trethewey, R.N. (2001). Gene discovery via metabolic profiling. Current Opinion in Biotechnology **12**(2): 135-138.

Trethewey, R.N., Krotzky, A.J. and Willmitzer, L. (1999). Metabolic profiling: a rosetta stone for genomics? Current Opinion in Plant Biology **2**(2): 83-85.

Trygg, J. and Lundstedt, T., Eds. (2007). Chemometrics in metabolomics - An introduction The Handbook of Metabonomics and metabolomics. Amsterdam, Elsevier Science B.V.

Tsai, S.Y., Huang, S.J., Lo, S.H., Wu, T.P., Lian, P.Y. and Mau, J.L. (2009). Flavour components and antioxidant properties of several cultivated mushrooms. Food Chemistry **113**: 578-584.

Tsai, S.Y., Wu, T.P., Huang, S.J. and Mau, J.L. (2007). Nonvolatile taste components of *Agaricus bisporus* harvested at different stages of maturity. Food Chemistry **103**(4): 1457-1464.

Tseng, Y.H. and Mau, J.L. (1999). Contents of sugars, free amino acids and free 5'-nucleotides in mushrooms, *Agaricus bisporus*, during post-harvest storage. Journal of the Science of Food and Agriculture **79**: 1519-1523.

Valentão, P., Lopes, G., Valente, M., Barbosa, P., Andrade, P.B., Silva, B.M., Baptista, P. and Seabra, R.M. (2005). Quantitation of nine organic acids in wild mushrooms. Journal of Agricultural and Food Chemistry **53**: 3626-3630.

Van Horen, L. (2008). Economic Developments in the Mushroom Industry - Grasping the Opportunities. The 6th International Conference on Mushroom Biology and Mushroom Products, Bonn, Germany.

Van Laere, A., Ed. (1995). Intermediary metabolism. The growing fungus. London, Chapman and Hall.

Vaz, J.A., Barros, L., Martins, A., Sa Morais, J., Vasconcelos, M.H. and Ferreira, I.C.F.R. (2011). Phenolic profile of seventeen Portuguese wild mushrooms. LWT - Food Science and Technology **44**(1): 343-346.

Vikram, A., Hamzehzarghani, H. and Kushalappa, A.C. (2005). Volatile metabolites from the headspace of onion bulbs inoculated with postharvest pathogens as a tool for disease discrimination. Canadian Journal of Plant Pathology **27**(2): 194-203.

Vikram, A., Lui, L.H., Hossain, A. and Kushalappa, A.C. (2006). Metabolic fingerprinting to discriminate diseases of stored carrots. Annals of Applied Biology **148**(1): 17-26.

Vikram, A., Prithiviraj, B., Hamzehzarghani, H. and Kushalappa, A.C. (2004). Volatile metabolite profiling to discriminate diseases of McIntosh apple inoculated with fungal pathogens. Journal of the Science of Food and Agriculture **84**(11): 1333-1340.

Vizhanyo, T. and Felfoldi, J. (2000). Enhancing colour differences in images of diseased mushrooms. Computers and Electronics in Agriculture **26**: 187-198.

Wagemaker, M., J.M., Eastwood, D.C., Welagen, J., van der Drift, C., Jetten, M.S.M., Burton, K., Van Griensven, L.J.L.D. and Op den Camp, H.J.M. (2007). The role of ornithine aminotransferase in fruiting body formation of the mushroom *Agaricus bisporus*. Mycological Research **111**(8): 909-918.

Wagemaker, M.J.M., Welboren, W., van der Drift, C., Jetten, M.S.M., Van Griensven, L.J.L.D. and Op den Camp, H.J.M. (2005). The ornithine cycle enzyme arginase from Agaricus bisporus and its role in urea accumulation in fruit bodies.

Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression **1681**(2-3): 107-115.

Wang, Z.Y., Soanes, D.M., Kershaw, M.J. and Talbot, N.J. (2007). Functional analysis of lipid metabolism in *Magnaporthe grisea* reveals a requirement for peroxisomal fatty acid β–oxidation during appressorium-mediated plant infection. The American Phytopathological Society **20**(5): 475-491.

Wangen, L.E. and Kowalski, B.R. (1989). A multiblock partial least squares algorithm for investigating complex chemical systems. Journal of Chemometrics **3**(1): 3-20.

Wannet, W.J., Aben, E.M., van der Drift, C., Van Griensven, L.J., Vogels, G.D. and Op den Camp, H.J. (1999). Trehalose phosphorylase activity and carbohydrate levels during axenic fruiting in three *Agaricus bisporus* strains. Current Microbiology **39**(4): 205-210.

Wannet, W.J.B., Van der Drift, C. and Op den Camp, H.J.M. (2000). Trehalose and mannitol metabolism in *Agaricus bisporus*. Science and Cultivation of Edible Fungi. L. J. L. D. Van Griensven. Horst, The Netherlands, Springer Netherlands.

Ward, J.L., Harris, C., Lewis, J. and Beale, M.H. (2003). Assessment of $^{1}$H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. Phytochemistry **62**(6): 949-957.

Watkins, S.M. and German, J.B. (2002). Metabolomics and biochemical profiling in drug discovery and development. Current Opinion in Molecular Therapeutics **4**: 224-228.

Weete, J.D., Ed. (1980). Lipid biochemistry of fungi and other organisms. New York, Plenum Press.

Weete, J.D., Furter, R., Hanseler, E. and Rast, D.M. (1985). Cellular and chitosomal lipids of *Agaricus bisporus* and *Mucor rouxii*. Canadian Journal of Microbiology **31**(12): 1120-1126.

Wells, J.M., Sapers, G.M., Fett, W.F., Butterfield, J.E., Jones, J.B., Bouzar, H. and Miller, F.C. (1996). Postharvest discoloration of the cultivated mushroom *Agaricus bisporus* caused by *Pseudomonas tolaasii*, *P*. 'reactans', and *P*. 'gingeri'. Postharvest Pathology and Mycotoxins **86**: 1098-1104.

Wells, T.K., Hammond, J.B.W. and Dickerson, A.G. (1987). Variations in activities of glycogen phosphorylase and trehalase during the periodic fruiting of the edible mushroom *Agaricus bisporus* (Lange) Imbach. New Phytologist **105**(2): 273-280.

Westerhuis, J.A., Kourti, T. and MacGregor, J.F. (1998). Analysis of multiblock and hierarchical PCA and PLS models. Journal of Chemometrics **12**: 301-321.

Will, M. and Guenther, D. (2007). Food Quality and Safety Standards. D. G. Fuer, Division 45 - Agriculture, Fisheries and Food.

Wilson, I.D., Nicholson, J.K., Castro-Perez, J., Granger, J.H., Johnson, K.A., Smith, B.W. and Plumb, R.S. (2005). High resolution 'ultra performance' liquid chromatography coupled to a-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. Journal of Proteome Research **4**(2): 591-598.

Wishart, D.S. (2008a). Applications of metabolomics in drug discovery and development. Drugs in R & D **95**(5): 307-322.

Wishart, D.S. (2008b). Metabolomics: applications to food science and nutrition research. Trends in Food Science & Technology **19**(9): 482-493.

Wishart, D.S., Tzur, D., Knox, C., Eisner, R., Guo, A.C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M.A.,

Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Amegbey, G., Block, D., Hau, D.D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G.E., MacInnis, G.D., Weljie, A.M., Dowlatabadi, R., Bamforth, F., Clive, D., Greiner, R., Li, L., Marrie, T., Sykes, B.D., Vogel, H.J. and Querengesser, L. (2007). HMDB: the Human Metabolome Database. Nucleic Acids Research **35**(suppl_1): D521-526.

Wold, S. (1995). Chemometrics; what do we mean with it, and what do we want from it? Chemometrics and Intelligent Laboratory Systems **30**(1): 109-115.

Wong, W.C., Fletcher, J.T., Unsworth, B.A. and Preece, T.F. (1982). A note on ginger blotch, a new bacterial disease of the cultivated mushroom, *Agaricus bisporus*. Journal of Applied Microbiology **52**(1): 43-48.

Workman, J., Ed. (2001). Methods and Interpretations in "Handbook of Organic Compounds". London, Academic Press.

Wu, H., Southam, A.D., Hines, A. and Viant, M.R. (2008). High-throughput tissue extraction protocol for NMR- and MS-based metabolomics. Analytical Biochemistry **273**: 204-212.

Wu, H.Q., Huang, X.L., Lin, X.S., Huang, F., Zhu, Z.X. and Ma, Y.-F. (2007). Gas chromatographic retention time rule and mass spectrometric fragmentation rule of fatty acids and Its application in food. Chinese Journal of Analytical Chemistry **35**(7): 998-1003.

Wurzenberger, M. and Grosch, W. (1984). The formation of 1-octen-3-ol from the 10-hydroperoxide isomer of linoelic acid by a hyperoxidase lyase in mushrooms (*Psalliota bispora*). Biochimica et Biophysica Acta (BBA) - Lipid and Lipid Metabolism **794**(1): 25-30.

Yamamoto, H., Yamaji, H., Abe, Y., Harada, K., Waluyo, D., Fukusaki, E., Kondo, A., Ohno, H. and Fukuda, H. (2009). Dimensionaltiy reduction for metabolome data using PCA, PLS, OPLS, and RFDA with differential penalties to latent variables. Chemometrics and Intelligent Laboratory Systems **98**: 136-142.

Yang, J.H., Lin, H.C. and Mau, J.L. (2001). Non-volatile taste components of several commercial mushrooms. Food Chemistry **72**(4): 465-471.

Yilmaz, N., Solmaz, M., Türkekul, I. and Elmastas, M. (2006). Fatty acid composition in some wild edible mushrooms growing in the middle Black Sea region of Turkey. Food Chemistry **99**(1): 168-174.

Yokokawa, H. and Mitsuhashi, T. (1981). The sterol composition of mushrooms. Phytochemistry **20**(6): 1349-1351.

Yu, J., Hu, S., Wang, J., Wong, G.K.-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. and Yang, H. (2002). A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. indica). Science **296**(5565): 79-92.

Yu, S.M. (1999). Cellular and genetic responses of plants to sugar starvation. Plant Physiology **121**: 687-693.

Yu, Y.X., Huang, N.B., Zhang, X.Y., Li, J.L., Yu, Z.Q., Han, S.Y., Lu, M., Van de Wiele, T., Wu, M.H., Sheng, G.Y. and Fu, J.M. (2011). Polybrominated diphenyl ethers in food and associated human daily intake assessment considering bioaccessibility measured by simulated gastrointestinal digestion. Chemosphere **83**(2): 152-160.

Zheng, L., Watson, D.G., Johnston, B.F., Clark, R.L., Edrada-Ebel, R. and Elseheri, W. (2009). A chemometric study of chromatograms of tea extracts by correlation optimization warping in conjunction with PCA, support vector machines and random forest data modeling. Analytica Chimica Acta **642**(1-2): 257-265.

Zhou, R., Su, S., Yan, L. and Li, Y. (2007). Effect of transport vibration levels on mechanical damage and physiological responses of Huanghua pears (Pyrus pyrifolia Nakai, cv. Huanghua). Postharvest Biology and Technology **46**(1): 20-28.

Zirrolli, J.A. and Murphy, R.C. (1993). Low-energy tandem mass spectrometry of the molecular ion derived from fatty acid methyl esters: A novel method for analysis of branched-chain fatty acids. Journal of the American Society for Mass Spectrometry **4**(3): 223-229

# PUBLICATIONS

## List of Publications

**Full-length papers:**

- O'Gorman, A., Barry-Ryan, C., Frias, J.M. (2011). Evaluation and identification of markers of damage in mushrooms (*Agaricus bisporus*) postharvest using a GC/MS metabolic approach. *Metabolomics*, online 11th March. DOI: 10.1007/s11306-011-0294-3.

- O'Gorman, A., Downey, G., Gowen, A.A., Barry-Ryan, C., Frias, J.M. (2010). Use of fourier transform infrared spectroscopy and chemometric data analysis to evaluate damage and age in mushrooms (*Agaricus bisporus*) grown in Ireland. *Journal of Agricultural and Food Chemistry*, 58 (13), 7770-7776.

- Gowen, A.A., O'Donnell, C.P., Taghizadeh, M., Gaston, E., O'Gorman, A., Cullen, P.J., Frias, J.M., Esquerre, C., Downey, G. (2008). Hyperspectral imaging for the investigation of quality deterioration in sliced mushrooms (Agaricus bisporus) during storage. *Sensing and Instrumentation for Food Quality and Safety*, 2 (3), 133-143.

**Conferences:**

- O'Gorman, A., Barry-Ryan, C., Frias, J.M. (2011). Evaluation and identification of markers of postharvest damage in mushrooms (*Agaricus bisporus*) using a GC/MS metabolomics approach. *International Congress on Engineering and Food (icef 11)*, May 22-26, Athens, Greece.

- O'Gorman, A., Downey, G., Barry-Ryan, C., Frias, J.M. (2010). Metabolic profiling of mushrooms (*Agaricus bisporus*) for the evaluation and identification of markers of postharvest damage. *Irish Chemometrics Society Meeting 2010*, November 5th, Dublin Institute of Technology (DIT), Kevin Street, Dublin, Ireland.

- O'Gorman, A., Frias, J.M., Barry-Ryan, C. (2009). Metabolic profiling of damaged and undamaged mushrooms (*Agaricus bisporus*) using GC/MS. *The 61st Irish Universities Chemistry Research Colloquium*, June 18-20, Dublin Institute of Technology, Dublin, Ireland.

- Gowen, A., Gaston, E., O'Gorman, A., Esquerre, C., Taghizadeh, M., Cullen, P.J., Barry-Ryan, C., O'Donnell, C., Downey, G., Frias, J. (2009). Post-harvest mushroom quality assessment by hyperspectral imaging and metabolomics. *All Ireland Mushroom Conference and Trade Show*, May 21, Monaghan, Ireland.

- O'Gorman, A., Frias, J., Barry-Ryan, C., Downey, G. (2008). Metabolic fingerprinting using fourier transform infrared spectroscopy & chemometrics to distinguish between damaged and undamaged mushrooms. *The 6th International Conference on Mushroom Biology and Mushroom Products*, September 30 – October 3, Bonn, Germany.

- Gaston, E., Frias, J.M., Cullen, P.J., O'Sullivan, C., O'Gorman, A., Barry-Ryan, C., Aguirre, L. (2008). Efficacy of a simple computer imaging system to differentiate and monitor the browning of mushrooms. *The 6th International Conference on*

*Mushroom Biology and Mushroom Products*, September 30 – October 3, Bonn, Germany.

- Gowen, A.A., O'Donnell, C., Taghizadeh, M., Gaston, E., O'Gorman, A., Cullen, P.J., Frias, J., Esquerre, C., Downey, G. (2008). Partial least squares regression for prediction of mushroom (*Agaricus bisporus*) slice quality during storage. *CAC2008, 11th Conference on Chemometrics in Analytical Chemistry*, June 30 – July 4, Montpellier, France.

- O'Gorman, A., Frias, J.M., Barry-Ryan, C., Downey, G. (2008). Metabolic fingerprinting using fourier transform infrared and chemometrics to distinguish between damaged and undamaged mushrooms (*Agaricus bisporus*). *The 60th Irish Universities Chemistry Research Colloquium*, June 11-15, University College Cork, Cork, Ireland.