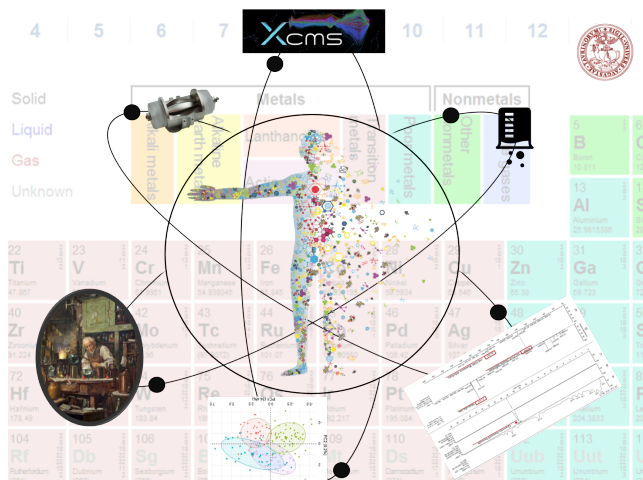




Università degli Studi di Torino

Doctoral School of Sciences and Innovative Technologies
PhD Program in Chemical and Materials Sciences XXXII Cycle
Thesis submitted in fulfilment of the requirements for the Degree of
Doctor of Philosophy

**Looking for new markers in biological and natural
samples. From targeted to untargeted
HPLC-HRMS based metabolomics**



Michael Zorzi

Supervisor: Prof. Claudio Medana

Candidate:

Michael Zorzi

Supervisor:

Prof. Claudio Medana

University of Turin

Department of Molecular Biotechnology and Health Sciences

Jury Members:

Prof. Giancarlo Aldini

University of Milan

Department of Pharmaceutical Sciences

Prof. Giuseppe Ciccarella

University of Salento

Department of Sciences, Biological and Environmental Technologies

Prof. Valter Maurino

University of Turin

Department of Chemistry

Head of the Doctoral School:

Prof. Alberto Rizzuti

University of Turin

Department of Humanistic Studies

PhD Programme Coordinator:

Prof. Bartolomeo Civalleri

University of Turin

Department of Chemistry

June 1st 2020

Contents

Author's declaration	V
Abstract	V
List of Acronyms	XI
List of Figures	XVII
List of Tables	XXIII
1. Omics sciences and metabolomics	1
1.1. Introduction	1
1.2. Metabolomics	3
1.2.1. History of metabolomics	3
1.2.2. Approaches in metabolomics	6
1.2.3. Applications for metabolomics	10
1.2.4. Analytical techniques in metabolomics analysis	11
2. Aims and outlines of PhD thesis	23
3. Instrumental apparatus	25
3.1. Sample preparation	26
3.2. Liquid chromatography	27
3.3. Mass Spectrometry	29
3.3.1. High Resolution Mass Spectrometry LTQ-Orbitrap TM	31
3.3.2. Tohoku University analytical platforms	33
3.3.3. SCIEX QTRAP [®] 5500	35

Contents

4. Validation methodology in metabolomics methods	39
4.1. Validation of targeted methods	39
4.2. Validation of untargeted methods	47
5. Targeted Analysis: the plastic pollution	55
5.1. Introduction	55
5.2. Aim of the work	59
5.3. Sample preparation and method settings	60
5.4. Results and discussion	65
5.5. Conclusion and future prospects	68
6. Targeted analysis: Uremic toxins molecules	75
6.1. Introduction	75
6.2. Aim of the work	77
6.3. Sample preparation and method settings	79
6.4. Results and discussion	81
6.5. Conclusion and future prospects	87
7. Targeted analysis: Quorum sensing molecules	95
7.1. Introduction	95
7.2. Aim of the work	97
7.3. Sample preparation and method settings	99
7.4. Results and discussion	115
7.5. Conclusion and future prospects	129
8. Targeted analysis: Lipids and hepatic steatosis in liver mice	135
8.1. Introduction	135
8.2. Aim of the work	137
8.3. Sample preparation and method settings	141
8.4. Results and discussion	145
8.5. Conclusion and future prospects	156
9. Semi-targeted analysis: Berries and phytochemicals	165
9.1. Introduction	165
9.2. Aim of the work	166
9.3. Sample preparation and method settings	167

9.4. Results and discussion	173
9.5. Conclusion and future prospects	180
10. Untargeted analysis: The Spacemice project	185
10.1. Introduction	185
10.2. Aim of the work	187
10.3. Sample preparation and method settings	188
10.4. Results and discussion	201
10.5. Conclusion and future prospects	214
11. Untargeted analysis: NRF2 activation in cancer cells	221
11.1. Introduction	221
11.2. Aim of the work	222
11.3. Sample preparation and method settings	223
11.4. Results and discussion	225
11.5. Conclusion and future prospects	234
12. Untargeted analysis: Detection of OQDS markers	239
12.1. Introduction	239
12.2. Aim of the work	240
12.3. Sample preparation, QC samples and method settings	241
12.3.1. Sampling, extraction protocols and QCs	241
12.3.2. HPLC-ESI-LTQ Orbitrap parameters	242
12.3.3. Data processing and analysis	245
12.4. Results and discussion	248
12.5. Conclusion and future prospects	253
13. Final remarks, conclusions and future research	261
A. Appendix A - Ionization Sources in Mass Spectrometry	265
B. Appendix B - Classifications of polyphenol and flavonoids	271
C. Appendix C - Classifications of lipids classes	273
D. Appendix D - Fatty acids and related masses list	275

Contents

E. Appendix E - LC-DAD polyphenol Characterization	279
-----------------------------------------------------------	------------


Author's declaration

I, Michael Zorzi, declare that, except where explicit references are made to the contribution of others by reference or acknowledgment, this dissertation is the result of my own work and that it has not been submitted, in whole or in part, in any previous application for a degree at the University of Turin or any other institution.

Furthermore, I declare that none of this work has been previously published elsewhere, with the exception of the articles I mentioned or the ones where I appear as author or co-author.

Io sottoscritto Michael Zorzi dichiaro che, fatta eccezione per quelle specifiche citazioni relative al contributo di altre persone, questa tesi è frutto del mio lavoro e non è stata presentata, né nella sua interezza né in parte, in nessun'altra domanda di laurea presso l'Università di Torino o presso qualsiasi altra Istituzione.

Dichiaro, inoltre, che nessuna parte di questo elaborato è stata pubblicata in precedenza altrove, ad eccezione degli articoli che ho citato o di quelli nei quali compaio come autore o coautore.

Signature: 

Date: June 1st 2020

Printed name: Michael Zorzi

Abstract

Nowadays, there is a specific approach able to achieve ambitious purposes: understanding metabolites' roles into the biological system; clarifying the origin of chemicals detected in natural products; discovering few or thousands of distinct species using only a single analytical method. This multi-faceted technique is called metabolomics.

Metabolomics is a complementary approach to genomics, transcriptomics, and proteomics. Perhaps, it is one of the most powerful tools for understanding the biological correlations with physiological or pathological conditions. In the last 15 years, this tool has become more and more attractive, although it is still a relatively new method compared to its other -omic siblings.

This tool could be applied in many fields: the food industry, the pharmaceutical area, the agricultural sector, etc. Furthermore, it's combined with different analytical techniques such as gas or liquid chromatography-mass spectrometry (GC-MS, LC-MS), or nuclear magnetic resonance (NMR). Data generated from MS or NMR metabolomics experiments are generally massive and accurate. Although statistical and bioinformatic platforms have made significant improvements to deal with this complexity, facilitating data exploration and interpretation.

Aim of the present work is the establishment of MS metabolomics approaches useful to investigate metabolites among biological and natural matrices. The analytical platform used consist of an HPLC-HRMS or UPLC-TQMS instrument and scheme for a full validation procedure are applied according to the different cases. LC-MS-based metabolomics strategies presented in my PhD thesis will be deeply discussed.

Overall, results obtained as in targeted as in untargeted approaches seems to be a reliable and efficient way to qualify and quantify metabolites in matrices of interest.

Key words: Mass Spectrometry, Analytical Chemistry, Untargeted Metabolomics, Targeted Metabolomics, High-Pressure Liquid Chromatography, HPLC-HRMS, omics sciences

Comprendere il ruolo dei metaboliti all'interno di un sistema biologico. Chiarire l'origine delle sostanze chimiche rilevate in prodotti naturali. Quantificare poche o migliaia di specie diverse per mezzo di un solo metodo analitico. Esiste una settore scientifico in grado di assolvere a tutti questi ambiziosi scopi: la metabolomica.

La metabolomica è un approccio complementare alla genomica, alla trascrittomica e alla proteomica ed è uno degli strumenti più potenti per comprendere le relazioni tra meccanismi biologici e condizioni fisio-patologiche. Negli ultimi 15 anni questo approccio è diventato sempre più attraente, sebbene comunque la metabolomica sia ancora una metodica di indagine relativamente nuova rispetto alle altre scienze -omiche. Inoltre, questo strumento può essere applicato nei campi più disparati (industria alimentare, area farmaceutica, industria agricola ecc.) grazie a differenti tecniche analitiche quali la gascromatografia o la cromatografia liquida accoppiate ad alla spettrometria di massa (GC-MS o LC-MS) o la Risonanza Magnetica Nucleare (NMR).

Sebbene i dati generati da esperimenti di metabolomica operati mediante le piattaforme appena citate siano grandi e accurati, le piattaforme statistiche e bioinformatiche hanno apportato miglioramenti significativi per la gestione della complessità del sistema, facilitando così l'esplorazione e l'interpretazione dei dati ottenuti. Lo scopo del presente lavoro di dottorato di ricerca è la creazione di approcci metabolici (mirati e globali) utilizzando la spettrometria di massa. Tali metodi, una volta messi a punto e validati, risulteranno quindi utili per studiare i metaboliti in diverse matrici, sia bi-

ologiche sia naturali.

La piattaforma analitica da me utilizzata si compone di diverse strumentazioni, tutte del tipo HPLC-HRMS o UPLC-TQMS. Inoltre, vengono applicati schemi per una procedura di validazione completa in base ai diversi casi in esame. Le strategie di metabolomica basate su LC-MS presentate nella mia tesi di dottorato saranno discusse approfonditamente, non prima però di aver introdotto ai lettori il mondo della metabolomica.

Nel complesso, i risultati ottenuti tramite approcci mirati e non mirati sono risultati efficienti per qualificare e quantificare i metaboliti predetti e nuovi all'interno delle matrici di interesse.

Parole chiave: Spettrometria di Massa, Chimica analitica, Metabolomica Untargeted, Metabolomica Targeted, Cromatografia liquida ad alta prestazione, HPLC-HRMS, scienze omiche

List of Acronyms

1. **ACN**: Acetonitrile
2. **ADME**: Absorption, Distribution, Metabolism, Excretion
3. **AFLD**: Alcoholic Fatty Liver Disease
4. **AGEs**: Advanced Glycation End Products
5. **AKI**: Acute Kidney Injury
6. **APCI**: Atmospheric Pressure Chemical Ionization
7. **API**: Atmospheric Pressure Ionization
8. **ASAP**: Atmospheric Solids Analysis Probe
9. **BPA**: Bisphenol A
10. **BMI**: Body Mass Index
11. **BPS**: Bisphenol S
12. **CE**: Collision Energy
13. **CE**: Cholesterol Ester
14. **CE-MS**: Capillary Electrophoresis Mass Spectrometry
15. **CKD**: Chronic Kidney Disease
16. **CXP**: Collision Cell Exit Potential
17. **DAD**: Diode Array Detector
18. **DAG**: Diacylglycerol
19. **DART**: Direct Analysis in Real Time
20. **DCM**: Dichloromethane

21. **DESI**: Desorption Electro-Spray Ionization
22. **DM**: Dry Matter
23. **DNA**: Deoxyribonucleic acid
24. **DPPH**: 2,2'-diphenyl-1-picrylhydrazyl radical scavenging activities
25. **DP**: Declustering Potential
26. **dQC**: Diluted Quality Control
27. **EC or ECN**: Equivalent Carbon Number
28. **EDCs**: Endocrine Disrupting Chemicals
29. **EDTA**: Ethylenediaminetetraacetic acid
30. **EI**: Electron impact Ionization technique
31. **ELISA**: Enzyme-linked immunosorbent assay
32. **EMC**: Enhanced Multiply Charged
33. **EP**: Entrance Potential
34. **EMS**: Enhanced Mass Spectrometry
35. **ESI**: ElectroSpray ionization
36. **EtOH**: Ethanol
37. **EUTox**: European Uremic Toxin group
38. **F.A.**: Fatty Acid
39. **FA**: Formic Acid
40. **FIA**: Flow injection analysis
41. **FID**: Flame Ionization Detector
42. **FL**: Fluorescence
43. **FRAP**: Ferric-Reducing Antioxidant Power
44. **FWHM**: Full width at half maximum
45. **GC**: Gas Chromatography
46. **GFR**: Glomerular Filtration Rate

47. **HESI**: Heated Electrospray Ionization
48. **HFAT**: High Fat Diet Fed Mouse
49. **HFRT**: High Fructose Diet Fed Mouse
50. **HILIC**: Hydrophilic interaction chromatography
51. **HPLC**: High-Performance Liquid Chromatography
52. **HMDB**: Human Metabolome DataBase
53. **HRMS**: High-Resolution Mass Spectrometry
54. **Isop**: Isopropanol or 2-propanol
55. **ISS**: International Space Station
56. **KEAP1**: Kelch ECH associating protein 1
57. **KEGG**: Kyoto Encyclopedia of Genes and Genomes
58. **KO**: Knock-Out
59. **JAXA**: Japan Aerospace Exploration Agency
60. **LB**: Lysogeny Broth
61. **JEM**: Japanese Experiment Module
62. **LDTD**: Laser Diode Thermal Desorption
63. **LLE**: Liquid-Liquid Extraction
64. **LLOQ**: Lower Limit of Quantification
65. **LOD**: Limit of Detection
66. **LOQ**: Limit of Quantification
67. **LTQ**: Linear Trap Quadrupole
68. **M9**: M9 Minimal medium bacteria culture
69. **Maf**: MusculoAponeurotic Fibrosarcoma
70. **MeOH**: Methanol
71. **METLIN**: Metabolite and Chemical Entity Database

72. **MRM**: Multiple Reaction Monitoring
73. **MS**: Mass Spectrometry
74. **MS/MS**: Tandem Mass Spectrometry
75. **MSⁿ**: Multi-stage Mass Spectrometry Analysis
76. **MODs**: Multiple Organ Dysfunction Syndrome
77. **NAFLD**: Non Alcoholic Fatty Liver Disease
78. **NASA**: National Aeronautics and Space Administration
79. **NASH**: Non Alcoholic Steato Hepatitis
80. **n.d.**: Not detected
81. **NL**: Neutral Loss
82. **NMR**: Nuclear Magnetic Resonance
83. **NRF2**: Nuclear factor erythroid 2-related factor 2
84. **N.S.**: Not Significant
85. **PC**: Principal Component
86. **PC**: Phosphatidylcholine
87. **PCPs**: Personal Care Products
88. **PDA**: Photodiode Array Detector
89. **PI**: Precursor Ion
90. **PI**: Phosphatidylinositol
91. **ppb**: Parts Per Billion
92. **ppm**: Parts Per Million
93. **ppt**: Parts Per Trillion
94. **PQS**: Pseudomas Quinolone Signal
95. **PS**: Phosphatidylserine
96. **QC**: Quality Control
97. **QIT**: Quadrupole Ion Trap
98. **QqQ**: Triple Quadrupole Mass Spectrometry
99. **RAGE**: Receptor for Advanced Glycation End-products

100. **RF**: Radio Frequency
101. **RNA**: Ribonucleic Acid
102. **ROS**: Reactive Oxygen Species
103. **RP**: Reverse phase Chromatography
104. **s.d.**: Standard Deviation
105. **SD**: Standard Diet
106. **SFC**: Supercritical Fluid Chromatography
107. **SIM**: Selected Ion Monitoring
108. **SPE**: Solid Phase Extraction
109. **SREBP1**: Sterol Regulatory Element-Binding Protein 1
110. **TAG**: Triacylglycerol
111. **TEAC**: Trolox Equivalent Antioxidant Capacity
112. **TIC**: Total Ion Current
113. **TMS**: Tetramethylsilane
114. **TPC**: Total Phenolic Content
115. **TSP**: Trimethyl-Silyl-Propionate
116. **TOF**: Time-of-Flight Mass Spectrometry
117. **ToMMO**: Tohoku Medical Megabank Organization
118. **TQMS**: Triple Quadrupole Mass Spectrometer
119. **UHPLC**: Ultra High Performance Liquid Chromatography
120. **UPLC**: Ultra High-Performance Liquid Chromatography
121. **UV-Vis**: Ultraviolet-Visible
122. **WT**: Wild Type
123. **XIC**: Extracting Ion Current

List of Figures

1.1.	The omics cascade	2
1.2.	Publications on PubMed concerning omics fields	7
3.1.	Schematics of a LTQ-Orbitrap TM	32
3.2.	Schematic representation of Synapt HDMS instrument	37
3.3.	Schematic representation of Thermo Scientific Q Exactive TM instrument	37
3.4.	Schematic representation of QTRAP [®] 5500 SCIEX instrument	37
4.1.	Flowchart of valuations in LC–MS target analysis	41
4.2.	Example of workflow for metabolite identification [17]	48
4.3.	Example of typical PCA scores plot for a data set	49
4.4.	Flowchart for global metabolomics LC–MS analysis	50
5.1.	Possible bisphenol analogues detectable in nature	56
5.2.	Chemical structures of Bisphenol A	56
5.3.	Chemical structures of Bisphenol S	58
5.4.	Chemical structures of glucuronide and sulfate conjugates of Bisphenol A. Figure has been taken from [12]	60
5.5.	TIC chromatogram of BPA/BPS standards solution	64
6.1.	Schematic example of few molecules belonging to uremic tox- ins' group	77
6.2.	Chemical structures for p-cresyl sulfate	78
6.3.	Chemical structures for indoxyl sulfate	78
6.4.	Result of the chromatographic run related to uremic toxins .	84
6.5.	p-cresyl sulfate levels over samples	88

6.6. Indoxyl sulfate levels over samples	88
6.7. HRMS XIC for different uremic toxins monitored	89
7.1. Cell density mediated by gene expression	97
7.2. N-acyl homoserine lactone generical structure	98
7.3. Quinolone signaling molecules generical structure	98
7.4. Quinolone signaling molecules standards	114
7.5. TIC PI AHLs bacteria culture	117
7.6. TIC PI AHLs bacteria culture WT; LB broth; PI method . .	119
7.7. TIC PI AHLs bacteria culture WT; LB broth; NL method . .	120
7.8. TIC PI AHLs bacteria culture WT; M9 broth; PI method . .	120
7.9. TIC PI AHLs bacteria culture WT; M9 broth; NL method . .	121
7.10. TIC PI AHLs bacteria culture RHII-; LB broth; PI method .	121
7.11. TIC PI AHLs bacteria culture RHII-; LB broth; NL method .	122
7.12. TIC PI AHLs bacteria culture RHII-; M9 broth; PI method .	122
7.13. TIC PI AHLs bacteria culture RHII-; M9 broth; NL method	123
7.14. TIC PI HQs bacteria culture WT; LB broth; PI method . . .	123
7.15. TIC PI HQs bacteria culture WT; M9 broth; PI method . . .	124
7.16. TIC PI HQs bacteria culture RHII-; LB broth; PI method . .	124
7.17. TIC PI HQs bacteria culture RHII-; M9 broth; PI method . .	125
7.18. TIC MRM AHLs sample	125
7.19. TIC NL AHLs sample	126
7.20. TIC PI AHLs sample	126
7.21. TIC MRM HQs sample	127
7.22. TIC MRM AHLs standard	127
7.23. TIC NL AHLs standard	128
7.24. TIC PI AHLs standard	128
7.25. TIC MRM HQs standard	129
8.1. Pathways of the analytes object of studies and relative cellular sections in which they take place	138
8.2. Oil red coloured histological sample of livers of mice subjected to three different diets. Figures are obtained by Professors Aragno's research group	139
8.3. Example of a TIC for sphingolipids standards mixture	147

8.4. Example of TAGs separation obtained through the HPLC- HRMS method for sphingolipids analysis	148
8.5. TAG MS spectra, fatty acids lost in the external positions . .	150
8.6. TAG MS spectra, fatty acids lost in the external positions . .	150
8.7. Histogram summarizing triglycerides mean values in SD, HFAT, and HFRT diets	152
8.8. Summary histogram for sphingolipid molecules among samples	156
8.10. Histogram summarizes of concentrations of sphingosine-1- phosphate, $\mu\text{g/L}$	159
9.1. Pictures of the cited berries. In top row, from left to right: white currant, blueberry, white gooseberry, and goji; in the bottom row, from left to right: blackberry, raspberry, black currant, red currant, and red gooseberry	167
9.2. Quercetin structure. It is a plant flavonol from the flavonoid group of polyphenols and it is one of the markers detected within all the samples under analysis	168
9.3. Scheme of the workflow applied for untargeted metabolomics analysis of lyophilized berries samples	170
9.4. Chromatograms and MS spectra of lyophilized berries sam- ples for quercetin-rhamnetil-hexoside	172
9.5. Histogram of the 70 different analytes detected in the eleven berry fruits investigated	179
10.1. Schematic report of Biocrates kits' application in scientific publications in different fields of research in the last 3 years [2]	189
10.2. Workflow provided by Biocrates Life Sciences AG	190
10.3. The two chromatographic gradients used for LC analyses (above in the figure) and FIA analyses (below in the figure) [2]	192
10.4. A detail of the sample extraction procedure: sampling of a portion of the liver's mouse	198
10.5. Cryostat: detail about the microtome portion	200
10.6. Cryostat tissue sections by placed onto a glass ready for DESI-MS	200

10.7. DESI – SYNAPT G2-Si High Definition Mass Spectrometry by Waters	200
10.8. DESI ionization source used for SpaceMice ambient mass spectrometry	200
10.9. Score plot of FIA Biocrates data	202
10.10 Loading plot of FIA Biocrates data	203
10.11 Score plot of FIA Biocrates data without sugars' group . . .	203
10.12 Histogram summarizing the variables' number applied to dis- criminate the different mice groups under examination	205
10.13 XIC and MS spectra of different analytes detected with the UHPLC-HRMS method	206
10.14 Score plot WT vs Nrf2KO in real samples obtained by UHPLC- HRMS data - reverse phase chromatography in positive ion mode. In blue are marked the WT samples, in black the NRF2KO samples	207
10.15 Score plot WT vs Nrf2KO in real samples obtained by UHPLC- HRMS data - direct phase chromatography in negative ion mode. In blue are marked the WT samples, in black the NRF2KO samples	207
10.16 S-plot WT vs Nrf2KO of real samples, reverse phase chro- matography in positive ion mode	208
10.17 S-plot WT vs Nrf2KO of real samples, direct phase chro- matography in negative ion mode	208
10.18 TIC and XIC chromatograms obtained through the analytical lipidomics method built up	210
10.19 Score plots and related S-plots generated through the results obtained in negative ion mode	211
10.20 DESI image of a kidney sample - positive polarity	215
10.21 DESI image of a kidney sample - positive polarity	215
10.22 DESI image of a kidney sample - positive polarity	215
11.1. The Keap1–Nrf2 system	222
11.2. Figures A - C extrapolated from the article [25]	227
12.1. XICs of internal standard - Positive Ion Mode	243

12.2. XICs of internal standard - Negative Ion Mode	243
12.3. Internal standard molecules in Xylella analyses	244
12.4. TICs and PDA from a real OQDS sample	246
12.5. PCA plots for negative polarity experiments	250
12.6. PCA plots for negative polarity experiments - zoom	251
12.7. PCA plots for positive polarity experiments	251
12.8. Box and whisker plots of MS/MS confirmed molecules	253
12.9. Box and whisker plots of MS/MS confirmed molecules	254
B.1. Systematic polyphenolic molecules classification	271
B.2. Main sub-classes structures of flavonoids.	272
C.1. General structure of glycerol-based lipids	274
C.2. General structure of sphingoid-based lipids	274
E.1. UV spectra of common polyphenols detected in plants	280

List of Tables

1.1.	Growth of metabolomics field over the years	7
1.2.	Comparison between most used analytical metabolomics platforms [41]	17
3.1.	Different possible chromatographic methods available [7] . . .	28
3.2.	Mass analyser parameters used in metabolomics studies . . .	30
5.1.	Chromatographic conditions for BPA and BPS, urine samples	63
5.2.	Chromatographic conditions for BPA and BPS, breast milk samples	63
5.3.	MS parameters for BPA and BPS urine and breast milk samples	65
5.4.	General characteristics of mothers enrolled in the project . .	66
5.5.	General characteristics of control women samples enrolled in the project	66
5.6.	General characteristics of babies enrolled in the project . . .	66
5.7.	Urinary free and conjugate BPA/BPS values detected in the groups analysed	67
6.1.	Instrumental equipment used in the HPLC-HRMS system . .	81
6.2.	Chromatographic and MS conditions - LTQ Orbitrap TM Thermo Scientific	82
6.3.	Instrumental equipment used in the UPLC-TQ MS platform .	82
6.4.	Chromatographic conditions - SCIEX QTRAP [®] 5500	82
6.5.	MS parameters - SCIEX QTRAP [®] 5500	83
6.6.	Schematic report of parameters obtained during validation procedures for uremic toxins analyses	83
6.7.	Concentration of p-cresyl sulfate during therapy, HPLC-HRMS	87

6.8. Concentration of indoxyl sulfate during therapy, HPLC-HRMS	87
6.9. Concentration of p-cresyl sulfate during therapy, UPLC-TQ MS	87
6.10. Concentration of indoxyl sulfate during therapy, UPLC-TQ MS	90
7.1. List of validated parameters for N-butanoyl-DL-homoserine lactone - 1 of 6	103
7.2. List of validated selectivity parameters for N-butanoyl-DL- homoserine lactone - 2 of 6	104
7.3. List of validated selectivity parameters for N-butanoyl-DL- homoserine lactone - 3 of 6	104
7.4. List of validated selectivity parameters for N-butanoyl-DL- homoserine lactone - 4 of 6	105
7.5. List of validated selectivity parameters for N-butanoyl-DL- homoserine lactone - 5 of 6	105
7.6. List of validated selectivity parameters for N-butanoyl-DL- homoserine lactone - 6 of 6	106
7.7. List of validated selectivity parameters for N-(3-oxododecanoyl)- L-homoserine lactone - 1 of 6	106
7.8. List of validated selectivity parameters for N-(3-oxododecanoyl)- L-homoserine lactone - 2 of 6	106
7.9. List of validated selectivity parameters for N-(3-oxododecanoyl)- L-homoserine lactone - 3 of 6	107
7.10. List of validated selectivity parameters for N-(3-oxododecanoyl)- L-homoserine lactone - 4 of 6	107
7.11. List of validated selectivity parameters for N-(3-oxododecanoyl)- L-homoserine lactone - 5 of 6	108
7.12. List of validated selectivity parameters for N-butanoyl-DL- homoserine lactone - 6 of 6	108
7.13. List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline - 1 of 6	108
7.14. List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline - 2 of 6	109

7.15. List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline - 3 of 6	109
7.16. List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline - 4 of 6	110
7.17. List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline - 5 of 6	110
7.18. List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline - 6 of 6	110
7.19. Recovery values in quorum sensing analyses	113
7.20. General instrument parameters used as in all UPLC-TQMS determinations	114
7.21. Specific AHLs NL and PI scan parameters used	114
7.22. Specific AHLs and HQs MRM scan parameters used	115
8.1. Physiological parameters of mice fed with different diets	141
8.2. Instrumental equipment used in the HPLC-HRMS TAG anal- ysis	143
8.3. Chromatographic and MS conditions - LTQ Orbitrap TM Thermo Scientific in TAG analysis	143
8.4. Instrumental equipment applied for UPLC-TQMS lipids anal- yses	145
8.5. HPLC and MS parameters for SCIEX QTRAP [®] 5500 lipids analyses	146
8.6. Sciex QTRAP [®] 5500 MRM parameters applied for sphin- golipids analyses	146
8.7. Triacylglycerols detected in sphingolipids' samples	149
8.8. Percentage distribution values of triglycerides in SD, HFAT, HFRT	151
8.9. Schematic report of parameters obtained during validation procedures for sphingolipids in UPLC-TMQS analyses	154
8.10. Concentration expressed in $\mu\text{g}/\text{L}$ of sphingolipids of interest among the three diets	155
8.11. LOD, LOQ, and LLOQ and validation parameters in sphin- golipids analyses	159

8.12. Concentration of sphingolipids introduced in UPLC-TMQS method in real samples. s.d.: standard deviation	160
9.1. Characteristics of berries under analysis	168
9.2. ProteoWizard parameters applied for berries .raw files con- version	170
9.3. Polyphenolic compounds identification by HPLC-PDA-HRMS in positive polarity	176
10.1. Metabolite classes detectable with AbsoluteIDQ [®] p400 HR Kit	190
10.2. MS Settings of Q-Exactive [™] Thermo instrument used dur- ing LC and FIA. All the parameters are imposed by Biocrates procedure	192
10.3. Performance characteristics of the Q-Exactive [™] Thermo in- strument	194
10.4. Schematic summary of the different molecules detected using the different instrumental approaches	209
10.5. Summary of the different analytes detected by untargeted lipidomics approach - Positive polarity	212
10.6. Summary of the different analytes detected by untargeted lipidomics approach - Negative polarity	213
11.1. Features detected in the four different assays: HILIC-negative, HILIC-positive, C18-negative, C18-positive	228
12.1. ProteoWizard parameters applied for Xylella .raw files con- version	246
12.2. Features identification and confirmation methods applied to achieve this task. CID: Compound Identification number; METLIN: Metabolite and Chemical Entity Database; MSI: Metabolomics Standard Initiative	254
A.1. Summary of ionization sources in MS	265
A.2. Comparison between DART and DESI sources	269
D.1. List of fatty acids and related masses	276

Omics sciences and metabolomics

1

1.1. Introduction

What is metabolomics? Before giving a proper answer to this question, we must first reply to another type of query: What is an omics science?

Depending on whom you ask, you will undoubtedly receive different answers. The evolution of the omics sciences has been so rapid that even the scientists involved in this field don't have a fixed definition. In general, the omics world is a big field of study that integrates multiple biological disciplines to focus on measurements of molecules with biological activities [1]. The main areas belonging to omics science are genomics, transcriptomics, proteomics, and metabolomics. The omics field is a boundless molecular puzzle, and only once the pieces are linked together, a larger and brighter and defined picture of the human body can predict the outcome. As reported in [38], the most crucial aspect is that “by studying the brick no one can learn anything about the design of the building nor the architect.”

So, before focusing more deeply into the metabolomics field, a brief introduction of the omics sciences cited before is necessary. The relationship between the different omics fields (genome, proteome, and metabolome) is summarized in figure 1.1.

Genomics

The ambitious task of genomics science is the discovery in a specific organism of all the possible sequences in the genome, where “genome” can be defined as the complete genes set inside a single cell.

The sequencing of the genes, however, is just the first step of a journey in genomics science. Once this task is achieved, the genomic sequence obtained will be useful in many different ways: to uncover the functions of genes (functional genomics); to compare the genes between different organisms

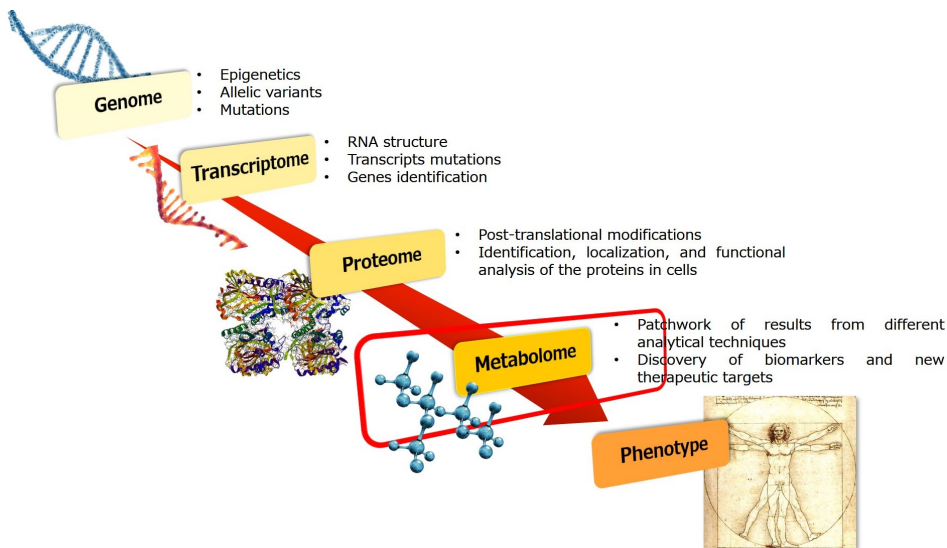


Figure 1.1.: *The omics cascade. Relationship between different omics fields, from genome to phenotype*

(comparative genomics); to generate the three-dimensional genes structures too via informatics tools (structural genomics).

Therefore, genomics is the first point to look at to understand omics sciences. The genotype, the set of genes in DNA of an organism, is responsible for the phenotype, or rather the composite of an organism's observable characteristics. However, not only the genotype but also the environment can influence the phenotype characteristics.

Transcriptomics

The complete set of RNA transcripts that are produced by the genome, under specific circumstances or within a particular cell, is called transcriptome. Therefore, transcriptomics using high-throughput methods, such as microarray analysis, pretends to study and clarify the transcriptome of a specific genome. Once the transcriptome is analysed, the comparison of transcriptomes allows the identification of genes that are differentially expressed in distinct cell populations or response to different treatments.

Proteomics

The tasks for which proteins are responsible in a cell are endless. The set of all proteins in a single cell is called proteome, and so proteomics can be defined as the science involved in the study of protein structures and functions. The proteome is not a fixed entity, but it changes continuously in response to environmental factors. So, the principal aim of proteomics studies is to discover how the structure and function of proteins allow them to carry out some of the essential processes of life. Furthermore, other proteomics tasks exist. The technique can also be used to develop a protein-network map of an organism, or it can also be applied to understand and clarify the difference between wild type and knock-out modified organisms.

1.2. Metabolomics

So, what is metabolomics? Metabolomics is just a big tile of this omics puzzle, and it can define as “the scientific analysis and study of the alterations of the metabolomes, where the metabolome is the full suite of metabolites present and detectable within an organism, through techniques applied to the characterization of fluids and tissue samples” [33]. Obviously, in scientific literature there is a plenty of definitions related to this term. As a matter of fact, it can be also defined as: “The large-scale study of small molecules within organisms [...]”[23], or “ From a biological functionality point of view, metabolomics is that bioanalytical tool that allows obtaining a picture of the metabolites of an organism in the course of a biological process [...]”[44]. So, definitions aside, to understand actually what metabolomics is, we must first analyse the historical context in which it was born.

1.2.1. History of metabolomics

From a historical point of view, to understand the meaning of this word and from which step in time that name started to be used, it becomes necessary to explore what is happened before 1949 (PubMed foundation’s year) looking not only to metabolomics exactly name but also to its related

acronyms like for example, *metabolic profiling* or *metabonomics*. Concerning human body aspects, what is essential is to look at which point in time the presence or the absence of metabolites in human fluids were significant from a medical point of view to diagnosing specific pathologies. In ancient times (2000–1500 B.C.) in China, diabetes was detected thanks to ants. Traditional Chinese doctors use ants to detect if the level of glucose in urine was high or not: if the concentration was high, ants were attracted by urine; otherwise, they avoided it [27, 38].

Around 200 A.C. in Greece, doctors to predict diseases investigating for the first-time *humours* (body fluids name at the time), which shows primary steps towards modern metabolomics. Later in 131 A.D., Galen developed a system of pathology combining humeral theories of Hippocrates with the Pythagorean philosophy. This new theory was unchallenged and remained the most used until the 17th-century [17].

Coming to nowadays, the idea that each individual might have a specific metabolic profile that could be reflected in its biological fluids was introduced by Roger Williams in the late 1940s [12]. Professor Williams demonstrated how, using paper chromatography, it was possible to reveal in human fluids like urine and saliva specific metabolic patterns associated with diseases like schizophrenia. However, the technique proposed by Professor Williams became feasible to quantitatively metabolites involved in metabolic profiles only through technological advancements that occurred in the 1960s and 1970s.

The term *metabolic profile* was introduced by Horning, et al. in 1971 [18] after they demonstrated that with a gas chromatography-mass spectrometry instrumentation was possible to conduct multicomponent analyses of compounds like steroids or drugs metabolites present in human urine and tissue extracts [31, 38]. Through the 1970s, different scientific groups were involved in this field (besides Horning group also Linus Pauling and Arthur B. Robinson were involved too), and they led the development of GC-MS methods able to detect metabolites in urine never detected before [16].

Concurrently to MS techniques, NMR spectroscopy, which was discovered in the 1940s by Nobel Prize for Physics Isidor Isaac Rabi, was making giant strides. In 1974, scientists demonstrated the utility of NMR techniques to

detect biomarkers in non-treated biological samples [19]. Ten years later, the efforts to utilize NMR for metabolomics mainly been driven by the laboratory of Jeremy K. Nicholson at Birkbeck College, University of London, and then at Imperial College London. As a matter of fact, in 1984, Nicholson showed ^1H NMR spectroscopy could be involved in the diagnose of the diabetes mellitus illness [25].

After 2000s, I.T. resources have begun to dominate. In 2005, for example, in the Scripps Research Institute, the first metabolomics web database, METLIN, [36] for characterizing human metabolites was developed. In the early days of its life, the METLIN database contained 10000 metabolites and some tandem mass spectral data too. After ten years of research in September 2015, METLIN counted in its database over 240000 metabolites as well as an incredible database of tandem mass spectrometry data in the metabolomics field for bacteria, animals, and plant organisms.

Two years later, METLIN's creation, another relevant web database, the Human Metabolome Project (HMDB), is created by Dr. David Wishart of the University of Alberta, Canada. The project completed the first draft of the human metabolome, and at first, it consists of a database of approximately 2500 metabolites, 1200 drugs, and 3500 food components [40]. Nowadays, The Human Metabolome Database (HMDB) has 279972 predicted MS/MS spectra for 98601 compounds metabolite uploaded in the library [42] and it aims to recognize all the metabolites present in humans in the next future [40].

Nowadays, metabolomics techniques have made impressive advantages, and new several metabolomics databases are available, each of them serving a different purpose. Their common goal is to organize metabolites in some order so that it becomes easy for researchers to spot and analyse the data. In addition to the aforementioned METLIN and HMDB, the following databases are free to use:

- MassBank [34]: a spectral database that has over 39000 entries.
- Lipid metabolites and pathways strategy (LIPID MAPS) [13]: it is the largest repository for lipid molecular structures.
- Madison metabolomics consortium database (MMCD) [28]: an on-line

database who has over 20000 data, a resource for mass spectrometry and nuclear magnetic resonance-based metabolomics research.

- Kyoto Encyclopedia of Genes and Genomes (KEGG) [24]: a Japanese database containing an endless harvest metabolic pathway.

An increasing focus in metabolomics research is now evident in academia, industry, and government, with more than 500 papers a year being published on this subject [37]. Therefore, a concrete and simple way to monitor how the metabolomics field is growing up is checking how many times during a year the name “metabolomics” is contained within the title of a scientific article published into international journals.

Using PubMed [6] and Google Scholar [15] as databases for this research, what founded is reported in table 1.1. As it is evident, from ten years now, metabolomics is a science in a continually growing. Furthermore, figure 1.2 shows the rate of publications in the genomics, proteomics and metabolomics disciplines. As it is clear, metabolomics is a relatively new application field constantly growing, but genomics and proteomics disciplines still arouse much interest. However, on closer inspection in this period, something has changed: if the rate of publication of metabolomics articles is constantly increasing, genomics and proteomics show a different behaviour. Furthermore, in time period considered the growth rate in publication of genomics, proteomics, and metabolomics articles are 8.24%, 12.31%, and 29.72% respectively¹. Therefore, this bodes well for future research in this area.

1.2.2. Approaches in metabolomics

Analytical chemistry strategies applied in metabolomics could be divided into three distinct approaches: untargeted, semi-targeted, and targeted techniques.

Generally, scientists resort to untargeted and semi-targeted methods for hypothesis-generating studies, while to the manners which targeted ap-

¹The growth rate value was calculated using the following equation: Growth rate,% = $100 \times [(\frac{pr}{ps})^{\frac{1}{n}} - 1]$, where n is the number of years considered, pr (present) is value of the most recent year taken into account and ps (past) is value of the first year of range considered.

Table 1.1.: *Growth of metabolomics field over the years. The numbers were obtained from PubMed and Google Scholar search using keywords “mass spectrometry” and “metabolomics”*

Year	Articles from Google Scholar [15]	Articles from PubMed [6]
2018	11300	2636
2017	17600	4114
2016	17300	3789
2015	16700	3128
2014 - 2011	52800 (13800 each year)	7353 (1843 each year)
2010 - 2006	28600 (5700 each year)	2437 (609 each year)
2005 - 2001	6950 (1400 each year)	265 (66 each year)
Before 2000	500 globally	—

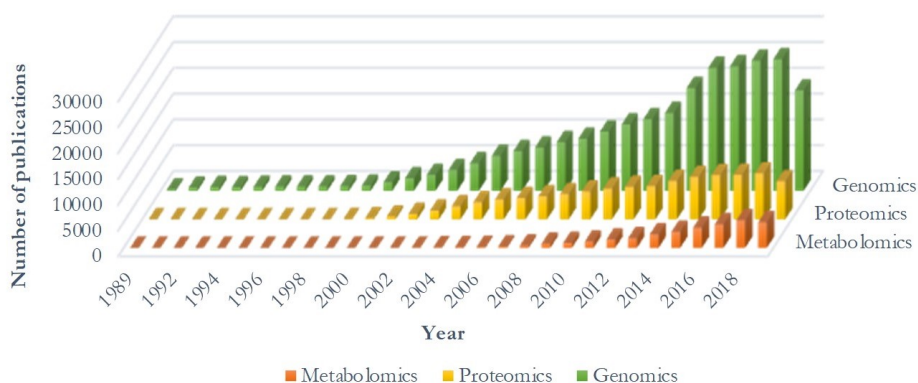


Figure 1.2.: *The number of publications on different omics fields (genomics, proteomics and metabolomics) on PubMed between 1989 and 2019*

proach are usually applied are related to detect single analytes and to prove novel discoveries of hypothesis-generating studies cited before [5]. Three most important differences between these approaches are:

- Amount of sample available for analysis;
- Quantity of metabolites detectable by a single analysis;
- Lower level of quantification of these metabolites in real samples.

The discriminating factor that leads to choosing a method rather than another depends on the type of analytical instruments available and the set objectives.

On the one hand, there are the non-targeted (or untargeted) analyses. The most appropriate way to detect changes in specific metabolite concentrations is the untargeted approaches; this technique aims to maximize the possible metabolites detectable during a unique analysis and, therefore, provide the opportunity to observe new biomolecules or unexpected changes. Typically, from hundreds to thousands of metabolites can be measured in a single analysis. However, it is impossible to think that just one single analytical technique (such as NMR or MS) can detect all the metabolites of the sample under analysis. The only desirable solution is, therefore, the combination of at least two different analytical approaches to maximize the number of analytes detectable and to increase the coverage of the metabolome.

In untargeted studies, sample preparation is an essential and time-consuming step. This moment involves the theoretically 100% extraction of all the metabolites from the sample (biological, inorganic, foods, etc.). Then, the extracted sample is analysed using an appropriate and established analytical method: mass spectrometry combined with liquid chromatography or NMR technique are available and valid tools for this task. For example, in a LC-MS analysis, the area under each metabolite's peak obtained via chromatographic separation can be easily used as the parameter in the statistical analysis to define the differences in concentration between different biological samples analysed. This procedure is called relative quantification and, if a full quantification is necessary, a calibration curve built-up is required. Subsequently, data analysis is required to uncover the biological

importance of each species detected, and the interpretation of these data need to be performed at the end of the experimental pathway. The main important aspect to remark is that in untargeted metabolomics, as suggested by this name, the chemical identity of each metabolite is not known at the begin of the analysis. Currently, the identification of metabolites is one of the most limitations aspects of untargeted approaches. Anyhow, fortunately for data scientists involved in these fields, some useful new strategies and software are available for helping them in the identification compounds and in the statistical analysis aspect too. These strategies, called “in silico approaches” help to unravel the complexity of the metabolome and mainly shed light on unknown metabolites [7]. In MS some crucial tools are for example ProgenesisQI (from Waters), MS-FINDER, CFM-ID, and MetFrag.

On the other hand, there are targeted studies. In these analyses, a relatively small and specific number of metabolites, typically up to twenty metabolites, is in the centre of attention. These metabolites are already chemically characterized and biochemically known before performing any data acquisition. Knowing what scientists are searching for, they can create specific analytical methods to obtain higher selectivity and sensitivity than global approaches. In this case, the quantification of the analytes is easily performed using internal standards and pure chemical standards to build up a calibration curve for each molecule of interest in the pathway studied. Therefore, a targeted study cannot be valid unless a chemical standard of the metabolite of interest is commercially available. The sample preparation in targeted methods needs always to be optimized at the maximum level to retain the metabolites of interest, to maximize their concentration in the extracted sample, and to remove all the possible matrix interferences due to biological species and analytical artefacts that are not carried through to the analysis. The last aspect to remark is that in this kind of approach, analysis of data and interpretation of biological significance is much simpler than non-targeted studies.

There is also a third way to operate a metabolomics analysis: semi-targeted approach. The semi-targeted methods work in a middle way, between untargeted and targeted strategies. These techniques aim to quantify hundreds of metabolites (as in an untargeted analysis) whose identity is

known before data acquisition (as in a targeted approach). Method typically applies one calibration curve for each of the metabolites or, at least, one curve for analytes with a similar chemical structure. So, in a semi-targeted approach, monitoring a hundred metabolites is quite common, and it requires tens of calibration curves for a full quantification.

1.2.3. Applications for metabolomics

If metabolites can be defined as “a spoken language, broadcasting signals from the genetic architecture and the environment” [20], metabolomics can provide a “functional readout of the physiological state” [14] in an organism.

The size of the metabolome is still disputed [35]. Alongside the variety of chemical classes, the physical properties of metabolites, and the variation in concentrations of the analytes depending on the matrix, it is necessary nowadays to employ a range of analytical techniques in metabolomics research to obtain data.

Nowadays, the primary tools to analyse many metabolites in a single analysis are NMR and mass spectrometry coupled to different chromatographic separation techniques. Unfortunately, due to the vast diversity between the analytes in terms of chemical structure and abundance, a unique technology to analyse the entire metabolome is still not available.

One big challenge in metabolomics is to extract and interpret all the possible useful information from a biological sample and discharge the vast amount of usefulness data produced by high-throughput analysers. The application of sophisticated statistical and multi-variant data analysis tools, including cluster analysis, is the key to complete this hard task.

To achieve the aims, four conceptual approaches in metabolomics can be used: targeted analysis, metabolite foot-printing, metabolic fingerprinting, and metabonomics. As already mentioned in 1.2.2, targeted analysis allows the determination and quantification of a small number of known metabolites (targets) using a specific analytical technique of best performance for the compounds of interest. Some examples of targeted metabolomics involve biological analysis: as a matter of fact, from urine or blood could be extracted different information which could be related to specific health

conditions or syndromes [10, 11, 41].

Moving to metabolite foot-printing, also described as the exo-metabolome, it is the analysis of the uptake of extracellular metabolites and the secretion of intracellular metabolites. Typically, the samples can be taken quickly via non-invasive approaches, and in this approach, the metabolites of interest are monitored over a fixed period, and, in the end, have compared each other. So, these experiments can differentiate, for example, physiological states of wild-type yeast and between yeast mutants [2] or another dome not strictly connected to metabolomics world [29].

Moreover, metabolite fingerprinting, also defined as intracellular metabolome, provides a snapshot of the global metabolism. Here a metabolic “signature” or mass profile of the sample of interest is generated and then compared in a large sample population to screen for differences between the samples. This approach, aims for an untargeted analysis of the metabolome, can be applied in different fields such as plants [21], microbes [8], the ways in which an organism interacts with its environment [4] or biological samples [22, 32]. To obtain a wide coverage of the metabolome, different analytical methodologies, such as LC-MS/MS, GC-MS, and or NMR or different extraction methods can be used.

Last but not least, metabonomics, subset of metabolomics, aims to show quantitative changes appearing in the metabolome organism due to a response to physiological stimuli or genetic perturbations [3]. Furthermore, the metabonomic approach put particular emphasis on the elucidation of differences between groups due to genetic modification, disease, and environmental stress along a precise period [39]. When, for example, the condition is not linked to a mutation, metabolites can be used as markers of illness or chronic exposure to toxins or drugs. In particular, as far as drugs are concerned, metabonomics approach can be used as a tool in the investigation of the absorption, distribution, metabolism, and excretion (ADME) of drugs [26, 30].

1.2.4. Analytical techniques in metabolomics analysis

To study the biological role of small-molecule metabolites within biological samples, not only a single instrument can be used to detect all metabo-

lites present in the metabolome. The low concentration of metabolites of interest and the different biological classes [43] such as lipids, peptides, or carbohydrates to which they belong makes global metabolomics studies a difficult task to be carried out. Owing to the complexity of the metabolome and the diverse properties of metabolites, each instrument has advantages and limitations, and the appeal to different analytical methods or multiple instruments approach is required to provide the most considerable number of metabolites information. Two complementary methods dominate metabolomics: nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). I will briefly summarize here the panorama of some analytical techniques applied in metabolomics studies, excluding the LC-MS methods, which I will treat more in detail in the following sections

Gas chromatography-mass spectrometry

The GC-MS technique is an analytical approach applied to the metabolites which have a low boiling point, and which will be present in the gas phase in the temperature range 50-350°C. The metabolites under analysis can possess this low pointing point by their biological nature or, otherwise, this property can be modified through a chemical reaction called derivatization. This process increases the number of metabolites detectable in a biological sample and is commonly applied in metabolomics studies.

The separation mechanism used in GC-MS consists of differential absorption and desorption on a stationary phase of vaporized metabolites transported by a mobile gas phase. In particular, the stationary phase is coated on the inside surface of a hollow silica capillary, and the mobile phase is a gas, for example, helium. The chromatographic columns used in GC techniques are silica capillaries of 10-60 m in length and an internal diameter of 100-500 μm . The stationary phase is constituted of siloxane or a similar chemical composition with varying percentages of chemical moieties. Operatively, samples are vaporized and introduced on to the GC column by an inert gas (the mobile phase) flowing continuously. By increasing the oven temperature (generally, the temperature of the gas chromatographic ovens typically ranges from 5°C to 400°C) the different analytes are eluted

from the column. Spectra of compounds eluted by the column are collected by the mass spectrometer, which identifies and quantifies all the chemicals according to their mass-to-charge ratio m/z . The formation of the ions is commonly obtained by electron impact ionization technique (EI). The ionization via EI is a high energy process where electrons hit molecules in a vacuum environment and release an electron from the molecule to form a positively charged ion. EI provides significant fragmentation of the molecular ion and produces a reproducible fragmentation pattern that can be employed for the identification of metabolites by searching mass spectral libraries. This is a crucial aspect that discriminates LC metabolomics analyses and GC metabolomics analyses: due to the higher reproducibility of the mass spectra in GC-MS analysis, the libraries for GC have higher accuracy than the LC libraries.

One of the most important factors to consider in GC-MS analyses is the resolution parameter, which could be expressed by the equation 1.1:

$$R_s = \left[\frac{\sqrt{N}}{4} \right] \times \left[\frac{\alpha - 1}{\alpha} \right] \times \left[\frac{k}{k + 1} \right] \quad (1.1)$$

The equation 1.1 explicates the effect of the chromatographic variables on the resolution and hence is an aid to column choice and optimization. Focusing more deeply on the equation, the different terms are: R_s = separation or resolution factor, N = theoretical plate number, α = selectivity factor, k = retention factor and generally tells us that:

- Sample components have to be retained by the stationary phase of the column to be separated ($k > 0$).
- k values must be different from each other so that the selectivity α is greater than 1. The greater the difference, the greater is the separation.
- If the N value is high enough, the peaks at the baseline are sufficiently narrow. When N is not sufficiently high, peaks overlap, and the separation is incomplete.

Capillary electrophoresis-mass spectrometry

Another analytical technique that combines a separation procedure with mass spectrometry is the capillary electrophoresis-mass spectrometry (CE-MS). The principle of the technique is the application of a high voltage electric field to obtain the separation of the ions present in a liquid electrolyte solution. The different electrophoretic mobility of the ion metabolites is the key to get their separation during the experiment. As a matter of facts, this electrophoretic mobility is linked to the charge (the most important parameter because no charge means no detection), the structure and the size of the analyte and so if metabolites under analysis have different values for these parameters their separation during a CE-MS experiment is possible.

The two cited hyphenated techniques provide the possibility to provide the separation of analytes and the ability to detect metabolites at low concentrations, commonly in the ranges of nanomoles/litre or micromoles/litre. However typical analysis times can provide separation of mixtures range from 10 to 120 minutes, and therefore sample throughput is lower than used in direct infusion mass spectrometry or the nuclear resonance techniques described below.

Finally, as reported before, MS coupled techniques such as GC-MS, LC-MS, and CE-MS all operate with different mechanisms to separate analytes and provide the detection of different sets of metabolites. However, many of them can be detected by more than one analytical technique. Therefore, applying two or more of these methods will provide the detection of a higher number of metabolites than the application of a single technique.

NMR: Nuclear Magnetic Resonance

The process of sample treatment, the quantification range of metabolite concentrations, the high level of experimental reproducibility, and the non-destructive nature of this spectroscopic technique has made NMR as one of the most exciting platforms for large-scale clinical metabolomics studies [9]. You can use as NMR as MS methods to identify and detect metabolites but both approaches have their pros and cons. In particular, NMR is a versatile

metabolic profiling tool, which doesn't need extensive sample preparation steps or reference compounds for labelling (compared to LC-MS or GC-MS techniques).

To fully understand how NMR can be applied in the metabolomics field is necessary to understand what kind of physical principles govern this analytical technique.

Nuclear Magnetic Resonance is a spectroscopy where an intrinsic property of atomic nuclei, the spin (Nuclear) is probed by the application of a large external magnet (Magnetic). As happened in all the spectroscopies techniques, NMR exploits a component of the electromagnetic radiation to promote transitions between nuclear energy levels (Resonance).

Every atom consists of three different elementary particles: the electrons which orbit around the nucleus, the protons and the neutrons which constitute the nucleus itself. In particular, the nucleus possesses an intrinsic property called spin (I). There are three possible cases:

- When the number of both neutrons and protons are odd, nuclei have integral spins (e.g., $I = 1, 2, 3 \dots$);
- When the total number of neutrons and protons is odd, nuclei have fractional or half-integer spins (e.g. $I = \frac{1}{2}, \frac{3}{2}, \frac{5}{2} \dots$);
- When the number of both neutrons and protons are even, nuclei have spin $I = 0$ (e.g. $^{12}\text{C}, ^{16}\text{O}, ^{32}\text{S}, \dots$).

Isotopes of particular interest for our purposes are ^1H , ^{13}C , ^{19}F and ^{31}P , all of which have $I = \frac{1}{2}$. A nucleus will have $2I + 1$ possible orientation (where I is the net spin). For instance, a $1/2$ spin nucleus can be orientated in two different positions. In the absence of an external magnetic field, these orientations have the same energy, and they are randomly oriented. During an NMR experiment, the sample is placed in an external magnetic field, called B_0 , which forces the bar magnets to align with (low energy) or against (high energy) the B_0 . In the presence of this external magnetic field, two-spin states exist, $+\frac{1}{2}$ and $-\frac{1}{2}$. A transition between the two spins state can be induced when there are different energy levels. Given that, a nucleus with a low energy orientation can be transferred to a higher energy

orientation. This energetic transfer can be measured, and therefore, the absorption of energy during this transition is the basis of NMR spectroscopy. The difference in energy between the two spin states is dependent on the external magnetic field strength and is always very small. Due to that, strong magnetic fields are required by NMR spectroscopy. Furthermore, this transfer of energy between low and high energy orientation is proportional to the number of nuclei possessed by the molecule, which provides in the end information about molecular quantification.

The last aspect to remark is that the nucleus magnetic field and the external magnetic field are not equal and electrons spinning around the nucleus can act as a shield for the nucleus magnetic field. The result of the difference between these magnetic fields is called nuclear shielding, which in the end, causes the shift of the signal frequency (chemical shift). The structure of the molecule affects this chemical shifting and these differences provide information for molecular identification. For ^1H NMR analysis, the chemical shift is usually measured relative to a reference compound, such as trimethylsilyl propionate (TSP) and tetramethylsilane (TMS).

Summary

The concepts expressed so far show once again how, in the development of metabolomics methods and strategies, the utilization of a specific analytical apparatus rather than another one may lead to spectacular failures or on the contrary to phenomenal discoveries. Each technique has pros and cons in the management, cost, flexibility, or detection.

All these parameters are briefly resumed in table 1.2.

Table 1.2.: Comparison between most used analytical metabolomics platforms [41]

	NMR	LC-MS	GC-MS	CE-MS
Advantage	Non-destructive High reproducibility Fast analysis Minimal sample preparation No derivatization No separation Good identification software Good identification database Liquid and solid compatibility Easy novel compound identification	Minimal sample size needed High sensitivity Flexible technique High number of metabolites detectable	Good sensitivity Robust Relatively inexpensive Modest sample size needed Good identification software and database Excellent separation reproducibility	High resolution Automation Applicable to large molecules Low sample volumes required
Disadvantage	Expensive instrumentation Cannot detect salt and inorganic ion Cannot detect no-protonated metabolites Larger analytical sample required Relatively low sensitive	Sample not recoverable Expensive instrumentation Slow (20-30 min/sample) Poor separation resolution vs. GC Poor robustness Limited database for identification Limited identification software Difficult novel compound identification	Slow (20-30 min/sample) Difficult novel compound identification Sample not recoverable Derivatization procedure required	Not suitable for molecules > 20 kDa Sample not recoverable Control of pH is critical

References of Chapter 1

- [1] NASA - National Aeronautics et al. *NASA's twins study explores space through you*. 2016. URL: <https://phys.org/news/2016-04-nasa-twins-explores-space.html> (cit. on p. 1).
- [2] J. Allen et al. "High-throughput classification of yeast mutants for functional genomics using metabolic footprinting". In: *Nature Biotechnology* 21, 692-696 (2003). DOI: 10.1038/nbt823 (cit. on p. 11).
- [3] D. Antcliffe et al. "Metabonomics and intensive care". In: *Critical Care* 20, 68 (2016). DOI: 10.1186/s13054-016-1222-8 (cit. on p. 11).
- [4] C. Bedia et al. *Comprehensive Analytical Chemistry*. Chapter Nineteen - Applications of Metabolomics Analysis in Environmental Research - Volume 82, Pages 533-582. Barcelona Spain: Elsevier, 2018 (cit. on p. 11).
- [5] L.G. Biesecker. "Hypothesis-generating research and predictive medicine". In: *Genome Research* 23(7), 1051-1053 (2013). DOI: 10.1101/gr.157826.113 (cit. on p. 8).
- [6] National Center for Biotechnology Information. *PubMed: US National Library of Medicine National Institutes of Health Search database*. 2020. URL: <https://www.ncbi.nlm.nih.gov/pubmed> (cit. on pp. 6, 7).
- [7] I. Blaženović et al. "Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics". In: *Metabolites* 8(2), 31 (2018). DOI: 10.3390/metabo8020031 (cit. on p. 9).
- [8] J. Borner et al. "A high-throughput method for microbial metabolome analysis using gas chromatography/mass spectrometry". In: *Analytical Biochemistry* 367, 143-151 (2007). DOI: 10.1016/j.ab.2007.04.036 (cit. on p. 11).
- [9] A.H. Emwas et al. "NMR Spectroscopy for Metabolomics Research". In: *Metabolites* 9(7), E123 (2019). DOI: 10.3390/metabo9070123 (cit. on p. 14).

-
- [10] S. Fujigaki et al. “Identification of serum biomarkers of chemoradiosensitivity in esophageal cancer via the targeted metabolomics approach”. In: *Future Medicin* 12(8), 146-158 (2018). DOI: 10.2217/bmm-2017-0449 (cit. on p. 11).
- [11] C.J. García et al. “Targeted Metabolomics Analysis and Identification of Biomarkers for Predicting Browning of Fresh-Cut Lettuce”. In: *Journal of Agricultural and Food Chemistry* 67(20), 5908-5917 (2018). DOI: 10.1021/acs.jafc.9b01539 (cit. on p. 11).
- [12] S.C. Gates et al. “Quantitative metabolic profiling based on gas chromatography”. In: *Clinical Chemistry* 24(10), 1663-1673 (1978) (cit. on p. 4).
- [13] LIPID MAPS Lipidomics Gateway. *LIPID MAPS Lipidomics Gateway*. 2020. URL: <https://www.lipidmaps.org/> (cit. on p. 5).
- [14] C. Gieger et al. “Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum”. In: *Plos Genetics* 4(11), e1000282 (2008). DOI: 10.1371/journal.pgen.1000282 (cit. on p. 10).
- [15] Google. *Google scholar: Stand on the shoulders of giants*. 2020. URL: <https://scholar.google.com> (cit. on pp. 6, 7).
- [16] W.J. Griffiths et al. “Mass spectrometry: from proteomics to metabolomics and lipidomics”. In: *Chemical Society Reviews* 38(7), 1882-1896 (2009). DOI: 10.1039/b618553n (cit. on p. 4).
- [17] K. Hollywood et al. “Metabolomics: current technologies and future trends”. In: *Proteomics* 6, 4716-4723 (2006). DOI: 10.1002/pmic.200600106 (cit. on p. 4).
- [18] E.C. Horning et al. “Metabolic profiles: gas-phase methods for analysis of metabolites”. In: *Clinical Chemistry* 17, 802-809 (1971) (cit. on p. 4).
- [19] D.I. Hoult et al. “Observation of tissue metabolites using ^{31}P nuclear magnetic resonance”. In: *Nature* 252, 285-287 (1974). DOI: 10.1038/252285a0 (cit. on p. 5).

- [20] M.C. Jewett et al. “Fungal metabolite analysis in genomics and phenomics”. In: *Current Opinion in Biotechnology* 17(2), 191-197 (2006). DOI: 10.1016/j.copbio.2006.02.001 (cit. on p. 10).
- [21] J.K. Kim et al. “Time-course metabolic profiling in *Arabidopsis thaliana* cell cultures after salt stress treatment”. In: *Journal of Experimental Botany* 58(3), 415-424 (2007). DOI: 10.1093/jxb/er1216 (cit. on p. 11).
- [22] T. Kind et al. “A comprehensive urinary metabolomic approach for identifying kidney cancer”. In: *Analytical Biochemistry* 363, 185-195 (2007). DOI: 10.1016/j.ab.2007.01.028 (cit. on p. 11).
- [23] A. Klassen et al. *Metabolomics: Definitions and Significance in Systems Biology*. Chapter: Metabolomics: Definitions and Significance in Systems Biology. In: Sussulini A. (eds); pp. 3-15. Berlin: Springer, Cham, 2017 (cit. on p. 3).
- [24] Kanehisa Laboratories. *KEGG: Kyoto Encyclopedia of Genes and Genomes*. 2020. URL: <https://www.genome.jp/kegg/> (cit. on p. 6).
- [25] E.M. Lenz et al. “Analytical Strategies in Metabonomics”. In: *Journal of Proteome Research* 6(2), 443-458 (2007). DOI: 10.1021/pr0605217 (cit. on p. 5).
- [26] J.C. Lindon et al. “The role of metabonomics in toxicology and its evaluation by the COMET project”. In: *Toxicology and Applied Pharmacology* 187(3), 137-46 (2003). DOI: 10.1016/S0041-008X(02)00079-0 (cit. on p. 11).
- [27] J. MacCracken et al. “From ants to analogues. Puzzles and promises in diabetes management”. In: *Journal of Postgraduate Medicine* 101, 138-150 (1997). DOI: 10.3810/pgm.1997.04.195 (cit. on p. 4).
- [28] National Magnetic Resonance Facility - Madison. *Madison-Qingdao Metabolomics Consortium Database*. 2020. URL: <http://mmcd.nmrfa.madison.wisc.edu/> (cit. on p. 5).

-
- [29] V. Mapelli et al. “Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology”. In: *Trends in Biotechnology* 21, 692–696 (2008). DOI: 10.1016/j.tibtech.2008.05.008 (cit. on p. 11).
- [30] A. Nicholls et al. “Use of metabonomics to identify impaired fatty acid metabolism as the mechanism of a drug-induced toxicity”. In: *Chemical Research in Toxicology* 27(2), 165-173 (2004). DOI: 10.1021/tx034123j (cit. on p. 11).
- [31] M.V. Nototny et al. “Biochemical Individuality Reflected in Chromatographic, Electrophoretic and Mass-Spectrometric Profiles”. In: *Journal of Chromatography B* 866, 26-47 (2007). DOI: 10.1016/j.jchromb.2007.10.007 (cit. on p. 4).
- [32] I. Parveen et al. “Application of gas chromatography-mass spectrometry metabolite profiling techniques to the analysis of heathland plant diets of sheep”. In: *Journal of Agricultural and Food Chemistry* 55, 1129-1138 (2007). DOI: 10.1021/jf062995w (cit. on p. 11).
- [33] Oxford University Press. *Oxford Dictionaries*. 2019. URL: <https://www.oxforddictionaries.com> (cit. on p. 3).
- [34] MassBank Project. *MassBank Europe - High Quality Mass Spectral Database*. 2020. URL: <https://www.http://massbank.eu> (cit. on p. 5).
- [35] L.D. Roberts et al. “Targeted Metabolomics”. In: *Current Protocols in Molecular Biology* 98(1), 30.2.1–30.2.24 (2012). DOI: 10.1002/0471142727.mb3002s98 (cit. on p. 10).
- [36] C.A. Smith et al. “METLIN: A Metabolite Mass Spectral Database”. In: *Therapeutic Drug Monitoring* 27, 6-10 (2005). DOI: 10.1097/01.ftd.0000179845.53213.39 (cit. on p. 5).
- [37] The Metabolomics Society. *Metabolomics:Introduction*. 2019. URL: <https://www.http://metabolomicssociety.org> (cit. on p. 6).
- [38] J. Van der Greef et al. “Symbiosis of chemometrics and metabolomics: past, present, and future”. In: *Journal of Chemometrics* 19, 376-386 (2005). DOI: 10.1002/cem.941 (cit. on pp. 1, 4).

- [39] F.A. Van Dorsten et al. “Metabonomics Approach To Determine Metabolic Differences between Green Tea and Black Tea Consumption”. In: *Journal of Agricultural and Food Chemistry* 54(18), 6929-6938 (2006). DOI: 10.1021/jf061016x (cit. on p. 11).
- [40] D.S. Wishart et al. “HMDB: the Human Metabolome Database”. In: *Nucleic Acids Research* 35, D521-D526 (2007). DOI: 10.1093/nar/gk1923 (cit. on p. 5).
- [41] D.S. Wishart. “Metabolomics: applications to food science and nutrition research”. In: *Trends in food science and technology* 19(9), 482-493 (2008). DOI: 10.1016/j.tifs.2008.03.003 (cit. on pp. 11, 17).
- [42] D.S. Wishart et al. “HMDB 4.0: the human metabolome database for 2018”. In: *Nucleic acids research* 46, D608-D616 (2018). DOI: 10.1093/nar/gkx1089 (cit. on p. 5).
- [43] A. Zhang et al. “Modern analytical techniques in metabolomics analysis”. In: *Analyst* 137(2), 293-300 (2011). DOI: 10.1039/C1AN15605E (cit. on p. 12).
- [44] T. Zhang et al. “Current Trends and Innovations in Bioanalytical Techniques of Metabolomics”. In: *Critical Reviews in Analytical Chemistry* 46(4), 342-351 (2015). DOI: 10.1080/10408347.2015.1079475 (cit. on p. 3).

Aims and outlines of PhD thesis

2

The principal aim of my PhD project was to elucidate and understand if metabolomics approaches could be applied to the samples commonly treated in the Lab I was applied (Prof. Claudio Medana's Lab - Department of Molecular Biotechnology and Health Sciences, University of Turin) to detect possible important targets for a specific illness or to unravel information about new, unexpected molecules in mechanisms related to human health or environmental samples. As already mentioned in chapter 1, the status of a biological system or the fundamental characteristics of specific samples belonging to different domains can be measured by using a LC-MS metabolomics approach. To accomplish this ambitious purpose, metabolomics approaches were developed and carried out on projects developed within the Lab where I spent my PhD training. So, depending on the request and samples under analysis, the best analytical method, considering analytical process from the treatment of sample to statistical analysis, was developed to maximize the information obtained from a single analysis, saving money and times.

LC-MS-based metabolomics strategies, instruments, and methods applied will be discussed one by one in the next chapters. In detail, my PhD thesis is structured in thirteen chapters, where the first four chapters will introduce the readers to the metabolomics world and will give all the necessary instruments to understand the following sections. As a matter of facts, from chapter 5 to chapter 13 of this thesis I will present results obtained in different fields using as targeted as untargeted metabolomics methods.

Chapter 1 provides an introduction concerning history, the most important techniques and main fields where metabolomics approaches are used. Moving on, the aims and purposes of my work are reported in chapter 2. Chapter 3 presents, on the contrary, an introduction to the LC-MS platforms I managed during my experience as in Italy as in Japan, where I spent

five months of PhD training. The last introductory chapter, chapter 4, concerns on the contrary on targeted and untargeted metabolomics methods I used, by emphasizing which aspects to pay attention to. From now on, each chapter of this thesis (chapters 5÷12) will outline a type of experience that I faced during my PhD program. Some of the experiments described have led to publications of scientific articles on international reviews. To obtain greater clarity and simplicity of reading, all these chapters will always be divided into these sections:

- Introduction;
- Aim of the work;
- Sample preparation and method settings;
- Results and discussion;
- Conclusion.

Moreover, from chapter 5 to chapter 7, all the experiences reported concerning targeted metabolomics approaches with biological fluids and, on the contrary, from chapter 10 to chapter 12, all the analyses are carried out by untargeted analysis. Chapter 8 deals with a semi-targeted metabolomics application, and it can be considered as a cross-link between targeted and untargeted worlds. Concluding, in chapter 13, I will sum up with results, goals achieved, and future perspectives.

Thanks to the approaches reported in this PhD thesis, the literature concerning the development of quality analytical methods in the metabolomics fields can be considered complete.

In this chapter, the metabolomics platform used all along my PhD experience will be introduced. Basic aspects of liquid chromatography and mass spectrometry will be introduced, and links between these techniques and metabolomics will be highlighted. This chapter serves to introduce and define concepts that will be used in later chapters.

As already reported in 1.2.4, different analytical platforms like nuclear magnetic resonance (NMR) and mass spectrometry (MS) coupled to separation techniques such as liquid chromatography (LC) or gas chromatography (GC) can be used to explore the metabolomics world. In the experiments described below, I used LC-MS techniques. These methods will be described since LC-MS offers superior sensitivity and selectivity compared to most other methods while also having a relatively high throughput and broad coverage of compounds.

All the methods developed in this thesis used the instrumental equipment reported below.

- HPLC:
 - HPLC Dionex Ultimate 3000 by Thermo Scientific;
 - Shimadzu Nexera X2 series from Shimadzu Corporation;
 - Acquity™ Ultra Performance LC I-class;
 - Nanospace Si 2 HPLC System - Shiseido.
- Mass spectrometry analysers:
 - High Resolution Mass Spectrometry HRMS LTQ Orbitrap™ by Thermo Scientific coupled with an ESI or APCI interface according to the analysis;
 - QTRAP® 5500 SCIEX coupled with an ESI or APCI interface;
 - Waters SYNAPT G2-S Mass Spectrometer coupled with ESI or DESI interface depending the purpose of the study;
 - Q Exactive Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer.

- Diode array detector:
 - Finningan Surveyor PDA Plus Detector.

3.1. Sample preparation

Sample preparation is always the first step for performing targeted metabolomics analyses as well as untargeted ones. The combinations of sample type (biological samples such as human blood or fresh leaves for examples) and the analytical platform used require a different approach in the preparation step. This treatment is probably the most time-consuming in analysis and the crucial one to extract the maximum analytical information. Due to the possible treatments, sample preparation could involve different practices: chemical digestion of the species operated by acids or alkaline solutions, liquid-liquid or solid-liquid extraction, dissolution of the sample in a proper solvent, simple dilution of the matrix, sample pre-concentration by SPE system, etc.

For these reasons, the perfect extraction method doesn't exist, but an appropriate sample preparation depends on the matrix which is the object of investigation. In particular, for the studies described in the next chapters, heterogeneous matrix will be discussed:

- Urine: these matrices often require soft or no treatment than a dilution. However, in the cases exposed later (chapter 5, urine need to be extracted and pre-concentrated to detect the analytes of interest.
- Plasma and serum: for our purposes, these kinds of matrices need to be extracted. It is recommendable to choose plasma because the risk of haemolysis and thrombocytolysis is generally lower than what happens with serum and samples are easier to be treated [5]. Unfortunately, in plasma samples, an anticoagulant agent needs to be added. Different kinds of anticoagulants are commercially available, but all of them can interfere with the detection of analytes of interest through ion-suppression. Barri and Dragsted [1] studied different anticoagulants and showed in their work how heparin plasma is the most appropriate anticoagulant for LC-MS metabolomics analysis. Other substances

such as citrate or EDTA can be used, but they cause a higher ionic suppression, which in turn interferes with quantification of compounds of interest.

- **Tissues:** monitor the distribution of an endogenous compound or confirm the presence/absence of a molecule in a specific site of action. These are two of the most important reasons why tissues are so studied in LC-MS metabolomics analysis. Disruption and the consequent homogenization of the matrix is always a required step. Homogenization can be conducted manually or mechanically (sonication or blending, for example), and, if required, a further enzymatic digestion step can be added. The nature of the tissue (fibrous or soft) plays a central role during setting up the best homogenization procedure.
- **Food:** the world of food analysis conducted via LC-MS would be borderless. From fruits to meats, from high fatty products to water, each matrix requires a different approach, also depending on which molecules we are interested in. Solvent extraction and solid-phase extraction (SPE)¹ can be used as a valid method to isolate or enrich target analytes from foods. The bottleneck of these techniques is the need to have trained staff that limits the statistical variability of this time-consuming procedure.

3.2. Liquid chromatography

By using the term chromatography, I refer to a variety of techniques based on the partitioning or the distribution of a sample (called solute) between the mobile phase where it is dissolved and a fixed or stationary phase. In

¹Solid-phase extraction is a method of sample preparation that separates and concentrates analytes from a solution, prior to analysis. This technique separates analytes using physical or chemical absorption in reaction with a solid media. The media is mounted on an absorbent material (usually a disk or a cartridge). The analytes present in the solution are retained in the media as the sample passes through the absorbent material. The analytes are then eluted from the media using a solvent in which the analytes are soluble. This is the solution retained for the analysis.

Some of the advantages of the solid-phase extraction compared to the liquid-liquid extraction are the less solvent usage, the less time and labour required, and the reduction in overhead costs.

Table 3.1.: *Different possible chromatographic methods available [7]*

Method	Mobile phase	Stationary phase	Retention varies with
Gas-liquid chromatography	Gas	Liquid	Molecular size/polarity
Gas-solid chromatography	Gas	Solid	Molecular size/polarity
Supercritical fluid chromatography	Supercritical fluid	Solid	Molecular size/polarity
Reverse phase chromatography	Polar liquid	Non polar liquid or solid	Molecular size/polarity
Normal phase chromatography	Less polar or non-polar liquid	Polar liquid or solid	Molecular size/polarity
Ion exchange chromatography	Polar liquid	Ionic solid	Molecular charge
Size-exclusion chromatography	Liquid	Solid	Molecular size
Hydrophobic-interaction chromatography	Polar liquid	Nonpolar liquid or solid	Molecular size/polarity
Affinity chromatography	Water	Binding sites	Specific structure

a chromatographic separation between the mobile and the stationary phase a huge series of equilibrations occur. The mobile phase may be either a gas (GC), a liquid (LC), or a supercritical fluid (SFC). On the contrary, the stationary phase may usually be a solid or a liquid. The different combinations of mobile stationary phases generate various techniques able to separate different kinds of molecules. These strategies are briefly resumed in table 3.1.

The eluted compounds obtained by one of the techniques considered in table 3.1 are transported by the mobile phase to the detector. Here they are recorded with a bell-shaped curve called the Gaussian curve. The signals are known as peaks, and the sum of all these entities is known as a chromatogram. The peaks give as qualitative as quantitative information about the sample.

For our purposes, the technique to take in consideration is the Reverse phase High-Pressure Liquid Chromatography (RP-HPLC). The HPLC is an analytical (and preparative too) technique where the liquid mobile phase is pumped through a column containing very finely packed particles. The two actors involved in the interaction are the chemical groups present on the surface of the particles and the analytes dissolved in the mobile phase. Different analytes will have different types of affinity for the stationary particles and thus will permit the separation during the chromatographic run. Most of the HPLC separations treated in this PhD work are RP-HPLC. In a reverse phase chromatography on the silica particles, hydrophobic carbon chains are linked. Different kinds of chains can be connected but for most of the studies conducted a C18 silica column is used (C18 means a chain of 18 atoms of Carbon long linked to the silica particles). The mobile phase used is now a very polar one like water, methanol, acetonitrile, or isopropanol.

This is called reverse phase because it is precisely the reverse of normal phase chromatography where mobile phases are non-polar, and particles of the stationary phase are polar ones.

3.3. Mass Spectrometry

The fundamental parts of a mass spectrometer are the ion source, the mass analyser, and the detector. Depending on the type of platform used, the properties of the samples, and the data required the nature of these components is variable. In any case, the fundamental challenge to the application of mass spectrometry to any class of analytes is the production of gas-phase ions of those species, and difficulties in producing gas-phase ions can prevent mass spectrometric analysis. The development of electrospray ionization (ESI), atmospheric-pressure chemical ionization (APCI) and ambient ionization techniques such as desorption electrospray ionization (DESI) or direct analysis in real-time (DART) revolutionized metabolomics analysis by MS. To break the train of thought concerning MS analysers, an introduction of the mentioned ionization sources is reported in appendix A.

For metabolomics research, different types of mass analysers are commonly used: triple quadrupole, time-of-flight (TOF), OrbitrapTM, quadrupoles, and ion traps, and each of them has specific characteristics. They can be used to separate all analytes in a sample during global (or full scan) analysis, or they can be used as a filter to deflect only specific ions towards the detector. In table 3.2 the specifications and the suitable applications fields of the different mass analyser were reported.

Table 3.2.: *Mass analyser parameters used in metabolomics studies. QIT: Quadrupole Ion Trap; LTQ: Linear Trap Quadropole; Q-q-Q: Triple quadrupole; Q-q-LIT: Triple quadrupole Linear trap; TOF: Time-of-Flight; TOF-TOF: Tandem time-of-flight; Q-q-TOF: Time-of-Flight triple quadrupole; FTICR: Fourier-transform ion cyclotron resonance*

Instrument	Mass resolution	Mass accuracy	Sensitivity	m/z range	Scan rate	Dinamic range	MS/MS capability	Ion source
QIT	1000	100-1000 ppm	picomole	50-2000	Moderate	1E3	MS ⁿ	ESI
LTQ	2000	100-500 ppm	femtomole	50-2000	Fast	1E4	MS ⁿ	ESI
Q-q-Q	1000	100-1000 ppm	attomole to femtomole	10-4000	Moderate	6E6	MS/MS	ESI
Q-q-LIT	2000	100-500 ppm	femtomole	10-3000	Fast	4E6	MS ⁿ	ESI
TOF	10000-20000	10-20 ppm	femtomole	No upper limit	Fast	1E4	N/A	MALDI
TOF-TOF	10000-20000	10-20 ppm	femtomole	No upper limit	Fast	1E4	MS/MS	MALDI
Q-q-TOF	10000-20000	10-20 ppm	femtomole	No upper limit	Moderate to fast	1E4	MS/MS	ESI; MALDI
FTICR	50000-750000	<2 ppm	femtomole	50-2000 or 200-4000	Slow	1E3	MS ⁿ	ESI; MALDI
LTQ-Orbitrap TM	30000-100000	<5 ppm	femtomole	50-2000 or 200-4000	Moderate to fast	5E3	MS ⁿ	ESI; MALDI

After the generation and their separation, the ions are revealed by a detector. Most often, these detectors are electron multipliers or micro-channel plates, and when an ion collides against them, they emit a cascade of electrons. To improve sensitivity and register a clear signal, this electron cascade must be amplified (or multiplied). The entire MS system must be under extreme vacuum (10^{-6} to 10^{-8} torr). As a matter of fact, the atoms of the gases if are not removed can collide with sample ions and alter their paths

For my purpose, the most interesting and essential MS platform is the LTQ OrbitrapTM. This analytical instrument captures ions around a central spindle electrode, and the values of their m/z are determined by the harmonic oscillation frequencies they have along the central spindle. OrbitrapTM can achieve extremely high sensitivity and high-resolution accurate mass (HRMS) of obtained mass spectra.

3.3.1. High Resolution Mass Spectrometry LTQ-OrbitrapTM

By using high-performance liquid chromatography (HPLC) systems is possible to obtain a chromatographic separation (and a compound detection too) of the analytes based on the time with which they are eluted from the analytical column. The retention time alone, however, cannot positively identify a molecule because many compounds can have, using the same experimental conditions, an identical retention time. Besides, the coelution of other compounds present in the sample can cause errors during quantification. The LTQ-MS spectrometer provides a high level of analytical specificity if compared to other detectors used with LC systems, such as a single-stage mass analyser mass spectrometer or UV-Vis detectors.

In addition to the mass-to-charge ratio (m/z) express with accurate mass measurements (achieving a precision of 10 ppm or better), this kind of MS platform can provide multiple levels of mass analysis. Each level of MS analysis adds a new dimension of specificity for unequivocal compound identification. In particular, the LTQ can provide these kinds of levels in mass analysis:

- Single-stage mass analysis. This kind of analysis provides molecular mass information. Typically, atmospheric pressure ionization like ESI

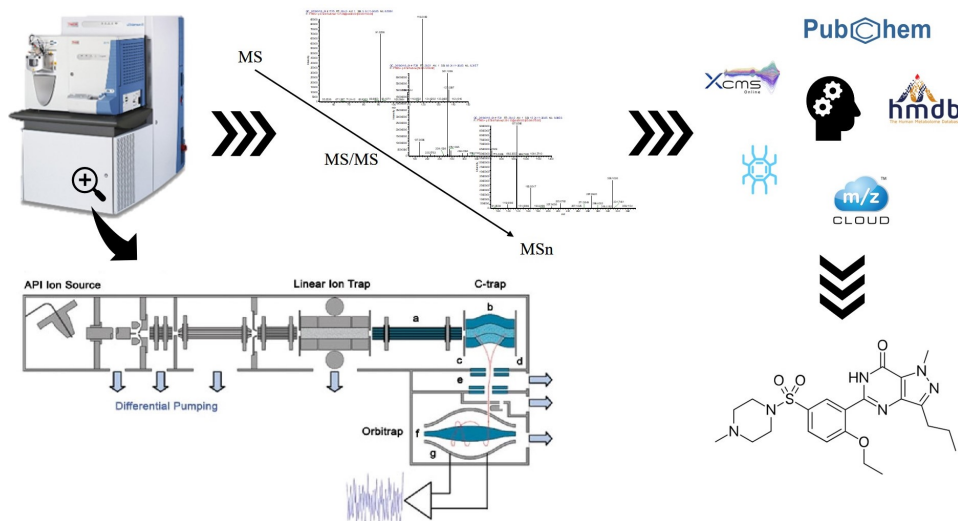


Figure 3.1.: Schematics of a LTQ-OrbitrapTM instruments and work-flow analysis [8]. In the zoom section are reported the main components of this instrument: (a) transfer octupole; (b) curved RF-only quadrupole (C-trap); (c) gate electrode; (d) trap electrode; (e) ion optics; (f) inner orbitrap electrode (central electrode); (g) outer orbitrap electrode

or APCI ionization produces mass spectra that provide information about the analyte of interest.

- Two-stage mass analysis, generally called tandem mass spectrometry or MS/MS experiments. They provide, on the contrary, structural information of the compound of interest by monitoring how precursor ion fragments react when exposed to a new stage of excitation. There are two types of MS/MS analysis:
 - Full-scan MS/MS monitors the production of all product ions from a specific precursor ion;
 - Selected reaction monitoring (SRM) MS/MS analysis monitors a specific reaction path, which is the production of a specific product ion from a specific precursor ion.
- Multi-stage mass analysis, MSⁿ. In this technique, the product ions from MS/MS (or a previous MS stage) are subjected to re-fragmentation. By providing deeper structural information, this approach is one of the most effective techniques in resolving, for example, heterogeneity or

branching pattern.

All these possibilities allow using this analytical platform for targeted as for untargeted metabolomics studies. As a matter of fact, in the first case LTQ system can detect a single analyte in the sample of interest with a limit of detection, variable from the target of interest, from mg/kg to $\mu\text{m}/\text{kg}$. On the contrary, in the second case, this system can - and must be used - for untargeted metabolomics studies thanks to the possibility of monitoring and creating dependent scan and MS^n experiments for all the molecules contained the samples, from the expected to the new ones. A general scheme of LTQ instrument with its work-flow is reported in figure 3.1. As underlined in this figure, once the spectra are obtained to attribute the proper name to the metabolite under analysis, a time-consuming step concerning online database research and spectra interpretation is necessary.

3.3.2. Tohoku University analytical platforms: Waters SYNAPT G2-S and Thermo Scientific Q Exactive OrbitrapTM

During my PhD project, I had the honour to join for five months Prof. Masayuki Yamamoto's group at Tohoku University, located in Tohoku Medical Megabank Organization, Department of Integrative Genomics [11, 12]. The projects I dealt with during my stay at the ToMMO Organization under the supervision of Prof. Daisuke Saigusa were two. The first one was called SpaceMice Project and the second one Lung cancer cell Project. I will discuss these projects in two separate chapters (chapter 9 and chapter 10, respectively). What I want to underline in this paragraph are instruments I used to conduct metabolomics analysis in Japan.

Instruments used were a Waters SYNAPT G2-S Mass Spectrometer coupled with ESI or DESI interface (depending on the purpose of the study) and a Q Exactive Plus Hybrid Quadrupole-OrbitrapTM Mass Spectrometer from Thermo Scientific. As already mentioned in the last lines of the section 3.3.1, also these kinds of analytical platforms express their best potential thanks to the possibility of conduct MS^n analysis over a large m/z range.

SYNAPT G2-S system (figure 3.2 is a high-resolution accurate mass MS/MS platform able to operate in two different modes: TOF mode and mobility-TOF mode. Time-of-Flight (ToF) is a mass spectrometry technique able to measure the mass-to-charge ratio of an ion via time of flight measurement in a field free region. The ions of the same charge, due to the acceleration in this known strength electrical field, have the same kinetic energy. The mass-to-charge ratio influences, on the contrary, the velocity of ions: if the charge is the same, heavier ions are related to lower speeds, and on the contrary, lighter ions mean an increase in velocity. Furthermore, ions with higher charges but the same mass will also increase in velocity.

Given that, each ion needs a specific time to reach a detector located at a known distance. This measurable time will depend on the speed of the ion, and therefore it can be related to the mass-to-charge ratio. From this ratio and known experimental parameters, each ion can be identified. For the SYNAPT G2-S TOF mode, mass range is 20 to 32000 m/z in High-resolution mode; the mass spectra can be acquired up to rate of 30 scans per second and the mass measurement accuracy of the instrument, in High-resolution mode, is better than 1 ppm².

Another important aspect to remark is the presence of Travelling wave (T-Wave) ion mobility to increase separation with high-efficiency. By using the T-Wave ion mobility, an additional dimension of separation can be obtained, based on the molecular size and shape. The range of benefits derived is huge: isomer separation, elimination of interferences, and generation of cleaner spectra are just a few examples. Furthermore, a wide range of experimental options is available thanks to the possibility of using different ionization sources, quickly interchangeable and ready to use: Electrospray ionization (ESI), Atmospheric Pressure Chemical Ionization (APCI), Atmospheric Solids Analysis Probe (ASAP), Desorption Electrospray Ionization (DESI), Direct Analysis in Real-Time (DART) and Laser Diode Thermal Desorption (LDTD). Concluding, all the cited parameters make the SYNAPT G2-S an instrument of choice for the analysis of both solid (ex. ambient Mass Spectrometry) and extracted matrices in the con-

²Based on 10 consecutive repeat measurements of the $[M+Na]^+$ ion of raffinose (m/z 527.1588), using the $[M+H]^+$ ion of leucine enkephalin (m/z 556.2771) as LockSpray Lockmass.

text of untargeted metabolomics studies.

The Q exactive Plus Hybrid Quadrupole-OrbitrapTM Mass Spectrometer from Thermo Scientific is a high-resolution mass spectrometer consisting of six main components: the ion source, an injection flatapole with mass resolving capabilities, a quadrupole mass filter for precursor ion selection, a curved linear trap (known as C-trap), a collision cell (new and not present in the instrument described in section 3.3.1) , and OrbitrapTM analyser for Fourier transform mass analysis. A schematic representation of Q ExactiveTM Plus mass spectrometer system is reported in figure 3.3. In this analyser, after ionization operated by the selected API source, the injection flatapole transmits the produced ions into the quadrupole. The injection flatapole works as a coarse pre-filtering of ions according to their mass-to-charge (m/z) ratios. Quadrupole rod can transmit and filter the ions according to m/z ratio. And then ions are moved into the C-trap, a system where the energy of the ions is dampened and where these species are accumulated for a selected time. Finally, the ions are then injected into the Orbitrap analyser. Furthermore, if MS/MS scans are necessary, the ions before entering in the OrbitrapTM are injected into the HCD cell for fragmentation. After fragmentation, the ions are transferred back into the C-Trap and finally detected by the Orbitrap analyser.

Q exactive Plus Hybrid Quadrupole-OrbitrapTM Mass Spectrometer too, thanks to the features described above, can be a fundamental tool for conduct untargeted analysis in metabolomics field.

3.3.3. SCIEX QTRAP[®] 5500

If we are not interested in untargeted metabolomics and if the metabolites, we want to detect in our samples have a shallow concentration the QTRAP[®] 5500 (and generally a triple quadrupole platform) is the suitable instrument to use. The QTRAP[®] 5500 LC/MS/MS system excel at metabolite identification, detection, and confirmation, for example, in low-level pesticides or biomarker quantification in biological samples. The principles of the technique are the same cited previously in 3.3. The only difference

is the analyser used to conduct the experiments. In a triple quadrupole mass analyser (also called QqQ or Q-q-Q- system), charged ions generated in the ion source enter the mass analyser. The analyser consists of three quadrupoles (or multipoles) in which two of them are identically each other and one (the q2) is bent to improve sensitivity and to reduce physical instrument encumbrance. While having the capability to function as a standard triple quad LC-MS/MS, QTRAP[®] system a linear ion trap (LIT) too, to perform a multitude of additional workflows beyond necessary multiple reaction monitoring (MRM). As a matter of facts, with this instrument, we can achieve these experiments: Precursor ion (PI), Neutral Loss (NL), Product Ion, MRM, Enhanced MS (EMS), Enhanced Multiply Charged (EMC), Enhanced Resolution, Enhanced Product Ion and MS3 (MS/MS/MS).

However, the most common set for analysis is the following:

- Q1: used as a filter for specific m/z (precursor ion);
- q2: used as a collision cell to fragment the precursor ion and generate product ions;
- Q3: set to specific m/z (SRM or MRM) or scan mode (product ion scan).

The scheme of this instrumentation is reported in figure 3.4. As we can see, a fourth quadrupole is present (called q0). It is a quadrupole array that works as ions lens and it transmits ions from the interface region to mass resolving quadrupole.

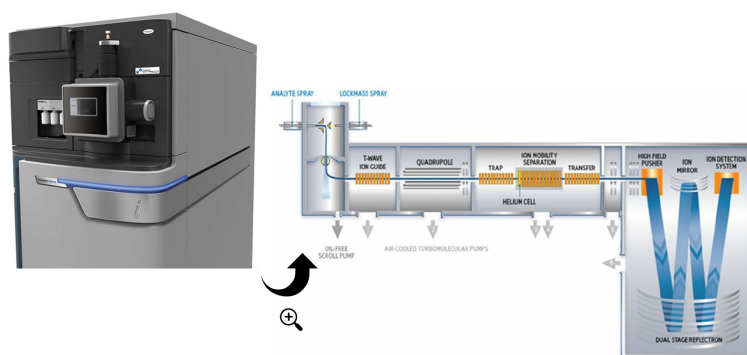


Figure 3.2.: Schematic representation of Synapt HDMS instrument indicating the major components of the instrument [2]

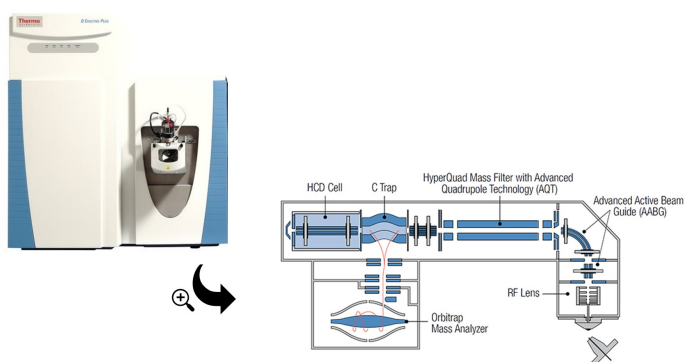


Figure 3.3.: Schematic structure of a Thermo Scientific Q Exactive™ Plus Hybrid Quadrupole Orbitrap™ Mass Spectrometer [8]

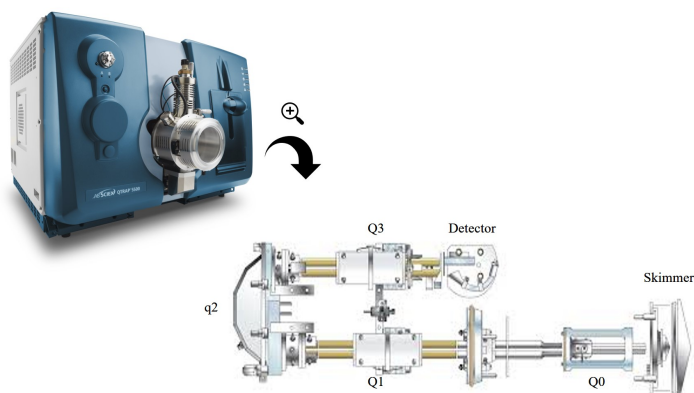


Figure 3.4.: Schematics of a QQQ QTRAP® 5500 SCIEX instrument [9]

References of Chapter 3

- [1] T. Barri et al. “UPLC-ESI-QTOF/MS and multivariate data analysis for blood plasma and serum metabolomics: effect of experimental artefacts and anticoagulant”. In: *Analytica Chimica Acta* 20(768), 118-128 (2013). DOI: 10.1016/j.aca.2013.01.015 (cit. on p. 26).
- [2] Waters Corporation. *Waters Corporation: The Science of What’s Possible*. 2019. URL: <https://www.waters.com> (cit. on p. 37).
- [3] J. Gross. *Spettrometria di Massa - Edizione Italiana a cura di Enrico Davoli e Claudio Medana*. Chapter 4. Napoli: EdiSES, 2016.
- [4] D.C. Harris. *Chimica analitica quantitativa - Edizione Italiana*. Milano: Zanichelli, 2017.
- [5] T. Hyotylainen et al. *Chromatographic Methods in Metabolomics*. Chapter 3. London: The Royal Society of Chemistry, 2013 (cit. on p. 26).
- [6] S. Mitra. *Sample Preparation Techniques in Analytical Chemistry*. Hoboken: Wiley-Interscience, 2007.
- [7] S. Nielsen. *Food Analysis*. Food Science Text Series, Chap. 27, pp. 475 - 495. New York, USA: Springer International publishing, 2017 (cit. on p. 28).
- [8] Thermo Fisher Scientific. *Thermo Fisher Scientific*. 2020. URL: <https://www.thermofisher.com> (cit. on pp. 32, 37).
- [9] DH Tech. Dev. Pte. Ltd. SCIEX. *SCIEX - The power of precision*. 2019. URL: <https://www.sciex.com> (cit. on p. 37).
- [10] D.A. Skoog et al. *Chimica analitica strumentale - Edizione Italiana a cura di L. Sabbatini*. Napoli: EdiSES, 2009.
- [11] Tohoku University. *Tohoku University: Creating global excellence*. 2019. URL: <http://www.tohoku.ac.jp/en/> (cit. on p. 33).
- [12] Tohoku University. *ToMMo - Tohoku Medical Megabank Organization*. 2019. URL: <https://www.megabank.tohoku.ac.jp/english/> (cit. on p. 33).

Validation methodology in targeted and untargeted methods

How can an analyst be sure about its elaborated targeted or untargeted metabolomics analytical method? What kind of parameters must be evaluated to consider acceptable the procedure for the desired purpose? The answer to all these questions is the reliability of analytical methods, or rather the validation process. Especially for a powerful and sophisticated analytical technique as LC-MS, it is imperative to assure that the method works (and will work) as expected, and the obtained results through these methods are reliable over time.

Unfortunately, nowadays, the setup of a guide that explains how validation should be carried out still not exist. But, the different validation parameters to evaluate and the acceptance criteria are in general the same: sensitivity, reproducibility over time, recovery, robustness, analyte stability, time, and cost are some criteria present in most of the published papers.

In the present chapter, I will report two different sections: the first one for targeted metabolomics analysis validation and the second one for untargeted validation. In this way, each time in the present PhD thesis I assert that the method has been validated, I refer to the practices here described.

4.1. Validation of targeted methods

The validation procedure is a time-consuming procedure that requires an established plan consisting of planned experiments as well as expected outcomes. The results obtained must then match with requirements from guidelines or other documents. LC-MS guidelines supply to analyst precise analytical protocols to perform and provide specific parameters' values, which, if not overcome, attest to the lower-quality of the analytical method developed.

Why must a method be validated? The development of a new analysis method from scratch or the demonstration of the equivalence of a standard

method with the in-method examination can be two reasons to validate an analytical method. So, the final purpose of the validation is to demonstrate without a doubt that the values of all the parameters under analysis are satisfactory. Three different cases can arise.

1. Methods must be applied in the scope of laws or directives. In this case, the requirements reported in the documents must be followed properly.
2. A client/industry/organization etc. requires specific requirements. The request can be as superior as inferior to those in the guidelines. Based on the analytical platform used, the analyst will conduct or refuse the validation procedure.
3. No external requirement. The analyst can set up the requirements freely based on the knowledge of the subject and the analytical platform of the Lab.

During my PhD experience, all the analyses I conducted (and analyses reported in this work) refer to the second and third points of this list. In the Lab for validation of targeted metabolomics analytical methods, several procedures were followed: EMA (European Medicine Agency [15]) protocol, FDA guidelines [11], and Accredia manuals [1].

Considering the last one cited, it's described in UNI CEI EN ISO/IEC 17025, "General requirements for the competence of testing and calibration laboratories. Application Guideline" [1]. Among all the different variables rising during a validation method, for our purpose it was decided to evaluate these parameters: linearity (DIFF%), sensitivity (LOD, LOQ, LLOQ), selectivity (SEL%), relative percentage error (Er%), and percent relative standard deviation (RSD%). Figure 4.1 summarizes the sequence we conduct in our Lab to validate the chosen analytical method.

Stability

The validation should start with the stability parameters by checking whether the selected chromatographic and MS conditions allow detection of analytes of interest [3]. For example, if the analyte decomposes during

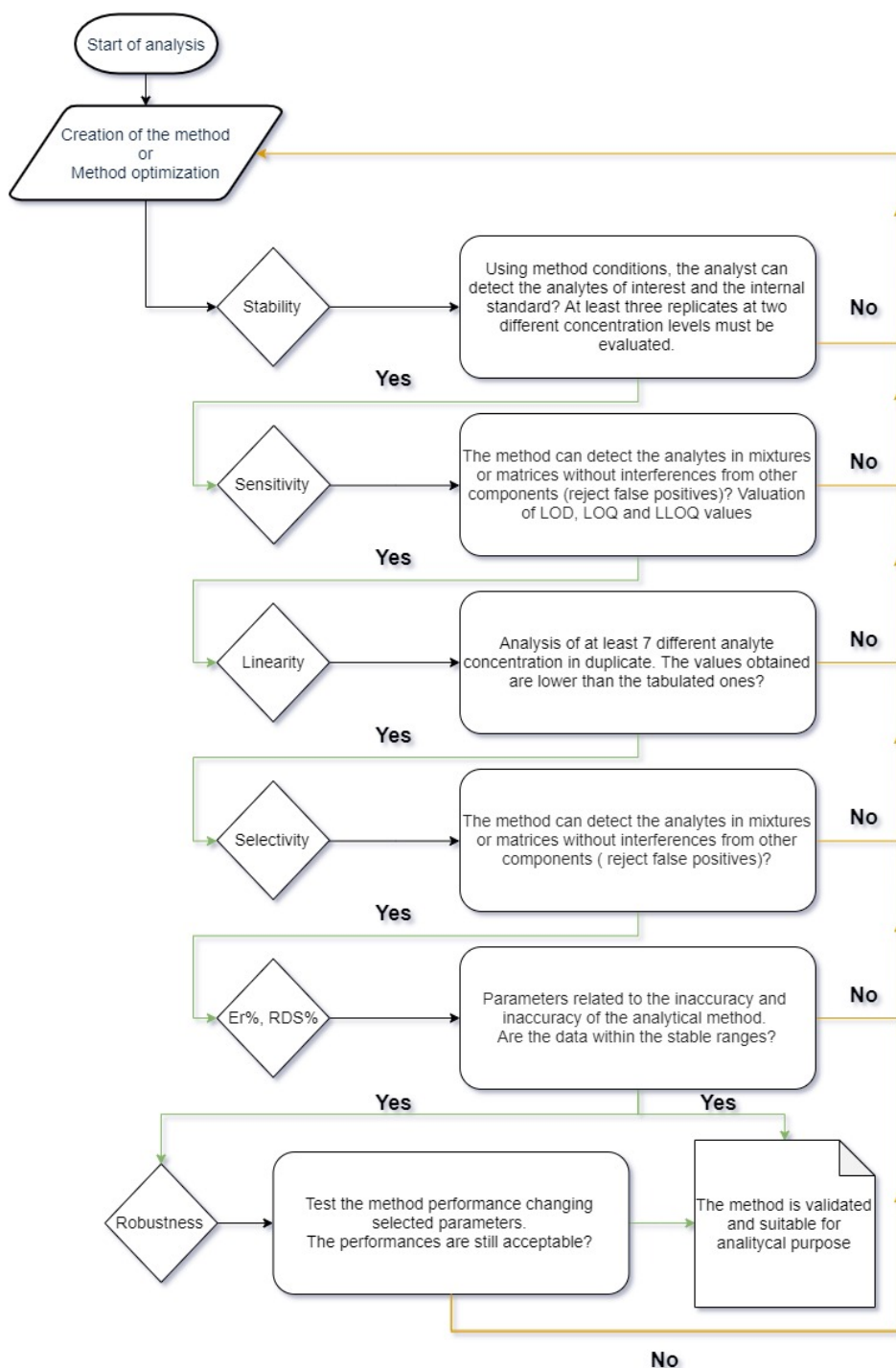


Figure 4.1.: Flowchart of the sequence of valuations in LC-MS target method validation in our Lab. All steps are explained in detail in section 4.1

the storage in the auto-sampler or it is eluted from the column in the dead time, no linear relation can be achieved, and the method needs improvement. At least three free replicates at two different concentration levels must be analysed to evaluate the stability of the method. If the retention time and the intensity of the peak are still the same over the three different injections we can consider the method stable.

Sensitivity: LOD, LOQ, LLOQ

Sensitivity is expressed through the parameters of Limit Of Detection (LOD), Limit Of Quantification (LOQ), and Lower Limit Of Detection (LLOQ). The LOD is the smallest amount or concentration of analyte in the test sample that can be reliably distinguished from zero [9, 10]. One of the purposes of LOD is to clarify the ability of the analytical method to detect the lowest level of analyte compared to other kinds of methods [19]. Depending on the samples and the methods used, different mathematical procedures can be used to calculate the LOD values [18].

On the contrary, Limit of quantification (LOQ) is defined as the lowest analyte concentration in the sample of analyte that can be determined with an acceptable repeatability and trueness [9, 10]. LOQ can also be called quantification limit, quantification limit, or limit of determination. In some cases, LOQ is called a wrongly lower limit of quantification. The lower limit of quantification LLOQ is the lowest concentration of analyte that produces an analytical signal, and it corresponds to the lowest point of the calibration curve. So, if LOD and LOQ are calculated mathematical values (can be calculated using the software too), the LLOQ is an experimental one.

The LOD and LOQ values can be calculated as $3\sigma/m$ and $10\sigma/m$ respectively. The value of the standard deviation, σ , is obtained through the equation 4.1, where n is the number of points used to construct the curve, y refers to the experimental peak area of one of the n samples, and y_0 are the areas recalculated through the calibration curve equation $y_0 = m_i x_0 + q_i$.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y - y')^2}{n - 2}} \quad (4.1)$$

The results obtained for LOD, LOQ, and LLOQ have to be interpreted. This aspect should be done in this way:

- Analyte of interest is not detected, or its concentration is under the LOD value. In this case, the analyte cannot be detected because under LOD and no data area reported¹;
- Concentration of the analyte is located between LOD and LOQ values. The analyte is present at level trace, but no data concerning quantification are reported;
- Level of analyte content is at or above LOQ. The result must be quantified.

Linearity

The linearity of an analytical procedure is the method's ability to obtain results that are directly proportional to the concentration of the analyte in samples within a specific and establish range [6]. To mathematically evaluate linearity, a medium calibration curve is constructed for each analyte. This curve is the average of at least three repetitions of a generic equation $y = m_i x + q_i$ where y is the area value of the peak, x is a concentration value and m_i and q_i the slope and the intercept respectively. Linearity is calculated through the equation 4.2². In equation 4.2, \bar{m} is the medium slope of the medium calibration curve, and if the experimental values are lower than 25.00, the linearity values can be considered proper for validation.

$$DIFF\% = \left| \frac{m_i - \bar{m}}{\bar{m}} \right| \times 100 \quad (4.2)$$

Selectivity

IUPAC defines selectivity as “the extent to which other substances interfere with the determination of a substance according to a given procedure”

¹It's important to remark that if the result is below LOD doesn't mean that there is no analyte in the sample but, on the contrary, that the method used is not able to detect the analyte in that particular sample. The concentration of zero never happen in analytical chemistry[12].

²This parameter, and the ones described in this section, are also discussed by the American Institution FDA [11].

[13]. In general, the larger the interferences, the less selective the method is. If a given method is 100% selective toward the target, the method can be defined as specific. Selectivity is calculated using the SEL% parameter obtained calculating the concentration of analyte in three matrices where the analyte should be absent, and three matrices positivized to the LLOQ level. The equation 4.3 clarifies how to calculate this parameter.

$$SEL\% = \left| \frac{100 \times \text{Peak area at retention time in the matrix}}{\text{Average of 3 area at Rt in positivized matrix}} \right| \quad (4.3)$$

Er% and RSD%

The relative percentage error Er% is a parameter through which it is expressed the inaccuracy of the method. It is determined through the analysis of three matrices positivized to the LLOQ concentration. On the contrary, the relative standard deviation (RSD%) parameter is related to the error with which the method's inaccuracy is expressed and it's calculated through the analysis of three matrices LLOQ positivized. The equations 4.4 and 4.5 report how to calculate Er% and RSD% respectively.

$$E_r\% = \left| 100 \times \left(\frac{[\text{Analyte conc.}]_{\text{predicted}} - [\text{Analyte conc.}]_{\text{experimental}}}{[\text{Analyte conc.}]_{\text{experimental}}} \right) \right| \quad (4.4)$$

$$RSD\% = \frac{\sigma \times 100}{\text{Average of 3 area at Rt in positivized matrix}} \quad (4.5)$$

If all the parameters analysed are considered acceptable by comparison with the data reported in the guidelines, then the method can be considered validated and applicable as a routine method.

A final consideration must be done. In some circumstances, there is a further parameter to be evaluated at the end of the validation development phase. It is the robustness of the method. It consists of the variation of the method performance characteristics to verify if the method's results remain

unaffected when slight variations are applied. In LC-MS analyses, typical modifications are the variation of pH mobile phase, the change in mobile phase composition (for example, instead of 5 mM ammonium acetate in H₂O use just H₂O), the substitution of the chromatographic column with another having the same characteristics, the variation of temperature or flow rate. If the values prove to be uncharged after the change, the method will be considered robust and applicable.

As already anticipated above, this kind of targeted validation is not the only one available. As a matter of fact, in 2019 the European Medicine Agency [15] has issued guidelines [2] for harmonized the validation of bioanalytical methods for “*chemical and biological drug quantification and their application in the analysis of study samples*”[2]. In one of the last projects I followed during my PhD research activity (targeted analysis of quorum sensing molecules, see chapter 7) validation procedure was conducted using the guidelines and parameters reported in this guide. For chromatographic methods, EMA underlines how a full validation should include different elements: selectivity, specificity, matrix effect, calibration curve response, operative linear range (from the lower limit of quantification LLOQ to the upper limit of quantification ULOQ), accuracy, precision, carry-over, and stability. The following bullet point list resumes the characteristic that each of the cited parameters must respect:

- Selectivity is evaluated using blank samples (matrix samples processed without the addition of an analyte or IS) obtained from at least six different sources. The evaluation of selectivity should demonstrate that no significant response attributable to interfering components is observed at the retention time(s) of the analyte or the IS in the blank samples. Responses detected and attributable to interfering components should not be more than 20% of the analyte response at the LLOQ and not more than 5% of the IS response in the LLOQ sample for each matrix.
- Specificity is the ability of a bioanalytical method to detect and differentiate the analyte from other 337 substances. Responses detected and attributable to interfering components should not be more than

20% of the analyte response at the LLOQ and not more than 5% of the IS response in the LLOQ sample.

- Matrix effect should be evaluated by analysing at least three replicates of low and high-quality control samples, each prepared using matrix from at least six different sources/lots. The accuracy should be within $\pm 15\%$ of the nominal concentration and the precision (percent coefficient of variation, %CV) should not be greater than 15% in all individual matrix sources/lots.
- Calibration standards should be prepared in the same biological matrix as the study samples. The calibration range is defined by the LLOQ, which is the lowest calibration standard, and the ULOQ, which is the highest calibration standard. A calibration curve should be generated with a blank sample, a zero sample (blank sample spiked with IS), and at least six concentration levels of calibration standards, including the LLOQ and the ULOQ. If back-calculated concentrations of each calibration standard are within $\pm 20\%$ of the nominal concentration at the LLOQ and within $\pm 15\%$ at all the other levels, the calibration curve can be considered valid. Furthermore, at least 75% of the calibration standards with a minimum of 6 calibration standard levels should meet the above criteria.
- QCs are intended to mimic study samples and should be prepared by spiking the matrix with a known quantity of analyte. During method validation, a minimum of four QCs concentration levels within the calibration curve range must be prepared: LLOQ, low-QC level, medium QC level, high QC level. QCs samples are essential for the evaluation of accuracy and precision. Within-run accuracy and precision should be evaluated by analysing at least five replicates at each QC concentration level in each analytical run. Between-run accuracy and precision should be evaluated by analysing each QC concentration level in at least 3 analytical runs over at least two days. The overall accuracy at each concentration level should be within $\pm 15\%$ of the nominal concentration, except at the LLOQ, where it should be within $\pm 20\%$. The precision (%CV) of the concentrations determined at each level should

not exceed 15%, except at the LLOQ, where it should not exceed 20%.

- For carry-over, during validation, it should be assessed by analysing blank samples after the calibration standard at the ULOQ. Carry-over in the blank samples following the highest calibration standard should not be greater than 20% of the analyte response at the LLOQ and 5% for the IS.
- For stability parameters, evaluations should be carried out to ensure that every step was taken during sample preparation, processing, and analysis, as well as the storage conditions used, do not affect the concentration of the analytes of interest. Different elements must be taken into account: stability of stock solutions, stability to freeze-thaw process, long-term matrix stability, etc.

4.2. Validation of untargeted methods

The global metabolomic profiling approach cannot be related to a validation procedure, such as the one described for targeted metabolomics. As a matter of facts, numerous aspects distinguish the two approaches [4, 17]: untargeted is a hypothesis-generating technique, it's focused on global analysis, and it proposes relative quantifications and qualitative identifications thanks to database comparison; on the contrary, targeted metabolomics focuses on defined molecules and aims by using reference standards to absolute quantification. So, since in untargeted metabolomics, hundreds and hundreds of molecules are detected over hundreds of samples [8], it is impossible to use internal standards or generate all the calibration curves. Furthermore, it is impossible to provide an estimation for the accuracy (how close the measurement conforms to the real value) or the precision (how close the results are to each other, in the same biological identity of the methods, and the definition of limits of detection (LOD), limits of quantification (LOQ), and linearity previously described (section 4.1) are no more valid.

How assured then, in a LC-MS untargeted analysis, the reliability of results? For these kinds of analysis, these criteria can accomplish very well to this purpose: reproducibility of retention times all along with the batch

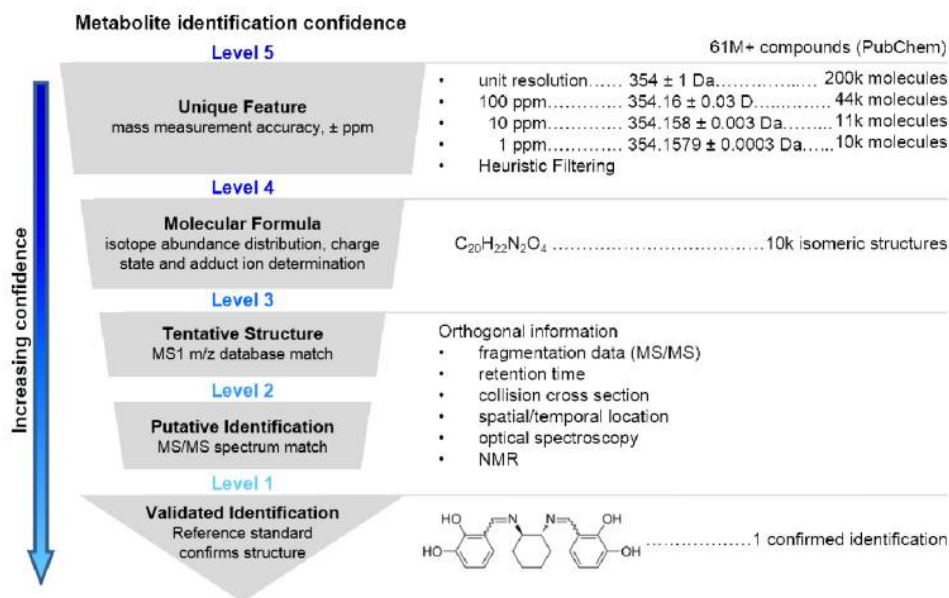


Figure 4.2.: Example of workflow for metabolite identification [17]

samples, accurate masses of molecules (essential for data alignment) [14], intensities of the detected peaks (essential to detect biological variation), online databases [20], and quality control (QCs) samples [5, 7]. Using all these parameters, we can project an efficient workflow (such as the one reported in figure 4.2), which, if respected, allows the analyst to obtain results with a high degree of reliability.

From the list of cited criteria, the most important one in global metabolomics analysis is the quality control (QC) sample. As a matter of fact, in untargeted metabolomics analyses the QC samples provide a mechanism to evaluate the quality of the data (the variance of the data) generated during the whole analytical workflow. The entire collection of the samples must (or should be) represented qualitatively and quantitatively by QC samples.

Different types of QCs samples can be prepared. In some experiments, they are constituted by a small aliquot of each biological sample and are withdrawn to create a pool of all the samples under analysis [16]. On the contrary, when it is not possible to create a QC pooled sample (e.g., limit number of samples or study samples collected over the years), commercially available QC samples can be purchased. This kind of QC samples are consti-

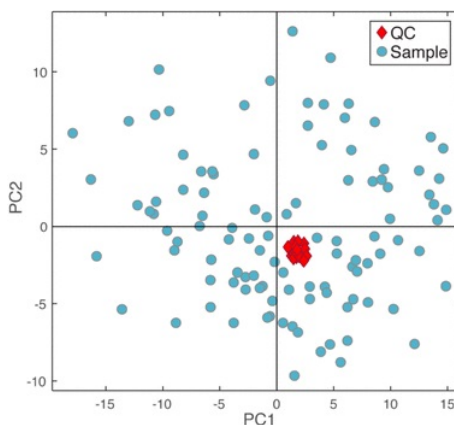


Figure 4.3.: *Example of typical PCA scores plot for a data set. The QC samples (red squares) are grouped as a single spot in the center of the graph. This is an indication of the good result achieved [5]*

tuted by a mixture of authentic chemical standards representing the target analytes [5].

Then, once the QCs preparation step is concluded, the QCs samples injections are intermittently done over all the sample batch to assess the variance observed in the data.

After the analytical study, if the calculated QC measures are within the acceptance criteria decided at the beginning of the process (for example, CV% under 20% or the intensity of a certain molecule over a certain value) for precision and accuracy, and the acquired data for the test samples are within the linear calibration range, the data be considered acceptable and the statistical analysis can be performed. Otherwise, the test had a negative outcome, and the extraction procedure or the analytical method (or both) must be reviewed before re-processing the sample batch. So, QC samples data are useful for the analyst to monitor and correct drifts in the signal, separate high- and low-quality data and integrate data belonging from different analytical experiments.

Generally, if QC samples work properly, a PCA plot such as the one reported in figure 4.3 is obtained. Specific QC preparation procedures applied in the Lab will be discussed and reported in chapters 9 to 12, where my untargeted metabolomics experiences are exposed and commented. However, a gen-



Figure 4.4.: *Flowchart for global metabolomics LC-MS method apply in our Lab. The different steps will be deeply discussed in chapters 9 to 12*

eral workflow idea is reported in figure 4.4. This procedure is an adaptation of the one reported in [16]. Most of the untargeted procedures I learned during my PhD experiences were achieved during my PhD visiting student period at the Tohoku University, under the supervision of Prof. Masayuki Yamamoto, Prof. Seizo Koshiha and Prof. Daisuke Saigusa.

In workflow 4.4 critical points are n° four and five, linked to the valuation of data obtained. This process is known as data processing. Data processing is the computational process of converting raw files generated through LC-MS platform into biological knowledge, including data deconvolution and

identification of metabolites detected.

Data deconvolution step, known as peak picking, is a laborious process to apply caused by the complexity and the variation of the dataset (for example, differences in mass-to-charge ratio, or in retention time). A LC-MS dataset is three dimensional: as a matter of fact, each metabolite is identified by (1) the retention time, (2) the mass-to-charge ratio, and (3) the chromatographic peak area. To simplify, on assume that for each metabolite, just one metabolite feature is present. The extracted ion chromatogram (XIC) of each analyte can be plotted manually by the operator. But thousands of XICs are present, and this process would take months of hard work. For this reason, this process is performed by specific software: one of the most popular metabolomics tool-analyser is XCMS [21], but many other software packages are available. Another issue is that the same metabolites can have slightly different mass-to-charge ratios or retention time among different samples. Aligning step is, therefore, a critical but indispensable process. In the end, the integration of data from different samples to build a single data matrix (representing the metabolite composition of all the samples) is constructed.

The final step is the chemical identification of metabolites, essential to perform data analysis and to define the biological importance of detected molecules. Metabolite identification process is done by comparing the results to databases (online database or private research laboratories libraries). Considering the online databases, one of the most important and used is the Human Metabolome Database. It contains information on endogenous and exogenous metabolites (more than 40000) present in humans. But this is not the one only available: KEGG database lists metabolites and groups of metabolites linked together in metabolic pathways; PubChem and ChemSpider, useful chemical databases reporting information not only about metabolites but chemicals in general; mass spectral libraries like mz-Cloud, METLIN, and MassBank contain accurate mass, and fragmentation mass spectral data for metabolites. All the cited tools are freely available to use.

References of Chapter 4

- [1] Accredia. *ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories. Application Guideline*. Milan, 2000 (cit. on p. 40).
- [2] European Medicin Agency. *ICH guideline M10 on bioanalytical method validation*. Chapters 2 and 3. London, United Kingdom: European Medicines Agency, 2019 (cit. on p. 45).
- [3] B. Babu et al. “Development and validation of stability indicating liquid chromatography tandem mass spectrometry method for the determination of dasatinib in bulk”. In: *International Journal of Research in Pharmaceutical Sciences* 8(4), 663-666 (2017). DOI: 10.1016/j.jpha.2013.12.002 (cit. on p. 40).
- [4] O. Begou et al. “Quality Control and Validation Issues in LC-MS Metabolomics”. In: *Methods In Molecular Biology - Springer* 1738, 15-26 (2018). DOI: 10.1007/978-1-4939-7643-0_2 (cit. on p. 47).
- [5] D. Broadhurst et al. “Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies”. In: *Metabolomics* 14, 72 (2018). DOI: 10.1007/s11306-018-1367-3 (cit. on pp. 48, 49).
- [6] E. De Hoffmann et al. *Mass Spectrometry: Principles and Applications, 3rd Edition*. Chapter 4. Hoboken: Wiley-Interscience, 2007 (cit. on p. 43).
- [7] D. Dudzik et al. “Quality assurance procedures for mass spectrometry untargeted metabolomics. A review”. In: *Journal of Pharmaceutical and Biomedical Analysis* 147, 149-173 (2018). DOI: 10.1016/j.jpba.2017.07.044 (cit. on p. 48).
- [8] W.B. Dunn et al. “The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans”. In: *Bioanalysis* 4(18), 2249-2264 (2012). DOI: 10.4155/bio.12.204 (cit. on p. 47).

-
- [9] H. Evard et al. "Tutorial on estimating the limit of detection using LC-MS analysis, part I: Theoretical review". In: *Analytica Chimica Acta* 942, 23-29 (2016). DOI: 10.1016/j.aca.2016.08.043 (cit. on p. 42).
- [10] H. Evard et al. "Tutorial on estimating the limit of detection using LC-MS analysis, part II: Practical aspects". In: *Analytica Chimica Acta* 942, 40-49 (2016). DOI: 10.1016/j.aca.2016.08.042 (cit. on p. 42).
- [11] U.S. Food et al. *U.S. Food and Drug Administration*. 2020. URL: <http://www.fda.gov/home> (cit. on pp. 40, 43).
- [12] M. Hecht et al. "Utilization of data below the analytical limit of quantitation in pharmacokinetic analysis and modeling: promoting interdisciplinary debate". In: *Bioanalysis Perspective* 10(15) (2018). DOI: 10.4155/bio-2018-0078 (cit. on p. 43).
- [13] A.D. McNaught et al. *IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*. Oxford: Blackwell Scientific Publications, 1997. DOI: 10.1351/goldbook. (cit. on p. 44).
- [14] G. Mullard et al. "A new strategy for MS/MS data acquisition applying multiple data dependent experiments on Orbitrap mass spectrometers in non-targeted metabolomic applications". In: *Metabolomics* 11: 1068 (2015). DOI: 10.1007/s11306-014-0763-6 (cit. on p. 48).
- [15] EU Agencies Network. *European Medicine Agencies - Science Medicines Health*. 2020. URL: <https://www.ema.europa.eu/en> (cit. on pp. 40, 45).
- [16] D. Saigusa et al. "Establishment of Protocols for Global Metabolomics by LC-MS for Biomarker Discovery". In: *Plos One* (2016). DOI: 10.1371/journal.pone.0160555 (cit. on pp. 48, 50).
- [17] A.C. Schrimpe-Rutledge et al. "Untargeted metabolomics strategies – Challenges and Emerging Directions". In: *Journal of The American Society for Mass Spectrometry* 27(12), 1897-1905 (2016). DOI: 10.1007/s13361-016-1469-y (cit. on pp. 47, 48).

- [18] A. Shrivastava et al. “Methods for the determination of limit of detection and limit of quantitation of the analytical methods”. In: *Chronicles of Young Scientists* 2(1), 21-25 (2011). DOI: 10.4103/2229-5186.79345 (cit. on p. 42).
- [19] M. Thompson et al. “Quality assurance schemes for analytical laboratories. Harmonized guidelines for single laboratory validation of methods of analysis (IUPAC technical report)”. In: *Pure and Applied Chemistry* 74(5), 835-855 (2002) (cit. on p. 42).
- [20] R.J.M. Weber et al. “Training needs in metabolomics”. In: *Metabolomics* 11: 784 (2015). DOI: 10.1007/s11306-015-0815-6 (cit. on p. 48).
- [21] XCMS. *The XCMS Online*. 2020. URL: https://xcmsonline.scripps.edu/landing_page.php?pgcontent=mainPage (cit. on p. 51).

Targeted Analysis: the plastic pollution

5

5.1. Introduction

The term oxidative stress indicates the alterations that occur in the biological tissues and cells when they are exposed to an excess of oxidizing agents [4]. In all aerobic organisms, the balance between oxidizing substances, including the Reactive Oxygen Species (ROS) and antioxidant defences, is called oxidative-reductive balance. Thus, if an imbalance between the ROS production and the effective antioxidant system is generated, rising a condition called oxidative stress [30]. Given that, all life forms conserve, within their cells, a reducing environment protected by enzymes able to maintain the reduced state through a constant supply of metabolic energy. Disorders of the normal redox state may cause toxic effects through the over-production of reactive chemical species (over the physiological threshold value) that damage (apoptosis) the cell components including proteins, lipids, and nucleic acids [29].

Most important oxidative stress exogenous sources are determined by environmental pollutants such as air pollutants, some types of food, toxic substances, and UV radiations [26]. Nowadays, the oxidative imbalance is caused not only by environmental exposure but also by the lifestyle; thus, among exogenous sources even diet, physical activities seem to cover an important role in the production and exposure to oxidant species [26]. Therefore, some of the most critical pollutants are Endocrine Disrupting Chemicals (EDCs), and among these must be counted the xenoestrogens bisphenol A (BPA) bisphenol S (BPS) having the most considerable attention from the scientific community, due to their effects on the human health [8]. The assessment of EDCs exposure is particularly important in the case of infants who, being in the early stages of development, and therefore, with still underdeveloped metabolic mechanisms and organs, appear to be more vulnerable and susceptible to the onset of potential stress and pathological

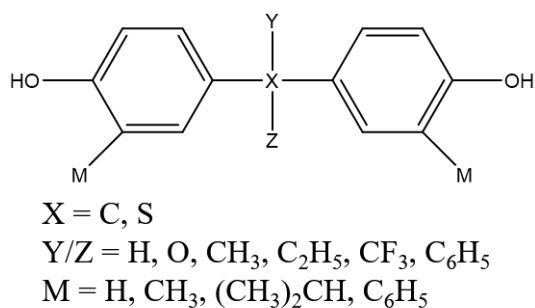


Figure 5.1.: *Possible bisphenol analogues detectable in nature*

conditions [17]. Moreover, the increasing concern about the late effects consecutive to early exposures identifies sensitive populations such as pregnant women and new-borns as a priority for biomonitoring studies [33]. Then, breast milk as mothers' and babies' urines appear as a relevant biological material for conducting such studies, giving access both to an estimated internal exposure level of the mothers and their babies and to an expected food exposure level of the breastfed newborn [5].

A final aspect to remark is that BPA and BPS are just two examples of the bisphenol analogues molecules that can act as EDCs. Figure 5.1 recaps the major bisphenol analogues most founded in the environment.

Bisphenol A

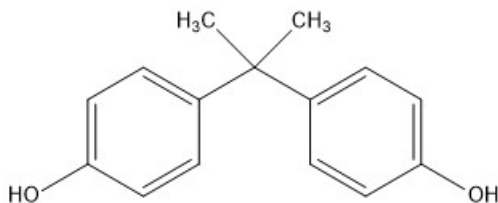


Figure 5.2.: *Chemical structures of Bisphenol A*

Bisphenol A (4,4'-isopropylidenediphenol), usually abbreviated as BPA, is a synthetic organic compound composed of two phenolic rings connected through a carbon bridge resulting from the reaction of 2 phenol molecules and acetone [31]. BPA is one of the most commonly produced and used

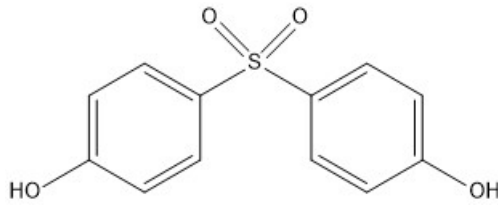
chemicals in the world: as a matter of fact, the annual global production of this molecule is about 8 million tons/year. Thus, it is that BPA can be easily found in everyday objects [16]: plastic bottles, toys, teats, medical and healthcare equipment, dental products, electronic devices, register receipts, books, ticket, newspapers, kitchen rolls, toilet paper, etc. [13, 21, 25]. The reason why is so used by industry is because of its properties such as thermal stability of BPA synthetic polymers, stabilizer in the production of vinyl chloride [25] and the low adsorption of moisture.

Usually, BPA can be detected in different biological matrices such as human urine, blood, cord blood, breast milk, fetal serum, semen, and placental tissue [10, 22, 32]. Among all these matrices, urine is considered one of the most appropriate body fluid to quantify the exposition [6]; actually, different researches [3, 27, 35] report how the urine sampled over different country and areas show level of this contaminant in the range of 75-90% and approximately 90% of the population had a detectable urinary BPA level (> 0.4 ng total BPA/mL) [18].

Why is BPA so dangerous? This plastic contaminant can bind to several kinds of receptors: estrogenic and androgenic receptors, aryl hydrocarbon receptor, and peroxisome proliferator-activated receptor which is associated with various hormones of the endocrine system and other systems of the body [14, 34, 37]. Therefore, BPA disrupts the function of various hormones, including sex hormones, leptin, insulin, and thyroxin. It may lead to adverse effects, especially on reproductive abilities, such as female infertility [7], male low sperm quality [19], and disruption the process of neurotransmissions [23]. That's means that during life stages such as neonatal period or childhood BPA can cause not complete brain development, reduction in survival (for concentration ≥ 500 mg/kg/day), a decrease in growth (≥ 300 mg/kg/day) and a delay on the onset of puberty (≥ 50 mg/kg/day) [24].

Bisphenol S

Bisphenol S (BPS, bis-(4-hydroxyphenyl)sulfone), has a similar chemical structure to BPA one: it contains a sulfone group (strong electron-

Figure 5.3.: *Chemical structures of Bisphenol S*

withdrawing ability) and two hydroxyl groups, two aspects that give a higher acidity and stability if compared to bisphenols. BPS was first synthesized as a dye in 1869 and is used as the substitute of BPA [15]. As reported in [1], the annual manufactured or imported rate of BPS in the European Economic Area is variable between 1000 and 10000 tons/year. The usage of BPS is not regulated and so it isn't very easy to specify exactly what kind of products contain it. However, BPS has been found in daily consumer products such as baby bottles and toys or personal care products [20]. In the same way of BPA, BPS can quickly enter the human body by ingestion (food or dust), inhalation (air or particle), and dermal absorption (dust or PCPs). Moreover, BPS is the primary substitute of BPA as a colour developer into thermal papers [16].

Therefore, humans are unfortunately widely exposed to BPS and this BPA-alternative has already been detected in 81% of urine samples sampled in the United States, China, and in six other Asian countries [11]. In [2, 10, 32], researchers demonstrated how the exclusive route for BPS human exposure is the adsorption through thermal papers and clothes. Compared to the high number of studies concerning BPA toxicity, the biological mechanisms ruling BPS toxic effects are much less investigated. Although, *in vitro* and *in vivo* experiments [28] indicate that BPS is an endocrine disruptor and may have potential health hazards, similar to that of BPA. Nowadays, a daily exposure dose or an oral maximum allowable dose of BPS for human exposure is not still established. However, if we look closely at their chemical structure, BPS and BPA are quite similar, and long-term exposure to BPS can have BPA similar effects on the development of the reproductive or the nervous system. Besides, as pregnancy is a particularly sensitive period

in terms of exposure to BPS, special attention must be paid to avoid exposure to this dangerous chemical (e.g., reducing the use of plastic products, PCPs, canned food, and dermal contact with thermal receipt papers)[36].

5.2. Aim of the work

The present work was conducted in collaboration with the Department of Public Health and Paediatrics - University of Turin - Prof. Roberto Bono group. The objective of this collaboration was to perform a complete oxidative stress profile analysis by monitoring different conditions of environmental and lifestyle habits. In particular, the main actors under investigation were children, mothers, and their newborns' babies, to better understand responses and adaptations to these different risk conditions.

Many risk conditions were analysed but, for our chemical points of view, we focused on the detection of bisphenol A and bisphenol S. Our purpose was to clarify if during pregnancy and the first days of breast-feeding may occur oxidative alterations caused by bisphenol A and bisphenol S in newborns, by detecting these analytes firstly in mothers' and babies' urines and in breast milks. These biological matrices are the most useful to monitor levels of these two contaminants in mothers and their breastfeeding babies. In fact, as already reported in 5.1, these pollutants are an endocrine disruptor, and they represent the onset of oxidative stress in newborns. Furthermore, they may be exposed to higher concentrations of free bisphenol if compared to adults (think about it: more or less, all the baby's products are plastic made). In this contest, studies show that the levels of unconjugated bisphenol in human adults' serum are typically low, while higher levels have been reported among newborns [17]. Also, the animal study demonstrates the passage of bisphenol through the placenta, with the subsequent passage in the umbilical cord and amniotic fluid of the fetus [9].

To accomplish this ambitious task, a UPLC-TQ MS targeted metabolomics procedure and an extraction protocol were built up. One important aspect to underline is that during Phase II metabolism Bisphenol A (and

related molecules) are mostly converted in BPA-glucuronide [12]. So, in order to quantify the real levels of these plasticizers within biological samples of interest, a de-glucuronidation reaction (β -glucuronidase/arylsulfatase) is required. By this way, the analyst will get a double benefit. On the one hand, fewer MRM transitions will be monitored (increment of the sensibility of the triple quadrupole); on the other hand, the difference between the samples treated with β -glucuronidase/arylsulfatase and the untreated ones will furnish the concentration of BPA-glucuronide present in the sample. The concentration values obtained are reported in table 5.7.

Now¹ both extraction sample methods (urine and milk matrices) are ready and validated, but only the results rising from urine matrix samples are already obtained. For this reason, I will report in 5.3 both extraction methods and instrumental parameters but in 5.4 only urine samples results.

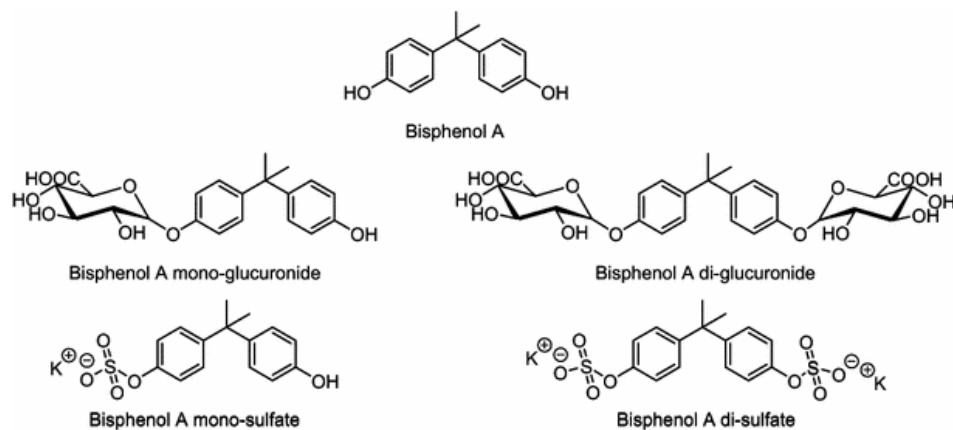


Figure 5.4.: Chemical structures of glucuronide and sulfate conjugates of Bisphenol A. Figure has been taken from [12]

5.3. Sample preparation and method settings

The epidemiological samples were selected to represent the population of mothers and their new-borns potentially and variously exposed to bisphenols. Subjects were recruited by consulting the register of subjects born

¹October 2019

during 2015-2017 at the University Unit of Neonatology of the Sant'Anna Gynaecological Hospital of Turin. At the end of the sampling, 100 mothers in physiological conditions with their babies and 50 mothers in pre-delivery circumstances potentially at risk (gestational diabetes, smoking addiction, hypertension, and hypothyroidism) with their babies were enrolled. Furthermore, to better evaluate the oxidative trend, a control population constituted of 50 women, aged between 25 and 35 years (same range of mothers considered), healthy and far, at least of two years, from any pregnancy was selected.

In the present project, urine and milk collection tubes, equipment used to handle and store samples, and each step in the chemical analysis procedure were screened to confirm the absence of BPA and BPS. So, from the sample collection until the mass spectrometry analysis the samples never entered in contact with any plastic material.

Extraction of total (Free + Conjugate) BPA and BPS from a urine sample and the extraction of free BPA and BPS from milk breast samples are different and they are reported below in two separate sections. In both extractions, internal standards have been used to compensate matrix effects: bisphenol-d8 and bisphenol-d16.

Protocol extraction of total BPA and BPS - urine samples

4 mL of urine previously collected are vortexed to homogenize the matrix. 2 mL are used for the determination of free bisphenols. In contrast, the other two are used for the quantification of bisphenol-glucuronides after 12 hours of incubation with 20 units of β -glucuronidase/arylsulfatase².

2 mL of urine are transferred in a tube pre-treated with methanol, vortexed, and acidified with HCl at pH 1. Subsequently, NaCl is added (for encouraging the salting-out process) together with 30 μ L of a standard solution containing BPS-d8 and BPA-d16 with a concentration of 0.10 and

²This analysis was conducted by Prof. Bono's group. After the incubation time with the enzyme β -glucuronidase /arylsulfatase, the extraction procedure is the same as the one described in detail above.

0.05 mg/L respectively.

750 μL of chloroform and 500 μL of acetone are added for the liquid-liquid extraction (LLE). Each sample is then vortexed and sonicated for 1 minute each. Finally, samples are centrifuged at room temperature at 1800 g for 20 minutes. The supernatant is collected and transferred to a new pre-treated vial. LLE is repeated twice adding 800 μL of chloroform each time. All the extracted supernatant obtained are collected and brought to dryness under nitrogen flux. The residue is suspended in 100 μL of a solution composed of ammonium acetate 5 mM in ultrapure H_2O and acetonitrile 5mM.

Protocol extraction of total BPA and BPS - breast milk samples

1 mL of breast milk is thawed, vortexed, and subsequently diluted with 9 mL of a mixture 80:20 Milli-Q water - methanol. Successively, positivize the sample with 30 μL of a standard solution containing BPS-d8 and BPA-d16 (concentration: 0.10 and 0.05 mg/L respectively). In parallel, an SPE C-18 cartridge (Strata C18-E 55 μm , 70 Å) was conditioned with 5 mL of methanol and immediately rebalanced with 5 mL of Milli-Q water. After loading the sample in the C-18 column, wash with 8 mL of Milli-Q water, and another one with 2 mL of a solution of methanol and H_2O (90:10) were carried out. Finally, the sample is eluted with 5 mL of methanol, brought to dryness under a gentle stream of nitrogen, and then re-suspended in 100 μL of a solution composed of ammonium acetate 5 mM in ultrapure H_2O 60% and acetonitrile 5 mM 40%.

Instrumental parameters

Instrumental analysis was carried out with a UPLC Shimadzu Nexera X2 System interfaced through an ESI source (Turbo VTM Ion) to a mass spectrometer SCIEX QTRAP[®] 5500 system. The chromatographic column was a Phenomenex Luna Omega C18 (1.6 μm , 100 mm \times 2.1 mm, 100 Å), and the mobile phase solvents for reverse phase analysis were $\text{CH}_3\text{COONH}_4$ 5 mmol in water as the weak solvent and $\text{CH}_3\text{COONH}_4$ 5 mmol in ACN as

Table 5.1.: *Chromatographic conditions for BPA and BPS, urine samples*

Injection	20 μ L		
Oven temperature	40° C		
Flux	350 μ L/min		
	Time (min)	CH ₃ COONH ₄ 5 mmol in H ₂ O	CH ₃ COONH ₄ 5 mmol in ACN
Chromatographic run	0	70%	30%
	5	0%	100%
	6	70%	30%
	11	70%	30%

Table 5.2.: *Chromatographic conditions for BPA and BPS, breast milk samples*

Injection	20 μ L		
Oven temperature	50° C		
Flux	350 μ L/min		
	Time (min)	CH ₃ COONH ₄ 5 mmol in H ₂ O	CH ₃ COONH ₄ 5 mmol in ACN
Chromatographic run	0	70%	30%
	2	70%	30%
	12	0%	100%
	15	0%	100%
	15.1	70%	30%
	22	70%	30%

the strong organic solvent. The final pH value of both solvents was corrected to 7.5-8.0 by adding a few drops of ammonia solution 33% to ensure a more significant presence of bisphenol A and bisphenol S in the deprotonated form.

The transitions considered during this research are shown in table 5.3. On the contrary, the chromatographic conditions used for urine samples and breast milk samples targeted metabolomics are reported in tables 5.1 and 5.2 respectively.

The result obtained by applying these instrumental parameters is visible in the figure 5.5, where a TIC chromatogram of a 10 μ g/kg standard solution is reported. As it's clear from the chromatogram, the BPS molecule generates a more intense chromatogram peak compared to the BPA one. So, a zoom-in is then necessary to appreciate the chromatographic peak generated by this analyte. Furthermore, this difference in the instrumental sensitivity towards the two analytes is then reflected on the different LLOQ that they possess: in fact, BPS has an LLOQ 5 times lower than BPA molecule (0.1 μ g/kg for BPS vs. 0.5 μ g/kg for BPA).

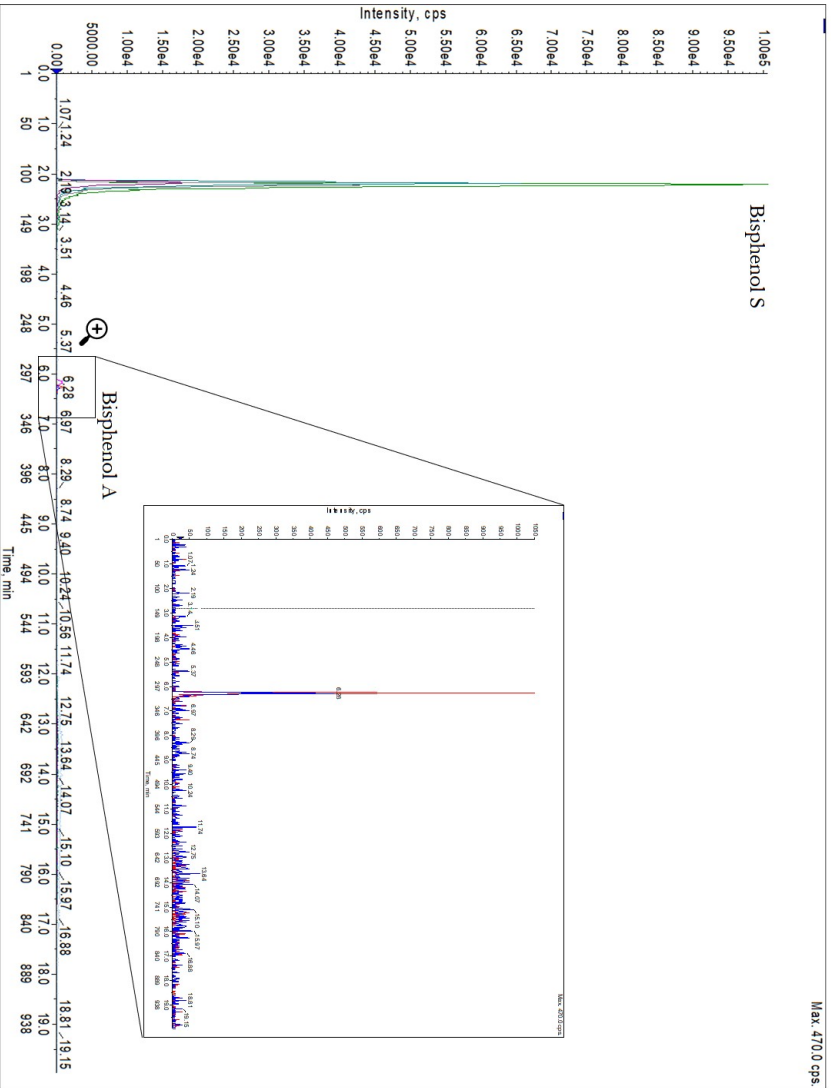


Figure 5.5.: TIC chromatogram of a 10 µg/kg standards solution. BPA peak, having a lower intensity compared to BPS, must be zoomed to be appreciated

Table 5.3.: *MS parameters for BPA and BPS urine and breast milk samples*

Source Temperature		400° C					
Polarity		Negative					
GS1; GS2		45:40					
Dwell time		100					
	Molecular ion (m/z)	Fragmented ion (m/z)	Retention time (min)	DP	EP	CE	CXP
BPA	227.0	212.0	6.20	-79	-9	-23	-14
		133		-79	-9	-30	-14
BPS	249.0	108.0	2.15	-75	-3	-35	-13
		92.0		-75	-3	-38	-13
BPA-d16	241.0	223.0	6.15	-62	-8	-25	-11
		142.0		-62	-8	-33	-18
BPS-d8	257.0	112.0	2.15	-120	-7	-34	-14
		96.0		-120	-7	-35	-14

5.4. Results and discussion

The parameters of the all subjects enrolled in the project are reported in tables 5.4 (mothers' group), (control group), and (babies group). In mothers' group to better clarify the possible exposure levels of candidates to bisphenols, each sample has been sub-grouped according to the presence or absence of some specific diseases (diabetes, hypertension, and hypothyroidism) influenced by the exposure to high levels of bisphenols. The two groups are called pathological (or pat.) and non-pathological (or non-pat.) group. The analysis shows differences statistically significant (non-parametric test: >0.05) between the two groups for weight in the pre-pregnancy period, BMI (Body Mass Index), and educational levels. In comparison, the other parameters are proved not to be statistically different among the two groups.

On the contrary, the children were just divided into the groups' boys and girls following their sex genders. Finally, in the control samples, no subgroups were adopted. For both groups, all the related parameters were tested and no statistically differences rise (non-parametric test: $p > 0.05$). Consequently, the samples should be considered as a homogeneous population.

In table 5.7 are reported the results concerning the BPA and BPS levels for mothers, controls, and children's urine samples, respectively. By observing these results, we can make some considerations.

First, the UPLC-TQ MS analytical method can detect in all the groups

Table 5.4.: *General characteristics of mothers enrolled in the project; N.S.: Not Significant, s.d.: standard deviation*

Parameters	Mothers			p-value
	Total	Pathological	Non pathological	
Number of samples	172	60	112	N.S.
Age (years) Mean \pm s.d.	33.6 \pm 4.93	33.9 \pm 5.0	33.4 \pm 4.9	N.S.
Height (m) Mean \pm s.d.	1.64 \pm 0.07	1.65 \pm 0.06	1.63 \pm 0.07	<0.05
Weight pre-pregnancy (kg) Mean \pm s.d.	63.8 \pm 13.6	70.9 \pm 17.1	60.1 \pm 9.4	<0.05
BMI mean \pm s.d.	23.7 \pm 4.8	25.9 \pm 6.1	20.5 \pm 3.3	<0.05
Scholarization level N. (%)	First level: 23 (13.4%)	First level: 9 (15.0%)	First level: 14 (12.6%)	<0.05
	Second level: 69 (40.1%)	Second level: 34 (56.6 %)	Second level: 35 (31.5 %)	
	Third level: 80 (46.5%)	Third level: 17 (28.4%)	Third level: 62 (55.9%)	
Ethnic group N. (%)	Caucasian: 157 (91.3%)	Caucasian: 53 (88.3 %)	Caucasian: 104 (92.8%)	N.S.
	Non-Caucasian: 15 (8.7%)	Non-Caucasian: 7 (11.7%)	Non-Caucasian: 8 (7.2%)	
Smoking habits N. (%)	Active 13 (7.5%)	Active 7 (11.6 %)	Active 6 (5.35 %)	N.S.
	Passive: 26 (15.1%)	Passive: 9 (15.0%)	Passive: 17 (14.2%)	
	Non-expose: 133 (77.4%)	Non-expose: 44 (73.3%)	Non-expose: 89 (79.5%)	

Table 5.5.: *General characteristics of control women samples enrolled in the project; N.S.: Not Significant, s.d.: standard deviation, p-value: probability value, N.: Number*

Parameters	Control sample	p-value
Number of samples	50	N.S.
Age (years) mean \pm s.d	28.8 \pm 4.83	N.S.
Height (m) mean \pm s.d	1.65 \pm 0.07	N.S.
Weight (kg) mean \pm s.d.	59 \pm 9.7	N.S.
BMI mean \pm s.d.	21.5 \pm 3.24	N.S.
Smoking habits N. (%)	Active: 6 (12%)	N.S.
	Passive: 12 (24%)	
	No exposure: 32 (64%)	

Table 5.6.: *General characteristics of babies enrolled in the project; N.S.: Not Significant, s.d.: standard deviation, p-value: probability value, N.: Number*

Parameters	Babies		p-value
	Boys	Girls	
Number of samples	99	73	N.S.
Length at birth (cm) Mean \pm s.d.	50.06 \pm 1.7	49.3 \pm 1.65	N.S.
Cranial Circumference at birth (cm) Mean \pm s.d.	34.4 \pm 1.10	33.9 \pm 1.10	N.S.
Weight at birth (kg) Mean \pm s.d.	3.34 \pm 0.39	3.24 \pm 0.41	N.S.
Breast feeding N. (%)	Yes: 81 (81.8%)	Yes: 60 (82.2%)	N.S.
	No: 18 (18.2%)	No: 13 (17.8%)	
Breast feeding + formula N. (%)	Yes: 69 (69.7%)	Yes: 54 (74.0%)	N.S.
	No: 30 (30.3)	No: 19 (26.0%)	
Pathologies at birth N. (%)	Yes: 19 (19.2%)	Yes: 21 (28.7%)	N.S.
	No: 80 (80.8%)	No: 52 (71.3%)	

Table 5.7.: *Urinary free and conjugate BPA/BPS values detected in the groups analysed; s.d.: standard deviation*

Group	Sub-group	BPA (ng/mg Creatinine) Mean (\pm s.d.)		BPS (ng/mg Creatinine) Mean (\pm s.d.)	
		Free	Conjugated	Free	Conjugated
Mothers	Patological	0.24 (\pm 0.42)	0.20 (\pm 0.41)	0.4 (\pm 0.18)	0.33 (\pm 0.19)
	Non Patological	0.23 (\pm 0.43)	0.19 (\pm 0.31)	0.36 (\pm 0.12)	0.17 (\pm 0.11)
	Total	0.24 (\pm 0.42)	0.20 (\pm 0.41)	0.38 (\pm 0.15)	0.25 (\pm 0.15)
	p-value			N.S.	
Control samples	50 samples	1.02 (\pm 1.38)	1.00 (\pm 1.37)	0.7 (\pm 0.14)	0.7 (\pm 0.15)
	p-value			N.S.	
Babies	Boys	0.24 (\pm 0.69)	0.63 (\pm 1.25)	0.17 (\pm 0.56)	0.15 (\pm 0.52)
	Girls	0.26 (\pm 0.52)	0.31 (\pm 0.92)	0.24 (\pm 0.10)	0.16 (\pm 0.44)
	Total	0.25 (\pm 0.61)	0.47 (\pm 1.09)	0.21 (\pm 0.33)	0.16 (\pm 0.48)
	p-value			N.S.	

analysed the BPA and BPS molecules. The method has been validated (see chapter 4 and the LLOQ values for BPA and BPS analytes are $0.5 \mu\text{g}/\text{kg}$ and $0.1 \mu\text{g}/\text{kg}$ respectively. If the values are compared with the results reported in table 5.7 the BPS free levels, BPA and BPS conjugated levels measured in the sampled subjects are close to the limit of detectability. Therefore, potentially misleading can occur.

Afterward regarding the exposure to endocrine disruptors, the urinary quantification of BPA and BPS in their free (active) and conjugated (inactive) forms did not put in evidence significant differences. The exposure is homogeneous among the mothers who have been hospitalized for around three days. The hospitalization, and so the impossibility of mothers to keep their personal and usual life habits, as well as food, make-up and environmental exposition, together with the short half-life (about 6h) of the considered analytes, allowed to "photograph" only the exposure of the last hours before sampling. That condition most likely does not coincide with their common exposure because the hospitalization interrupted and changed for some days the usual life-habits and, probably, also the exposure to bisphenols. On the contrary, in the control group, a typical situation closer to a conventional lifestyle was recorded, even if related to the last pre-sampling hours. The exposure appears a little higher although characterized by an efficient hepatic clearance since the levels of bisphenol (especially BPA) conjugated and then metabolized exceeds the share of the free and still active.

Next, the comparison between mothers BPA free levels and those of their babies does not highlight a statistically significant difference. Still, a little tendency among children with higher levels of BPA free appears. Differently, the urinary BPA-conjugate levels of the mothers show statistical differences ($p < 0.005$) if compared with those of their respective newborns. This aspect can be considered as a proof of how these contaminants can pass from the mother to the baby during the pregnancy. However, the difference between the urinary BPA-conjugate levels in mothers compared to their babies is significant. Newborns, given their still underdeveloped enzymatic system, are not able to metabolize toxic substances with the same speed and efficiency of adults.

Finally, if we focus only on the levels of BPA and BPS in the babies' group, there are no statistical differences between boys' and girls' sub-groups. This means that the child's sex doesn't influence the migration of these contaminants through the placenta. So, newborns are not significantly exposed to free BPA if compared to their mothers, although the maternal uterus, but especially the placenta, constitute a predisposing environment to greater exposure given the passage, although partially, through the umbilical cord [17].

5.5. Conclusion and future prospects

In conclusion, the developed UPLC-MRM MS method to quantify BPA and BPS in human samples was robust, selective, and sensitive. As already reported in the literature, bisphenol A and bisphenol S are plastic pollutants diffused everywhere with which the willing or unaware humankind must deal. Generally, BPA levels are higher than BPS ones, and even newborn babies, too, who are assumed to have never encounter plastic tools, show traces of these pollutants in their urine. So, thanks to the results obtained, it's clear how biological monitoring can provide new information for the evaluation of individual and population risk.

The critical future perspective of this work is to quantify BPA/BPS levels in breast milk (as described in 5.3, the analytical method is already definitely set up). Once the results are obtained, another important aim is to compare the BPA/BPS milk breast levels with the urine ones in order to clarify whether during breastfeeding period babies are exposed again to these oxidative stress exogenous sources or not.

References of Chapter 5

- [1] European Chemistry Agency. *Bisphenol S: Registration Data*. 2020. URL: <https://echa.europa.eu> (cit. on p. 58).
- [2] Liao C. et al. “Bisphenol S in urine from the United States and seven Asian countries: occurrence and human exposures”. In: *Environmental Science & Technology* 46, 6515-6522 (2012). DOI: 10.1021/es301334j (cit. on p. 58).
- [3] A.M. Calafat et al. “Exposure of the U.S. population to bisphenol A and 4-tertiary-octylphenol: 2003-2004”. In: *Environmental Health Perspectives* 116(1), 39–44 (2008) (cit. on p. 57).
- [4] A. Corti et al. “Agenti e meccanismi di stress ossidativo nella patologia umana”. In: *Ligand Assay* 14(1), 9-16 (2009) (cit. on p. 55).
- [5] Y. Deceuninck et al. “Determination of bisphenol A and related substitutes/analogues in human breast milk using gas chromatography-tandem mass spectrometry”. In: *Analytical and Bioanalytical Chemistry* 407(9), 2485-2497 (2015). DOI: 10.1007/s00216-015-8469-9 (cit. on p. 56).
- [6] W. Dekant et al. “Human exposure to bisphenol A by biomonitoring: methods, results and assessment of environmental exposures.” In: *Toxicology and Applied Pharmacology* 228(1), 114-134 (2008). DOI: 10.1016/j.taap.2007.12.008. (cit. on p. 57).
- [7] E. Diamanti-Kandarakis et al. “Endocrine-disrupting chemicals: an Endocrine Society scientific statement”. In: *Endocrine Reviews* 30, 293-342 (2009). DOI: 10.1210/er.2009-0002 (cit. on p. 57).
- [8] M. Ejaredar et al. “Bisphenol A exposure and children’s behavior: A systematic review”. In: *Journal of Exposure Science and Environmental Epidemiology* 27(2), 175-183 (2017). DOI: 10.1038/jes.2016.8 (cit. on p. 55).
- [9] Gauderat G. et al. “Prediction of human prenatal exposure to bisphenol A and bisphenol A glucuronide from an ovine semi-physiological toxicokinetic model”. In: *Scientific Reports - Nature* 7(1), 15330 (2017). DOI: 10.1038/s41598-017-15646-5 (cit. on p. 59).

-
- [10] T. Geens et al. “Assessment of human exposure to Bisphenol-A, Triclosan and Tetrabromobisphenol-A through indoor dust intake in Belgium”. In: *Chemosphere* 76(6), 755-760 (2009). DOI: 10.1016/j.chemosphere.2009.05.024 (cit. on pp. 57, 58).
- [11] G. Ginsberg et al. “Does rapid metabolism ensure negligible risk from bisphenol A?” In: *Environmental Health Perspectives* 117(11), 1639-1643 (2009). DOI: 10.1289/ehp.0901010 (cit. on p. 58).
- [12] K.L. Ho et al. “Glucuronide and Sulfate Conjugates of Bisphenol A: Chemical Synthesis and Correlation Between Their Urinary Levels and Plasma Bisphenol A Content in Voluntary Human Donors”. In: *Archives of Environmental Contamination and Toxicology* 73, 410-420 (2017). DOI: 10.1007/s00244-017-0438-1 (cit. on p. 60).
- [13] Y.Q. Huang et al. “Bisphenol A (BPA) in China: a review of sources, environmental levels, and potential human health impacts”. In: *Environment International* 42, 91–99 (2012). DOI: 10.1016/j.envint.2011.04.010 (cit. on p. 57).
- [14] T. Iso et al. “DNA damage caused by bisphenol A and estradiol through estrogenic activity”. In: *Biological and Pharmaceutical Bulletin* 29(2), 206-210 (2006). DOI: 10.1248/bpb.29.206 (cit. on p. 57).
- [15] Glausiusz J. “Toxicology: The plastics puzzle”. In: *Nature* 508(7496), 306-308 (2014) (cit. on p. 58).
- [16] S. Jenkins et al. “Oral exposure to bisphenol a increases dimethylbenzanthracene induced mammary cancer in rats”. In: *Environmental Health Perspectives* 117(6), 910–915 (2009). DOI: 10.1289/ehp.11751 (cit. on pp. 57, 58).
- [17] K.K. Jung et al. “Differential regulation of thyroid hormone receptor-mediated function by endocrine disruptors”. In: *Archives of Pharmacal Research* 30(5), 616-623 (2007). DOI: 10.1007/bf02977657 (cit. on pp. 56, 59, 68).
- [18] J.S. Lakind et al. “Temporal trend in Bisphenol A exposure in the UNited States from 2000 - 2012 and factors associated with BPA exposure: spot samples and urine dilution complicate data interpreta-

- tion”. In: *Environmental Research* 142, 84–95 (2015). DOI: 10.1016/j.envres.2015.06.013 (cit. on p. 57).
- [19] X. Li et al. “4-Nonylphenol, bisphenol-A and triclosan levels in human urine of children and students in China, and the effects of drinking these bottled materials on the levels”. In: *Environmental International* 52, 81-86 (2013). DOI: 10.1016/j.envint.2011.03.026 (cit. on p. 57).
- [20] C. Liao et al. “Bisphenol analogues in sediments from industrialized areas in the United States, Japan, and Korea: spatial and temporal distributions”. In: *Environmental Science & Technology* 46(21), 11558-65 (2012). DOI: 10.1021/es303191g (cit. on p. 58).
- [21] C. Liao et al. “Determination of free and conjugated forms of bisphenol A in human urine and serum by liquid chromatography-tandem mass spectrometry”. In: *Environmental Science & Technology* 46(9), 5003-9 (2012). DOI: 10.1021/es300115a (cit. on p. 57).
- [22] C. Liu et al. “Exposure to bisphenol A disrupts meiotic progression during spermatogenesis in adult rats through estrogen-like activity”. In: *Cell Death & Disease* 4:e676 (2013). DOI: 10.1038/cddis.2013.203. (cit. on p. 57).
- [23] Y. Masuo. “Neurotoxicity of endocrine disruptors: possible involvement in brain development and neurodegeneration”. In: *Journal of Toxicology and Environmental Health, Part B* 14(5-7), 346-369 (2013). DOI: 10.1080/10937404.2011.578557 (cit. on p. 57).
- [24] K. Mendonca et al. “Bisphenol A concentrations in maternal breast milk and infant urine”. In: *International Archives of Occupational and Environmental Health* 87(1), 13-20 (2014). DOI: 10.1007/s00420-012-0834-9 (cit. on p. 57).
- [25] S-H. Nam et al. “Bisphenol A migration from polycarbonate baby bottle with repeated use”. In: *Chemosphere* 79(9), 949-952 (2010). DOI: 10.1016/j.chemosphere.2010.02.049 (cit. on p. 57).

-
- [26] A. Phaniendra et al. “Free Radicals: Properties, Sources, Targets, and Their Implication in Various Diseases”. In: *Indian Journal of Clinical Biochemistry* 30(1), 11–26 (2015). DOI: 10.1007/s12291-014-0446-0 (cit. on p. 55).
- [27] C. Pirard et al. “Exposure of the U.S. population to bisphenol A and 4-tertiary-octylphenol: 2003-2004”. In: *Environmental International* 48, 78–83 (2012). DOI: 10.1016/j.envint.2012.07.003 (cit. on p. 57).
- [28] J.R. Rochester et al. “Bisphenol S and F: a systematic review and comparison of the hormonal activity of bisphenol A substitute”. In: *Environmental Health Perspectives* 123, 643-650 (2015). DOI: 10.1289/ehp.1408989 (cit. on p. 58).
- [29] G. Siciliano et al. “Antioxidant capacity and protein oxidation in cerebrospinal fluid of amyotrophic lateral sclerosis”. In: *Journal of Neurology* 254(5), 575-80 (2017). DOI: 10.1007/s00415-006-0301-1 (cit. on p. 55).
- [30] P. Sompol et al. “A neuronal model of Alzheimer’s disease: An insight into the mechanisms of oxidative stress-mediated mitochondrial injuries”. In: *Neuroscience* 153(1), 120-130 (2008). DOI: 10.1016/j.neuroscience.2008.01.044 (cit. on p. 55).
- [31] S. Srivastava et al. “Bisphenol A: a threat to human health?” In: *Journal of Environmental Health* 77(6), 20–26 (2015) (cit. on p. 56).
- [32] L.N. Vandenberg et al. “A round robin approach to the analysis of bisphenol A (BPA) in human blood samples”. In: *Environmental Health* 13(1), 25 (2014). DOI: 10.1186/1476-069X-13-25. (cit. on pp. 57, 58).
- [33] L.N. Vandenberg et al. “Urinary, circulating, and tissue biomonitoring studies indicate widespread exposure to bisphenol A”. In: *Environmental Health Perspectives* 118(8), 1055-70 (2010). DOI: 10.1289/ehp.0901716 (cit. on p. 56).
- [34] Y.B. Wetherill et al. “Xenoestrogen action in prostate cancer: pleiotropic effects dependent on androgen receptor status”. In: *Cancer Research* 65(1), 54-65 (2005) (cit. on p. 57).

- [35] N.K. Wilson et al. “An observational study of the potential exposures of preschool children to pentachlorophenol, bisphenol-A, and nonylphenol at home and daycare”. In: *Environmental Research* 103(1), 9-20 (2007) (cit. on p. 57).
- [36] L.H. Wu et al. “Occurrence of bisphenol S in the environment and implications for human exposure: A short review”. In: *Science of the Total Environment* 615, 87-98 (2018). DOI: 10.1016/j.scitotenv.2017.09.194 (cit. on p. 59).
- [37] A. Ziv-Gal et al. “Bisphenol A inhibits cultured mouse ovarian follicle growth partially via the aryl hydrocarbon receptor signaling pathway”. In: *Reproductive Toxicology* 42, 58-67 (2013). DOI: 10.1016/j.reprotox.2013.07.022 (cit. on p. 57).

Targeted analysis: Uremic toxins molecules

6.1. Introduction

People suffering from Chronic Kidney Diseases (CKD) have a progressive and irreversible loss of renal function over time [14]. This pathology is primarily caused by obesity, followed by oldness and secondary symptoms of other diseases (diabetes or hypertension) to [1, 14]. The CKD, therefore, implies that a wide variety of solutes, usually extracted by the organism through the urine, gradually accumulates in the blood over time. In case of correlation between the accumulation of these solutes and the onset of clinically recognizable disorders, then it is possible to attribute to these molecules the name uremic toxins [14].

The first step in treating the CKD is to identify the causes that determined this state and, if it's possible, remove them. However, chronic renal failure is, unfortunately, a progressive disease, and the damage proceeds independently of the original causes that led to it [13]. For this reason, a continuous removal treatment over time of these metabolites is essential for CKD patients who have a compromised renal function [4]. These treatments consist of repeated dialysis therapy over time or, in severe cases, renal transplantation.

Dialysis is a therapy whose objectives are the elimination of waste products, the removal of excess fluids, and the balance of electrolytes and nutrients in the body [2]. In this process, the fundamental role is interpreted by a semipermeable membrane (the filter), which purifies the blood from all waste substances. Although dialysis is now an established technique that shall be subject to monitoring, it is possible that during this treatment, the patients' contract infections in which the symptoms are not necessarily shown in a short time [7].

In this work, I developed two different methods for the determination of uremic toxins in nephrological patients' plasma. Using a HPLC-HRMS I carried out a qualitative analysis to understand which uremic toxins are present in plasma. On the contrary, using the UPLC-TQ MS I created a quantitative and fast method for routine analysis. The aims of this work were firstly the development of HPLC-HRMS and UPLC-TQ MS methods for the determination and quantification of p-cresyl sulfate and indoxyl sulfate uremic toxins in plasma blood and secondly the determination of other uremic toxins by untargeted analysis exploiting the HPLC-HRMS validated method.

Uremic toxins identities' are well known nowadays [10, 20]. These solutes, which under normal conditions are excreted by the healthy kidney, possess heterogeneous properties. First at all, they belong from different chemical groups: ribonucleotides, guanidines, polyols, purines, etc. Secondly, they have various molecular weights: from hundreds of Dalton to ten thousands Dalton. Finally, their normal concentrations in blood could range from 32.0 ng/L (methionine-enkephalin) up to 2.3 g/L (urea). An example of some of the most studied uremic toxins molecules are reported in figure 6.1.

To understand which mechanisms link uremic toxins and CKD patients, we can focus our attention on the two most monitored molecules in this context: indoxyl sulfate and p-cresyl sulfate. It has been demonstrated that indoxyl sulfate serum levels increase with declining kidney function and this elevation contributes to the progression of kidney injury [8, 9]. Various pathophysiological mechanisms are involved in this process. An experimental study using a rat model of CKD demonstrated that the administration of indoxyl sulfate accelerates glomerular sclerosis, which could be considered the most important mechanism of the kidney injury progression [15]. Furthermore, this uremic toxin plays an important role in the pathophysiology of acute kidney injury by influencing the NO-dependent pathway and decreasing the number of endothelial progenitor cells (EPCs) which delays recovery from kidney injury [9, 23].

Looking at p-cresyl sulfate, as well as indoxyl sulfate, has adverse effects upon the kidney. The experimental study proposed by Sung and colleagues

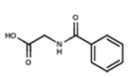
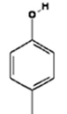
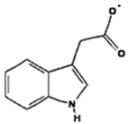
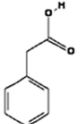
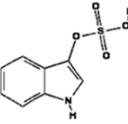
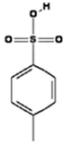
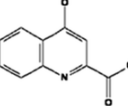
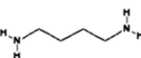
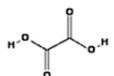
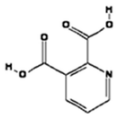
Solute	Structure	Group	Solute	Structure	Group
hippuric acid		hippurates	p-cresol		phenols
indole-3-acetic acid		indoles	phenylacetic acid		phenols
indoxyl sulfate		indoles	p-toluensulfonic acid		phenols
kynurenic acid		indoles	putrescine		polyamines
oxalate			quinolinic acid		indoles

Figure 6.1.: Schematic example of few molecules belonging to uremic toxins' group

[17] showed that p-cresyl sulfate induces significant cellular immune and inflammatory responses including activation of the TGF- β signalling pathway. Furthermore, in [18], a higher level of p-cresyl sulfate in bloodstream activates renin-angiotensin system¹ which, in turns, increases the amount of oxidative stress by stimulating leukocytes response. All these changes lead to the same goal: progression of kidney fibrosis.

6.2. Aim of the work

The following bullet points can resume the thesis work:

- The main purpose of this work is the development, in collaboration with a clinical research group, of UPLC-TQ MS (mass spectrometer: QTRAP[®] 5500 system) and HPLC-HRMS (mass spectrometer: LTQ-Orbitrap[™]) targeted metabolomics methods able to detect and quantify, in human plasma samples, uremic toxins responsible for vascular

¹Renin-angiotensin system, commonly known as RAS, has been known for more than a century as a cascade that regulates body fluid balance and blood pressure.

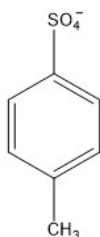


Figure 6.2.: *Chemical structures for p-cresyl sulfate*

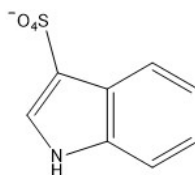


Figure 6.3.: *Chemical structures for indoxyl sulfate*

problems and high mortality in CKD patients [3, 12]. The uremic toxins of interest from a medical point of view are p-cresyl sulfate (figure 6.2) and indoxyl sulfate (figure 6.3).

- A second objective consists in comparing the results obtained with the different instrumental configurations used. The aim is the identification of their strengths and weaknesses in the analysis of these molecules.
- Patients performed dialysis using, in separated period of time, filters with different characteristics (Toray[®] filter and Filtryzer[®] filter [19]). The therapy consisted of three weekly dialysis sessions of three hours each using a filter, then stop in the treatment for a week, and finally three weekly dialysis sessions again using the other filter. Therefore, a further aim of the thesis is to verify if there are statistically significant differences in the performance of one or the other filter in the removal of p-cresyl sulfate and indoxyl sulfate from the bloodstream.
- As previously reported (section 3.3.1), thanks to the LTQ-Orbitrap[™] platform is possible to conduct untargeted metabolomics analysis. A final purpose of this work is to monitor, in the same extracted samples, the presence of other kinds of uremic toxins, through the same method developed for p-cresyl and indoxyl sulfate.

6.3. Sample preparation and method settings

CKD plasma samples are analysed in this work. As previously reported (section 6.2) the purpose of the analysis is to understand if the use of a polysulfone filter rather than a PMMA filter could affect the elimination of p-cresyl sulfate and indoxyl sulfate from the blood of patients under dialysis therapy.

The matrices studied are plasma samples of dialysis patients A.S.L. City of Health and Science of Turin and CN1 of Savigliano. The patients are involved in the study based on different criteria such as the adult age, the non-hospitalization of the patients during treatment, the absence of further pathologies beyond the CKD, and, obviously, the consent of the patients to the treatment of their blood for scientific purposes.

Each sample is characterized by a combination of letters and numbers. To indicate the day in which it was obtained, the letter D (Day) is used followed by a number between 1 and 4. This number indicates the different collection days according to the following bullet point:

- D1: first week, withdrawal made on Monday;
- D2: first week, withdrawal made on Friday. Wednesday of first week patients undergo dialysis without withdrawing a blood sample for analysis;
- D3: second week, withdrawal made on Tuesday;
- D4: second week, withdrawal made on Saturday. Thursday of the second week patients are subjected to dialysis without withdrawing a blood sample for analysis.

On the contrary, to indicate the time at which the sample is obtained during the dialysis session, the letter N is used followed by a number between 1 and 7. The meaning of numbers paired to N is now reported:

- N1: Sample of serum taken pre-dialysis;

- N2: Sample taken from arterial access one hour after the start of dialysis;
- N3: Sample taken from venous access one hour after the start of dialysis;
- N4: Sample taken from arterial access two hours after the start of dialysis;
- N5: Sample taken from venous access two hours after the start of dialysis;
- N6: Sample taken from arterial access four hours after the start of dialysis (end of dialysis);
- N7: Sample taken from venous access four hours after the start of dialysis (end of dialysis).

As soon as the blood samples are collected in the hospital, they are treated to obtain the plasma. The standard protocol for the collection and the separation of plasma from blood was performed by trained personal in the hospital environment, and it is as follows:

1. Withdraw 30 mL of peripheral blood using a syringe *vacuette*®.
2. Transfer the blood into a special 50 mL tube containing anticoagulant (for 30 mL of whole blood, 300 μ L EDTA 0.5 M in aqueous solution pH 8 are necessary). Avoid abnormal movements of the tubes to reduce cell lysis to a minimum.
3. Centrifuge the sample 500 g (Scilogex D3024R[®] centrifuge), 4° C 10 minutes.
4. Transfer the supernatant containing the plasma into a 15 mL Falcon tube, avoiding the closest portion to the lymphocyte ring.
5. Centrifuge again the sample 500 g, 4° C 10 minutes.
6. Withdraw the supernatant containing the plasma and aliquot it, if necessary, freeze at -80° C.

Table 6.1.: *Instrumental equipment used in the HPLC-HRMS system*

HPLC	Chromatographic column Phenomenex Luna 3 μm C18(2) 100 \AA , LC Column 100 x 4.6 mm	Ultimate 3000 Dionex
HRMS		LTQ Orbitrap TM Thermo Scientific
Ionization source and parameters		ESI Capillary temperature 250° C Flow rate sheath gas 30.0 arbitrary units (a.u.) Flow rate auxiliary gas 15.0 arbitrary units (a.u.) Capillary voltage 8.5 V Source voltage 4.5 kV Tube lens 60 V

Then, once the plasma samples are ready to be analysed arrived, they must be treated using the following protocol steps:

1. Thaw the sample.
2. Once they are liquid, vortex the plasma samples for 30 s to uniform the matrix.
3. Withdraw 100 μL and place it in a 3 mL tube.
4. Add 300 m μL of cold methanol.
5. Centrifuge the sample 8000 g, 4° C 10 minutes.
6. Treated samples are then diluted according to the instrument used for the analysis. This difference in treatment is due to the different sensitivity of the two instruments used.

One of the aims is to verify the difference in sensitivity that two analytical platforms have. Furthermore, we want to understand which are the pros and cons of using one tool rather than the other for this specific analysis. In the tables 6.1, 6.2, 6.3, 6.4, and 6.5 the different instrumental parameters and the different chromatographic conditions used for LTQ OrbitrapTM and for QTRAP[®] 5500 are reported. All these parameters have been optimized to have the maximum analytical sensitivity in the separation and detection of the analytes of interest.

6.4. Results and discussion

Once the methods have been validated for the two MS instruments (see table 6.6), the plasma samples of the patients are analysed. Chromatogram

Table 6.2.: *Chromatographic and MS conditions - LTQ Orbitrap™ Thermo Scientific*

Injection	20 μ L		
Flux	200 μ L/min		
	Time (min)	CH ₃ COONH ₄ 5 mmol in water	CH ₃ OH
Chromatographic run	0	95 %	5 %
	18	50 %	50 %
	28	5 %	95 %
	30	5 %	95 %
	31	95 %	5 %
	45	95 %	5 %
PDA: UV - Vis	Start of acquisition wavelength = 200 nm End of acquisition wavelength = 600 nm		
Polarity	Negative		
Mass range	100 - 1000 m/z		
Resolution	30000 (FWHM)		

Table 6.3.: *Instrumental equipment used in the UPLC-TQ MS platform*

UPLC	Shimadzu Nexera X2 UPLC Column Phenomenex 1.7 μ m Gemini C18 100 Å, LC Column 50 x 2.1 mm
TQMS	SCIEX Qtrap® 5500 System
Ionization source	ESI (Turbo ion spray©)

Table 6.4.: *Chromatographic conditions - SCIEX QTRAP® 5500*

Injection	3 μ L		
Flux	300 μ L/min		
Oven temperature	40° C		
	Time (min)	CH ₃ COONH ₄ 5 mmol in water	CH ₃ OH
Chromatographic run	0	95%	5%
	4	90%	10%
	6	0%	100%
	7	95%	5%
	15	95%	5%

Table 6.5.: *MS parameters - SCIEX QTRAP® 5500*

Polarity	Negative						
Source temperature	400° C						
GS1; GS2	50; 50						
Dwell time	70 for indoxyl sulfate, 180 for all the other transitions						
	Molecular ion (<i>m/z</i>)	Fragmented ion (<i>m/z</i>)	Retention time (min)	DP	EP	CE	CXP
p-cresyl sulfate	187.0	107.0	6.2	-45	-8	-30	-10
		80.0		-45	-8	-32	-8
p-cresyl sulfate-d7	194.0	114 .0	6.1	-63	-3	-26	-6
		80.0		-63	-3	-21	-11
indoxyl sulfate	212.0	132.0	3.3	-37	-6	-24	-8
		80.0		-37	-6	-25	-10
indoxyl sulfate-13C	218.0	138.0	3.3	-72	-7	-80	-12
		80.0		-72	-7	-24	-25

Table 6.6.: *Schematic report of parameters obtained during validation procedures for uremic toxins analyses*

	Indoxyl sulfate			p-cresyl sulfate		
	Validation	Theoretical value	Experimental value	Validation	Theoretical value	Experimental value
LTQ - Orbitrap®	LOD	✗	20 µg/Kg	LOD	✗	25 µg/Kg
	LOQ	✗	70 µg/Kg	LOQ	✗	90 µg/Kg
	LLOQ	✗	300 µg/Kg	LLOQ	✗	300 µg/Kg
	Selectivity, %	<30.00	1.77	Selectivity, %	<30.00	3.70
	Linearity, %	<25.00	4.84	Linearity, %	<25.00	4.95
	Relative Standard Deviation in LLOQ, %	<25.00	23.80	Relative Standard Deviation, %	<25.00	22.20
QTRAP® 5500	Validation	Theoretical value	Experimental value	Validation	Theoretical value	Experimental value
	LOD	✗	0.3 µg/Kg	LOD	✗	0.03 µg/Kg
	LOQ	✗	1 µg/Kg	LOQ	✗	0.1 µg/Kg
	LLOQ	✗	20 µg/Kg	LLOQ	✗	0.5 µg/Kg
	Selectivity, %	<30.00	8.39	Selectivity, %	<30.00	8.30
	Linearity, %	<25.00	16.56	Linearity, %	<25.00	18.10
Relative Standard Deviation in LLOQ, %	<25.00	7.91	Relative Standard Deviation, %	<25.00	4.97	

reported in figure 6.4 is an example of how a MRM chromatographic run looks like.

Tables 6.7, 6.8, 6.9 and 6.10 report the p-cresyl sulfate and indoxyl sulfate concentrations during the various days of dialysis. Based on the values shown in these tables, some interesting considerations can be made. Firstly, it is noted that the physiological levels of p-cresyl sulfate and indoxyl sulfate decrease over two weeks thanks to the treatment. Furthermore, comparing the values of D2N7 with those of D3N1 samples, it is clear how a stop in the treatment of only four days causes an increase in the levels of these uremic toxins up to values, in certain cases, near or even higher than those obtained in the pre-dialysis analysis.

Secondly, thanks to Pearson's nonparametric χ^2 test it is demonstrated that the hypothesis H_0 formulated for this test (no statistically significant difference between the populations of data taken into consideration) is ac-

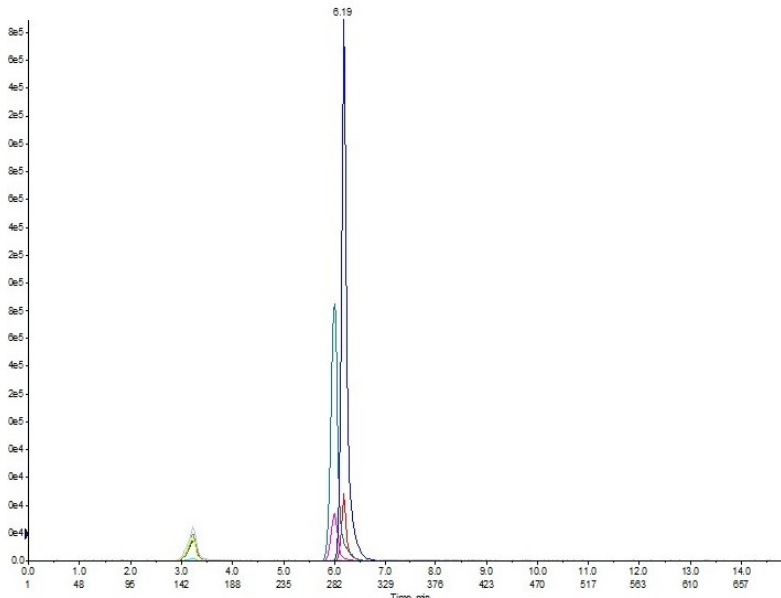


Figure 6.4.: *Result of the chromatographic run related to uremic toxins*

ceptable².

The results provided by the two tools used are not statistically different. The use of the LTQ-OrbitrapTM or QTRAP[®] 5500 to carry out this type of analysis will depend, therefore, on what the operator wants to obtain: if you wish to gain mass spectra in high resolution you must necessarily refer to the LTQ-Orbitrap[®]; on the contrary, if you wish to have high sensitivity and you want to perform routine analysis in a short time (thanks to the UPLC configuration) it is preferable to use the QTRAP[®] platform. Finally, using the Student t parametric test for the comparison between the various data, it emerges that also in this case the hypothesis H₀ formulated (the differences in the results obtained with the two filters are attributable to the case) is valid. Therefore, the use of the Toray[®] filter or the Filtryzer[®] filter does

²The mathematical equation for χ^2 test is:

$$\chi^2 = \sum_{i=1}^n \frac{(e_i - o_i)^2}{e_i} \quad (6.1)$$

where e_i is the tabulated value, o_i is the experimental value and n is the number of the samples. Once the α values is fixed (in this case, $\alpha=0.05$, degrees of freedom = 3) if χ^2 is higher than the tabulated value hypothesis H_1 is accepted while rejecting hypothesis H_0 .

not significantly affect the removal of these toxins from the bloodstream.

To simplify the visualization of the results obtained by this targeted metabolomics analysis, refer to the histograms reported in figures 6.5 and 6.6. In these graphs on the x-axis are reported the different patients reported for privacy as a number or a letter: from 1 to 6 for the illness ones, A and B for control and healthy patients. On the y-axis are reported the four samples taken into consideration: D1N1 (first day of the first week, first sample of the day), D2N7 (Last day of the first week, last sample of that day), D3N1 (first day of the second week, first sample of the day), D4N7 (Last day of the second week, last sample of that day). Finally, on z-axis are reported the concentration levels for the specific uremic toxin.

An aspect to clarify is the use of internal standards during analysis. For the quantification of these uremic toxins, stable isotopes labelled internal standards are used to compensate matrix effects. Two different internal standards have been used: the p-cresyl sulfate potassium salt [D7] standard used to compensate the effects of the matrix for p-cresyl sulfate and the [13C6]-Indoxyl sulfate potassium salt standard for indoxyl sulfate. These substances can offset matrix effects, assuming that the internal standard (a compound very similar but not identical to the chemical species of interest) suffers same loss. The ability of a stable isotope-labelled internal standard to compensate for matrix effects is strictly connected to the co-elution with the unlabelled compound³. For this reason, the internal standard is a compound that must behave most similarly concerning the chemical species of interest in the samples, given that the effects of sample preparation must be as similar as possible to the signal of the internal standard and of the analysis in the sample. Hence the reason for the use of this type of standard: adding known quantities of analyte(s) of interest is a distinct technique called standard addition, which is performed to correct for matrix effects. Finally, the relationship between the two signals obtained (analyte vs. internal standard) is then used to calculate the final concentration of the analyte

³This doesn't always happen due to the deuterium isotope effect [6, 16, 21]. This effect is thought to be caused by the change in lipophilicity of the molecule when one or more hydrogens are replaced with deuterium. Different articles in literature [11, 21, 22] observed this aspect upon different matrices such as plasma or urine

in the sample from the calibration curve. Tables for isotopes labelled internal standards are not reported because the parameters used for their analysis are quite similar: in the case of Orbitrap system, the parameters are the same; in the other case, the triple quadrupole platform, the only difference is in the m/z of ions monitored.

Once the targeted metabolomics analysis of p-cresyl sulfate and indoxyl sulfate has been completed, an untargeted metabolomics analysis of other uremic toxins listed in the literature [4, 14] is carried out upon the samples analysed via HPLC-HRMS platform. The choice of which uremic toxins to seek was made based on the levels reported in the literature [14]. Untargeted metabolomics analysis shows the extraction of the HRMS ionic signals for the various uremic toxins monitored. As we can see, for some of them (for example, molecules with m/z 265.1163), the chromatographic peaks extracted in high resolution seem to confirm the presence. The most promising peaks to verify using standards, except indoxyl sulfate and p-cresyl sulfate, are for both polarities: m/z 265.1163 (phenylacetylglutamine), m/z 239.0914 (3-Carboxy-4-Methyl-5-Propyl-2-Furanpropionic Acid, CMPF), m/z 178.0499 (hippuric acid), *136.0427* (homocysteine), and an unknown compound $C_8H_{15}N_4O_2$ m/z 199.0311.

It is an important goal: by using a single extraction procedure, different types of uremic toxins can be detected. If nowadays only p-cresyl sulfate and indoxyl have a medical interest, in the future, the possibility to obtain information upon other uremic toxins levels in samples already studied will be amazing. For example, nowadays CMPF is markedly accumulated in the serum of CKD patients, but cannot be removed at the common haemodialysis approaches because of its strong albumin binding (this is probably the reason why its chromatographic peak is so well detectable). On the contrary, hippuric acid is easily removed by HD, with a 60% decrease of the free fraction [5, 14].

However, despite the use of MS2 and MS3 spectra and the use of online databases, the certainty of the presence of these toxins can only be obtained after having analysed, with the same instrument and in the same instrumental conditions, a certified standard of that specific analyte.

Table 6.7.: Concentration of *p*-cresyl sulfate during therapy, HPLC-HRMS

Sample	<i>p</i> -cresyl D1N1 mg/L	<i>p</i> -cresyl D2N7 mg/L	<i>p</i> -cresyl D3N1 mg/L	<i>p</i> -cresyl D4N7 mg/L
Patient 1	64.05	51.74	86.72	43.36
Patient 2	148.16	73.62	174.71	111.87
Patient 3	51.39	24.10	68.69	17.55
Patient 4	71.95	33.92	118.17	63.10
Patient 5	109.25	70.74	99.30	45.03
Patient 6	54.28	24.14	31.72	15.52
Control A	5.53	✗	✗	✗
Control B	6.80	✗	✗	✗

Table 6.8.: Concentration of indoxyl sulfate during therapy, HPLC-HRMS

Sample	Indoxyl sulfate D1N1 mg/L	Indoxyl sulfate D2N7 mg/L	Indoxyl sulfate D3N1 mg/L	Indoxyl sulfate D4N7 mg/L
Patient 1	40.13	13.37	53.51	26.16
Patient 2	39.83	13.08	49.54	25.87
Patient 3	20.28	8.38	28.66	11.76
Patient 4	28.24	12.28	33.32	22.92
Patient 5	42.92	34.25	47.92	37.78
Patient 6	45.78	29.36	40.55	20.51
Control A	1.58	✗	✗	✗
Control B	3.75	✗	✗	✗

As it is clear from figure 6.7, it is evident that some of the expected analytes could not be detected by this approach. This aspect paves the way for future implementation at work, i.e. the creation of an adequate panel composed only by LTQ-detectable analytes, thus converting the initial untargeted approach into a semi-targeted one.

6.5. Conclusion and future prospects

In conclusion, retracing the various objectives set in section 6.2, the results obtained in the present study are shown below.

Table 6.9.: Concentration of *p*-cresyl sulfate during therapy, UPLC-TQ MS

Sample	<i>p</i> -cresyl D1N1 mg/L	<i>p</i> -cresyl D2N7mg/L	<i>p</i> -cresyl D3N1 mg/L	<i>p</i> -cresyl D4N7 mg/L
Patient 1	61.96	52.40	83.07	38.87
Patient 2	130.55	72.99	158.91	113.53
Patient 3	61.75	20.01	77.86	21.50
Patient 4	66.69	27.17	108.40	52.76
Patient 5	102.34	62.73	85.66	52.10
Patient 6	56.85	27.01	34.52	15.87
Control A	5.13	✗	✗	✗
Control B	7.08	✗	✗	✗

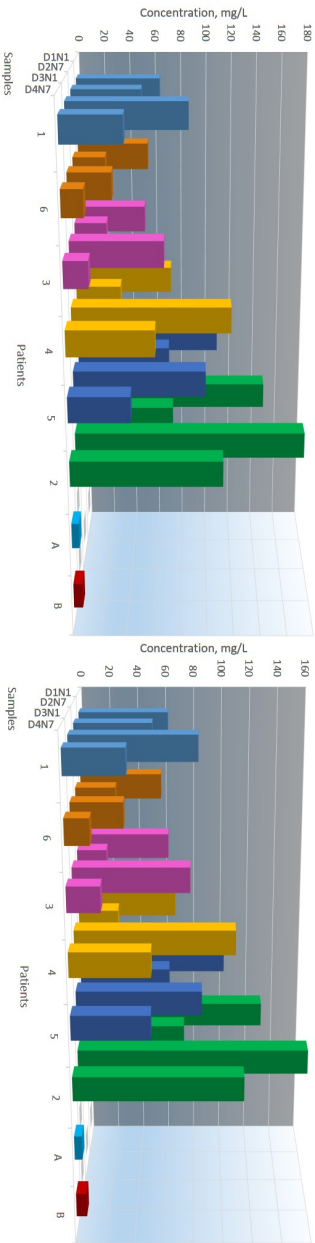


Figure 6.5.: *p*-cresyl sulfate levels over samples (1-6). On the left LTQ-Orbitrap™ platform results, on the right QTRAP® 5500 results

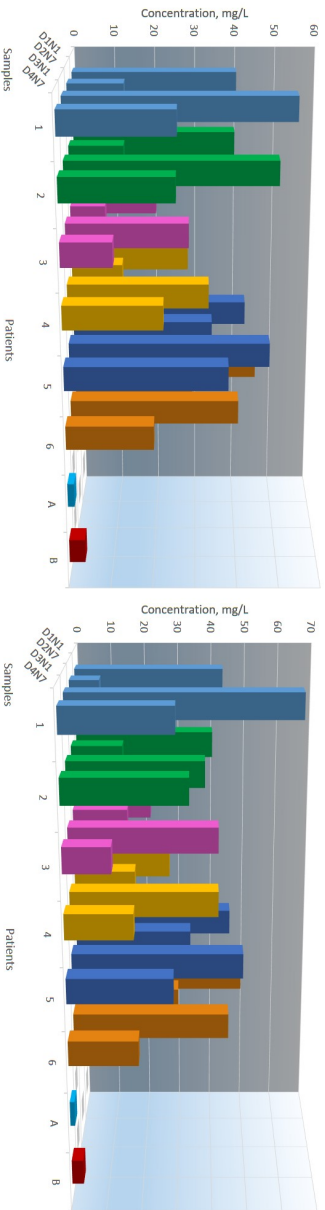


Figure 6.6.: *p*-cresyl sulfate levels over samples (1-6). On the left LTQ-Orbitrap™ platform results, on the right QTRAP® 5500 results

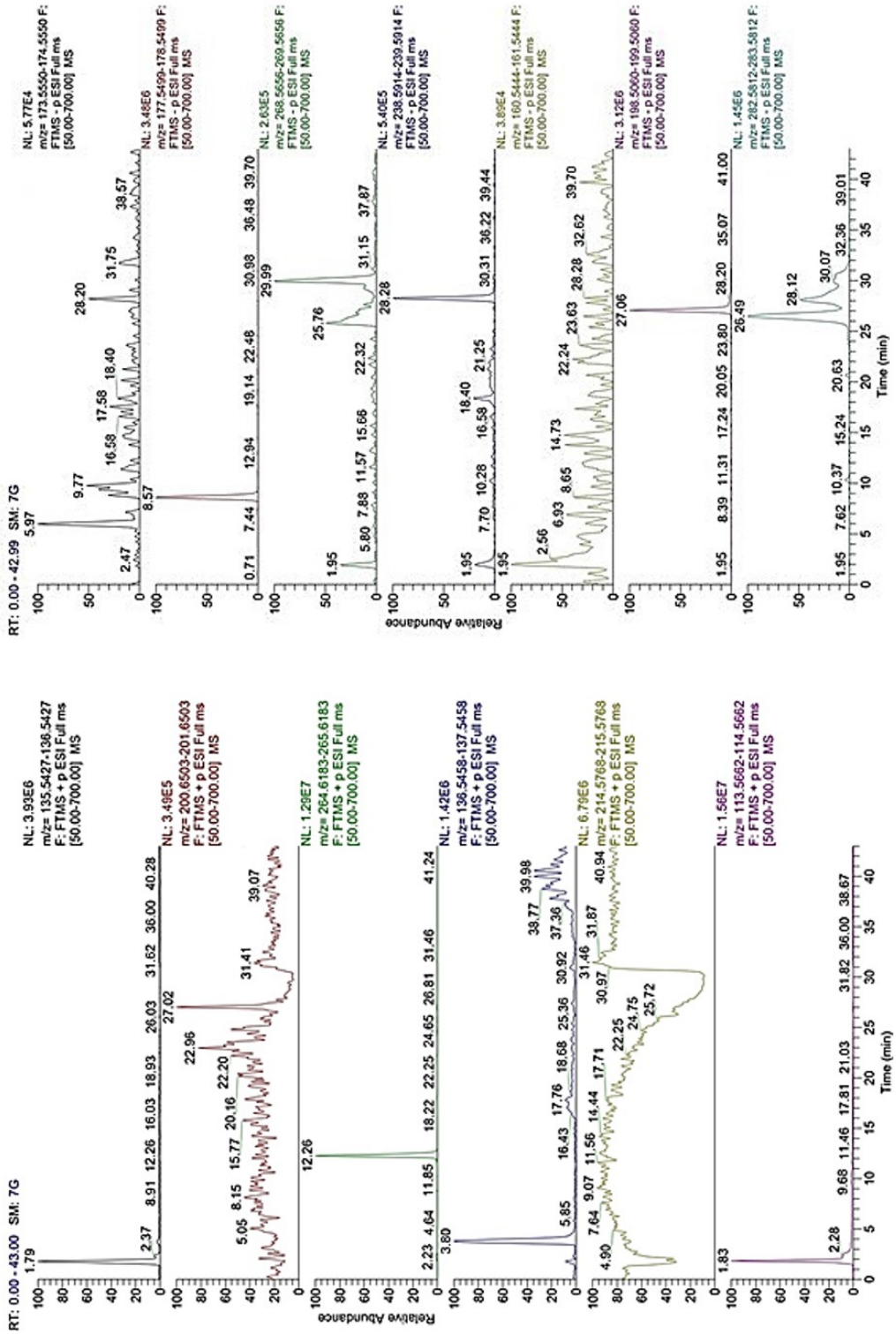


Figure 6.7.: HRMS ionic signals extracted for different uremic toxins monitored by LTQ-Orbitrap™ platform

Table 6.10.: *Concentration of indoxyl sulfate during therapy, UPLC-TQ MS*

Sample	Indoxyl sulfate D1N1 mg/L	Indoxyl sulfate D2N7 mg/L	Indoxyl sulfate D3N1 mg/L	Indoxyl sulfate D4N7 mg/L
Patient 1	43.22	8.40	64.68	30.12
Patient 2	40.28	14.56	37.92	33.66
Patient 3	21.75	15.64	41.74	13.05
Patient 4	27.37	17.56	41.82	18.70
Patient 5	45.90	33.80	48.96	29.31
Patient 6	49.57	30.11	44.84	19.56
Control A	1.31	✗	✗	✗
Control B	3.29	✗	✗	✗

1. The LTQ-OrbitrapTM and QTRAP[®] 5550 platforms employed in this work provide comparable and no statistically different results. The reason to prefer LTQ-OrbitrapTM or QTRAP[®] 5550 depends on some factors. First, the difference between sensitivities. In general, QTRAP[®] systems (and all triple quadrupole platform) allows the quantification of analytes with higher sensitivity than the LTQ-OrbitrapTM.

The use of QTRAP[®] 5550 is therefore not essential for these patients, given the high concentrations of toxins present in their plasma samples. However, thanks to the coupling to a UPLC system, the time required for a single analysis is lower compared to HPLC-HRMS time analysis. So, the management costs of the instrument are considerably contained. Given that, this analytical platform is handy in the case of routine analysis. On the contrary, the use of LTQ-OrbitrapTM is essential if you want to get it MS spectra with high accuracy through high resolution, for discriminate between molecules with similar mass (resolution 30000 at 500 Da without internal calibration). This tool is fundamental in the case of untargeted metabolomics analysis. Time machine and management costs, compared to those of the UPLC-TQ MS, are higher.

2. Toray[®] and Filtryzer[®] filters do not show statistically significant differences during dialysis treatment. The use of one rather than the other filter must be evaluated primarily from a medical point of view rather than a statistical one.
3. Whatever the analytical platform used, all the statistical parameters are taken into account during the method validation fall whitening the values established a priori. The method can, therefore, be considered

validated and applicable for targeted and untargeted metabolomics analysis.

4. Regarding the two uremic toxins studied, the dialysis treatment leads to concrete benefits. As a matter of fact, if it is repeated three times a week with sessions of three hours each, the concentrations of p-cresyl sulfate and indoxyl sulfate over two weeks of therapy have halved in most of the patients.
5. Thanks to untargeted metabolomics approach different uremic toxins molecules could be monitored in CKD patients. At the moment they could be only considered as preliminary results and further measurements are still necessary. However, this method is a promising tool for future medical development: as a matter of facts, in the future cause progression in dialysis technologies, the detection of new markers (e.g. hippuric acid, CMPF, phenylacetyl-glutamine) will be surely fundamental for medical purposes.

The results here reported are just the tip of an iceberg. Uremic toxins project is still evolving and expanding. From late 2017 until nowadays, thanks to the routine method built upon UPLC-TQ MS platform has been used for the analysis of more than three hundred samples collected in different hospital in North-Est of Italy. The original purpose of the project (development of an analytical targeted metabolomics method able to detect and quantify, in human plasma samples, two specific uremic toxins) is evolved into a new one: monitor of the levels of p-cresyl sulfate and indoxyl sulfate over different hospitals (Savigliano, Verbania, Biella, Novara, Genova, Versiglia, Torino, and Cuneo) where different types of dialysis (hemodiafiltration and hemoperfusion) and different types of filters (Toray[®] filter, Filtryzer[®] filter and Solacea[®] filter are used).

In this case, the samples are collected, differently from what happened previously, only once in a month during the dialysis therapy. In this way, by having a concrete number of samples, we can conduct as statistical analysis (monitor the levels of these analytes over a vast territory) as select the suitable treatment to adopt for CKD patients.

References of Chapter 6

- [1] G. Alloatti et al. *Fisiologia dell'uomo*. Chap. 2. Milano: Edi-Ermes, 2010 (cit. on p. 75).
- [2] F. Barreto et al. "Dialysis elimination of waste products excess fluids balance electrolyte". In: *Nutrition in Clinical Practice* 20(2), 192-201 (2005). DOI: 10.1177/0115426505020002192 (cit. on p. 75).
- [3] F. Barreto et al. "Serum indoxyl sulphate is associated with vascular disease and mortality in chronical kidney disease patients". In: *Clinical Journal of the American Society of Nephrology* 4, 1551-1558 (2009). DOI: 10.2215/CJN.03980609 (cit. on p. 78).
- [4] C. Basile et al. "Tossine uremiche: il caso dei protein-bound compounds". In: *Giornale Italiano di Nefrologia* 27, 498-507 (2010) (cit. on pp. 75, 86).
- [5] A. Dhondt et al. "The removal of uremic toxins". In: *Kidney International* 58(76), S47-S59 (2000). DOI: 10.1046/j.1523-1755.2000.07606.x (cit. on p. 86).
- [6] S. Di Palma et al. "Evaluation of the Deuterium Isotope Effect in Zwitterionic Hydrophilic Interaction Liquid Chromatography Separations for Implementation in a Quantitative Proteomic Approach". In: *Analytical chemistry* 83(21), 8352-8356 (2011). DOI: 10.1021/ac2018074 (cit. on p. 85).
- [7] F.H.Bender et al. "Prevention of infectious complications in peritoneal dialysis: best demonstrated practices". In: *Kidney International* 70(103), S44-S54 (2006). DOI: 10.1038/sj.ki.5001915 (cit. on p. 75).
- [8] H. Fujii et al. "Role of oxidative stress and indoxyl sulfate in progression of cardiovascular disease in chronic kidney disease". In: *Therapeutic Apheresis and Dialysis* 15(2), 125-128 (2011). DOI: 10.1111/j.1744-9987.2010.00883.x (cit. on p. 76).
- [9] H. Fujii et al. "Role of Uremic Toxins for Kidney, Cardiovascular, and Bone Dysfunction". In: *Toxins* 10(5), 202 (2018). DOI: 10.3390/toxins10050202 (cit. on p. 76).

-
- [10] European Uremic Toxin Work Group. *EuTox - European Uremic Toxins*. 2020. URL: <https://www.uremic-toxins.org/> (cit. on p. 76).
- [11] M. Jemal et al. “Liquid chromatography/tandem mass spectrometry methods for quantitation of mevalonic acid in human plasma and urine: method validation, demonstration of using a surrogate analyte, and demonstration of unacceptable matrix effect in spite of use of a stable isotope analog internal standard”. In: *Rapid Communications in Mass Spectrometry* 17(15), 1723-1734 (2003). DOI: 10.1002/rcm.1112 (cit. on p. 85).
- [12] S. Liabeuf et al. “Free p-cresyl sulphate is a predictor of mortality in patients at different stages of chronic kidney disease”. In: *Nephrology Dialysis Transplantation - Oxford Journal* 25, 1183-1191 (2010). DOI: 10.1093/ndt/gfp592 (cit. on p. 78).
- [13] L. Maschio G. Oldrizzi et al. “Is There a “Point of No Return” in Progressive Renal Disease?” In: *Journal of the American Society of Nephrology* 2(4), 832-840 (1991) (cit. on p. 75).
- [14] T. Niwa. *Uremic Toxins*. Chap. 1, sec. 1.2. Hoboken: Wiley, 2012 (cit. on pp. 75, 86).
- [15] T. Niwa et al. “Indoxyl sulfate, a circulating uremic toxin, stimulates the progression of glomerular sclerosis”. In: *Journal of Laboratory and Clinical Medicine* 124(1), 96-104 (1994) (cit. on p. 76).
- [16] K.A. Smith et al. “Selecting a Structural Analog as an Internal Standard for Quantification of 6-Methylmercaptopurine by LC-MS/MS”. In: *The Journal of Applied Laboratory Medicine* 3(3), 384-396 (2018). DOI: 10.1373/jalm.2018.026187 (cit. on p. 85).
- [17] C.Y. Sung et al. “p-Cresol sulfate and indoxyl sulfate induce similar cellular inflammatory gene expressions in cultured proximal renal tubular cells”. In: *Nephrology Dialysis Transplantation* 28, 70-78 (2013). DOI: 10.1093/ndt/gfs133 (cit. on p. 77).
- [18] C.Y. Sung et al. “Uremic toxins induce kidney fibrosis by activating intrarenal renin-angiotensin-aldosterone system associated epithelial-to-mesenchymal transition”. In: *PLoS ONE* 7:e34026 (2012). DOI: 10.1371/journal.pone.0034026 (cit. on p. 77).

- [19] INC. Toray Industries. *Toray, Innovation by Chemistry*. 2019. URL: <http://www.toray.com/> (cit. on p. 78).
- [20] R. Vanholder et al. “Review on uremic toxins: Classification, concentration, and interindividual variability”. In: *Journal of the International society of nephrology* 63(5), 1934-1943 (2003). DOI: 10.1046/j.1523-1755.2003.00924.x (cit. on p. 76).
- [21] S. Wang et al. “Does a stable isotopically labeled internal standard always correct analyte response? A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma”. In: *Journal of Pharmaceutical and Biomedical Analysis* 43, 701-707 (2007). DOI: 10.1016/j.jpba.2006.08.010 (cit. on p. 85).
- [22] J. Weiling. “LC-MS-MS experiences with internal standards”. In: *Chromatographia* 55(1),S107-S113 (2002). DOI: 10.1007/s10337-011-2090-7 (cit. on p. 85).
- [23] V.C. Wu et al. “In acute kidney injury, indoxyl sulfate impairs human endothelial progenitor cells: Modulation by statin”. In: *Angiogenesis* 16, 609-624 (2013). DOI: 10.1007/s10456-013-9339-8 (cit. on p. 76).

Targeted analysis: Quorum sensing molecules

7

7.1. Introduction

Quorum sensing (QS) could be translated literally with *detection of the quorum* and it is the biochemical mechanism by which bacteria regulate gene expression in accordance with population density through the application of signal molecules [4, 21]. The QS communication system is used by bacteria in processes like diseases or infections and gives bacteria an extraordinary evolutionary complexity. As a matter of fact, the communication abilities offered by quorum sensing are highly useful for bacteria populations. These abilities, which include a widespread activity among living beings such as behaviour synchronization, allow bacteria populations survive and thrive: they can act as a single organism, they gain access to resources otherwise precluded, or they assure to the species better chances of survival [20, 21]. Quorum sensing is of interest because it plays a central role in different infectious disease processes [2, 12, 14].

One of the most studied bacteria linked to quorum sensing molecules is a marine bacterium called *Vibrio Fischeri*. The *vibrio fischeri* bacterium has the special property to be in particular cases bioluminescent. The light-emitting reaction of bacteria as *Vibrio fischeri* involves a luciferase-catalysed oxidation of reduced flavin mononucleotide [11]. What is interesting about *Vibrio fischeri* and bio luminesced bacteria is when primitive organisms like bacteria exploit their light-emission property. When the bacteria are alone, so when they were in dilute suspension, they made no light. But when they grew to a certain cell number all the bacteria turned on light simultaneously. With proper signaling systems, bacterial communities can behave as pseudo-multicellular organisms, such as by biofilm formation, virulence, and production of antibiotics [10].

Standard QS pathways consist of bacteria populations, signal molecules, and behavioural genes. The signal molecules, known as autoinducers,

are secreted into the environment by bacteria or the purpose of cell-to-cell communication and they gradually increase in concentration as the bacteria population grows [21]. After reaching a threshold level concentration (the quorum, where the molecules' name came from), molecules become detectable to bacteria populations, which then activate corresponding response genes that regulate various behaviours, such as virulence expression (see 7.1).

Gram-positive and Gram-negative bacteria use different communication-ways. The Gram-positive bacteria exploit oligopeptides, while the Gram-negative bacteria communicate among themselves using different kinds of autoinducers [19, 21]. The main ones are N-acyl homoserine lactones (AHLs). These molecules are characterized by a γ -lactonic cycle which is N-acylated in α position and the acyl-chain (indicated as R-chain in figure 7.2) can differentiate through its length. The chain length is the signal specificity factor for bacteria and, generally, the chain contains between 4 and 14 carbon atoms [15]. Furthermore, the presence or the absence of an -oxo or a -hydroxy group linked to the third carbon is another element of distinction.

In some human opportunistic pathogens, as *Pseudomonas aeruginosa*, the secretion of QS molecules like N-(3-oxododecanoyl)-L-homoserine lactone (3-oxo-C12- AHL) or N-butanoyl-L-homoserine lactone (C4-AHL) depends from the regulatory circuits systems (Las or Rhl systems [15]). In bacteria those systems control expression of different virulence genes in a population density-dependent [5, 6, 13]. These molecules can be conveniently analysed by HPLC-MS techniques[3, 5, 6], using neutral loss scan methods (monitoring the neutral loss of a lactone cycle molecule) or precursor ion scan methods (monitoring, on the contrary, the formation of an ionized cycle): these kinds of analysis allow to detect – in a single step – the whole class of AHLs and can be applied to different kinds of samples (from bacterial culture supernatants to biological fluids in patients affected by bacterial infections).

In addition to AHLs, Gram negative bacteria use other types of signaling molecules. For example, the pathogen bacterium *Pseudomonas aeruginosa* can communicate through a hydroxyquinolone signaling molecule (HQs), 2-

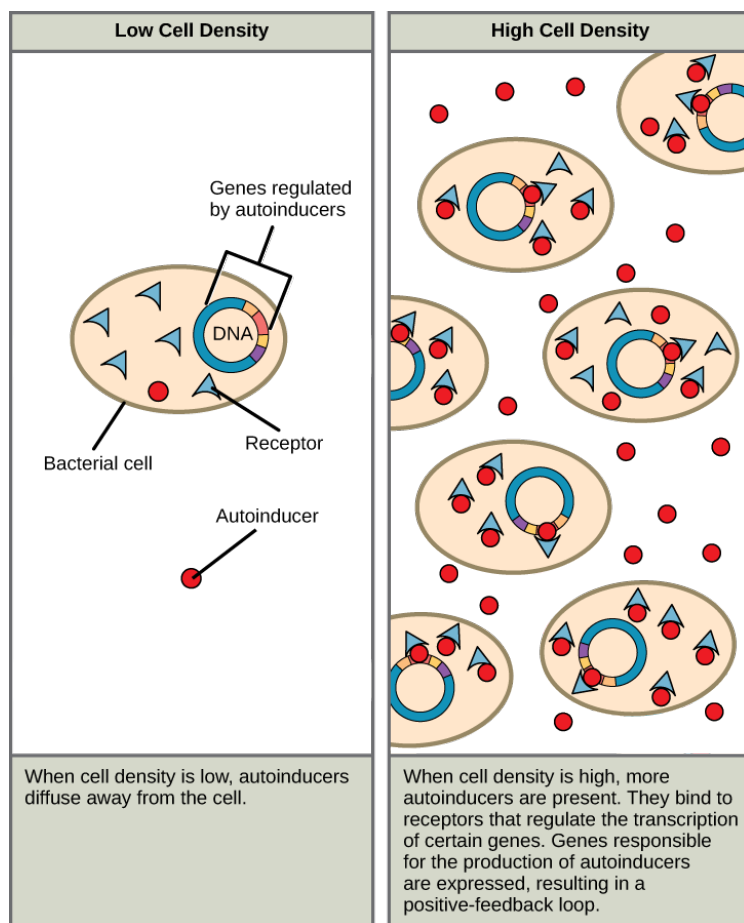


Figure 7.1.: *Cell density mediated by gene expression in quorum sensing mechanism (e.g. virulence expression). Figure drawn from [9]*

heptyl-3-hydroxy-4-quinolone and it releases in the extracellular milieu also its immediate biosynthetic precursor, 2-heptyl-4-hydroxyquinoline (HHQ) [13]. The general structure for these molecules is reported in figure 7.3. HPLC-MS techniques cited before can also be used to recognize HQs signaling molecules.

7.2. Aim of the work

Purpose of this work is to build up and validate a method for the quantification of quorum sensing molecules in different matrices (cellular cultures

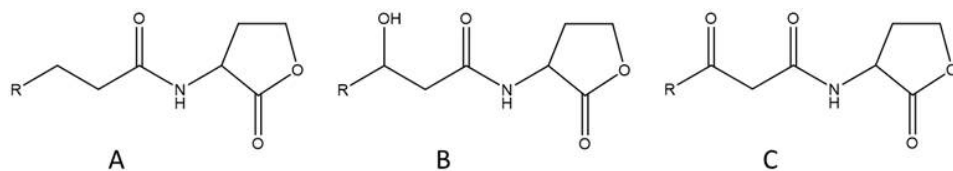


Figure 7.2.: *N*-acyl homoserine lactone general structure. A: Non-substituted *N*-acylhomoserine-*L*-lactone (*C_n*-HSL) acyl chain; B: *N*-(3-Hydroxyacylhomoserine)-*L*-lactone (3-OH-*C_n*-HSL) acyl chain; C: *N*-(3-oxoacylhomoserine)-*L*-lactone (3-oxo- *C_n*-HSL) acyl chain. The length chain *R*- is variable in carbon atoms in the range C4-C14.

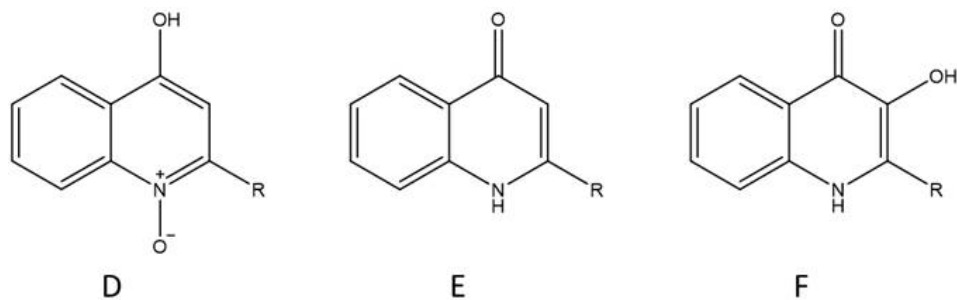


Figure 7.3.: Quinolone signaling molecules general structure. D: 2-Alkyl-4-hydroxyquinolone *N*-oxide with a chain length variable (C7-C9); E: 2-Alkyl-4(1H)-quinolone alkyl chain (the length chain is variable, C7-C11); F: 2-Alkyl-3-hydroxy-4(1H)-quinolone, with an alkyl chain length variable from C7 until C11. For PQS the length of the *R*-chain in position 2 is a C9 aliphatic chain

and biological samples) by UPLC-TQMS (QTRAP[®] 5500 System by Sciex coupled to a Shimadzu Nexera X2) platform, applying multiple reaction monitoring, neutral loss, and precursor ion experiments to detected through targeted and untargeted approaches the low concentration of quorum sensing signaling molecules in samples of interest.

The final aim will be to apply this validated method for the identification of specific biomarkers involved in the sepsis-related multi-organ failure, associated with high mortality. By operating as in targeted (MRM scan method) as in untargeted (NL and PI modes) ways, this work can furnish a precious and indispensable tool to medical laboratories for confirming the AKI disease when it is still at an early stage. Furthermore, the relationship QS signal molecules and microvesicles could pave the way to the identification and synthesis of new potential targets in medical fields.

An aspect to underline is that bacteria and plasma samples were analysed at different times. This led, therefore, to study the bacterial samples (the first to have been treated) only by untargeted NL and PI methods for the quantification of AHLs and HQs species. Contrarily, plasma samples had been analysed subsequently (when, unfortunately, bacterial sample could be no more re-analysed), and they were submitted to all the previously described scan methods.

7.3. Sample preparation and method settings

Sample preparation: chemicals and instrumentation

As representatives of AHLs and HQs classes -heptyl-4-hydroxyquinoline, N-(3-oxododecanoyl)-L-homoserine lactone, N-hexanoyl homoserine lactone d3, and N-butanoyl-L-homoserine lactone are used in this work. They were purchased from Sigma-Aldrich (Milan, Italy) and applied in all experiments without any kind of further purification treatment. All aqueous solutions were prepared with ultrapure water Millipore Milli-Q[™] system. All organic solvents (chloroform and methanol hypergrade for LC-MS), and formic acid were purchased from VWR International (Milan, Italy).

For what concerning biological sample preparation procedure, we analysed plasma samples from volunteers' AKI patients. All the patients in-

volved in the present study express their consent and express their will based on their awareness of the proposed research upon their biological fluids, freely deciding whether to accept it. Plasma samples were taken during dialysis sessions to which patients undergo several times a week. 200 μL of plasma was spiked with N-Hexanoyl-L-homoserine lactone-d3 at 200 $\mu\text{g}/\text{L}$. Successively, it was extracted twice with 1 mL of ethyl acetate. After each addition of organic solvent, the sample was centrifuged at $5000 \text{ g} \times 5 \text{ min}$ to separate the two phases. Then, the organic fractions were collected and dried under a stream of N_2 heated to 40°C . Finally, the residue was dissolved in 0.1 mL of HCOOH 0.1% in water/ACN 60:40. However, the 2x pre-concentration obtained is a fundamental step given the average concentration levels of the molecules of our interest in the samples of interest.

Moving now to bacterial cultures' samples, different *Pseudomonas aeruginosa* PAO1 cultures were analysed:

1. wild type bacterial culture, grown in the LB culture medium (rich in nutrients);
2. wild type bacterial culture, grown in the M9 culture medium (low in nutrients);
3. RhII- mutant bacterial culture, grown in the LB culture medium (rich in nutrients);
4. RhII- mutant bacterial culture, grown in the M9 culture medium (low in nutrients).

The *Pseudomonas aeruginosa* RhII- mutant has been rendered unable to produce the RhII synthetase protein and therefore is unable to produce C4-HSL. The creation of bacterial cultures was performed by Prof. Orlandi at the University of Insubria. The extraction procedure described for plasma samples is the same one used for bacterial cultures extraction step. As happened before, also in this case, a 2x pre-concentration is essential to assure the detection of molecules of interest within the samples.

Validation of the analytical methods

The validation procedure upon the UPLC-TQMS platform was performed according to the European Medicine Agency (EMA)[1] and Eurachem guidelines [7, 17]. The calibration curves were run using quorum sensing molecules free matrix (plasma belonging from healthy people) by performing standard addition method. For method validation, different parameters were evaluated such as selectivity, intra- and inter-run precision and accuracy, the linearity of calibration curves, the lower limit of quantification LLOQ, and recovery at LLOQ values.

For the validation of AHLs molecules quantification we selected N-butanoyl-L-homoserine lactone and N-(3-oxododecanoyl)-L-homoserine lactone as molecules class representative. On the contrary, for the HQs molecules, 2 heptyl 3,4 dihydroxyquinoline was used as a representative class molecule. The results of the validation parameters are reported, for N-butanoyl-DL-homoserine lactone, from table 7.1 to table 7.5; for N-(3-oxododecanoyl)-L-homoserine lactone from table 7.7 to table 7.11; finally, for 2-heptyl-3,4-dihydroxyquinoline from table 7.13 to table 7.17.

Two different aspects must be remarked. The first one is related to the differences in linearity ranges between MRM, NL, and PI methods. If for MRM the range goes from 0.4 to 400 $\mu\text{g/L}$, for NL and PI the values of LLOQ are respectively 5 and 1 $\mu\text{g/L}$. Furthermore, for the PI method the point at 400 $\mu\text{g/L}$ had to be eliminated during the validation phase because, observing the data, the linearity range was no longer maintained. So, compared with the neutral loss and precursor ion analysis, the MRM approach provided superior sensitivity, selectivity, and linearity range. Analogously, in HQs MRM detection 300 and 400 $\mu\text{g/L}$ calibration points need to be eliminated for the same reason.

The second aspect to underline is linked to 2-heptyl-3,4-dihydroxyquinoline validation. As reported in [16] and mentioned above, *Pseudomonas aeruginosa* produces an autoinducer (2-heptyl-3,4-dihydroxyquinoline) distinct from homoserine species. So, since this opportunistic pathogen is related to infections at different sites within the body, we have as an aim to monitor (and eventually measure) its concentration within plasma samples. To do

that, the MRM approach is the only one that guarantees levels of sensitivity and selectivity suitable for the purpose. As a matter of fact, as already mentioned before, if in cellular cultures the HQs levels could be detected through PI approaches, in real plasma sample only the MRM approach could succeed in the goal.

Moving to validation tables some aspects need to be clarified. Tables 7.1, 7.7, and 7.13 resume the linearity of calibration curve (expressed as coefficient of correlation R^2) and the different detection limits required by EMA and Eurachem guidelines: LOD, LOQ, LLOQ, and ULOQ. The determination coefficients (R^2) and linear regression analysis gave values always $R^2 > 0.990$. Furthermore, the slopes of calibration curves were highly reproducible among different analytical sessions.

Tables 7.2, 7.8, and 7.14 go on with selective parameters (the carry-over effect). This phenomenon is studied by comparing the signal of the molecule of interest in blank sample after the injection of the highest concentration point of the calibration curve. Blank samples could be defined as samples obtained from a matrix that is free of the analytes of interest. The analytical method is selective for analytes of interest since all the areas obtained in blank samples are lower than 20% of the corresponding LLOQ area.

Tables 7.3, 7.9, and 7.15 resume the calculated concentration in $\mu\text{g/L}$ of the four different curves obtained in four different analytical sessions. The calibration curves for all scan methods under validation were linear in the selected range. The slope values were highly reproducible among different analytical sessions. To consider the curves acceptable for validation purposes, each must consist of a minimum of seven points equally distributed in the chosen linearity range. Furthermore, the acceptability range (the lower and the upper permissible values), the average concentration ($\mu\text{g/L}$) for each level, and the standard deviation SD are reported.

Tables 7.4, 7.10, and 7.16 contain the validated parameters concerning accuracy, precision and BIAS for N-butanoyl-L-homoserine lactone, N-(3-oxododecanoyl)-L-homoserine lactone, and 2-heptyl-3-hydroxy-4(1H) quinolone respectively. Parameters obtained successfully respect validation purposes expressed by EMA. The mean accuracy of the back-calculated concentrations for MRM, NL, and PI analysis ranged from 97.5% to 113% in plasma

Table 7.1.: *List of validated parameters for N-butanoyl-DL-homoserine lactone. Linearity of calibration curve expressed as coefficient of correlation R² (> 0.990). LOD: Limit of Detection; LOQ: Limit of Quantification; LLOQ: Lower Limit of Quantification; ULOQ: Upper Limit of Quantification*

Scan method	Linearity R ²	LOD ($\mu\text{g/L}$)	LOQ ($\mu\text{g/L}$)	LLOQ ($\mu\text{g/L}$)	ULOQ ($\mu\text{g/L}$)
MRM	0.9993	1.02	3.39	0.4	400
NL	0.9989	0.77	2.55	5.0	400
PI	0.9985	0.10	0.32	1.0	300

samples. On the contrary, the precision of back-calculated concentrations – expressed as coefficient of variation (CV%) – was acceptable and successfully respect validation purposes expressed by EMA. Furthermore, as reported in EMA protocol, at least of 4 concentrations’ levels in the monitored range must be evaluated for validation purposes. In our case, only MRM can be appreciated over the entire range (LLOQ for NL and PI are 1 and 5 $\mu\text{g/L}$, respectively). Finally, the mean inter-day vs. intra-day validation parameters are similar and results are within EMA guidelines limits for all concentrations taken into account.

In tables 7.5, 7.6, 7.10, 7.12, 7.17, and 7.18 are described parameters concerning stability test. According to EMA, stability of the analyte is evaluated using both low- and high-level QC samples in three replicates. The investigation of stability should cover short-term stability at room temperature or sample processing temperature (the so-called bench-top stability); freeze-thaw stability; long-term freezer stability. For our purposes, because the samples are not stored in a freezer for periods longer than 24h, it was decided not to proceed with the evaluation of the long-term freezer stability. Parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between 85.0 <STAB% <115.0 and that, therefore, plasma matrix is stable as to operative conditions temperatures (tables 7.6, 7.12, and 7.18) as to freezing and thawing cycles (tables 7.5, 7.11, and 7.17), without affecting the concentration of the analyte.

In table 7.19 evaluation of internal standard recovery of the extraction procedure is made. For plasma samples the mean recovery, considering as LLOQ as ULOQ concentrations, stands at 80% on average.

Table 7.2.: *List of validated selectivity parameters for N-butanoyl-DL-homoserine lactone. Analytical method is selective for analytes of interest since all the areas obtained are lower than 20% of the corresponding LLOQ area. Blank samples are samples obtained from a matrix that is free of the analytes of interest*

Blank samples	Area N-butanoyl-DL-homoserine lactone			20% LLOQ area – average		
	MRM	NL	PI	MRM	NL	PI
1	4.82x10 ²	1.75x10 ³	5.94x10 ³			
2	4.85x10 ²	3.67x10 ³	1.00x10 ⁴			
3	1.27x10 ³	1.88x10 ³	6.03x10 ³	2.63x10 ³	1.33x10 ⁴	1.54x10 ⁵
4	4.31x10 ²	2.28x10 ³	8.88x10 ³			
5	1.29x10 ³	2.89x10 ³	6.12x10 ³			
6	3.72x10 ²	2.43x10 ³	4.77x10 ³			

Table 7.3.: *List of validated calibration curve parameters for N-butanoyl-DL-homoserine lactone. Parameters obtained successfully respect validation purposes. Concentration values marked in bold, as indicated in the EMA protocol, cannot be included in acceptability range ($\pm 15\%$ of nominal concentration, $\pm 20\%$ for LLOQ level). However, since this principle must be satisfied at least by 75% of calibration points with a minimum of six concentration levels, straight lines can be considered valid usable for validation purposes. Furthermore, as indicated in the Eurachem protocol, after appropriately calculating the statistical regression, the residues were calculated. Since the residues show a random distribution around zero, it is possible to confirm that the response range within which you work is linear*

Nominal concentration ($\mu\text{g/L}$)	Real concentration (recalculated, $\mu\text{g/L}$)			Acceptability range						SD		
	Average from 4 different curves			Lower limit ($\mu\text{g/L}$)			Upper limit ($\mu\text{g/L}$)			MRM	NL	PI
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
0.4	0.45	x	x	0.32	x	x	0.48	x	x	0.02	x	x
1	1.06	x	1.06	0.85	x	0.85	1.15	x	1.15	0.05	x	0.10
5	5.33	5.65	5.11		4.25			5.75		0.40	0.54	0.42
10	9.99	10.87	10.06		8.5			11.5		1.27	1.83	1.26
50	53.54	52.43	50.99		42.5			57.5		1.97	1.08	2.69
100	97.51	102.31	105.94		85			115		4.49	4.86	3.92
200	187.36	201.76	204.40		170			230		18.69	7.56	9.51
300	301.57	295.36	298.44		255			345		3.75	5.84	1.90
400	403.21	405.39	x	340	340	x	460	460	x	5.36	3.96	x

Table 7.4.: List of validated parameters concerning accuracy, precision and BIAS for *N*-butanoyl-DL-homoserine lactone. Parameters obtained successfully respect validation purposes expressed by EMA. As reported in EMA protocol, at least 4 concentrations (as low as high values) con must be evaluated. In our case, only MRM can be appreciated over the entire range (LLOQ for NL and PI are 1 and 5 $\mu\text{g/L}$, respectively)

Nominal concentration ($\mu\text{g/L}$)	Accuracy (%) intraday			Precision (%) intraday			Bias (%) intraday		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
0.4	113.344	✗	✗	4.344	✗	✗	13.344	✗	✗
1	105.698	✗	106.013	4.572	✗	9.762	5.698	✗	6.013
5	106.683	112.987	102.177	7.400	9.519	8.205	6.683	8.673	2.177
200	97.511	100.880	102.199	4.602	3.747	4.655	6.322	0.880	2.199
300	100.522	98.454	99.480	1.243	1.976	0.635	0.522	1.546	0.520
400	100.801	101.348	✗	1.330	0.977	✗	0.801	1.348	✗
Nominal concentration ($\mu\text{g/L}$)	Accuracy (%) inter-day			Precision (%) inter-day			Bias (%) inter-day		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
0.4	110.887	✗	✗	4.283	✗	✗	11.917	✗	✗
1	103.648	✗	104.749	3.681	✗	9.603	4.737	✗	5.850
5	104.404	111.434	101.051	6.351	9.233	8.042	5.668	8.355	2.152
200	95.525	99.719	101.018	3.521	3.669	4.522	5.173	0.882	2.140
300	99.421	97.428	98.490	1.278	2.099	0.807	0.601	1.665	0.621
400	99.893	100.200	✗	1.359	1.020	✗	0.880	1.324	✗

Table 7.5.: List of validated parameters concerning freeze-thaw stability test for *N*-butanoyl-DL-homoserine lactone. As for MRM as for NL, parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between 85.0 <STAB% <115.0 and that, therefore, plasma matrix is stable to freezing and thawing cycles, without affecting the concentration of the analyte

Nominal concentration ($\mu\text{g/L}$)	MRM	t_0		MRM	t_{cycle}		Average stability %, STAB%		
		NL	PI		NL	PI	MRM	NL	PI
LLOQ	0.472	5.697	1.098	0.430	4.843	0.994	97.575	86.600	89.109
	0.438	6.329	1.132	0.453	6.189	0.986			
	0.474	5.748	1.128	0.465	4.426	1.012			
ULOQ	396.903	403.941	299.024	411.702	361.690	285.120	101.645	89.513	95.877
	401.507	401.878	298.664	403.206	365.709	281.110			
	402.711	403.633	295.422	405.861	355.197	289.982			

Table 7.6.: List of validated parameters concerning freeze-bench top stability test for *N*-butanoyl-*L*-homoserine lactone. Parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between $85.0 < STAB\% < 115.0$ and that, therefore, plasma matrix shows a proper bench-top stability without affecting the concentration of the analyte

Concentration level ($\mu\text{g/L}$)	t0			Bench-top stability			Average stability %, STAB%		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
LLOQ	0.472	5.697	1.098	0.441	4.918	0.942	94.531	90.411	86.121
	0.438	6.329	1.132	0.412	5.953	0.959			
	0.474	5.748	1.128	0.4555	5.222	0.991			
ULOQ	396.903	403.941	299.024	395.091	397.901	292.011	98.935	98.320	97.257
	401.507	401.878	298.664	397.121	398.199	284.888			
	402.711	403.633	295.422	396.091	393.019	291.663			

Table 7.7.: List of validated parameters for *N*-(3-oxododecanoyl)-*L*-homoserine lactone. Linearity of calibration curve expressed as coefficient of correlation R^2 (> 0.990). LOD: Limit of Detection; LOQ: Limit of Quantification; LLOQ: Lower Limit of Quantification; ULOQ: Upper Limit of Quantification

Scan method	Linearity R^2	LOD ($\mu\text{g/L}$)	LOQ ($\mu\text{g/L}$)	LLOQ ($\mu\text{g/L}$)	ULOQ ($\mu\text{g/L}$)
MRM	0.9994	0.52	1.74	0.4	400
NL	0.9961	0.50	1.67	5	400
PI	0.9957	0.14	0.46	1	300

Table 7.8.: List of validated selectivity parameters for *N*-(3-oxododecanoyl)-*L*-homoserine lactone. Analytical method is selective for analytes of interest since all the areas obtained are lower than 20% of the corresponding LLOQ area. Blank samples are samples obtained from a matrix that is free of the analytes of interest

Blank samples	Area <i>N</i> -(3-oxododecanoyl)- <i>L</i> -homoserine lactone			20% LLOQ area – average		
	MRM	NL	PI	MRM	NL	PI
1	3.45×10^3	8.74×10^3	9.94×10^3			
2	4.67×10^3	1.14×10^4	8.77×10^3			
3	5.12×10^3	1.54×10^4	1.00×10^4			
4	4.31×10^3	1.33×10^4	8.60×10^3	1.09×10^4	2.69×10^6	8.72×10^5
5	3.27×10^3	9.64×10^3	9.57×10^3			
6	4.99×10^3	2.01×10^4	1.19×10^4			

Table 7.9.: List of validated calibration curve parameters for *N*-(3-oxododecanoyl)-*L*-homoserine lactone. Parameters obtained successfully respect validation purposes. Concentration values marked in bold, as indicated in the EMA protocol, cannot be included in acceptability range ($\pm 15\%$ of nominal concentration, $\pm 20\%$ for LLOQ level). However, since this principle must be satisfied at least by 75% of calibration points with a minimum of six concentration levels, straight lines can be considered valid usable for validation purposes. Furthermore, as indicated in the Eurachem protocol, after appropriately calculating the statistical regression, the residues were calculated. Since the residues show a random distribution around zero, it is possible to confirm that the response range within which you work is linear

Nominal concentration ($\mu\text{g/L}$)	Real concentration (recalculated, $\mu\text{g/L}$)			Acceptability range						SD		
	Average from 4 different curves			Lower limit ($\mu\text{g/L}$)			Upper limit ($\mu\text{g/L}$)					
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
0.4	0.40	X	X	0.32	X	X	0.48	X	X	0.03	X	X
1	0.95	X	1.11	0.85	X	0.85	1.15	X	1.15	0.07	X	0.16
5	5.06	5.35	5.41		4.25			5.75		0.34	0.14	0.30
10	9.87	9.19	10.15		8.5			11.5		0.44	0.40	1.35
50	55.70	53.01	49.96		42.5			57.5		1.49	3.14	3.54
100	93.28	93.55	99.23		85			115		0.80	7.78	9.29
200	201.79	199.59	197.72		170			230		2.94	8.43	16.19
300	301.64	284.19	301.11		255			345		4.60	8.48	7.44
400	398.85	408.49	X	340	340	X	460	460	X	2.57	6.62	X

Table 7.10.: List of validated parameters concerning accuracy, precision and BIAS for *N*-(3-oxododecanoyl)-*L*-homoserine lactone. Parameters obtained successfully respect validation purposes expressed by EMA. As reported in EMA protocol, at least 4 concentrations (as low as high values) con must be evaluated. In our case, only MRM can be appreciated over the entire range (LLOQ for NL and PI are 1 and 5 $\mu\text{g/L}$, respectively)

Nominal concentration ($\mu\text{g/L}$)	Accuracy (%) intraday			Precision (%) intraday			Bias (%) intraday		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
0.4	99.026	X	X	8.585	X	X	0.974	X	X
1	95.050	X	111.284	7.114	X	14.134	4.950	X	8.170
5	101.198	106.019	108.170	6.680	2.591	5.587	1.198	6.904	1.481
200	100.898	99.795	98.858	1.458	4.224	8.190	0.898	0.205	1.142
300	100.546	94.729	100.371	1.524	2.982	2.470	0.546	5.271	0.371
400	99.713	102.123	X	0.644	1.621	X	0.287	2.123	X

Nominal concentration ($\mu\text{g/L}$)	Accuracy (%) inter-day			Precision (%) inter-day			Bias (%) inter-day		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
0.4	96.800	X	X	4.443	X	X	0.155	X	X
1	94.221	X	109.976	7.130	X	13.787	2.501	X	7.926
5	100.157	104.661	106.792	2.591	2.529	5.576	0.245	6.628	1.456
200	92.808	98.594	97.799	0.550	4.151	7.984	1.082	0.302	3.113
300	102.423	95.834	100.011	1.490	2.922	5.811	0.735	5.150	1.174
400	96.557	98.871	X	0.793	2.001	X	0.379	1.999	X

Table 7.11.: *List of validated parameters concerning freeze-thaw stability test for N-(3-oxododecanoyl)-L-homoserine lactone. As for MRM as for NL, parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between $85.0 < STAB\% < 115.0$ and that, therefore, plasma matrix is stable to freezing and thawing cycles, without affecting the concentration of the analyte*

Nominal concentration ($\mu\text{g/L}$)	t_0			t_{cycle} (3 cycles of freeze-thaw procedure)			Average stability %, STAB%		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
LLOQ	0.426	5.580	1.329	0.426	4.843	0.968	105.931	90.062	87.375
	0.396	5.243	0.887	0.344	5.189	0.889			
	0.388	5.244	1.095	0.398	4.426	0.975			
ULOQ	395.633	397.204	300.589	402.251	361.690	286.191	101.158	88.585	90.186
	397.263	411.222	267.895	398.851	365.709	249.569			
	400.257	414.158	252.251	405.861	355.197	207.325			

Table 7.12.: *List of validated parameters concerning freeze-bench top stability test for N-(3-oxododecanoyl)-L-homoserine lactone. Parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between $85.0 < STAB\% < 115.0$ and that, therefore, plasma matrix shows a proper bench-top stability without affecting the concentration of the analyte*

Concentration level ($\mu\text{g/L}$)	t_0			Bench-top stability			Average stability %, STAB%		
	MRM	NL	PI	MRM	NL	PI	MRM	NL	PI
LLOQ	0.472	5.697	1.098	0.441	4.918	0.942	91.695	92.183	86.091
	0.426	5.580	1.329	0.401	5.011	1.079			
	0.396	5.243	0.887	0.391	4.981	0.802			
ULOQ	0.388	5.244	1.095	0.319	4.811	0.949	98.646	96.103	97.834
	395.633	397.204	300.589	393.101	391.088	298.327			
	397.263	411.222	267.895	389.781	390.818	260.237			

Table 7.13.: *List of validated parameters for 2-heptyl-3,4-dihydroxyquinoline. Linearity of calibration curve expressed as coefficient of correlation $R^2 (> 0.990)$. LOD: Limit of Detection; LOQ: Limit of Quantification; LLOQ: Lower Limit of Quantification; ULOQ: Upper Limit of Quantification*

Scan method	Linearity R^2	LOD ($\mu\text{g/L}$)	LOQ ($\mu\text{g/L}$)	LLOQ ($\mu\text{g/L}$)	ULOQ ($\mu\text{g/L}$)
MRM	0.999	0.03	0.09	0.4	200

Table 7.14.: *List of validated selectivity parameters for 2-heptyl-3,4-dihydroxyquinoline. Analytical method is selective for analytes of interest since all the areas obtained are lower than 20% of the corresponding LLOQ area. Blank samples are samples obtained from a matrix that is free of the analytes of interest.*

Blank sample	Area 2-heptyl-3,4-dihydroxyquinoline 20% LLOQ area – average	
	MRM	MRM
1	1,01x10 ³	1,44x10 ³
2	1,37x10 ³	
3	1,19x10 ³	
4	1,11x10 ³	
5	1,29x10 ³	
6	1,22x10 ³	

Table 7.15.: *List of validated calibration curve parameters for 2-heptyl-3,4-dihydroxyquinoline. All parameters obtained successfully respect validation purposes. Concentration values marked in bold, as indicated in the EMA protocol, cannot be included in acceptability range ($\pm 15\%$ of nominal concentration, $\pm 20\%$ for LLOQ level). However, since this principle must be satisfied at least by 75% of calibration points with a minimum of six concentration levels, straight lines can be considered valid usable for validation purposes. Furthermore, as indicated in the Eurachem protocol, after appropriately calculating the statistical regression, the residues were calculated. Since the residues show a random distribution around zero, it is possible to confirm that the response range within which you work is linear.*

Nominal concentration ($\mu\text{g/L}$)	Real concentration (recalculated, $\mu\text{g/L}$)		Acceptability range		SD
	Average from 4 different curves		Lower limit ($\mu\text{g/L}$)	Upper limit ($\mu\text{g/L}$)	
	MRM	MRM	MRM	MRM	
0.4	0,409	0.32	0.48	0,019	
1	1,089	0.85	1.15	0,146	
5	4,897	4.25	5.75	0,084	
10	9,464	8.5	11.5	1,608	
50	49,746	42.5	57.5	1,368	
100	99,598	85	115	1,469	
200	200,291	170	230	0,953	
300	301.64	255	345	4.60	
400	398.85	340	460	2.57	

Table 7.16.: *List of validated parameters concerning accuracy, precision and BIAS for 2-heptyl-3,4-dihydroxyquinoline. Parameters obtained successfully respect validation purposes expressed by EMA. As reported in EMA protocol, at least 4 concentrations (as low as high values) con must be evaluated.*

Nominal concentration ($\mu\text{g/L}$)	Accuracy (%) intraday	Precision (%) intraday	Bias (%) intraday
	MRM	MRM	MRM
0.4	107.118	6.273	2.231
5	91.565	4.711	2.101
50	92.947	6.902	0.147
200	100.387	0.164	0.146

Nominal concentration ($\mu\text{g/L}$)	Accuracy (%) inter-day	Precision (%) inter-day	Bias (%) inter-day
	MRM	MRM	MRM
0.4	89.159	5.319	3.257
5	92.793	4.823	1.083
50	94.333	3.837	2.149
200	99.383	1.162	1.145

Table 7.17.: *List of validated parameters concerning freeze-thaw stability test for 2-heptyl-3,4-dihydroxyquinoline. Parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between $85.0 < \text{STAB\%} < 115.0$ and that, therefore, plasma matrix is stable to freezing and thawing cycles, without affecting the concentration of the analyte.*

Nominal concentration ($\mu\text{g/L}$)	t_0	t_{cycle} (3cycles of freeze – thaw procedure)	Average stability %, STAB%
	MRM	MRM	MRM
LLOQ	0.405	0.385	97.58
	0.401	0.399	
	0.385	0.378	
ULOQ	197.773	188.152	98.71
	194.917	190.518	
	201.123	199.811	

Table 7.18.: *List of validated parameters concerning freeze-bench top stability test for 2-heptyl-3,4-dihydroxyquinoline. Parameters obtained successfully respect validation purposes. Based on results, average stability % value at each concentration level is in the range between $85.0 < \text{STAB\%} < 115.0$ and that, therefore, plasma matrix shows a proper bench-top stability without affecting the concentration of the analyte*

Concentration level ($\mu\text{g/L}$)	t_0	bench-top stability	Average stability %, STAB%
	MRM	MRM	MRM
LLOQ	0.405	0.301	91.409
	0.401	0.385	
	0.385	0.4	
ULOQ	197.773	191.032	97.185
	194.917	194.998	
	201.123	190.91	

Instrumental parameters

Separation and analysis of all analytes were achieved upon a UPLC-TQMS platform. HPLC consist of a Shimadzu Nexera X2 ultra high-performance liquid chromatography UPLC system (Shimadzu, Kyoto Japan) coupled for identification and quantitation to a SCIEX QTRAP[®] 5500 LC-MS/MS system mass spectrometer (SCIEX, Darmstadt, Germany). This triple quadrupole was equipped with a Turbo V[™] Source which utilized nitrogen and air as the sheath and reagent gas.

For all AHLs UPLC-TQMS analyses, LC system operated in a gradient mode with a flow rate of 200 $\mu\text{g}/\text{L}$. The analytical column used was a Phenomenex Luna C18 reverse-phase analytical column (150 \times 2.1 mm i.d., 3 μm particles). The sample injection volume was 10 μL . Rinse solution consisted in 2-propanol HPLC-MS grade; rinse volume injected was 500 μL performed at 30 $\mu\text{L}/\text{sec}$ before and after aspiration and rinse time was set to 10 sec. Autosampler and oven temperatures were set at 15° C and 45° C respectively for all the duration of the analysis.

To evaluate ionization of QS compounds, ad initial direct FIA injection into QTRAP[®] 5500 mass spectrometer was performed. N-(3-oxododecanoyl)-L-homoserine lactone, N-butanoyl-L-homoserine lactone, and 2-heptyl-4-hydroxyquinoline were used as reference standard to build up MRM and NL methods (figure 7.4). Furthermore, N-Hexanoyl-L-homoserine lactone-d3 molecule was used as internal standard. This evaluation indicated that the electrospray ionization (ESI) in positive polarity, creating $[\text{M}+\text{H}]^+$ ions, was much more effective in production of a signal respect than the negative electrospray ionization. ESI-MS conditions were optimized separately by infusing the desired compound at a concentration of 0.3 $\mu\text{g}/\text{mL}$, into mass spectrometer by a syringe at a constant rate of 15 $\mu\text{g}/\text{L}$.

For MRM, NL, and PI AHLs analyses, the HPLC column was routinely equilibrated before use for 20 minutes. Aqueous eluent of high-performance liquid chromatography was formic acid 0.1% in water. The organic component of the mobile phase was formic acid 0.1% in acetonitrile, HPLC-MS grade. The ratio of the previous aqueous mixture to acetonitrile was first set at 60:40. LC run was in a gradient mode so that the mixture was changed to 0:100 during the first 19 min. This ratio was maintained for 10 min,

then it is changed to 60:40 at 30 min and was held there for 20 min for re-equilibration of the system. Chromatographic run time was globally 50 min for each analysis.

Observing chromatograms reported from figure 7.6 to figure 7.13, it is possible to note how bacterial cultures samples and plasma samples were analysed with two different chromatographic gradients. Both LC runs were in a gradient mode so that the solvents' mixture was at the beginning 60:40 as mentioned previously. For bacterial cultures samples, the mixture was changed to 0:100 during the first 35 min. This ratio was maintained for 10 min, then it returns in one minute to initial 60:40 conditions and it was held there for 24 min in order to obtain in 70 minutes a proper system's re-equilibration. On the contrary, for plasma sample, LC run was in gradient mode as well and initial solvents' percentages were changed to 0:100 during the first 19 min. This ratio was maintained for 10 min, then it is changed to 60:40 at 30 min and was held there for 20 min for re-equilibration of the system. Chromatographic run time was globally in this case 50 min for each analysis.

The difference between the two chromatographic gradients is essentially linked to two reasons:

- Bacterial culture samples provide a greater number of chromatographic peaks. Therefore, to obtain a better chromatographic separation of these entities, a slower percentage solvents' variation during gradient elution must be applied.
- For biological samples, considering their number and the limited quantity of QS molecules present on average, they could be analysed by a faster gradient. In this way, the method does not lose sensitivity and becomes more efficient from an economical point of view.

Besides, MRM analysis of 2-heptyl-3-hydroxy-4-quinolone consists in acetonitrile 0.1% formic acid as organic solvent and 2-picolinic acid 2 mM/formic acid 0.1% in water. 2-picolinic acid acting as a bidentate chelator (towards elements such as zinc, copper or iron [8]) prevents peak distortion caused by 2-heptyl-4-hydroxyquinoline, an iron chelator molecule [18]. To perform chromatographic separation, ratio of aqueous and organic solvent was set at

Table 7.19.: *List of validated parameters concerning recovery parameters for N-hexanoyl-L-homoserine lactone-d3 internal standard.*

Concentration level ($\mu\text{g/L}$)	Means (3 replicates)	SD (%)	CV (%)
LLOQ – MRM mode	92 %	3 %	5 %
LLOQ – NL mode	90 %	11 %	10 %
LLOQ – PI mode	88 %	9 %	9 %
ULOQ – MRM mode	95 %	4 %	6 %
ULOQ – NL mode	91 %	8 %	11 %
ULOQ – PI mode	89 %	7 %	13 %

80:20 with a flow rate of 250 $\mu\text{g/L}$. In this case, gradient mode changes to 0:100 in 12 min, then in one minute it returns to initial conditions. Finally, the solvent ratio 80:20 was held for 10 min for re-equilibrating the system. So, for 2-heptyl-4-hydroxyquinoline molecule 22 minutes global chromatographic is necessary.

One aspect that I want to reiterate is the impossibility of combining all the methods mentioned in a single analysis. QTRAP[®] 5500 system allows the analyst to create a global method composed of different experiments. We first tried to create this comprehensive method, but several factors such as loss of sensitivity (e.g., LLOQ in MRM method was 10 $\mu\text{g/L}$) or distortion of the chromatographic peaks forced us to use separate methods to obtain satisfactory results. So, if you desire to explore all the possibilities described so far, each sample is analysed four times: MRM AHLs analysis, NL AHLs analysis, PI AHLs analysis, and MRM 2-heptyl-3-hydroxy-4(1H)-quinolone analysis (the last one mentioned requires changing of solvents too).

All instrument parameters settings are shown in table 7.20, 7.21, and 7.22.

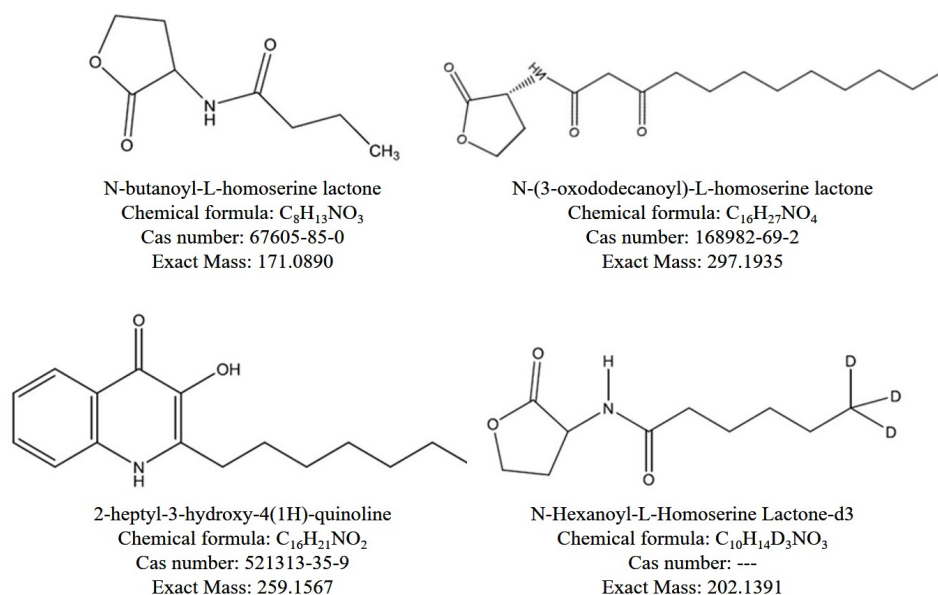


Figure 7.4.: Structures of quinolone signaling molecules used as standards in the present work

Table 7.20.: General instrument parameters used as in all UPLC-TQMS determinations

Instrumental parameters	Values
Curtain gas (CUR)	26
Collision gas (CAD)	Medium
Ion spray voltage (IS)	5500
Ion spray temperature (TEM)	500
Ion source gas 1 (GS1)	45
Ion source gas 2 (GS2)	50
Polarity	Positive

Table 7.21.: Specific AHLs NL and PI scan parameters used. Parameters in first row are related to NL scan method; the ones reported in the second row referring to PI method. Declustering potential; EP: entral potential; CE: collision energy; CXP: collision cell exit potential

Loss of (Da) δm	Start (Da)	Stop (Da)	DP (Volts)	EP (Volts)	CE start (Volts)	CE stop (Volts)	CXP start (Volts)	CXP stop (Volts)
101.000	150.000	350.000	109.000	10.000	15.000	25.000	9.000	13.000
102.000	150.000	350.000	110.000	10.000	15.000	25.000	9.000	13.000

Table 7.22.: *Specific AHLs and HQs MRM scan parameters used. DP: Declustering Potential; EP: Entrance Potential; CE: Collision Energy; CXP: Collision Cell Exit Potential*

Analyte	Precursor ion (Da)	Product Ion (Da)	Time (msec)	DP (Volts)	EP (Volts)	CE (Volts)	CXP (Volts)
N-3-oxo-dodecanoyl-L-homoserine lactone	298.2	240.2	700	109.0	10.0	26.9	12
N-3-oxo-dodecanoyl-L-homoserine lactone	298.2	197.2	700	109.0	10.0	20.9	20
N-3-oxo-dodecanoyl-L-homoserine lactone	298.2	102.1	700	109.0	10.0	15.4	13
N-butanoyl-DL-homoserine lactone	172.1	102.2	70	49.0	10.0	12.0	13.0
N-butanoyl-DL-homoserine lactone	172.1	71.1	70	49.0	10.0	15.0	8.0
2-heptyl-3-hydroxy-4(1H)-quinolone	260.0	188.0	1200	290.0	14.0	42.1	10.0
2-heptyl-3-hydroxy-4(1H)-quinolone	260.0	175.0	1200	290.0	14.0	38.6	27.0
2-heptyl-3-hydroxy-4(1H)-quinolone	260.0	147.0	1200	290.0	14.0	49.2	13.0
N-hexanoyl-L-homoserine lactone-d3	203.2	102.1	70	65.0	6.0	23.0	6.0
N-hexanoyl-L-homoserine lactone-d3	203.2	74.1	70	65.0	6.0	20.0	7.0

7.4. Results and discussion

In the method development, we optimized the chromatographic separation of several lipophilic QS compounds on RP-HPLC columns and studied the sensitivity of different ionization/acquisition modes. The use of TQMS instrumentation is essential and unavoidable given the concentrations $\mu\text{g/L}$ that analytes of interest recorded in plasma human samples analysed.

The first samples we want to expose and discuss are the bacterial cultures. As already anticipated in section 7.1, AHLs species in bacterial cultures are monitored only through NL and PI untargeted scan methods. Considering AHLs species in bacterial samples, all Total Ion Current (TIC) chromatograms of these samples are shown in the figures 7.6÷7.13. As we can see, regarding the results the samples of bacterial cultures grown in poor soil (M9) have as expected fewer peaks than those grown in rich soil (LB). The background noise of the chromatograms acquired in NL results is lower than the ones acquired in PI. The alleged reason why is that the interfering species capable of losing a neutral molecule at 101 m/z are rarer than those capable of producing a fragment ion at 102 m/z . Thanks to these analyses we can monitor the presence of unexpected QS molecules. As a matter of facts, selecting from the chromatograms the value of m/z of the molecular ions $[\text{M}+\text{H}]^+$ of interest it was possible to identify different AHLs molecules thanks to untargeted approach: C4-AHL, C6-AHL, 3-oxo-C10-AHL, 3-oxo-C12-AHL, and C12-AHL. Within these molecules, the 3-oxo-C12-AHL is the one with the highest concentration levels in all the samples analysed. As example, the histogram bar 3D graph reported in figure 7.5 shows clearly how

C12-3oxo-AHL is the most abundant molecule if compared to the ones cited before in *Pseudomonas aeruginosa* PAO1 wild types and mutants grown in LB broth medium.

Furthermore, it has been observed – in three different samples – that C4-HSL has 2 different peaks and C6-HSL has different 3 peaks. This can be explained by assuming the presence of different kinds of isomers within the side chain.

Finally, the elution time of the analytes is consistent too: as the length of the tail (and therefore the hydrophobicity) of AHLs molecules become bigger, the retention time increases as well. On the contrary, if the length of the tail is the same, the presence of a carboxyl group decreases the retention time (3-oxo-C12-HSL is eluted 14 minutes earlier than C12-HSL). As far as AHLs are concerned, MS/MS study evidenced then common losses useful for precursor ion analysis (generation of gamma-butyrolactone-2-ammonium ion) or neutral loss analysis (elimination of 2-amino-gamma-butyrolactone) of N-acyl homoserine lactone QS derivatives.

Considering HQs molecules in bacterial cultures. TIC chromatograms of the samples are shown in the figures 7.14÷7.17. As we can appreciate, for HQs' detection in bacterial cultures only PI experiments are reported. As already anticipated in the introduction section, in the early stages of the study, the detection HQs class was performed via PI scan method monitoring the formation of a radical ion (m/z 175.0633) containing the quinolone bicycle. Later, since for biological samples this method would not have sufficiently sensitive, it was decided to vary the method and considering the more sensitive MRM approach. So, here it is the reason why for plasma samples only MRM chromatograms will be shown and discussed.

So, figures from 7.14÷7.17 show the chromatograms related to the four types of bacterial cultures of *Pseudomonas aeruginosa* PAO1 analysed (wild type and RhlI- mutant; in LB and M9 medium). As expected, the TIC chromatograms of samples of bacterial cultures grown in poor soil (M9) have fewer peaks than those grown in rich soil (LB). As done before, extracting from the TIC chromatograms the m/z values of the molecular ions $[M+H]^+$ different HQs species could be identified with success by this untargeted ap-

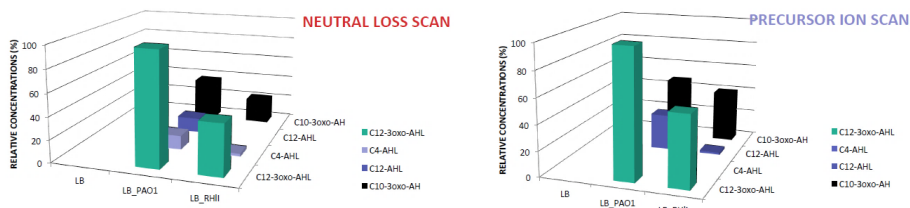


Figure 7.5.: *Quantitation of AHLs produced by wild type vs. mutant Pseudomonas aeruginosa PAO1 in lysogeny broth (LB) medium*

proach: C2-HQ, C3-HQ, C4-HQ, C5-HQ, C6-HQ, C6:1-HQ, C7-HQ, C7:1-HQ, C8-HQ, C8:1-HQ, C9-HQ, C9:1-HQ, C11-HQ, and C11:1-HQ. The elution sequence of the analytes is consistent with AHLs trend: as the length of the tail of HQs increases, as the retention time increases too. As reported before, when the length of the tail being is the same, the presence of an unsaturation decreases the time for the molecule to be eluted by the column (C4:1-HQ is eluted 4 minutes earlier than C4-HQ). From these HQs analysis different aspects rise. Firstly, we can observe that a higher number of HQs species can be detected compared to AHLs PI method. Secondly, both the wild type and the mutant species RhII- of the bacterium *Pseudomonas aeruginosa* PAO1, if grown in the LB rich soil, produces a wide range of HQs, in which the chain length is approximately C2÷C11 and may contain an unsaturation. C7-HQ is the hydroxyquinolone produced the most, while the others are produced in minimal quantities. In poor M9 soil, however, both the wild type and the RhII- mutant species, only the C7-HQ species was detected. In this regard, we advance the hypothesis that the mutated species, unable to produce C4-HSL, intensify the production of this signal molecule.

After discussion of bacteria samples, we move to plasma samples' results. Considering AHLs, the Total Ion Current (TIC) chromatograms of a real sample are shown in the figures 7.18, 7.19, and 7.20. As we can see, the background noise of the chromatograms acquired in NL results is higher than the ones acquired in PI. The possible reason why is that the interfering species capable of losing a neutral molecule at 101 m/z are more abundant

than those capable of producing a fragment ion at 102 m/z . Thanks to these analyses, by extracting from the chromatograms m/z values of the molecular ions $[M+H]^+$, it was possible to identify different AHLs molecules: C4-AHL, 3-oxo-C10-AHL, 3-oxo-C12-AHL, and C12-AHL. The results are consistent with metabolism: the presence of the hydroxy group could be intrinsic to the structure of the signal molecule or have been established by the patient's metabolism. Within these molecules, generally, the C4-AHL is the one with the highest concentration levels in all the samples analysed. For example, looking closely to figure 7.20 we can observe at 21.90 minutes and 27.42 minutes peaks generated by 3-oxo-C10-AHL and C12-AHL species respectively. On the contrary, in figure 7.19 (NL scan method for AHLs analytes' detection) the peak at 27.4 minutes is still present but the peak at 21.9 minutes is disappeared. This aspect is a further proof of the lower sensitivity of NL scan method compared to PI one.

Finally, the elution time of the analytes is consistent: as the length of the tail (and therefore the hydrophobicity) of AHLs molecules become bigger, the retention time increases as well. On the contrary, if the length of the tail is the same, the presence of a carboxyl group decreases the retention time (3-oxo-C12-HSL is eluted 14 minutes earlier than C12-HSL). As far as AHLs are concerned, MS/MS study evidenced then common losses useful for precursor ion analysis (generation of gamma-butyrolactone-2-ammonium ion) or neutral loss analysis (elimination of 2-amino-gamma-butyrolactone) of N-acyl homoserine lactone QS derivatives.

If we look now at MRM chromatograms' samples, they exhibit, compared to NL and PI ones, advantages in terms of high sensitivity and accuracy, and they well suited for the detection of low AHLs and HQs molecules. In fact, MRM chromatograms have a background noise practically nil, and they are a clear demonstration, thanks also to data obtained during validation, of the ability of MRM methods to detect quantities of a few $\mu\text{g/L}$ of quorum sensing molecules.

Figures 7.22, 7.23, and 7.24 represent 300 $\mu\text{g/L}$ standards' chromatograms for MRM, NL, and PI AHLs molecules. We focus now our attention on the intensity, expressed in cps, of each AHLs' chromatogram. If we divide the intensity of the most intense peak present in each chromatogram by

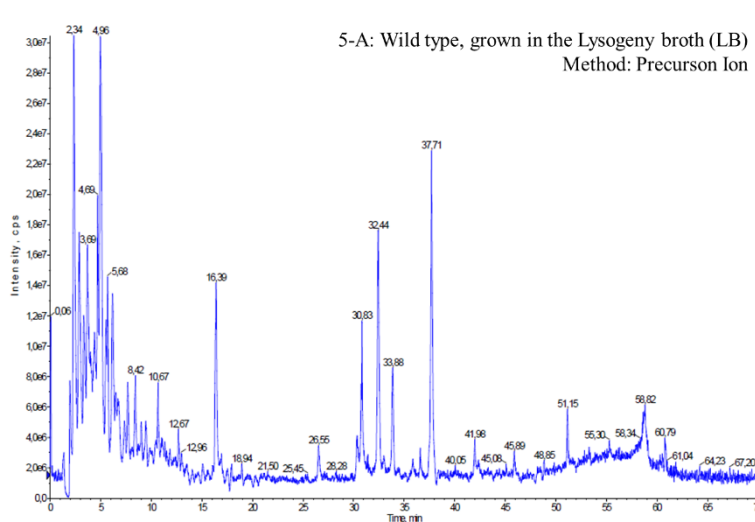


Figure 7.6.: *Example of Total Ion Current chromatograms of AHLs molecules in wild type bacteria cultures' samples; LB broth; PI method*

its background noise, we have further proof of how the MRM mode is more sensitive. The ratios for MRM, NL, and PI cases are 1.00×10^6 , 1.00×10^1 , and 1.00×10^1 respectively.

Moving on to HQs molecules, the TIC chromatograms of the samples are shown in the figure 7.21 (plasma of AKI patient under hemoperfusion treatment) and figure 7.25 ($300 \mu\text{g}/\text{L}$ standard sample). Surprisingly, although few known bacterial species use this signal molecules class for quorum sensing mechanism, C7-HQ species is detected among all the samples under exam. This aspect paves the way for future analyses aimed at quantifying other HQs species (in which the chain length is approximately from C2 to C11) into different matrices.

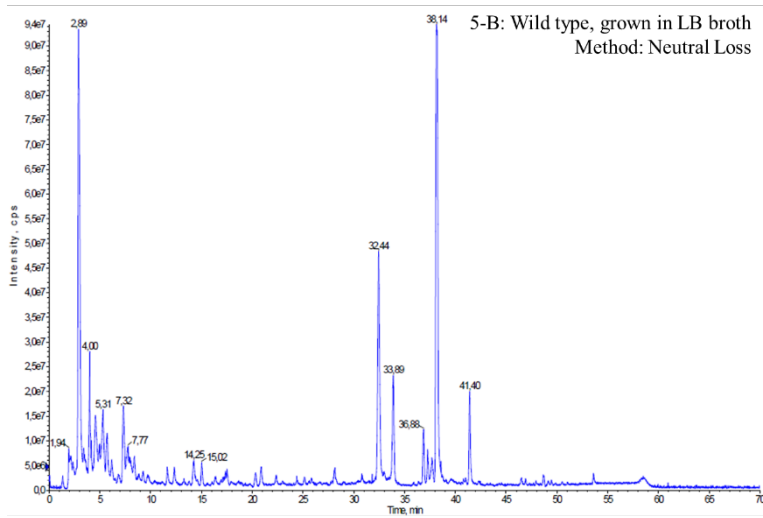


Figure 7.7.: Example of Total Ion Current chromatograms of AHLs molecules in wild type bacteria cultures' samples; LB broth; NL method

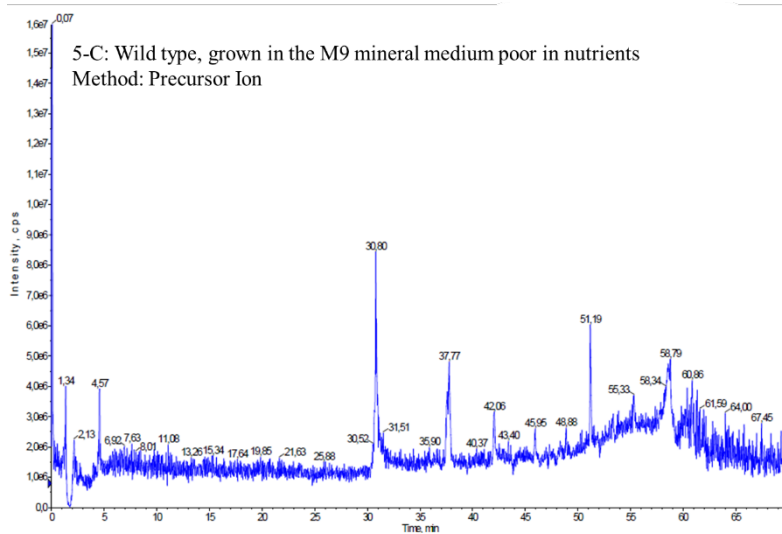


Figure 7.8.: Example of Total Ion Current chromatograms of AHLs molecules in wild type bacteria cultures' samples; M9 broth; PI method

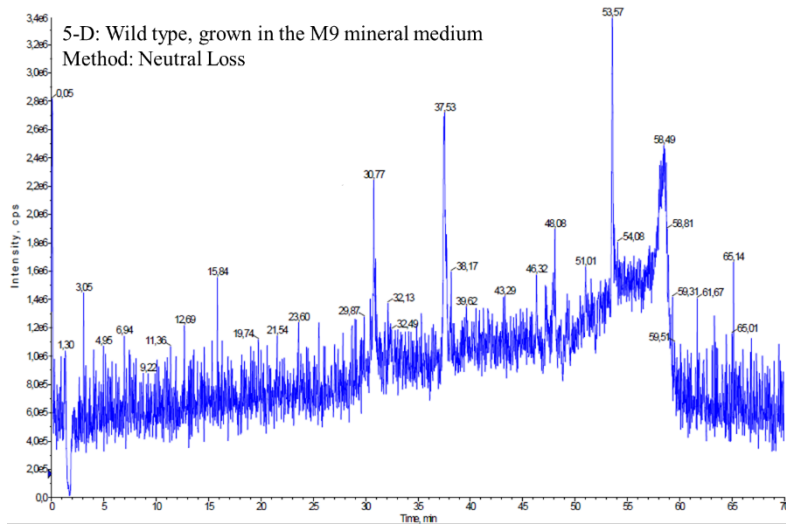


Figure 7.9.: Example of Total Ion Current chromatograms of AHLs molecules in wild type bacteria cultures' samples; M9 broth; NL method

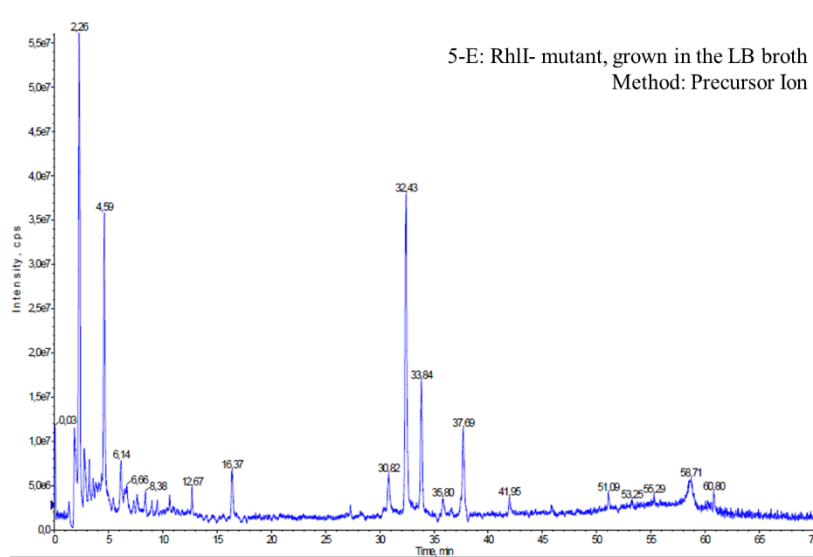


Figure 7.10.: Example of Total Ion Current chromatograms of AHLs molecules in RHLI- mutant bacteria cultures' samples; LB broth; PI method

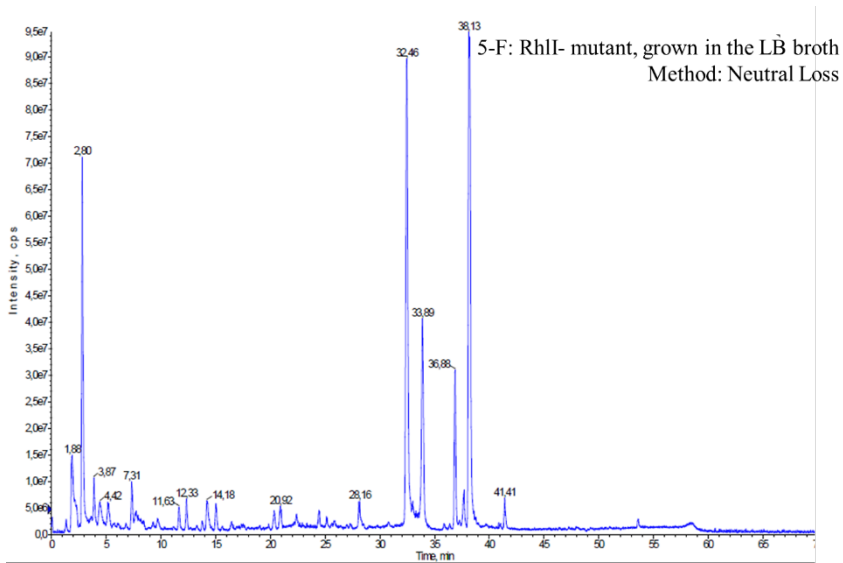


Figure 7.11.: Example of Total Ion Current chromatograms of AHLs molecules in RhII- mutant bacteria cultures' samples; LB broth; NL method

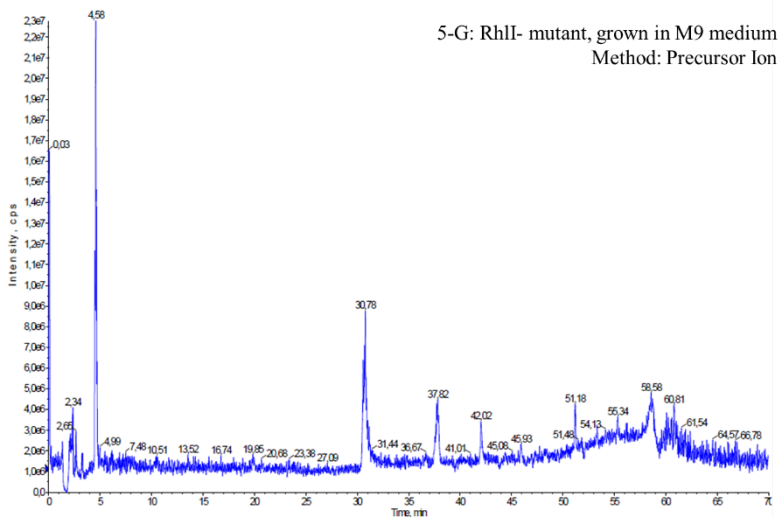


Figure 7.12.: Example of Total Ion Current chromatograms of AHLs molecules in RhII- mutant bacteria cultures' samples; M9 broth; PI method

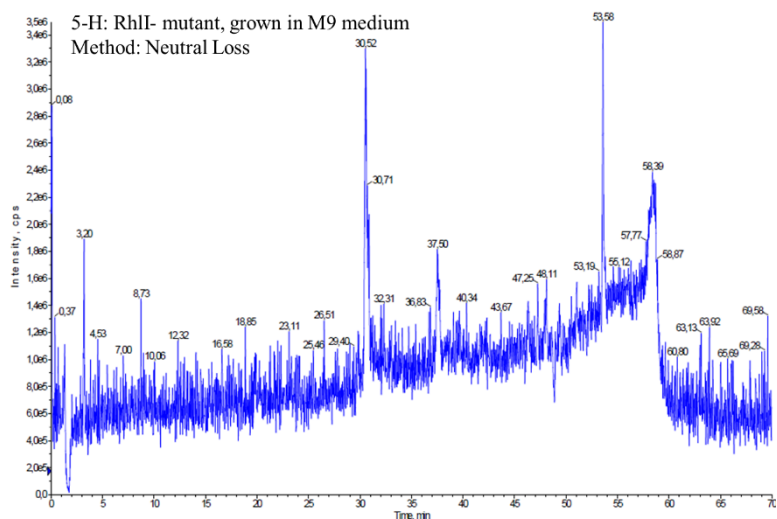


Figure 7.13.: Example of Total Ion Current chromatograms of AHLs molecules in RhlI- mutant bacteria cultures' samples; M9 broth; NL method

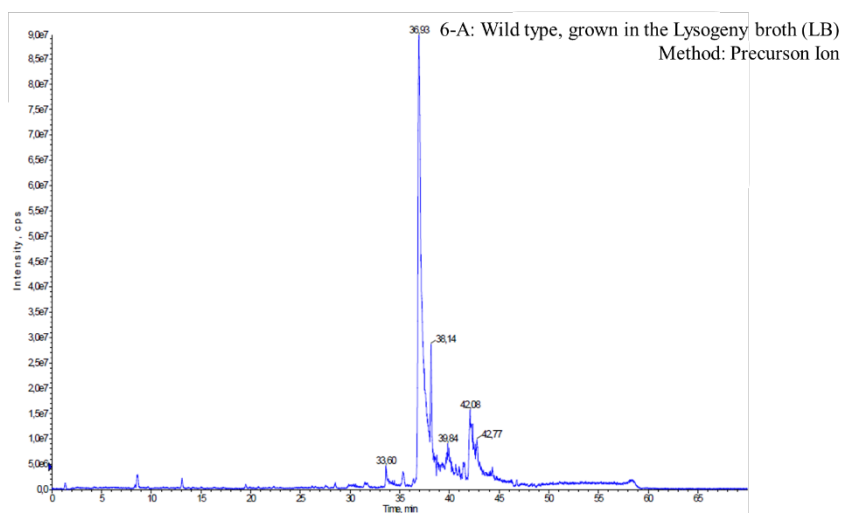


Figure 7.14.: Example of Total Ion Current chromatograms of HQs molecules in wild type mutant bacteria cultures' samples; LB broth; PI method

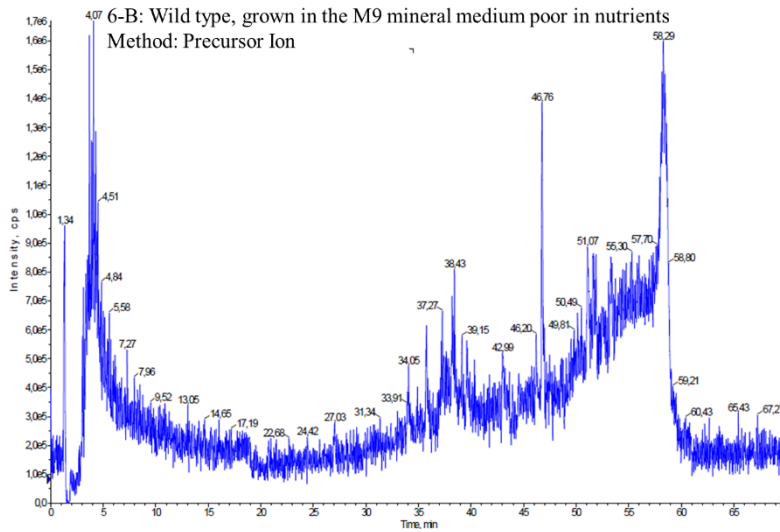


Figure 7.15.: Example of Total Ion Current chromatograms of HQs molecules in wild type mutant bacteria cultures' samples; M9 broth; PI method

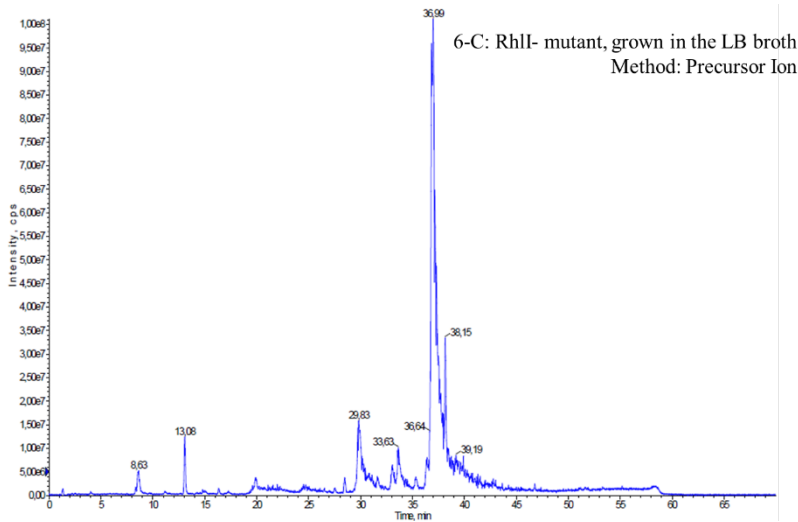


Figure 7.16.: Example of Total Ion Current chromatograms of HQs molecules in RhlI- mutant bacteria cultures' samples; LB broth; PI method

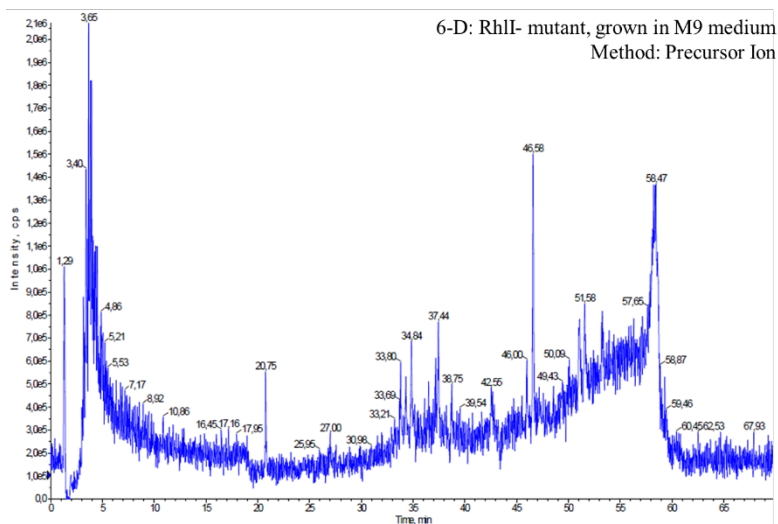


Figure 7.17.: Example of Total Ion Current chromatograms of HQs molecules in RHLI-mutant bacteria cultures' samples; M9 broth; PI method

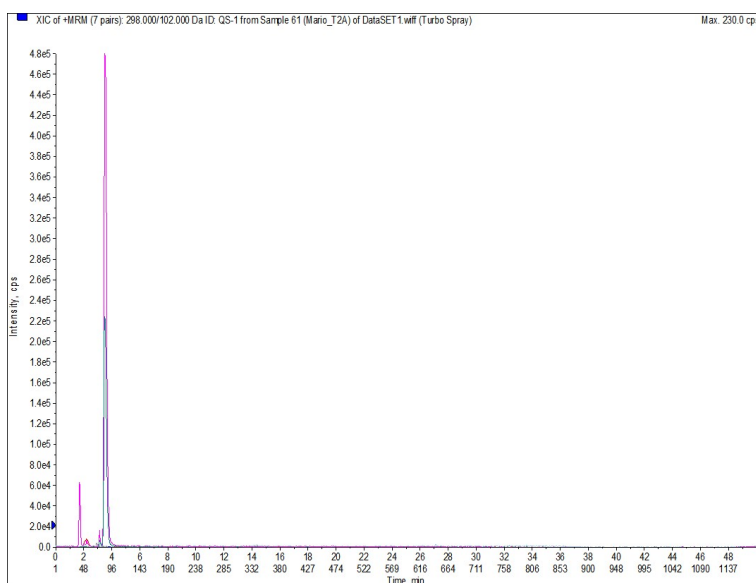


Figure 7.18.: Example of Total Ion Current chromatograms of AHLs molecules in real plasma sample via MRM method

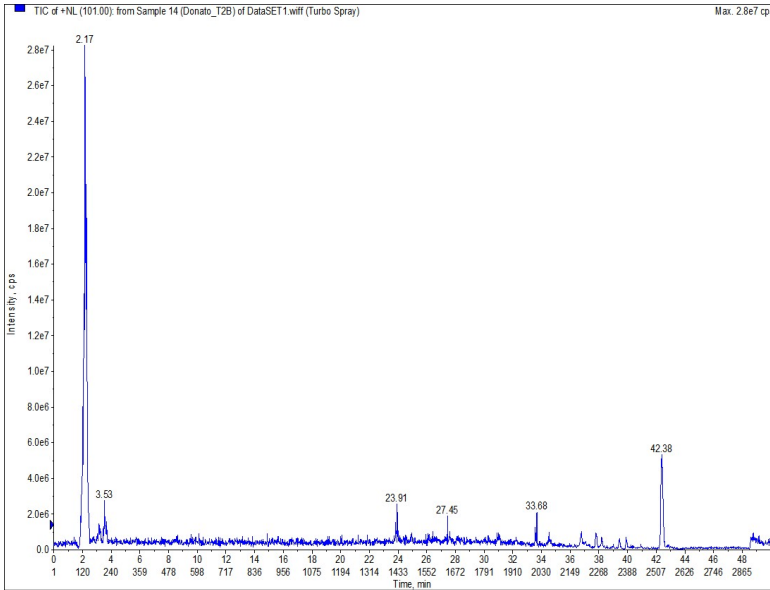


Figure 7.19.: Example of Total Ion Current chromatograms of AHLs molecules in real plasma sample via NL method

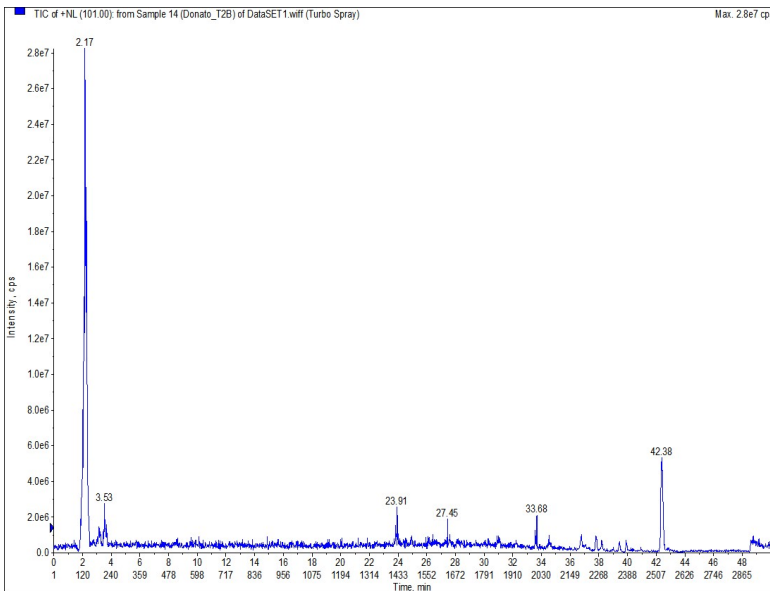


Figure 7.20.: Example of Total Ion Current chromatograms of AHLs molecules in real plasma sample via PI method

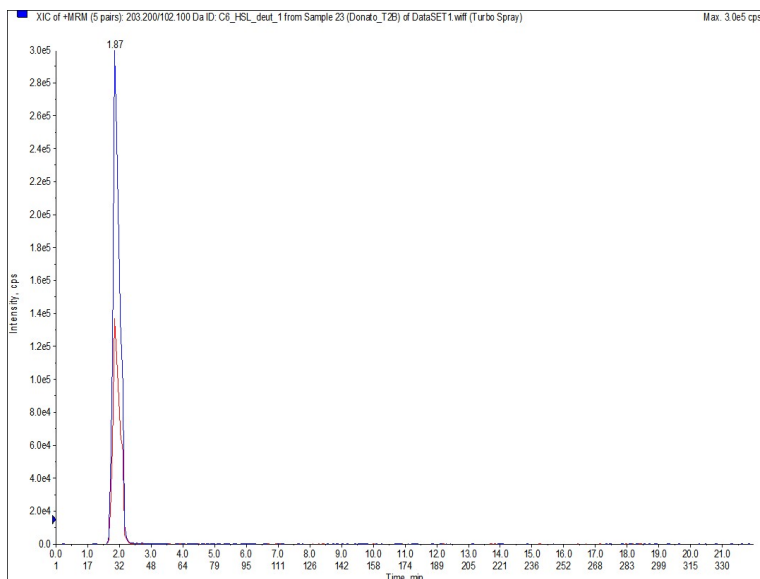


Figure 7.21.: Example of Total Ion Current chromatograms of HQs molecules in real plasma sample via MRM method

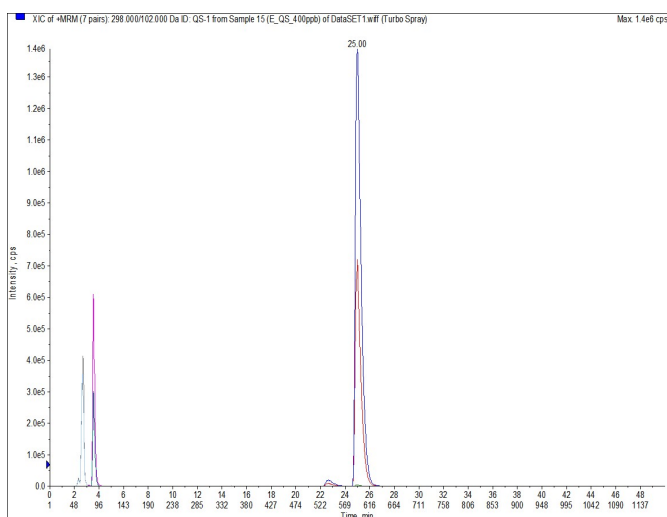


Figure 7.22.: Example of Total Ion Current chromatograms of AHLs molecules in standard sample via MRM method

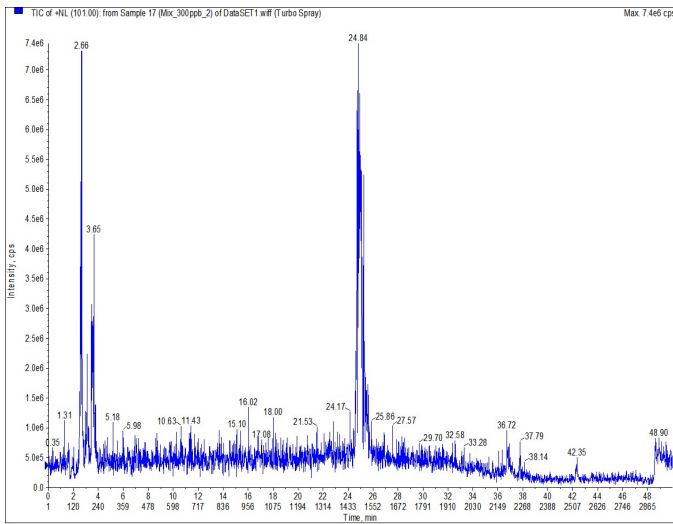


Figure 7.23.: Example of Total Ion Current chromatograms of AHLs molecules in standard plasma sample via NL method

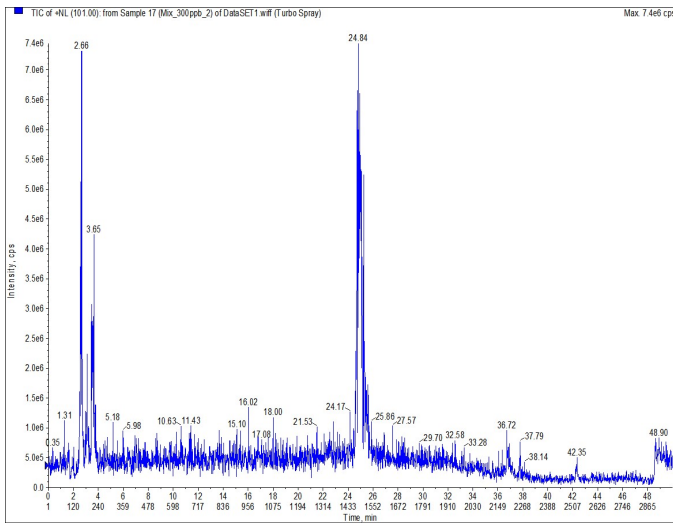


Figure 7.24.: Example of Total Ion Current chromatograms of AHLs molecules in standard plasma sample via PI method

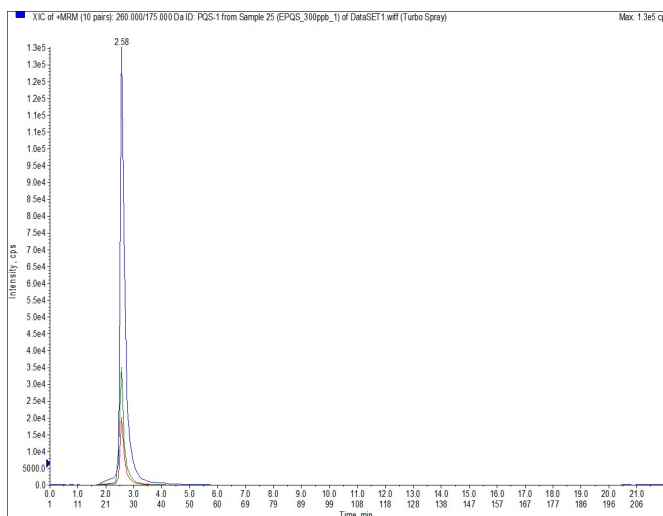


Figure 7.25.: *Example of Total Ion Current chromatograms of HQs molecules in standard plasma sample via MRM method*

7.5. Conclusion and future prospects

In conclusion, a fast and selective UPLC-TQMS has been developed and applied, in an early stage of the study, to wild-type and mutant bacteria samples and, in a second stage, to biological plasma samples. Thanks to untargeted NL and PI methods and MRM targeted ones different AHLs and HQs molecules has been analysed.

The UPLC-TQMS/MS analytical procedure has been fully validated using MRM, NL, and PI scan methods. All the MS/MS approaches show to be fast and selective for the rapid quantification of bacterial small molecules in biological fluids. Preliminary experimental data indicate that the method is suitable for the detection of low concentration of AHLs and HQs in bacterial cultures and biofluids in the early stages of the illness. As a matter of fact, using the NL and PI scan methods it is possible to discover new kinds of species whose presence within the sample was not predictable at the beginning of the analysis. On the contrary, as supported by the results previously showed, MRM approach is suitable for the detection of low AHLs and HQs levels within biological samples. In addition, the comparison between bacteria cultures samples and biological samples highlights even more

how high sensitivity methods are required for the plasma samples, given the concentration of few $\mu\text{g}/\text{L}$ of the molecules of interest within the sample.

As for future perspectives, different questions must be examined. First, the validation procedure of NL and PI approached for HQs species need to be investigated; secondly, which molecules can be studied through AHLs and HQs MRM scan methods to unravel new quorum sensing mechanism; finally, verify if the methods discussed in the present work could be applied to other samples like urine or blood samples.

References of Chapter 7

- [1] European Medicines Agency. *ICH Topic Q 2 (R1) Validation of Analytical Procedures: Text and Methodology*. European Medicines Agency, 1995 (cit. on p. 101).
- [2] T. and others Bjarnsholt. “Why chronic wounds will not heal: A novel hypothesis”. In: *Wound Repair and Regeneration* 16, 2–10 (2008). DOI: 10.1111/j.1524-475X.2007.00283.x (cit. on p. 95).
- [3] S. R. Campagna et al. “Direct quantitation of the quorum sensing signal, autoinducer-2, in clinically relevant samples by liquid chromatography tandem mass spectrometry”. In: *Analytical Chemistry* 81, 6374–6381 (2009). DOI: 10.1021/ac900824j (cit. on p. 96).
- [4] X. Chen et al. “Structural identification of a bacterial quorum-sensing signal containing boron”. In: *Nature* 415, 545–549 (2002). DOI: 10.1038/415545a (cit. on p. 95).
- [5] S. De Bentzmann et al. “The *Pseudomonas aeruginosa* opportunistic pathogen and human infections”. In: *Environmental Microbiology* 13, 1655–1665 (2011). DOI: 10.1111/j.1462-2920.2011.02469.x (cit. on p. 96).
- [6] T. R. De Kievit et al. “Role of the *Pseudomonas aeruginosa* las and rhl quorum-sensing systems in rhlI regulation”. In: *Journal of Bacteriology* 212, 106–111 (2002). DOI: 10.1128/jb.179.18.5756-5767.1997 (cit. on p. 96).
- [7] Eurachem. “The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics”. In: *Eurachem publication* 1 (2014). DOI: 978-91-87461-59-0 (cit. on p. 101).
- [8] R.S. Grant et al. “The Physiological Action of Picolinic Acid in the Human Brain”. In: *International Journal of Tryptophan Research* 2, 71–79 (2009). DOI: 10.4137/ijtr.s2469 (cit. on p. 112).
- [9] LumenWayMaker. *Signaling in Bacteria*. 2020. URL: <https://courses.lumenlearning.com/wm-biology1/chapter/reading-signaling-in-bacteria/> (cit. on p. 97).

- [10] A. Mohammad et al. “Quorum sensing: A noble target for antibacterial agents”. In: *Avicenna Journal of Medicine* 2(4), 97–99 (2012). DOI: 10.4103/2231-0770.110743 (cit. on p. 95).
- [11] K. H. Nealson et al. “Bacterial bioluminescence: its control and ecological significance”. In: *Microbiology and Molecular Biology Reviews* 43, 496–518 (1979) (cit. on p. 95).
- [12] R. P. Novick et al. “Quorum Sensing in Staphylococci”. In: *Annual Review of Genetics* 42, 541–564 (2008). DOI: 10.1146/annurev.genet.42.110807.091640 (cit. on p. 95).
- [13] C. A. Ortori et al. “Simultaneous quantitative profiling of N-acyl-homoserine lactone and 2-alkyl-4(1H)-quinolone families of quorum-sensing signaling molecules using LC-MS/MS”. In: *Analytical and Bioanalytical Chemistry* 399, 839–850 (2011). DOI: 10.1007/s00216-010-4341-0 (cit. on pp. 96, 97).
- [14] C. T. Parker et al. “Cell-to-cell signalling during pathogenesis”. In: *Wound Repair and Regeneration* 11, 363–369 (2009). DOI: 10.1111/j.1462-5822.2008.01272.x (cit. on p. 95).
- [15] J. P. Pearson et al. “Roles of *Pseudomonas aeruginosa* las and rhl quorum-sensing systems in control of twitching motility”. In: *Journal of Bacteriology* 197, 5756–5767 (1997) (cit. on p. 96).
- [16] E.C. Pesci et al. “Quinolone signaling in the cell-to-cell communication system of *Pseudomonas aeruginosa*”. In: *Proceedings of the National Academy of Sciences* 96, 11229–11234 (1999). DOI: 10.1073/pnas.96.20.11229 (cit. on p. 101).
- [17] E. Rogatsky et al. “Bioanalytical Method Validation Guidance”. In: *Journal of Chromatography B Anal. Technol. Biomed. Life Sci* 1043, 25 (2017) (cit. on p. 101).
- [18] P. Turnpenny et al. “Bioanalysis of *Pseudomonas aeruginosa* alkyl quinolone signalling molecules in infected mouse tissue using LC-MS/MS; and its application to a pharmacodynamic evaluation of MvfR inhibition”. In: *Journal of Pharmaceutical and Biomedical Analysis* 139, 44–53 (2017). DOI: 10.1016/j.jpba.2017.02.034 (cit. on p. 112).

- [19] K. Vendeville A. and Winzer et al. “Making ‘sense’ of metabolism: Autoinducer-2, LuxS and pathogenic bacteria”. In: *Nature Reviews Microbiology* 3, 383–396 (2005). DOI: 10.1038/nrmicro1146 (cit. on p. 96).
- [20] C. M. Waters et al. “QUORUM SENSING: Cell-to-Cell Communication in Bacteria”. In: *Annual Review of Cell and Developmental Biology* 21, 319–346 (2005). DOI: 10.1146/annurev.cellbio.21.012704.131001 (cit. on p. 95).
- [21] F. Xu et al. “Quantitative determination of AI-2 quorum-sensing signal of bacteria using high performance liquid chromatography–tandem mass spectrometry”. In: *Journal of Environmental Science* 52, 204–209 (2017). DOI: 10.1016/j.jes.2016.04.018 (cit. on pp. 95, 96).

Targeted analysis: Lipids and hepatic steatosis in liver mice

8

8.1. Introduction

The hepatic steatosis disease

The human liver is a 1.5 kilograms gland, the largest of the human body [22]. Its functions are multiple and mostly essential: regulation of glycogen storage, decomposition of red blood, or cells. Among these, the role of primary importance is covered by the sorting and synthesis of fats [11, 22]. When the lipid content of the liver exceeds 5% of its weight, the accumulation of triglycerides and sphingosines inside the hepatocytes increases. Consequently, an inflammation of the liver called fatty liver or hepatic steatosis rises [2, 12].

Many liver disorders show up as a common denominator the steatosis. These disorders can be classified according to the etiology [24]:

1. steatosis from increased fat intake, caused for example by hyperlipidic diet [10] or diabetes mellitus [1];
2. steatosis from increased endogenous fat synthesis, which alcohol abuse [19] or use of barbiturates [20] are the main culprits;
3. steatosis from reduced fat disposal and burning, linked to hypoproteic diet [6], vitamin B12 deficiency [21] etc.

In scientific articles and medical fields, the first steatosis cited is defined by the acronyms NAFLD (Non-Alcoholic Fatty Liver Disease) or NASH (Non-Alcoholic Steato Hepatitis). On the contrary, type 2 steatosis is called AFLD -Alcoholic Fatty Liver Disease. In the end, type 3 is much rare, and it's not generally identified by an acronym [12, 19]. It is important to remark that, over time, an inflamed liver can become fibrous and hardened. This condition, which is universally known as hepatic cirrhosis [22], is a decisive step towards liver failure and death, and it's not only linked to alcohol abuse.

The liver shows malaise signs only at a very advanced illness stage [12]. As a matter of fact, because of this asymptomatic nature, more than 90% of people with fatty liver occasionally discover this disorder [14]. In Italy and Western European countries, the hepatic steatosis syndrome affects 25% of the standard population and rises 50-90% levels in overweight-obesity subjects [17]. Therefore, monitoring this disease and the causes that generate it becomes of primary importance for human health to predict metabolic dysfunctions [4].

Sphingolipids and sugar diets in hepatic steatosis disease

Ceramide is considered the hinge of sphingolipid metabolism (figure 8.1), and an enzymatic equilibrium is continually respected to maintained correct levels of sphingosine, sphingosine-1-phosphate, and ceramides. In addition to ceramide, in the diagram shown in figure 8.1, another critical role is that of sphingosine-1-phosphate. This molecule has a bioactive receptor role, and most of the effects mediated by this molecule derive from its interaction with specific receptors located onto plasma membranes: nowadays, five have been identified (generally called SPR, sphingosine-1-phosphate receptors).

Levels of sphingosine-1-phosphate within the cell are modulated through the enzyme's regulation in biosynthetic and degradative pathways. Therefore, the sphingosine-1-phosphate levels can potentially be altered by influencing the expression of its synthesis enzyme, the sphingosine kinase (SK) [13]. The SK1 isoform is involved in the inflammatory response or steatosis diseases such as non-alcoholic steatohepatitis [7]. In metabolic disorders induced by unbalanced diets, the altered balance of the ceramide / sphingosine / sphingosine-1-phosphate system contributes to the onset of cellular metabolic damage. In a diet rich in saturated fatty acids, high levels of ceramides are observed. These molecules interfere with mitochondrial metabolism and induce the activation of SREBPs1 factors [2]. Furthermore, ceramides activate apoptosis and interfere with cell differentiation, senescence, migration, and adhesion.

The correction of alterations that compromise cellular functionality is possible by inhibiting the activity of SK1 or enhancing that of ceramidase. The AGE/RAGE signal pathway can activate SK1 and inhibit ceramidase

[18]: therefore, by acting on AGE with anti-glycating products, such as pyridoxamine, there is a reduction in AGE production and overexpression of RAGE in tissues and plasma. Consequently, the action of SK1 should be inhibited, favouring the production of sphingosine [23].

8.2. Aim of the work

This project has taken the first steps as an evolution of previous researches carried out at the Department of Clinical and Biological Sciences of the University of Turin by Professor Manuela Aragno and Professor Raffaella Mastrocola. Their research's activity in this field has already led to exciting publications [15, 16].

This project aimed to clarify the effects of sugary drinks consumption over time. To do that, mice received 15% (wt/vol) fructose or 15% (wt/vol) glucose in water to drink for 30 weeks, resembling human habit to consume sugary drinks [16]. The objective data obtained were increased levels for glycemia, glucose intolerance, and the appearance of hepatic steatosis in fructose-drinking mice. Furthermore, the effects of a fructose-rich diet (HFRT) were similar to the ones obtained through a fat-rich diet (HFAT). The comparison between oil red-coloured histological samples is reported in figure 8.2.

The liver of mice subjected to a standard diet is morphologically healthy (the number of lipid vesicles is adequate). The livers of mice fed with HFAT and HFRT diets show a very high inclusion of triglycerides in the liver tissue; moreover, the accumulation morphology in the two diets is different: the fat vesicles in the liver HFAT are smaller and more distributed; on the contrary, HFRT accumulations are less spread and more abundant.

So, to clarify which sphingolipids characterize the lipid vesicles, the purpose of this new work is the development of two different analytical methods for the detection and the quantification of triacylglycerols, ceramides, and sphingosines in cases of induced non-alcoholic hepatic steatosis in mice. To monitor triacylglycerols levels, an HPLC-HRMS (LTQ OrbitrapTM platform) method was built up; au contraire, to detect ceramides and sphin-

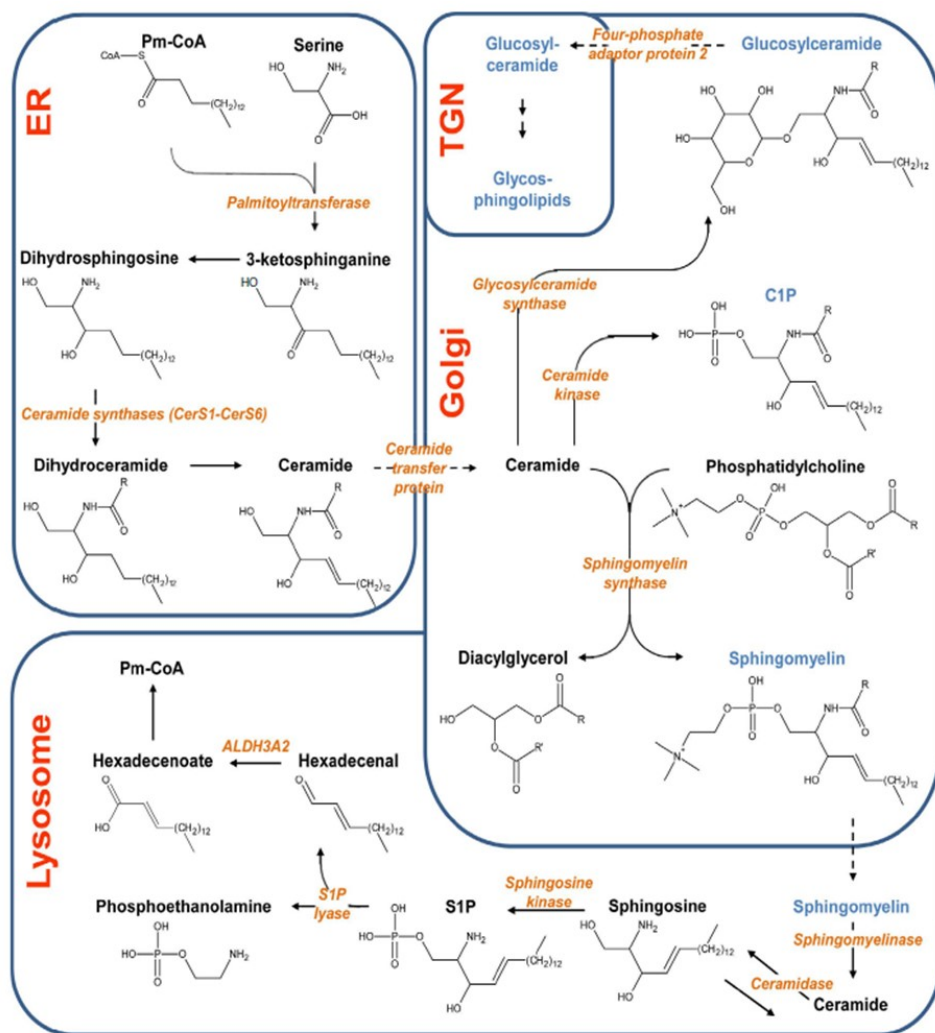


Figure 8.1.: Pathways of the analytes object of studies and relative cellular sections in which they take place. Adapted by [9]

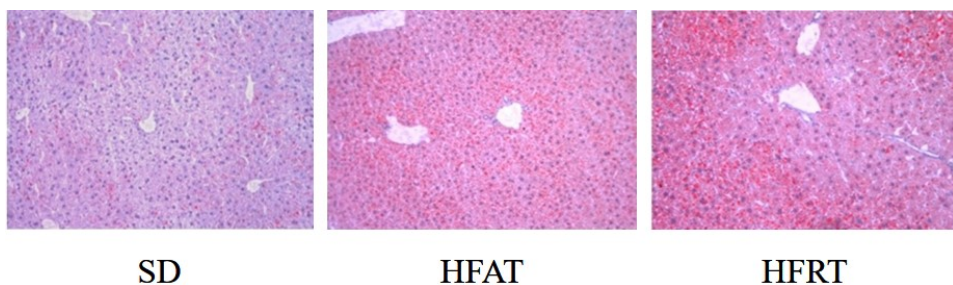


Figure 8.2.: *Oil red coloured histological sample of livers of mice subjected to three different diets. Figures are obtained by Professors Aragno's research group*

gossines molecules, an UPLC-TQMS QTRAP[®] 5500 system was used. The project in its entirety includes also the development of a GC-MS method for the quantification of derivatized fatty acids. Since the main purpose of this thesis is the discussion of results obtained through LC-MS techniques, the ones obtained via GC-MS platform are not reported.

The non-alcoholic hepatic steatosis samples were obtained by feeding three groups of mice with three different but isocaloric diets (more details will be provided in section 8.3):

- standard diet (SD);
- fat-rich diet (HFAT) in which 40% of the caloric intake derives from lard, consisting almost entirely of triglycerides of stearic acid;
- fructose-rich diet (HFRT) in which 60% of caloric intake is provided by fructose¹.

From the researches previously carried out by Professor Manuela Aragno and Professor Raffaella Mastrocola, they have seen how the mice livers fed with HFAT and HFRT diets have a very high inclusion of triglycerides in

¹A question may arise. Why can a sugar-based diet have the same effects as a fat-based diet in the onset of hepatic steatosis? Fructose is a monosaccharide; therefore, it cannot be stored as it is in the liver. When the organism has an energy surplus, the liponeogenesis pathway is then activated. The cell firstly converts the sugars into acetyl-CoA, then into palmitic acid, and finally converts it into ceramides and phosphatidylcholines. Over time, the fats produced by this metabolic pathway accumulate in the liver, generating steatosis caused by a decrease in β -oxidation. Excessive fructose consumption, however, is not only accompanied by hepatic steatosis but also cell oxidative stress, inflammatory response, and insulin resistance.

the liver tissue. So, the first objectives of this work are to reveal which of these triglycerides are present and to identify which of them can be used as a marker in different types of steatosis.

The reason why two different analytical instruments are applied is to be found in the just listed objectives. In the ceramides and sphingosines groups (which are subclasses of the sphingolipids family), only some molecules present at low concentrations are interesting to monitor from our point of view. So, an UPLC-TQMS platform provides the best solution to our problems. On the contrary, in the triacylglycerols analysis an untargeted approach via HPLC-HRMS platform is suitable. As a matter of fact, in the liver samples, the presence of molecules belonging to the class of triglycerides is well documented, but their exact composition is not fully known.

Following these unbalanced diets, the increase of lipids in plasma and tissues leads to an altered equilibrium of the ceramide-sphingosine-sphingosine-1-phosphate system, which in turn contributes to the onset of metabolic damage to the cellular level. The attention is paid in particular on sphingosine-1-phosphate, to understand its role in the activity of enzymes that regulate the equilibrium of the ceramide/sphingosine system. It is possible to correct the alterations that compromise cellular functionality by inhibiting the activity of sphingosine kinase 1² or enhancing that of ceramidase. The AGE/RAGE signal pathway is able to activate SK1 and to inhibit ceramidase. Therefore, by acting on AGEs with anti-glycating products, such as pyridoxamine, there is a reduction in AGE production and over-expression of RAGE in the tissues and plasma of mice fed with the HFAT diet, with consequent inhibition of SK1 action, favouring the production of sphingosine. Pyridoxamine was added to the diet in part of the individuals in both groups. Consequently, a further objective (described in the conclusion section) linked to the sphingosines' detection is to prove how sphingosine-1-phosphate levels decrease in the liver in the case of mice fed with pyridoxamine.

²The sphingosine kinase 1 or SK1 is a human enzyme present in cytosolic protein able to phosphorylate sphingosine molecule to sphingosine-1-phosphate

Table 8.1.: *Physiological parameters of mice fed with different diets. SD: Standard diet, HFAT: fat-rich diet, HFRT: fructose-rich diet*

Biological parameter	SD	HFAT	HFRT
Daily Caloric Intake (Kcal/Die/Mouse)	9.6 ± 0.4	12.1 ± 0.7	10.2 ± 0.3
Fats Caloric Intake (Kcal/Die/Mouse)	0.43 ± 0.05	5.46 ± 0.32	0.46 ± 0.01
Sugars Caloric Intake (Kcal/Die/Mouse)	0.00 ± 0.00	2.06 ± 0.12	6.13 ± 0.26
Increase in body weight (g)	8.5 ± 1.6	17.2 ± 4.6	8.3 ± 1.5
Liver weight (% of the body weight)	3.5 ± 0.2	2.7 ± 0.4	3.6 ± 0.1
Epididymis weight (% of the body weight)	1.2 ± 0.4	6.1 ± 0.9	1.5 ± 0.4
Systolic pressure (mm Hg)	112 ± 7	106 ± 11	113 ± 3
Glycemia (mg/dL)	69 ± 19	120 ± 21	91 ± 11
Hyperinsulinemia (mg/dL)	85.8 ± 5.3	106.0 ± 12.6	86.6 ± 7.1
Plasma-triglycerides (mg/dL)	30.7 ± 6.6	86.3 ± 27.8	45.5 ± 5.0
Plasma-cholesterol (mg/dL)	77.2 ± 5.7	267.1 ± 94.6	115.3 ± 12.5

8.3. Sample preparation and method settings

The liver samples under study come from mice that have been used for twelve weeks on various and isocaloric diet regimens. As already reported in the previous section 8.2, the diets were a standard diet (SD), a fat-rich diet (HFAT), and a fructose-rich diet (HFRT). The table 8.1 shows in the first three lines parameters concerning the regular consumption and total caloric intake of different diets. Comparing those values, there are no significant differences in the daily consumption of food, and the total caloric intake. However, it's important to remark how, in the HFAT and HFRT, diets almost all of the daily calorie intake belong from fats (lard) and sugars (fructose), respectively. Looking at the other rows of the table, it's clear how HFAT and HFRT diets have resulted in significant alterations in glycidic and lipid metabolism that can be detected as an increase in fasting glycemia, appearance of hyperinsulinemia, and increase in plasma triglyceride and cholesterol levels. The alterations are much more marked in the HFAT then in the HFRT diet. Furthermore, only the HFAT diet determines a marked increase in body weight and the relative quantity of epididymal adipose tissue, associated with a significant reduction in liver mass, compared to the SD diet.

All the feeding process and the measurements of these biological parameters here conduct by Prof. Aragno and Prof. Mastrocola's group at the

Department of Clinical and Biological Sciences of the University of Turin [15, 16]. Once the biological parameters were taken, mice were sacrificed. Then, the homogenated livers were frozen with liquid nitrogen and sent to our Laboratory where they were stored at -25°C until analysis.

From now on, this present chapter and the next one will be divided into two separate sections according to the different platforms used to perform analyses: the HPLC-HRMS section and the UPLC-TQMS section. As mentioned above, onto the same liver samples, two different approaches were applied to maximize the information obtainable from these matrices.

Sample preparation and method settings: HPLC-HRMS

The triglycerides (TAG) were separated from the homogenate liver through a liquid phase extraction:

1. 200 μL of homogenate liver (furnished already prepared by the Department of Clinical and Biological Sciences of the University of Turin) are added in a 10 mL pyrex tube. The sample's container should be in pyrex or, at least, it must be made of a material that, in contact with chloroform, is not dissolved.
2. Mix 1 mL of chloroform/methanol in a proportion of 2:1 to the sample;
3. Shake on the vortex for 30 seconds to homogenate the matrix;
4. Centrifuge at 4500 g for 15 minutes at room temperature;
5. Withdraw the supernatant separated from the matrix and place it in a glass vial for the analysis.

The HPLC parameters and the MS settings are reported in tables 8.2 and 8.3.

Sample preparation and method settings: UPLC-TQMS

The sphingolipids were extracted from the matrix through a liquid-liquid extraction requiring numerous analytical steps, affecting the final extraction yield. The liquid-liquid extraction used is a slightly variation of Folch's

Table 8.2.: *Instrumental equipment used in the HPLC-HRMS TAG analysis*

HPLC	Ultimate 3000 Dionex Chromatographic column Phenomenex Luna 3 μm C18(2) 100 \AA , LC Column 100 x 4.6 mm
HRMS	LTQ Orbitrap TM Thermo Scientific
Ionization source and parameters	APCI
	Capillary temperature 250° C
	Vaporizer temperature 450° C
	Flow rate sheath gas 30.0 arbitrary units (a.u.)
	Flow rate auxiliary gas 15.0 arbitrary units (a.u.)
	Capillary voltage 10.0 V
	Source voltage 4.1 kV
	Tube lens 40 V
	Discharge current 5 MA

Table 8.3.: *Chromatographic and MS conditions - LTQ OrbitrapTM Thermo Scientific in TAG analysis*

Injection	5 μL		
Flux	200 $\mu\text{L}/\text{min}$		
Chromatographic run	Time (min)	Acetonitrile-Dichloromethane (95:5 v/v)	Dichloromethane-Acetonitrile (95:5 v/v)
	0	80 %	20 %
	30	80 %	20 %
	40	50 %	50 %
	45	0 %	100 %
	50	0 %	100 %
	51	80 %	20 %
61	80 %	20 %	
Polarity	Positive		
Mass range	200 - 1500 m/z		
Resolution	30000 (FWHM)		

method [5]. Folch's extraction procedure is one of the most popular methods for isolating lipids from biological samples [3]. Our extraction method includes the following steps:

1. Place 50 μL of the homogenate liver in a 10 mL test tube;
2. Add to the sample 2 mL of 0.1% trifluoroacetic acid in chloroform/methanol (ratio 1:2);
3. Positivize the sample by using the three isotope-labelled standards at a concentration of 250 $\mu\text{g/L}$;
4. Shake 30 seconds on Vortex;
5. Add 0.5 mL of chloroform;
6. Shake 30 seconds on Vortex;
7. Add 0.5 mL of water;
8. Shake 30 seconds on Vortex;
9. Centrifuge the sample 15 minutes at 3500 g, room temperature;
10. After the centrifugation step, three different phases are present: the organic phase (subnatant), the liver residue (floating at the interface), and the aqueous supernatant. Withdraw the organic phase from the bottom of the test tube and place it in a new one;
11. Re-extract the aqueous phase twice. Use the same conditions as before (Vortex time, centrifugation settings), but instead of 0.5 mL, add 1 mL of chloroform as extraction solvent. After each re-extraction procedure collect the organic phase in the same tube;
12. Bring to dryness under N_2 flow at room temperature the organic phases collected;
13. Add to the dry residue 0.1 mL of chloroform/methanol solution (ratio 9:1);
14. Transfer to a glass vial for analysis.

Table 8.4.: *Instrumental equipment applied for UPLC-TQMS lipids analyses*

UPLC	Shimadzu Nexera X2 UPLC Column Phenomenex 1.7 μm Kinetex EVO C18 100 \AA , LC Column 50 x 2.1 mm
TQMS	SCIEX QTRAP [®] 5500 system
Ionization source	ESI (Turbo ion spray $\text{\textcircled{C}}$)

This research project has been underway for more than two years. During this time, the TQMS analytical method was gradually improved, both in terms of sensitivity and in the number of analytes quantifiable through a single chromatographic run. In its latest version, the analytes monitored are: (d18:1) sphingosine, (d17:1) sphingosine, (d18:1/16:0) ceramide, (d18:1/17:0) ceramide, glucosyl (β) ceramide (d18:1/16:0), glucosyl (β) ceramide (d18:1/12:0), sphingosine-1-phosphate, dihydroceramide (d18:0/16:0), sphingomyelin (d18:1/d16:0), and sphingomyelin (d18:1/d12:0). Within these ten compounds, four molecules were used as internal standards to quantify the extraction recovery (in table 8.6 are marked in orange): (d17:1) sphingosine, (d18:1/17:0) ceramide, glucosyl(β)ceramide (d18:1/12:0), and sphingomyelin (d18:1/d12:0). It is a matter of general knowledge that isotopic-labelled standards should be used to quantify properly the extraction yield. Unfortunately, for our purpose, the isotopic-labelled molecules of interest are not commercially available. Therefore, the cited synthetic molecules which have a similar structure and are not present in biological samples were used. The tables 8.4, 8.5, and 8.6 resume all the optimized parameters used.

By applying the parameters in only ten minutes is possible to obtain very good chromatographic separations. As example, figure 8.3 shows the TIC of a mix of 8 of the 10 standards listed in table 8.6.

8.4. Results and discussion

Results and discussion: HPLC-HRMS

Thanks to these settings, a correct chromatographic separation could be achieved. The chromatographic separation of triglycerides takes place based on the number of equivalent carbon atoms (EC): first elute the triacylglyc-

Table 8.5.: HPLC and MS parameters for SCIEX QTRAP® 5500 lipids analyses

Injection	3 μ L		
Flux	400 μ L/min		
Oven temperature	40° C		
Chromatographic run	Time (min)	ACN/H ₂ O(20:80 v/v), 0.1% FA	ACN/Isopropanol (20:80 v/v), 0.1% FA
	0	70%	30%
	1	70%	30%
	2.5	30%	70%
	4	20%	80%
	5	20%	80%
	6.5	10%	90%
	7.0	0%	100%
	7.1	70%	30%
10.0	70%	30%	
Source temperature	350° C		
Ion spray voltage	5000 V		
Polarity	Positive		
GS1; GS2	40 ; 50		
Curtain gas	20		

Table 8.6.: Sciex QTRAP® 5500 MRM parameters applied for sphingolipids analyses. The molecules reported in orange are the isotopic-labelled internal standards useful for quantification step. The standards reported above the double line in the middle of the page are those added to the method most recently and in the UPLC-TQMS results section will not be discussed

Analyte	Precursor Ion	Product Ion	Retention time (min)	DP (V)	EP (V)	CE (V)	CXP (V)
(d18:1) sphingosine	300.3	282.3 211.2	1.40	90 90	10 10	13 15	20 18
(d17:1) sphingosine	286.3	268.3 238.2	1.35	100 100	10 10	15 21	15 23
(d18:1/16:0) ceramide	538.5	520.5 264.3	3.85	100 100	10 10	15 30	26 24
(d18:1/17:0) ceramide	552.5	534.5 264.5	4.15	70 70	9 9	14 40	18 25
Glucosyl(β)ceramide (d18:1/16:0)	700.6	682.6 264.3	2.55	60 60	10 10	17 45	30 30
Glucosyl(β)ceramide (d18:1/12:0)	644.5	626.5 264.3	2.30	60 60	10 10	20 40	25 30
Sphingosine-1-phosphate	380.3	264.3 247.2	3.40	80 80	10 10	21 15	15 16
Dihydroceramide (d18:0/16:0)	540.5	302.3 284.3	3.55	150 150	8 8	26 28	13 10
Sphingomyelin (d18:1/d16:0)	703.6	685.6 184.1	3.20	110 110	8 8	24 28	16 5
Sphingomyelin (d18:1/d12:0)	647.7	184.3 629.6	3.00	100 100	9 9	40 25	14 7

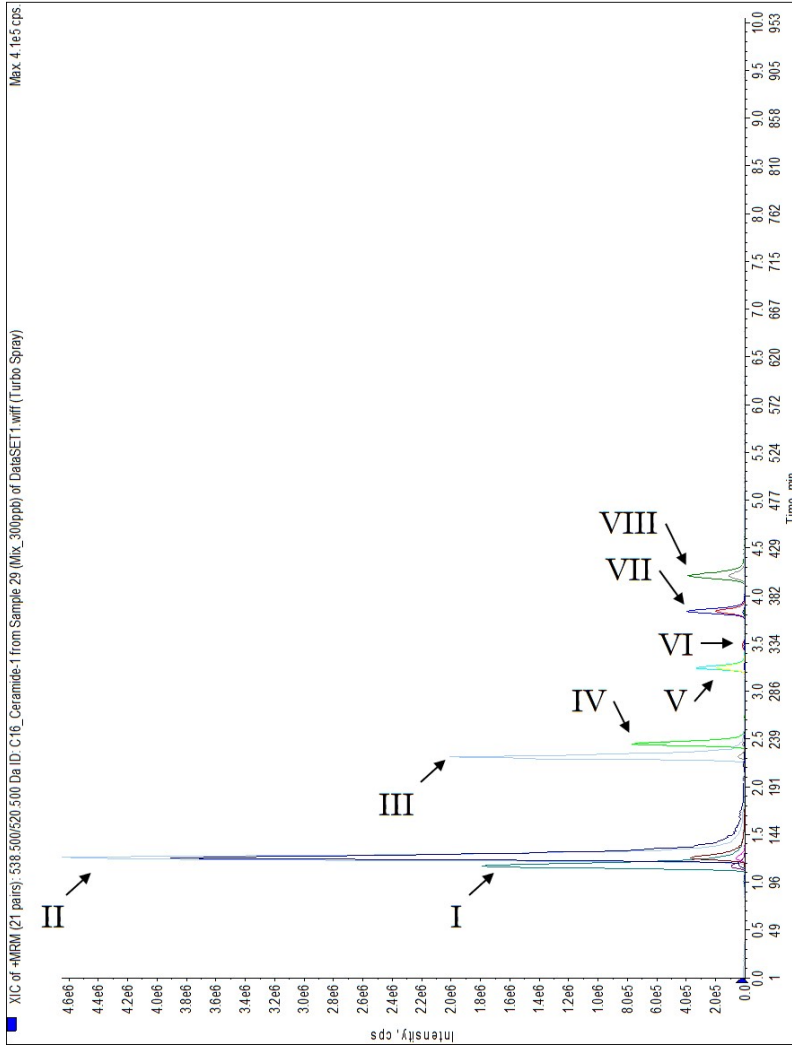


Figure 8.3.: Example of a TIC of a 300 $\mu\text{g}/\text{kg}$ standard mix. I: (d17:1) sphingosine; II: (d18:1) sphingosine; III: Glucosyl(β)ceramide (d18:1/12:0); IV: Glucosyl(β)ceramide (d18:1/16:0); V: Sphingosine-1-phosphate; VI: Dihydroceramide (d18:0/16:0); VII: (d18:1/16:0) ceramide; VIII: (d18:1/17:0) ceramide

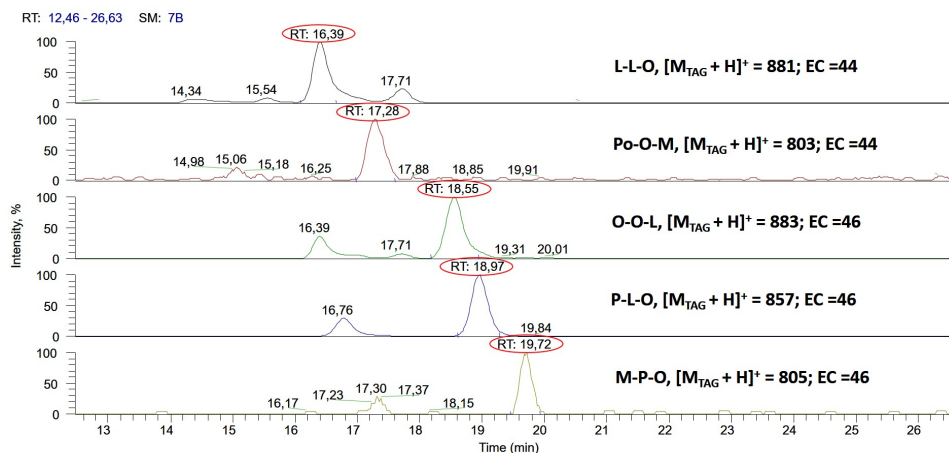


Figure 8.4.: Example of TAGs separation obtained through the HPLC-HRMS method in sampled SD5. The retention time of TAG considered are circled in red and their molecular weights and equivalent carbon atoms are reported on the right

erols that have a low EC number and, in the case of equivalence in carbon atoms equivalent number, the first eluted will be the TAGs of greater molecular weight. The figure 8.4 shows the result of the chromatographic separation. In this figure, the names of TAGs are reported using the abbreviations of fatty acids composing the triacylglycerol structure. In appendices D the table D.1 furnishes a complete overview of fatty acids: systematic and common names, number of carbon atoms, number of double bonds, etc.

The identification of TAGs in samples is conducted via a “restricted” global metabolomic approach. As a matter of fact, at the beginning of the analysis we know from literature and previous analyses that in biological samples such as homogenate livers, triacylglycerols molecules are present. The unknown aspect is which kind and how many of them are present in our samples. The identification of the different triacylglycerols was conducted through the analysis of the mass spectra corresponding to each chromatographic peak in the chromatographic run; based on the intensity of the fragments generated in the source by simple thermal degradation, it was possible to identify the fatty acids that are present in triglyceride and respective positions on the glycerol skeleton. The typical fragmentation of triacylglycerols consists in the loss of bound fatty acids in the external po-

Table 8.7.: *Triacylglycerols detected in samples. In the last column are reported the fragment ions generated from the parental ion, in order from the highest to the lowest intensity peak*

Triacylglycerol	Retention time (min)	$[M_{TAG} + H]^+$	$[M_{DAG} + H]^+$
L-L-O	16.5	881.7	599.5 - 601.5 - 601.5
PO-O-M	17.3	803.7	575.5 - 549.5 - 521.5
O-O-L	18.5	883.7	601.5 - 601.5 - 603.5
P-L-O	19.00	857.7	601.5 - 575.5 - 577.5
M-P-O	19.7	805.7	577.5 - 549.5 - 523.5
O-O-O	20.8	885.7	603.5 - 603.5 - 603.5
P-P-O	21.7	833.7	577.5 - 577.5 - 551.5
O-O-P	21.2	859.7	603.5 - 577.5 - 577.5
P-O-S	21.2	861.7	605.5 - 577.5 - 579.5
O-O-S	23.1	887.7	603.5 - 605.5 - 605.5

sitions (Sn-1 and Sn-3 positions); the ion signals generated are the most intense. This can be justified by the fragmentation mechanism proposed by Holcapek in [8]: the loss of the fatty acid in position sn-1 and sn-3 leads to the formation of a six atoms ring, a structure much more stable compared to a five atoms ring generated through the loss of fatty acid in Sn-2 position. The figures 8.5 and 8.6 are practical examples of triacylglycerols identification by mass spectrum interpretation. In the mass spectrum reported in figure 8.5 the parent ion protonated TAG has 859.7636 m/z and the product ions having a m/z ratio equal to 577.5204 ($[M_{DAG} + H - 282.2432]$) and 603.5301 ($[M_{DAG} + H - 256.2335]$), due to the neutral loss of an oleic acid and a palmitic acid respectively. These fragment ions are the most intense and are related to the loss of fatty acids in the Sn1 and Sn3 positions: therefore, simply the fatty acid linked to the glycerol skeleton at the Sn2 position can be deduced by difference, in this case, oleic acid. On the contrary, for the mass spectrum shown in figure 8.6, except for m/z 885.7842 parent ion mass, only a single fragment ion at m/z 603.5308 is present: this means that there is a loss of oleic acids both in the Sn1 and Sn3 positions; the fatty acid bound in the Sn2 position is deduced by difference.

Using TAG common name, the most abundant TAG detected in the samples are reported in table 8.7.

The inter-individual variability of data and extraction yields is not known. So, in each sample to operate the comparison, a percentage TAG value is

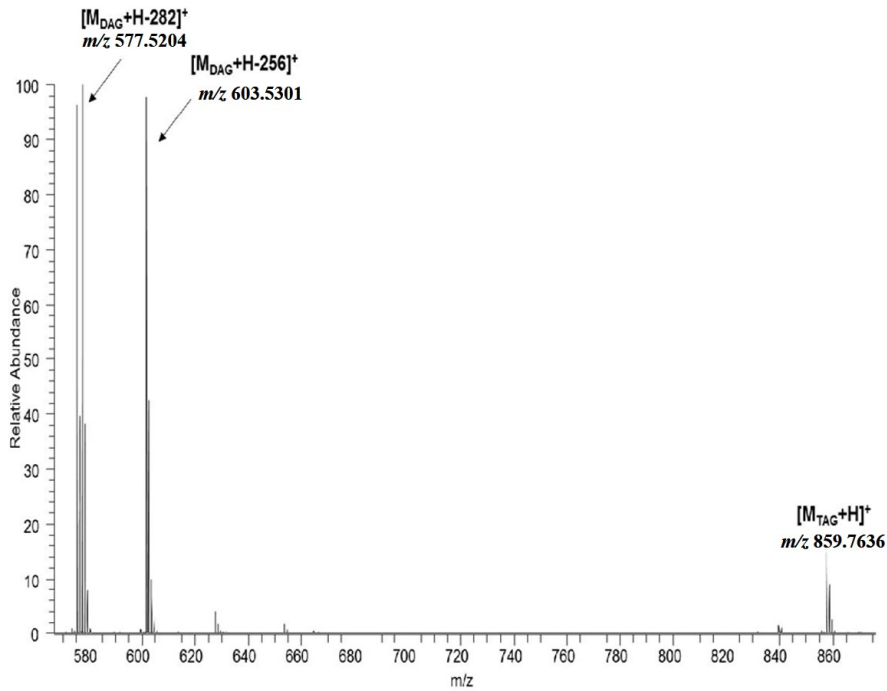


Figure 8.5.: Real case of a TAG mass spectrum. Fatty acids lost in the external positions (Sn-1 and Sn-3 positions) are different

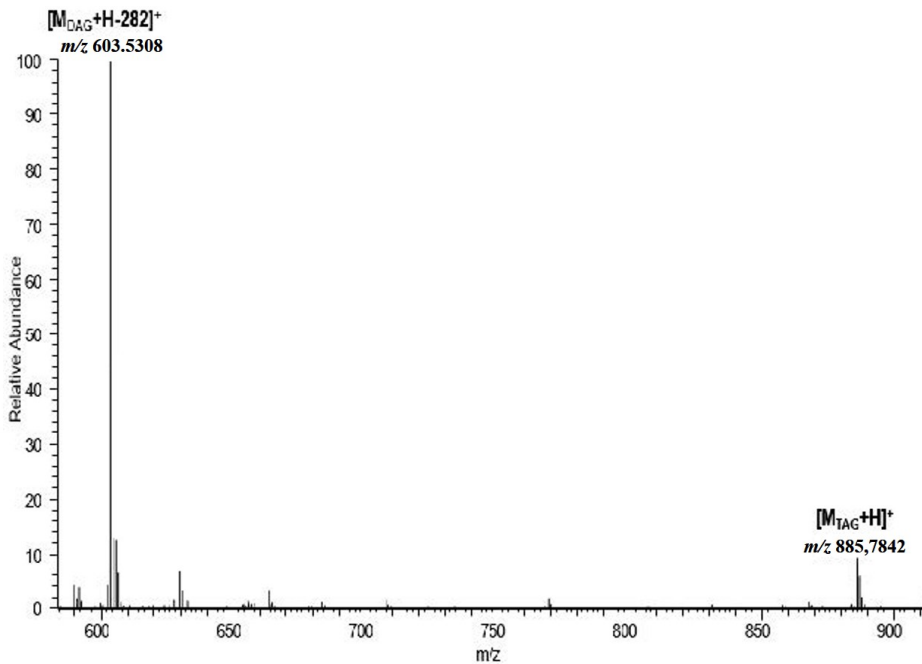


Figure 8.6.: Real case of a TAG mass spectrum. Fatty acids lost in the external positions (Sn-1 and Sn-3 positions) are the same molecule

Table 8.8.: *Percentage distribution values of triglycerides in SD, HFAT, HFRT. The number next to the diet name refers to the mice number attributed by the Department of Clinical and Biological Sciences Lab. The last three rows represent the means values for each TAG molecule. Furthermore, within each diet family the sum all all the mean values is 100%*

Sample	Triglycerides, % distribution values									
	L-L-O	Po-O-M	O-O-L	P-L-O	M-P-O	O-O-O	P-P-O	O-O-P	P-O-S	O-O-S
SD 1	32.73	0.44	21.23	37.26	0.12	2.71	0.51	6.46	1.20	0.34
SD 2	24.88	0.58	18.91	40.17	0.16	5.20	0.68	8.38	1.56	0.48
SD 5	23.59	0.79	20.04	36.40	0.32	7.75	1.15	10.42	1.94	0.90
SD 53	24.36	0.78	17.57	41.44	0.15	3.09	0.63	8.27	1.46	0.45
SD 54	29.26	0.55	19.52	42.73	0.20	2.79	0.68	6.52	1.34	0.41
SD 55	25.67	0.54	19.74	42.21	0.09	2.18	0.58	6.53	1.14	0.32
HFAT 60	7.47	6.12	12.83	35.57	1.11	8.94	2.09	22.75	4.19	1.93
HFAT 61	6.39	5.79	14.68	34.05	2.38	9.74	2.50	24.76	4.51	2.21
HFAT 62	7.55	4.89	8.21	34.79	2.02	9.67	2.30	24.04	4.46	2.06
HFAT 63	10.85	10.37	11.21	29.69	3.58	10.36	2.25	16.89	3.14	1.66
HFAT 69	11.89	8.30	9.21	34.82	1.97	7.02	2.25	18.73	3.39	1.42
HFAT 71	9.99	7.08	9.31	39.03	2.82	7.34	2.11	17.64	3.21	1.46
HFAT 93	8.00	6.09	13.15	25.02	1.60	6.14	0.96	32.37	6.02	0.67
HFAT 94	11.37	7.56	8.71	33.23	1.37	7.73	2.43	21.33	3.54	1.73
HFAT 96	11.38	6.54	12.95	38.83	2.31	7.34	1.79	19.68	3.61	1.58
HFAT 97	11.04	6.84	9.52	39.33	2.43	8.18	1.68	18.32	3.15	1.51
HFAT 98	13.48	8.90	13.11	33.83	3.24	8.33	2.37	15.82	2.58	1.35
HFAT 100	15.70	9.00	11.69	35.62	1.81	7.06	1.92	13.04	2.13	1.03
HFRT 58	5.96	1.25	8.74	57.44	3.88	4.61	1.59	16.46	2.94	0.53
HFRT 60	18.55	0.93	18.98	34.24	2.37	10.76	0.99	12.23	2.01	0.72
HFRT 70	18.36	1.43	20.17	37.91	2.20	9.40	2.02	9.85	4.81	0.84
HFRT 87	14.19	0.37	18.82	42.28	2.07	4.17	2.06	12.11	3.31	0.61
HFRT 88	22.55	0.36	23.15	34.65	2.08	4.09	3.51	8.60	2.60	0.41
HFRT 90	23.74	0.34	24.42	31.80	2.13	5.30	3.55	8.67	2.62	0.42
HFRT 93	24.20	0.49	24.67	31.14	2.16	4.48	2.68	8.19	2.55	0.43
HFRT 94	3.54	0.55	30.21	40.83	2.19	10.24	2.84	10.96	4.00	0.63
SD mean	26.75	0.61	19.65	40.03	0.17	3.55	0.71	6.98	1.54	0.60
HFAT mean	10.45	7.50	11.73	33.58	1.99	8.16	2.15	21.50	2.41	2.09
HFRT mean	16.39	0.79	19.85	38.79	2.51	5.51	2.91	8.68	4.01	0.58

calculated³. This value is independent by absolute concentration and allows us to compared triglyceride abundance in samples belonging from different diets. The table 8.10 reports all the percentage distributions of triglycerides in SD, HFAT, and HFRT. In the last three lines, the average percentage distribution values of triglycerides in SD, HFAT, and HFRT are reported.

Focusing on these last rows, some considerations can be made. A sum-

³This value was calculated by dividing the area of the single TAG with the sum of all the areas of TAGs in the sample, then multiplying it by 100. For example, if the TAG area is 1.05×10^6 and the sum of all the TAGs area in that sample is 1×10^8 , the value obtained for that TAG is 10.5 %

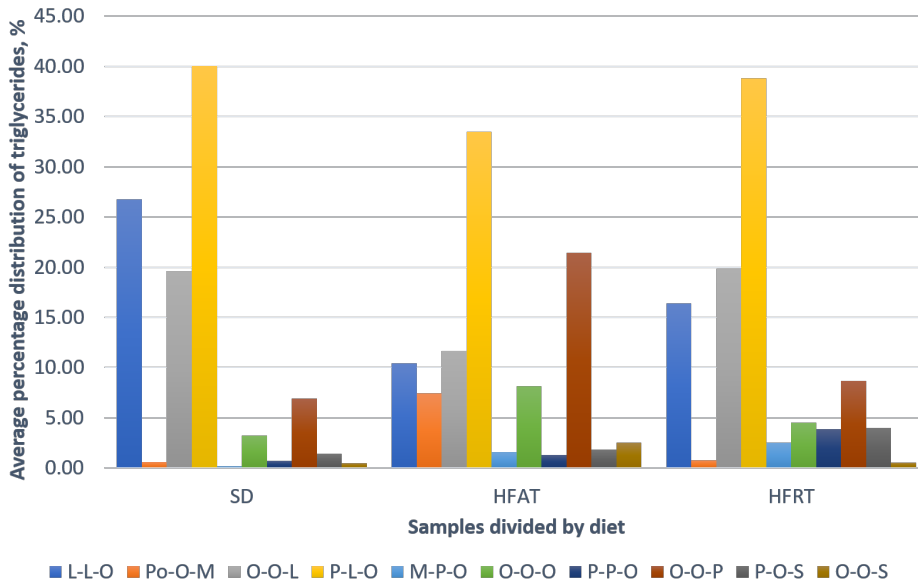


Figure 8.7.: *Histogram summarizing triglycerides mean values in SD, HFAT, and HFRT diets*

marizing histogram based on the mean values is reported in figure 8.7. The histogram is divided into three different portions, and each of them refers to a single diet program (SD, HFAT, HFRT). The sum of all the percentage values within each family is 100%.

All the triglycerides values monitored show a percentage increase in the HFRT diet, although this diet did not include fat administration; this first result can be considered as a molecular confirmation of the data obtained during the histological examination, that is the presence of hepatic steatosis in excessive fructose diets. As already mentioned before, the energy surplus furnished by sugars activates the liponeogenesis pathway, which, in turn, increases the production of fats and promotes hepatic steatosis.

Furthermore, the percentage increase in HFAT compared to HFRT is different and indicates that the causes of the two hepatic steatoses illness are quite distinct. Focusing on the histogram 8.7, it's possible to notice how for some TAGs levels are quite the same among diets and for other the levels are different. For example, P-L-O TAG (the highest yellow bar in the histogram) has more or less the same concentration in the homogenized

livers. On the contrary, Po-O-M could be used as a marker for the HFAT diet, because it has <1% values in SD and HFRT diets. Focusing on O-O-O and O-O-P it's possible to notice how certain TAGs selectively increase in distributions (in this case, TAGs belonging from oleic acid metabolism).

The differences are, therefore, attributable to the diet. Understanding why concentrations of specific analytes increase or decrease depending on the diet will undoubtedly require more biochemical studies in the future.

Results and discussion: UPLC-TQMS

As mentioned above, this research project has been underway for more than two years. The method of analysis in its latest formulation involves the analysis of ten different analytes. However, only a limited number of samples were analysed in this way. Therefore, wishing to report data of a larger number of samples in order to make valid statistical considerations, only results concerning (d18:1) sphingosine, (d17:1) sphingosine, (d18:1/16:0) ceramide, (d18:1/17:0) ceramide, glucosyl (β) ceramide (d18:1/16:0), and glucosyl (β) ceramide (d18:1/12:0) have been reported. These molecules are, in fact, the analytes that, since the first studies carried out, have always been quantified. In table 8.6 all the MRM parameters are however reported for completeness of the information (the analytes not considered in these results section are the ones above the double line in the table 8.6).

The table 8.9 reports the parameters that allow us to say that the method is validated and applicable as a routine method.

The table 8.10 resumes the different sphingolipids levels detected in livers samples. The bar graph in the figure 8.8 shows the same results of the table 8.10 but in a more user-friendly way.

The instrumental methods have shown excellent reproducibility. In all samples treated with the SD diet, the (d18:1) sphingosine levels are generally shallow and close to the LOD. In contrast, for the HFRT and HFAT samples, the levels of this analyte are five and ten times higher, respectively. These findings demonstrate that altered diets that bring a number of lipids higher than the average have repercussions on the liver leading to dyslipidemia. On the contrary, the concentrations of ceramides do not

Table 8.9.: Schematic report of parameters obtained during validation procedures for sphingolipids in UPLC-TMQS analyses

(d18:1) sphingosine		
Parameter	Theoretical value	Experimental value
LOD	X	0.17 $\mu\text{g}/\text{Kg}$
LOQ	X	0.57 $\mu\text{g}/\text{Kg}$
LLOQ	X	1 $\mu\text{g}/\text{Kg}$
Selectivity, %	<30.00	4.45
Linearity, %	<25.00	0.91
Relative Standard Deviation, %	<25.00	4.86
(d18:1/16:0) ceramide		
Parameter	Theoretical value	Experimental value
LOD	X	0.28 $\mu\text{g}/\text{Kg}$
LOQ	X	0.93 $\mu\text{g}/\text{Kg}$
LLOQ	X	1 $\mu\text{g}/\text{Kg}$
Selectivity, %	<30.00	28.03
Linearity, %	<25.00	16.33
Relative Standard Deviation, %	<25.00	10.40
Glucosyl(β)ceramide (d18:1/16:0)		
Parameter	Theoretical value	Experimental value
LOD	X	0.25 $\mu\text{g}/\text{Kg}$
LOQ	X	0.84 $\mu\text{g}/\text{Kg}$
LLOQ	X	1 $\mu\text{g}/\text{Kg}$
Selectivity, %	<30.00	23.84
Linearity, %	<25.00	9.67
Relative Standard Deviation, %	<25.00	11.20

Table 8.10.: Concentration expressed in $\mu\text{g/L}$ of sphingolipids of interest among the three diets. These levels take into consideration the dilution factor calculated starting from the standard isotopically marked charges (reported in braces under the corresponding analyte). The sample are named using the combination of the specific diet and the mouse number used in the Lab during feeding period. So, the liver sample belonging from mouse number 62 who followed an HFAT diet is called "HFAT, 62"

Sample name	(d18:1) sphingosine, [$\mu\text{g/L}$] {(d17:1) sphingosine}	(d18:1/16:0) ceramide, [$\mu\text{g/L}$] {(d18:1/17:0) ceramide}	Glucosyl(β)ceramide (d18:1/16:0), [$\mu\text{g/L}$] {glucosyl(β)ceramide (d18:1/12:0)}
SD 1	<LOD	740.03	353.21
SD 2	6.36	598.05	339.82
SD 5	<LOD	558.98	226.38
SD 53	<LOD	615.26	356.54
SD 54	15.48	564.94	429.61
SD 55	4.21	464.13	292.25
HFAT 60	<LOD	515.76	319.12
HFAT 61	<LOD	379.15	321.83
HFAT 62	<LOD	650.68	296.89
HFAT 63	<LOD	600.36	379.68
HFAT 69	<LOD	690.02	309.13
HFAT 71	<LOD	389.24	286.83
HFAT 93	108.41	121.72	312.33
HFAT 94	193.52	955.03	477.18
HFAT 96	51.09	489.18	491.64
HFAT 97	4.52	522.19	324.56
HFAT 98	8.33	109.98	431.89
HFAT 100	82.23	647.36	328.98
HFRT 57	<LOD	892.33	169.12
HFRT 58	15.64	728.45	265.13
HFRT 60	12.64	448.47	212.67
HFRT 66	53.16	877.69	399.52
HFRT 67	11.91	381.96	244.55
HFRT 70	53.93	574.47	320.13
HFRT 87	<LOD	345.64	310.89
HFRT 88	<LOD	272.73	263.82
HFRT 92	<LOD	363.88	237.60
HFRT 93	34.19	170.65	273.62
HFRT 94	<LOD	401.21	265.92
SD mean	8.68	332.97	590.23
HFAT mean	74.68	356.67	505.89
HFRT mean	30.25	269.36	496.13

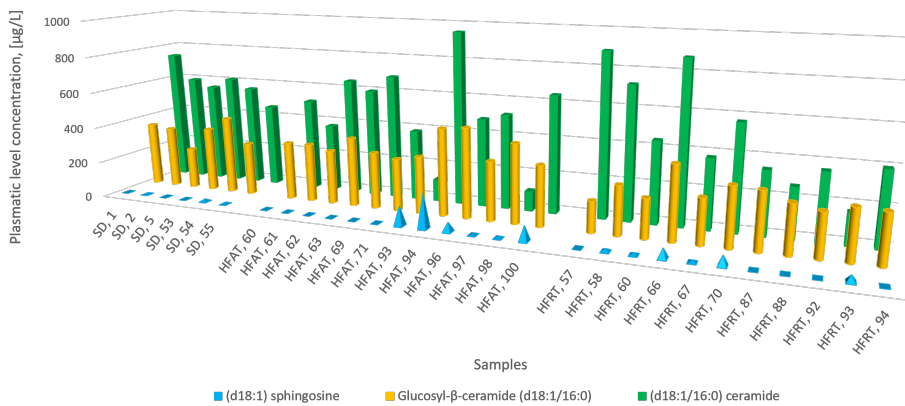


Figure 8.8.: Summary histogram with the concentrations [$\mu\text{g/L}$] of (d18:1) sphingosine (reported as blue pyramids), (d18:1/16:0) ceramide (reported as yellow cylinders), and glucosyl (β) ceramide (d18:1/16:0) (reported as green parallelepipeds) among samples

show a significant variation between the SD and HFRT distribution. The ceramides of palmitic acid are probably not the only product that accumulates in the liver leading to hepatic steatosis disease. The data obtained from the quantification of ceramide (d18:1/16:0) in the HFAT diet is bucking the trend: the distribution is significantly lower than that obtained from the SD diet. An excessive quantity of lipids introduced into the diet probably depresses the metabolic pathway of liponeogenesis.

8.5. Conclusion and future prospects

The present thesis work was developed from the objective data of the appearance of hepatic steatosis in fructose-drinking mice (figure 8.2).

The quantification of triacylglycerols performed by HPLC-HRMS confirmed the histological data. The lipids levels in livers belonging to the HFRT diet are higher than those found in the SD diet. Furthermore, the variations in the concentrations of TAGs are different from those found in the HFAT diet; this positive feedback seems to confirm the different causes of steatosis in the two altered diets.

Based on what has been achieved in not mentioned GC-MS analyses⁴ we proceed to build up an UPLC-TQMS method to quantify what seemed to be the primary product of palmitic acid metabolism: ceramides. The molecules considered were ceramide of palmitic acid, glucosylceramide of palmitic acid, and sphingosine (d18:1). The instrumental methods, obtained for the quantification of these molecules, have shown an excellent reproducibility and have been validated. Unfortunately, the concentrations of ceramides do not show a significant variation between the SD and HFRT distribution. The ceramides of palmitic acid are, therefore, not the last or the only product that accumulates in the liver leading to dyslipidemia. Moreover, the data obtained by ceramide (d18:1/16:0) quantification in HFAT diet is bucking the trend: the distribution is significantly lower than the one obtained from the SD diet. This result could be interpreted as follow: an excessive quantity of lipids introduced into the diet depresses the metabolic pathway of liponeogenesis.

Results so far obtained have been reported and discussed in the present chapter. However, the research project has evolved and channelled into two different tracks: on one hand the detection of sphingosine-1-phosphate, in the elucidation of biochemical mechanisms related to prediabetic pathology; on the other hand, the incrementation of metabolic products detectable by one single analysis (sphingomyelins, phosphatidicolins, phosphatidylethanolamines) in order to have a complete picture of the quantitative levels of the analytes involved in ceramide/sphingosine/sphingosine-1-phosphate balance scheme briefly resumed in figure 8.9.

For the first purpose mentioned, exciting results are already achieved. In this case, new samples were prepared. Mice were divided into four different groups: mice fed with a SD diet; mice fed with a SD diet +

⁴In GC-MS analyses were monitored fatty acids levels in liver homogenate samples. The purpose was to check the increase or the decrease of individual fatty acids in the samples. TAGs analyses cannot obtain this analytical data due to the simultaneous presence of three fatty acids in each triglyceride. Quantification confirmed, firstly, the data already collected by the HPLC-HRMS approach is consistent, and a different increase in lipids in HFAT and HFRT diets occurred. Secondly, it was possible to identify a specific analyte, palmitic acid, as a marker of HFRT diet.

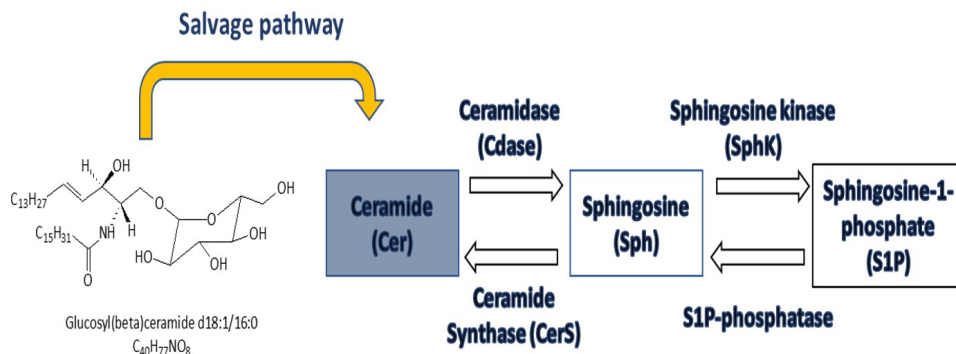


Figure 8.9.: *Ceramide/sphingosine/sphingosine-1-phosphate balance scheme with salvage pathway highlighted*

sphingosine-1-phosphate (called SD+P); mice fed with an HFAT diet; mice fed with an HFAT diet + sphingosine-1-phosphate (called HFAT+P). As already anticipated in table 8.6, sphingosine-1-phosphate was added to the analytical method to detected and quantified. By comparing the new samples, an effective decrease in the hepatic level of sphingosine-1-phosphate concentrations were found in mice fed with HFAT with pyridoxamine compared to those supplied with the only diet rich in fat (histogram 8.10. Therefore, this result demonstrates how the pyridoxamine introduced in a diet has a proven anti-glycating activity, acting as a quencher for advanced glycation end-products (AGEs).

However, the low number of samples analysed, the small but important variations in the present and past diets (for example, the same type of food but not the same), and the different sampling period of new and previous mice livers cannot allow us to compare in a statistically significant way the samples. That's why I mention these results only in the conclusion section.

Considering now the second research track mentioned above, results are not lacking. As already mentioned in table 8.6 new analytes are added to the UPLC-TQMS analytical method: sphingomyelin (d18:1/d16:0) and dihydroceramide (d18:0/16:0). The LOD, LOQ, and LLOQ values for these analytes and the previously cited ones are summarized in table 8.11.

For these experiments, samples consist of homogenized mice liver. In this case, as happened for the first purpose mentioned, mice were divided

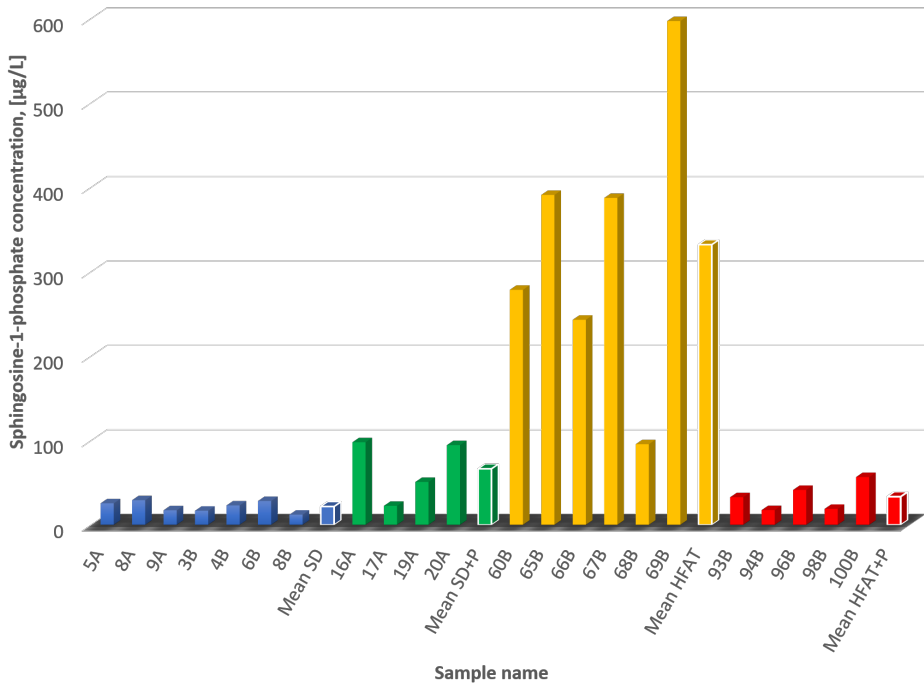


Figure 8.10.: Histogram summarizes of concentrations of sphingosine-1-phosphate expressed in $\mu\text{g/L}$. Bars' colour reflects family from which samples belong: blue for SD, green for SD+P, yellow for HFAT, red for HFAT+P. White-edge bars represent instead average values for the samples belonging to a specific diet

Table 8.11.: LOD, LOQ, and LLOQ values obtained during validation procedures for new sphingolipids introduced in UPLC-TMQS method

Analyte	LOD, ($\mu\text{g/L}$)	LOQ, ($\mu\text{g/L}$)	LLOQ, ($\mu\text{g/L}$)
Sphingosine-1-phosphate	1.63	5.44	2.5
Dihydroceramide (d18:0/16:0)	6.45	19.49	10
Sphingomyelin (d18:1/d16:0)	5.68	18.95	10

Table 8.12.: *Concentration of sphingolipids introduced in UPLC-TMQS method in real samples. s.d.: standard deviation*

Analyte	Diet	Average values, $\mu\text{g/L}$	s.d., $\mu\text{g/L}$
Sphingosine-1-phosphate	SD	15.64	4.62
	SD+P	31.76	13.09
	HFAT	409.87	104.62
	HFAT+P	55.89	21.87
Dihydroceramide (d18:0/16:0)	SD	15.11	3.25
	SD+P	31.11	11.11
	HFAT	74.44	27.65
	HFAT+P	29.75	14.10
Sphingomyelin (d18:1/d16:0)	SD	198.17	80.81
	SD+P	206.77	43.10
	HFAT	310.29	198.97
	HFAT+P	269.43	8.24

into four groups: SD diet, SD+P diet, HFAT diet, and HFAT+P diet. The results, due to the low number of samples analysed, variations in the sampling period of new and previously treated mice livers cannot allow us to compare globally all samples analysed in a correct statistical way. However, looking at the results obtained (tables 8.11 and 8.12) we can assume that thanks to the statistical parameters taken into consideration, the method could be defined validated and applicable. Furthermore, the results show once again how a high fat diet (HFAT) causes an imbalance in sphingolipids equilibrium. Pyridoxamine, with its antiglycation activity, can rebalance this altered equilibrium and acts as a RAGE antagonist. Moreover, the onset of nonalcoholic steatohepatitis (NASH) is mediated by the studied analytes.

References of Chapter 8

- [1] M. Bellan et al. “Severity of Nonalcoholic Fatty Liver Disease in Type 2 Diabetes Mellitus: Relationship between Nongenetic Factors and PNPLA3/HSD17B13 Polymorphisms”. In: *Diabetes & Metabolism Journal* 43(5), 700-710 (2019). DOI: 10.4093/dmj.2018.0201 (cit. on p. 135).
- [2] D. Eberlé et al. “SREBP transcription factors: Master regulators of lipid homeostasis”. In: *Biochimie* 86(11), 839-848 (2004). DOI: 10.1016/j.biochi.2004.09.018 (cit. on pp. 135, 136).
- [3] L.F. Eggers et al. *Encyclopedia of Lipidomics*. Section: Liquid Extraction, Folch. In: Wenk M. Dordrecht: Springer, 2016 (cit. on p. 144).
- [4] E. Fabbrini et al. “Hepatic steatosis as a marker of metabolic dysfunction”. In: *Nutrients* 7(6), 4995-5019 (2015). DOI: 10.3390/nu7064995 (cit. on p. 136).
- [5] J. Folch. “The isolation of phosphatidyl serine from brain cephalin, and identification of the serine component”. In: *Journal of Biological Chemistry* 139(2), 973-974 (1941) (cit. on p. 144).
- [6] S.P. Gomes et al. “Stereology shows that damaged liver recovers after protein refeeding”. In: *Nutrition* 38, 61-69 (2017). DOI: 10.1016/j.nut.2017.02.010 (cit. on p. 135).
- [7] N.C. Hait et al. “Sphingosine kinases , sphingosine 1-phosphate, apoptosis and diseases”. In: *Biochimica et Biophysica Acta* 1758(12), 2016-2026 (2006). DOI: 10.1016/j.bbamem.2006.08.007 (cit. on p. 136).
- [8] M. Holcapek et al. “Characterization of triacylglycerol and diacylglycerol composition of plant oils using high-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry”. In: *Journal of Chromatography* 1010(2), 195-215 (2003). DOI: 10.1016/S0021-9673(03)01030-6 (cit. on p. 149).
- [9] M.M. Hussain et al. “Mechanisms involved in cellular ceramide homeostasis”. In: *Nutrition & Metabolism* 9(1), 71 (2012). DOI: 10.1186/1743-7075-9-71 (cit. on p. 138).

- [10] Burlamaqui I.M. et al. “Hepatic and biochemical repercussions of a polyunsaturated fat-rich hypercaloric and hyperlipidic diet in Wistar rats”. In: *Archives of Gastroenterology* 48(2), 153-8 (2011) (cit. on p. 135).
- [11] S. Kersten. “Mechanisms of nutritional and hormonal regulation of lipogenesis”. In: *EMBO Reports* 2, 282-286 (2001). DOI: 10.1093/embo-reports/kve071 (cit. on p. 135).
- [12] J.M. Kneeman. “Secondary causes of nonalcoholic fatty liver disease”. In: *Therapeutic Advances in Gastroenterology* 5(3), 199–207 (2012). DOI: 10.1177/1756283X11430859 (cit. on pp. 135, 136).
- [13] H. Le Stunff et al. “Generation and metabolism of bioactive sphingosine-1-phosphate”. In: *Journal of Cellular Biochemistry* 92(5), 882-899 (2004). DOI: 10.1002/jcb.20097 (cit. on p. 136).
- [14] C. Loguercio et al. “Non-alcoholic fatty liver disease: a multicentre clinical study by the Italian Association for the Study of the Liver”. In: *Digestive and Liver Disease* 36(6), 398-405 (2004). DOI: 10.1016/j.dld.2004.01.022 (cit. on p. 136).
- [15] R. Mastrocola et al. “Accumulation of advanced glycation end-products and activation of the SCAP/SREBP lipogenetic pathway occur in diet-induced obese mouse skeletal muscle”. In: *PLoS One* 10(3), e0119587 (2015). DOI: 10.1371/journal.pone.0119587 (cit. on pp. 137, 142).
- [16] R. Mastrocola et al. “Advanced glycation end products promote hepatosteatosis by interfering with SCAP-SREBP pathway in fructose-drinking mice”. In: *American Journal of Physiology: Gastrointestinal and Liver Physiology* 305(6), G398-407 (2013). DOI: 10.1152/ajpgi.004502012 (cit. on pp. 137, 142).
- [17] L. Pimpin et al. “Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies”. In: *Journal of Hepatology* 69(3), 718-735 (2018). DOI: 10.1016/j.jhep.2018.05.011 (cit. on p. 136).

-
- [18] A.M. Schmidt et al. “Receptor for age (RAGE) is a gene within the major histocompatibility class III region: implications for host response mechanisms in homeostasis and chronic disease”. In: *Frontiers in Bioscience* 6:D1151-60 (2001). DOI: 10.2741/schmidt (cit. on p. 137).
- [19] M. Sengupta et al. “Inhibition of Hepatotoxicity by a LXR Inverse Agonist in a Model of Alcoholic Liver Disease”. In: *Translational Science - ACS Pharmacology* 1(1), 50-60 (2018). DOI: 10.1021/acsptsci.8b00003 (cit. on p. 135).
- [20] J.T. Sessions Jr. et al. “The effect of barbiturates in patients with liver disease”. In: *Journal of Clinical Investigation* 33(8), 1116-1127 (1954). DOI: 10.1172/JCI102985 (cit. on p. 135).
- [21] T. Sugihara et al. “Falsely Elevated Serum Vitamin B12 Levels Were Associated with the Severity and Prognosis of Chronic Viral Liver Disease”. In: *Yonago Acta Medica* 60(1), 31-39 (2017) (cit. on p. 135).
- [22] P. Treuting et al. *Comparative Anatomy and Histology (Second Edition) - A Mouse, Rat, and Human Atlas*. Chapter 13, pages 229-239. Cambridge: Academic Press, 2018 (cit. on p. 135).
- [23] S.F. Yan et al. “The RAGE Axis: a Fundamental Mechanism Signaling Danger to the Vulnerable Vasculature”. In: *Circulation Research* 106(5), 842-853 (2010). DOI: 10.1161/CIRCRESAHA.109.212217 (cit. on p. 137).
- [24] A.R. El-Zayadi. “Hepatic steatosis: A benign disease or a silent killer”. In: *World Journal of Gastroenterology* 14(26), 4120–4126 (2013). DOI: 10.3748/wjg.14.4120 (cit. on p. 135).

Semi-targeted analysis: Phytochemical characterization of small cultivated berries

9.1. Introduction

The term berry has a different meaning depending on the application field. Generally, berries are edible fruit with specific characteristics: brightly coloured, sour, or sweet, juicy fruits without a stone inside. Blueberry, strawberry, or blackcurrant are just a few examples. On the contrary, from a scientific point of view, berries are «fruits produced from the ovary of a single flower»[19]. In these terms, grapes, bananas, and tomatoes can be defined berries too. In the present work, each time we will generally refer to berries we only apply to the fruits commonly definition.

Thanks to their antioxidant, anti-inflammatory, and antimicrobial activities [15], the popularity of small berries fruits has rapidly increased in Western countries given to their composition and health-promoting properties related to the prevention or the onset delay of chronic age-related diseases [2]. As such, berries are now classified as superfruits¹. These beneficial effects are provided by a structurally varied range of bioactive compounds [27].

Increased consumer interest in berry health-promoting effects has also been recorded in Italy [8], particularly in the North-Western regions, where berry farmers have more than triplicated the extension of cultivated areas compared with the last decade. Nowadays, the total purchased fruit of these specialty foods in Italy is only 1.5% with an average of 2.7 annual purchases per family [18]; on the contrary, in the United Kingdom, for example, they represent 7.6% of the total purchased fruit with 10.4 annual purchases per

¹“The foods that supply an abundance of nourishing natural vitamins, minerals, healthy fats, amino acids, plant enzymes, antioxidants and phytonutrients which can be considered beneficial for health and well-being” [19].

family [11].

Currently, at Italian regional level, 80% berry production is dedicated to blueberries (*Vaccinium corymbosum*) and 14% to raspberries (*Rubus idaeus*) [8, 18], while the remaining part is intended for the cultivation of black currants (*Ribes nigrum*), red currants (*Ribes rubrum*), white currants (*Ribes pallidum*), blackberry (*Rubus fruticosus*), white and red gooseberries (*Ribes grossularia L.*) and more recently goji (*Lycium barbarum L.*). Moreover, these valuable fruits can be grown in lands otherwise abandoned due to poor productivity [20].

9.2. Aim of the work

The phytochemical characterization of each sample provides useful knowledge about berries, such for example, the levels of bioactive compounds that could be used as a natural antioxidant and preservative in food processing. To obtain this information, after the berry samples extraction, different procedures were conducted: evaluation of fatty acid profile, total phenolic - anthocyanin - flavonoid contents, ferric-reducing antioxidant power (FRAP) and a Dionex HPLC coupled to two different detectors: an HRMS LTQ OrbitrapTM set in positive ion mode and a Photodiode-Array Detection PDA. Hereafter, the analytical platform used for these analyses will be named with the acronym HPLC-PDA-ESI HRMS.

This ambitious project was carried-on in collaboration with Dr. Peiretti and Dr. Gai from CNR ISPA of Grugliasco (TO). So, from my point of view, the study aimed to compare thanks to HPLC-PDA-ESI HRMS platform the bioactive compounds and antioxidant activities of eleven different cultivated berry fruits extracts (harvested during the 2016 summer season in the Piedmont Region) by using a semi targeted metabolomics approach.

As reported in section 1.2.2, one of the three ways to operate metabolomics analyses can be the semi-targeted approach. This approach fits perfectly with the purpose of our study. Before proceeding with sample extraction and analysis, the identity of molecules of interest is already known (anthocyanins, flavonoids - targeted approach). Still, the numbers of entities to detect (could be ten or two-hundred) and their chemical characteristics



Figure 9.1.: *Pictures of the cited berries. In top row, from left to right: white currant, blueberry, white gooseberry, and goji; in the bottom row, from left to right: blackberry, raspberry, black currant, red currant, and red gooseberry*

(the molecules are free or glycosylated) are unknown (untargeted approach).

To help to orientate the reader in the polyphenolic molecule classification and the sub-classes structures of flavonoids mainly present in fruit extracts, some useful schemes are reported in the appendix B.

9.3. Sample preparation and method settings

Samples: from sampling to extraction

Raspberries, black currants, red currants, white currants, white and red gooseberries, blackberry (Loch Tay cultivar), goji, and three cultivars of blueberries fruits (Duke, Blue Ray, and Misty) were purchased from local farms (figure 9.1 and table 9.1).

Upon arrival in the laboratory, berries were instantaneously freeze-dried using a laboratory lyophilizer (5 Pascal, 4 Trezzano sul Naviglio, Italy) and then finely ground and vacuum-packed and stored at -20°C until analysis. The sample preparation consists of several steps. Firstly, 2 g of lyophilized sample are extracted with 10 mL of 80:20 (v/v) methanol:water with formic acid (1%) using a Polytron tissue homogenizer (Type PT 10-35; Kinematica GmbH, Luzern, Switzerland) for 1 minute. Next, the extracted samples are centrifuged at 4000 g for 20 minutes at room temperature [3]. Then, pellets were re-extracted, and supernatants were pooled and then evaporated

Table 9.1.: *Scientific name, site, dry matter (DM, g/kg fresh matter) and extract yield of the berries. The extract yields expressed as milligrams of extract per gram (dry weight) of fruits*

Sample	Scientific name	Latitude	Longitude	DM	Yield
Blackberry, Loch Tay	<i>Rubus fruticosus</i>	45°35'N	8°04'E	284.5	178.1
Black currants	<i>Ribes nigrum</i>	45°35'N	8°04'E	753.5	241.9
Blueberries, Blue Ray	<i>Vaccinium corymbosum</i>	45°35'N	8°04'E	498.8	264.5
Blueberries, Duke	<i>Vaccinium corymbosum</i>	45°03'N	7°30'E	589.0	197.7
Blueberries, Misty	<i>Vaccinium corymbosum</i>	45°35'N	8°04'E	473.2	366.5
Goji	<i>Lycium barbarum</i>	45°04'N	7°43'E	769.9	266.3
Raspberries	<i>Rubus idaeus</i>	45°03'N	7°30'E	249.2	451.6
Red currants	<i>Ribes rubrum</i>	44°32'N	7°65'E	655.3	162.6
Red gooseberries	<i>Ribes grossularia</i>	45°03'N	7°30'E	691.1	186.9
White currants	<i>Ribes pallidum</i>	44°82'N	7°97'E	655.7	178.1
White gooseberries	<i>Ribes grossularia</i>	44°79'N	7°37'E	866.4	120.2

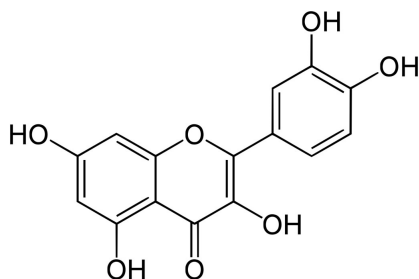


Figure 9.2.: *Quercetin structure. It is a plant flavonol from the flavonoid group of polyphenols and it is one of the markers detected within all the samples under analysis*

under vacuum at room temperature using a Speedvac (SC210A; Savant Instruments, Farmingdale, NY, USA). At last, residues were re-suspended in 10 mL of methanol/formic acid (99:1, v/v). Aliquots were stored at -80°C before analysis.

Analytical method

The work reported in this chapter is part of broader research that resulted in the publication of a future article² entitled “Bioactive compounds, antiox-

²The mentioned article is under peer-reviewing in March, 2020.

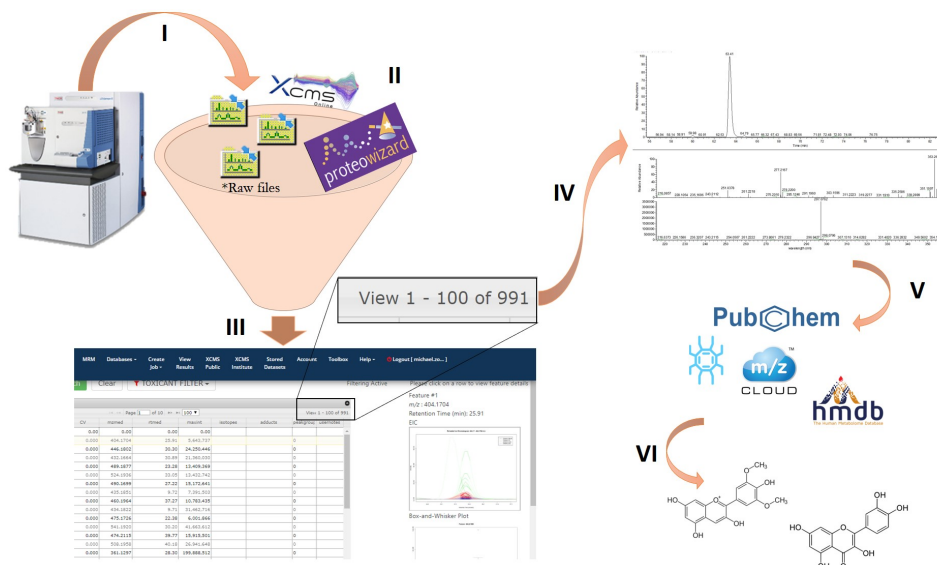
idant and antiradical activity of small Berries cultivated in the northwest of Italy”. As we read in the article on the samples extracted so many types of tests were performed: determination of Total Phenolic Content (TPC), total flavonoids and total anthocyanins; three different antioxidant capacity assays ((TEAC, FRAP, and DPPH); HPLC-PDA-ESI HRMS in positive ion mode and fatty acids composition by GC-FID. However, since the purpose of my research project focused on mass spectrometry techniques, I will only report this kind of measurements in these sections.

The analytical platform used to characterize the extracts was a HPLC operated in reverse phase mode coupled to a PDA-ESI HRMS scanning in positive ion mode. Based on the information obtained from the UV-VIS spectrum, the accurate mass of precursor ions, MS/MS experiments, and the use of online databases such as XCMS [28] the main anthocyanins and polyphenols were identified and quantified. The figures 9.3 9.4 summarize the two different workflows applied. In figure 9.3 the raw data obtained thanks to the HRMS platform (number I) are firstly converted by Proteowizard software in mzXML files and secondly are elaborated by XCMS online software (number II).

All the different parameters to monitor and the suitable XCMS setting to use are the same as the ones reported and discussed in chapter 12. In table 9.2 are reported the different parameters useful to convert .raw files into the mzXML format. Once the proper files are generated, they were uploaded into XCMS online software. The XCMS parameters applied for the multi-job processing were set as follows: centWave for feature detection ($\Delta m/z=30$ ppm, minimum and maximum peak width 30 and 120 s, respectively; S/N threshold = 6; Integration method = 1; Pref. peaks = 5; Pref intensity = 100000; obiwarped settings for retention time correction (profStep=1.0); parameters for chromatogram alignment, including mzwid = 0.015, minfrac = 0.5, and bw = 5; parameters for statistics includes Kruskal-Wallis non-parametric, perform post-hoc analysis = true, significance p-value threshold = 0.005, fold change threshold = 1.5; finally, parameters for annotation, including searching for isotopes, ppm error=5 and m/z , and absolute error=0.05. All samples behaving as an outlier in PCA

Table 9.2.: *ProteoWizard parameters applied for berries .raw files conversion*

Filter	Parameters
PeakPicking	Cwz: snr=1.0 peak Space=0.1 msLevel=1-1
LockmassRefiner	ESI+: mz=291.0843; tol=0.005
msLevel	1-2

Figure 9.3.: *Scheme of the workflow applied for untargeted metabolomics analysis of lyophilized berries samples*

plots were removed to smoothen the distribution and PCA plot was then recalculated.

The results obtained (number III, screenshot of the XCMS report table) show clearly the high number of possible analytes present within the sample. At this point, the analyst verifies the validity of the results directly into the .raw file (number IV; for example, he monitors if the matches proposed by the software can have scientific meaning or he checks if the molecules detected are or no background noise). To accomplish this task, he resorts to online databases and personal knowledge (number V). At the end of the process, the presence or the absence of a certain molecule will be demonstrated (number VI).

The second approach used, which integrates with the first one, starts directly from the data obtained through the HRMS instrumentation. From the chromatograms monitored in Full Mass (figure 9.4, line A) is possible to estimate the presence of more abundant substances. The structural details that need to reach a reasonable identification are obtained from UV spectra³, Full Mass spectra in high resolution (figure 9.4, raw D) and from MS/MS fragmentation spectra (figure 9.4, raw E). Once a molecule of interest is found (for example, m/z 741.2235; figure 9.4, raw B) the high-resolution data allow the calculation of the raw formula and the number of rings and double bonds, while fragmentation schemes highlight the neutral parts of the molecule that are lost and the product ions that survive. Finally, a comparison between the hypothesized molecule with standards (if available), literature, or databases must be done to have full certainty about the attribution. So, by combining the two approaches, the determination of a specific analyte whitening the sample is double-checked.

Returning to the analytical platform, two instrument setups were built up. On one hand for polyphenols (phenolic acids, flavanols, flavonols, and anthocyanins) analyses, the method consisted of a Dionex Ultimate 3000 HPLC system equipped with a PDA detector (Thermo Scientific Surveyor, Milan, Italy), coupled with a High-Resolution Mass Spectrometer LTQ-OrbitrapTM (Thermo Scientific, Milan, Italy) through an ESI interface operating in positive ion mode.

Furthermore, in these analyses two different instrumental conditions were adopted to detect anthocyanins and other flavonoid compounds. A procedure described in literature [27] with slight modifications was used. For anthocyanin compounds a RP C18 column (Varian Pursuit C18, 150 × 2.0 mm, 3 μm particle size; Agilent, Milan, Italy) at 200 μL/min flow rate was used. The elution solvents adopted were formic acid 0.1% in methanol (B) and water (A). The gradient profile was 0-6 min from 10 to 15% of B, 6-12 min from 15 to 25% B, 12-16 min from 25 to 30% B, an isocratic step to 30% of solvent B for 14 minutes, and finally 30-42 min from 30 to 100%

³For further information concerning the process to characterize polyphenols through HPLC-PDA system refers to the appendix E

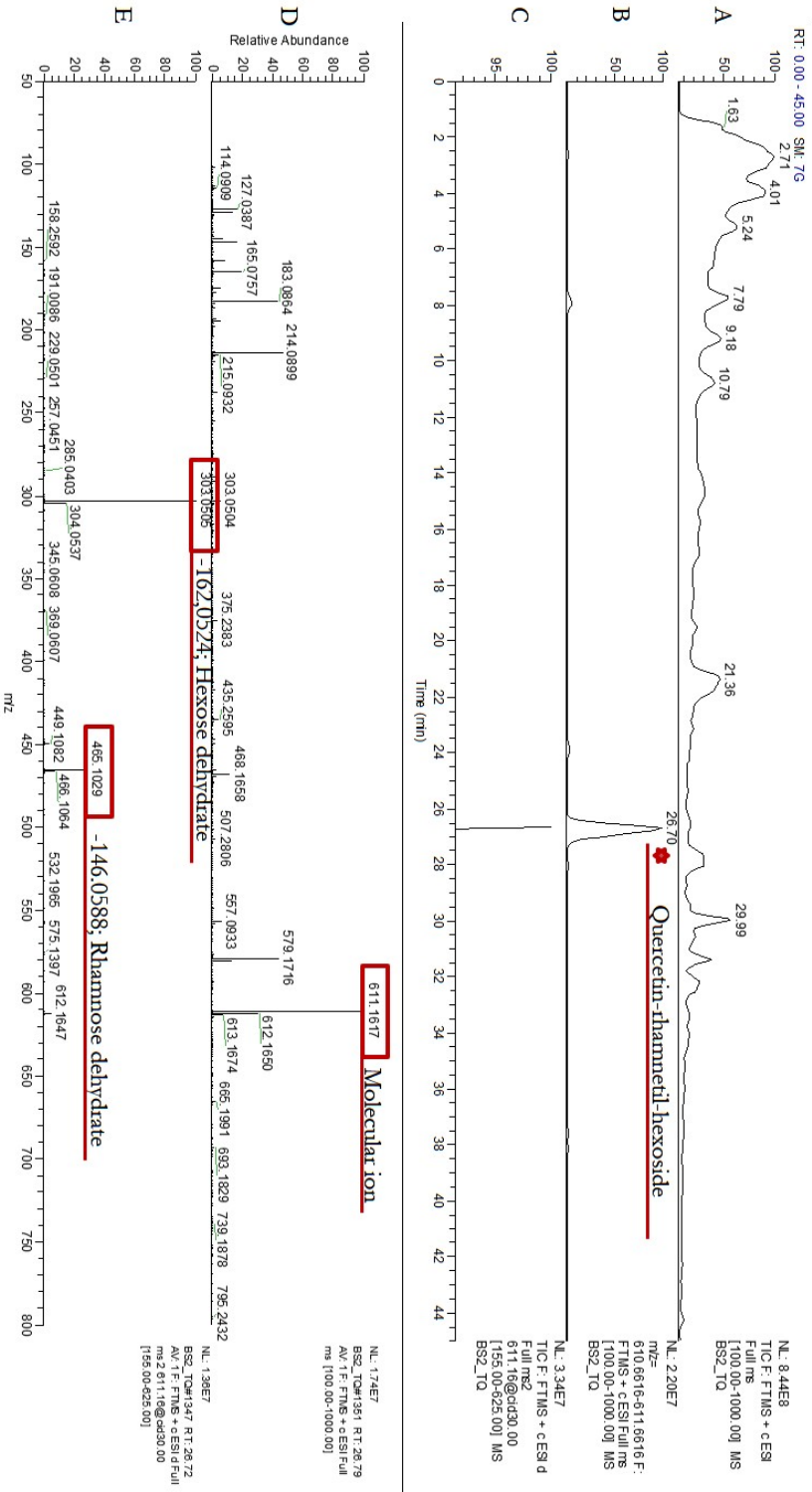


Figure 9.4.: Example of chromatograms and MS spectra obtained during targeted metabolomics analysis of lyophilized berries samples for a specific analyte (the quercetin-rhamnetil-hexoside). Raw A: TIC; raw B: XIC of m/z 611.1617; raw C: MS2 chromatogram of m/z 611.1617; raw D: MS spectrum of spotted peak in line B; raw E: Full MS2 of m/z 611.1617

B. Injection volume was 20 μL . The tuning parameters used for the ESI source were: capillary temperature 270° C, the flow rate of sheath gas and auxiliary gas set at 35.0 and 15.0 arbitrary units, capillary voltage 8.0 V, source voltage 4.5 kV and tube lens 65 V. Full scan spectra were acquired in positive ion mode in the range 250-1000 m/z with the resolution of 30000 (FWHM). MS^n spectra were acquired in the range between ion trap cut-off and precursor ion m/z values.

On the other hand, for other polyphenol compounds (phenolic acids, and flavonols) a biphenyl stationary phase (Pinnacle DB BiPh, 150 \times 2.1 mm, 3 μm particle size; Resteck, Milan, Italy) at 200 $\mu\text{L}/\text{min}$ flow rate was used. The elution solvent adopted was methanol (B) and ammonium acetate 5mM (A). The gradient profile was 0-3 min to 2% of B, 3-60 min from 2 to 62% B, 60-65 min from 62 to 100% B. The tuning parameters used for the ESI source were: capillary temperature 270° C, the flow rate of sheath gas and auxiliary gas set at 35.0 and 15.0 arbitrary units, capillary voltage 16.0 V, source voltage 3.5 kV and tube lens 55 V. Full scan spectra were acquired in positive ion mode in the range 100-1000 m/z with the resolution of 30000 (FWHM). MS^n spectra were acquired in the range between ion trap cut-off and precursor ion m/z values. Both cited methods reveal satisfactory linearity in the range 0.01 (LLOQ) – 500 $\mu\text{g}/\text{g}$.

9.4. Results and discussion

The identification and quantification of polyphenols in fruits matrix are conducted mainly through HPLC-PDA-MS techniques [26]. The identification of phenolic compounds was investigated employing the information obtained by UV-VIS spectra, the accurate mass of precursor ions and tandem MS product ions (second approach reported, 9.3), as well the online databases, the XCMS results and the literature (the first approach proposed, 9.3). The UV-Vis spectra of the compounds were recorded between 200 and 650 nm: as a matter of fact, flavonols and anthocyanins have a maximum absorption at 520 nm and hydroxycinnamic acids at 260 nm.

The extracts were classified based on the content of the various polyphenols. The results showed that the berries with the highest and the lowest

polyphenols abundance were blueberries (Cv. Misty) and white gooseberries, respectively. Indeed, within the berry family, blueberries and blackberries contain high quantities of antioxidants species as anthocyanin or compounds [7, 29]. All the different compounds detected or tentatively identified, their $[M + H]^+$, the occurrence within the samples, and the most important MS/MS fragments reported in order of decreasing intensity are shown in the table 9.3. Up to 70 different compounds have been detected in the samples. Most of them belong to the flavonols class: mono and di-glycosylated compounds of myricetin, kaempferol, and quercetin are the most abundant. As reported in literature [22] flavonols as quercetin, myricetin, and kaempferol and their aglycons are the most abundant ones in fruits like berries. The glycosides quercetin species represent, between all the detected molecules, the highest percentage among flavonols family in all the berries analysed [16]. Many of these phytochemicals aglycons compounds detected exist as mono-, di- and ter-glycosidic polyphenols, and the sugar units are linked in different positions on the skeletons of the molecules. The presence of these analytes is known: the flavonols and their glycosylated analogues are potent antioxidants that the plants synthesize to protect themselves from reactive oxygen species [25].

Glucose, galactose, and arabinose are the most frequently occurring sugars bounded to polyphenol species to generate the different aglycons [16]. In our analysis, given the impossibility of attributing a correct name to the sugar linked to the flavonoid of interest, generic names like “hexose” or “pentose” are reported.

On the contrary, if we consider other analytes belonging to the phenolic acids or coumarins families, they are present in lower concentration and not in all berry’s extracts. As far as anthocyanins classes are concerned, anthocyanidins of cyanidin, delphinidin, malvidin, and petunidin were recorded in all berries with the exceptions of goji, white currant, and white gooseberry 9.3. Among fruits, berries are one of the richest sources of these analytes in nature [29], and results are consistent with previous research [16]: white currant is one of the berries with the lowest amount of total flavonols molecules. Cyanidin was the most commonly occurring anthocyanidin found in five fruits (blackberry, black currant, raspberry, red cur-

rant, and red gooseberry), meanwhile, malvidin and petunidin glycosides were detected the most in blueberries. Delphinidin glycosides were the dominant form of anthocyanins in black currant and blueberry fruits. These results are confirmed by [13]. The authors identified different anthocyanins such as delphinidin-3-O-galactoside, delphinidin-3-O-glucoside, petunidin-3-O-glucoside, and malvidin-3-O-galactoside in blueberries and delphinidin-3-O-rutinoside and cyanidin-3-O-rutinoside as the significant anthocyanin's species in black currant. However, it is essential to remember that the polyphenol concentrations among the berries and foods can significantly differ depending on various factors such as genetic, technological factors, season, cultivars species, and growing location[7].

Finally, a schematic histogram of the data obtained is reported in figure 9.5. Each chromatogram bar represents the single berry analysed. The height of each bar is instead provided by the sum of all the areas of the chromatographic peaks of the polyphenols found for each single species. This graphical view shows clearly how the berries with the highest concentration in polyphenolic compounds are the blueberries and the blackberries fruits [7, 29]. On the contrary, white gooseberry and red currant samples are the ones with the lowest polyphenolic amount. Finally, all the results founded are according to scientific literature [23].

Table 9.3.: Polyphenolic compounds identification by HPLC-PDA-HRMS in positive polarity

Analyte	[M+H] ⁺	Occurrence in samples	MS/MS fragments	Identification by
Kaempferol	287.0555	All the samples	137.0232, 269.2270	Standard
Epicatechin	291.0868	BB, BC, DB, LB, RC, WC	139.0388, 273.0764	Standard
Catechin	291.0868	BC, BB, DB, LB, RA, RC, RG, WC	139.0388, 273.0765	Standard
Quercetin	303.0604	All the samples	257.0450, 285.0401	Standard
Coumaric acid hexoside 1	327.0868	BC, GO, RA, RC, RG, WC, WG	295.0607, 165.0547	Literature [3, 24, 25]
Coumaric acid hexoside 2	327.0868	BC, RG	295.0607, 165.0547	Literature [3, 24, 25]
Caffeoyl-hexose 1	343.1029	BB, BC, MB, RC, RG, WC, WG	325.0714, 191.0342, 181.0861	Literature [25]
Caffeoyl-hexose 2	343.1029	BB, MB, RC, RG, WG	325.0714, 191.0342, 181.0861	Literature [25]
Chlorogenic acid 1	355.1030	LB, RG, WG,	163.0389	MS [4]
Chlorogenic acid 2	355.1030	BB, BC, DB, GO, MB, RA	163.0389	Literature [4]
Ferulic acid hexoside 1	357.1187	BB, MB	163.0391, 195.0655	Literature [9]
Ferulic acid hexoside 2	357.1187	BB, MB	163.0391, 195.0655	MS
Cyanidin-pentoside	419.1134	LB, RG, WG	287.0556	Literature [3])
Kaempferol-rhamnoside	433.1135	MB, WC, WG	287.0555	MS
Delphinidin-pentoside	435.0927	BB, DB, MB	303.0505	Literature [3]
Quercetin-arabinoside	435.0927	BB, DB, LB, MB, RA, RG	303.0608	Literature [25]
Naringenin-hexoside	435.1291	BB, BC, LB, RA, RC	273.0971	MS
Kaemperol-hexoside	449.0880	BC, RA, RC, WC	287.0557	Literature [25]
Petunidin-pentoside	449.1082	BB, DB, MB	317.0661	Literature [1, 3]
Cyanidin-hexoside (Chrysanthemine)	449.1084	BC, LB, RG, WG	287.0555	Literature [3]
Kaemperol-glucuronide 1	463.1032	LB	287.065	MS
Kaemperol-glucuronide 2	463.1032	LB	287.065	MS
Malvidin-arabinoside	463.1239	BB, DB, MB	331.0818	Literature [3]
Rhamnetin-rhamnoside	463.1239	GO, WC	317.2076	MS
Quercetin-hexoside	465.1031	All the samples	303.0608	Literature [25]; [14]

Continues in next page

Continues from previous page

Analyte	[M+H] ⁺	Occurrence in samples	MS/MS fragments	Identification by
Delphinidin-hexoside 1	465.1032	BB, DB, GO, MB, WC	303.0506	Literature [3]
Delphinidin-hexoside 2	465.1032	BB, BC, DB, GO, MB	303.0506	Literature [3]
Ellagic acid acetyl-xyloside	477.0829	LB	301.0461, 179.0054	MS
Quercetin-glucuronide	479.0825	LB	303.0608	Literature [1, 25]
Rhamnetil-glucuronide	479.1186	BB, MB, RG, WC	317.2076	Literature [1, 25]
Petunidin-hexoside 1	479.1188	BB, DB, MB	317.0662	Literature [4]
Petunidin-hexoside 2	479.1188	BB, MB	317.0662	Literature [4]
Myricetin-hexoside	481.0980	BB, BC, DB, LB, MB, RA, RC, RG WC, WG	319.0572	Literature [25]
Malvidin-hexoside 1	493.1343	BB, DB, MB	331.0818	MS
Malvidin-hexoside 2	493.1343	MB	331.0818	MS
Quercetin acetyl-hexoside 1	507.1303	BB, MB, RA	303.0505	Literature [10]
Quercetin acetyl-hexoside 2	507.1303	BB, DB, LB, MB	303.0506	Literature [10]
Cynarine 1	517.1345	BB, DB, GO, MB	499.1238	MS
Cynarine 2	517.1345	MB	499.1238	MS
Cynarine 3	517.1345	GO	499.1238	MS
Kaemperol-malonyl-hexoside	535.1086	BC, LB, MB, RC, WC	287.0650, 491.1348	MS
Malvidin-acetylhexoside	535.1452	BB, MB	331.0819	Literature [4]
Quercetin-malonyl-hexoside	551.1219	BB, BC, DB, LB, MB, RC, WC	303.0607	Literature [25]
Rhamnetil-malonyl -hexoside	565.1194	BB, BC, DB, LB, MB, RC, WC	317.2075	MS
Pelargonidin-sambubioside	565.1760	RC	271.2430	MS
Catechin dimer	579.1505	BB, BC, DB, LB, RG, WC	291.087, 427.1169	MS
Catechin dimer	579.1505	BB, DB, RA, WC	291.087, 427.1169	MS
Cyanidin-hexosil-pentoside	581.1708	BC, RA	449.0394, 419.1182	MS
Cyanidin-rutinoside	595.1664	BC, LB, RA, RG,	287.0556	Literature [12]
Kaemperol-rhamnetil-hexoside	595.1665	GO, MB, RG, WC	287.0650, 449.1230	MS
Quercetin-xylosil-hexoside	597.1663	LB, RC, RG, WC, WG	303.0605, 465.0927, 435.1341	MS
Delphinidin-rhamnetin-hexoside	611.1514	BC, RA	303.0554, 465.1079	MS

Continues in next page

Continues from previous page

Analyte	[M+H] ⁺	Occurrence in samples	MS/MS fragments	Identification by
Kaempferol -dihexoside	611.1616	All the samples	287.0652, 449.123	Literature [12]
Quercetin-rhamnetil-hexoside	611.1616	BB, DB, RA, MB	303.0608, 465.1187	MS [1]
Quercetin-galloyl-hexoside	617.1146	RA	303.0605, 465.1028	MS
Rhamnetin-hexosil-rhamnoside	625.1771	All the samples	317.0663, 479.1186	Literature [25]
Quercetin-dihexoside	627.1567	DB, LB	303.0605	Literature ([25])
Myricetin-hexosil-rhamnoside	627.1567	BC, RC, RG, WC, WG	319.0570, 465.103	MS
Syringetin-rutinoside 1	655.2036	MB	357.0975, 623.1766, 389.1237	Literature [6, 25]
Syringetin-rutinoside 2	655.2036	MB	357.0975, 623.1766, 389.1237	Literature [6, 25]
Cyanidin-sambubiosil-rhamnetile1	727.235	BC, RC	287.0555, 433.3021	MS
Cyanidin-sambubiosil-rhamnetile 2	727.235	BC, RC	287.0555, 433.3020	MS
Kaemperol-dirhamnetil-hexoside	741.2253	RA, RC	287.0557, 595.1662, 449.1082	MS
Quercetin-xylosil-rhamnosil-hexoside	743.2025	LB, RA, RC, RG, WC, WG	303.0506, 611.1613, 465.1031	MS
Kaemperol-rhamnosil-dihexoside	757.2202	GO	449.1080, 595.1661, 287.0556	MS
Quercetin-dirhamnosil-hexoside	757.2202	GO, LB, RC, RG, WG	611.1614, 303.0506	Literature [25]
Cyanidin-rhamnetin-dihexoside	757.2456	RA	595.1930, 433.1404, 287.0556	MS
Quercetin-dihexosil-rhamnoside	773.2153	GO, RA, RG, WC	465.1030, 627.1561, 611.1612, 303.0506	Literature [25]
Myricetin-hexosil-dirhamnoside	773.2153	RC, WG	319.0455, 627.1563	MS
Myricetin-dihexosil-rhamnoside	787.2309	GO, RG, WG	479.1186, 641.1716, 317.0662	MS

The table ends here

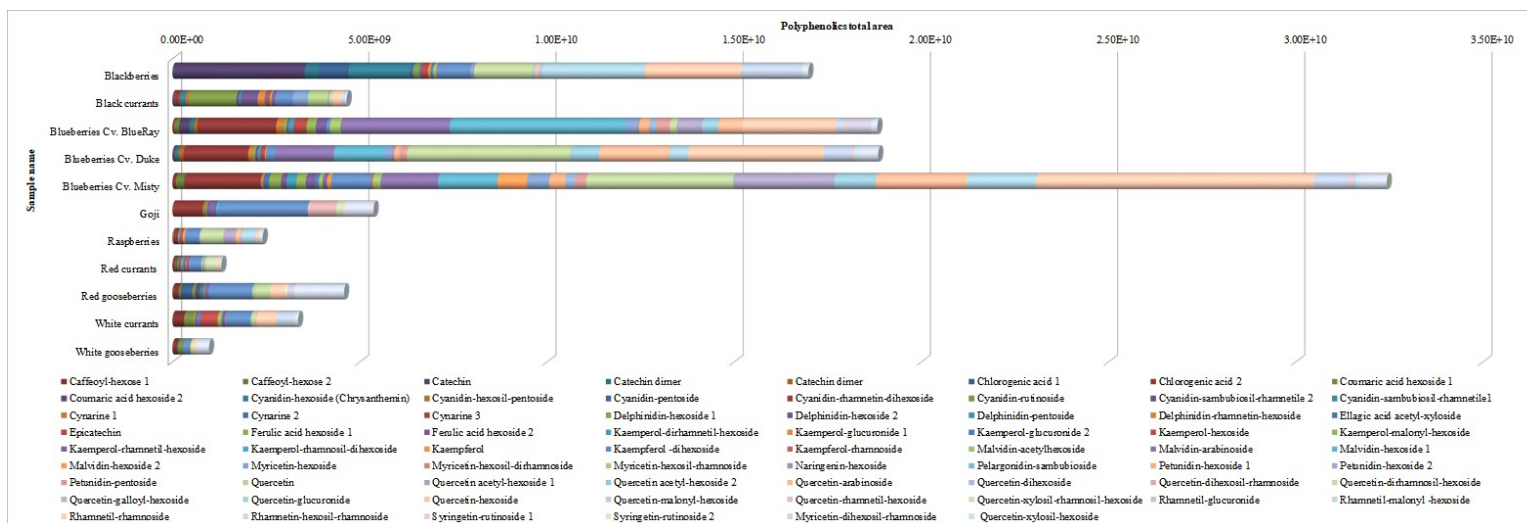


Figure 9.5.: Histogram of the 70 different analytes detected in the eleven berry fruits investigated

9.5. Conclusion and future prospects

In this project, we evaluated the bioactive compounds, antioxidant and antiradical activities, and fatty acid profile of extracts of eleven berries cultivated in North-Western Italy, a geographical area where berry production has a high economic interest.

The HPLC-PDA-ESI HRMS global metabolomics analysis and antioxidant and antiradical activity tests have been able to generate exciting results. As a matter of fact, these results showed that berry fruits have a demonstrated antioxidant capacity due to their constituents. Furthermore, they are rich in flavonoids and fatty acids. Among the investigated berries, black currants and blueberries are the most valuable fruits (best antioxidant and antiradical activity), and their properties could be useful for food and health purposes.

In conclusion, the methods established can be applied for future studies on the detection of polyphenols and anthocyanins analytes in similar matrices. Furthermore, one of the future purposes of this work is to analyse wild berry fruits (black mulberry, elderberry, lingonberry, cornelian cherry, hawthorn, and rosehip) too.

These fruits are often incorporated in traditional medicine and cuisine [5, 17], and they are superior to many commercially grown small fruits in terms of antioxidant potency and antioxidant concentration [5, 21]. So, we aim to compare wild and cultivated berry grown in North-Western Italy and monitor the levels of the analytes of interest in these matrices.

References of Chapter 9

- [1] C. Ancillotti et al. “Polyphenolic profiles and antioxidant and antiradical activity of Italian berries from *Vaccinium myrtillus* L. and *Vaccinium uliginosum* L. subsp. *gaultherioides* (Bigelow) S.B. Young”. In: *Food Chemistry* 204, 176-184 (2016). DOI: 10.1016/j.foodchem.2016.02.106 (cit. on pp. 176–178).
- [2] E. Balogh et al. “Application of and correlation among antioxidant and antiradical assays for characterizing antioxidant capacity of berries”. In: *Scientia Horticulture* 125(3), 332-336 (2010). DOI: 10.1016/j.scienta.2010.04.015 (cit. on p. 165).
- [3] G. Borges et al. “Identification of flavonoid and phenolic antioxidants in black currants, blueberries, raspberries, red currants, and cranberries”. In: *Journal of Agriculture and Food Chemistry* 58(7), 3901-3909 (2010). DOI: 10.1021/jf902263n (cit. on pp. 167, 176, 177).
- [4] A. Brito et al. “Anthocyanin characterization, total phenolic quantification and antioxidant features of some Chilean edible berry extracts”. In: *Molecules* 19(8), 10936-55 (2014). DOI: 10.3390/molecules190810936 (cit. on pp. 176, 177).
- [5] B. Dinda et al. “*Cornus mas* L. (cornelian cherry), an important European and Asian traditional food and medicine: Ethnomedicine, phytochemistry and pharmacology for its commercial utilization in drug industry”. In: *Journal of Ethnopharmacology* 193, 670-690 (2016). DOI: 10.1016/j.jep.2016.09.042 (cit. on p. 180).
- [6] R. Fang et al. “Enhanced profiling of flavonol glycosides in the fruits of sea buckthorn (*Hippophae rhamnoides*)”. In: *Journal of Agriculture and Food Chemistry* 61(16), 3868-3875 (2013). DOI: 10.1021/jf304604v (cit. on p. 178).
- [7] G. Giovanelli et al. “Comparison of polyphenolic composition and antioxidant activity of wild Italian blueberries and some cultivated varieties”. In: *Food Chemistry* 112(4), 903-908 (2009). DOI: 10.1016/j.foodchem.2008.06.066 (cit. on pp. 174, 175).

-
- [8] V. Girgenti et al. “Chemical characterization and antioxidant evaluation of muscadine grape pomace extract”. In: *Sustainability* 8(10), 1027-1043 (2016). DOI: 10.3390/su8101027 (cit. on pp. 165, 166).
- [9] F Ieri et al. “Phenolic composition of “bud extracts” of *Ribes nigrum* L., *Rosa canina* L. and *Tilia tomentosa* M.” In: *Journal of Pharmaceutical and Biomedical Analysis* 115, 1-9 (2015). DOI: 10.1016/j.jpba.2015.06.004 (cit. on p. 176).
- [10] G.H. Jang et al. “Characterization and quantification of flavonoid glycosides in the *Prunus* genus by UPLC-DAD463QTOF/MS”. In: *Saudi Journal of Biological Sciences* 25(8), 1622-1631 (2018). DOI: 10.1016/j.sjbs.2016.08.001 (cit. on p. 177).
- [11] G. Koutsimanis et al. “Influences of packaging attributes on consumer purchase decisions for fresh produce”. In: *Appetite* 59(2), 270-280 (2012). DOI: 10.1016/j.appet.2012.05.012 (cit. on p. 166).
- [12] A. Lavola et al. “Bioactive polyphenols in leaves, stems, and berries of Saskatoon (*Amelanchier alnifolia* Nutt.) cultivars”. In: *Journal of Agriculture and Food Chemistry* 60(4), 1020-1027 (2012). DOI: 10.1021/jf204056s (cit. on pp. 177, 178).
- [13] S.G. Lee et al. “Contribution of anthocyanin composition to total antioxidant capacity of berries”. In: *Plant Foods for Human Nutrition* 70(4), 427-432 (2015). DOI: 10.1007/s11130-015-0514-5 (cit. on p. 175).
- [14] P. Liu et al. “Phenolic compounds in hawthorn (*Crataegus grayana*) fruits and leaves and changes during fruit ripening”. In: *Journal of Agriculture and Food Chemistry* 59(20), 11141-11149 (2011). DOI: 10.1021/jf202465u (cit. on p. 176).
- [15] G. Manganaris et al. “Berry antioxidants: small fruits providing large benefits”. In: *Journal of the Science of Food and Agriculture* 94(5), 825-833 (2014). DOI: 10.1002/jsfa.6432 (cit. on p. 165).
- [16] M. Mikulic-Petkovsek et al. “HPLC-MSn identification and quantification of flavonol glycosides in 28 wild and cultivated berry species”. In: *Food Chemistry* 135(4), 2138-2146 (2012). DOI: 10.1016/j.foodchem.2012.06.115 (cit. on p. 174).

- [17] S.F. Navabi et al. “Antibacterial Effects of Cinnamon: From Farm to Food, Cosmetic and Pharmaceutical Industries”. In: *Nutrients* 7(9), 7729-48 (2011). DOI: 10.3390/nu7095359 (cit. on p. 180).
- [18] Corriere Ortofrutticolo. *Piccoli frutti, raddoppiati i consumi in 10 anni*. 2019. URL: <http://www.corriereortofrutticolo.it/2019/06/05/piccoli-frutti-raddoppiati-i-consumi-in-10-anni-ecco-come-gestire-il-boom/> (cit. on pp. 165, 166).
- [19] Oxford University Press. *Oxford Dictionaries*. 2019. URL: <https://www.oxforddictionaries.com> (cit. on p. 165).
- [20] L.J. Rowland et al. *Genetics, Genomics and Breeding of Berries*. Chap. 12, “Blueberry” pages 1-32. Boca Raton, Florida: Folta, & C. Kole (Eds.), 2011 (cit. on p. 166).
- [21] D. Samec et al. “Antioxidant potency of white (*Brassica oleracea* L. var. capitata) and Chinese (*Brassica rapa* L. var. pekinensis (Lour.)) cabbage: The influence of development stage, cultivar choice and seed selection”. In: *Scientia Horticulturae* 128(2), 78-83 (2011). DOI: 10.1016/j.scienta.2011.01.009 (cit. on p. 180).
- [22] C.L.L. Saw et al. “The berry constituents quercetin, kaempferol, and pterostilbene synergistically attenuate reactive oxygen species: Involvement of the Nrf2-ARE signaling pathway”. In: *Food Chemistry and Toxicology* 72, 303–311 (2014). DOI: 10.1016/j.fct.2014.07.038 (cit. on p. 174).
- [23] P. Skenderidis et al. “Chemical properties, fatty-acid composition, and antioxidant activity of goji berry (*Lycium barbarum* L. and *Lycium chinense* Mill.) fruits”. In: *Antioxidants* 8(3), 60 (2019). DOI: 10.3390/antiox8030060 (cit. on p. 175).
- [24] S. Skrovankova et al. “Bioactive compounds and antioxidant activity in different types of berries”. In: *International Journal of Molecular Sciences* 16, 24673–24706 (2015). DOI: 10.3390/ijms161024673 (cit. on p. 176).

- [25] Y. Tian et al. “Phenolic compounds extracted by acidic aqueous ethanol from berries and leaves of different berry plants”. In: *Food Chemistry* 220, 266–281 (2017). DOI: 10.1016/j.foodchem.2016.09.145 (cit. on pp. 174, 176–178).
- [26] K. Tzima et al. “Qualitative and quantitative analysis of polyphenols in Lamiaceae plants. A review”. In: *Plants (Basel)* 7(2), pii: E25 (2018). DOI: 10.3390/plants7020025 (cit. on p. 173).
- [27] X. Wang et al. “Exploring Perceptions of Raspberries and Blueberries by Italian Consumers”. In: *Food Chemistry* 123(4), 1156–1162 (2010). DOI: 10.1016/j.foodchem.2010.05.080 (cit. on pp. 165, 171).
- [28] XCMS. *The XCMS Online*. 2020. URL: https://xcmsonline.scripps.edu/landing_page.php?pgcontent=mainPage (cit. on p. 169).
- [29] Q. You et al. “Comparison of anthocyanins and phenolics in organically and conventionally grown blueberries in selected cultivars”. In: *Food Chemistry* 125(1), 201–208 (2018). DOI: 10.1016/j.foodchem.2010.08.063 (cit. on pp. 174, 175).

Untargeted analysis: The Spacemice project

10

10.1. Introduction

This chapter and the following chapter 11 will concern experiences matured during my PhD period abroad at the Tohoku University under the supervision of Prof. Daisuke Saigusa, Prof. Seizo Koshiba, and Prof. Masayuki Yamamoto. So, before deal with the different analyses and results achieved, a brief introduction upon this institution and the hosting department is deserved.

The Tohoku University and the ToMMO department

Tohoku University is the third oldest Imperial University in Japan, and it is considered one of the three most prestigious in the country and the 36th best University in the World [26]. It is located in Sendai, the second-largest city north of Tokyo. From its foundation, the University closed only in 2011 for a few months due to the Great East Japan Earthquake and tsunami in Tohoku Region. Most campuses were unaffected by the earthquake, but some buildings suffered damage [26]. The university has five main campuses located in Sendai city. Students are generally split across these campuses by subject. For what concerns me, I was applied to the Seiryō Campus, where there are located the Department of Medicine (inside the Tohoku University Hospital) and the Department of Dentistry. In particular, I was employed in Tohoku Megabank Medical Organization, directed by Prof. Masayuki Yamamoto, in the Integrative Genomics Group of Prof. Seizo Koshiba under the supervision of Prof. Daisuke Saigusa.

Tohoku University - Tohoku Medical Megabank Organization (ToMMO) was founded to establish an advanced medical system to foster the reconstruction from the Great East Japan Earthquake in 2011 [27]. This organization was established on February 1, 2012, by Tohoku University as a 10-years project. The principal aim of this institution is to develop a biobank that combines medical and genome information during the pro-

cess of rebuilding the community medical system and supporting health and welfare in the Tohoku area. With the combination of medical information and genome information, Tohoku Medical Megabank Project is constructing an exceptional biobank in the Tohoku region, which contributes to the restoration of medical services in the disaster area and revitalizes related industries.

Tohoku Medical Megabank Organization consists of 6 research departments, a Department of Public Relations and Planning, a Department of Administration, 25 groups which promote the project individually, an Ethics Committee, and an Advisory board for Tohoku Medical Megabank Organization. In my case, I was applied in the Department of Integrative Genomics – Genomic and Omics Analyses Sector, Group of Omics Analysis headed by Prof. Seizo Koshihara. The aims which guide the group are the discovery of useful biomarkers through the global identification and quantification (i.e., multi-omics studies) of proteins and small compounds found in blood or urine specimens provided by cohort-study participants. Such multi-omics studies can contribute to the advancement of personalized prevention and treatment of diseases, as well as identification of disease mechanisms and development of novel therapeutics.

As already anticipated in chapters 1 and 2 during my period at Tohoku Medical Megabank Organization, I could learn about the untargeted metabolomics analyses conducted by high-performance liquid chromatography coupled to high-resolution mass spectrometry (HPLC-HRMS) techniques.

The SpaceMice Project

While staying in space, a human being faces different medical risks, so-called space stresses [3]. Space is an unfriendly and dangerous place to operate. For space travelers deadly dangers are always behind the corner: the micro-gravity environment interference the normal signal transduction between cells and causes a density dropping at over 1% per month in bones [1, 18]; the exposition to cosmic radiation increases the lifetime or dying from cancer [21]; isolation of people in small spaces, far away from their families and friends; oxidative stress caused by exposition to space radiation (in the ISS, “astronauts receive over ten times the radiation than what’s naturally

occurring on Earth” [1]), etc.

To elucidate the potential risks of space-life, and to adopt the best practices to avoid the space stresses, this project focuses its attention upon the Nrf2 transcription factor [22, 29]. The nuclear factor erythroid-2-related factor 2 (Nrf2) is a transcription factor of the leucine zipper family, and Keap1 (Kelch-like ECH-associated protein 1) is its specific repressor [13]. These two proteins mediate cellular response to oxidative stress and electrophilic xenobiotics [20]. So, Nrf2 and Keap1 factors are expected to play an essential role in the defensive body mechanism against space stress.

The mission is intended to return six Nrf2-deficient mice (called Nrf2KO) and six native mice to Earth after being bred on International Space Station (ISS), clarify the contribution of activated Nrf2 contributes to biological defence to space stress, and demonstrate the effectiveness of Nrf2-inducing agent in reducing risks in space.

The space mission took place for 31 days, from April 4 to May 5, 2018. During this period, the mouse breeding was conducted inside a special module, made by JAXA - Japan Aerospace Exploration Agency, the so-called Pressurized Module of Japanese Experiment Module (JEM) *Kibo*[®]¹. In the end, on May 7, 2018, the cage containing the space-bred mice was handed over from NASA to JAXA, which subsequently confirmed the survival of all mice. This mission has brought the world’s first achievement of the long-duration stay in space and return all alive of gene-knockout mice

10.2. Aim of the work

The SpaceMice project is an ambitious study program whose purpose is to discover the highest amount of new potential biomarkers linked to oxidative stress, useful to discriminate between the wildtype and the Nrf2KO mice groups.

The samples under analysis consist of different kinds of tissues belong to different groups of space-bred mice. The plasma and the blood were taken

¹The Kibo[®] module, which means hope in Japanese, is Japan’s first human-rated space facility and the Japan Aerospace Exploration Agency’s (JAXA’s) EARLY contribution to the International Space Station (ISS) program. Currently (2019), A wide variety of scientific, medical, and educational experiments are conducted on Kibo[®]. As a part of the ISS, Kibo[®] provides extensive opportunities for space environment utilization [3].

from the mice before and after sending them in the space. On the contrary, organs' mice (brain, kidney, liver, spleen) are collected directly at Tohoku University, and they are compared with organs of mice nourished in the same way but breeding at the University and not on the ISS.

To fulfil this project, it's quite easy to understand that the application of a single analysis cannot be enough. So, untargeted metabolomics approaches using HPLC-HRMS techniques were used:

- Semi-targeted analysis coupling Biocrates kits with UHPLC-HRMS instrument;
- General untargeted metabolomics analysis using UHPLC-HRMS instrument;
- General untargeted metabolomics analysis using ambient mass spectrometry techniques;
- General untargeted lipidomics analysis using UHPLC-HRMS instrument.

For each of the bullet points listed here, the detailed results obtained by each approach will be treated in separated sub-sections within sections 10.3 and 10.4 .

10.3. Sample preparation and method settings

Semi-targeted analysis: Biocrates kit

The SpaceMice samples are extremely costly, precious, and no reproducible ones. So, from the analytical method applied to biomedical research, we require at least high levels of accuracy, reproducibility of results. Furthermore, the analytical method must be already validated before proceeding and must be viable with the low-matrix amount (10 μ L per sample or less such be a condition required by the method). The targeted metabolomics kits by Biocrates provides these benefits. They enable researchers to obtain quantitative information on up to hundreds of metabolites, with high accuracy, as well as reproducibility over time and between laboratories

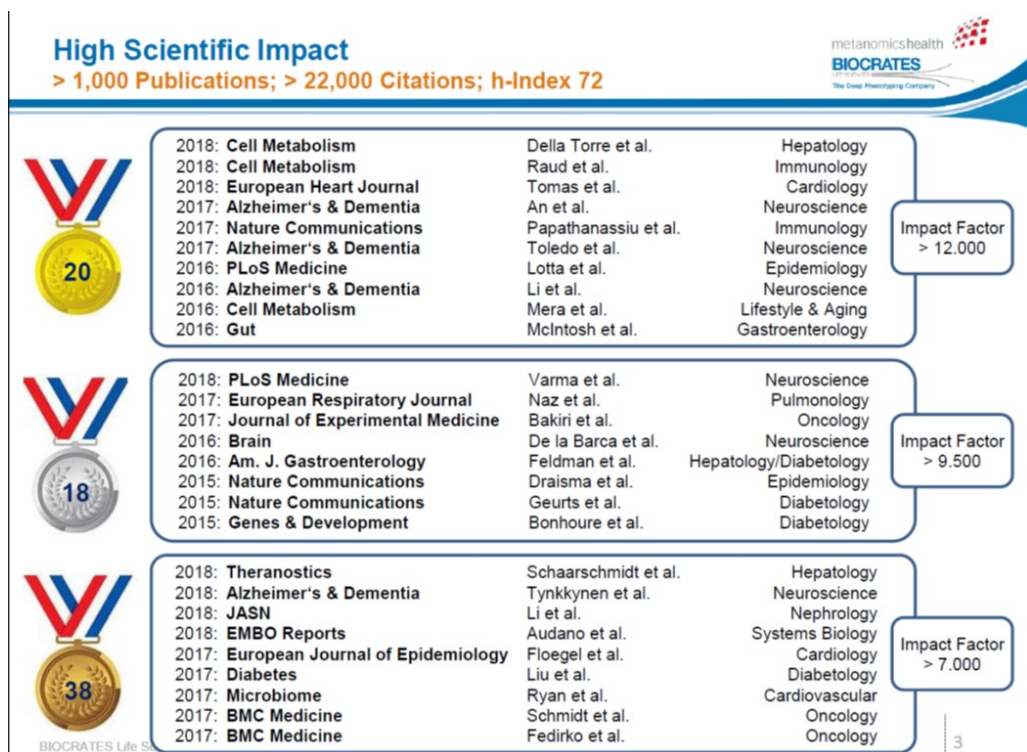


Figure 10.1.: Schematic report of Biocrates kits' application in scientific publications in different fields of research in the last 3 years [2]

[2]. Thanks to these kits, a piece of comprehensive and truthful information about the metabolic phenotype of a subject can be obtained in just a few minutes per sample. Moreover, the kits rely on thoroughly validated, quality-controlled analytical methods that are optimized for their purpose, and the LC-MS instruments can be easily combined for analytical purposes [12, 14, 16].

In literature, among the last few years, these kinds of kits are massively used by researchers. As reported in figure 10.1, targeted metabolomics kits have contributed to hundreds of scientific publications in many different fields of research, not only linked to mass spectrometry.

For our purpose, from all the Biocrates kit available, we used the one called AbsoluteIDQ[®] p400 HR Kit. The AbsoluteIDQ[®] p400 HR Kit is a complete solution for broad lipid and metabolic profiling on high resolution, accurate mass (HRMS) Q Exactive[™] mass spectrometers. It provides

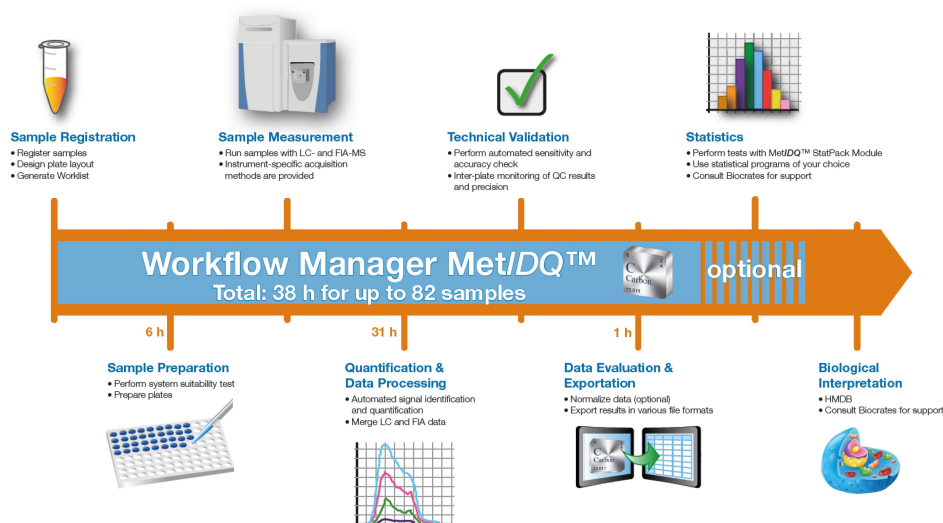


Figure 10.2.: Workflow provided by Biocrates Life Sciences AG

Table 10.1.: Metabolite classes detectable with AbsoluteIDQ® p400 HR Kit. The numbers in brackets are the number of analytes detectable in that specific class

	Small Molecules	Polar Lipids	Neutral Lipids
	Amino Acids (21)	Phosphatidylcholines (172)	Acylcarnitines (55)
	Biogenic Amines (21)	Lysophosphatidylcholines (24)	Diglycerides (18)
	Hexoses (1)	Sphingomyelins (31)	Triglycerides (42)
		Ceramides (9)	Cholesteryl Esters (14)
Total analytes detectable	43	236	129

quantification of a large range of analytes, with high inter-laboratory, inter-instrument, and longitudinal reproducibility[2]. This kit covers up to 408 metabolites belonging from eleven different metabolite classes, which are known to be relevant in a multitude of pathophysiological processes. The different metabolites detectable are report in table 10.1.

The operative workflow to follow to obtain statistically significant data is furnished directly together Biocrates Kit: reagents & consumables, quality controls, process quantification guidance, etc. The kit is furnished in one box, ready-to-use, and user-friendly. A scheme of this workflow is reported in figure 10.2.

Thanks to the use of this kit, we guarantee to determine theoretically a high number of analytes (more than 400 molecules)², using only 10 μ L of

² AbsoluteIDQ® p400 HR Kit is the Biocrates with the highest number of analytes theo-

plasma sample required.

The solvents used were different between LC and FIA modes. For LC part, solvent A (1000 mL) was 1000 mL water + 2 mL formic acid, and solvent B (500 mL) 500 mL acetonitrile + 1 mL formic acid. On the contrary, for FIA part 290 mL of MeOH, HPLC grade was mixed with 1 mL of a solution called FIA Mobile Phase Additive. No more pieces of information are available about this additive. For both methods, wash solvent is a mixture composed of 25% acetonitrile, 25% methanol, 25% isopropanol, 25% water; the seal wash was a 10% methanol, 90% H₂O. The two HPLC gradient used for LC and FIA chromatographic runs are reported in figure 10.3, directly extrapolated from the Biocrates User manual. Due to the trade secret applied by Biocrates, no information about the LC column furnished for the analyses are available. The only HPLC parameters known by the operator are the flux (0.8 mL/min in LC part; 0.2 mL/min in FIA part), the injection volume (5 μ L in LC, 20 μ L in FIA), and the column oven temperature (50° C in both methods). Concerning HPLC and FIA setups, a last but fundamental aspect must be remarked: during LC analyses, the chromatographic column must be connected to the system and stored in the column oven; during FIA mode, the autosampler of HPLC section is directly connected to the source of MS instrument. In this case, as also suggested by the kit, less is the distance (and the length of the peek tube) between the autosampler and the ion source better are the results.

On the contrary, the MS settings are known and are resumed in table 10.2. However, for a complete extraction procedure step-by-step and all the HPLC-HRMS parameters (calibration MS procedures, standard calibration curves, etc.), consult the extended manual on the Biocrates website [2].

The global metabolomics analysis: UHPLC-HRMS

To conduct the untargeted metabolomics analysis an UHPLC Dionex UltiMate 3000 coupled with a Q ExactiveTM Hybrid Quadrupole-Orbitrap Mass Spectrometer (both from Thermo Scientific) were used. Furthermore,

retically detectable by one single analysis. By the way, the company provides different kits, like AbsoluteIDQ[®] p180 Kit, AbsoluteIDQ[®] Stero17 Kit, or Biocrates[®] Bile Acids Kit depending on the research's purpose or the budget available.

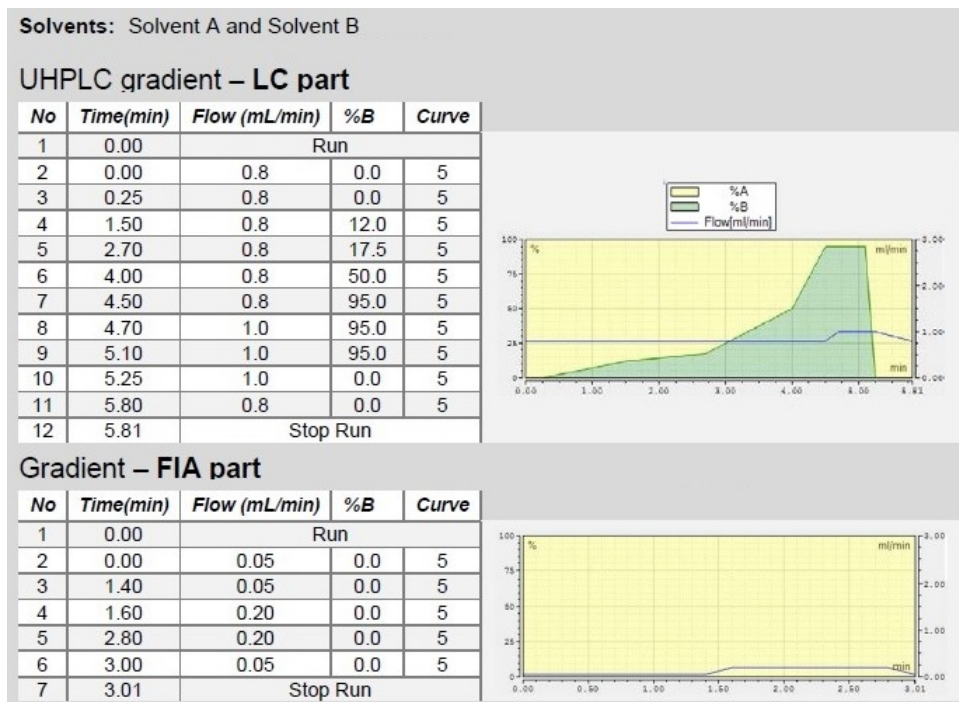


Figure 10.3.: The two chromatographic gradients used for LC analyses (above in the figure) and FIA analyses (below in the figure) [2]

Table 10.2.: MS Settings of Q-Exactive™ Thermo instrument used during LC and FIA. All the parameters are imposed by Biocrates procedure

Option	Parameter	LC analysis Value	FIA analysis Value
Scan parameters	Scan type	Full MS	Full MS
	Scan Range [m/z]	100.0 to 800.0	100.0 to 1000.0
	Fragmentation	None	None
	Resolution	70000	70000
	Polarity	Positive/Negative	Positive/Negative
	Microscan	1	1
	Lock Masses	Off	Off
	Maximum Injection time, ms	250	250
HESI Source	Sheath gas flow rate	60	15
	Aux gas flow rate	30	30
	Sweep gas flow rate	1	1
	Spray voltage [kV]	3.00	2.50
	Spray Current [μA]	X	X
	Capillary Temperature [$^{\circ} C$]	300	300
	S-lens RF level	90	60
Aux gas heater temp [$^{\circ} C$]	550	120	

to elaborate the raw data obtained, the ProgenesisQI software from Waters Society was employed. On the contrary, for the sample preparation procedure, an established protocol for Global Metabolomics by LC-MS was used (the method was built up and validated by Prof. Daisuke Saigusa; for further information check [23]). So, for all the samples preparation, I followed exactly this protocol: extraction procedure, QCs sample preparation, calibration of the system, etc.

The sample preparation was automated by a robot system in order to avoid possible analyst's errors. The robot is a Microlab[®] STARlet robot system (Hamilton, Reno, NV), a Rack Runner (Hamilton) cooling water circulation apparatus, a Cool mini (GL Science, Tokyo, Japan), an ultrasonic bath (Kyowa Irika, Kanagawa, Japan), and an auto-controlled centrifuge (Hettich, Tuttlingen, Germany). The automated steps in sample preparation are as follows:

1. Transfer 50 μL of each plasma sample in a 96-well plate;
2. Add to each sample 150 μL of methanol containing 0.1% formic acid;
3. Vortex a mixing for 5 minutes at room temperature;
4. Homogenization in an ultrasonic bath for 5 min, room temperature;
5. Centrifuge the samples: 6440 g, 20 min, at 4°C;
6. 100 μL of the supernatant was transferred to Sirocco[®] Protein Precipitation;
7. 3 times washing step with 100 μL methanol containing 0.1% formic acid;
8. Centrifuge the supernatants: 3000 g, 5 min, at 4°C;
9. 100 μL of water containing 0.1% formic acid was added to each study sample.

For the quality control samples (fundamental in an untargeted metabolomics analysis in order to uniform the data) the sample preparation provides 30 μL of each study sample after the automated sample extraction

Table 10.3.: *Performance characteristics of the Q-ExactiveTM Thermo instrument (under defined conditions)*

Performance Characteristics	
Resolving power	Up to 140,000 @ m/z 200
Mass range	50 to 6,000 m/z
Scan rate*	Up to 12 Hz at resolution setting of 17,500 @ m/z 200
Mass accuracy *	Internal: <1 ppm RMS
	External: <3 ppm RMS
Sensitivity	Full MS: 500 fg Buspirone on column S/N 100:1
	SIM: 50 fg Buspirone on column S/N 100:1
Polarity switching	One full cycle in <1 sec (one full scan positive mode and one full scan negative mode at resolution setting of 35,000)
Multiplexity	Up to 10 precursors/scan
Analog inputs	One (1) analog input (0 – 1 V) One (1) analog (0 – 10 V)

are mixed together. The resulting mixture obtained is transferred into a well for the study quality control (SQC). From this quality control, the progressive diluted quality controls (dQC) are generated by continuous dilution with 50% methanol (water/methanol = 50/50, v/v %) containing 0.1% formic acid. For the d2QC (QC diluted twice), a 2-fold dilution is operated; then, 4-fold for d4QC, 8-fold for d8QC, and 16-fold for 16QC.

Coming back to mass spectrometry method, a Thermo Scientific Q ExactiveTM hybrid quadrupole-Orbitrap mass spectrometer was used. It combines the possibility of high-performance quadrupole precursor selection with the high-resolution Orbitrap detection. Thanks to that, this instrument could be used for a wide range of qualitative and quantitative applications. Its high scan speed and spectral multiplexing capabilities make it fully compatible with UHPLC and fast chromatography techniques. Some of its most important parameters used for untargeted metabolomics analysis are resumed in table 10.3.

Two different experiments took place: as a matter of fact, HPLC analyses were done as in reverse phase (C18 columns) as in direct phase (HILIC columns). Furthermore, each experiment was conducted as in positive as in negative polarity. So, four distinguished analyses were conducted. For all the untargeted metabolomics (C18 and HILIC analyses) an Acquity HSS T3

(150 mm \times 2.1 mm i.d., 1.8 μ m - Waters) analytical column was used. This aspect could result strange, but as reported in [10]: “[...] the Acquity HSS T3 is a universal, silica-based bonded phase compatible with 100% aqueous mobile phase and should be your first choice when developing separations for polar and non-polar compounds.”

The mobile phases were 0.01% FA in H₂O(solvent A) and 0.01% FA in ACN (solvent B). The oven temperature was set at 40° C. The reverse phase gradient was: 1% B (0.0-1.0 min), 1-99% B (1.0-8.0 min), 85% B (8.0-13.0 min), 1% B (13.0 – 15.0 min), injection volume: 2 μ L. On the contrary, the HILIC gradient was: 100% B (0.0-1.0 min), 100-1% B (1.0-8.0 min), 15% B (8.0-13.0 min), 100% B (13.0 – 15.0 min), injection volume: 2 μ L. The MS parameters used were the same reported in 10.2, “LC analysis” column.

After having reported the HPLC-HRMS parameters, we focus the attention upon the statistical software used to elaborate the raw data, Progenesis QI. Progenesis QI is small molecule discovery analysis software created by Waters company for LC-MS data. The most fundamental steps within the Progenesis QI analysis (not strictly applicable only to this analysis but generally valid) are reported below in the next list.

1. Import data: Progenesis QI supports data formats produced by all the LC-MS machines systems commonly used for untargeted compound analysis: Waters, Thermo, Agilent, Sciex, and Bruker instruments are the most common ones. Furthermore, it also supports cross-vendor file-formats, including mzXML, mzML, and NetCDF. Data can be in profile or centroid form, be high resolution or low resolution, and have positive, negative, or mixed ionization polarity.
2. Select your possible adducts presented in the samples.
3. Ion intensity maps: After being imported in the software, each run in the experiment is shown as an ion intensity map, which is representative of the sample’s MS signal by m/z and retention time.
4. Run alignment: The software can combine data from multiple MS runs. This is a required task for comparative abundance profiling studies. It enables the comparison of different experimental conditions using a high number of replicates. To combine and compare results

from different runs, Progenesis QI aligns them with compensating for between-run variation in the chromatography.

5. Peak picking: To ensure consistent peak picking and matching across all data files, an aggregate data set is created from the aligned runs. It contains all peak information from all sample files, allowing the detection of a single map of compound ions. This map is then applied to each sample, giving 100% matching of peaks with no missing values. The ultimate purpose is to generate reliable results using accurate multivariate statistical analysis.
6. Compound results: This is the step where the results of quantification and identification are automatically brought together. All the compound ions are automatically deconvoluted to provide accurate quantitation of each compound. After detection, the ion abundance measurements are normalized, so we can make comparisons between the runs and find compounds of biological interest. Ion intensity maps, 3D views, mass spectra, and chromatograms are displayed for each compound ion on all runs to provide quality assurance of peak picking and alignment.
7. Identify Compounds: Once founded a list of detected compound ions that need identification, MetaScope allows us to search for compound identifications based not only on neutral mass and retention time but also MS/MS fragmentation data.
8. Multivariate statistical analysis and integrate results with another bio-informatics.

The untargeted lipidomics analysis: UHPLC-HRMS

The sample preparation procedure described in the previous section was not focused on lipid extraction. So, a new extraction procedure and a new instrumental method were applied. The final purpose is to obtain, by untargeted lipidomics analysis, information upon lipids abundance and presence in the mice samples. The methods used for extraction and quantification

of the different lipid classes is a slightly modified method based on the one tested and validated by Prof. Saigusa in a previous article [24].

Briefly, as soon as the selected organ is collected from the mice, it is placed into a sample tub and then frozen using liquid nitrogen. For each analysis, a portion of the organ is sectioned and placed in a test tube (2.0 mL). The sectioning procedure was conducted upon refrigerated metallic support (previously disinfected with EtOH 70% in H₂O), and, for cutting, a new and disinfected cutter was employed. The figure 10.4 shows this step of extraction procedure. The metallic plate is placed in an icebox to keep it continuously refrigerated.

A variable amount of acidified solvent (0.1% formic acid - volume fraction - in methanol) was added to the sample as extraction solvent. Indeed, the volume added in each sample corresponds to approximately ten times the weight of the cut section obtained. For example, 200 μ L of the aforementioned solvent was added for in sample having a weight of 20.097762 g (the Mettler Toledo xp56 analytical balance provided with six decimal places has been used).

After the addition of the formic acid, the mixture was homogenized for 30 s using a Micro Smash MS-100R (TOMY) (3,000 \times rpm at 4° C). Subsequently, centrifugation at 16,400 \times g for 20 min at 4° C and a deproteinization operated by a Sirocco instrument - Waters occurred.

Finally, further centrifugation using the same conditions happened. The supernatant obtained is now ready for UHPLC-HRMS analyses. The instrument settings for HPLC consist of a NANOSPACE SI-2 (Shiseido) as LC device coupled with a L-Column2 ODS (2.0 x 100 mm. 2 μ m; CERI) as analytical column. The aqueous mobile phase was 0.1% formic acid in H₂O-ACN - 1.0 M ammonium formate in H₂O= 60:40:1 and the organic solvent was 0.1% formic acid in 2-propanol - ACN = 90:10. The oven temperature was set at 45° C, flow rate 200 μ L/min, injection volume was 2 μ L and mobile phase gradient applied to obtain chromatographical separation was set as follow: 30% B (0.0-2.0 min), 30-100% B (2.0-20.0 min), 100% B (20.0-30.0 min), 30% B (30.0 – 35.0 min). Wash liquid and wash liquid port were H₂O and 90% ACN/H₂O respectively.

Moving to the MS section, the MS device used was a Q ExactiveTM Hybrid



Figure 10.4.: *A detail of the sample extraction procedure: sampling of a portion of the liver's mouse*

Quadrupole-Orbitrap Mass Spectrometer equipped with a HESI ionization source and operating as in positive as in negative polarity. The scan range was set in the range 350.0 - 1050.0 m/z , resolution was 70000, and source voltage was set at 3500 V for positive polarity analyses, -2500 V for negative ones. The capillary temperature was set at 275° C, and auxiliary and sheath gas pressures were maintained at 45 psi and 10 psi respectively during all the analyses.

The untargeted metabolomics analysis: Ambient mass spectrometry techniques

The analysis of the chemical composition of the inks on a paper, to locate and to quantify ceramides molecules in a tissue sample, monitor the adulteration of food: these are just a few examples of possible applications of ambient mass spectrometry in researches or daily life aspects. Generally, the condition sine qua non to analyse samples by mass spectrometry is that the analytes need to be ionized and present in a gas phase. If usually the ionization process requires high pressure, in ambient mass spectrometry the molecules are ionized under the regular pressure in an open air space.

Furthermore, in ambient mass spectrometry applications, minimal (or no) treatment of the sample before analysis is required [4, 9].

Nowadays, more than thirty different techniques [6, 28] concerning desorption (operated by laser or plasma) and ionization (like electrospray or chemical ionization) methods are available [11]. Amongst all these possibilities, desorption electrospray ionization (DESI) and direct analysis in real-time (DART) are two most widely adopted for analysis of analytes whitening biological samples like organs or tissues by metabolomics untargeted MS approaches [7, 8, 25].

For the target we set, different organs belonging from space mice were analysed via the DESI-MS approach. The MS used was a quadrupole time of flight MS (QTOF/MS) DESI – SYNAPT G2-Si High Definition Mass Spectrometry by Waters. The samples cannot be analysed untreated, but cryosections must be prepared. The instrument used was a cryostat, which is essentially a microtome operating in a cold ambient. The sections obtained by microtome has slicing sections between 10 and 20 μm . Figures 10.5 and 10.6 show the microtome and the cryosections obtained respectively. On the contrary, the QTOF instrument and the DESI source are shown in figures 10.7 and 10.8.

Unfortunately, during my experience at the ToMMO Laboratory, I was applied only to the initial phases of the development and method's improvement (sample preparation, different analysis tests upon QTOF platform, etc.). At present (September 2019), the DESI-MS method is developed, but the results obtained are still being processed. The instrumental conditions are therefore not reported. However, recently by the ToMMO research group, an article has been published [17]. In this article, a metabolic profiling of tumor and nontumor regions in human livers was performed, exploiting more or less the same DESI-MS conditions applied in the SpaceMice project. So, if necessary, used these conditions as long as the SpaceMice ambient mass spectrometry's article will be published.

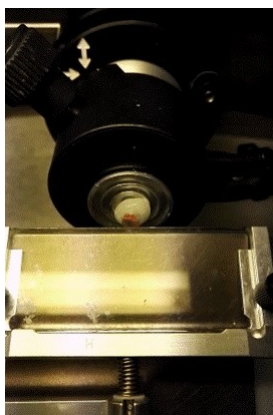


Figure 10.5.: *Cryostat: detail about the microtome portion*



Figure 10.6.: *Cryostat tissue sections by placed onto a glass ready for DESI-MS*



Figure 10.7.: *DESI – SYNAPT G2-Si High Definition Mass Spectrometry by Waters*

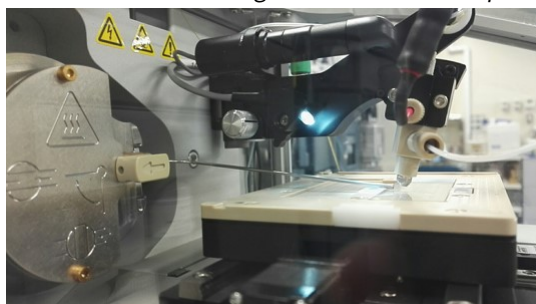


Figure 10.8.: *DESI ionization source used for SpaceMice ambient mass spectrometry*

10.4. Results and discussion

Semi-targeted analysis: Biocrates kit

The Biocrates AbsoluteIDQ[®] p400 HR Kit plans to perform as FIA as LC analyses. However, as regards the LC results, when I left ToMMO group the LC data analysis was still in progress. Therefore, only the FIA results obtained are reported.

Thanks to FIA analysis, 312 analytes of the theoretically expected 408s have been identified and quantified. The molecules detected belong from different families: aminoacids, acylcarnitines, triglycerides, phosphatidylcholines, and ceramides. In this study, PCA was used to extract relevant information from FIA samples analysed. The score plot reported in figure 10.9 is the projection of the FIA data onto the PCs. The Score Plot is a scattered plot that involves the projection of the data onto the PCs in two dimensions. The first PC is the one with the minimal total distance between the data and their projection onto the PC. In this way, the variance of the projected points assumes the maximum value. The second (and so on) PCs are selected in an analogue way, with an additional requirement: each PC must be uncorrelated with all the previous ones. Thus, the PCs are selected to provide a new space of uncorrelated 'variables' which best carry the variation in the original data and in which to represent the original samples more succinctly.

In 10.9 different groups are represented:

- Keap1KD: Plasma sample from mice with a depletion of the Keap1 gene. This gene has been shown to interact with Nrf2. Under quiescent conditions, Nrf2 is anchored in the cytoplasm through binding to Keap1, which, in turn, facilitates the ubiquitination and subsequent proteolysis of Nrf2. Such sequestration and further degradation of Nrf2 in the cytoplasm are mechanisms for the repressive effects of Keap1 on Nrf2. Because Nrf2 activation leads to a coordinated antioxidant and anti-inflammatory response, and Keap1 represses Nrf2 activation, Keap1 has become a desirable drug target.
- Nrf2KO: plasma sample of Nrf2KO mice sent in Space and collected

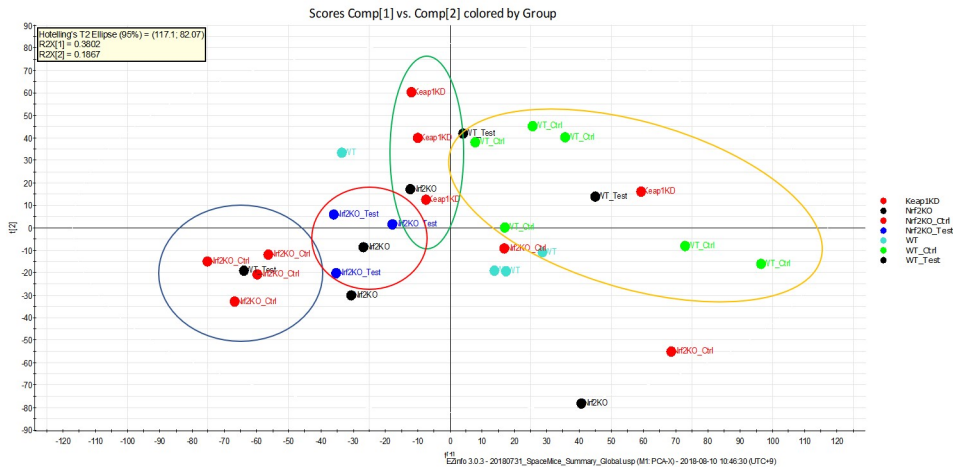


Figure 10.9.: *Score plot of FIA Biocrates data*

after they landed on the Earth.

- Nrf2KO-Ctrl: plasma sample of Nrf2KO mice sent in Space and collected before they have been sent in the Space.
- Nrf2KO-Test: plasma sample of Nrf2KO mice never sent in Space. They lived on the Earth in the same conditions of food, night and daylight, etc. as space on the ISS Station.
- WT: plasma sample of WT mice sent in Space and collected after they landed on the Earth.
- WT-Ctrl: plasma sample of WT mice sending in Space and collected before they have been sent in the Space.
- WT-Test: plasma sample of WT mice never sent in Space. They lived on the Earth in the same conditions of food – night and daylight etc. as space on the ISS Station.

In figure 10.9 a good separation between all the groups is already present. That's means that some of the analytes detected by the FIA method can discriminate between mice send in space (oxidative stress subjects), and the ones seed on the Earth (low levels of oxidative stress). Subsequently, to check if some of the analytes detected during FIA analyses could have a

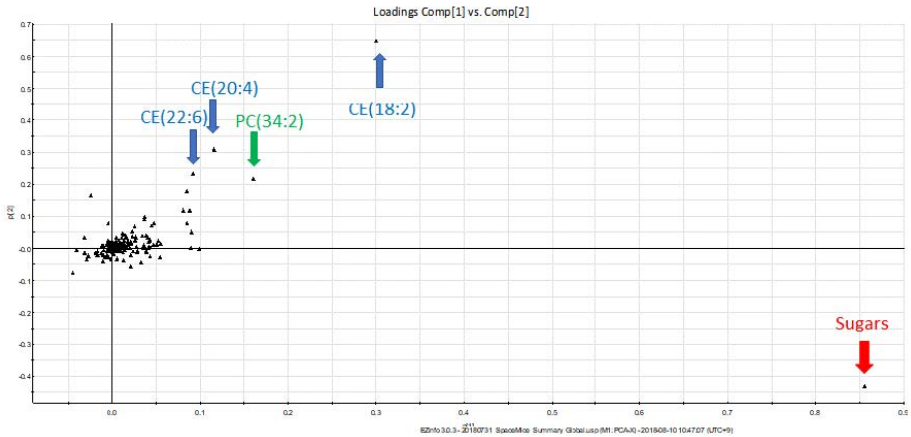


Figure 10.10.: Loading plot of FIA Biocrates data

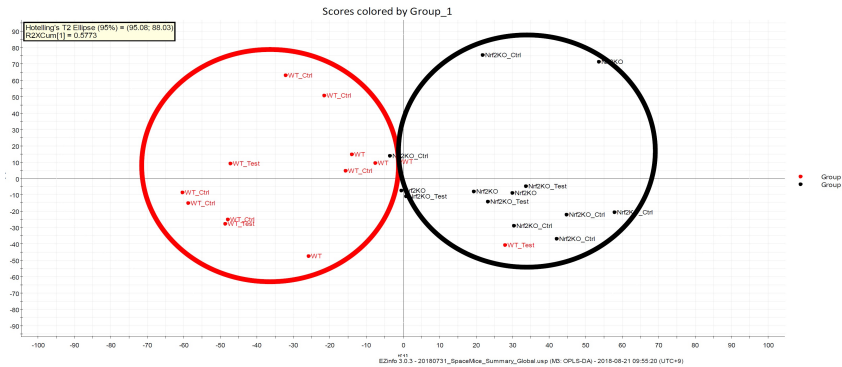


Figure 10.11.: Score plot of FIA Biocrates data without sugars' group

positive or negative influence on the results, a loading plot (figure 10.10) was constructed. From this loading plot is clear how the sugars variable is an out-layer if compared to the other variables. So, it must be eliminated and not taken into consideration. If, after the elimination of this out-layer family, we reconstruct a new score plot the result (see figure 10.11) shows a much more clear separation between WT and Nrf2KO groups.

By looking at loadings plot 10.10, only some molecules belonging to specific groups contribute to discriminate WT and Nrf2KO mice groups. As a matter of fact, lower is the distance between the centre of the axis, and the variable position lower is the variable's important in the discrimination process. So, after monitor which are the most important analytes able to discriminate WT mice correctly from Nrf2KO mice, we finally demon-

strated that the best discrimination results rising using only molecules belonging only from these four classes: phosphatidylcholines (PC, 142 analytes), aminoacids (AA, 32 analytes), sphingomyelins (SM, 26 analytes), and cholesteryl esters (CE, 13 analytes). In this way, the total number of molecules taken into account goes to 213 analytes. The histogram in figure 10.12 summarizes the number of discriminating analytes case-by-case. For example, observing the leftmost histogram, it is evident that molecules that have the greatest effect on discrimination between the Nrf2KO Ctrl and WT Ctrl groups are those belonging to the PC family, followed by the CE family and so on. In general, we can say that PCs are the ones that most discriminate among the groups considered. It is important to emphasize that, within an analytes' group, not all molecules are always present in all the samples. In fact, in the leftmost histogram, only 29 of the 142 possible molecules are present and useful for the discrimination process.

Furthermore, the unsaturated compounds like CE (18:2), CE (20:4), or PC (34:2) have always a lower concentration in Nrf2KO samples compared to WT samples. This is a new aspect with no evidence in the literature and could be in the next future the main subject for an original and interesting article. Moreover, specific compounds show a significant difference and they need to be investigated in detail to understand the reason for the difference between groups. The compounds are 1-palmitoyl-2-oleoyl-SN-glycero-3-phosphocholine, cholesteryl oleic acid ester, glycine, histidine, and histamine molecule.

The untargeted metabolomics analysis: UHPLC-HRMS

Using the protocol and the settings reported in section 10.3 we analysed the homogenized liver samples of mice. Upon these samples, we conducted four different UHPLC-HRMS analyses in order to investigate all the spectrum of possible metabolites: reverse phase chromatography in positive and negative mode and HILIC chromatography in positive and negative mode. The results here reported will not mention about the reverse phase chromatography in the negative mode because the samples are still under analysis.

As an example, the TIC chromatogram and three different XIC chroma-

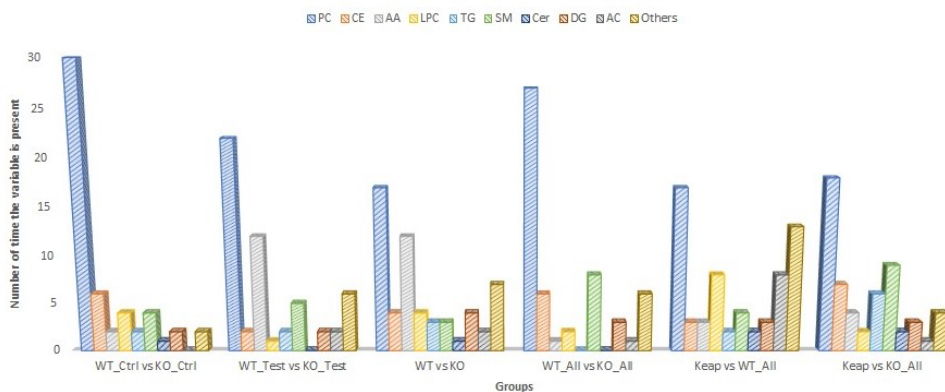


Figure 10.12.: *Histogram summarizing the variables' number applied to discriminate the different mice groups under examination*

tograms (related to their respectively MS spectrum) are reported in figure 10.13. The peaks extracted for the different m/z values are well shaped, narrow, and have a proper intensity.

This quantity of data such as that obtained during these experiments cannot be processed manually. As a matter of fact, in untargeted metabolomics analyses the use of software like Progenesis QI is essential to obtain data. Indeed, this software can, for example, combine data from different chromatograms and find clusters, trends or, on the contrary, demonstrate that the samples are completely independent of each other. Focusing on Progenesis, by applying the bullet point list previously reported in 10.3, it's possible to extrapolate interesting data. For example, the score plots of data belonging to C18 positive ion mode and HILIC negative ion mode are reported in figures 10.14 and 10.15 respectively. In both plots, as happened for the score plot obtained by Biocrates FIA data, different variables detected can be clearly divided the samples into two distinct groups: WT group and Nrf2KO group. Furthermore, in order to select the variables that best discriminate between the two groups under examination, the relative S-plots are generated (figures 10.16 and 10.17). In an S-Plot the horizontal component will capture variation between the groups and the vertical dimension will show the variation that exists within the groups considered. So, these plots are a visualization of the covariance between the metabolites and the class of designation.

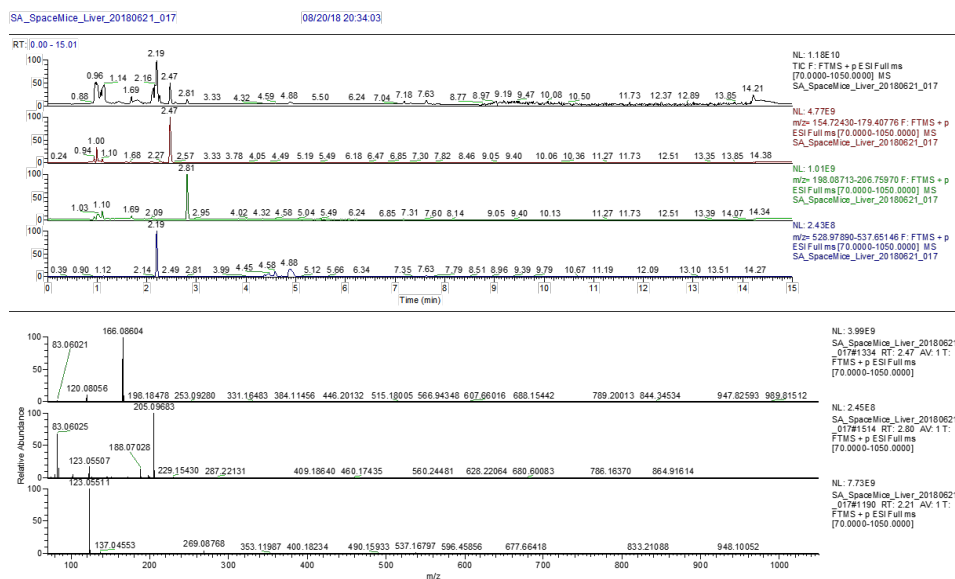


Figure 10.13.: XIC and MS spectra of different analytes detected with the UHPLC-HRMS method

Thanks to the interpretation of these plots it's possible to detect new molecules that can be used as markers able to divide WT by KO samples. Furthermore, the most critical aspect to underline is that these new molecules are different from those found using Biocrates kit. The most crucial molecules detected through these analyses are resumed in table 10.4. However, it's fundamental to remark that the molecules' plausibility detected by the software have been checked by using online database (such as HMDB, METLIN, or Chempidder) and by critical and scientific sense: for example, in HILIC mode if the peak related to a non-polar analyte rises in the re-condition time of the chromatographic run, I will reject the match also if Progenesis QI gives me back a score match of 99% because I know that the data is not scientifically true.

The untargeted lipidomics analysis: UHPLC-HRMS

The lipidomics approach generated interesting results. As a matter of fact, some of the molecules detected through this method have also been detected and quantified by the Biocrates approach. Therefore, this is a double demonstration of the validity of the untargeted method we used and

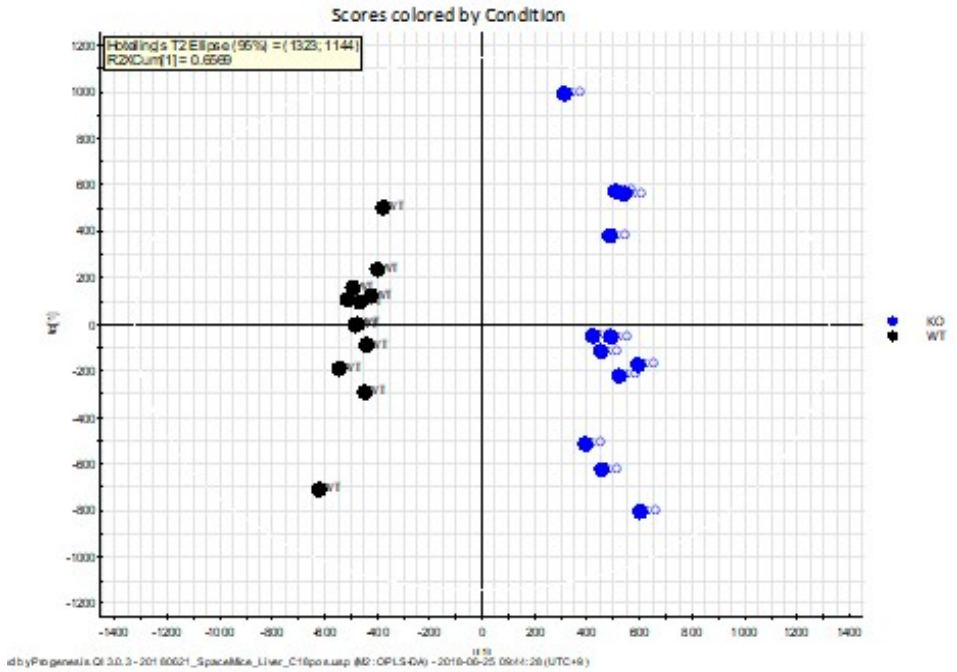


Figure 10.14.: Score plot WT vs Nrf2KO in real samples obtained by UHPLC-HRMS data - reverse phase chromatography in positive ion mode. In blue are marked the WT samples, in black the NRF2KO samples

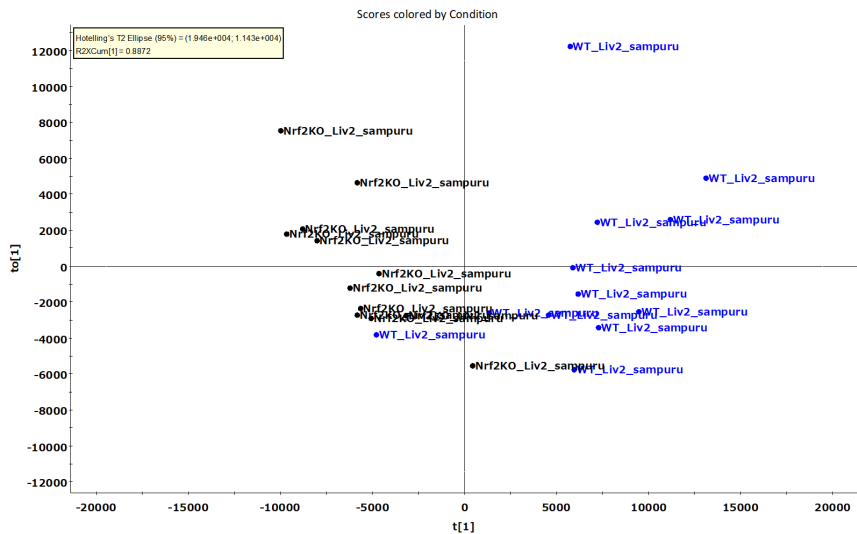


Figure 10.15.: Score plot WT vs Nrf2KO in real samples obtained by UHPLC-HRMS data - direct phase chromatography in negative ion mode. In blue are marked the WT samples, in black the NRF2KO samples

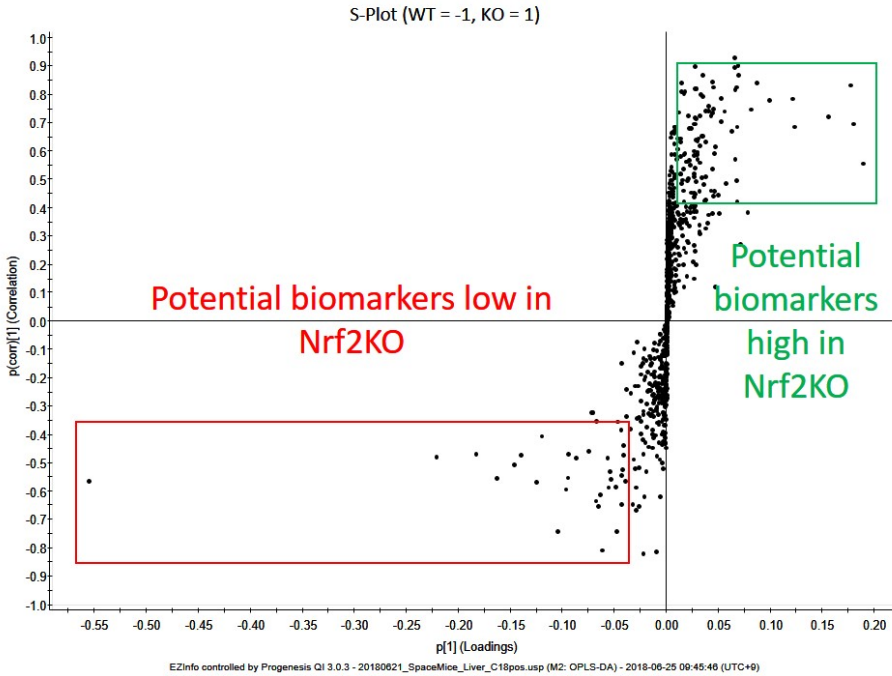


Figure 10.16.: *S*-plot WT vs Nrf2KO of real samples, reverse phase chromatography in positive ion mode

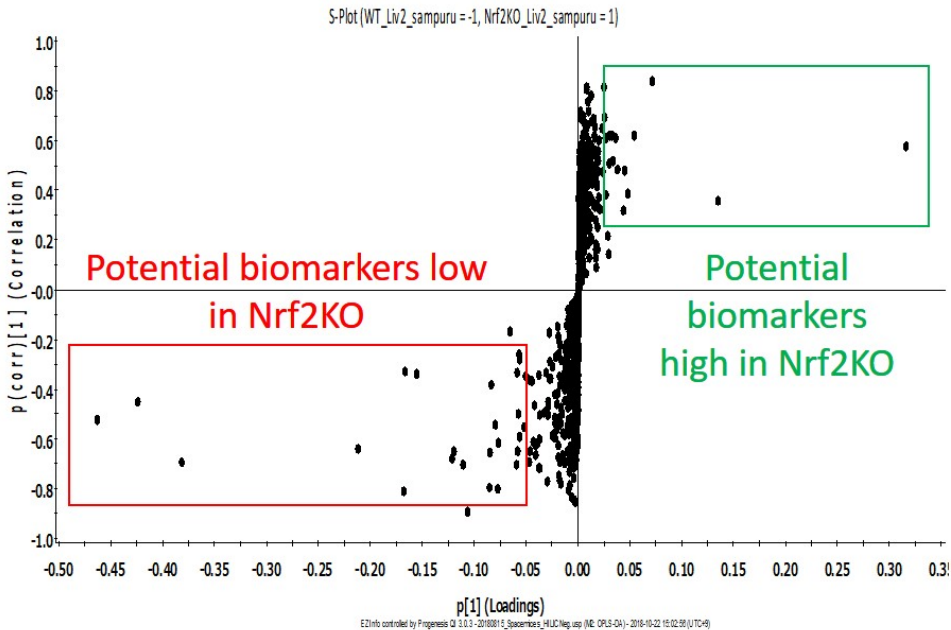


Figure 10.17.: *S*-plot WT vs Nrf2KO of real samples, direct phase chromatography in negative ion mode

Table 10.4.: *Schematic summary of the different molecules detected using the different instrumental approaches*

HILIC positive ion mode results		
Identifications	Adducts	Fragmentation Score %
2,4-Diaminobutyric acid	$[M+H-H_2O]^+$	76.0
2-octenoylglycine	$[M+Na]^+$	81.6
3-Methylcytosine	$[M+H-H_2O]^+$	61.4
Alanyl-Histidine	$[M+H-H_2O]^+$	73.3
Alanyl-Threonine	$[M+H]^+$	75.0
Sphingosine	$[M+H]^+$	82.3
Muramic acid	$[M+H-H_2O]^+$	81.7
HILIC negative ion mode results		
Identifications	Adducts	Fragmentation Score, %
Adducts	$[M-H-H_2O]^-$	52.9
Fragmentation Score %	$[M-H-H_2O]^-$	75.2
Hydroxybutyric acid	$[M-H-H_2O]^-$	53.5
PI(18:0/20:4)	$[M-H]^-$	78.3
PS(16:0/22:6)	$[M-H]^-$	66.2
C18 positive ion mode results		
Identifications	Adducts	Fragmentation Score, %
Oxoproline	$[M+H]^+$	64.7
17a-hydroxypregnenolone	$[M+H-H_2O]^+$	75.9
19-Hydroxydeoxycorticosterone	$[M+H-H_2O]^+$	66.9
2,4-Dimethylperidine	$[M+H]^+$	49.7
4-butyl-5-ethylthiazole	$[M+NH_4]^+$	31.5
Hydroxyprogesterone	$[M+H]^+$	75.5
Tetrahydrocortisone	$[M+H]^+$	58.5

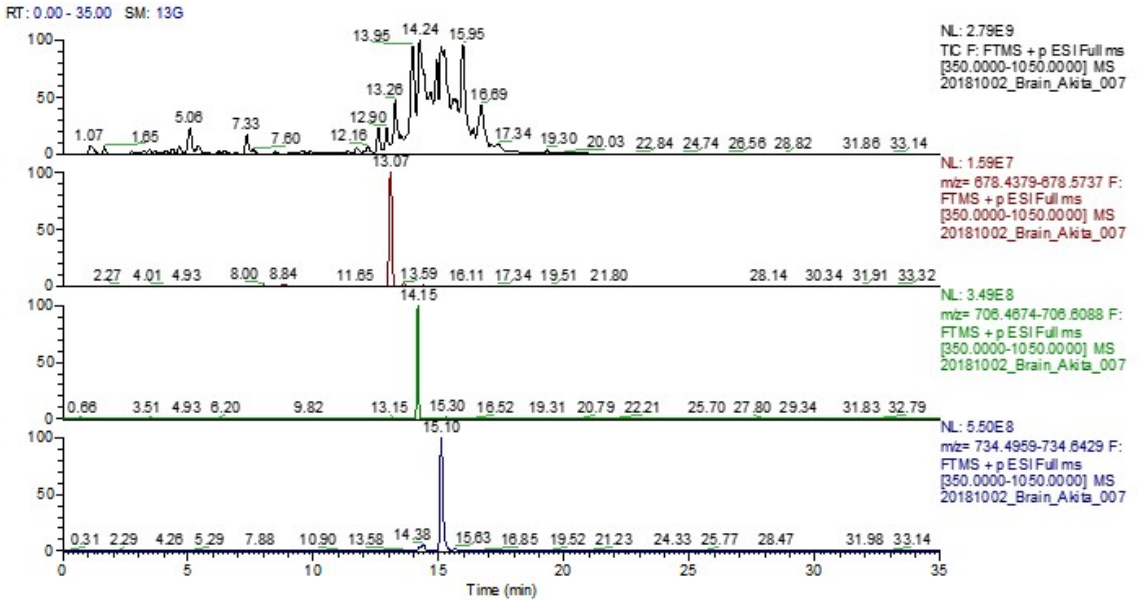


Figure 10.18.: *TIC and XIC chromatograms obtained through the analytical lipidomics method built up. The three extracted m/z values are, from the top to the bottom, 678.5058 (PC(28:0)), 706.5353(PC(30:0)), and 734.5638 (PC(32:0))*

a further confirmation of the validity of the Biocrates Kit. As example, the TIC chromatogram and three different XIC chromatograms of lipid molecules detected through the analytical method in positive ion mode are reported in figure 10.18. The m/z peaks extracted are well shaped, narrow, and show a good intensity. Furthermore, from the TIC we can notice that the richest portion in signals of the TIC chromatogram is between the tenth minute (61% B) and the twentieth minute of analysis (100%). This is in line with what is reported in the literature [5, 15, 19]: in fact, molecules such as TG or PC tend to be eluted, operating with conditions similar to ours, with organic percentages over 50%.

In figure 10.19 are reported as example four score plots (and the related S-plots) concerning the comparison (performed in two) of the various groups of samples analysed through negative polarity mode. The groups considered are Nrf2KO mice send in space (called Nrf2KO-Space), Nrf2KO mice feed at Tohoku University (Nrf2KO-Earth), mice with no depletion of the Nrf2 gene send in space (WT-Space), and wild type mice grew at Tohoku University

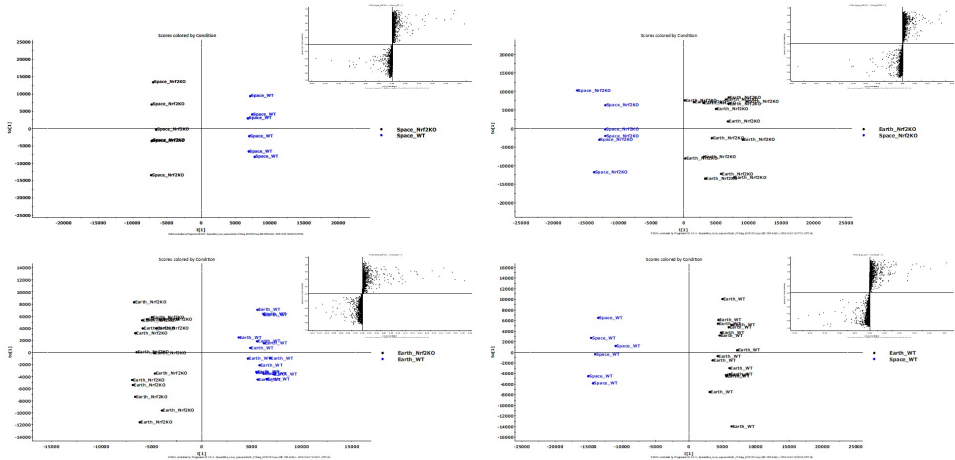


Figure 10.19.: *Score plots and related S-plots generated through the results obtained in negative ion mode. For all of them, the distinction between the two groups under consideration is always clear*

(WT-Earth). We can see, starting from the graph in the upper left (figure 10.19) and then proceeding clockwise: Space NRF2KO vs. Space WT, Space NRF2KO vs. Earth NRF2KO, Earth NRF2KO vs. Earth WT, and Space WT vs. Earth WT. The interesting result is that independently of the combination looked at, the data analysis allows always to discriminate the two groups under examination.

From the S-plots generated, it is possible to understand which are the metabolites able to discriminate the best between the different groups. The results of this analysis, conducted as on positive polarity as on the negative ones, are summarized in tables 10.5 and 10.6 respectively. Thanks to both methods, more than thirty different lipid molecules could be detected.

Table 10.5.: Summary of the different analytes detected by untargeted lipidomics approach - Positive polarity

Compound	m/z	Charge	Retention time (min)	Identifications	Adduct	Formula	Frag.mass,%	Mass error (ppm)
1	468.3076	1	3.02	PC(14:0/0:0)	$[M + H]^+$	C ₂₂ H ₄₆ NO ₇ P	85.18	9
2	496.3391	1	4.81	PC(16:0/0:0)	$[M + H]^+$	C ₂₄ H ₅₀ NO ₇ P	91.01	7
3	494.3235	1	3.28	PC(16:1/0:0)	$[M + H]^+$	C ₂₄ H ₄₈ NO ₇ P	76.53	6
4	482.3599	1	5.7	PC(O-16:0/0:0)	$[M + H]^+$	C ₂₄ H ₅₂ NO ₆ P	57.65	6
5	524.3702	1	7.02	PC(18:0/0:0)	$[M + H]^+$	C ₂₆ H ₅₄ NO ₇ P	88.65	9
6	522.3549	1	5.14	PC(18:1/0:0)	$[M + H]^+$	C ₂₆ H ₅₂ NO ₇ P	77.65	5
7	520.3391	1	3.69	PC(18:2/0:0)	$[M + H]^+$	C ₂₆ H ₅₀ NO ₇ P	61.81	7
8	508.3753	1	5.97	PC(O-18:1/0:0)	$[M + H]^+$	C ₂₆ H ₅₄ NO ₆ P	67.67	8
9	754.5383	1	12.92	PC(20:4/14:0)	$[M + H]^+$	C ₄₂ H ₇₆ NO ₈ P	71.91	-2
10	758.5703	1	14.22	PC(20:2/14:0)	$[M + H]^+$	C ₄₂ H ₈₀ NO ₈ P	70.72	-9
11	568.3389	1	3.25	PC(22:6/0:0)	$[M + H]^+$	C ₃₀ H ₅₀ NO ₇ P	83.90	9
12	542.3234	1	2.71	PC(20:5/0:0)	$[M + H]^+$	C ₂₈ H ₄₈ NO ₇ P	64.19	7
13	544.339	1	3.54	PC(20:4/0:0)	$[M + H]^+$	C ₂₈ H ₅₀ NO ₇ P	80.91	8
14	546.3546	1	4.25	PC(20:3/0:0)	$[M + H]^+$	C ₂₈ H ₅₂ NO ₇ P	66.01	-2
15	760.5848	1	14.97	PC(16:0/18:1)	$[M + H]^+$	C ₄₂ H ₈₂ NO ₈ P	82.63	3
16	780.554	1	13.34	PC(22:5/14:0)	$[M + H]^+$	C ₄₄ H ₇₈ NO ₈ P	77.43	-2
17	784.5857	1	14.46	PC(16:0/20:3)	$[M + H]^+$	C ₄₄ H ₈₂ NO ₈ P	65.12	-10
18	786.6013	1	15.15	PC(22:2/14:0)	$[M + H]^+$	C ₄₄ H ₈₄ NO ₈ P	50.91	-17
19	788.6168	1	15.8	PC(22:1/14:0)	$[M + H]^+$	C ₄₄ H ₈₆ NO ₈ P	61.71	-6
20	810.6007	1	14.97	PC(22:4/16:0)	$[M + H]^+$	C ₄₆ H ₈₄ NO ₈ P	85.93	0
21	812.6162	1	15.34	PC(22:2/16:1)	$[M + H]^+$	C ₄₆ H ₈₆ NO ₈ P	86.39	-9
22	832.5856	1	13.79	PC(18:1/22:6)	$[M + H]^+$	C ₄₈ H ₈₂ NO ₈ P	68.54	-5
23	834.6011	1	14.72	PC(18:0/22:6)	$[M + H]^+$	C ₄₈ H ₈₄ NO ₈ P	59.01	-3
24	836.6165	1	15.29	PC(18:0/22:5)	$[M + H]^+$	C ₄₈ H ₈₆ NO ₈ P	63.04	-11
25	502.2923	1	3.69	PE(20:4/0:0)	$[M + H]^+$	C ₂₅ H ₄₄ NO ₇ P	80.00	5
26	742.5745	1	14.69	PC(O-16:0/18:3)	$[M + H]^+$	C ₄₂ H ₈₀ NO ₇ P	93.02	0
27	744.5898	1	14.84	PC(O-16:0/18:2)	$[M + H]^+$	C ₄₂ H ₈₂ NO ₇ P	67.64	4
28	766.5748	1	14.47	PC(O-16:1/20:4)	$[M + H]^+$	C ₄₄ H ₈₀ NO ₇ P	84.43	-3
29	796.6218	1	15.53	PC(O-16:0/22:4)	$[M + H]^+$	C ₄₆ H ₈₆ NO ₇ P	82.39	-3
30	794.606	1	14.64	PC(O-18:1/20:4)	$[M + H]^+$	C ₄₆ H ₈₄ NO ₇ P	72.98	-2
31	788.6165	1	15.8	PC(20:1/16:0)	$[M + H]^+$	C ₄₄ H ₈₆ NO ₈ P	84.31	-1
32	806.5698	1	13.75	PC(20:2/18:4)	$[M + H]^+$	C ₄₆ H ₈₀ NO ₈ P	75.06	-3
33	808.5848	1	14.08	PC(20:2/18:3)	$[M + H]^+$	C ₄₆ H ₈₂ NO ₈ P	69.55	3
34	678.5061	1	12.87	PC(14:0/14:0)	$[M + H]^+$	C ₃₆ H ₇₂ NO ₈ P	60.32	8
35	706.538	1	14.00	PC(14:0/16:0)	$[M + H]^+$	C ₃₈ H ₇₆ NO ₈ P	65.66	2
36	734.5693	1	14.96	PC(14:0/18:0)	$[M + H]^+$	C ₄₀ H ₈₀ NO ₈ P	68.93	2
37	762.6006	1	15.84	PC(14:0/20:0)	$[M + H]^+$	C ₄₂ H ₈₄ NO ₈ P	80.34	2
38	675.5442	1	12.82	SM(d18:1/14:0)	$[M + H]^+$	C ₃₇ H ₇₅ N ₂ O ₆ P	61.43	-6
39	689.5598	1	13.47	SM(d18:1/15:0)	$[M + H]^+$	C ₃₈ H ₇₇ N ₂ O ₆ P	66.78	-6
40	703.5742	1	14.07	SM(d18:1/16:0)	$[M + H]^+$	C ₃₉ H ₇₉ N ₂ O ₆ P	77.90	7
41	731.6059	1	15.10	SM(d18:1/18:0)	$[M + H]^+$	C ₄₁ H ₈₃ N ₂ O ₆ P	72.05	3
42	759.6368	1	15.97	SM(d18:1/20:0)	$[M + H]^+$	C ₄₃ H ₈₇ N ₂ O ₆ P	74.44	7
43	787.6678	1	16.73	SM(d18:1/22:0)	$[M + H]^+$	C ₄₅ H ₉₁ N ₂ O ₆ P	86.54	10
44	701.5596	1	12.97	SM(d18:1/16:1)	$[M + H]^+$	C ₃₉ H ₇₇ N ₂ O ₆ P	69.01	-4
45	729.5903	1	14.19	SM(d18:1/18:1)	$[M + H]^+$	C ₄₁ H ₈₁ N ₂ O ₆ P	84.83	2
46	757.6216	1	15.19	SM(d18:1/20:1)	$[M + H]^+$	C ₄₃ H ₈₅ N ₂ O ₆ P	79.79	2
47	785.6524	1	16.03	SM(d18:1/22:1)	$[M + H]^+$	C ₄₅ H ₈₉ N ₂ O ₆ P	85.59	7
48	813.6843	1	16.61	SM(d18:1/24:1)	$[M + H]^+$	C ₄₇ H ₉₃ N ₂ O ₆ P	88.20	1
49	841.7157	1	17.39	SM(d18:1/26:1)	$[M + H]^+$	C ₄₉ H ₉₇ N ₂ O ₆ P	91.99	0
50	675.5434	1	12.87	SM(d16:1/16:0)	$[M + H]^+$	C ₃₇ H ₇₅ N ₂ O ₆ P	83.33	2
51	729.5902	1	14.19	SM(d16:1/20:1)	$[M + H]^+$	C ₄₁ H ₈₁ N ₂ O ₆ P	89.46	3

Table 10.6.: *Summary of the different analytes detected by untargeted lipidomics approach - Negative polarity*

Compound	m/z	Charge	Retention time (min)	Identifications	Adduct	Formula	Frag.mass,%	Mass error (ppm)
1	832.5984	1	15.73	PS(21:0/18:0)	$[M - H]^-$	C ₁₅ H ₃₈ NO ₁₀ P	68.8	-10
2	850.5517	1	13.64	PC(22:6/20:5)	$[M - H]^-$	C ₅₀ H ₇₈ NO ₈ P	59	14
3	738.4995	1	14.26	PE(16:0/20:4)	$[M - H]^-$	C ₄₁ H ₇₄ NO ₈ P	73.5	-11
4	804.5673	1	14.85	PS(18:0/19:0)	$[M - H]^-$	C ₄₃ H ₈₁ NO ₁₀ P	50.4	-7
5	764.5145	1	14.30	PE(20:4/18:1)	$[M - H]^-$	C ₄₃ H ₇₆ NO ₈ P	67.8	-5
6	752.5149	1	14.72	PE(17:0/20:4)	$[M - H]^-$	C ₄₂ H ₇₆ NO ₈ P	79	-13
7	846.6148	1	16.24	PS(18:0/22:0)	$[M - H]^-$	C ₄₆ H ₉₀ NO ₁₀ P	66.8	-6
8	865.4933	1	11.45	PG(22:6/22:6)	$[M - H]^-$	C ₅₀ H ₇₅ O ₁₀ P	85.1	-7
9	745.4947	1	14.03	PG(18:2/16:0)	$[M - H]^-$	C ₄₀ H ₇₅ O ₁₀ P	76.1	-7
10	478.2892	1	5.30	LysoPE(18:1)	$[M - H]^-$	C ₂₃ H ₄₆ NO ₇ P	78.7	-3
11	790.5312	1	14.92	PE(22:6/18:)	$[M - H]^-$	C ₄₅ H ₇₈ NO ₈ P	59.6	-15
12	747.5100	1	14.70	PG(18:1/16:0)	$[M - H]^-$	C ₄₀ H ₇₇ O ₁₀ P	67	-9
13	885.5408	1	14.53	PI(18:0/20:4)	$[M - H]^-$	C ₄₇ H ₈₃ O ₁₃ P	67.6	-10
14	843.5094	1	12.30	PG(22:6/20:3)	$[M - H]^-$	C ₄₈ H ₇₇ O ₁₀ P	77.4	-10
15	773.525	1	14.95	PG(18:2/18:0)	$[M - H]^-$	C ₄₂ H ₇₉ O ₁₀ P	69.2	-6
16	452.2738	1	4.97	LysoPE(16:0)	$[M - H]^-$	C ₂₁ H ₄₄ NO ₇ P	72.7	-9
17	818.5836	1	15.41	PS(18:0/20:0)	$[M - H]^-$	C ₄₄ H ₈₆ NO ₁₀ P	63.4	-9
18	819.509	1	12.89	PG(18:1/22:6)	$[M - H]^-$	C ₄₆ H ₇₇ O ₁₀ P	81.6	-11
19	748.4838	1	13.44	PE(22:6/15:0)	$[M - H]^-$	C ₄₂ H ₇₂ NO ₈ P	82.6	-11
20	714.5001	1	14.45	PE(18:2/16:0)	$[M - H]^-$	C ₃₉ H ₇₄ NO ₈ P	78.6	-10
21	524.2729	1	3.16	LysoPE(22:6)	$[M - H]^-$	C ₂₇ H ₄₄ NO ₇ P	89.4	-4
22	831.4935	1	12.72	PI(18:2/16:1)	$[M - H]^-$	C ₄₃ H ₇₇ O ₁₃ P	77.2	-10
23	771.5091	1	13.33	PG(18:2/18:1)	$[M - H]^-$	C ₄₂ H ₇₇ O ₁₀ P	83.6	-8
24	476.2736	1	3.85	LysoPE(18:2)	$[M - H]^-$	C ₂₃ H ₄₄ NO ₇ P	79.9	-5
25	716.5159	1	15.16	PE(16:0/18:1)	$[M - H]^-$	C ₃₉ H ₇₆ NO ₈ P	76.1	-15
26	500.2734	1	3.63	LysoPE(20:4)	$[M - H]^-$	C ₂₅ H ₄₄ NO ₇ P	85.6	-9
27	769.4936	1	12.48	PG(18:1/18:3)	$[M - H]^-$	C ₄₂ H ₇₅ O ₁₀ P	84.4	-11
28	452.2740	1	4.56	LysoPE(16:0)	$[M - H]^-$	C ₂₁ H ₄₄ NO ₇ P	84	-9
29	885.4804	1	11.97	PG(18:2/22:6)	$[M - H]^-$	C ₄₆ H ₇₅ O ₁₀ P	79.5	-10
30	480.3049	1	6.80	LysoPE(18:0)	$[M - H]^-$	C ₂₃ H ₄₈ NO ₇ P	88.8	-7

The untargeted metabolomics analysis: Ambient mass spectrometry techniques

The results obtained concern only the analysis of some organs' mice not belonging to the SpaceMice project. In fact, before analysing these precious and not reproducible samples, the method must be developed and refined. The samples analysed are kidneys belonging from an 8-week old mouse, GSK360A.

Figure 10.20 shows four different 50 μ m slides obtained from the same kidney sample. In the slides was monitored the presence of m/z 347.1059: the more the colour turns to yellow, the higher the concentration of that particular analyte is; on the contrary, the more the gradation tends to the blue, the amount of analyte is lower. In the figure 10.20 some scratches are present. They may be due to the speed of movement of the probe above the sample and the type of extraction solvent used. The elimination of this problem will, therefore, be a future objective. The figures 10.21 and 10.22 are just other examples to demonstrate how the method seems to be valid

to detect molecules of interest in real samples.

10.5. Conclusion and future prospects

The ultimate purpose of global metabolomics analyses can be expressed by three verbs: identification, quantification, tracking. While the detection and the quantification of metabolites through targeted techniques can be achieved by many laboratories, the comprehensive metabolomics approaches able to empower biomarkers' discovery are extremely difficult to conduct.

The SpaceMice project has succeeded in providing a comprehensive screening of the vast Metabolome reign. The described results show how different approaches (Biocrates kits, lipidomics analyses, DESI, etc.) are needed for a full and global comprehension of the biological system of interest. Retracing the arguments to the order in which they were treated, I will summarize the main results achieved and prospects. As for the Biocrates kit, its simplicity of use and the quantity/quality of the results provided make it an essential tool for global metabolomics analysis. The results shown are only relative to the FIA approach. More than 300 analytes have been identified and quantified, but only 213 analytes will be useful for the statistical analysis. Between all these analytes, only molecules belonging from phosphatidylcholines, aminoacids, sphingomyelins, and cholesteryl esters can be successfully used to discriminate within groups. A future perspective will be certain to verify the data obtained in C18 mode runs to understand first how many analytes have been identified by this approach and, secondly, how many of the analytes already detected by FIA have been confirmed again by LC mode. Moving now to the untargeted metabolomics analysis, we can draw interesting conclusions. The approach used allowed to obtain results both in HILIC and C18 cases. The molecules reported in table 10.4 have a double meaning: on the one hand, they are a proof of the validity of the method used; on the other hand, they are different from the molecules identified by the Biocrates kit. The two approaches are then complementary and provide non-redundant information. Therefore, both must be applied to have a complete overview of the sample being analysed.

The lipidomics approach's results can be interpreted similarly. As a mat-

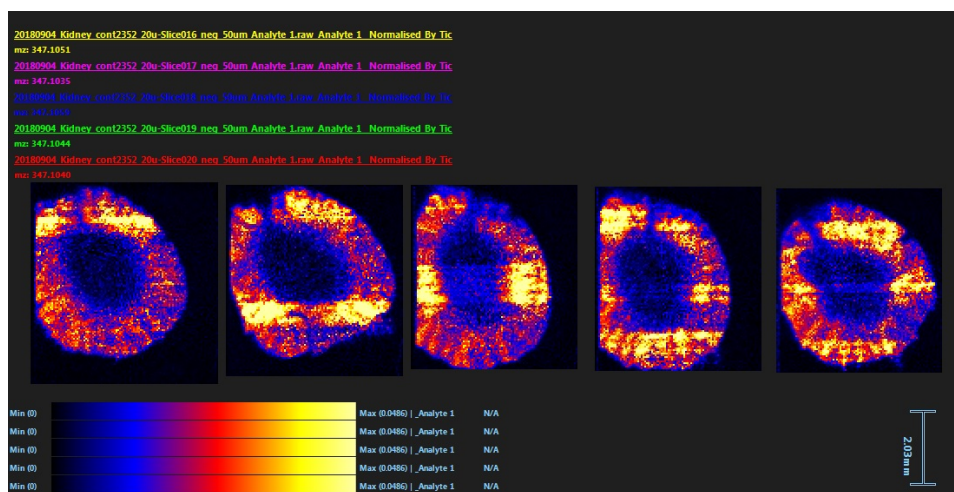


Figure 10.20.: *DESI image of a kidney sample - positive polarity*

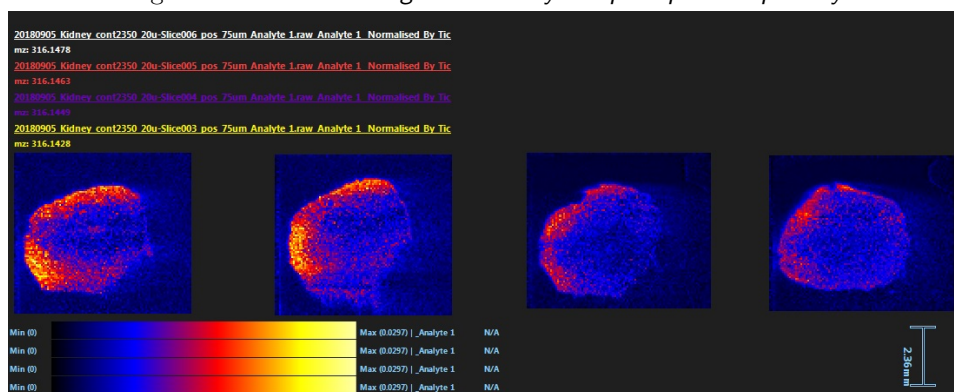


Figure 10.21.: *DESI image of a kidney sample - positive polarity*

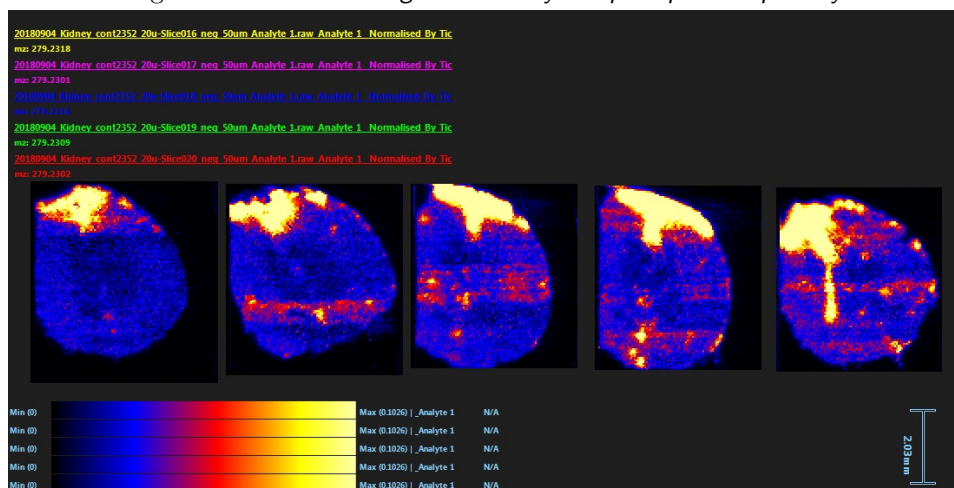


Figure 10.22.: *DESI image of a kidney sample - positive polarity*

ter of fact, more than 30 different lipids compounds detected, and only some of them are the same molecules already detected by Biocrates kit. So, also in this case, as in the previous ones, the information provided by the analysis is not redundant but, on the contrary, to have a general understanding of the problem, it must be taken into account. Finally, the ambient mass spectrometry techniques provide an additional view not given by any of the previous analyses. If, in fact, the previous analyses allowed to say “how much” of a specified analyte was present, the ambient mass spectrometry techniques allowed instead to say “where is” that specific molecule located in the tissue/organ. Furthermore, by overlapping the different sections of the same organ, it is possible to construct a 3D distribution of the analyte, implementing, even more, the knowledge.

References of Chapter 10

- [1] NASA - National Aeronautics et al. *NASA's twins study explores space through you*. 2016. URL: <https://phys.org/news/2016-04-nasa-twins-explores-space.html> (cit. on pp. 186, 187).
- [2] Biocrates Life Sciences AG. *Biocrates Life Sciences: The deep phenotyping company*. 2018. URL: <http://www.biocrates.com> (cit. on pp. 189–192).
- [3] Japan Aerospace Exploration Agency. *JAXA's 3rd Mouse breeding mission finished*. 2018. URL: http://iss.jaxa.jp/en/kiboexp/news/180719_mouse.html (cit. on pp. 186, 187).
- [4] C. Black et al. “The current and potential applications of Ambient Mass Spectrometry in detecting food fraud”. In: *Trends in Analytical Chemistry* 82, 268-278 (2016). DOI: 10.1016/j.trac.2016.06.005 (cit. on p. 199).
- [5] C. Byrdwell et al. “Quantitative Analysis of Triglycerides Using Atmospheric Pressure Chemical Ionization-Mass Spectrometry”. In: *Lipids* 31(9), 919-935 (1996). DOI: 10.1007/BF02522685 (cit. on p. 210).
- [6] H. Chen et al. “What can we learn from ambient ionization techniques?” In: *Journal of the American Society for Mass Spectrometry* 20(11), 1947-63 (2009). DOI: 10.1016/j.jasms.2009.07.025 (cit. on p. 199).
- [7] K. Chughtai et al. “Mass Spectrometric Imaging for biomedical tissue analysis”. In: *Chemical Reviews* 110(5), 3237-3277 (2010). DOI: 10.1021/cr100012c (cit. on p. 199).
- [8] C.S. Clendinen. “Ambient Mass Spectrometry in Metabolomics”. In: *The Analyst* 142(17) (2017). DOI: 10.1039/C7AN00700K (cit. on p. 199).
- [9] R.G. Cooks et al. “Detection Technologies. Ambient mass spectrometry”. In: *Science* 311(5767), 1566-70 (2006). DOI: 10.1126/science.1119426 (cit. on p. 199).
- [10] Waters Corporation. *Waters Corporation: The Science of What's Possible*. 2019. URL: <https://www.waters.com> (cit. on p. 195).

-
- [11] M.J. Culzoni et al. “Ambient mass spectrometry technologies for the detection of falsified drugs”. In: *MedChemComm* 1(5), 9-19 (2014). DOI: 10.1039/C3MD00235G (cit. on p. 199).
- [12] I.M. Di Gangi et al. “Metabolomic profile in pancreatic cancer patients: a consensus-based approach to identify highly discriminating metabolites”. In: *Oncotarget* 7(5), 5815-5829 (2016). DOI: 10.18632/oncotarget.6808 (cit. on p. 189).
- [13] C Gonçalo et al. “Activation of Nrf2 by H₂O₂: De Novo Synthesis Versus Nuclear Translocation”. In: *Methods in Enzymology* 28, 157-171 (2013). DOI: 10.1016/B978-0-12-405881-1.00009-4 (cit. on p. 187).
- [14] D. Hampel et al. “Human Milk Metabolomics Using Biocrates AbsoluteIDQ® p180 Kit Assay (OR06-04-19)”. In: *Current Developments in Nutrition* 3(1), 36 (2019). DOI: 10.1093/cdn/nzz036.0R06-04-19 (cit. on p. 189).
- [15] J.J. Kamphorst et al. “Liquid chromatography – high resolution mass spectrometry analysis of fatty acid metabolism”. In: *Analytical Chemistry* 83(23), 9114-9122 (2011). DOI: 10.1021/ac202220b (cit. on p. 210).
- [16] T. Koal et al. “Standardized LC–MS/MS based steroid hormone profile-analysis”. In: *Journal of Steroid Biochemistry and Molecular Biology* 129(3-5), 129-138 (2012). DOI: 10.1016/j.jsbmb.2011.12.001 (cit. on p. 189).
- [17] N. Koshi et al. “Identification of novel biomarkers of hepatocellular carcinoma by a high definition mass spectrometry; UHPLC-QTOF/MS and DESI-MSI”. In: *Rapid communications in Mass Spectrometry - Wiley* Not available yet (2019). DOI: 10.1002/rcm.8551 (cit. on p. 199).
- [18] T. Lang et al. “Towards human exploration of space: the THESEUS review series on muscle and bone research priorities”. In: *NPJ Microgravity* 3, 8 (2017). DOI: 10.1038/s41526-017-0013-0 (cit. on p. 186).

- [19] W. Lu et al. “Metabolomic Analysis via Reversed-Phase Ion-Pairing Liquid Chromatography Coupled to a Stand Alone Orbitrap Mass Spectrometer”. In: *Analytical Chemistry* 82(8), 3212-3221 (2010). DOI: 10.1021/ac902837x (cit. on p. 210).
- [20] W.O. Osburn et al. “Nrf2 signaling: an adaptive response pathway for protection against environmental toxic insults”. In: *Mutation Research - Genetic Toxicology and Environmental Mutagenesis* 659(1-2), 31-39 (2008). DOI: 10.1016/j.mrrev.2007.11.006 (cit. on p. 187).
- [21] L.E. Peterson et al. “Adjustment of lifetime risks of space radiation-induced cancer by the healthy worker effect and cancer misclassification”. In: *Heliyon* 1(4), e00048 (2015). DOI: 10.1016/j.heliyon.2015.e00048 (cit. on p. 186).
- [22] M. Qiang. “Role of Nrf2 in Oxidative Stress and Toxicity”. In: *Annual Review of Pharmacology and Toxicology* 53, 401-426 (2013). DOI: 10.1146/annurev-pharmtox-011112-140320 (cit. on p. 187).
- [23] D. Saigusa et al. “Establishment of Protocols for Global Metabolomics by LC-MS for Biomarker Discovery”. In: *Plos One* (2016). DOI: 10.1371/journal.pone.0160555 (cit. on p. 193).
- [24] D. Saigusa et al. “Simultaneous Quantification of Sphingolipids in Small Quantities of Liver by LC-MS/MS”. In: *Mass Spectrometry Society of Japan* 3(S0046), 1-8 (2014). DOI: 10.5702/massspectrometry.S0046 (cit. on p. 197).
- [25] J. Takyi-Williams et al. “Ambient ionization MS for bioanalysis: recent developments and challenges”. In: *Bioanalysis* 7(15), 1901-23 (2015). DOI: 10.4155/bio.15.116 (cit. on p. 199).
- [26] Tohoku University. *Tohoku University: Creating global excellence*. 2019. URL: <http://www.tohoku.ac.jp/en/> (cit. on p. 185).
- [27] Tohoku University. *ToMMo - Tohoku Medical Megabank Organization*. 2019. URL: <https://www.megabank.tohoku.ac.jp/english/> (cit. on p. 185).

- [28] G.J. Van Berkel et al. “Established and emerging atmospheric pressure surface sampling/ionization techniques for mass spectrometry”. In: *Journal of Mass Spectrometry* 43(9), 1161-80 (2008). DOI: 10.1002/jms.1440 (cit. on p. 199).
- [29] C.J. Wruck et al. “Role of oxidative stress in rheumatoid arthritis:insights from the Nrf2-knockout mice”. In: *Annals of the Rheumatic Diseases* 70(5), 844-50 (2010). DOI: 10.1136/ard.2010.132720 (cit. on p. 187).

Untargeted analysis: Impacts of NRF2 Activation in non Small Lung Cancer Cell Lines

11.1. Introduction

Cancer cells can alter metabolic activities compared to healthy cells to provide useful nutrients for aggressive proliferation and survival [9]. So, the detection of a re-programmed metabolism is always considered as a marker for cancer [13, 22]. These cancer cells' metabolic activities consist of self-advantages, as incorporation of nutrients into the biomass or increase in antioxidant and detoxification abilities [2, 9]. Besides, different new studies show how the altered metabolism of cancer cells modulates the tumor micro-environment [23, 26].

The major defence mechanism applied by multicellular organisms to counteract reactive oxygen species (ROS) is the KEAP1-NRF2 system [16, 29]. The key proteins within this defence mechanism are the transcription factor NRF2 (nuclear factor erythroid 2-related factor 2) that binds together with small Maf (musculoaponeurotic fibrosarcoma) proteins to the antioxidant species, and Keap1 (Kelch ECH associating protein 1), a repressor cysteine-rich protein that binds to NRF2 and promotes its degradation via a ubiquitin-proteasome pathway in non-stressful conditions. [16]. A scheme of Keap1–NRF2 system is reported in figure 11.1. In normal conditions, NRF2 is constantly ubiquitinated through Keap1 and degraded in the proteasome. After the exposure to electrophiles or oxidative stress species, the repressor Keap1 protein is inactivated. Then, stabilized NRF2 is accumulated in the nucleus of the cell and activates many cytoprotective genes.

By protecting cells from various insults and pathological conditions, the activation of the NRF2 is therefore essential for human health [28]. In contrast to the protecting role of the NRF2 factor, cancer cells often hijack

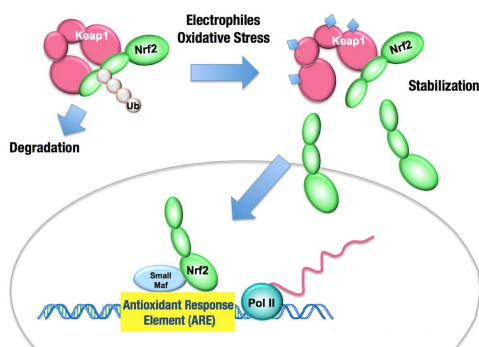


Figure 11.1.: *The Keap1–Nrf2 system. NRF2 is constantly ubiquitinated (Ub, ubiquitin) through Keap1. When ROS species interact with Keap1–NRF2, Keap1 is inactivated. Successively, NRF2 accumulates in the nucleus of the cell (the environment surrounded by the grey line) and activates many cytoprotective genes. Figure is extrapolated from the article [20]*

the KEAP1-NRF2 system [21].

Different shreds of evidence show how the aberrantly activated NRF2 in cancer cells promotes their malignant progression and causes an NRF2 addiction too [18, 19]. Recently results published by the ToMMO department show how, under the influence of the microenvironment, NRF2 actively promotes tumorigenesis by inducing cytokines and prostaglandin-metabolizing enzymes [17]. The results demonstrate the presence of a communication system between cancer cells and their microenvironments.

11.2. Aim of the work

In this study, we chose to analyse the nutrients and metabolites that act as signalling molecules for cancer cells to communicate with their microenvironments, characterizing the metabolic signatures of the inside and outside of NRF2-addicted cancer cells in comparison with those of NRF2-normal cancer cells. In this way, the study aims to detect possible metabolic features useful for evaluating the NRF2 activation status in healthy tissues during the early stage of the illness. To accomplish this ambitious task, a global metabolomics analysis (conducted by UHPLC-QTOF/MS and LC-FTMS systems) suits perfectly.

11.3. Sample preparation and method settings

Method settings, analytical platforms, and results achieved reported in chapter 11 are only an extract of what was obtained and published in [25]. Considering the whole project, at first, to clarify mutation profiles, an exome analysis upon cell lines of non-small-cell lung cancers (NSCLCs) was conducted. NSCLC cell lines were then classified into three groups (by comparing mutation profiles and transcriptomes): NRF2-high, NRF2-middle, and NRF2-low cell lines. Successively, two metabolomics approaches were applied to extract the maximum information from NSCLC cell lines samples: a targeted metabolomics (T-Met) analysis was conducted for the cell lysates and culture supernatants, and a global metabolome (G-Met) analysis was either conducted only upon the culture supernatants. For this PhD thesis, only the aspects concerning the G-Met approach will be considered.

For what concerning sample preparation, all NSCLC cell lines were cultured in RPMI 1640 medium supplemented with 10% FBS with penicillin/streptomycin under 5.0% CO₂ at 37° C. Culture supernatants were lately collected and kept at -80°C until sample preparation and analysis. The whole extraction procedure was operated by a Microlab STARlet robot system (Hamilton, Reno, NV) to avoid to minimize the possible random errors made by an operator. The sample extraction is the one reported in [24]. 50 µL of sample is extracted with 150 µL of methanol containing 0.1% formic acid. After mixing for 5 min and 5 min in the ultrasonic bath, the sample is centrifugated at 6,440 g for 20 min at 4° C. 120 µL of the supernatant was transferred to another 96-well plate and diluted with 120 µL of 0.1% formic acid in water. The sample to prepared is ready to be analysed by UHPLC-QTOF/MS and LC-FTMS systems. The quality control samples were created too following the same method already explained in chapter 10 and reported in [24].

The UHPLC-QTOF/MS system consisted of an Acquity Ultra Performance LC I-class system (Waters Corp., Milford, MA, US) and a Synapt G2-Si QTOF MS with an electrospray ionization system (Waters Corp., Manchester, UK) operated in both positive and negative ion modes. LC separation was performed using a C18 column (Acquity HSS T3; 150 mm

× 2.1 mm i.d., 1.8 μm particle size; Waters Corp., Taunton, MA, US) with a gradient elution of solvent A (water containing 0.01% formic acid) and solvent B (acetonitrile containing 0.01% formic acid) at 0.4 mL min⁻¹. The data were collected using MassLynx v4.1 software (Waters Corp.). The LC-FTMS system consisted of a NANOSPACE SI-II HPLC (Osaka Soda, Kyoto, Japan) and a Q Exactive Orbitrap MS with a heated-ESI-II source (Thermo Fisher Scientific, San Jose, CA) operated in both positive and negative ion mode analysis. LC separation was performed using a HILIC column (ZIC,R-pHILIC; 100 mm × 2.1 mm i.d., 5 μm particle size; Sequant, Darmstadt, Germany) with a gradient elution of solvent A (10 mmol L⁻¹ ammonium bicarbonate in water, pH 9.2) and solvent B (acetonitrile) at 0.3 mL min⁻¹. The data were collected using Xcalibur v4.1 software (Thermo Fisher Scientific). Using this approach, four different assays are generated (two different assays for each analytical platform, both conducted as in positive as in negative ion modes).

The data obtained from Thermo platform as from Waters Corp. system were processed with Progenesis QI software from Waters Corp. for peak picking, alignment, and normalization to produce peak intensities for tR and m/z data pairs. First, the features were selected based on their coefficient of variation (CV) and according to the inverse correlation of the dilution fold and the peak intensity with our normalization process to exclude contaminants following the previous report¹⁷. Second, the features, which were significantly higher in the Nrf2 high-activated (High) group than in the low-activated (Low) group as the candidate functional biomarker, were extracted with a p-value ($p < 0.05$) calculated by the Wilcoxon rank sum test. Then, the extracted features were identified by matching the results of the MS spectra, MS/MS spectra, isotopic ratio, mass accuracy, and chromatographic tR with the compounds in the Human Metabolome Database (HMDB, [4]), the LIPID MAPS Structure Database (LMSD, [12]), mzCloud ([14]) and ChemSpider ([5]) from the lowest p-value in each assay. Finally, the histograms were created to compare the relative intensity of each compound among the High, Low, and medium as a necessary condition of the culture medium. Statistical significance was determined by Wilcoxon rank sum test for the metabolites in the H and L cell groups with mean

of the three replicate measurements for each cell line, satisfying a Bonferroni adjustment ($p < 6.25 \text{ E-}4$, $p < 1.04 \text{ E-}3$, $p < 1.16 \text{ E-}3$, and $p < 8.47 \text{ E-}4$) for measurement mode HILIC-negative, HILIC-positive, C18-negative, C18-positive, respectively, by using the R package (ver. 3.4.2).

11.4. Results and discussion

In a previous study [17] has been demonstrated how the immune response is suppressed in tumours of NRF2-activated cancer cells. That's means that NRF2 helps cancer cells to evade the anti-tumor immunity actions of the biological system. In this work, we focused on the metabolites consumed from and excreted to the microenvironments by cancer cells. Cancer and non-cancer cells in their microenvironments compete for a specific metabolite, and alternatively, cancer cells excrete a specific metabolite to instruct the cells in their microenvironment. The global metabolomics approach applied to culture supernatants of NSCLC cell lines (11 of NRF2-high cells and 7 of NRF2-low cells) highlight interesting aspects. The detected metabolites could be categorized into seven groups upon the different concentrations between NRF2-high and NRF2-low cancer cell lines (figure 11.2).

Lipids and their derivatives are the predominant species of the metabolites that were higher in NRF2-high cells than in NRF2-low cells (figure 11.2, section A). For example, N-acetylsphingosine-1-phosphate (figure 11.2, section B, panel iii), glycerophosphocholine (figure 11.2, section B, panel iv), lysophosphatidic acid (LPA) (18:0) (figure 11.2, section B, panel vi), 10,16-dihydroxypalmitic acid (figure 11.2, section B, panel viii), 3-hexenedioic acid (figure 11.2, section B, panel x), 11-hydroxy-3,21-dioxoolean-12-en-28-oic acid (figure 11.2, section B, panel xv), and 3-acetamido-3-aminohexanoic acid (figure 11.2, section B, panel xvii) were lipid metabolites that were likely to be exported from NRF2-high cells. NRF2-high cells consume larger amounts of triglyceride (TG) (18:2/20:4/21:0) (figure 11.2, section C, panel iv) from the culture media than NRF2-low cells,

In contrast, dipeptides were lower in NRF2-high cells than in NRF2-low cells (Figure A). For example, phenylalanyl-arginine was excreted from the NSCLC cell lines with low NRF2 activity but not from those with high

NRF2 activity (figure 11.2, section C; panel i). These results imply that NRF2-high cells do not excrete dipeptides, but rather, utilize them more efficiently than NRF2-normal cells. A last interesting result to remark concerns the hypoxanthine: in facts, this molecule has been detected in the culture supernatants of NRF2-high cells in two independent measurement methods: C18-negative mode and C18-positive mode (figure 11.2, section B, panels v and xii).

All these molecules could be considered as suitable candidates for novel biomarkers that indicate the activation status of NRF2. However, the features cited in figure 11.2 are not the only ones detected by the four analytical methods used. In table 11.1 the 194 different molecules detected in the four different assays are reported. So, what's reported in the table 11.1 shows once more how the different methods developed provide valid data, in correlation with results already present in literature.

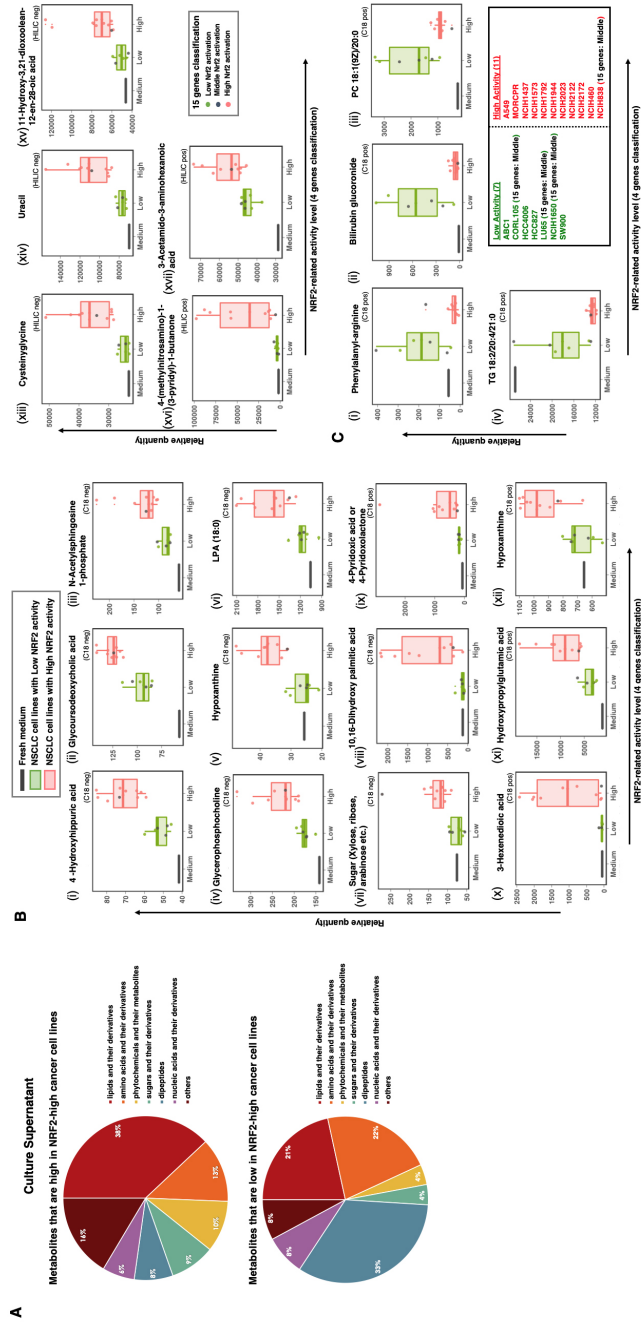


Figure 11.2.: Figure A: Pie charts representing the different categories in which the detected metabolites are divided among NRF2-high and NRF2-low cancer cell lines; Figure B: box and whisker plots of analytes high expressed in NRF2-high cancer cells samples; Figure C: box and whisker plots of analytes low expressed in NRF2-high cancer cells samples. The figures are extrapolated from the article [25]

Table 11.1.: Features detected in the four different assays: HILIC-negative, HILIC-positive, C18-negative, C18-positive. The ✓ mark indicates where the specific analyte has been detected. All the 194 detected molecules are reported

Compound name	HILIC- POS Presence	HILIC- POS Reference	HILIC- NEG Presence	HILIC- NEG Reference	C18-POS Presence	C18-POS Reference	C18-NEG Presence	C18-NEG Reference
Hypoxanthine	✓				✓		✓	
22:2 Cholesteryl ester					✓			
2-Acetyl-5-tetrahydroxybutyl imidazole (THI)					✓	[10]		
3-Hexenedioic acid					✓	[15]		
3-methylglutaconyl-CoA					✓			
4-Pyridoxic acid or 4-Pyridoxolactone					✓	[11]		
5-hydroxy-tryptophan					✓			
Acetylcysteine					✓			
Acetylserine or Glutamic acid					✓			
Arginine-Glutamine					✓			
Arginyl Isoleucine					✓			
Asparaginy-Tyrosine					✓			
Bilirubin glucuronide					✓	[30]		
Caffeic acid 3-sulfate					✓	[7]		
CoA(24:5(6Z,9Z,12Z,15Z,18Z))					✓			
Creatine					✓			
Cysteinyl-Proline					✓			
Deoxycortisol					✓	[7]		
Galalpha1-4Galbeta1-4Glcbeta-Cer(d18:1/26:0)					✓			
Galalpha1-4Galbeta1-4Glcbeta-Cer(d18:1/26:1)					✓			
Gamma-glutaminy-Triptophan					✓			
Gamma-Glutamylcysteine					✓			
Glutaminy-Phenylalanine					✓			
Glycerophosphocoline					✓	[3]	✓	
Glycocholic acid					✓	[7]		
GPCho(18:3/5:0)					✓			
Guanine or hydroxy-adenine					✓			
Histidiny-Hydroxyproline					✓			
Hydroxypropyl-Glutamate					✓			
Isobutyryl-L-carnitine					✓			
Isoleucyl-Lysine					✓			
Isoleucyl-Threonine or Prolyl-Valine					✓	[27]		

Continues in next page

Continues from previous page

Compound name	HILIC- POS Presence	HILIC- POS Reference	HILIC- NEG Presence	HILIC- NEG Reference	C18-POS Presence	C18-POS Reference	C18-NEG Presence	C18-NEG Reference
Lactosylceramide(d18:0/24:0)					✓			
L-gamma-glutamyl-L-isoleucine					✓			
L-Histidine					✓	[27]		
L-isoleucyl-L-proline					✓			
Methyladenosine					✓			
N-Acetyl-L-methionine					✓			
N-Acetylsphingosine-1-phosphate					✓	[8]	✓	
N-palmitoyl methionine					✓			
Oxalosuccinic acid					✓	[7]		
Palmitoyl carnitine					✓	[7]		
PC(10:0/0:0)					✓			
PC(18:1(9Z)/20:0)					✓	[3, 7]		
PG(18:3(9Z,12Z,15Z)/0:0)					✓			
Phenylacetamide or Acetylarlylamide or Aminoacetophenone					✓			
Phenylalanine-Tyrosine					✓	[27]		
Phenylalanyl-arginine					✓	[27]		
Sphinganine-1-Phosphate					✓			
Sphingosine-1-Phosphate					✓			
TG(18:2/20:4/21:0)					✓			
Threoninyl-Phenylalanine					✓	[27]	✓	
Threoninyl-Tyrosine					✓	[27]		
Thyroxine					✓	[27]		
Trimethyl-2E-hendecenoyl-CoA					✓			
Tryptophan-Arginine					✓			
Valyl-Leucine					✓	[27]		
3-Acetamido-3-aminohexanoic acid	✓							
4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone	✓							
4-Dimethylalyltryptophan	✓							
5-Aminovaleric acid or isomers	✓							
5-Hydroxykynurenamine	✓							
5-Methylcytosine	✓							
5'-S-Methyl-5'-thioinosine	✓							
6-Methylcytosine	✓							
Aminohippuric acid	✓							
Beta-Alanyl-arginine	✓	[27]						
Biotin	✓							

Continues in next page

Continues from previous page

Compound name	HILIC- POS Presence	HILIC- POS Reference	HILIC- NEG Presence	HILIC- NEG Reference	C18-POS Presence	C18-POS Reference	C18-NEG Presence	C18-NEG Reference
Creatinine	✓	[8]						
Cystine	✓		✓					
Cytidine	✓							
Desthiobiotin	✓							
Ethylendiaminatetracetic acid	✓							
Gamma-Aminobutyric acid	✓							
Gamma-Glutamylglutamine	✓	[6]						
Glycine-Lysine	✓							
Imidazolemethanol	✓							
L-beta-Aspartyl-L-arginine	✓							
L-Citrulline	✓							
L-Histidyl-L-Leucine	✓	[27]						
L-leucyl-L-proline	✓	[8]		[27]				
L-Serine	✓	[27]						
Methionine	✓							
n(2)-Succinyl-L-arginine	✓	[27] *						
N-(4-Oxobutyl)-L-glutamine	✓	[6]						
O-amino-L-homoserine	✓							
Octanoyl-L-carnitine	✓							
PA(20:0/0:0)	✓							
PC(16:0/9:0(CHO))	✓							
Pyrazole or Imidazole	✓							
S-Formylglutathione	✓							
SM(d18:1/16:0)	✓							
SM(d18:1/22:0)	✓							
S-Prenyl-L-Cysteine	✓							
Taurine	✓	[8]		[27]				
Tyramine	✓							
(2R,3S,4R,5R,6S)-3-Amino-2,4,5,6-tetrahydrocyclohexanone			✓					
(2R,3S,4R,5S)a-amino-2,3,4-trihydroxycyclohexanone			✓					
(2S)-2-(beta-D-Glucopyranosyloxy)-3-methyl-3-butenenitrile			✓					
(Nitrosulfonyl)diazene			✓					
(R)-3-(4-Hydroxyphenyl)-2-hydroxypropionic acid			✓					
10,16-Dihydroxy palmitic acid							✓	
11-Hydroxy-3,21-dioxolean-12-en-28-oic acid			✓					
18-oxocortisol							✓	

Continues in next page

Continues from previous page

Compound name	HILIC- POS Presence	HILIC- POS Reference	HILIC- NEG Presence	HILIC- NEG Reference	C18-POS Presence	C18-POS Reference	C18-NEG Presence	C18-NEG Reference
2'-Deoxy-5-(hydroxymethyl)cytidine-5'-(trihydrogen phosphate) diphosphate			✓					
2'-Deoxy-5-(hydroxymethyl)cytidine-6'-(trihydrogen phosphate) diphosphate			✓					
2'-deoxy-5'-inosinic acid			✓					
2'-Deoxy-5-methylcytidine			✓					
2-Hydroxy-8-methyl-2H-chromene-2-carboxylic acid			✓					
2-Hydroxyglutarate							✓	
3-[[5S,6R)-5,6-Dihydroxy-1,3-cyclohexadien-1-yl]propionic acid			✓					
3-dehydroshikimic acid			✓					
4-dehydroshikimic acid			✓					
4-hydroxy-9-methoxy-7H-furo[3,2-g]chromen-7-one			✓					
4-Hydroxyhippuric acid							✓	
4-O-beta-D-xylo-Hexopyranosyl-3-ulose-D-glucopyranose			✓					
4-oxo-D-Norvaline			✓					
4-oxoproline			✓	[8]				
5-Hydroxyindoleacetic acid							✓	
5'-Phosphoribosyl-N-formylglycinamide			✓					
6-Acetamido-3-oxohexanoic acid			✓					
Adrenic acid			✓					
Alanyl-Leucine							✓	[27]
Amino(3,5-dihydroxyphenyl)acetic acid			✓					
Arachidonic acid				[1]			✓	
Asparaginy-Hydroxyproline							✓	
Aspartyl-Tyrosine							✓	[27]
Beta-Cytryl-glutamic acid							✓	
beta-D-glucuronamide			✓					
Butyryl carnitine							✓	
Butyryl dihydrogen phosphate			✓					
Citric acid			✓					
Citrulline			✓					
Cysteineglutathione disulfide							✓	
Cysteinyglycine			✓					
Cysteinyglycone disulfide			✓					
Dehydroascorbic acid							✓	
D-Glucopyranose			✓					

Continues in next page

Continues from previous page

Compound name	HILIC- POS Presence	HILIC- POS Reference	HILIC- NEG Presence	HILIC- NEG Reference	C18-POS Presence	C18-POS Reference	C18-NEG Presence	C18-NEG Reference
Didehydro-D-gluconic acid			✓					
Dihydrotricetin			✓					
Dodecanedioic acid							✓	
Galactopyranose molecule			✓					
Galactopyranosyl-D-mannopyranose			✓					
Gamma glutamylglutamic acid			✓					
Gamma-Glutamyl-L-Leucine							✓	[27]
Gamma-Glutamyl-Tyrosine							✓	[27]
Gluconolactone			✓					
Galactopyranuronic acid		✓						
Glucosyl sphingosine							✓	
Glutamic acid			✓	[27]				
Glutamylphenylalanine							✓	
Glutaric acid			✓					
Glycoursodeoxycholic acid							✓	
Glycyl-Phenylalanine							✓	
Homocysteine							✓	
Hydroxydodecanedioic acid							✓	
Hydroxyoctanoic acid							✓	
Hydroxyvalproic acid							✓	
Inosinic acid			✓					
Isomaltose			✓					
Kojic acid			✓					
Leucyl-L-proline							✓	
Linsodamine			✓					
Methylphosphate			✓					
Mimosine			✓					
N-Acetylglucosamine 6-sulfate			✓					
N-cyclopropylammelide			✓					
Nimesulide			✓					
N-Methylserotonin							✓	
Organic acid like xylonic acid			✓					
Oxoglutaric acid			✓					
PA(18:0/0:0)							✓	
Palmitoyl threonine							✓	
PC(14:0/2:0)							✓	

Continues in next page

Continues from previous page

Compound name	HILIC- POS Presence	HILIC- POS Reference	HILIC- NEG Presence	HILIC- NEG Reference	C18-POS Presence	C18-POS Reference	C18-NEG Presence	C18-NEG Reference
PC(18:1/0:0)							✓	
Phosphohydroxypyruvic acid							✓	
Prostaglandin G1							✓	
Purine riboside			✓					
Pyridoxal			✓					
Ribulofuranose			✓					
S-Adenosyl-L-homocysteine			✓					
Scrophulein (as know as Cirsimaritin)			✓					
Succinic acid			✓				✓	
Sugar like fructopyranose			✓					
Sugar like ribose							✓	
Theophylline			✓					
Threoninyl-Leucine							✓	
Tridecynoic acid							✓	
Uracil			✓					
Valyl-Proline							✓	

The table ends here

11.5. Conclusion and future prospects

In this work, four different assays (HILIC-negative, HILIC-positive, C18-negative, C18-positive) were applied to detect possible analytes useful to evaluating the NRF2 activation status in healthy tissues during the early stage of cancer pathology. One hundred ninety-four different molecules belonging necessarily from lipid and dipeptides classes were detected. The abundance of lipid molecules in NRF2-high cells is a clear symptom of how lipid metabolism is enhanced in these cancer cells. On the contrary, peptide structures are characteristic of NRF2-low cancer cells. So, this study provides invaluable information for the exploration of new metabolic nodes that are unique to NRF2-activated NSCLC for the development of anti-cancer drugs that are effective for improving the clinical outcome of NRF2-activated NSCLC patients.

References of Chapter 11

- [1] E.G. Armitage et al. “Metabolomics in cancer biomarker discovery: Current trends and future perspectives”. In: *Journal of Pharmaceutical and Biomedical Analysis* 87, 1-11 (2014). DOI: 10.1016/j.jpba.2013.08.041 (cit. on p. 231).
- [2] T. Bohn et al. “Tumor immunoevasion via acidosis-dependent induction of regulatory tumor-associated macrophages”. In: *Nature Immunology* 19, 1319-1329 (2018). DOI: 10.1038/s41590-018-0226-8 (cit. on p. 221).
- [3] B. Callejón-Leblic et al. “Metabolic profiling of potential lung cancer biomarkers using bronchoalveolar lavage fluid and the integrated direct infusion/ gas chromatography mass spectrometry platform”. In: *Journal of Proteomics* 145, 197-206 (2016). DOI: 10.1016/j.jprot.2016.05.030 (cit. on pp. 228, 229).
- [4] Canada Foundation for Innovation Canadian Institutes of Health Research et al. *The Human Metabolome Database*. 2019. URL: <http://www.hmdb.ca> (cit. on p. 224).
- [5] Royal Society of Chemistry. *ChemSpider - Search and share chemistry*. 2020. URL: <https://www.chemspider.com> (cit. on p. 224).
- [6] G.Y. Chen et al. “Rapid quantification of glutaminase 2 (GLS2)-related metabolites by HILICMS/MS”. In: *Analytical Biochemistry* 539, 39-44 (2017). DOI: 10.1016/j.ab.2017.10.002 (cit. on p. 230).
- [7] Y. Chen et al. “Metabolomic profiling of human serum in lung cancer patients using liquid chromatography/hybrid quadrupole time-of-flight mass spectrometry and gas chromatography/mass spectrometry”. In: *Journal of Cancer Research and Clinical Oncology* 141, 705-718 (2015). DOI: 10.1007/s00432-014-1846-5 (cit. on pp. 228, 229).
- [8] M. Ciborowski et al. “Development of LC-QTOF-MS method for human lung tissue fingerprinting. A preliminary application to non-small cell lung cancer”. In: *Electrophoresis* 38, 2304-2312 (2017). DOI: 10.1002/elps.201700022 (cit. on pp. 229-231).

-
- [9] R.J. DeBerardinis et al. “Fundamentals of cancer metabolism”. In: *Science Advances* 2(5), e1600200 (2016). DOI: 10.1126/sciadv.1600200 (cit. on p. 221).
- [10] C. Fabiani et al. “2-Acetyl-5-tetrahydroxybutyl imidazole (THI) protects 661W cells against oxidative stress”. In: *Naunyn-Schmiedeberg’s Archives of Pharmacology* 390, 741-751 (2017). DOI: 10.1007/s00210-017-1374-3 (cit. on p. 228).
- [11] L. Galluzzi et al. “Effects of vitamin B6 metabolism on oncogenesis, tumor progression and therapeutic responses”. In: *Oncogene* 32, 4495-5004 (2013). DOI: 10.1038/onc.2012.623 (cit. on p. 228).
- [12] LIPID MAPS Lipidomics Gateway. *LIPID MAPS Lipidomics Gateway*. 2020. URL: <https://www.lipidmaps.org/> (cit. on p. 224).
- [13] D. Hanahan et al. “Hallmarks of Cancer: The Next Generation”. In: *Cell* 144(5), 646-674 (2011). DOI: 10.1016/j.cell.2011.02.013 (cit. on p. 221).
- [14] Slovakia HighChem LLC. *MzCloud - Advanced Mass Spectral Database*. 2019. URL: <https://www.mzcloud.org> (cit. on p. 224).
- [15] S. Hory et al. “A metabolomic approach to lung cancer”. In: *Lung Cancer* 74, 284-292 (2011). DOI: 10.1016/j.lungcan.2011.02.008 (cit. on p. 228).
- [16] E. Kansanen et al. “The Keap1-Nrf2 pathway: Mechanisms of activation and dysregulation in cancer”. In: *Redox Biology* 1(1), 45-49 (2013). DOI: 10.1016/j.redox.2012.10.001 (cit. on p. 221).
- [17] H. Kitamura et al. “IL-11 contribution to tumorigenesis in an NRF2 addiction cancer model”. In: *Oncogene* 36(45), 6315-6324 (2017). DOI: 10.1038/onc.2017.236 (cit. on pp. 222, 225).
- [18] H. Kitamura et al. “NRF2 addiction in cancer cells”. In: *Cancer Science* 109(4), 900-911 (2018). DOI: 10.1111/cas.13537 (cit. on p. 222).
- [19] Y. Mitsuishi et al. “Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming”. In: *Cancer Cell* 22(1), 66-79 (2012). DOI: 10.1016/j.ccr.2012.05.016 (cit. on p. 222).

- [20] Y. Mitsuishi et al. “The Keap1–Nrf2 system in cancers: stress response and anabolic metabolism”. In: *Frontiers in Oncology* 2, 200 (2012). DOI: 10.3389/fonc.2012.00200 (cit. on p. 222).
- [21] H. Motohashi et al. “L-49 - Bright and dark sides of KEAP1-NRF2 system in carcinogenesis”. In: *Free Radical Biology and Medicine* 120(1), S18 (2018). DOI: 10.1016/j.freeradbiomed.2018.04.078 (cit. on p. 222).
- [22] K.C. Patra et al. “Hexokinase 2 Is Required for Tumor Initiation and Maintenance and Its Systemic Deletion Is Therapeutic in Mouse Models of Cancer”. In: *Cancer Cell* 24(2), 213-228 (2013). DOI: 10.1016/j.ccr.2013.06.014 (cit. on p. 221).
- [23] K. Renner et al. “Metabolic Hallmarks of Tumor and Immune Cells in the Tumor Microenvironment”. In: *Frontiers in Immunology* 8, 248 (2017). DOI: 10.3389/fimmu.2017.00248 (cit. on p. 221).
- [24] D. Saigusa et al. “Establishment of Protocols for Global Metabolomics by LC-MS for Biomarker Discovery”. In: *Plos One* (2016). DOI: 10.1371/journal.pone.0160555 (cit. on p. 223).
- [25] D. Saigusa et al. “Impacts of NRF2 Activation in Non-Small-Cell Lung Cancer Cell Lines on Extracellular Metabolites”. In: *Cancer Science* (2019). DOI: 10.1111/cas.14278 (cit. on pp. 223, 227).
- [26] H. Sekine et al. “Tumors sweeten macrophages with acids”. In: *Nature Immunology* 19(12), 1281-1283 (2018). DOI: 10.1038/s41590-018-0258-0 (cit. on p. 221).
- [27] M. Sugimoto et al. “Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles”. In: *Metabolomics* 6, 78-95 (2010). DOI: 10.1007/s11306-009-0178-y (cit. on pp. 228–232).
- [28] X.J. Xiao-Jun Wang et al. “Nrf2 protects human bladder urothelial cells from arsenite and monomethylarsonous acid toxicity”. In: *Toxicology and Applied Pharmacology* 225(2), 206-213 (2007). DOI: 10.1016/j.taap.2007.07.016 (cit. on p. 221).

- [29] M. Yamamoto et al. “The KEAP1-NRF2 System: a Thiol-Based Sensor-Effector Apparatus for Maintaining Redox Homeostasis”. In: *Physiological Reviews* 98(3), 1169-1203 (2018). DOI: 10.1152/physrev.00023.2017 (cit. on p. 221).
- [30] C. Zhou et al. “Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study”. In: *The Lancet Oncology* 12, 735-742 (2011). DOI: 10.1016/S1470-2045(11)70184-X (cit. on p. 228).

Untargeted analysis: Detection of OQDS markers. An untargeted metabolomics approach via HPLC-HRMS

12.1. Introduction

2013 will be recorded as *annus horribilis* for olive trees of Salento Peninsula in the Apulia region, south-eastern Italy [7]. Starting from this year until nowadays olive trees of Salento Peninsula began to decline and die with a condition of unknown etiology called “olive quick decline syndrome” (OQDS) [19]. OQDS is a destructive disorder characterized by the presence of leaf scorching and desiccation of twigs and small branches that in the early stages prevail in the upper part of the canopy, then extend to the rest of the crown, which acquires a burned look. The more severely affected plants are heavily pruned by the growers to favour new growth, which, however, does not give the expected results, and the trees desiccate in a short while [3].

The onset of this economically-devastating disease has been associated with *Xylella fastidiosa* (Xf) infection, which is a formidable Gram-negative plant pathogen¹ so far absent in Europe that has already caused enormous damage in the United States and South America [5]. This bacteria forms biofilms inside xylem vessels leading to its occlusion impairing water and mineral salt uptake [17]. This bacterium seems to be responsible for the de-

¹Bacteria are classified by the colour they turn after a chemical reaction called Gram staining. Gram-negative bacteria stain red when this process is used while Gram-positive pathogens turn blue. Bacteria stain differently because their cell walls are different. Gram-negative bacteria are enclosed in a protective capsule. Under the capsule, Gram-negative bacteria possess another membrane that protects them against antibiotics actions. When disrupted, this membrane releases toxic - endotoxins - which contribute to the severity of symptoms during infections.

velopment of some other economically significant diseases, including Pierce's disease of grapevine [10]; leaf scorch of almond, oleander, and coffee [9]; citrus variegated chlorosis; and other diseases of crop, forest, and landscape plants [12]. In the case of olive, the damage may be aggravated by the presence of fungi of different genera, *Phaeoacremonium* and *Phaemoniella* in particular, but also *Pleuromotomophora* and *Neofusicoccum*, which colonize and necrotize the sapwood [18]. Whether the OQDS is a result of a co-infection by other pests (fungi and-or insects) present on olive in Apulia region or related solely to the bacterial strain is yet unknown [19].

In this scenario, it is becoming more critical than ever to have an analytical method able to generate an early diagnosis of the disease. To date, it is possible to detect Xf's presence in olive trees by serological (ELISA²) and molecular (PCR) methods [2, 16].

OQDS is a multifactorial disease resulting in an unknown and very complex pathosystem. Given that, it is necessary to have tools to recognize the infected trees not only by detecting the bacterium, but also other characteristic molecular features involved in early infection state, such as molecules involved in a stress response.

12.2. Aim of the work

Metabolomics is one of the "omics" approaches that can be used to acquire comprehensive information on the composition of a metabolite pool to provide a functional screenshot of the cellular state at a molecular level. By employing the metabolomic approach, it is possible:

- to identify the factors involved in the induction of infection effectors;
- to understand the host response to the infection since the pathogenic attack involves an interaction between two partners that can only be partially assessed if focusing only on one of them.

The main goal of this work is to build an untargeted metabolomics approach. Thanks to this method, the analyst will detect OQDS molecular markers able to discriminate between healthy and infected trees. This

²It is an analytical biochemistry assay applied to detect the presence of a ligand like a protein in a liquid sample using antibodies directed against the protein to be measured.

global metabolomics approach, by measuring levels of thousands of metabolite features in a single analysis, suits entirely for our purposes thanks to the comprehensive information on the composition of a metabolite pool it can provide [20, 21].

However, HPLC-DAD-ESI HRMS metabolomics method [13] is just a starting point: in facts, devised biomarkers panel could be then exploited by an analytical laboratory through targeted metabolomics approaches. These methods require a priori knowledge of metabolites of interest but are cheaper, faster, and more accurate if compared to untargeted ones [27].

In addition to mass spectrometry, software like XCMS online and Statistica 6.0 were exploited for clustering pools of samples deriving from healthy (HP) and infected (OP) olive trees. Online databases allowed for the final identification of the principal metabolites involved in the host-pathogen interactions, which further MS/MS experiments can confirm.

12.3. Sample preparation, QC samples and method settings

12.3.1. Sampling, extraction protocols and QCs

The sample collection and preparation were exploited by the University of Salento under the supervision of Prof. Giuseppe Ciccarella. Healthy samples (HP) and samples with OQDS symptoms (OP) were collected directly from olive trees in different areas of the Salento Peninsula and Liguria region. The leaves collected from each sample tree were immediately shock-frozen with liquid nitrogen to block all metabolic processes and transferred to the laboratory for the extraction. It is fundamental to quench the metabolism as soon as possible, and a shock freezing using liquid nitrogen is the most common and efficient method to inactivate the metabolism and preserve all the metabolites.

Once the samples arrived at the laboratory, leaves were immersed in liquid nitrogen then manually grounded with a pestle and a mortar (pre-cooled and filled with liquid nitrogen to keep the cold chain and prevent degradation of the samples with the heat generated by the use of the mortar). 300

mg of fine powder were extracted with 1.2 mL of ethyl acetate in 1.5 mL Eppendorf tubes, sonicated for 15 min and centrifuged at 15000 g³ for 20 min. 500 μ L of the supernatants were transferred in new Eppendorf tubes and the solvent evaporated under a stream of nitrogen. The addition of 1 mL of water/acetonitrile 50:50 (v/v) to the dry extract was followed by sonication (15 min) and centrifugation (12000 rpm, 20 min). The supernatants were transferred into HPLC vials and analysed by a HPLC-HRMS platform. The choice of the extraction solvent is, therefore, essential and in our study ethyl acetate was chosen because it has been previously reported to be the best extraction solvent in terms of the number of metabolites with significant chemical and structural diversity detected by MS [1, 4, 11, 15]. Before conducting the analyses, each sample is spiked with three different internal standards: synephrine (168.1006 m/z pos, 166.0946 m/z neg), catechin (291.0845 m/z , pos, 289.0790 m/z neg), and galangine (271.0585 m/z pos, 269.0528 m/z neg). The use of three different internal standards (eluting by the column at different retention times) is a fundamental condition for having good results during the data alignment processing. XIC chromatograms of these standards for positive and negative ion modes are reported in figures 12.1 and 12.2 respectively.

QCs assess and ensure that the analytical method created is performed appropriately and meets the criteria defined a priori. In our case, the QC sample is a pooled sample in which a small aliquot (20 microliters) of each extracted sample under analysis is mixed in a 10-mL tube. By this way, pooled QC created represents the matrix as the metabolites' composition of *Xylella* samples. Last but not least aspect to remark: all the injections were performed in the same day-batch

12.3.2. HPLC-ESI-LTQ Orbitrap parameters

Analyses were performed on a HPLC-DAD-ESI HRMS operating as in positive as in negative ion mode. The instrument setup for all the analyses

³To convert RCF (relative centrifugal force, called g-force) to rpm (revolutions per minute) we can use the equation 12.1:

$$RCF = 1.12 \times \text{Radius of rotor} \times \left(\frac{\text{rpm}}{1000}\right)^2 \quad (12.1)$$

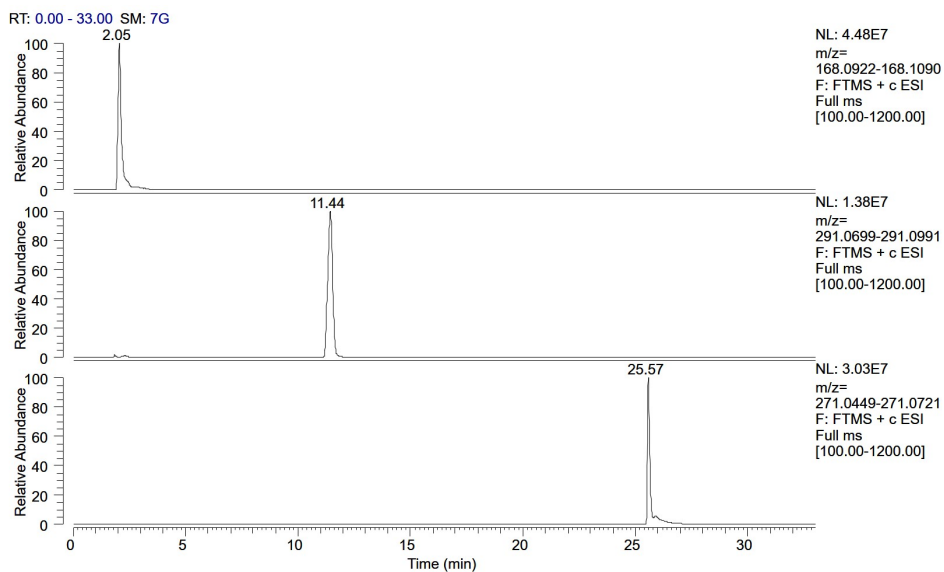


Figure 12.1.: *Extracted ion chromatograms (XIC) of three different internal standards applied for data alignment processing. From the top to the bottom: synephrine (168.1006 m/z), catechin (291.0845 m/z) and galangine (271.0585 m/z)*

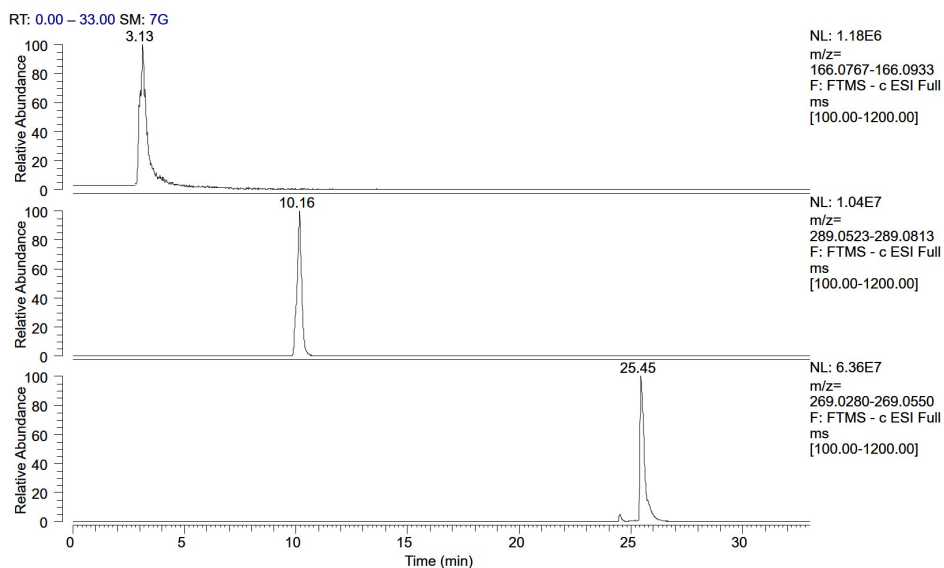


Figure 12.2.: *Extracted ion chromatograms (XIC) of three different internal standards applied for data alignment processing. From the top to the bottom: synephrine (166.0946 m/z), catechin (289.0790 m/z) and galangine (269.0528 m/z)*

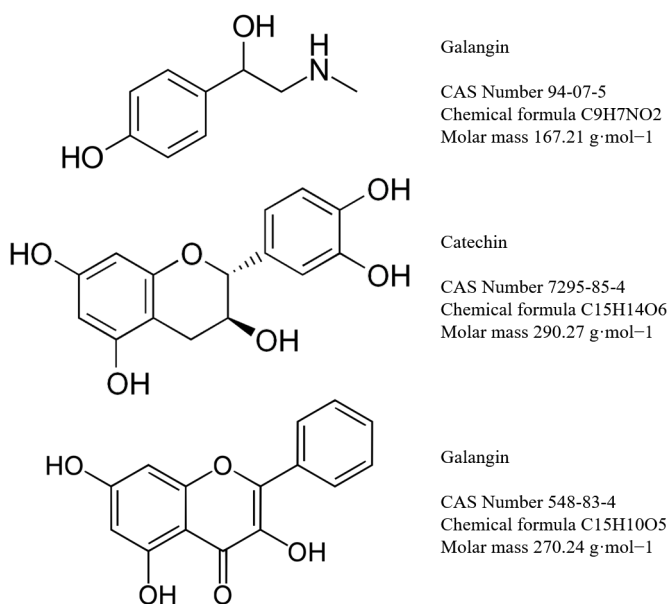


Figure 12.3.: *Chemical structures of synephrine, catechin, and galangin molecules. These analytes are the ones used in xylella studies for data normalisation*

consisted in a Dionex Ultimate 3000 HPLC system equipped with a solvent vacuum degasser (Thermo Scientific, Milan, Italy), a DAD Surveyor detector (Thermo Scientific, Milan, Italy), coupled with a High-Resolution Mass Spectrometer LTQ-Orbitrap (Thermo Scientific, Milan, Italy) through an electrospray ionization (ESI) interface. The polarities adopted for the ESI source were as positive as negative ion modes.

The chromatograms were recorded using the Thermo Xcalibur 2.1.0 software (Rev. SP1 1160). A Gemini NX-C18 column from Phenomenex (Gemini NX-C18, 2.0 × 150 mm, 3.0 μm, 110 Å, Phenomenex, Bologna, Italy) was used to obtain the chromatographic separation of the extracts. The mobile phases were 0.1% formic acid in H₂O(A) and acetonitrile (B) for positive ion mode analyses and ammonium acetate 0.005M H₂O(C) and acetonitrile (B) in negative ion mode. Gradient elution was set as follows: linear gradient from 5 to 20% B in 8 min, then 20% B held for 4 min, linear gradient from 20% to 30% B from 12 to 20 min, then in 2 min reaches 100% B, in min the percentage of acetonitrile returns to initial conditions and the 5% B is held for 10 min (from 23 to 33 min) to assure the correct equilibration of

the column at the initial conditions percentages. The flow rate was set at 200 $\mu\text{L}/\text{min}$. Global run time was 33 min. The column temperature was set at 25° C. Sample injection volume was 10 μL .

DAD acquisitions were carried out in the range of 200-800 nm. The HRMS system was operated in both positive and negative ionization modes. ESI tuning parameters were set as follows: the capillary voltage was 16 V (ESI+) and -13 V (ESI-); tube lens was set at 5 V (ESI+) and -36 V (ESI-); source voltage was set to 5.0 kV (ESI+) and 3.5 kV (ESI-); sheath gas and aux gas flow rate are respectively 35 and 20 arbitrary units in both methods; the spray current was set at 0.05 μA ; the capillary temperature was 270° C. The mass spectrometer was operated in full-scan mode in the m/z range 50-1200 (ESI+) or 100-1200 (ESI-), with a resolution of 30000 in FTMS mode. Tandem mass (MS/MS) spectra were automatically performed in the m/z range 50-1200 (ESI+) or 100-1200 (ESI-), using the automatic dependent scan function. The collision energy was set at 30 (arbitrary units) for all of the MS^n acquisitions.

All chromatograms and spectra were stored in centroid mode. Xcalibur (Thermo Scientific, Bremen, Germany) software was used both for acquisition and for elaboration and calculation.

12.3.3. Data processing and analysis

Once the quality of the TIC chromatograms obtained has been examined (by manually monitoring the presence over all the sample injected of internal standards and checking the loading pump pressure for each analysis) raw LC-MS data files are converted into *mzXML* files using the open-source software ProteoWizard – MSConvert. An example of TIC chromatogram obtained through the method discussed above is reported in figure 12.4. The conversion is an obliged process: as a matter of fact, files generated by HPLC-ESI-LTQ platform are *.raw format, which is incompatible with XCMS. The exact parameters applied in ProteoWizard software are reported in table 12.1.

After conversion, the files are submitted to XCMS online software [26, 28] in five different datasets:

1. fifteen HP_Liguria_healthy branch trees samples;

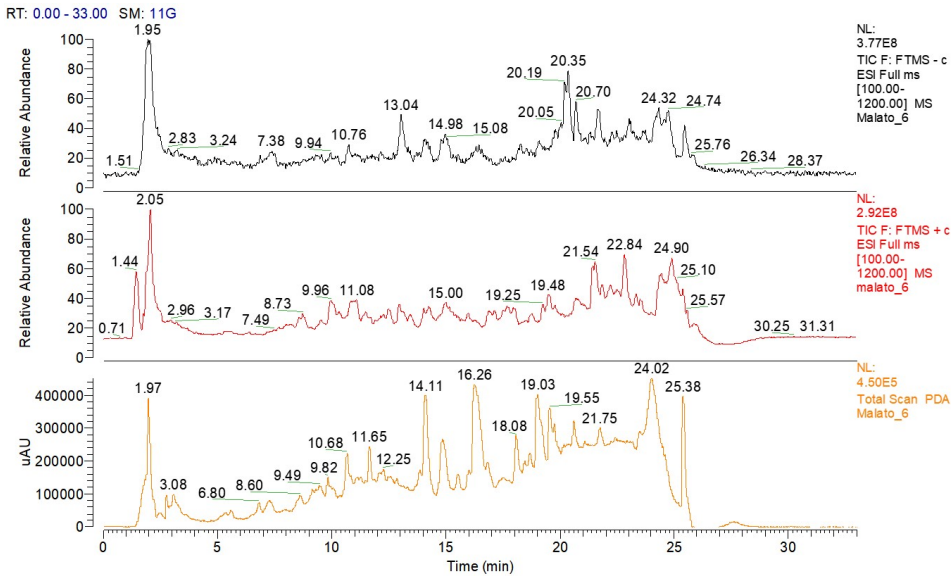


Figure 12.4.: Total ion chromatograms and PDA of a real OQDS sample. From the top to the bottom: TIC for positive ion mode, TIC for negative ion mode, PDA chromatogram

Table 12.1.: ProteoWizard parameters applied for *Xylella* .raw files conversion

Filter	Parameters
PeakPicking	Cwz: snr=1.0 peak Space=0.1 msLevel=1-1
LockmassRefiner	ESI-: mz=289.0790, mz=269.0528, mz=166.0946; tol=0.005
msLevel	ESI+: mz=291.0845, mz=271.0585, mz=168.1006; tol=0.005 1-2

2. fourteen HP_Apulia_ healthy branch trees samples;
3. thirteen HP_Apulia_desiccated branch trees samples;
4. twenty-six OP samples;
5. twenty quality control samples.

As reported previously, for the quality control (QC) samples, 20 microliters of each study sample were collected after the sample extraction, and the samples were mixed in a 10-mL tube and vortex for 1 min to homogenize the matrix.

Frequency of QC injections [14] was set by evaluating earlier publications, focusing particularly upon [23]. HP and OD samples were injected in randomized run order in the same day-batch. QC injections are essential for data correction and to obtain optimal signal correction results; at least one QC injection every four sample injections is required [23]. Additionally, at the beginning of the analysis, at least ten consecutive injections of QC sample are necessary to prime the column. Furthermore, it is strictly suggested to avoid injection of blank samples since the difference between blank and sample response could lead to analytical drifting in data in representation (PCA) and chromatographic elution-related reproducibility issues [23]. Online software processes all files using appropriate algorithms to filter the signals from background noise, corrects the drifts of the retention time between samples and align all total ion chromatograms (TICs) using the three different internal standards added to each sample, and finally check the quality of the correction procedure by overlapping all TICs. The parameters applied for the processing of the pairwise/multi-job (as negative as positive data analysis) were set as follows: centWave for feature detection ($\Delta m/z=30$ ppm, minimum and maximum peak width 30 and 120 s, respectively; S/N threshold = 6; Integration method = 1; Pref. peaks = 5; Pref intensity = 100000; obiwrap settings for retention time correction (profStep=1.0); parameters for chromatogram alignment, including mzwid = 0.015, minfrac = 0.5, and bw = 5; parameters for statistics includes Kruskal-Wallis non-parametric, perform post-hoc analysis = true, significance p-value threshold = 0.005, fold change threshold = 1.5; finally, parameters for annotation, including searching for isotopes, ppm error=5 and

m/z , and absolute error=0.05. All samples behaving as an outlier in PCA plots were removed to smoothen the distribution and PCA plot was recalculated. Identification of single metabolites was achieved by MS/MS spectrum match with available online databases such as METLIN [25], MoNA (<https://mona.fiehnlab.ucdavis.edu/>), UNPD (<https://omictools.com/unpd-tool>). XCMS csv output file was processed using MetaboAnalyst (<http://metaboanalyst.ca>) [29], an open bioinformatics platform for the implementation of statistical analysis of high-throughput metabolomic studies. Furthermore, XCMS data mining process was implemented by a principal component analysis (PCA) able to discriminate OP from the HP samples. PCA was conducted through the Statistica 6.0 software. To accomplish this task, only the variables identified by XCMS with a P values lower than 0.005 have been selected.

12.4. Results and discussion

Metabolomics is perhaps one of the most powerful tools for understanding the biological correlations with physiological or pathological conditions. Although the data sets generated from the processing of metabolomics experiments are generally extensive, fast and accurate, statistical and bioinformatic platforms, have made significant improvements to deal with this complexity facilitating data exploration and biological relationships interpretation. The samples obtained from the plant extracts were analysed by the HPLC-HRMS platform described in the experimental section. Samples analysed belong to four distinguished groups: OP_Apulia_infected samples, HP_Liguria_healthy branch trees, HP_Apulia_healthy branch trees, HP_Apulia_desiccated branch trees. OP_Apulia_infected samples are samples extracted at Salento University from plants infected by *Xylella fastidiosa* pathogen; HP_Liguria_healthy branch trees and HP_Apulia_healthy branch trees are leaves samples extracted from healthy olive trees grown in Pornassio (Imperia, Liguria region) and Salento (Apulia region) respectively; HP_Apulia_desiccated branch trees consist of leaves samples picked up from a desiccated branch of a healthy olive tree. Use of samples belonging to these groups allows understanding, in the case of group separa-

tion in PCA plots, if the features detected could be real markers for OQDS pathology or, on the contrary, they are markers to discriminate olive trees via spatial location (Apulia or Liguria regions) or dryness of the branch tree.

The data mining process operated by XCMS and Statistica 6.0 software reveals exciting features significantly up and down-regulated in OQDS infected olive trees samples ($P < 0.05$). Features could be used as putative markers to discriminate against the HP and OP samples. To have a clear and full idea of the most significant variations between the two groups, results can be appreciated by PCA analysis.

Thanks to PCA, the dimensionality of multivariate data can be reduced with a minimal loss in information. The PCA plots obtained in XCMS online software are shown in figures 12.5, 12.6 and 12.7, where each plot is related to a specific MS polarity (figure 12.5 for negative ion mode, figure 12.6 is a zoom of the PCA plot reported in figure 12.5 and figure 12.7 for positive ion mode). In our analysis, the first principal component's explained variance is, respectively, for positive and negative ion modes, 38% and 48%. The second principal components explain variance of respectively 10% (ESI+ mode) and 15% (ESI- mode). In PCA score plot related to negative ion mode three different clusters could be visualized (figure 2A, coloured circles): a cluster for infected samples (light blue triangle) a cluster for QC samples (red triangles), and a cluster for healthy olive trees samples (blue dots, dark blue diamonds, and green squares). Besides, positive ion mode PCA shows four groups: as happened before, three groups concern infected samples (light blue triangle), QC samples (red triangles), and healthy olive trees samples (blue dots and dark blue diamonds). A fourth group collects, in this case, healthy samples coming from dead branch trees (green ellipse, green squares). Also in this case, the grouping allows discriminating infected samples from QCs and healthy samples. The main difference resides in the QCs group location, which in positive ion mode PCA is slightly slid towards HP groups while, on the contrary, in negative ion mode PCA beneath the OP samples group. In order to exclude foliage desiccation progress as a variable in healthy-infected discrimination, we further introduced a desiccated set of samples; from data upcoming from the analysis we can assess



Figure 12.5.: *PCA plots for negative polarity experiments*

that features detected as putative markers are specific for OQDS disease and not implied in foliage desiccation process as seen in box and whiskers plots (figures 12.8 and 12.9).

For both PCA plots, the direction that best separates the HP and OP clusters is on the first principal component. Analysis in ESI+ and ESI- modes allowed for the identification of key molecules that were expressed in a significantly different manner between the groups of healthy and sick leaves. This work of molecular recognition is propaedeutic to potential biological markers identification for infection. In this view, several identified molecules could be considered as markers and used to develop a new analytical method based on a targeted metabolomic approach. This would pave the way to ensure early diagnosis of the infection. In this work, only those features that exhibited the most significant alterations between controls and diseased samples ($p < 0.00001$) were analysed in detail. Putative metabolite identification was based on accurate mass (within 10 ppm) and MS/MS data match against MS/MS spectra available on METLIN. When it was not present, we attempted a molecular recognition by the pattern of the experimental fragmentation spectrum available in online databases. Table

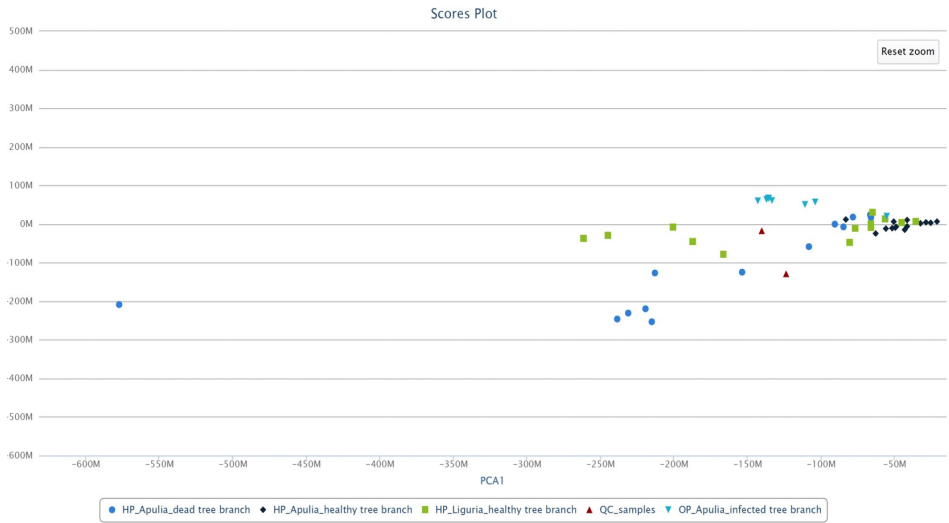


Figure 12.6.: *PCA plots for negative polarity experiments - zoom*

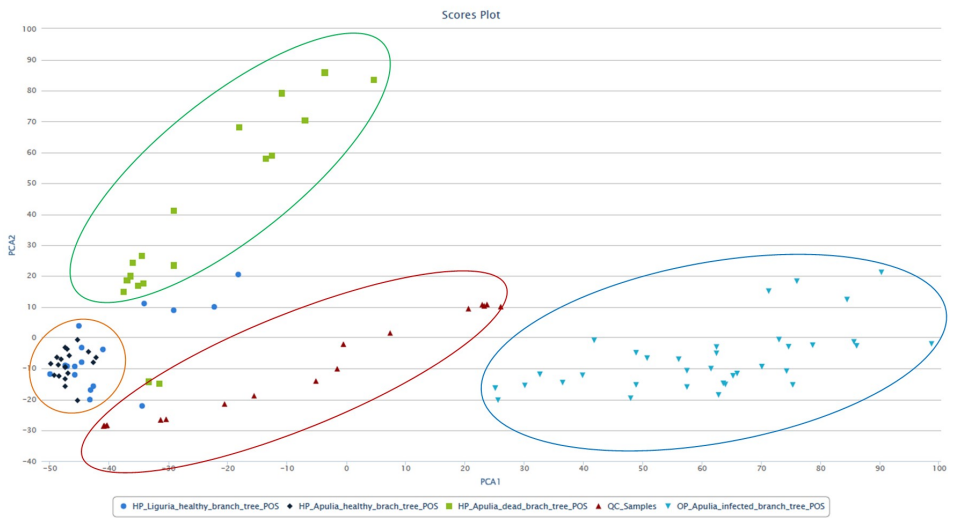


Figure 12.7.: *PCA plots for positive polarity experiments*

12.2 summarizes several features with high intensity and high significance identified through METLIN software and MS/MS spectra; in addition, for each compound annotation level is showed according to MSI (Metabolomic Standard Initiative) [8, 24].

The table 12.2 summarises several features with high intensity and high significance identified through METLIN software and MS/MS spectra. Furthermore, dopaxanthin, gibberellin O- β -D-glucoside, and homovanillic acid show high loading scores in PCA analysis (results obtained by statistical assay) so that they can be useful to discriminate HP and OP samples. As an example, gibberellin O- β -D-glucoside and ferulic acid have been identified as a stress compounds implicated in various environmental responses [22]. So, these molecules are an example of the validity of our global metabolomics method, and they could represent a valid biological marker of the infection. In table 12.2 other molecules can be proof of the validity of the method: quinic acid, vanillic acid, and ferulic acid. These phenolics are the products of secondary metabolism in plants, providing essential functions in the reproduction and the growth of the plants, acting as defence mechanisms against pathogens, parasites, and predators, as well as contributing to the colour of plants.

Considering such a significant difference in levels between HP and OP group, we can hypothesize that olive trees secrete these molecules as a defence mechanism. So, these analytes could be regarded as useful diagnostic markers of plant disease. Besides, other molecules are known to be highly expressed in olive leaves samples [6] have been detected by our global metabolomics analysis: oleanolic acid, oleoside, oleuropein glucoside, and quercetin. These analytes' profile does not show variation between OP and HP samples. Thus, they are not to be considered as early identification features, but they make their purpose as negative controls. As a matter of fact, these analytes could be used to confirm that the samples analysed are necessarily needed from the olive trees.

Identified molecules also confirmed by further MS/MS support data are showed also in their distribution via graphical box and whisker representation (figures 12.8 and 12.9); notably, variance found in QC mixes has a lower value than OP and HP's suggesting the excellent quality of data obtained.

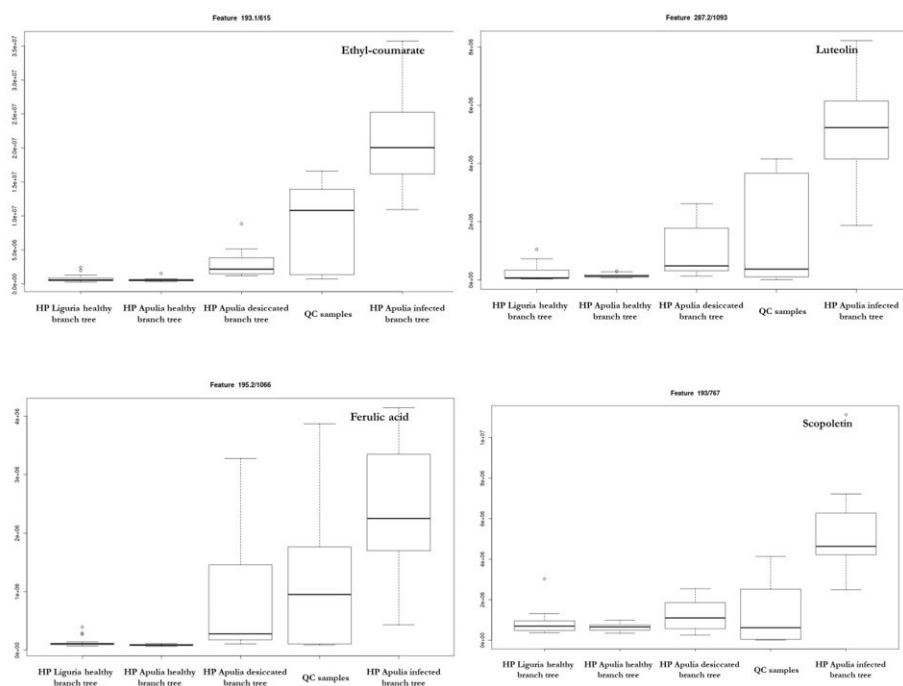


Figure 12.8.: *Box and whisker plots of MS/MS confirmed molecules*

In box and whisker plots reported in figures 12.8 and 12.9) five different groups are reported: the first three on the left are the ones related to the healthy samples; then the fourth group is referred the QC samples; finally, the last one on the right is the one correlated to the infected samples.

12.5. Conclusion and future prospects

An untargeted metabolomics approach was applied to olive leaves samples to understand the main differences between healthy plants and plants with OQDS like symptoms. To this end, we followed an extraction procedure with ethyl acetate and developed a simple, accurate high-resolution mass-based analysis method that could detect the broadest range of metabolites. Interestingly, several putative markers were identified with literature and mass database searches and confirmed with tandem MS experiments. The results of the multivariate analysis showed a clustering of the two pools of samples (HP vs. OP) based on two principal components (PC1 and PC2).

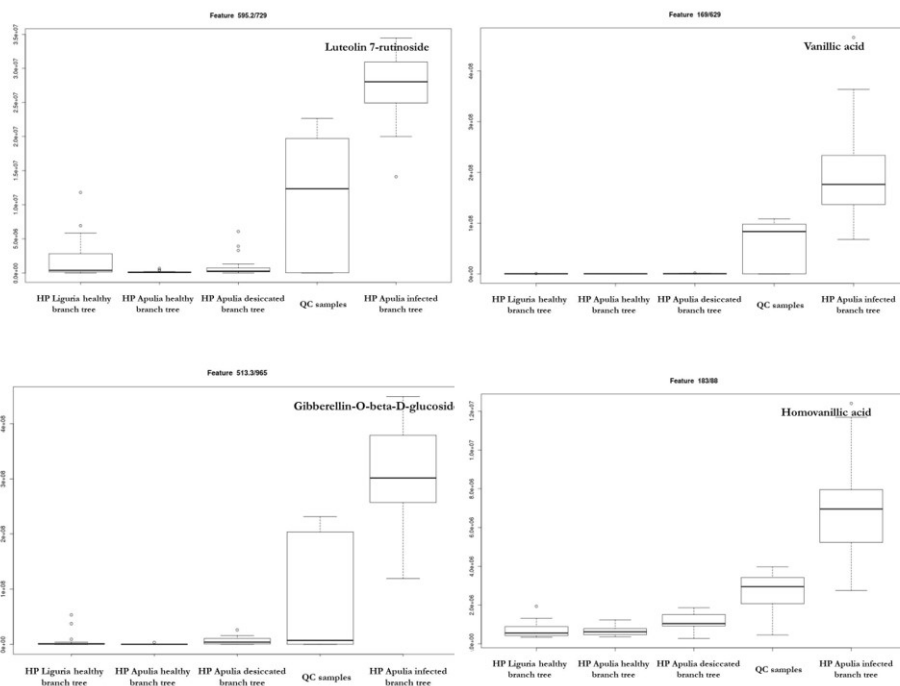


Figure 12.9.: Box and whisker plots of MS/MS confirmed molecules

Table 12.2.: Features identification and confirmation methods applied to achieve this task. CID: Compound Identification number; METLIN: Metabolite and Chemical Entity Database; MSI: Metabolomics Standard Initiative

Feature	PubChem CID	METLIN confirmation	MS/MS confirmation	High loading feature	MSI level
5,5'-Dihydroxy-pentamethoxyflavone 5"-glucoside	44259888	✓			2
Indole-3-acetyl-myo-inositol	440152	✓			2
Epicatechin 5-O-beta-D-glucopyranoside-3-benzoate	44257094	✓	✓		1
Cyanidin 3-(6"-malyglucoside)	44256741	✓	✓		1
Maximaisoflavone	21676205	✓			2
Quercetin-diacetylramnosyl-glucoside	44259307	✓			2
Apiumoside	131752524	✓			2
Ramontoside	14825498	✓			2
Aplysiatoxin	46173823	✓			2
p-Hydroxycoumaroyltryptophan	24891366	✓			2
Moschamindole	25233004	✓			2
Trehalose	7427	✓			2
Vanillic Acid	8468	✓	✓	✓	1
Apigenin-7-O-glucoside	5280704	✓	✓		1
Quinic Acid	6508	✓	✓		1
Luteolin	5280445	✓	✓	✓	1
Luteolin-7-rutinoside	44258082	✓	✓	✓	1
Ferulic Acid	445858	✓	✓	✓	1
Dopaxanthin	35438589	✓	✓	✓	2
Gibberellin O-β-D-glucoside	16019961	✓	✓	✓	1
Homovanillic Acid	1738	✓	✓	✓	1
Scopoletin	5280460	✓	✓	✓	1
Ethyl coumarate	676946	✓	✓	✓	1
Hypoxanthine	135398638	✓			2
3-Hydroxyisovalerylcarnitine	71464474	✓			2

Twenty-five different organic compounds high express in OP group were detected. Thanks to Metlin and MS/MS experiments, their exact molecular structure is unveiled, and they may be considered as markers for early diagnosis of OQDS. This would pave the way for a targeted and feasible analytical approach aimed to the detection of early infection state-related molecules for all the research and routine laboratories who cannot afford a global metabolomics instrumentation setup.

However, a truly comprehensive analysis of the plant metabolite pool is not easily feasible due to the large number of primary and secondary metabolites in any given plant species. Each analytical technology has advantages and limitations, and no one can cover the whole metabolome due to the chemical diversity of metabolites and their broad dynamic range in cellular abundance [27]. Consequently, different extraction techniques and combinations of analytical methods should be employed in attempts to achieve adequate metabolite coverage [13].

In the future, the primary goal is to obtain a more significant number of features by implementing diverse extraction methods and by merging results upcoming from MS to other techniques more oriented to identification and qualitative analysis like NMR spectroscopy. More in-depth and more accurate knowledge of olive metabolome and its infection-related differences can be provided by reiterating the analysis of different years in order to verify whether the data obtained by analysing samples belonging to a single vintage remain unchanged by analysing olive trees of different years or not.

References of Chapter 12

- [1] E.D. Beaulieu et al. “Characterization of a Diffusible Signaling Factor from *Xylella fastidiosa*”. In: *MBio - American Society for Microbiology* 4 (2013). DOI: 10.1128/mBio.00539-12 (cit. on p. 242).
- [2] B. Bextine et al. “Evaluation of Methods for Extracting *Xylella fastidiosa* DNA from the Glassy-Winged Sharpshooter”. In: *Journal of Economic Entomology* 97, 757-763 (2004). DOI: 10.1093/jee/97.3.757 (cit. on p. 240).
- [3] C. Carridi et al. “Isolation of a *Xylella fastidiosa* strain infecting olive and oleander in Apulia, Italy”. In: *Journal of Plant Pathology* 96, 425-429 (2014). DOI: 10.4454/JPP.V96I2.024 (cit. on p. 239).
- [4] A.V. Colnaghi Simionato et al. “Characterization of a putative *Xylella fastidiosa* diffusible signal factor by HRGC-EI-MS”. In: *Journal of Mass Spectrometry* 42, 1375-1381 (2007). DOI: 10.1002/jms.1325 (cit. on p. 242).
- [5] V.S. Da Silva et al. “Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains”. In: *BMC Genomics* 8, 4474 (2007). DOI: 10.1186/1461.2164.8.474 (cit. on p. 239).
- [6] L. Di Donna et al. “Secondary metabolites of *Olea europaea* leaves as markers for the discrimination of cultivars and cultivation zones by multivariate analysis”. In: *Food Chemistry* 121, 492-496 (2010). DOI: 10.1016/j.foodchem.2009.12.070 (cit. on p. 252).
- [7] A.M. D’Onghia et al. “*Xylella fastidiosa* & the Olive Quick Decline Syndrome (OQDS) A serious worldwide challenge for the safeguard of olive trees”. In: *Options Méditerranéennes : Série A. Séminaires Méditerranéens* 121, 172 (2017). DOI: <http://om.ciheam.org/om/pdf/a121/a121.pdf> (cit. on p. 239).
- [8] O. Fiehn et al. “The metabolomics standards initiative (MSI)”. In: *Metabolomics* 3, 175-178 (2007) (cit. on p. 252).

- [9] R. Hernandez-Martinez et al. “Differentiation of Strains of *Xylella fastidiosa* Infecting Grape, Almonds, and Oleander Using a Multi-primer PCR Assay”. In: *Plant Disease* 90, 1382-1388 (2006). DOI: 10.1094/PD-90-1382 (cit. on p. 240).
- [10] D.L. Hopkins et al. “*Xylella fastidiosa*: Cause of Pierce’s disease of grapevine and other emergent diseases”. In: *Plant Disease* 86, 1056-1066 (2002). DOI: 10.1094/PDIS.2002.86.10.1056 (cit. on p. 240).
- [11] M. Ionescu et al. “Diffusible Signal Factor (DSF) Synthase RpfF of *Xylella fastidiosa* Is a Multifunction Protein Also Required for Response to DSF”. In: *Journal of Bacteriology* 195, 5273-5284 (2013). DOI: 10.1128/JB.00713-13 (cit. on p. 242).
- [12] J.D. Janse et al. “*Xylella fastidiosa*: its biology, diagnosis, control and risks”. In: *Journal of Plant Pathology* 92, S35-48 (2010) (cit. on p. 240).
- [13] T.F. Jorge et al. “Mass spectrometry-based plant metabolomics: Metabolite responses to abiotic stress”. In: *Mass Spectrometry Review* 35, 620-649 (2016). DOI: 10.1002/mas.21449 (cit. on pp. 241, 255).
- [14] M.A. Kamleh et al. “Optimizing the Use of Quality Control Samples for Signal Drift Correction in Large-Scale Urine Metabolic Profiling Studies”. In: *Analytical Chemistry* 84, 2670-2677 (2012). DOI: 10.1021/ac202733q (cit. on p. 247).
- [15] S. Lindow et al. “Production of *Xylella fastidiosa* Diffusible Signal Factor in Transgenic Grape Causes Pathogen Confusion and Reduction in Severity of Pierce’s Disease”. In: *Molecular Plant-Microbe Interactions* 27, 244-254 (2014). DOI: 10.1094/MPMI-07-13-0197-FI (cit. on p. 242).
- [16] G. Loconsole et al. “Detection of *Xylella Fastidiosa* in olive trees by molecular and serological methods”. In: *Journal of Plant Pathology* 96, 7-14 (2014). DOI: 10.4454/JPP.V96I1.041 (cit. on p. 240).
- [17] L.L.R. Marques et al. “Characterization of Biofilm Formation by *Xylella fastidiosa* In Vitro”. In: *Plant Disease* 86, 633-638 (2002). DOI: 10.1094/PDIS.2002.86.6.633 (cit. on p. 239).

- [18] G.P. Martelli et al. “The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency”. In: *European Journal of Plant Pathology* 144, 235-243 (2016). DOI: 10.1007/s10658-015-0784-7 (cit. on p. 240).
- [19] F. Nigro et al. “Fungal species associated with a severe decline of olive in southern Italy”. In: *Journal of Plant Pathology* 95(3), 668 (2013). DOI: 10.4454/JPP.V95I3.034 (cit. on pp. 239, 240).
- [20] G.J. Patti et al. “A View from Above: Cloud Plots to Visualize Global Metabolomic Data”. In: *Analytical Chemistry* 85, 798-804 (2013). DOI: 10.1021/ac3029745 (cit. on p. 241).
- [21] G.J. Patti et al. “Metabolomics: the apogee of the omics trilogy”. In: *Nature Reviews Molecular Cell Biology* 13, 263-269 (2012). DOI: 10.1038/nrm3314 (cit. on p. 241).
- [22] A. Piotrowska et al. “Conjugates of abscisic acid, brassinosteroids, ethylene, gibberellins, and jasmonates”. In: *Phytochemistry* 72, 2097–2112 (2011). DOI: 10.1016/j.phytochem.2011.08.012 (cit. on p. 252).
- [23] D. Saigusa et al. “Establishment of Protocols for Global Metabolomics by LC-MS for Biomarker Discovery”. In: *Plos One* (2016). DOI: 10.1371/journal.pone.0160555 (cit. on p. 247).
- [24] L.W. Summer et al. “Proposed Minimum Reporting Standards for Chemical Analysis Chemical Analysis Working Group (CAWG) - Metabolomics Standards Initiative (MSI)”. In: *Metabolomics* 3(3), 211-221 (2007). DOI: 10.1007/s11306-007-0082-2 (cit. on p. 252).
- [25] R. Tautenhahn et al. “An accelerated workflow for untargeted metabolomics using the METLIN database”. In: *Nature Biotechnology* 30, 826-828 (2012). DOI: 10.1038/nbt.2348 (cit. on p. 248).
- [26] R. Tautenhahn et al. “XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data”. In: *Analytical Chemistry* 84, 5035-5039 (2012). DOI: 10.1021/ac300698c (cit. on p. 245).
- [27] S. Wang et al. “The Structure and Function of Major Plant Metabolite Modifications”. In: *Molecular Plant* 12, 899-91 (2019). DOI: 10.1016/j.molp.2019.06.001 (cit. on pp. 241, 255).

- [28] XCMS. *The XCMS Online*. 2020. URL: https://xcmsonline.scripps.edu/landing_page.php?pgcontent=mainPage (cit. on p. 245).
- [29] J. Xia et al. “MetaboAnalyst 3.0 - making metabolomics more meaningful”. In: *Nucleic Acids Research* 43, W251-W257 (2012). DOI: 10.1093/nar/gkv380 (cit. on p. 248).

Final remarks, conclusions and future research

13

This PhD thesis describes the development of robust and valid methods for global and targeted metabolomics analyses via multiple LC-MS platforms. In comparison to other omics techniques, metabolomics is the one who provides direct monitoring of phenotype, which makes it a scientific research field with borderless potential.

In the first chapters, I had laid the foundations of the metabolomics world to prepare the reader for analytical results reported from chapter five to chapter twelve. In particular, I described Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR) spectroscopy-based metabolomics platforms. These analytical techniques are the only ones able to supply comprehensive and quantitative analysis of different types of metabolites among various matrices such as biological or natural samples.

In the results' chapters, different methods and concepts were developed and deeply discussed. High-resolution MS methods focus on improved normalization and data processing for untargeted analyses. For targeted metabolomics, such as BPA and BPS detection in urine samples or quorum sensing molecules, the driving force of these projects was the improvement of extractions and detection of markers generally present in low concentration among matrices of interest.

So, in the light of the content of these chapters, I will expose in the next section the main contributions my thesis can bring to the scientific community, together with suggestions and pointers to future research.

Contributions of the thesis

The main contributions of the thesis are:

1. The establishment of validated protocols for metabolomics analysis utilizing as High-Resolution Mass Spectrometry HRMS as Triple Quadrupole Mass Spectrometer TQMS. Once again, I need to remark how metabolomics should not be seen as a stand-alone analytical technique, but, on the contrary, results obtained should be incorporated

in a larger biochemical context investigated through complementary techniques.

2. The establishment of effective workflows as for routine Lab-analysis as for new and non-previously analysed samples.
3. For physiology and pathology studies, I demonstrated how MS-based metabolomics studies are the most useful techniques for characterization of new potential biomarkers never monitored before. Furthermore, samples of interest, such as urine or plasma, are usually obtained in a minimally invasive manner.

The last consideration is related to the title of this thesis. The reader may ask the question of whether quali- and quantification in targeted and untargeted mass spectrometry experiments were significantly improved by this work, and are these methods suitable for metabolomics research? The answer is complicated as the developed methods indeed improved the quantification of specific metabolites or the detection of unexpected ones and allowed them to identify and deal with the sources of analytical variation. Still, more experiments/analysis and validation setup are required.

Future research

There is no doubt that several lessons in terms of research techniques have been learned during the conduct of the research in this thesis. As already highlighted in the discussions and conclusions of individual chapters, many future experiments and studies have suggested future works.

As first piece of advice for future improvements, I recommend incrementing numbers of samples. It is a good practice to validate a hypothesis and to check if the trend over an increased population is still valid or not. For example, in berries analyses (chapter 9), by increasing the number of berries belonging to the same species, it could corroborate the thesis according to which certain polyphenols are specific only for some berries.

Furthermore, some of the experiments I discuss in previous chapters (e.g. chapters 5, 6, and 8) are currently work in progress. They require further measurements, validation protocols appliance, and application of statistical analyses to have a clearer view.

For some projects, I was involved it would be fascinating to monitor other kinds of metabolites and their changes among different samples. For example, in BPA and BPS analysis (chapter 5), possible future research is to conduct an untargeted metabolomics analysis upon breast-milk or urine samples to monitor metabolites concentration (e.g., oxytocin, sugars, sphingolipids, etc.) to check if an increase in levels of BPA and BPS molecules is related to an increase or decrease of other analytes. In that case, these metabolites could, therefore, become markers of absence or presence of BPS/BPS (or plasticizers) in body fluids.

As final advice, I would like to give to all those will read or use my thesis is to do science, using the brain in a counterintuitive but correct way. In fact, in the scientific field, it is fundamental always to check whether a hypothesis is false. Often our brain operates with the so-called confirmation bias (search for a confirmation), trying to demonstrate the phenomenon just discovered. Instead, every scientist should prove first of all that his theory is incorrect. Sometimes I probably haven't applied this way of thinking, and I apologize to my readers. I hope whoever succeeds me can correct where I went wrong.

Appendix A - Ionization Sources in Mass Spectrometry

A

Whatever their configuration, mass spectrometers always comprise two basic components: an ionization source and a mass analyzer. The fundamental principle of mass spectrometers is that molecules must be charged to be detected (mass-to-charge ratio, m/z). Once formed, ions are electrostatically guided from the ion source into the mass analyzer where they are separated according to m/z and finally detected.

The creation of ions is therefore essential for this technique. To fulfil this task, different ionization techniques could be applied. Generally speaking, any mass analyzer can be coupled to any ionization source, and both are critical to a properly executed mass spec experiment. As a result, both must be considered when making a purchasing decision.

In table A.1 the main ionization sources related to this work, their acronyms, and method of ionization are reported. Subsequently, for each of the cited ionization sources a brief description will be furnished.

EI - Electron ionization

While ESI, APCI, and APPI are used strictly with LC, EI, or electron impact ionization, is a gas chromatographic (GC) ionization technique. Electron ionisation involves the production of electrons in the mass spectrometer ion source with energies sufficient (70 eV generally) to induce ion formation in neutral gas-phase molecules of sample, through electron repulsion. It

Table A.1.: *Summary of ionization sources, their acronyms, and the method of ionization*

Ionization source	Acronym	Method of ionization
Electron ionization	EI	Electron beam and electron transfer
Chemical ionization	CI	Proton transfer
Electrospray ionization	ESI	Evaporation of charged droplets
Atmospheric pressure chemical ionization	APCI	Corona discharge and proton transfer
Matrix-assisted laser desorption ionization	MALDI	Photon absorption and proton transfer
Desorption electrospray ionization	DESI	Combination of desorption and evaporation of charged droplets
Direct analysis in real time	DART	Glow discharge and proton transfer

is the basis for one of the most efficient mass spectrometry methods for identifying a given organic compound. This technique is considered to be a hard ionization method (high fragmentation) because of the high-energy electrons used to generate ions of interest.

This leads to extensive fragmentation, which contributes to the structural determination of unknown compounds. EI is useful for organic compounds with molecular weights below 600.

CI - Chemical ionization

CI, or chemical ionization, is also used exclusively for GC applications. Compared to EI, CI could be defined as a "soft" ionization technique that produces ions with little excess energy. As a result, less fragmentation is observed in the mass spectrum. Since this increases the abundance of the molecular ion, the technique is complimentary to 70 eV EI.

In Chemical Ionization the source is enclosed in a small cell with openings for the electron beam, the reagent gas and the sample. The reagent gas is added to this cell at approximately 10 Pa (0.1 torr) pressure and it is ionized with an electron beam to produce a cloud of ions. When analyte molecules are introduced into this cloud of ions, the reagent gas ions donate a proton to the analyte molecule and produce MH^+ ions. The energetics of the proton transfer is controlled by using different reagent gases: e.g., ammonia, methane, or isobutane.

CI has the advantage of being somewhat less harsh than EI. Therefore, CI is very useful for molecular mass determination thanks to the intrinsic stability of adduct ions.

ESI - Electrospray ionization

Electrospray Ionization (ESI) is linked to the name of John Fenn, who won the 2002 chemistry Nobel for his discovery of this new soft ionization method. ESI is a liquid phase process that generates a fine mist of charged droplets. The sample inlet typically is coupled to an in-line liquid chromatography (LC) system. When the mobile phase (liquid) exists from the

LC module, it passes into the sprayer, and it becomes subjected to a "voltage offset" (in the range 2000 V - 3000 V). This produces a high electrical charge on the surface of the drops, which then pass into a heated environment full of dry nitrogen gas. As the droplets evaporate, the ion density increases and becomes so high that charge transfers to the molecule, and the molecular ion gets lifted out.

As for APCI, ESI works in both positive and negative ion modes, and can produce ions with a wide range of charge states (e.g., +1, +2, +3, ...), making it amenable to detection and study of high molecular weight species.

APCI - Atmospheric pressure chemical ionization

APCI, or atmospheric pressure chemical ionization, is an ESI variant that uses a "corona discharge"—basically, an electrified needle—to induce ionization of the solvent, which in turn reacts with the sample molecules to induce a chemical reaction resulting in an ionized sample molecule. All the reaction involved in the ionization process takes place at atmospheric pressure.

APCI source usually consists of three parts: a nebulizer probe, an ionization region with a corona discharge needle, and an ion-transfer region under intermediate pressure. The analyte in solution is introduced from a direct inlet probe or a liquid chromatography (LC) eluate into a pneumatic nebulizer with a flow rate 0.2–2.0 mL/min. In the heated nebulizer, the analyte coaxially flows with nebulizer N₂ gas to produce a mist of fine droplets. By the combination effects of heat and gas flow, the emerged mist is converted into a gas stream. Once the gas stream arrives in the ionization region under atmospheric pressure, molecules are ionized at corona discharge which is 2 to 3 kV potential different to the exit counter-electrode.

APCI tends to favour relatively small, neutral, or hydrophobic compounds, such as steroids, lipids, and non-polar drugs, typically imparting a charge of +1.

MALDI - Matrix-assisted laser desorption ionization

The Matrix-assisted laser desorption ionization, simply known as MALDI, is a non-fragmenting ionization technique that uses laser energy to ionize molecules off a plate target. In contrast to the previously cited ionization techniques, MALDI requires a different protocol during sample preparation. Sample must be mixed with an organic matrix material (such as alpha-cyano-4-hydroxycinnamic acid), spotted onto the plate, and dried.

This sample is then irradiated with the laser source; upon irradiation, it absorbs laser energy and heats up, essentially transferring its energy to the molecule of interest and producing ionized molecules capable of being detected. The vast majority of ions produced in this manner have a charge of +1: as MALDI pioneer and developer Michael Karas used to say, "in MALDI, the $[M+H]^+$ ions are the lucky survivors." So, unlike ESI or APCI, MALDI approach does not work in both ion modes polarities.

DESI and DART - Desorption electrospray ionization and Direct analysis in real time

Both techniques differ slightly from the previous ionization sources described. In fact, these sources belong to the "ambient mass spectrometry" field, a subdiscipline of mass spectrometry that enables direct, high-throughput, surface analysis of native samples. DESI and DART are complementary techniques and could be seen as the other side of the coin of ESI and APCI respectively. DART, for example, makes use of ionization mechanisms that predominantly follow atmospheric-pressure chemical ionization (APCI) pathways, but in an open-air format. DESI, on the other hand, makes use of desorption mechanisms that involve a continuous solid-liquid extraction process, while capitalizing on the known ionization mechanisms of ESI.

DESI could be seen as the synthesis of electrospray (ESI) and desorption (DI) ionization approaches. Ionization takes place by directing an electrically charged mist to the sample surface that is located a few millimetres from the probe. The electrospray mist is pneumatically directed at the

Table A.2.: *Comparison between DART and DESI sources*

	DART	DESI
Detection of large MW. analytes	Requires extra heating, may induce unwanted fragmentation	Readily achievable, even for large proteins
Spectral background	Simpler	Complex
Robustness towards ion-source geometric configuration	Simpler, standardized geometry	Sensitivity and spatial resolution depend a great experimental variables
Sample throughput	Very high	Very high
Sensitivity	Depends on different variables	Depends on different variables

sample where subsequent splashed droplets carry desorbed, ionized analytes. After ionization, the ions travel through air into the atmospheric pressure interface which is connected to the mass spectrometer. The ionization mechanisms for low molecular weight molecules (the most interesting ones from metabolomics point of view) show different possibilities: charge transfer between a solvent ion and an analyte on the surface; charge transfer between a gas phase ion and analyte on the surface; charge transfer between a gas phase ion and a gas phase analyte molecule.

In contrast to the liquid spray used by DESI, DART is an ion source that produces excited-state species from gases such as helium, argon, or nitrogen that ionize atmospheric molecules. So, DART ion source contains a dry stream containing excited state species. Analytes with low ionization energy may be ionized directly. The DART ionization process can produce positive or negative ions depending on the potential applied to the exit electrode. As for DESI, also for DART approach it does not need a specific sample preparation, so it can be used for the analysis of solid, liquid and gaseous samples in their native state.

A quick comparison between these ambient mass spectrometry ionizations sources is reported in table A.2.

Appendix A - Classifications of polyphenolic and flavonoids molecules

B

Phenol compounds can be divided in four major classes of polyphenols: flavonoids, lignans, phenolic acids and stilbenes. The figure B.1 reports a systematic polyphenolic molecules classification. The small berries treated in chapter 9 are rich in flavonoids and some of its sub-classes such as flavonols or anthocyanidins.

The scheme figured in figure B.2 reports the basic flavonoid skeleton of flavonoids and the main sub-classes belonging from this group.

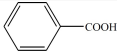
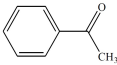
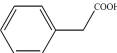
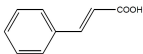
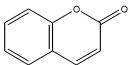
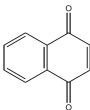
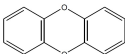
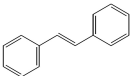
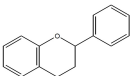
Skeleton	Classification name	Basic chemical structure
C ₆ -C ₁	Phenolic acids	
C ₆ -C ₂	Acetophenones	
C ₆ -C ₃	Phenylacetic acid	
C ₆ -C ₃	Hydroxycinnamic acids	
C ₆ -C ₄	Coumarins	
C ₆ -C ₄	Naphthoquinones	
C ₆ -C ₁ -C ₆	Xanthenes	
C ₆ -C ₂ -C ₆	Stilbenes	
C ₆ -C ₃ -C ₆	Flavonoids	
---	Lignin	---

Figure B.1.: *Systematic polyphenolic molecules classification*

As clarify by our analyses to, flavonoids are widely distributed in plants. They are the actors responsible for plant pigments or flower coloration.

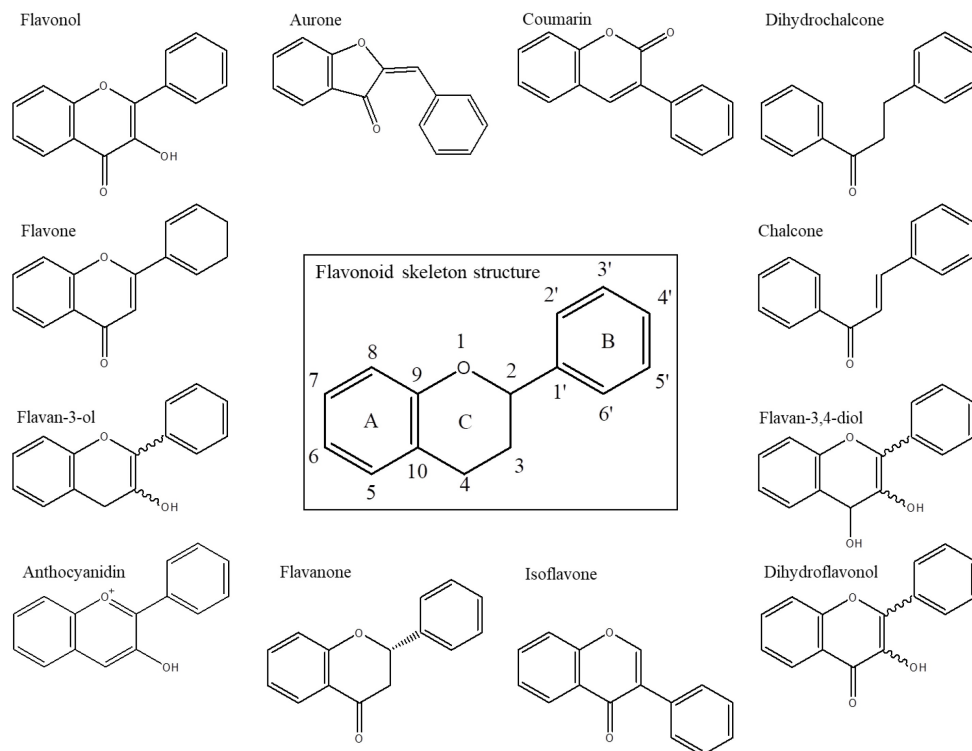


Figure B.2.: *Main sub-classes structures of flavonoids.*

Furthermore, flavonoids are not only linked to aesthetic aspects: in fact, they are involved in UV filtration and symbiotic nitrogen fixation. Not considering berries, other important sources of polyphenols are black tea, citrus, wine, etc.

Appendix C - Classifications of lipids classes

C

Lipids are the major constituents of animal and plant cells. Their common characteristic is the solubility in non-polar solvents, but not in water. The general structure of lipids molecules is a polar hydrophilic part (the head portion) and a non-polar hydrophobic part, which all belong to amphiphilic molecules. But some of the lipids like triacylglycerols in fats or oils or cholesterol are constituted above all by non-polar components. The Lipid MAPS consortium [1] proposes a lipid classification system, where all the lipids molecules are classified in eight different categories: fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterol lipids, prenol lipids, saccharolipids, and polyketides.

To better understand how the consortium has proposed the classification, it is sufficient to observe the cellular lipid molecular species. As a matter of facts, most of them is just a combination of different “building blocks”. If, for example, we look at glycerophospholipids we can distinguish a central portion (the glycerol group) linked to three different building blocks.

Similarly, if the species possess a sphingoid-based instead of glycerol they are called sphingoid-based lipids. Two general schemes of the different possible lipids molecules can be founded in figures C.1 and C.2.

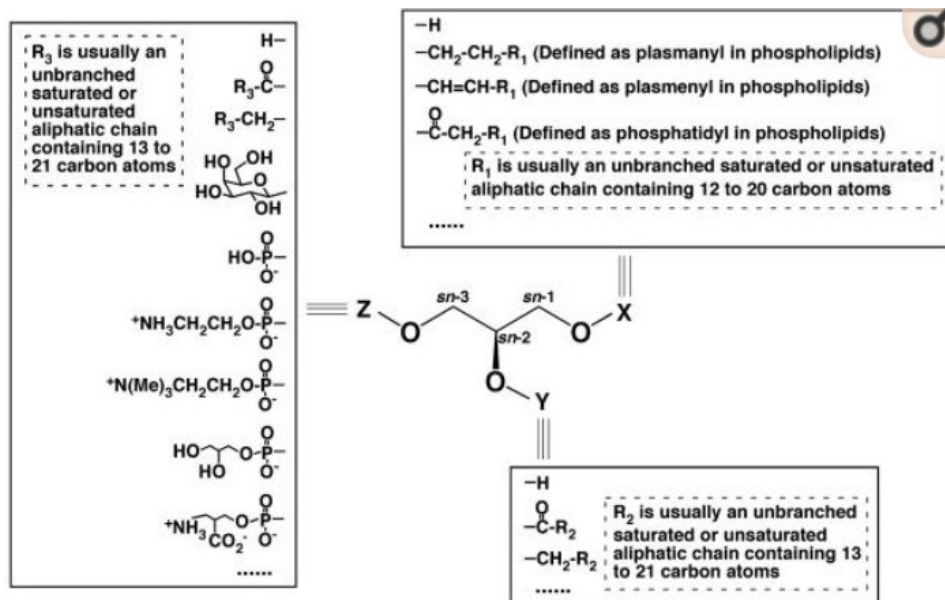


Figure C.1.: General structure of glycerol-based lipids. Three blocks are connected to the hydroxy groups of glycerol [2]

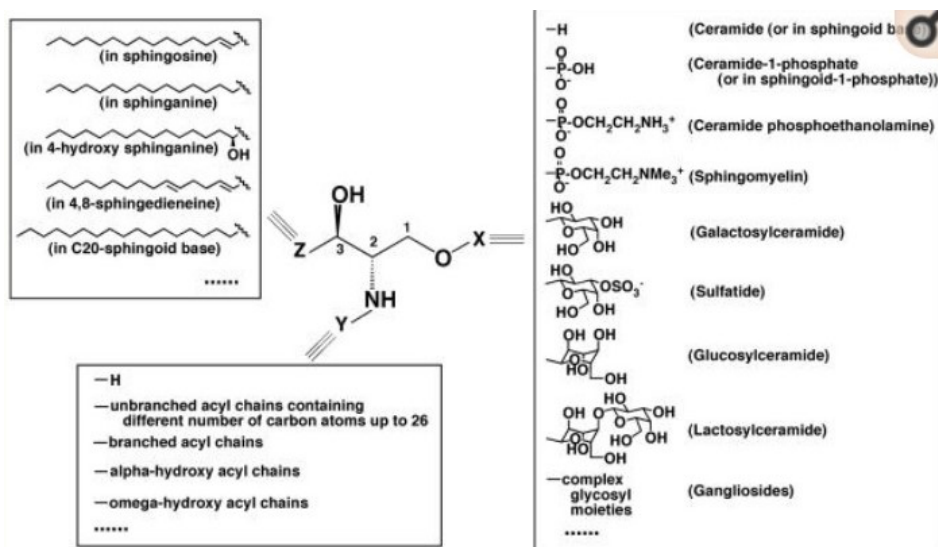


Figure C.2.: General structure of sphingoid-based lipids. The building block X represents a different polar moiety, the block Y represents fatty acyl chains, and the building block Z represents the aliphatic chains in all of possible sphingoid bases [2]

Appendix D - Fatty acids and related masses list

D

Fatty acids play a key role in metabolism: they are the major constituents of membranes, represent the fuel of metabolic process, and act as a barrier for the body against heat and electricity. Fatty acids are carbon chains with a methyl group at one end of the molecule (designated omega, ω) and α carboxyl group at the other end. The carbon atom next to the carboxyl group is called the α carbon, and the subsequent one the β carbon. The most common fatty acids contain from ten up to twenty-two carbon atoms. Furthermore, fatty acids can be also be saturated (no double bonds are present) or unsaturated (one or multiple double bonds are present all along the carbon chain). This variability in the composition of fatty acids becomes even greater when, by combining three fatty acids, Triglycerides (TAGs) are generated. To help the reader of this thesis in establishing the m/z values of any triglyceride of interest, table D.1 has been created. The table lists the individual fatty acids ordered from the smallest (Butanoic acid) to the heaviest (Hexatriacontadienoic acid). In addition, are reported too: the common fatty acid name, the number of carbon atoms, the number of double bonds present in each fatty acid, the fatty acid m/z in high resolution, the methyl ester m/z in high resolution, the RCO^+ m/z in high resolution, and the RCOO^- m/z in high resolution. Combining the different m/z values any TAG m/z compound can be obtained.

For example, the TAG P-O-Ln (Palmitic acid - Oleic acid - α -Linolenic acid) has a $[\text{M}+\text{H}]^+$ of 855.7454. This value can be obtained by this sum: 256.2402 (m/z of palmitic acid) + 282.2559 (m/z of oleic acid) + 278.2246 (m/z of α -linolenic acid) + 92.0468 (m/z of glycerol) + 1.0079 (m/z of H adduct) - 54.0300 (m/z of three molecules of water loss when the hydroxyl groups of the glycerol join the carboxyl groups of the fatty acid to form ester bonds). If, on the contrary, the analyst is focused on the $[\text{M}-\text{H}]^-$ value, instead of sum + 1.0079 just subtracted it from the total.

Table D.1.: List of fatty acids and related masses. The table belongs to [1]. In the table C: number of carbon atoms; U: number of double bonds; EC: Equivalent Carbon Number

Systematic Fatty Acid Name	Common Fatty Acid Name	Abbreviation	C	:	U	EC	Fatty Acid Mass	Methyl Ester	[RCO]+	[RCOO]-
Butanoic acid	Butyric acid	Bu	4	:	0	4	88.0524	102.0681	71.0491	87.0452
Pentanoic acid	Valeric acid	V	5	:	0	5	102.0681	116.0837	85.0653	101.0603
Hexanoic acid	Caproic acid	Co	6	:	0	6	116.0837	130.0994	99.0810	115.0759
Heptanoic acid	Enanthic acid	En	7	:	0	7	130.0994	144.1150	113.0966	129.0916
Octanoic acid	Caprylic acid	Cy	8	:	0	8	144.1150	158.1307	127.1123	143.1072
Nonanoic acid	Pelargonic acid		9	:	0	9	158.1307	172.1463	141.1279	157.1229
Decanoic acid	Capric acid	Ca	10	:	0	10	172.1463	186.1620	155.1436	171.1385
Hendecanoic acid	Undecylic acid		11	:	0	11	186.1620	200.1776	169.1592	185.1542
Dodecanoic acid	Lauric acid	La	12	:	0	12	200.1776	214.1933	183.1749	199.1698
Tridecanoic acid	Tridecylic acid		13	:	0	13	214.1933	228.2089	197.1905	213.1855
Tetradecanoic acid	Myristic acid	M	14	:	0	14	228.2089	242.2246	211.2062	227.2011
Pentadecanoic acid	Pentadecylic acid		15	:	0	15	242.2246	256.2402	225.2218	241.2168
Hexadecanoic acid	Palmitic acid	P	16	:	0	16	256.2402	270.2559	239.2375	255.2324
c-9-Hexadecenoic acid	Palmitoleic acid (n-7)	Po	16	:	1	14	254.2246	268.2402	237.2218	253.2168
t-9-Hexadecenoic acid	Palmitelaic acid (n-7)	Pe	16	:	1	14	254.2246	268.2402	237.2218	253.2168
Heptadecanoic acid	Margaric acid		17	:	0	17	270.2559	284.2715	253.2531	269.2481
Octadecanoic acid	Stearic acid	S	18	:	0	18	284.2715	298.2872	267.2688	283.2637
c-9-Octadecenoic acid	Oleic acid (n-9)	O	18	:	1	16	282.2559	296.2715	265.2531	281.2481
t-9-Octadecenoic acid	Elaidic acid (n-9)	El	18	:	1	16	282.2559	296.2715	265.2531	281.2481
c-9,12-Octadecadienoic acid	Linoleic acid (w-6)	L	18	:	2	14	280.2402	294.2559	263.2375	279.2324
c-9,12,15-Octadecatrienoic acid	α -Linolenic acid (w-3)	Ln	18	:	3	12	278.2246	292.2402	261.2218	277.2168
c-6,9,12-Octadecatrienoic acid	γ -Linolenic acid (w-6) = GLA	g-Ln	18	:	3	12	278.2246	292.2402	261.2218	277.2168
9c,11t,13t-Octadecatrienoic acid	α -Eleostearic acid	Eo	18	:	3	12	278.2246	292.2402	261.2218	277.2168
9t,11t,13t-Octadecatrienoic acid	β -Eleostearic acid		18	:	3	12	278.2246	292.2402	261.2218	277.2168
8c,10t,12c-Octadecatrienoic acid	Jacaric acid	Ja	18	:	3	12	278.2246	292.2402	261.2218	277.2168
9t,11t,13c-Octadecatrienoic acid	Catalpic acid	Ct	18	:	3	12	278.2246	292.2402	261.2218	277.2168
c-6,9,12,15-Octadecatetraenoic acid	Stearidonic acid (w-3)	St	18	:	4	10	276.2089	290.2246	259.2062	275.2011
Nonadecanoic acid	Nonadecylic acid		19	:	0	19	298.2872	312.3028	281.2844	297.2794
Eicosanoic acid	Arachidic acid	A	20	:	0	20	312.3028	326.3185	295.3001	311.2950
c-9-Eicosenoic acid	Gadoleic acid (n-11)	G	20	:	1	18	310.2872	324.3028	293.2844	309.2794

Continues in next page

Continues from previous page

Systematic Fatty Acid Name	Common Fatty Acid Name	Abbreviation	C	:	U	EC	Fatty Acid Mass	Methyl Ester	[RCO]+	[RCOO]-
c-11-Eicosenoic acid	Gondoic acid (n-9)	Go	20	:	1	18	310.2872	324.3028	293.2844	309.2794
c-11,14-Eicosadienoic acid	DihomoLinoleic acid (w-6)	C20:2	20	:	2	16	308.2715	322.2872	291.2688	307.2637
c-11,14,17-Eicosatrienoic acid	Bis-homo- α -Linolenic acid (w-3), ETA	C20:3	20	:	3	14	306.2559	320.2715	289.2531	305.2481
c-8,11,14-Eicosatrienoic acid	Bis-homo- γ -Linolenic acid (w-6)	C20:3	20	:	3	14	306.2559	320.2715	289.2531	305.2481
c-5,8,11-Eicosatrienoic acid	Mead Acid	C20:3	20	:	3	14	306.2559	320.2715	289.2531	305.2481
c-5,8,11,14-Eicosatetraenoic acid	Arachidonic acid (w-6)	Ao	20	:	4	12	304.2402	318.2559	287.2375	303.2324
c-8,11,14,17-Eicosatetraenoic acid	Eicosatetraenoic acid (w-3)	C20:4	20	:	4	12	304.2402	318.2559	287.2375	303.2324
c-5,8,11,14,17-EicosaPentaenoic acid	EPA (w-3)	Ep	20	:	5	10	302.2246	316.2402	285.2218	301.2168
Heneicosanoic acid	Heneicosylic acid		21	:	0	21	326.3185	340.3341	309.3157	325.3107
Docosanoic acid	Behenic acid	B	22	:	0	22	340.3341	354.3498	323.3314	339.3263
c-13-Docosenoic acid	Erucic acid (n-9)	E	22	:	1	20	338.3185	352.3341	321.3157	337.3107
c-13,16-Docosadienoic acid	Docosadienoic acid (w-6)	C22:2	22	:	2	18	336.3028	350.3185	319.3001	335.2950
c-5,13,16-Docosatrienoic acid	Eranthic acid (w-6)	C22:3	22	:	3	16	334.2872	348.3028	317.2844	333.2794
c-7,10,13,16-Docosatetraenoic acid	Adrenic acid (w-6)	C22:4	22	:	4	14	332.2715	346.2872	315.2688	331.2637
c-7,10,13,16,19-DocosaPentaenoic acid	DPA (w-3)	Dp	22	:	5	12	330.2559	344.2715	313.2531	329.2481
c-4,7,10,13,16,19-DocosaHexaenoic acid	DHA (w-3)	Dh	22	:	6	10	328.2402	342.2559	311.2375	327.2324
Tricosanoic acid	Tricosylic acid	C23:0	23	:	0	23	354.3498	368.3654	337.3470	353.3420
Tetracosanoic acid	Lignoceric acid	Lg	24	:	0	24	368.3654	382.3811	351.3627	367.3576
c-15-Tetracosenoic acid	Nervonic acid (n-9)	N	24	:	1	22	366.3498	380.3654	349.3470	365.3420
c-15,18-Tetracosadienoic acid	Tetracosadienoic acid (w-6)	C24:2	24	:	2	20	364.3341	378.3498	347.3314	363.3263
c-12,15,18-Tetracosatrienoic acid	Tetracosatrienylic acid	C24:3	24	:	3	18	362.3185	376.3341	345.3157	361.3107
c-9,12,15,18-Tetracosatetraenoic acid	Tetracosatetraenylic acid	C24:4	24	:	4	16	360.3028	374.3185	343.3001	359.2950
c-6,9,12,15,18-Tetracosapentaenoic acid (or 21)	Tetracosapentaenylic acid	C24:5	24	:	5	14	358.2872	372.3028	341.2844	357.2794
c-6,9,12,15,18,21-Tetracosahexaenoic acid	Tetracosahexaenylic acid (w-3)	C24:6	24	:	6	12	356.2715	370.2872	339.2688	355.2637
Pentacosanoic acid	Hyenic acid	C25:0	25	:	0	25	382.3811	396.3967	365.3783	381.3733
Hexacosanoic acid	Cerotic acid	Ce	26	:	0	26	396.3967	410.4124	379.3940	395.3889
c-17-Hexacosenoic acid	Ximenic acid	Xi	26	:	1	24	394.3811	408.3967	377.3783	393.3733
Hexacosadienoic acid	Hexacosadienylic acid	C26:2	26	:	2	22	392.3654	406.3811	375.3627	391.3576
Hexacosatrienoic acid	Hexacosatrienylic acid	C26:3	26	:	3	20	390.3498	404.3654	373.3470	389.3420
Hexacosatetraenoic acid	Hexacosatetraenylic acid	C26:4	26	:	4	18	388.3341	402.3498	371.3314	387.3263
Hexacosapentaenoic acid	Hexacosapentaenylic acid	C26:5	26	:	5	16	386.3185	400.3341	369.3157	385.3107
Hexacosahexaenoic acid	Hexacosahexaenylic acid	C26:6	26	:	6	14	384.3028	398.3185	367.3001	383.2950
Heptacosanoic acid	Carboceric acid	C27:0	27	:	0	27	410.4124	424.4280	393.4096	409.4046

Continues in next page

Continues from previous page

Systematic Fatty Acid Name	Common Fatty Acid Name	Abbreviation	C	:	U	EC	Fatty Acid Mass	Methyl Ester	[RCO]+	[RCOO]-
Octacosanoic acid	Montanic acid	Mo	28	:	0	28	424.4280	438.4437	407.4253	423.4202
Octacosenoic acid	Octacosenylic acid	C28:1	28	:	1	26	422.4124	436.4280	405.4096	421.4046
Octacosadienoic acid	-	C28:2	28	:	2	24	420.3967	434.4124	403.3940	419.3889
Nonacosanoic acid	Nonacosylic acid	C29:0	29	:	0	29	438.4437	452.4593	421.4409	437.4359
Triacontanoic acid	Melissic acid	C30:0	30	:	0	30	452.4593	466.4750	435.4566	451.4515
c-21-Triacontenoic acid	Lumequeic acid	C30:1	30	:	1	28	450.4437	464.4593	433.4409	449.4359
		C30:2	30	:	2	26	448.4280	462.4437	431.4253	447.4202
Henatriacontanoic acid	Henatriacontylic acid	C31:0	31	:	0	31	466.4750	480.4906	449.4722	465.4672
Dotriacontanoic acid	Lacceroic acid	C32:0	32	:	0	32	480.4906	494.5063	463.4879	479.4828
Dotriacontenoic acid	Dotriacontenylic acid	C32:1	32	:	1	30	478.4750	492.4906	461.4722	477.4672
Dotriacontadienoic acid	-	C32:2	32	:	2	28	476.4593	490.4750	459.4566	475.4515
Triatriacontanoic acid	Psyllic acid	C33:0	33	:	0	33	494.5063	508.5219	477.5035	493.4985
Tetraatriacontanoic acid	Gheddic acid	C34:0	34	:	0	34	508.5219	522.5376	491.5192	507.5141
Tetraatriacontenoic acid	Tetraatriacontenylic acid	C34:1	34	:	1	32	506.5063	520.5219	489.5035	505.4985
Tetraatriacontadienoic acid	-	C34:2	34	:	2	30	504.4906	518.5063	487.4879	503.4828
Pentatriacontanoic acid	Ceroplactic acid	C35:0	35	:	0	35	522.5376	536.5532	505.5348	521.5298
Hexatriacontanoic acid	Hexatriacontylic acid	C36:0	36	:	0	36	536.5532	550.5689	519.5505	535.5454
Hexatriacontenoic acid	Hexatriacontenylic acid	C36:1	36	:	1	34	534.5376	548.5532	517.5348	533.5298
Hexatriacontadienoic acid	-	C36:2	36	:	2	32	532.5219	546.5376	515.5192	531.5141

The table ends here

Appendix E - LC-DAD polyphenol Characterization

E

The polyphenol subclasses can have distinguished each other from the specific chromophore groups present in their structures. We can define as chromophore a portion of a molecule responsible for its colour. The colour that is seen by our eyes is the one not absorbed within a certain wavelength spectrum of visible light.

Molecules with chromophore groups can be easily detected by LC-DAD analysis, generating UV spectra. In most cases the UV spectrum may (focusing on the spectral shapes) reveal information of the possible structural features within the sub-group.

Of course, quality of the UV spectra is also dependent on the solution in which the spectra are recorded and on the concentration of the analytes. With LC-DAD-MS instruments the biggest problem in this respect is related to formic acid that is one of the most commonly used eluents in polyphenol analysis. The problem is especially pronounced if aqueous formic acid concentrations are exceeded above 0.1%. In these cases, very dilute LC peaks do not give very useful UV spectra as formic acid distorts the spectra at 190-230 nm region.

The figure E.1 could be used as a guide in the UV spectra interpretation: focusing on the maximum of absorption it's possible to distinguish the main polyphenolic species. Then, if an MS system is combined to the DAD system, the hypothesis generated by UV spectra will be confirmed.

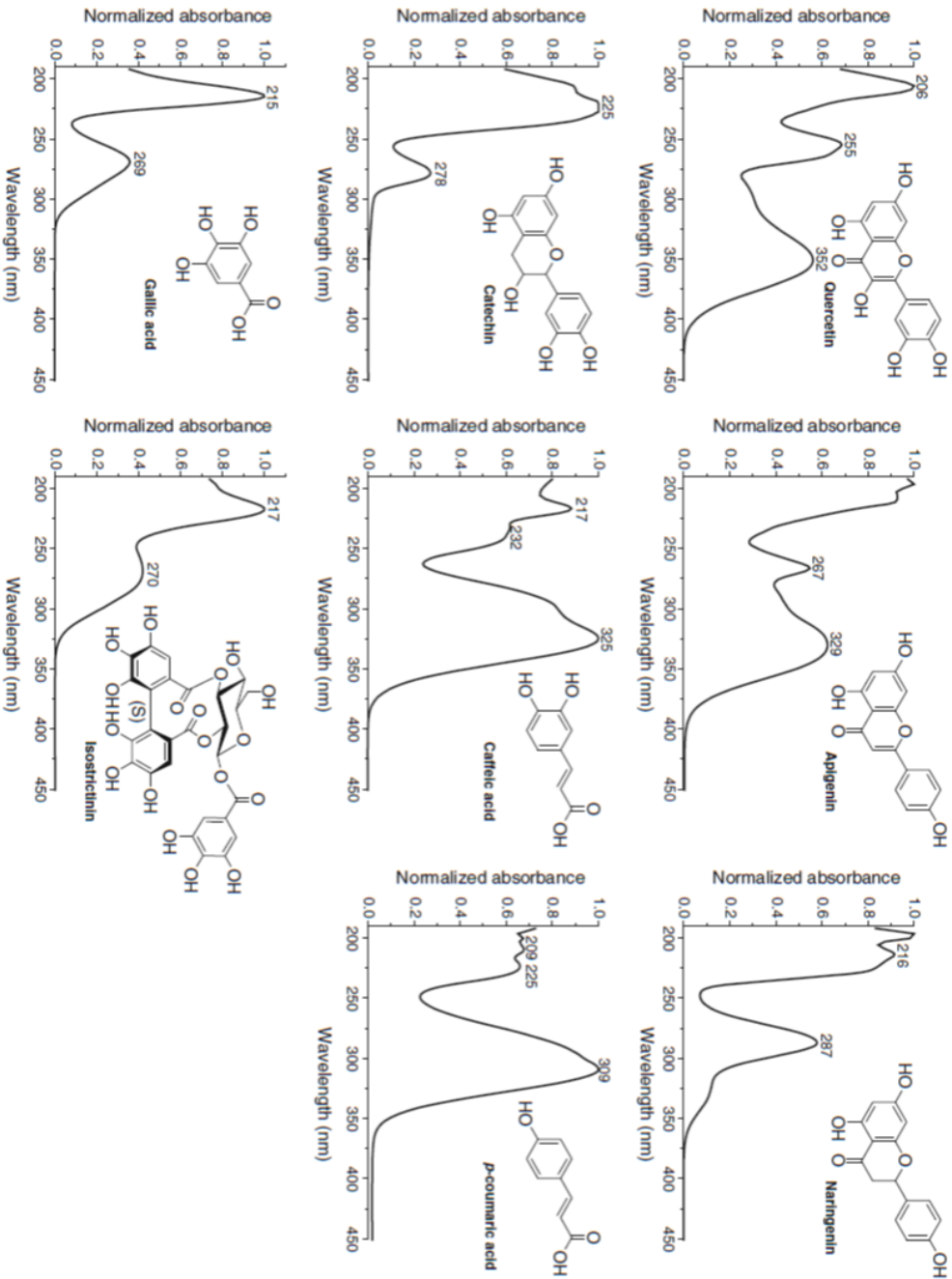


Figure E.1.: UV spectra of common polyphenols detected in plants