AN INVESTIGATION INTO THE CIRCADIAN REGULATION OF METABOLISM IN MAMMALS IN RELATION TO SEX, HEALTH, AND DISEASE.

by

THOMAS PAUL MICHAEL HANCOX

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham

September 2023

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Circadian rhythms describe endogenous ~24 h cycles in behaviour and biological processes that are regulated by the circadian system. The circadian system, akin to an internal clock, regulates an organism's behaviours and biological processes about the day so that they occur at the most optimal time. Over the course of the last decade a new field has emerged applying metabolomic platforms to study the circadian regulation of metabolism in humans. Interest in this field has grown as epidemiological evidence has shown metabolic disorders, such as obesity and type 2 diabetes, are enriched in populations at risk of circadian misalignment e.g., shift workers. This epidemiological evidence is suggestive of a link between the two systems. Prior studies in the field have attempted to characterise circadian metabolite rhythms in healthy populations to better understand the nature of this relationship between the two systems and their co-regulation of one another. In addition, studies simulating shift work and sleep deprivation, which can induce circadian misalignment, have been performed to investigate how these factors disrupt the regulation between these two systems. Studying circadian rhythms in humans requires specialist facilities and demanding protocols that control for external factors that can influence the circadian system and require multiple samples, e.g., blood, to be taken across a >= 24 h time course. As a result, cohorts from circadian studies are often small in number and favour male participants to remove the menstrual cycle, another biological rhythm, as a conflicting variable.

First and foremost, this thesis sought to review the current knowledgebase of this emergent field and identify avenues of future research. Identified gaps in prior research included the aforementioned lack of female representation with no prior human circadian/metabolomics study having investigated a balanced male and female cohort. Further, it was made apparent that there was little diversity in the ultra-high performance liquid chromatography mass spectrometry (UHPLC-MS) assays applied in the field with the vast majority favouring reversed phase assays in positive ion mode.

Second, this thesis acted on the results of the review to apply the most comprehensive untargeted UHPLC-MS study to data, encompassing four assays. These four UHPLC-MS assays were applied to high resolution (2 hourly, 32 h time course) time series samples collected from the largest mixed sex cohort in the field to date, containing 16 males and 16 females. This resulted in the largest list of annotated rhythmic metabolites assembled to date and highlighted lipids; specifically, glycerolipids, fatty acids, and phospholipids, as metabolite classes strongly regulated by the circadian system. Further, it is the first study capable of directly comparing male and female metabolite rhythms as the male and female cohorts were subject to the same experimental design from the point of recruitment through to sample and data analysis. These comparisons facilitated novel observations such as a greater number of rhythmic lipids in females versus male counterparts.

Finally, the data handling and rhythm analysis workflows developed as part of the human circadian study were applied to other time series metabolomic datasets. In this instance to investigate alterations in circadian regulation of metabolism of Huntington's disease at different ages within a large mammal (sheep) model. Application of the rhythm analysis workflow identified subtle shifts in phospholipid rhythms in sheep with Huntington's disease relative to controls which appear to progress with age. These perturbations were observed in pre-symptomatic animals and further our understanding of the pathology of Huntington's disease at this early stage.

Acknowledgements

Gratitude to the Biotechnology and Biological Sciences Research Council (BBSRC) for funding the project via the Midlands Integrative Biosciences Training Partnership (MIBTP).

I wish to thank my primary supervisor Prof. Warwick Dunn for his tutelage and guidance during my time of study at the University of Birmingham. Your passion for high quality research and willingness to teach guided me through the challenges of undertaking a PhD. It has been a privilege to work as a member of your team these past few years. I would also like to extend my thanks to co-supervisors Dr. Robert Dallman and Prof. Debra Skene. I could not have asked for better mentors to introduce me to the field of chronobiology, thank you both for always being so willing to share your knowledge, experience, and signposting opportunities within the field. These thanks are extended to members of your respective research teams, past and present. Special thanks to:

Lukas, Annie, Judith, Will and the fourth-floor team at Birmingham for your advice and assistance preparing me for a PhD in metabolomics. Dr Catherine Winder for her technical expertise and guidance in the lab. Despite the difficulties brought about by social distancing you always remained patient and ready to assist.

Laura for supervising my first PhD lab rotation, and the rest of the Dallmann group at Warwick for helping introduce me to field of chronobiology.

The University of Surrey chronobiology team, with special mentions to Prof. Daan van der Veen, Dr Silke Kiessling, Dr. Matt Spick, Namrata, and Ting for all the discussions and advice relating to rhythm analysis, ultradian/circadian rhythms and being excellent collaborators.

I would also like to extend my gratitude to Dr. Aneika Leney and members of the Leney and Cooper groups on the 8th floor. Our joint group meetings brought me a new level of appreciation for mass spectrometry and its applications outside of metabolomics. Special thanks to members of both teams,

past and present, in particular: Danielle, Jas, Jedd, Kish, and Matt. Despite our projects operating in different disciplines you were all always willing to offer meaningful insight, encouragement and discussion. Whenever I got frustrated with broken spaghetti code or weary from staring at endless data tables you were all there to lend an ear or a snack (do you get the fig?).

Last but certainly not least I wish to thank my future wife, Emma. Your endless encouragement, supply of tea, and belief in me even when I fail to believe in myself has always kept me moving forward and striving to be my best.

Contents

1. Introduction **1**

1.1. Chronobiology and circadian rhythms **2**

1.1.1. Mammalian clocks and their molecular components **2**

1.1.2. The SCN, melatonin & cortisol: the clock and its hands **7**

1.1.3. Circadian desynchrony and health outcomes **11**

1.2. Metabolomics **14**

1.2.1. Overview **14**

1.2.2. Experimental workflows applied in metabolomics studies **17**

1.2.2.1. Untargeted metabolomics **17**

1.2.2.2. Targeted metabolomics **18**

1.2.2.3. Untargeted UHPLC-MS workflow **21**

1.2.3. Platforms **29**

1.3. Analytical techniques **31**

1.3.1. Chromatography **31**

1.3.1.1. Overview **31**

1.3.1.2. Mechanisms of separation – HILIC **35**

1.3.1.3. Mechanisms of separation – Reversed phase chromatography **37**

1.3.2. Forgoing chromatography – alternative approaches **39**

1.3.3. Mass spectrometry **40**

1.3.3.1. General concepts **40**

1.3.3.2. Electrospray ionisation **43**

1.3.3.3. Mass analysers & ion selection **46**

1.3.3.4. Quadrupole mass analysers **47**

1.3.3.5. Orbitrap mass analysers **50**

1.3.3.6. Tandem mass spectrometry **54**

1.4. Raw data processing **55**

1.4.1. Deconvolution(pre-processing) **55**

1.5. Metabolite annotation and identification **59**

1.5.1. Levels of metabolite identification **59**

1.5.2. Application of m/z, RT, and MS/MS in metabolite annotation **61**

1.6. Conclusion **66**

1.6.1. Research Objectives **66**

2. Materials and methods **70**

2.1. Methods – Chapter 3 **71**

2.1.1. Literature search and inclusion criteria **71**

2.2. Methods – Chapter 4 **72**

2.3. Methods – Chapters 5 & 6 **74**

2.3.1. Participant recruitment & sample collection **74**

2.3.1.1. Recruitment & pre-study **74**

2.3.1.2. Cohort **75**

2.3.1.3. Constant routine **76**

2.3.1.4. Sample storage **77**

2.3.1.5. Plasma melatonin and cortisol profiles **77**

2.3.2. Metabolomics – UHPLC-MS assays **78**

2.3.2.1. Chemicals **78**

2.3.2.2. Instrumentation **78**

2.3.2.3. Quality control samples **78**

2.3.2.4. Sample preparation **79**

2.3.2.5. Lipidomics UHPLC-MS assay **80**

2.3.2.6. HILIC UHPLIC-MS assay **81**

2.3.2.7. Data deconvolution & quality controls **82**

2.3.2.8. Metabolite annotation **83**

2.3.3. Rhythm analysis **87**

2.3.3.1. Z-score **87**

2.3.3.2. DLMO correction **87**

2.3.3.3. Applied rhythm analysis algorithms **88**

2.3.3.4. Multivariate analysis & descriptive statistics **89**

2.3.3.5. Cohort changes **90**

2.4. Methods – Chapter 7 **91**

2.4.1. Animal husbandry & sample collection/analysis **91**

2.4.1.1. Animals **91**

2.4.1.2. Pre-study & sample collection **91**

2.4.1.3. Sample storage **92**

2.4.1.4. Plasma melatonin and cortisol profiles **92**

2.4.1.5. Plasma urea profiles **92**

2.4.2. Metabolomics – (semi)-targeted assay **92**

2.4.2.1. Quality control **93**

2.4.3. Rhythm analysis **93**

3. Time of day variation of the human metabolome **95**

3.1. Introduction **96**

3.2. Literature search – Results and commentary **99**

3.2.1. Circadian variation **107**

3.2.2. Sleep deprivation and prolonged wakefulness **108**

3.2.3. Shift work **111**

3.2.4. Diurnal (24 h) rhythms **112**

3.2.5. Health status **114**

3.2.6. Diet composition **116**

3.2.7. Morning versus evening studies **117**

3.2.8. Summary of the field **118**

3.3. Investigating metabolite rhythms – the next steps **120**

3.4. Chapter summary **123**

4. Development of data handling & analysis tools **125**

4.1. Introduction **126**

4.2. Results and Discussion **130**

4.3. Chapter summary **142**

5. Investigating metabolite rhythms of non-polar metabolites via a UHPLC-MS lipidomics assay applied to time series plasma samples collected under constant routine from males and females – a discovery based study. **145**

5.1. Introduction **146**

5.2. Results **150**

5.2.1. Data processing outcomes **150**

5.2.2. Group level rhythm analysis **150**

5.2.3. Individual level analysis **161**

5.3. Discussion **164**

5.4. Summary **171**

6. Investigating metabolite rhythms of water-soluble metabolites via a UHPLC-MS HILIC assay applied to time series plasma samples collected under constant routine from males and females – a discovery based study. **174**

6.1. Introduction **175**

6.2. Results **177**

6.2.1. Data processing outcomes **177**

6.2.2. Group level rhythm analysis **177**

6.2.3. Individual level rhythm analysis **186**

6.3. Discussion **188**

6.4. Chapter summary **193**

7. An investigation into circadian regulation of metabolites, in Huntington's disease via a new sheep model organism, applying a commercially available targeted assay **196**

7.1. Introduction **197**

7.2. Results **202**

7.2.1. Data processing outcomes **202**

7.2.2. Unsupervised multivariate analysis **202**

7.2.3. Group level rhythm analysis **204**

7.2.4. Pathway enrichment analysis **212**

7.3. Discussion **213**

7.4. Chapter summary **221**

Thesis Summary **223**

Reference List **231**

Appendix **268**

Appendix A **268**

Appendix B **275**

Appendix C **276**

Appendix D **297**

Appendix E **301**

List of Figures

1.1: A simplified schematic of the mammalian circadian system, from photic input to circadian output– reproduced from [16]. **4**

1.2: A simplified schematic of the core mammalian clock 'machinery' - reproduced from [17]. **6**

1.3: An example profile of melatonin in human males and females, reproduced from [64]. **10**

1.4: Melatonin profiles (top) demonstrating the shift in DLMO for (A) daytime shift condition vs baseline and (B) night-time shift condition vs baseline and cortisol profiles (bottom) showing the shift in cortisol rhythms between day and night-time shift conditions (C) relative to clock time and (D) circadian Time i.e., DLMO. Reproduced from [70]. **13**

1.5: An overview of the major steps that define the experimental design of a typical metabolomics experiment, reproduced from [121]. **21**

1.6: A depiction of the various analyte-column interactions that can influence selectivity in RPC or HILIC columns, reproduced from [184]. **32**

1.7: A simplified schematic of a UHPLC system interfacing with a mass spectrometer, reproduced from [196]. **34**

1.8: A schematic representing the assembly of a Q Exactive Plus Orbitrap mass spectrometer coupled to an atmospheric pressure electrospray ionisation source (see Section 1.3.3.2.) reproduced from Thermo Fisher [223]. **42**

1.9: A schematic illustrating ion formation in an ESI source, reproduced from [231]. **45**

1.10: A schematic illustrating the orientation and the opposing composite electric field maintained between pairs of quadrupoles, reproduced with addition of the axes from [234]. **48**

1.11: A schematic illustrating the QqQ operation of a triple quadrupole, reproduced from [239]. **49**

1.12: Cross section of the C-trap ion accumulation device and the Orbitrap mass analyser with an example of an ion trajectory (A), reproduced from [226], and example traces for a detected single and calculated frequency (B-D), reproduced from [227]. **53**

1.13: An illustration of peak grouping and integration within XCMS. **58**

1.14: (A) An example MS/MS spectra (top) compared to a reference spectrum (bottom) for a precursor ion of *m/z* 205.09692 with multiple structural isomers (B-C, reproduced from PubChem). **65**

2.1: Infographic providing details on sample collection and the constant routine protocol – further details in [64]. **77**

4.1: An illustrated workflow of the R coded pipeline. **131**

5.1: Principal components analysis in males (n = 10, green) and females (n = 10, red) for the lipidomics assay in (A) positive ion mode and (B) Negative ion mode, post quality filtering. **151**

5.2: Group level rhythm analysis outcomes (A) Number of rhythmic metabolite features by sex (males: n = 10, females n =10) (B) acrophase (peak time) by sex per lipid class (C) stacked

bar chart displaying % of rhythmicity in annotated (lipid) metabolites by sex per lipid class, with accompanying table displaying absolute number of metabolite features. **153**

5.3: Distribution of annotated metabolite features acrophase relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) male lipidomics assay, (B) female lipidomics assay. Male cohort n = 10, female cohort n =10. **155**

5.4: Male annotated rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged (male: n = 10) changes in relative abundance (red = high, blue = low) relative to dim light melatonin onset (DLMO, -6 : 24 h from left to right) and organised by metabolite class. **156**

5.5: Female annotated rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged (female: n = 10) changes in relative abundance (red = high, blue = low) relative to dim light melatonin onset (DLMO, -6 : 24 h from left to right) and organised by metabolite class. **157**

5.6: (A) – bar chart displaying number of rhythmic metabolites per individual determined by individual level rhythm analysis, (B) – Box and whisker plot of individual level analysis outcomes. **163**

6.1: Principal components analysis of metabolites in males (n = 10, green) and females (n = 10, red) for the HILIC assay in (A) positive ion mode and (B) Negative ion mode, post QC filtering. **178**

6.2: Group level rhythm analysis outcomes, number of rhythmic metabolite features by sex (males: n = 10, females n =10). **180**

6.3: Acrophase distributions relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) male HILIC assay, (B) female HILIC assay. Male cohort n = 10, female cohort n =10. For example, 25 metabolite features peaked 12 h after DLMO for the female group. **182**

6.4: Male (A, left) and female (B, right) annotated rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged (male: n = 10, females: n = 10) changes in relative abundance (red = high, blue = low) relative to dim light melatonin onset (DLMO, -6 : 24 h from left to right). **183**

6.5: (A) – bar chart displaying number of rhythmic metabolites per individual determined by individual level rhythm analysis, (B) – Box and whisker plot of individual level analysis outcomes. **187**

7.1: Principal components analysis of all metabolites across all time points for (A) 2015 control cohort (n = 13, red) & HD (n = 14, green) and (B) 2017 HD control cohort (n = 14, red) & HD cohort (n = 16, green). **203**

7.2: Group level rhythm analysis outcomes, number of rhythmic metabolites by time point (2015, 2017) and group (control = white, HD = blue). **204**

7.3: Reproduced in its entirety from [282] Comparison of rhythmic metabolites in control sheep to the same metabolites in HD sheep. Rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged changes in metabolite concentrations (red = high, blue = low) relative to dim light melatonin onset (DLMO) in 5-(control group, A; HD group, B) or 7- (control group, C; HD group, D) year old sheep. **206**

7.4: Acrophase distributions relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) Control sheep 2015, (B) HD sheep 2015. Control cohort n = 13, HD cohort n =14. **210**

7.5: Acrophase distributions relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) Control sheep 2017, (B) HD sheep 2017. Control cohort n = 14, HD cohort n =16. **211**

List of Tables

1.1. A summary of the advantages and disadvantages of targeted and untargeted workflows, reproduced from [142]. **20**

2.1: Adduct list submitted to BEAMSpy. **85**

2.2: Neutral losses list submitted to BEAMSpy. **85**

2.3: Isotope list submitted to BEAMSpy. **86**

3.1: Summary of literature search parameters and outcomes, reproduced from [78]. **99**

3.2: A brief summary of study design, cohort details and results with regard to observed time of day variation of metabolites for relevant studies analysing plasma/serum, reproduced from [78]. **101**

5.1: Number of metabolite features (a) following XCMS raw data processing, (b) following quality filtering and (c) following group rhythm analysis. **152**

5.2: A list of putatively identified metabolite features with an acrophase difference of >2 h between males and females. **154**

5.3: A list of putatively identified rhythmic (lipid) metabolites observed in both our group level analysis and the group level analysis of previous studies. **158**

5.4: A list of putatively identified rhythmic (lipid) metabolites also identified as differentially rhythmic between males and females. **160**

6.1: Number of metabolite features (a) following XCMS raw data processing, (b) following quality filtering and (c) following group rhythm analysis. **179**

6.2: A list of putatively identified metabolite features with an acrophase difference of >2 h between males and females. **180**

6.3: A list of putatively identified rhythmic water-soluble metabolites observed in both our group level analysis and previous studies. **184**

6.4: A list of putatively identified rhythmic (lipid) metabolites also identified as differentially rhythmic between males and females. **185**

6.5: Pathway enrichment analysis results for rhythmic putatively identified metabolites (non-lipid). **185**

7.1: A list of metabolites with an acrophase difference of >2 h between control and HD sheep at 2015 (age five) or 2017 (age seven) time points. **205**

7.2: A list of rhythmic metabolites also identified as differentially rhythmic between Control and HD sheep, 2017 (age seven) cohort. **207**

7.3: Pathway enrichment analysis results for rhythmic metabolites (non-lipid & lipid) in the 2015 (age five) cohort. **212**

7.4: Pathway enrichment analysis results for rhythmic metabolites (non-lipid & lipid) in the 2017 (age seven) cohort. **213**

List of abbreviations

ANOVA – Analysis of variance

ARNT – Aryl hydrocarbon receptor nuclear translocator

ATGL – Adipose triacylglycerol lipase

AUC – Area under the curve

BMAL1 – Basic helix-loop-helix ARNT like 1

bw – bandwidth

CAG – Cysteine-adenine-guanine

CAR – Acylcarnitine

CCGs – Clock-controlled genes

Cer – Ceramide

CID – Collision induced dissociation

CK – Casein kinase

CL – Cardiolipin

CLOCK – Circadian locomotor output cycles kaput

CR – Constant routine

CRAN – Comprehensive R archive network

CRY1/2 – Cryptochrome circadian regulator1 / Cryptochrome circadian regulator2

CV – coefficient of variation

DC – direct current

DESI – Desorption electrospray ionisation

DG – Diacylglyceride

DI-MS – Direct infusion mass spectrometry

DLMO – Dim light melatonin onset

DODR – Detection of differential rhythmicity

DS – Dayshift

ECHO – Extended circadian harmonic oscillator (ECHO)

EIC – Extracted ion chromatogram

ESI – Electrospray ionisation

FA – Fatty acyl

FA(X:X) – Fatty acid

FFA – Free fatty acid (cycle)

FIA-MS – Flow injection mass spectrometry

FT-MS – Fourier transform mass spectrometer

FT-ICR – Fourier transform ion cyclotron resonance

FWHM – Full width at half maximum

GC-MS – Gas chromatography mass spectrometry

GL – Glycerolipid

GluCer – Glucocyl-ceramide

GM3 – Ganglioside mannoside 3

HCD – High carbohydrate diet

HCD cell – Higher energy C-trap dissociation cell

HD – Huntington's disease

Hex2cer – Dihexosylceramide

HFD – High fat diet

HILIC – Hydrophillic interaction chromatography

HMDB – Human metabolome database

HSL – Hormone sensitive lipase

HTT – Huntingtin

ICU – Intensive care unit

IPA – Isopropanol

IPO – Isotopologue Parameter Optimization

iPRGCs – Intrinsically photosensitive retinal ganglion cells

KEGG – Kyoto Encyclopaedia of Genes and Genomes

KNN – K-nearest neighbour

LC – Liquid chromatography

LLOQ – Lower limit of detection

LOESS – Locally Estimated Scatter-plot Smoother

LOQ – Limit of quantification

m/z – Mass to charge ratio

MALDI – Matrix-assisted laser desorption ionisation

MeSH – Medical subject headings

MESOR – Midline estimating statistic of rhythm

mHTT – Mutant Huntingtin

MOSAIC – Multi-omics selection with amplitude independent criteria (MOSAIC)

MS – Mass spectrometry

MS/MS (MS2) – Tandem mass spectrometry

MSI – Metabolomics standards initiative

NCBI – National center for biotechnology information

NeuAcHexCer – N-Acetylneuraminic acid hexosylceramide

NMR – Nuclear magnetic resonance

NS – Nightshift

OB – Obese

OPLS-DA – Orthogonal partial least squares – discriminant analysis

OVT73 – *Ovis aries* transgene 73, a sheep model organism for Huntington's disease

OW – Overweight

PA – Phosphatidic acid

PC(X:X) – Phosphatidylcholine

PC – Principal component

PCA – Principal component analysis

PE – Phosphatidylethanolamine

PER1/2 – Period1 / Period2

PG – Phosphatidylglycerol

PI – Phosphatidylinositol

ppm – Parts per million

PS – Phosphatidylserine

PSQI – Pittsburgh sleep quality index

q – Quadrupole operating in RF mode only

Q – Quadrupole operating with a static field overlayed with an RF field

QC – Quality control samples

QqQ – Triple quadrupole

RAIN – Rhythmicity analysis incorporating non-parametric methods

REV-ERBa – Reverse-erythroblastosis

RF – Radio frequency

RHT – Retinohypothalamic tract

ROI – Regions of interest

RORa – Retinoic acid receptor-related orphan receptor

RPC – Reversed phase chromatography

RRE – ROR / Rev-ERB - response element

RSD – Relative standard deviation

RT – Retention time

SCN - Suprachiasmatic nucleus

SM – Sphingomyelin

SP – Sphingolipid

ST – Sterol lipid

T2DM – Type 2 diabetes mellitus

TCA – Tricarboxylic acid cycle

TG – Triacylglyceride

ToF – Time of flight

UHPLC-MS - Ultra-high performance liquid chromatography coupled to mass spectrometry

UPLC – Ultra performance liquid chromatography

VDA – Venn diagram analysis

XCMS – Extensible computational mass spectrometry

CHAPTER 1

INTRODUCTION

1.1. Chronobiology and circadian rhythms

1.1.1. Mammalian clocks and their molecular components

In mammals, the circadian system is a network of endogenous oscillators, or clocks, comprised of a transcription-translation autoregulatory feedback loop of "core clock genes". These self-regulating clocks temporally regulate expression of specific genes, known as clock-controlled genes (CCGs), which in turn regulate further downstream processes and behaviours generating robust cycles with a ~24 h period otherwise known as circadian rhythms [1]. Such 'rhythms' are typically modelled as sinusoidal waves which include but are not limited to: melatonin secretion [2], regulation of sleep/wake cycle in conjunction with the sleep homeostat [3,4], core body temperature [5], and insulin secretion and sensitivity [6]. Regulating processes such as these around a ~24 h cycle is considered advantageous, from an evolutionary standpoint, as the temporal regulation can facilitate specific behaviours to consistently occur at the most favourable time given that there is a robust and consistent cue which can be anticipated [7,8]. Earth's ~24 h light/dark cycle is both anticipable and a near universal cue across ecosystems, therefore organisms adapt their behaviours (e.g., feeding/fasting, sleep/wake) to adhere to this cycle. By definition circadian rhythms are self-sustaining [9], however, the frequency of the internal biological clock does not exactly match the frequency of the solar day; in fact, the length of circadian frequencies differs between individuals favouring >~24 h lengths [10]. Therefore, maintenance of the phase relationship (synchronised timing) between internal circadian time and the external light/dark cycle requires the system to have a degree of plasticity. Some cues, such as light, can directly influence the regulation of the circadian system, advancing or delaying its natural frequency to achieve resonance. Such cues are referred to as Zeitgeber (time giver). Synchronisation of a circadian clock to an external cycle, such as the light/dark cycle, is known as entrainment.

In the absence of Zeitgeber the circadian clock is described as free running, as it adheres to its natural frequency often depicted with actograms displaying the daily 'drift' in sleep/wake times relative to social time due to the differing frequencies of the endogenous circadian frequency (~24 h) to that of the solar day (24 h).

The majority of cells in the body exhibit an autonomous circadian clock capable of generating overt ~24 h rhythms. Thus, individual tissues/organs can exhibit specific and localised temporal regulation of CCGs and downstream processes resulting in differential responses to Zeitgeber such as feeding/fasting cycles. However, some form of entrainment is required to maintain synchrony across this multitude of clocks in different tissues and by extension maintain optimal timing of biological processes [11]. In mammals, the suprachiasmatic nucleus (SCN), otherwise referred to as the 'master pacemaker' or 'central clock', receives photic information via the retinal hypothalamic tract (RHT). Photic input occurs through intrinsically photosensitive retinal ganglion cells (iPRGCs) which make up non-image forming vision in humans and contain the photo pigment melanopsin which is most sensitive to blue light (peak absorbance with wavelengths at ~480 nm) [12]. iPRGCs relay photic input to the SCN via the RHT, with the photic input initiating a signalling cascade which can ultimately modulate the core clock machinery and induce phase shifts (advancing or delaying SCNs circadian rhythm). The SCN, through numerous mechanisms such as autonomic innervation, endocrine signalling, and eliciting specific behavioural responses will synchronise peripheral clocks. When in harmony the system will maintain robust temporal regulation of biological processes resulting in consistent circadian outputs. Outputted behaviours such as increased appetite, will result in feeding, which is a Zeitgeber for peripheral clocks like the liver. The consequences of these outputs, such as the act of feeding, may influence specific aspects of metabolism, in this instance energy balance, which can feedback and influence the circadian

system [11, 13-15], as illustrated in Figure 1.1. A common metaphor used to describe this hierarchical relationship describes the SCN as a conductor which then directs the orchestra (peripheral clocks) maintaining synchrony within the whole system.

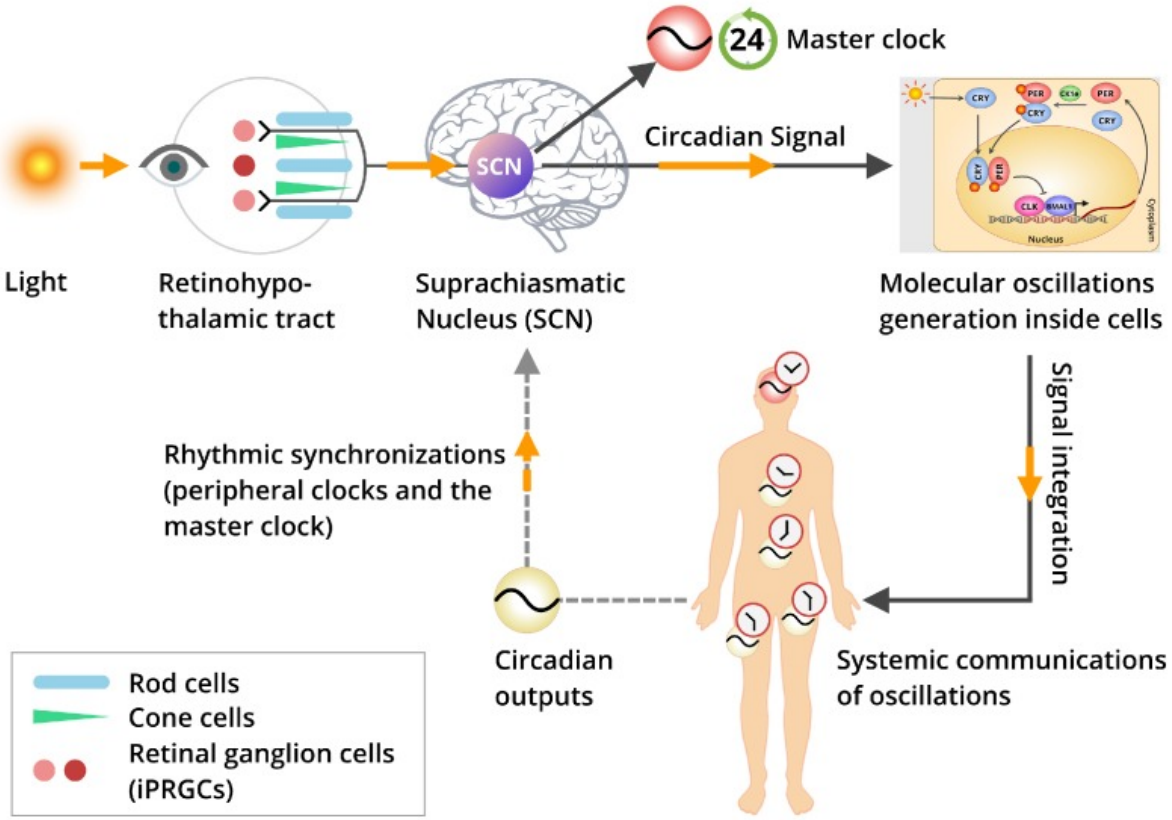


Figure 1.1: A simplified schematic of the mammalian circadian system, from photic input to circadian output– reproduced from [16].

As detailed above the mammalian circadian system functions on a basis of negative feedback with delay due to a transcription-translation autoregulatory feedback loop of core clock genes. The key interactions of these clock genes can be portrayed as two simple interacting loops (Figure 1.2, [17]). The first of these loops, referred to here as the core feedback loop involves *CLOCK-BMAL1* (circadian locomotor output cycles kaput – basic helix-loop-helix ARNT (aryl hydrocarbon receptor nuclear translocator) like 1) and *PER-CRY* (period – cryptochrome circadian regulator). The genes *CLOCK* and *BMAL1* are expressed at an early time in relation to the solar day. The products of these genes, CLOCK and BMAL1, are transcription factors that form a heterodimer and bind to the E-box regions of the clock genes: *PER1*, *PER2*, *CRY1*, *CRY2 and* CCGs to enhance transcription. The PER and CRY proteins will begin to accumulate later in the day, towards the evening, and form a complex with casein kinases CK1δ/ε. This PER/CRY complex translocates into the nucleus at night and acts to repress expression of *CLOCK/BMAL1* in turn repressing the transcription of *PER*/*CRY* genes. This repression is short lived however due to targeted ubiquitylation and degradation of the complex; once removed the cycle is then free to repeat itself the proceeding morning thus creating a system of negative feedback with delay. The second of these loops, referred to here as the auxiliary feedback loop involves REV-ERB and ROR. Driven by the CLOCK/BMAL1 complex, REV-ERBα (Reverse-erythroblastosis) and RORα (Retinoic acid receptor-related orphan receptor) comprise part of a separate negative feedback cycle, such that, elevated levels of one ultimately leads to a decline in the other. REV-ERBα competes with RORα for binding at RRE (ROR / Rev-ERB - response elements) with RORs binding to enhance expression and REV-ERB repressing expression of associated genes. This rhythmic cycling of REV-ERB and RORα feeds back to the core loop via enhanced/repressed expression of *BMAL1*, modulating it in such a way *BMAL1* is expressed rhythmically and in antiphase to *PER* expression, helping

maintain the functionality of the core loop as described above. The action of REV-ERB and RORα on RRE can also regulate expression of CCGs.

Taking the above into account, circadian rhythms can be defined as follows: Circadian rhythms are biological processes with a near 24 h period that are temperature compensated, resulting in no change to the rate of circadian oscillations across a significant physiological range of temperatures, which persist in the absence of Zeitgeber (free running conditions) but can entrain (synchronise) to Zeitgeber, ergo, they are self-sustaining [9, 18,19]. Any ~24 h rhythms not meeting these definitions, i.e., which do not/are not known to persist under free running conditions, are referred to as diurnal rhythms.

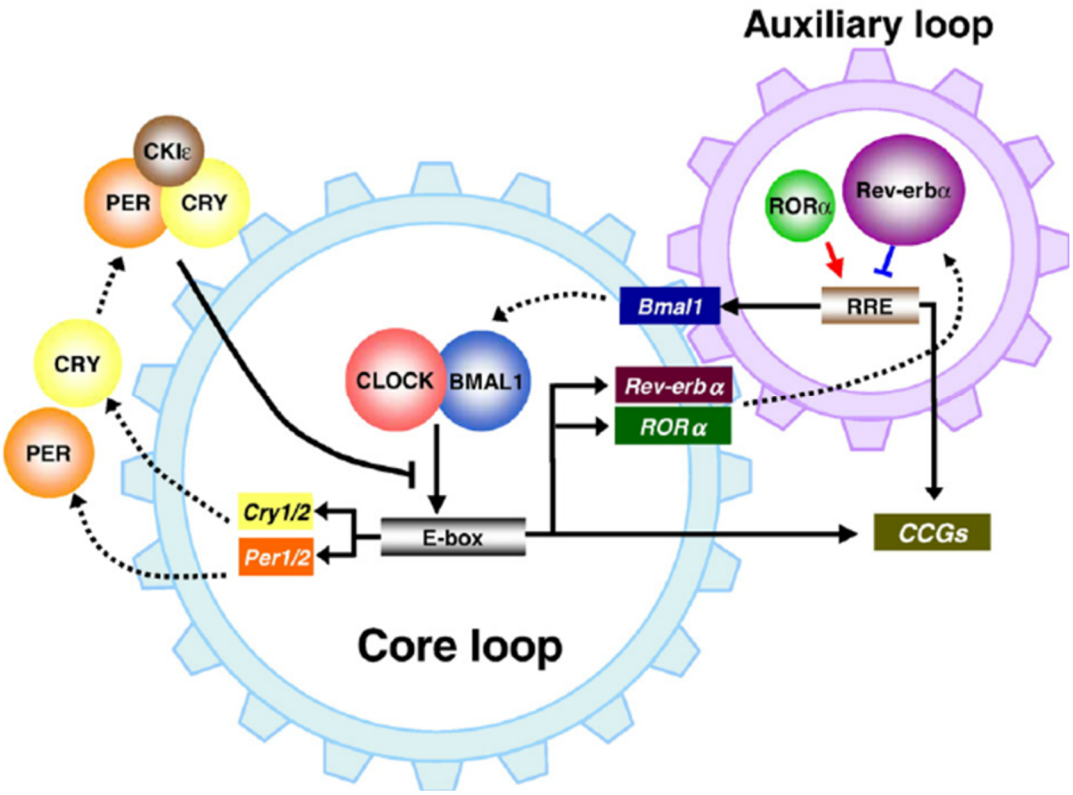


Figure 1.2: A simplified schematic of the core mammalian clock 'machinery' - reproduced from [17]. ARNT – Aryl hydrocarbon receptor nuclear translocator, BMAL1 – basic helix-loop-helix ARNT like 1, CCGs - Clock-controlled genes, CK1 – Casein kinase 1 eta, CLOCK – circadian locomotor output cycles kaput, CRY1/2 – Cryptochrome circadian regulator, PER1/2 – Period, REV-ERBα – Reverse-erythroblastosis, RORα – Retinoic acid receptor-related orphan receptor, RRE – ROR / Rev-ERB - response element.

1.1.2. The SCN, melatonin & cortisol: the clock and its hands

In hindsight, it is easy to see how SCN is suited to the role as the "master pacemaker" as it is the only internal clock in mammals directly entrained by, and therefore synchronised to, light/dark cycles. However, the initial proposal and acceptance of the SCN as pacemaker in mammals covers several decades of research [20]. The relationship between the visual and circadian systems, via the RHT, was first established in the early 1970s [21-23]. Within this timeframe the SCN was first reported and put forward as an endogenous (circadian) oscillator/pacemaker of mammalian circadian rhythms [24-26]. This concept of the SCN as a pacemaker was strengthened by observations which showed the SCN to have a persistent endogenous rhythm when 'isolated' in vivo [27] and to maintain its inherent endogenous oscillations in vitro [28-30]. Any remaining doubt about the SCNs role as master pacemaker was dispelled upon demonstrating that circadian rhythms could be perturbed by ablation of the SCN but then restored by transplantation of a donor SCN in animal models [31]. Not only were circadian rhythms restored post transplantation, but recipients inherited the circadian frequency of the donors, which were purposefully selected mutant strains with shorter circadian frequencies [32]. Whilst the concept of the SCN as the master pacemaker was generally accepted by the 1990s, it is far from the final discovery made expanding our understanding of the mammalian clock. One such further revelation being the discovery that individual cells can exhibit endogenous oscillations of varying frequencies [33].

If the SCN is akin to a clock for mammals then melatonin represents one of the hands in which the time is read, representing the gold standard marker of circadian phase. Such a statement is dangerously innocuous as it fails to convey the decades of research that established melatonin's significance within the circadian system. Melatonin was first isolated in 1958 [34].

Its diurnal regulation was established in the early 1960s when light was first demonstrated to suppress melatonin synthesis [35-37] and later in 1980 it was shown to be capable of complete suppression of melatonin secretion in humans [38]. Light is not the only regulatory mechanism of melatonin synthesis, shortly after the SCN was first proposed as an endogenous pacemaker the SCN was also shown to regulate key pathway enzymes in melatonin synthesis from serotonin [39,40] and establishing the initial link between these two components of the clock. At the same time binding of melatonin to hypothalamic membranes was first described [41], and then further refined to the SCN in the hypothalamus [42]. Melatonin receptor sites were later mapped within the hypothalamic region in mammals [43,44]. The presence of these receptors on the SCN corroborated with prior evidence that melatonin could act as an entraining agent to the mammalian circadian system [45] by providing a mode of action in which melatonin could directly interact with the SCN. By the turn of the millennium melatonin was conclusively shown as an entraining agent of circadian rhythms, or more specifically outputs of the circadian system such as sleep, as it could do so under free running conditions within sighted and blind individuals [46-48]. This body of evidence established that melatonin was not only regulated by the master pacemaker, the SCN, but was capable of feedback to entrain circadian outputs such as the sleep/wake cycle [49].

Cortisol is another important hormone within the mammalian circadian system and maintains the ongoing theme of an extended history to reach our modern day understanding of its role within the clock. Edward Kendall is often credited as the discoverer of cortisol in 1937, then known as ‘compound F’ and later 17-hydroxycorticosterone, due to his group’s elucidation of its structure [50]. By the early 1950s the diurnal plasma/urinary excretion profile of cortisol and related compounds were documented [51] leading onto further studies into the diurnal variation of corticosteroids in health and disease [52, 53]. Similar to melatonin, light was

reported as a potential synchronising agent of cortisol rhythms independent of the sleep wake cycle in the 1960s [54] and was believed to be regulated, at least in part by the SCN. This was due to the loss of a daily corticosterone rhythm in SCN ablated rats [25], this loss of rhythm could also occur due to lesions in the fornix region of the brain [55]. The SCN as a regulator of cortisol would also explain the prior observation of light being a regulator of cortisol's daily rhythms, similar to melatonin, as it had already been established the SCN receives photic input. Perhaps driven by the potential of joint regulation by the SCN, subsequent studies began to document the phase angle of entrainment between melatonin and cortisol under a range of experimental conditions. Over the proceeding decades melatonin and cortisol rhythms were demonstrated to remain phase locked in: night shift workers [56, 57], across seasons in healthy adults [58], in prolonged absence of natural daylight [59] and in blind individuals [60], however, the dynamic of this relationship may change with age [61], changes in mental health [62], or under 'isolated, confined and extreme' environments [63]. The phase of melatonin & cortisol is also reported to be the same between sexes, though as discussed by Honma *et al*, there is not a complete consensus within the literature [64, 65]. This stable phase relationship under a range of conditions and between sexes is likely by design rather than coincidence given that both melatonin and cortisol are regulated by a single entity, the SCN. The nature of this regulation affords the mammalian clock two temporally segregated yet inextricably linked hormones in which circadian outputs can be organised around; melatonin preceding the night and cortisol preceding the day. So reliable are these markers that melatonin is considered the gold standard (bio)marker of internal circadian time with other biological rhythms measured relative to dim light melatonin onset (DLMO) [64]. Here, DLMO is defined as the time at which melatonin concentration exceeded 25% of the peak (highest) measured melatonin concentration and is measured

independently for each participant. For example, if melatonin concentration peaked at 80 pg/ml for a given participant/group then the time point at which melatonin exceeded 25% of this value (20 pg/ml) is defined as DLMO for that individual/group as illustrated below (Figure 1.3).

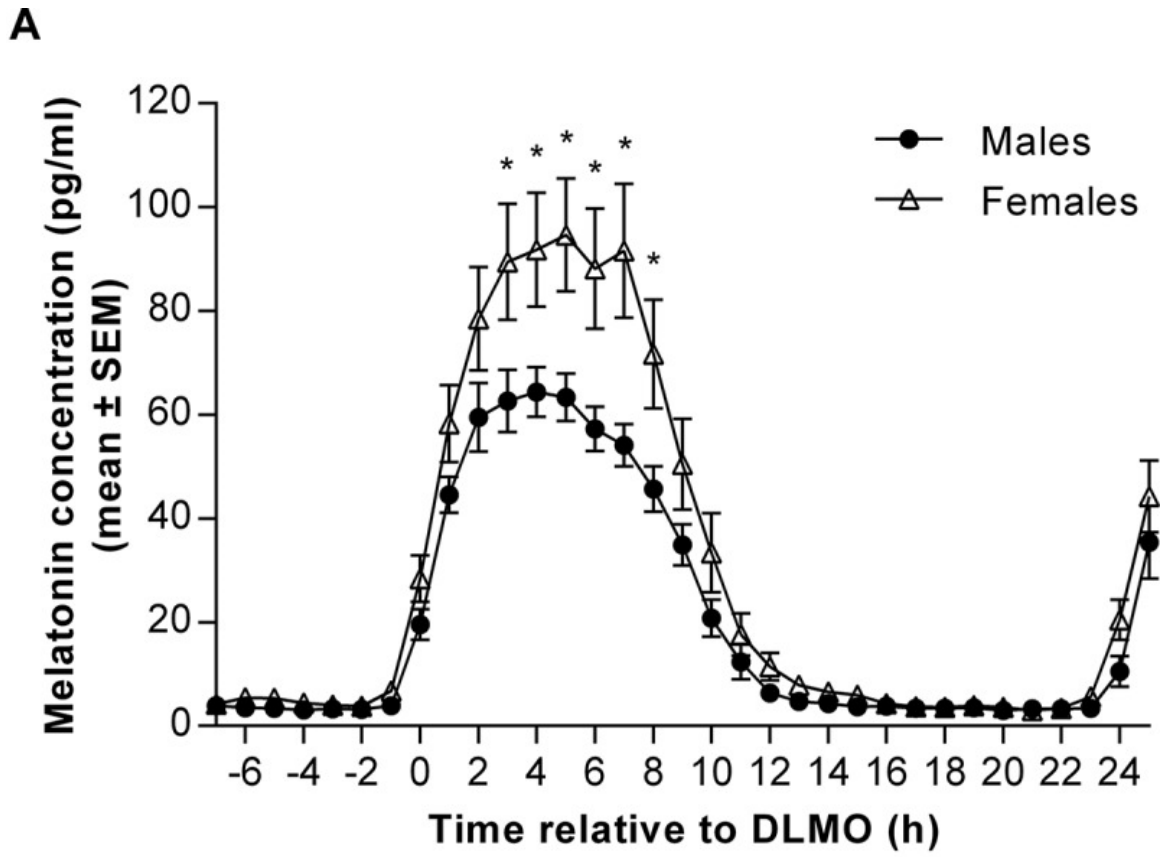


Figure 1.3: An example profile of melatonin in human males and females, reproduced from [64].

1.1.3. Circadian desynchrony & health outcomes

The mammalian circadian system is a robust biological time piece that, by nature of its evolution, is capable of limited adjustments to its phase to synchronise with daily changes in the solar day. Whilst it is a sufficient time piece for the natural environment the system evolved in, the system is fallible within the confines and self-imposed expectations of modern human society. The circadian system can be perturbed, resulting in circadian misalignment, should a Zeitgeber such as: light/dark, feeding/fasting, and in turn sleep, become irregular or erratic in their timing and or length of exposure [66-68]. For example, exposure to artificial light late in the solar day can attenuate night time melatonin levels which in turn can delay circadian outputs such as sleep contributing to circadian misalignment [69]. Further, simulated shift work studies have demonstrated a significant shift in DLMO and cortisol rhythms [70] resulting from night shift relative to baseline measurements (Figure 1.4). Development of artificial light has allowed humans, a diurnal species, to overcome the restrictions enforced by our natural light/dark cycle and birthed a 24 h society in which other key behavioural cycles such as sleep/wake and feeding/fasting can become wholly uncoupled from the natural cycle established by the solar day which we have evolved alongside. Shift work is one such example where humans may deviate from typical 'diurnal' behaviour; the consequences of which are not fully understood [69] but is an active area of research. This research activity having been prompted due to the negative health implications associated with shift work and its relevance to a large proportion of western populations. Shift workers, i.e., individuals who follow non-traditional atypical work shifts outside of 9am-5pm, are estimated to comprise 21% of the workforce in Europe (54 million workers as of 2015 [71]) and 16% of the workforce in the USA (23.5 million workers, as of 2017 [72]). Therefore, millions of individuals could be affected by circadian misalignment. Circadian misalignment,

and by association shift work, has been linked to numerous chronic conditions such as: metabolic disease, hormone dependent cancers, cardiovascular disease and poor mental health [73]. Furthermore, circadian regulated genes have been found to be enriched amongst disease-associated genes when compared to non-circadian controlled genes [74]. Individual metabolic processes and their relationship with the clock are well documented, with respect to circadian regulation [75-77]. However, the association between circadian misalignment and onset of metabolic disorders in addition to disruption of energy balance inducing circadian misalignment in a mouse model [14] suggests a much broader and complex relationship between the circadian system and the tissue-specific/global metabolism in mammals. This need to observe changes in metabolism and the metabolome beyond the level of select pathways, and better understand its interaction with the human circadian system (circadian rhythms) and behavioural/environmental cycles (diurnal rhythms) has resulted in a new interdisciplinary field which applies metabolomics (and other omics) platforms to analyse multiple samples collected from humans across a >24h time series under a variety of conditions; as reviewed here [78].

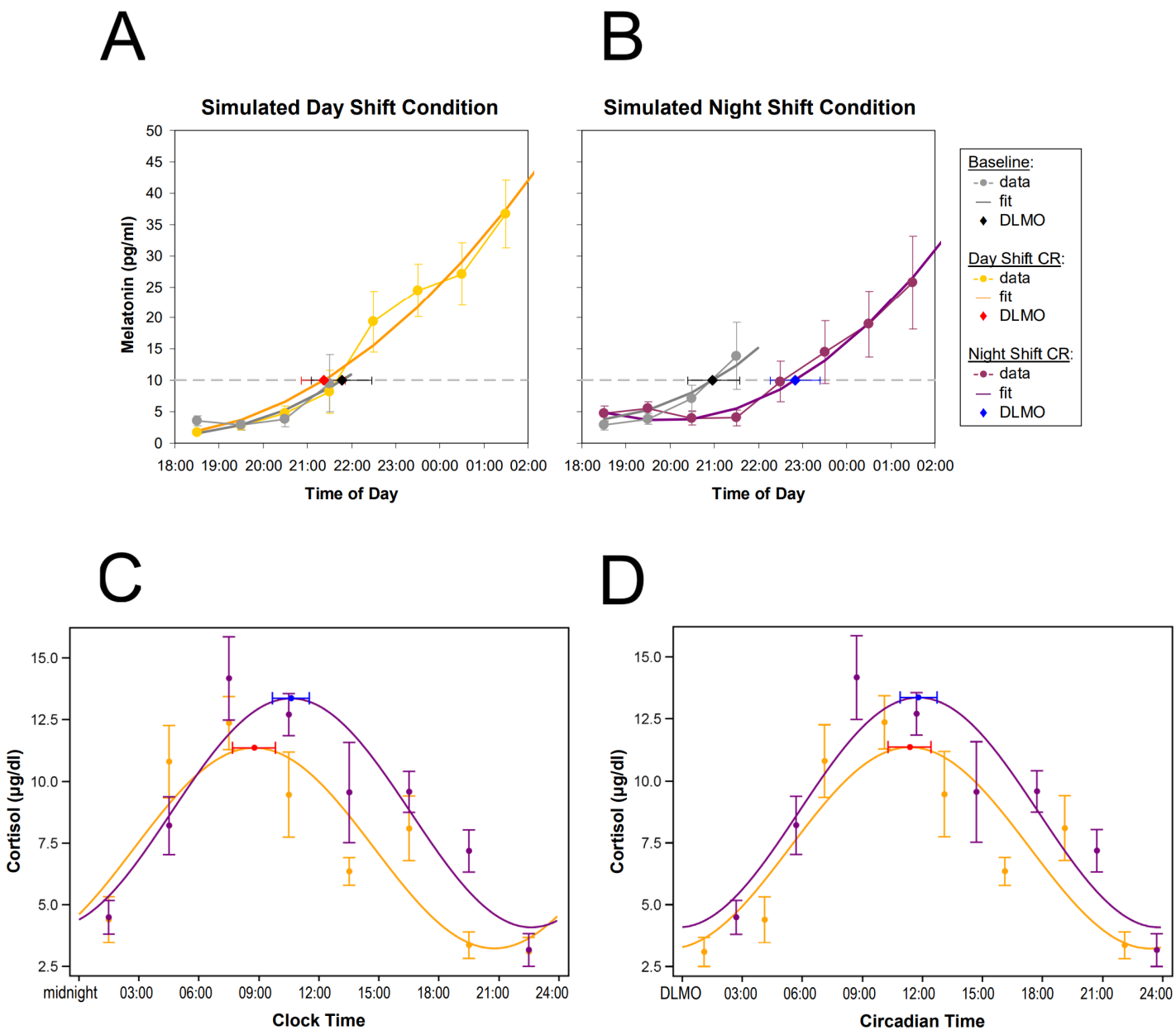


Figure 1.4: Melatonin profiles (top) demonstrating the shift in DLMO for (A) daytime shift condition vs baseline and (B) night-time shift condition vs baseline and cortisol profiles (bottom) showing the shift in cortisol rhythms between day and night-time shift conditions (C) relative to clock time and (D) circadian Time i.e., DLMO. Reproduced from [70]. Night-time shift pattern causes a significant delay to DLMO relative to baseline condition whilst no significant shift in DLMO was observed in the daytime conditions. A similar delay to the cortisol acrophase (peak time) was also observed during the night-time condition yet remained in phase (synchronised) relative to DLMO such that the relative peak times of cortisol and melatonin did not change in either condition only their times relative to clock (social) time.

1.2. Metabolomics

1.2.1. Overview

A metabolite can be defined as a product or reactant of metabolism and is commonly applied as a catch all term to describe small (< ~2000 Da) naturally occurring molecules within a biological system [79-81]. However, this definition can be blurred when considering molecules such as peptides, e.g., glutathione, which can also be metabolically active. Furthermore, not all metabolites are synthesised and derived within an organism, i.e., endogenous metabolites, and may be sourced external to the organism, i.e., exogenous metabolites. Exogenous metabolites can include food derived metabolites, such as 'essential amino acids' like leucine in humans which cannot be synthesised endogenously but can be metabolised, or pollutants and other chemical/biological agents which may or may not be naturally occurring [79, 81-83]. As with other chemical compounds, metabolites can be defined and classified based upon similar chemical structural e.g., amino acids, similar physiochemical properties as is the case with lipids which can be broadly defined by their lack of solubility in water [83], and by their biological function (e.g. presence in a metabolic pathway). The chemical taxonomic system of nomenclature, as outlined by IUPAC [84], can be employed to more accurately convey the metabolite(s) class being referred to though many research groups use ClassyFire [85]. Other naming formats, chemical identifiers including SMILES and InChI [86] and nomenclature exist, for example in the naming of lipids [87]. In this thesis metabolites and lipids are defined under the single term metabolites.

The complete complement of metabolites in a sample type, tissue, or organism is defined as the metabolome. The term metabolome is therefore context dependent as the whole organism metabolome will differ in the number and concentration of metabolites to the

blood, urine or salivary metabolomes. Metabolomes vary in size by sample type ranging from 10s to 1000s of metabolites, though unlike the genome, the metabolome cannot be easily sequenced and many metabolites detected have no chemical structure assigned or structures are only just being assigned [88]. A number of public metabolite databases have been created to efficiently collate and disseminate metabolite information relating to: physiochemical properties, biological function and species observed. Such databases include the Human Metabolome Database (HMDB) [89] which contains 220,945 metabolite entries that are confirmed or expected to be found within humans. LIPIDMAPS [90], a database dedicated to compiling information on lipids, has 47,825 unique entries inclusive of known and *in silico* generated structures across multiple species. These databases are still considered incomplete due to the large quantity of 'unknown' metabolites in metabolomics studies and the vast number of theoretical metabolite structures that could exist [89, 91, 92], resulting in additional challenges in metabolite annotation discussed below (see Section 1.5.).

Metabolites exhibit a high degree of diversity in their chemical and physical properties and in the multiple biological functions they have such as being components in macromolecules (proteins, RNA, DNA), energy storage/release (e.g., ATP), signalling molecules (e.g., ketone bodies [93]) , or cofactors (e.g., iron and heme [94]) [79, 83]. To fulfil these roles, metabolites undergo a multitude of enzymatic-mediated chemical (or otherwise known as metabolic) reactions and transformations. It is this complex network of chemical reactions that comprise metabolism, subsections of which are known as metabolic pathways which have a defined input/output [79]. Metabolic pathways do not act in isolation as individual metabolites or cofactors may take part in multiple reactions and pathways within metabolism [95]; for example, pyruvate synthesised in the glycolysis pathway is fed into the citric acid cycle (Tricarboxylic acid cycle (TCA)) as acetyl-CoA. The Kyoto Encyclopaedia of Genes and Genomes

(KEGG) has a database containing 352 such pathways in humans, mapping the metabolites and enzymes that make up a given metabolic process [96, 97]. Changes in metabolism may occur to compensate for the changing needs of the organism under varying internal or environmental conditions; such changes in metabolism can be inferred from the metabolome. Further, changes in metabolite concentrations may occur as a result of other biological processes, for example apoptosis. Apoptosis is a form of programmed cell death which releases metabolite signalling molecules into the surrounding tissue of the dying cell promoting apoptotic induced proliferation amongst other responses [98].

The metabolome is considered the functional level most representative of an organism's phenotype due to its dynamic response to external stimuli (e.g., lifestyle: diet [99, 100], exercise [101-103]), physiology (e.g., age [104], sex [105]), and internal regulation as has been identified in genome-scale metabolic models [106-109]. It is important to note that metabolism occurs at the level of individual cells and tissues which possess their own microenvironments, e.g., metabolic changes occurring within/around a malignant tumour can be different to that in a different part of the tissue which is healthy, and subject to differing epigenetic regulation [107, 110]. Therefore, no one biological sample is representative of the global whole organism metabolome; however, biofluids such as blood and urine are routinely used in metabolomics studies due to their minimal invasiveness to collect and acting as a proxy to monitor changes in multiple metabolically active organs (i.e., liver/kidney) [79, 111, 112].

Metabolomics is a sub discipline of systems biology and is the study of the complete complement of metabolites within a biological system [79, 81, 83, 113-119, 120, 121]; it is often employed to study metabolism and human health. Applications of metabolomics

technologies include but are not limited to: toxicology, forensics and anti-doping [122-125]; drug discovery [126-129]; chemical ecology [130]; bioengineering [131]; agriculture, food security and food authentication [132-136]; and biomedicine [137] in addition to metabolic profiling and biomarker discovery. Despite the wide adoption and application of metabolomics it is not without its challenges, metabolites exhibit a wide array of structural and physiochemical diversity and are present in tissues which may vary greatly in their physiology, e.g., blood, liver tissue, cell lysate, creating unique challenges in extracting, isolating and measuring the metabolome. To overcome these challenges metabolomics has evolved into a diverse platform of hardware and software tools to meet the varying needs of the aforementioned disciplines which employ metabolomics technologies. The nature of a metabolomics analysis will therefore differ dependent on: *a priori* knowledge of the field being investigated; the biological question being asked and the required sample type; study design in which samples were collected such as single or paired collections; and the analytical platform(s) and workflow employed that are appropriate for the sample and question being asked [79, 80]. The design of two workflows, untargeted & targeted, and commonly used metabolomic platforms are discussed below along with a general overview of the experimental design of metabolomics (see Section 1.2.2.).

1.2.2. Experimental workflows applied in metabolomics studies

1.2.2.1. Untargeted metabolomics

Untargeted metabolomic assays are a generally unbiased methodology designed to provide relative quantification (typically a normalised peak area) and global coverage of the sample metabolome being studied, a concept more fully explored in the literature [79, 138, 139] and summarised below. The terms unbiased and global are applied figuratively as no single

analytical assay is truly unbiased nor provides complete coverage of the metabolome dependent on sample preparation and employed analytical techniques [140]. However, multiple assays used in conjunction with one another provide greater coverage through extraction and detection of a larger number of metabolites, with a wide range of physicochemical properties. Further details on metabolomics assays are provided below (see Section 1.3.). The relative changes in metabolite abundance between multiple test conditions is typically subject to statistical (e.g., t-test, one-way Analysis of variance (ANOVA)) and/or multivariate analyses (e.g., principal component analysis (PCA), orthogonal partial least squares – discriminant analysis (OPLS-DA)) to infer metabolite-related changes. Due to the wealth of data collected from untargeted studies and not relying on *a priori* knowledge of the biological system being investigated, they are typically employed as 'discovery/learning' studies for hypothesis generation. The newly formulated hypotheses can subsequently be subject to an alternative workflow for validation and translation using a targeted metabolomics strategy.

1.2.2.2. Targeted metabolomics

Targeted metabolomic assays are designed to focus on a subset of metabolites that are chemically characterised and annotated [79, 114]. Whilst untargeted assays optimise wide coverage of the metabolome, targeted assays are optimised to enrich and quantify the metabolite(s) of interest by removal of undesirable elements in a sample such as high abundance metabolites/proteins/salts [114, 115]. Such separation is possible through exploitation of physiochemical properties, e.g., molecular weight, structure, polarity, solubility, acidity/basicity, volatility [79] of the targeted metabolites [114]. Targeted assays have traditionally focussed on <20 metabolites [79] as enriching a larger range of metabolites

may result in sub optimum recovery due to the increased complexity of extraction [115]. The increased complexity of sample preparation in targeted assays versus that of untargeted is but one reason why *a priori* knowledge of the targeted metabolites is required. A second being the targeted selection and isolation of analytical signals, such as mass to charge ratio (*m/z*) in MS, corresponding to the targeted metabolite panel [114, 115, 140]. As targeted assays are focussed on a smaller subset of metabolites it is possible to construct calibration curves for the entire panel of metabolites being investigated for the purpose of absolute quantification, an unfeasible feat within an untargeted assay [79]. Calibration curves would typically employ a dilution series of each metabolite within the sample matrix to be analysed, e.g., plasma, with the instrument response/analytical signal measured and graphed across the dilution series. An isotopically-labelled internal standard or standards can be included at the same concentration in all samples and calibration solutions to normalise for small experimental errors (e.g. $^{13}C_6$ glucose could be used as an internal standard for the quantification of glucose). The molarity of an 'unknown' can then be interpolated from its corresponding intensity signal [114, 115]. Should multipoint calibration curves for each metabolite not be feasible then single point calibration for each metabolite or metabolite class may be employed [141]. When complete quantification of all metabolites is not obtainable but has been achieved to a degree for a subset, the workflow may be described as semi-targeted. Semi-targeted workflows occupy the middle ground between a complete lack of quantification observed in untargeted assays and the complete quantification observed in targeted assays. Therefore, the semi-targeted approach merges the advantages of untargeted and targeted studies by providing greater coverage than a targeted approach and a degree of quantification lacking in the untargeted approach. A summary of the advantages and disadvantages of both untargeted and targeted workflows is provided below in Table 1.1.

Table 1.1: A summary of the advantages and disadvantages of targeted and untargeted workflows, reproduced from [142].

Workflow	Advantages	Disadvantages
Targeted	• Low limit of detection • Usually quantitative • Data analysis and interpretation are easier • Metabolite data can be connected with pathways	• Limited number of compounds can be targeted • Non-targeted compounds are not considered • Purified standards of targeted compounds are required for quantification
Untargeted	• Unbiased and comprehensive • High-throughput • Allows the discovery of new compounds not expected to be in the sample or not expected to be associated with the biological question	• Semi-quantitative • Larger number of false positives and false negatives • Many unknowns • Data interpretation can be challenging

The greatest strengths of untargeted assays are therefore: broad coverage, ease of implementation, hypothesis generating, at the cost of demanding processing and data analysis. Whilst the greatest strengths of targeted assays are: quantification, reproducibility, less laborious data handling, but requires *a priori* knowledge and bespoke method optimisation [89, 91, 114, 115, 138, 143, 144].

1.2.2.3. Untargeted UHPLC-MS workflow

Below is a brief overview of a typical metabolomics workflow and will primarily be in the context of performing ultra-high performance liquid chromatography coupled to mass spectrometry (UHPLC-MS) assays in an untargeted workflow on human serum/plasma and summarised by Figure 1.5.

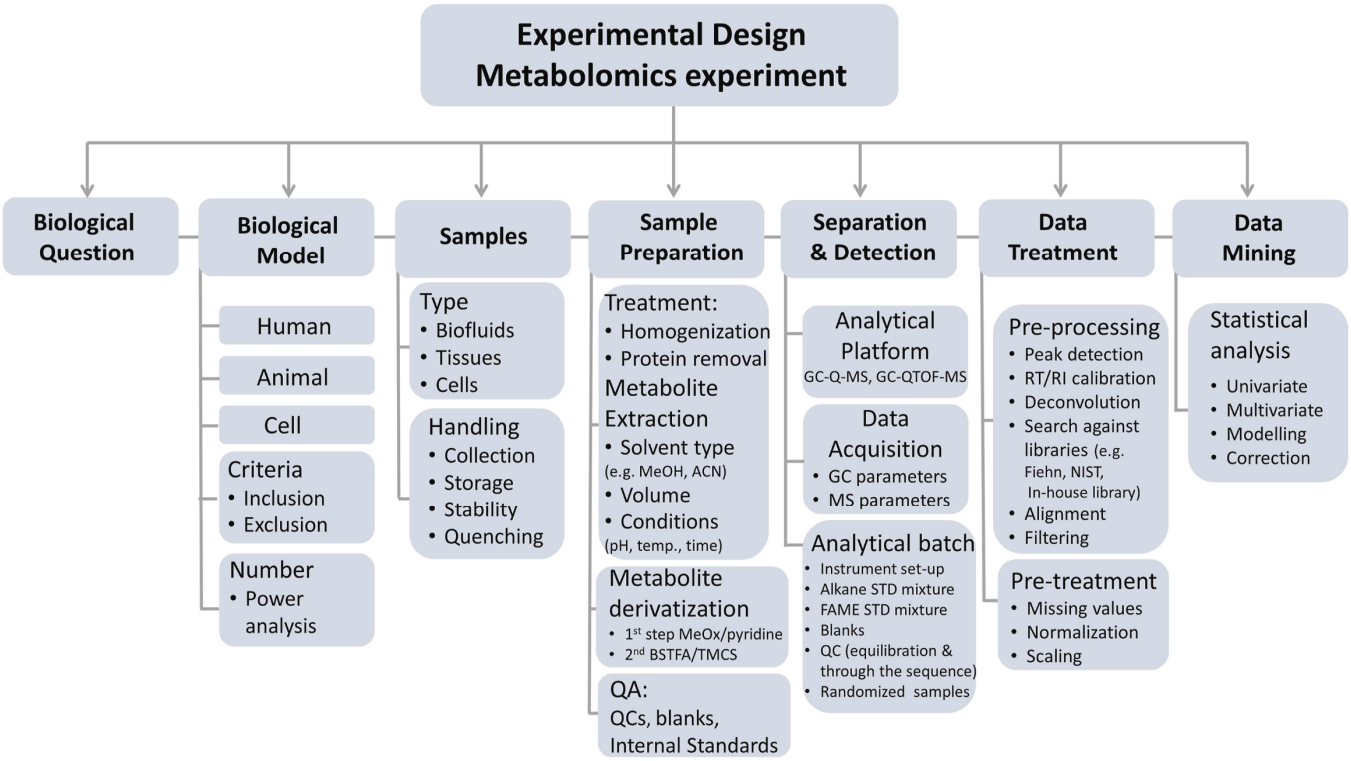


Figure 1.5: An overview of the major steps that define the experimental design of a typical metabolomics experiment, reproduced from [121].

Metabolomics represents a broad range of tools and methodologies for the analysis of metabolites. A good research question in metabolomics should therefore consider which of these tools (e.g., liquid chromatography, tandem mass spectrometry) and workflows ((semi)targeted versus untargeted) are best able to address the posed question. Commonly applied applications of metabolomics studies involve investigating the effect of a treatment and measure the metabolome before and after; alternatively, they or focus on the identification of biomarkers [79]. After formulating a biological question to be tested the next step in a metabolomics study is the considered curation of a cohort population. When dealing with small cohorts, as is common practice in circadian studies or pilot studies [78], variation and confounding factors within the population need to be reduced to preserve the integrity of any analysis. This may be done by age and BMI matching control/test cohorts, among other variables where possible such as ethnicity and sex, to avoid inadvertently introducing a selection bias [79, 145]. In large scale biomarker/population studies variation of the cohort is not as easily controlled; thus, demographic information is important in stratifying the population to elucidate causal relationships between a disease and metabolite biomarker as opposed to correlating the biomarker to a confounding factor. This can pose a major challenge for biomarker studies as a 'perfect' metabolite biomarker should exhibit a large enough change in a specific phenotype that far exceeds the natural variation observed in the population, yet such biomarkers may not always exist for complex diseases such as cancer [81, 145]. After characterising the desired population, the next step is designing the sampling protocol.

The number of samples collected, population size sampled from, and population demographics are dictated by the research question being asked. For example, if the research question relates to the effect of a treatment an investigator may opt to collect longitudinal

(paired) samples before and after treatment from a control and test group (examples include: [146-148]); longitudinal samples can also be employed to monitor disease progression or severity (examples include: [149-151]). Alternatively, if the research question relates to biomarker discovery or metabolite profiling single sampling can suffice but large population sizes of a control and test(s) cohorts would be favourable when scaling up from pilot to validation studies [79]. For the purpose of monitoring circadian metabolite rhythms, a high sampling resolution (2 – 6 h) across an extended period (>24) would be required [78]. Metabolomics is capable of analysing a range of tissues and biofluids to investigate localised responses to stimuli or investigate pathogenesis [81], however, tissue biopsies may not always be an option pre-mortem or be deemed too invasive. As previously stated, this is why biofluids such as blood are routinely used in metabolomics studies due to their minimal invasiveness and acting as a proxy to monitor changes in multiple metabolically active organs (i.e., liver/kidney) [79, 111, 112]. Sample selection becomes a balancing act between ease of collection, relevance to the biological question and possessing the means to process the sample for analysis.

Once blood is collected from a participant it is mixed with an anticoagulant such as heparin or EDTA prior to being subject to centrifugation to separate the plasma (supernatant) and the blood cell pellet. Dependent on when the sample is expected to be analysed it will be frozen at -80^{o}C (long term storage), - 20^{o}C (short term storage) with freeze thaw cycles kept to a minimum, as this can cause degradation of the sample metabolome [152]. Metabolite extraction is typically performed on wet ice with samples stored at ~4^{o}C prior to analysis. Despite being considered an ‘easy’ sample to prepare blood samples may be subject to poor handling procedures, either due to lack of required equipment or not following SOPs, causing delays in coagulating and quenching blood samples (preventing further metabolic reactions)

and resulting in poor quality samples [152]. Due care and attention is therefore warranted at this step and all subsequent steps such as metabolite extraction to maintain quality control.

Metabolite extraction protocols can broadly be defined as monophasic or biphasic (liquid liquid) extractions and will differ dependent upon sample type, workflow ((semi)targeted/untargeted) and assay (e.g., reversed phase chromatography (RPC)/Hydrophilic interaction liquid chromatography (HILIC) discussed further below); no single method is capable of extracting all metabolites from a sample [153]. Biphasic extractions traditionally employed chloroform which is an immiscible organic solvent, now methyl tert-butyl-ether is often used in lieu, alongside methanol and water. Addition of this extraction solution to a serum/plasma sample results in the formation two layers, hence the moniker biphasic, the top layer being an organic layer enriched with (non-polar) lipids and an aqueous layer enriched with polar metabolites [154, 155]. Biphasic extractions have been reported to exhibit reduced extraction performance compared to monophasic extraction methods [155]. Furthermore, monophasic extraction methods utilise UHPLC-MS compatible solvents, such as methanol in polar extractions or isopropanol for lipid extractions, which avoid the need for drying and reconstitution of samples which is required in biphasic extractions [155]. If global coverage of the sample is the objective, as is the case for untargeted assays, and a high throughput method is required monophasic extractions are the favourable choice assuming there is adequate sample volume to perform multiple (polar, non-polar) monophasic extractions.

Metabolomics studies often collect 100's or 1000's of samples that require analysis, making it logistically unfeasible for a single operator to analyse all samples on a single instrument in a single uninterrupted run whilst maintaining system performance. Therefore, samples are

typically organised into several blocks known as ‘batches’, i.e., an uninterrupted run of samples analysed sequentially on a single platform by the same operator [156, 157]. However, different batches can exhibit independent and differing analytical variation which, if not addressed, becomes a confounding factor when merging datasets from different batches [156, 157]. Specialist software and methodology to address these issues exist and examples are shown here [143, 158-160] and are incorporated in metabolomic analysis platforms such as Metaboanalyst [161]. A further concern prior to data collection is the collinearity between sample/patient demographics and analytical performance (e.g., changes in mass analyser sensitivity) over the course of a batch. For example, analysing samples in order of age of the participant they were collected from is ill advised as any linear trends observed in the data may be the result of linear changes in analytical performance and not representative of phenotypic changes in the samples. To circumvent such problems in metabolomics studies the order of sample preparation and analysis is typically randomised to avoid such collinearities [143]. A common method of monitoring system stability and performance within and between batches is the use of quality control (QC) samples in conjunction with process blanks, the application of which is discussed by [79, 157, 162, 163], and summarised below. Process blanks contain the solvents applied within the metabolite extraction procedure but otherwise lack the biological material being analysed. Process blanks can be applied to identify metabolite features unrelated to the biological sample but associated with the extraction process, i.e., metabolite features derived from the solvents/additives. QC samples comprise several variations but the most common is the use of pooled QC samples that consist of equal small aliquots of all biological samples (serum/plasma) to be analysed, or a subset of the biological samples in larger studies. QC samples made in this manner represent a collection of homogenous samples representative of the biological samples. Ergo, when

analysed periodically throughout a batch they act as a measure of analytical variability by observing any changes in analytical response between subsequent QC samples. Such changes would initially be evident by a sudden change in the number of detected metabolite peaks, a reduction in the sum intensity measured for a sample, or shifts in retention time (RT). Dependent on the nature of the analytical variability QC samples can be used as reference points to apply complex normalisation methods such as: retention time correction (see Section 1.4.1.) to accommodate for small changes in retention time, batch correction (see above) to correct for inter-batch effects, and metabolite feature filtering. Feature filtering is a commonly applied practise in metabolomics to improve data quality with common steps applied outlined by [163]. These typical steps include filtering out metabolite features that breach one of the following parameters: features must be detected in >=70% of QC samples, features must have a relative standard deviation (coefficient of variation) < 30, any feature whose average intensity in the biological samples is below the average intensity of the process blank samples as a percentage assigned by the user, e.g., if 0.1 this is 10% or a factor of <10x). By applying these filters metabolite features that exhibit high variability across QC samples, are not present in the vast majority of QC samples or have a prominent presence in process blanks are removed, thus improving the quality of the dataset. QC samples may also be employed to condition LC columns or circumvent the need for replicate injections of biological injections.

Once blood samples have been collected and metabolites extracted, randomisation applied and batches determined, it is time to analyse the samples on the metabolomics platform of choice. Plasma/serum are complex biological samples and may require separation, e.g., via chromatography, prior to analysis to more accurately detect and measure metabolites present in the sample. Further details pertaining to chromatography (see Section 1.3.1.), mass

spectrometry platforms (see Section 1.2.3.) and their mode of operation (see Section 1.3.3.) are provided below. Metabolomics data collected on UHPLC-MS systems is three dimensional, containing information on retention time (relating to chromatographic separation), *m/z*, and signal intensity for thousands of detected metabolite signals (peaks) which can collectively be referred to as a feature intensity matrix. Processing of this complex data requires specialist software to (1) convert vendor format files to an opensource compatible file format and (2) preprocess the data ready for statistical analysis. The parameters applied in these steps can affect the outputted metabolite data. Raw data processing is discussed in greater detail below (see Section 1.4.). Furthermore, in an untargeted study these metabolite peaks have no known identity at the point of data acquisition and thus require annotation and identification to affirm the compounds that correspond to detected signals (see Section 1.5.).

Once processed the resulting feature intensity matrix(ices) can be subject to statistical analyses which follow well defined community guided workflows discussed in detail across numerous bioinformatics articles with examples shown here, [164-166], and a brief overview of key methods given below. Typical methods of analyses may include: Volcano plots (fold change analysis combined with T-test) to discern metabolite (features) that differ significantly between groups of interest and gauge the relative change in intensity/concentration between groups. If the number of groups >2 or there are multiple test variables (e.g., treatment & time) then (two way) ANOVA may be applied in lieu to determine significant differences between groups. Uni/bi-variate statistics are useful in identifying significant differences between groups based on pre-defined variables such as phenotype, thus drawing attention to metabolites of potential interest. However, the metabolome is a dynamic system influenced by a multitude of variables which may or may not have been considered within the experimental design (e.g., age, BMI, ethnicity, diet, medical history) that can contribute to the

observed inter/intra group variance. Multivariate dimensionality reduction methods such as PCA transforms complex data sets by projecting them onto a limited number of principal components (PC). The first principal component can be imagined as a line through the data which represents the maximum co-variance between any pair of datapoints, thus, it represents a linear component through the dataset capturing the most variance and represents a potentially meaningful pattern. The second principal component is uncorrelated to the first, i.e., when represented geometrically it is perpendicular to PC1, but other than this restriction is a linear line through the data which represents maximum co-variance between any pair of datapoints. PCs are ranked in order of variance 'captured' with PC1 representing the trend with greatest variance across the dataset, then PC2 and so on. It is typical to present the data as a 2D PCA plot wherein datapoints which are more similar will cluster together and those that are less similar have a greater distance apart along one or both plotted PCs. This allows for group outliers to be identified and, dependent on the groupings applied, can help identify the variables that contribute most to this variance [167]. Whilst group labels can be applied to data points, they do not inform the construction of the PCs, therefore PCA is considered an unsupervised data exploration tool useful for assessing data quality. Similar multivariate analysis, such as OPLS-DA, is a supervised multivariate technique that projects datapoints onto two components. Unlike PCA, components in OPLS-DA are plotted to maximise inter-group variation (typically plotted on the x axis) and intra group variation (y axis). Techniques such as OPLS-DA are therefore useful to identify discriminatory variables that explain inter-group differences, e.g., biomarkers differentiating healthy vs disease groups, or sources of intra-group variation. However, its use is cautioned as OPLS-DA and similar techniques are designed to force group separation requiring the model to be validated (e.g., by group separation observable in PCA) on a per case basis to

avoid known issues such as discriminatory variables being derived from noise [168]. To assist in conveying results from an otherwise complex dataset it is common within metabolomics to employ various data visualisation methods which include but are not limited to the aforementioned PCA/OPLS-DA plots, heatmaps (with or without clustering), boxplots, and various graph types.

1.2.3. Platforms

UHPLC-MS is the most widely used platform within metabolomics [119, 128, 169, 170]. Within the literature the latter may also be referred to as Ultra-performance liquid chromatography (UPLC), a trademark of Waters (Waters Corporation, MA, USA) or inaccurately abbreviated to LC-MS when still referring to UPLC/UHPLC-MS technologies; only the term UHPLC-MS will be used throughout this thesis. UHPLC-MS systems are so favoured due to their greater sensitivity i.e., their ability to detect and measure low abundance metabolites, versus alternative platforms such as nuclear magnetic resonance (NMR) which were more widely used prior to the advent of UHPLC-MS systems. Furthermore, UHPLC-MS systems can utilise multiple available liquid chromatography (LC) assays facilitating retention and analysis of metabolites across a broad range of polarities and hydrophobicity/hydrophilicity. It is this flexibility and greater coverage of the metabolome over previously popular platforms, such as gas chromatography mass spectrometry (GC-MS), which have seen UHPLC-MS become the most used platform since its debut in 2004 [119, 140, 169-172]. When coupled to the high accuracy of MS instrumentation, i.e., the ability to accurately measure mass, chemical formula can be derived for a wide array of metabolites in a sample [83]. However, accurate mass is insufficient information to derive metabolite structures; tandem mass spectrometry (MS/MS or MS^2) employs methods such as collision induced dissociation (CID, further details

provided below in section 1.4.2.4.) to fragment the analyte of interest and acquire additional structural information [83, 119]. However, the success and reliability of producing fragmentation spectra is platform dependent and reliant on databases/spectral libraries for metabolite identification [138, 173]. NMR offers lower coverage than MS methods, yet its strengths overcome some of the limitations coupled to MS. NMR yields greater structural information suitable for compound identification, even between isomers, in addition to: offering greater reproducibility, the option of real time analysis on living cells and cultures due to its non-destructive nature, and being more easily quantifiable vs MS. Due to these strengths NMR, alongside GC-MS, was one of the first platforms used within metabolomics and is still applied in the field.

As NMR [174, 175] and GC-MS [121, 176] are not directly pertinent to the following chapters their underlying principles and applications are not discussed here, though they are outlined in the previous works referenced [121, 174-177]. Further details on chromatography and MS in relation to UHPLC-MS systems, as applied to metabolomics, are provided below (see Section 1.3.).

1.3. Analytical Techniques

1.3.1. Chromatography

1.3.1.1. Overview

Liquid chromatography, as detailed by [178-183] and summarised below, can broadly be described as a methodology which separates out a multi-component mixture into its separate components by exploiting the unique physiochemical properties of the components and of the chromatography system (chromatography column, mobile phases etc). Separation typically occurs by the continuous absorption and then release of components at different rates on an immobilised surface (stationary phase) as the components pass through the column. The stationary phase composes a range of embedded chemical functional groups that can interact (e.g., via hydrogen bonding, Figure 1.6) with components (analytes). Analytes are carried over the stationary phase by a liquid mobile phase which may change in composition over the course of the chromatography run to progressively desorb (elute) analytes previously bound to the stationary phase. For example, for a C_{18} reversed phase assay of urine two mobile phases are applied, typically water (weaker eluting solvent) and acetonitrile (stronger eluting solvent). To elute analytes being more strongly bound to the stationary phase a gradient elution is applied where the composition of the mobile phase changes from high percentage of water/low percentage of acetonitrile to high percentage of acetonitrile/low percentage of water. The rate of change is programmable and can be linear or curved. The rate of desorption is dependent on the propensity of the analyte to interact with the stationary and mobile phases with analytes most weakly bound to the stationary phase eluting first while analytes interacting more strongly with the stationary phase eluting later and therefore achieving separation.

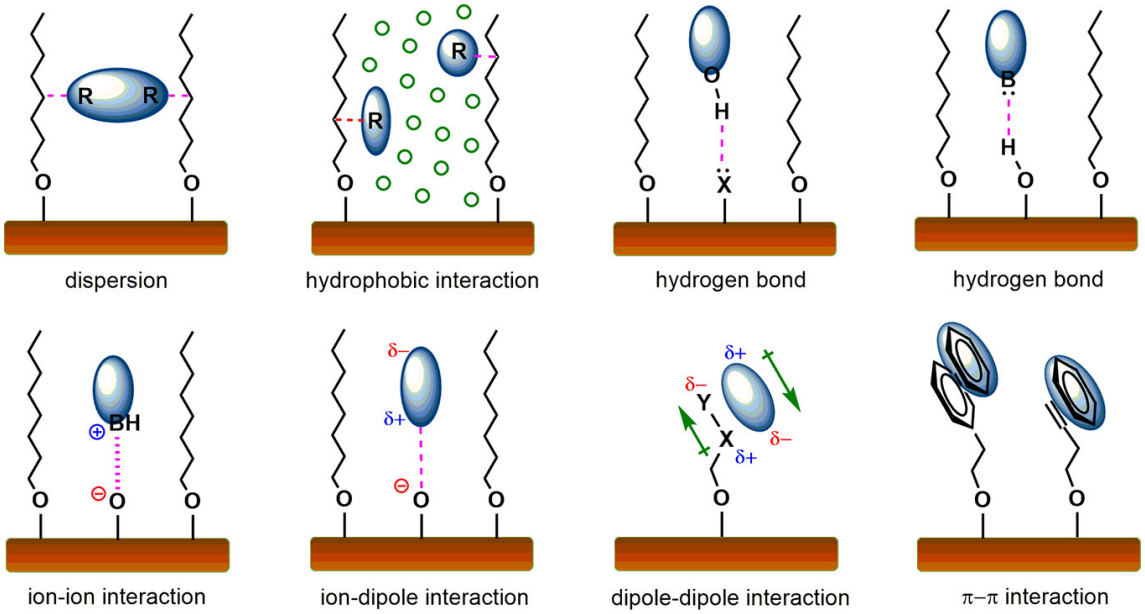


Figure 1.6: A depiction of the various analyte-column interactions that can influence selectivity in RPC or HILIC columns, reproduced from [184].

(London) Dispersion / van der Waals forces: Intermolecular forces between the analyte/stationary face as a result of temporary dipole moments. Interaction tends to be stronger among larger (non-polar) molecules [185].

Hydrophobic interaction: (aqueous) Solvent induced interactions affecting non-polar (hydrophobic) molecules that promote their spontaneous aggregation e.g., non-polar analytes to non-polar stationary phase in reversed phase chromatography (see below, Mechanisms of separation – Reversed phase chromatography) [186].

Hydrogen bond: Intermolecular forces between an atom with a lone pair of electrons and a hydrogen atom covalently bonded to a more electronegative atom (O, F, N), e.g., H_2O. A prominent interaction in HILIC chromatography [187] (see below, Mechanisms of separation – HILIC) and can contribute to retention in reversed phase chromatography [188].

Ion-ion interaction: Intermolecular attraction between two opposingly charged ions

Ion-dipole interaction: Intermolecular attraction between an ion and oppositely charged dipole moment induced by the ion.

Ion interactions may feature dependent on the stationary phase and column pH, e.g., at a neutral pH silanol groups present in a C18 column can carry a negative charge but at low pH they are protonated and neutral (see below, Mechanisms of separation – Reversed phase chromatography) [188].

π-π interaction: Van der Waals forces between unsaturated aromatic compounds, e.g., with a phenyl-hexyl stationary phase utilised in reversed phase chromatography (see Section 1.3.1.3.).

Similar to metabolomics, this initial broad definition belies the complexity and wide-ranging applications of chromatography that have developed since its initial inception and growth over the course of the 20th century. Although liquid chromatography is the most widely applied today in chemical analysis (and metabolomics), there are other types including gas chromatography [119, 121, 189], thin-layer chromatography [190, 191], ion exchange chromatography [192] and supercritical fluid chromatography [193, 194]. Capillary electrophoresis, although not defined theoretically as chromatography, is a component separation technique which can be coupled to mass spectrometry and is applied in metabolomics studies. As powerful separation techniques, chromatography has been 'hyphenated' to numerous analytical (detection) technologies such as spectroscopic techniques [195] and perhaps more well known, mass spectrometry.

UHPLC (Figure 1.7) was commercially introduced in 2005 by Waters Ltd and helped to revolutionise the use of liquid chromatography coupled to MS in metabolomics studies. UHPLC operates with smaller chromatography particle sizes (typically <2.5µm compared to 5µm for the traditional HPLC columns) and along with higher flow rates provides higher chromatographic resolution and higher sensitivity compared to traditional HPLC columns; for example, chromatographic peak widths are typically 3-10 seconds wide at baseline compared to typically greater than 30 seconds for HPLC. Before UHPLC-MS was introduced commercially, metabolomics studies predominantly applied GC-MS and NMR spectroscopy. As UHPLC-MS is the most widely applied metabolomics platform the following sections will focus on the two most frequently used UHPLC assays, reversed phase chromatography (RPC) and hydrophilic interaction chromatography (HILIC). Of importance to note is that RPC and HILIC assays are complementary to each other and both are typically applied together in metabolomics studies to allow the study of water-soluble metabolites (HILIC) and lipids (RPC)

for most sample types. Sometimes an RPC assay using water and methanol or acetonitrile is applied instead of HILIC assays, especially for the study of urine.

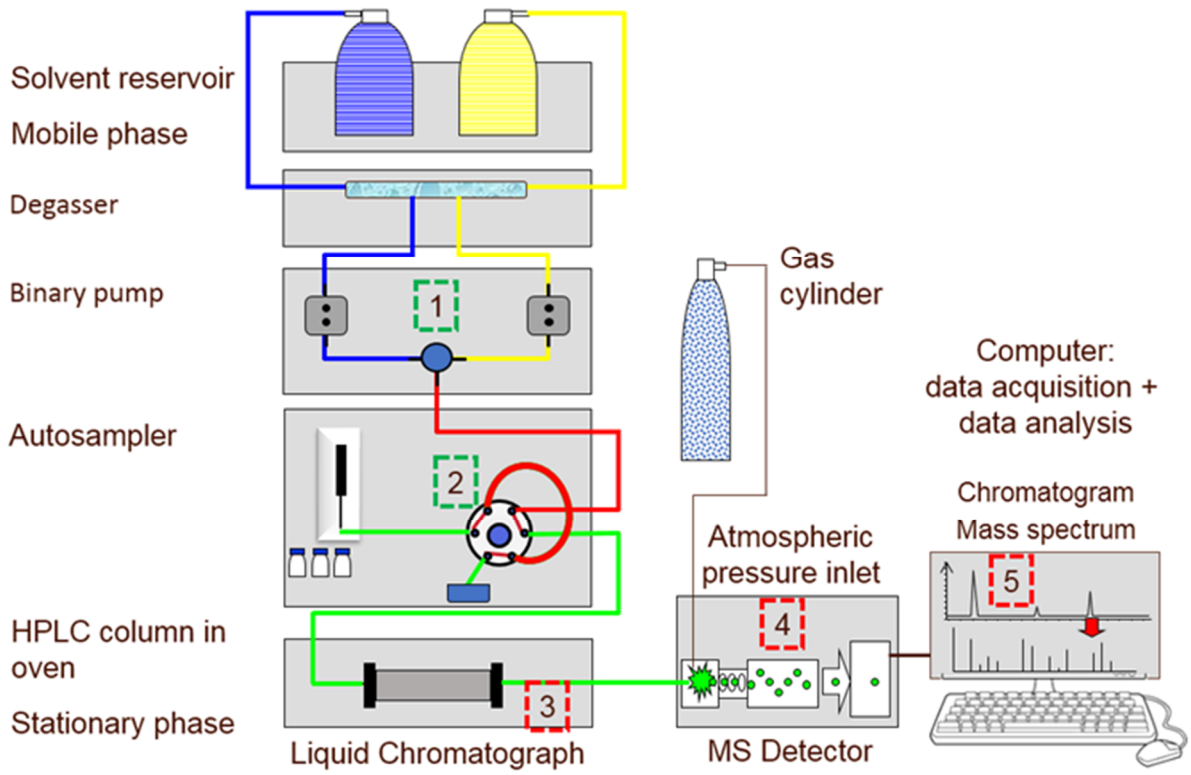


Figure 1.7: A simplified schematic of a UHPLC system interfacing with a mass spectrometer, reproduced from [196]. (1) The required ratio of solvent is drawn from the reservoir by the binary pump system prior to being mixed. This ratio is user defined and can change over the course of the UHPLC. (2) The autosampler aspirates the user defined volume of sample then injects the sample volume into the now mixed mobile phase via a switching valve. (3) The sample and mobile phase solution is pumped through the chromatography column, housed in a temperature-controlled unit, wherein analytes will experience varying degrees of retention and be separated. (4) The eluted solution from the chromatographic column enters the mass spectrometer (see Section 1.3.3. for further details) detector of choice and (5) chromatograms and mass spectra are collected from the connecting workstation terminal.

1.3.1.2. Mechanisms of separation – HILIC

The stationary phase is considered the most influential factor in determining selectivity in HILIC separation, e.g., silica based stationary phases show lower hydrophilic selectivity than polymer based stationary phases [197]. HILIC stationary phases are typically bare silica or silica derivatised with different polar groups (e.g., amide, diol) [198], the surface may be acidic, neutral, basic dependent on the specific structure of the stationary phase e.g., acidic silica to encourage ion exchange interactions [187] or stronger anionic/cationic retention of analytes under low/high pH mobile phase conditions [199]. The two mobile phases applied in HILIC assays are typically acetonitrile and water with modifiers included in each mobile phase (e.g., ammonium formate or ammonium acetate or formic acid). For example, the gradient elution applied starts with 95% acetonitrile and ends with a composition of 50/50 acetonitrile/water. If a larger water content is applied then the adsorbed water layer, discussed below, can break apart and reduce analyte absorption. The primary retention mechanism in HILIC is a phenomenon referred to as partitioning. With a high ratio of organic solvent (acetonitrile) and a small ratio of water in the mobile phase, water interacts more strongly with the polar stationary phase of the HILIC column (e.g., bare silica with silanol functional groups or amine and amide functional groups) [197, 198]. Thus, interactions between water, as opposed to acetonitrile, and silanol groups are favoured resulting in water adsorbing onto the silanol groups and forming an adsorbed water layer. [198]. The thickness of the adsorbed water layer can be influenced by pH and mobile phase modifiers which may alter the charge state of the stationary phase affecting its interaction with water molecules [197] and, somewhat counterintuitively, the water layer formed is larger with a higher organic acetonitrile:H_2O ratio in the mobile phase [198]. Partitioning is described by the differential distribution of an analyte between the adsorbed water layer and organic rich mobile phase.

The more hydrophilic the analyte the greater the effect of partitioning, favouring analyte interaction and retention with the adsorbed water layer [197, 198]. Separation under HILIC conditions is deemed multimodal (i.e., multiple retention mechanisms) with some examples shown in Figure 1.6 [198] and may relate to the poor reproducibility in elution and modelling of retention times [197]. Secondary mechanisms may include hydrogen donor interactions between neutral polar molecules, weak electrostatic interactions dependent on pH or under high organic conditions and hydrogen bonding [197, 198]. Thus, these secondary interactions are partially dependent on the chemical structure of the stationary phase and analytes and the mobile phase employed [187].

A 'hierarchy' of mobile phase features that affect selectivity are mobile phase pH, organic solvent and modifier type and concentration, as detailed by Guo in their review [197]. Mobile phase pH is considered the most important parameter as it can influence the charge state of ionisable analytes and charges present on the stationary phase surface which may alter the interactions between analytes and the stationary phase. Modifier type and concentration was deemed to have the least influence on selectivity but is potentially a significant factor for retention as it may influence the thickness of the adhered water layer on the stationary phase through modulation of electrostatic interactions between the stationary and mobile phases and will determine mobile phase pH. As the primary mechanism of separation in a HILIC column relies on partitioning into a water adsorbed layer it is an ideal assay to investigate water-soluble metabolites [200-202], polar metabolites [198, 203], including polar lipids [204].

1.3.1.3. Mechanisms of separation – Reversed phase chromatography

The most frequently applied stationary phase is a C_{18} reversed phase which is a C_{18} alkane chain chemically bonded to the stationary phase particle (silica or polymer). However, other types of reversed phase stationary phases are used including C_8, C_{30} and biphenyl columns. Unlike HILIC assays, RPC assays can use a range of different mobile phases. For water-soluble metabolites water is used as a weak eluting solvent and either methanol or acetonitrile are used as the strong eluting solvent with the gradient elution operating from high percentage water/low percentage organic solvent to high percentage organic solvent/low percentage water. The study of lipids normally applies a water/organic solvent mixture as the weak eluting solvent and isopropanol as the strong eluting solvent. RPC is the most commonly applied chromatographic method within analytical chemistry, as such, there are a wide range of columns available, similar to HILIC, that exhibit large differences in selectivity [184]. The separation mechanisms of RPC are still subject to further investigation and review as evidenced by the recent publications on the subject matter. There has been some contention on the primary mechanism of separation and retention in RPC. One mechanism relies on partitioning, not to be mistaken with the partitioning layer of water within HILIC. In this context partitioning refers to analytes immersing themselves within but not adsorbing to the bonded phase, for example a C_{18} alkyl chain bonded to a silica anchor which is the most commonly used RPC column [184]. The other mechanism occurs at the bonded phase/solvent interface with adsorption of analytes directly onto the bonded phase [184]. With alkyl chains as the stationary phase, organic modifiers within the mobile phase may adsorb onto the ends of the alkyl chains creating an enriched organic region separating the bonded phase region to the 'water-rich' mobile phase [205]. These three distinct regions contribute to retention of semi-polar or non-polar analytes either through: unfavourable (hydrophobic) interactions of

the analyte with the hydro-organic mobile phase (Solvophobic theory) and/or driven by favourable (lipophilic) interactions between the analyte and bonded phase (Revisionist theory) [184, 205, 206]. The mechanism that contributes greatest to retention is dependent upon the analyte and the stationary phase. For example, on C_8 columns non-polar analytes are primarily retained through adsorption while with C_{18} columns partitioning and adsorption are the primary mechanisms, yet for non-polar analytes adsorption is the primary mechanism of retention regardless of alkyl chain length (C_8 or C_{18}) [184].

Despite ongoing discussion on the underlying theories that govern RPC, the selectivity of different stationary phases is well understood resulting in a large commercial market of RPC columns with numerous applications [184]. Like HILIC, the stationary phase is an important factor in the selectivity of RPC. For example, alkyl chains (C_4, C_8, C_{18}) bonded to silica are primarily employed to separate and retain: hydrophobic compounds, non-polar/polar lipids via the mechanisms described above [184]. However, phenyl-hexyl bound to silica exhibit stronger π-π (non-covalent) interactions with aromatics which facilities their retention, but an altered bonded phase-mobile phase interface attenuates the retention of other analyte classes including aromatic-containing analytes [184, 207]. It has also been shown that analyte retention is greater with an increase of the hydrophobicity of the solutes, hydrophobicity of the stationary phase surface and polarity of the mobile phase [184]. Retention is also dependent on other parameters such as temperature, pH, pressure, surface heterogeneity, and ionic strength of the mobile phase [205]. Despite favouring retention of hydrophobic metabolites RPC has been applied to urine and plasma to investigate water-soluble metabolites with low polarity such as short chain fatty acids [208]. As RPC favours the retention of semi-polar and non-polar metabolites in contrast to HILIC favouring retention of polar metabolites the two assays offer complementary coverage of the metabolome when

applied in tandem [209, 210]. A RPC assay utilising a C18 column, the most used and optimal choice for investigation/retention of lipids/non-polar molecules [184], and a HILIC assay utilising a hydrophilic stationary phase, such as amide facilitating fast and efficient retention of highly polar metabolites [211], would provide broad coverage for the sample of choice making it an ideal combination for untargeted metabolomic workflows. This tandem approach has been applied to investigate water-soluble metabolites in urine [212], in addition to faeces blood and saliva [213].

1.3.2. Forgoing chromatography – alternative approaches

Chromatographic separation is not a mandatory step of utilising MS analysis with some advantages offered from purposefully mitigating the process. Such methods, known as direct infusion mass spectrometry (DI-MS) or flow injection analysis mass spectrometry (FIA-MS) do not apply a separation step such as chromatography. As a consequence, ions of the same mass/chemical formula are detected as a single ion (metabolite peak) which cannot differentiate between stereo/structural isomers [214, 215]. The lack of chromatographic separation has additional disadvantages such as increased ion suppression, i.e., analytes 'competing' for charge within the ionisation source resulting in reduced signals for less volatile metabolites [177, 215]. However, the lack of chromatography facilitates ease of sample preparation allowing for faster analysis and can be more easily automated with robotic systems to create a high throughput system for large sample sets [177]. Standalone MS platforms have been applied to study human tissue and blood samples [177] human urine [214] and saliva [216]; furthermore they have been applied in other fields such as quality control and counterfeit detection in medications and supplements, analysis and detection of

pesticides [215] and is one of the primary platforms utilised by commercial kits such as the Biocrates AbsoluteIDQ p180 kit (Biocrates, Life Science AG, Innsbruck, Austria).

1.3.3. Mass spectrometry

1.3.3.1. General concepts

Mass spectrometry is an analytical technique capable of qualitatively or quantitatively measuring the abundance of molecules such as metabolites, peptides or proteins which are ionised and detected as ions related to the structurally intact or fragmented analyte [217]. Mass spectrometers do not measure mass directly, but return a measurement known as mass-to-charge ratio (*m/z*) [79]. For example, protonated tryptophan, represented as $[M+H]^+$ (mass of compound + mass of hydrogen), may have a measured *m/z* of 205.0976 (205.0976 / 1) whilst doubly protonated tryptophan $[M+2H]^{2+}$ may have a measured *m/z* of 103.0527 (206.1054 / 2). In either instance the mass of tryptophan can be inferred if the adduct is known, in this instance the addition of hydrogen, yielding a mass of ~204.0898 Da which is the monoisotopic mass of tryptophan. Due to the small size of metabolites they typically form singly charged ions [79], however, adduct formation is dependent on the method of ionisation and type of hyphenated technique used (e.g., UHPLC) [218-221].

An example of a mass spectrometry set up is shown below (Q Exactive Plus Orbitrap MS, Figure 1.8) with a detailed breakdown of key components provided in the proceeding sections and a brief overview, informed by the manufacturer's user manual [222], given here. Once a sample has been introduced into the mass spectrometer, either via direct injection or post chromatography (Figure 1.7), the sample is ionised within the source (see Section 1.3.3.2. for further details). The resulting ions can be manipulated by electromagnetic fields as they enter the mass spectrometer, which operates under a vacuum below 1×10-6 ba, via the capillary.

The mass spectrometer consists of several 'ion optic' devices (lenses and -poles) whose purpose is to focus and redirect the ion beam (see Section 1.3.3.4. for further details on quadrupoles). Any neutral molecules (e.g., gas, solvent) will not be re-directed by the bent flatapole whilst charged ions will undergo a 90^{o} turn into the first mass analyser/filter (e.g., Transfer multipole). The ions then enter the C-trap (see Section 1.3.3.5. for further details), a bent quadrupole, that ejects discrete packets of ions into the terminal mass analyser (Orbitrap, see Section 1.3.3.5. for further details). Alternatively, ions may be subject to fragmentation (see Section 1.3.3.6. for further details) prior to mass analysis in the Orbitrap by transferring from the C-trap to the higher-energy C-trap dissociation (HCD) cell before re-entering the C-trap and being ejected into the Orbitrap. Ions of the same mass form rings which both orbit around the Orbitrap spindle and oscillate along its axis. These oscillations induce an electric current that a detection system can interpret into relative intensity of the signal and *m/z*.

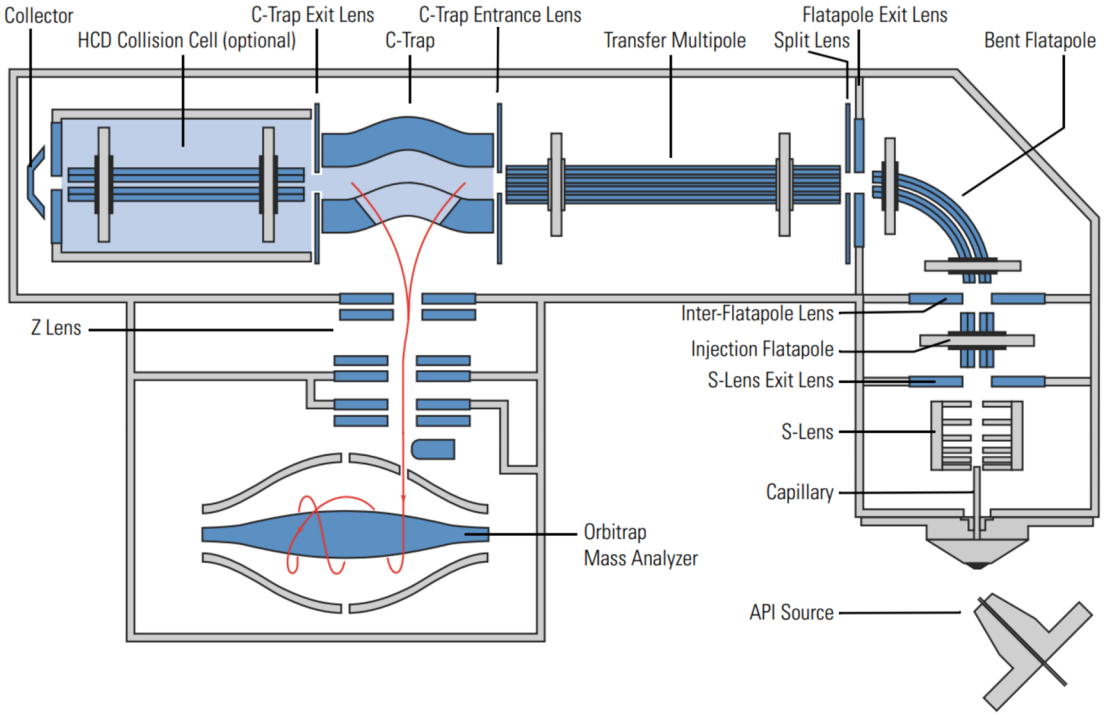


Figure 1.8: A schematic representing the assembly of a Q Exactive Plus Orbitrap mass spectrometer coupled to an atmospheric pressure electrospray ionisation source (see Section 1.3.3.2.) reproduced from Thermo Fisher [223]. The device contains a higher energy collision dissociation (HCD) cell, which functions on the same principles of collision induced dissociation, for MS/MS applications (see Section 1.3.3.6.). The mass analyser is an Orbitrap system (see Section 1.3.3.5.) which is fed ion packets from a C-trap (see Section 1.3.3.5.).

1.3.3.2. Electrospray ionisation

Ionisation of molecules is a vital stage within MS as charged particles are responsive to electrical, RF and magnetic fields allowing them to be accelerated, reflected, spatially isolated (trapped) or spatially separated based on their *m/z*. [224-227]. Metabolites are not equally ionisable and may favour a specific method of ionisation such as 'spray-based' atmospheric electrospray ionisation (ESI) techniques with either positive/negative polarity or 'laser-based' matrix-assisted laser desorption ionisation (MALDI) [118, 217, 228]. As previously mentioned, the employed method of ionisation can result in multiple adducts and so a single metabolite may have several different *m/z* values recorded by a mass spectrometer which can complicate metabolite identification (see Section 1.5.2.). Examples of ionisation methods include but are not limited to: 'laser based' ionisation methods such as MALDI and its variants — liquid atmospheric pressure-MALDI, matrix-assisted laser desorption electrospray ionisation; non-laser or spray based methods such as — rapid evaporative ionisation, matrix-assisted ionisation vacuum and 'spray based' ionisation methods such as — desorption electrospray ionisation (DESI), nano-DESI, or ESI, with the aforementioned methods reviewed elsewhere [218, 228]. Use of laser-based ionisation methods, i.e., MALDI and its variations, see prominent use in the field of proteomics but less so in metabolomics and so these methods are not discussed further; instead, the proceeding chapters focus on ESI as it is the most prominent form of ionisation employed alongside UHPLC-MS systems. The underlying mechanisms of ESI are detailed by [229, 230] and inform the summary below and in Figure 1.9. In a typical ESI setup (Figure 1.9) a liquid solvent, such as the mobile phase from a UHPLC system, is passed to a capillary (spray capillary). Dependent upon the polarity of the ESI source, the spray capillary is attached to a positive (positive ion mode) or negative (negative ion mode) terminal directed towards an oppositely charged counter electrode situated

around an orifice leading to the vacuum region of the mass spectrometer. There is a large potential difference (~2-3kV) applied between the capillary and counter electrode supplied by a direct current. The chamber in which the capillary and counter electrode occupy (ESI chamber) operates at atmospheric pressure as the ambient gas assists with solvent evaporation (see below) and maintaining a vacuum within the chamber would be difficult due to the high flow of liquid ejected from the capillary. Further, the mass spectrometer operates at a high vacuum pressure creating a favourable pressure differential between the ESI chamber and MS which assists entry of gas phase ions into the mass spectrometer. The electric field, as applied to the spray capillary, is strongest at its tip and causes polarisation of the liquid solvent. Under positive ion mode this results in the accumulation of positive ions at the meniscus of the liquid solvent forming a Taylor cone which possess an elliptical shape at the spray capillary aperture. The ions present depend on the sample and UHPLC assay applied, for example, NH_4^+ ions may accumulate in positive ion mode due to the presence of ammonium acetate additives in UHPLC mobile phases along with other potential ion species [220, 221]. Two opposing forces influence the shape of the meniscus, surface tension acts to retract the liquid to within the spray capillary whilst electrostatic forces such as attraction of ions at the meniscus to the oppositely charged counter electrode and liquid flow draw the meniscus out of the spray capillary. From the Taylor cone is emitted a fine jet of droplets which contain like charged ions. As these droplets are in flight the solvent starts to evaporate due to thermal energy provided by the ambient gas within the ESI chamber reducing the volume and diameter of the droplet. This reduced volume results in an increased field density at the droplet surface, i.e., increased localised charge, until the repulsion between like charged ions overcomes the surface tension of the droplet. The point at which this occurs is known as the Rayleigh limit which defines the theoretical maximum charge that a droplet of

known electrical permittivity, surface tension and radius can possess. Upon approaching/reaching the Rayleigh limit additional Taylor cones form on the droplet emitting secondary droplets of smaller size in a process referred to as coulombic fission (explosion). This process repeats itself with the secondary droplets and so on until individual ions, no longer in solvent but in the gas phase, enter the mass spectrometer. ESI is considered a sensitive enough method of ionisation that very low concentrations of analyte (10^{-7} – 10^{-3} mol/L) can be present in the liquid solvent but still be successfully ionised and enter the mass spectrometer. The ability for ESI to ionise a range of biologically relevant molecules at low concentrations and be easily coupled to UHPLC systems has resulted in its popularity as the standard ionisation method of choice for UHPLC-MS based metabolomic applications.

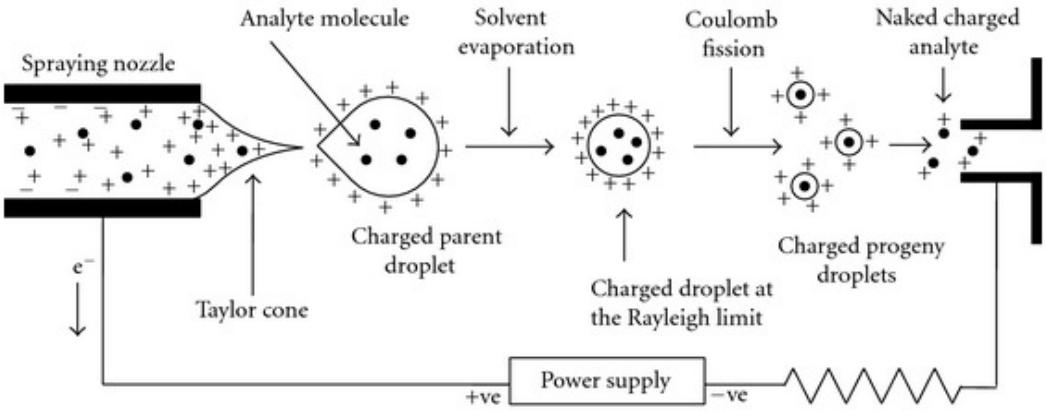


Figure 1.9: A schematic illustrating ion formation in an ESI source, reproduced from [231].

1.3.3.3. Mass analysers & ion selection

Post ionisation, analyte ions are transferred to a mass analyser and detector. Keeping in theme with mass spectrometry there are a range of methods and technologies that can be used in tandem or isolation with one another that serve to filter(select)/fragment(MS/MS)/trap and measure the mass of these ions. Types of mass analysers include: the quadrupole and its variants (multipoles) [232-234], linear ion traps [232-234] alternative ion traps such as the Fourier transform ion cyclotron resonance (FTICR) [233, 235] and Orbitraps [226, 233] as well as Time of Flight (ToF) [224, 233] to name a few, with cited literature summarising their application or mechanisms. Mass spectrometers can be constructed with multiple mass analysers, known as hybrid instruments. These include the triple quadrupole mass spectrometer applied for targeted assays and the combination of low mass resolution quadrupole mass analysers with high mass resolution ToF or Orbitrap mass analysers for untargeted metabolomics assays. Due to the high mass resolution, mass accuracy and sensitivity offered by ToF and Orbitrap mass spectrometers and their successful coupling to UHPLC platforms utilising ESI sources they have become favoured mass analysers applied in metabolomics [226, 236, 237]. Just as different ion sources have their own advantages and disadvantages so too do mass analysers. Metrics used to report mass analyser capability include mass accuracy, defined as the ratio of the *m/z* measurement error to the true *m/z* measured in parts per million (ppm) and resolving power defined as the ability to resolve (differentiate) between two similar measured *m/z* values (peaks) [227]. Detected *m/z* peaks have definable characteristics with regards to their shape (width, maximum height) discussed in greater detail below (see Section 1.4.1.), thus, resolving power is typically quoted with reference to these characteristics [238]. One such metric for resolving power is full width at half maximum (FWHM) calculated as $m/\Delta m$ where m is the mass and Δm the peak width

(*m/z* range) at half peak intensity [238]. If the resolution of a mass analyser is too low, adjacent *m/z* peaks may not be adequately resolved which can negatively impact mass accuracy and result in a reduced ability to infer a chemical formula and chemical identity (annotation) for observed *m/z* peaks [227]. Accurate mass measurements are therefore reliant on sufficient mass resolution. However, the mass resolution of a mass analyser can be dependent on the *m/z* ratio of the molecule being analysed. For example, within Fourier transform mass spectrometers (FTMS), e.g., cyclotron and Orbitrap mass analysers discussed more below (see Section 1.3.3.5.). Therefore, resolution is typically quoted for specific *m/z* values [227]. Due to their wide application in metabolomics and relevance to further chapters only the theory and application of Orbitrap mass analysers, in conjunction with quadrupoles and variants, are discussed further.

1.3.3.4. Quadrupole mass analysers

Whilst quadrupole mass analysers represent the lowest mass resolution analysers, capable of selecting for ions of nominal mass but unable to acquire an accurate measure of *m/z*, they remain one of the most widely used mass analysers [234]. A linear quadrupole mass analyser (quadrupole mass filter) consists of four rods, either circular or hyperbolic in shape, arranged equidistance and parallel to one another in a diamond formation. These four rods apply a static electric field (direct current (DC) potential) overlayed with an alternating radiofrequency (RF) field. Opposing rods possess equal polarity and magnitude of the RF field with each pair of rod's static field being of the opposing polarity and the RF field being 180^{o} out of phase (Figure 1.10) such that one pair of rods is positively charged and the other negatively charged at any given moment with the polarity of the rods switching at a user defined frequency [233, 234]. Ions travelling through the quadrupole are subject to the sum

of forces from both the static and alternating fields. However, velocity is maintained within the z axis (through the quadrupole) as the static/alternating fields are perpendicular to the direction of motion in the x and y field. Therefore, velocity is subject to change in the x and y planes only [233].

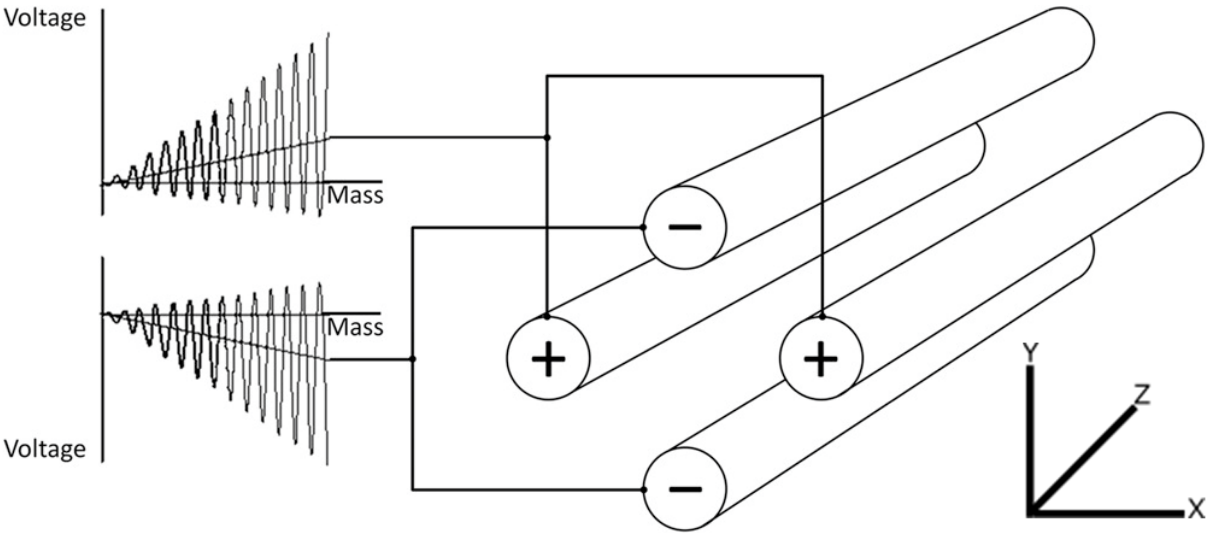


Figure 1.10: A schematic illustrating the orientation and the opposing composite electric field maintained between pairs of quadrupoles, reproduced with addition of the axes from [234]. Dependent on the assigned frequency of the oscillating field the quadrupole pairs will alternate between positive and negative polarity. Ions with a stable (corkscrew) trajectory travel through the quadrupole along the z axis.

The trajectory of ions in the x/y plane of a quadrupole are defined by the Mathieu and Paul equations which relates the relative position of an ion in the x/y plane to time, the static and RF fields and the *m/z* of the ion in question [233, 234]. These equations can be applied to display areas of stability within the x/y planes i.e., areas in which ions of a given *m/z* will have a stable trajectory and pass through the quadrupole. Modulation of the electric fields can discriminate which *m/z* will maintain a stable trajectory and pass through the quadrupole and those that have an unstable trajectory and will not pass through, colliding with and discharging on the quadrupole rods [233]. By exploiting this principle, the quadrupole is able

to function as a mass filter to transmit specific *m/z* packets through or to operate (typically in RF only mode) to focus and transmit all *m/z* packets. Finally, multiple quadrupoles can be applied in succession, as is observed in triple quadrupole mass spectrometers (QqQ) wherein one mode of operation it to have the first quadrupole (Q1) select a precursor ion based on its *m/z* which is user defined, the second quadrupole (q2) (RF mode) is a CID cell that fragments the selected ion and allows all fragments to enter quadrupole (Q3) which then allows all or a subset of the fragment ions produced by the CID to pass through the quadrupole to an ion detector [233, 234] as illustrated in Figure 1.11 (for further details on fragmentation see Section 1.3.3.6.).

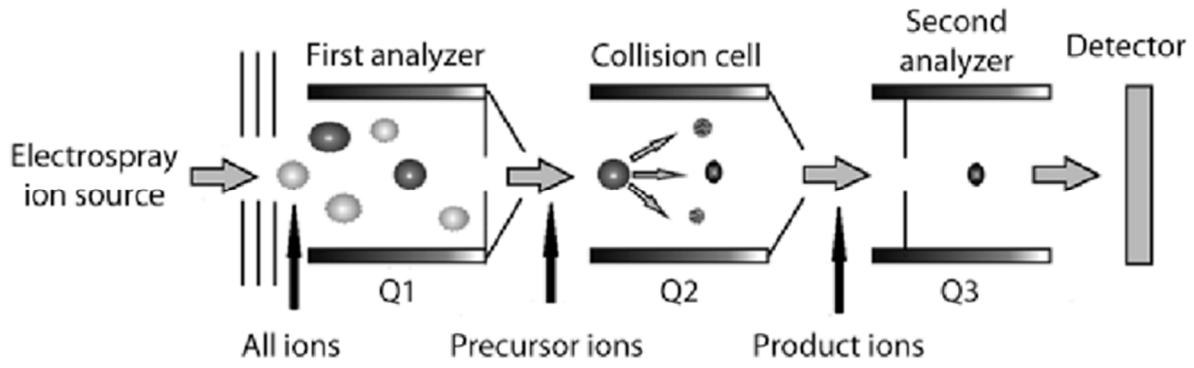


Figure 1.11: A schematic illustrating the QqQ operation of a triple quadrupole, reproduced from [239]. Q1/Q3 operate as mass filters to permit ions of a specified *m/z* range to pass through the quadrupole whilst q2 is operating in RF mode to function as a collection cell to induce fragmentation of the precursor ions.

1.3.3.5. Orbitrap mass analysers

Orbitrap mass analysers provide high mass resolution capabilities and offer different mass resolutions on the same instrument compared to other mass analysers which offer just one mass resolution. Orbitrap mass analysers are typically operated at mass resolutions greater than 60,000 (FWHM at *m/z* 200), dependent on the hybrid instrument used, and exceed the mass resolutions available on most ToF instruments. A quadrupole may also be applied as a mass filter, to select for ions with a specific *m/z* prior to fragmentation (see Section 1.3.3.6.) or act as an ion focussing device to assist in effective transmission of ions to the Orbitrap [232].

The theory and applications of Orbitrap mass analysers are detailed in the literature [225-227] and summarised below. An Orbitrap requires ions to be delivered in discrete pulses whilst UHPLC coupled to an ESI source produces a continuous stream of ions. Prior to injection into the Orbitrap mass analyser, ions are directed into a C-trap which is a bent quadrupole operating in RF-only mode and filled with a bath gas (typically helium). The purpose of the bath gas is to induce collisional cooling of the ions to reduce their kinetic energy. Once the ions are cooled the RF voltage of the C-trap is reduced (ramped down) prior to a high voltage pulse being applied across the C-trap. This pulse ejects the ions as a small ion packet orthogonally to the curved axis of the C-trap directing them to dedicated bent ion optics that focus the ions on the aperture into the Orbitrap mass analyser (as illustrated in Figure 1.12). Orbitraps consist of three electrodes exerting two independent electric fields (Figure 1.12). The central electrode is spindle shaped and exerts a radial electric field of opposing polarity to the ions that induces their stable circular motion about the spindle. The outer electrodes exert a field that induces motion in the axial plane (along the central electrode as opposed to

around it). Ions are purposefully injected into the Orbitrap mass analyser away from the spindle electrode's centre (equator) as the axial field strength at this point is zero. The axial field strength increases uniformly in opposing directions away from the equator of the spindle. By opting to inject ions away from the equator they experience a force from the axial electric field and oscillate across the axial plane of the spindle. Ions decelerate as they approach the 'poles' of the spindle, are repelled and accelerated back across the equator only to decelerate and have their direction reversed at the opposing end. These oscillations are mass-dependent harmonic oscillations and are independent of initial velocities and (entry) coordinates of the ions. The frequency of axial oscillations is a reliable measure to infer *m/z*. Ions of the same *m/z* form tight packets, or rings, orbiting the central spindle. These oscillations are defined by the following equation wherein the angular frequency, measured in radians per second, equates to the square root of an ion's charge divided by its mass multiplied by a constant k (proportional to the potential difference between central and outer electrodes).

Ω (angular frequency) = Square root(k/(charge/mass))

Ions with a lower *m/z* exhibit a higher harmonic frequency relative to higher *m/z* ions. Individual rings of oscillating ions induce an electric current, of a frequency matching the axial oscillations, detected by an amplifier. These detected signals are in the time domain and are interpreted as relative intensity against time for all oscillating ions within the Orbitrap. A fast Fourier transform is applied to this complex waveform to convert the data to the frequency domain and yield the individual frequencies of each ion ring. The data can now be plotted as intensity against frequency or by using the above equation to calculate *m/z* from the frequency, intensity against *m/z* can be used to plot a mass spectrum. An important limitation

to note about the Orbitrap is that the resolving power is proportional to the frequency of axial oscillations of ions and acquisition time or inversely proportional to the square root of the *m/z*. Therefore, lower *m/z* ions with a higher frequency observed within a fixed acquisition time benefit from increased resolution. This limitation is not shared by other mass analysers, such as TOFs, nor does it strongly hinder metabolomics studies which are focussed on small molecules of <1500 Da in size. Also, the longer the acquisition time the higher the mass resolution observed; this allows the Orbitrap mass spectrometers to be operated at different mass resolutions and the resolution to be applied can be chosen by the user. Most current tribrid instruments can operate at mass resolutions of 500,000 – 1,000,000. The mass resolution is preferred to be as high as possible but the number of data points across a chromatographic peak also has to be considered so that the peak is accurately plotted and so mass resolutions of 60,000 – 120,000 are typically applied in metabolomics applications.

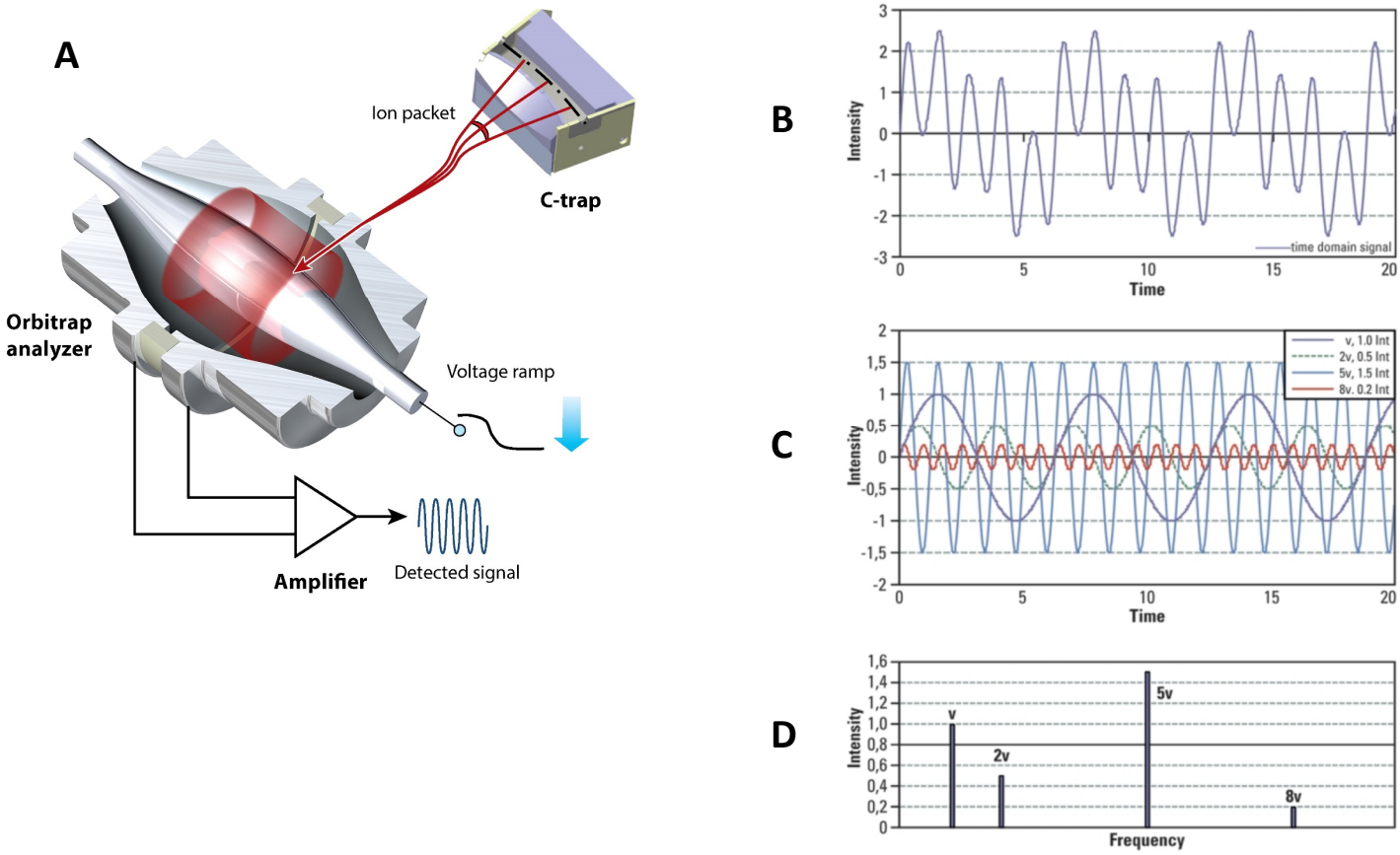


Figure 1.12: Cross section of the C-trap ion accumulation device and the Orbitrap mass analyser with an example of an ion trajectory (A), reproduced from [226], and example traces for a detected signal and calculated frequency (B-D), reproduced from [227]. The detected signal is in the time domain (B) and is a composite signal of multiple detected waveforms of varying frequencies, which results from the ion packets of differing *m/z* (C), these frequencies can be calculated via a reverse Fourier transform of the composite signal (D).

1.3.3.6. Tandem mass spectrometry

High resolution mass spectrometry is capable of measuring the *m/z* of an ion to high enough accuracy to elucidate a metabolite's chemical formula, which is the first step in annotating (putatively identifying) detected analytes, discussed further below (see Section 1.5.). However, a chemical formula provides insufficient information to derive the chemical structure of an analyte for all but the simplest of metabolites. Tandem mass spectrometry (MS/MS) is a technique that is applied to acquire greater structural information through fragmentation of the precursor ion (intact metabolite) and measuring the mass of its product ions (fragment ions) and inferring their chemical formula [83]. A commonly employed method of fragmentation is collision induced dissociation (CID) which subjects precursor ions to an electric potential (2-10 keV High energy, 1-200 eV low energy) in a specialised cell containing a neutral inert gas such as helium or argon [83, 240]. The kinetic energy imparted to the precursor ions by the electric potential induces ion acceleration and collisions with the inert gas during which some of the kinetic energy is transferred to the internal energy store of the precursor ion. As the ions are in a vacuum and so energy cannot be efficiently released by gas-ion collisions, internal energy is lost through bond breaking (fragmentation) and rearrangements. The type and efficiency of which product ions are produced is dependent on the collision energy and activation time (length of time the potential is applied) [83, 240, 241]. The MS/MS spectrum is dependent on the chemical structure, contains the relative intensities and *m/z* of product ions and is reproducible for a given metabolite assuming similar methods/parameters of fragmentation are applied. Thus, acquired spectra are akin to a 'mass spectrometric fingerprint' and can be compared to the spectra of chemical standards analysed using the same experimental methods and which are either available within public available mass spectral libraries, e.g., METLIN, or in 'in house' libraries to aid identification

[83]. However, should some product ions go undetected/unidentified or the precursor ion is a novel metabolite, a closely matching spectra may not be found within a mass spectral library and even the MS/MS data will be insufficient to enable the metabolite's structure to be defined [83]. As discussed above, in a triple quadrupole set up (QqQ) the second quadrupole acts as a CID cell and operates in RF mode so does not filter fragment ions based on *m/z*. Ions then pass through the third quadrupole which acts as a mass filter to select fragment(s) of a specific *m/z* to pass through to a detector. Within an Orbitrap device a separate collision cell (HCD cell) is used to perform the fragmentation with ion fragments passed back to the C-trap prior to ejection into the Orbitrap mass analyser (Figure 1.8).

1.4. Raw data processing

1.4.1. Deconvolution(pre-processing)

Data acquired by a mass spectrometer is multidimensional. For UHPLC-MS it is three-dimensional (*m/z*, (retention) time and response). One 'raw' data file is acquired for each sample analysed and so tens-to-hundreds of raw data files can be constructed. In their raw state these data files are incompatible to be used directly with multivariate and statistical methods of data analysis [242]. For UHPLC-MS raw data processing, otherwise known as deconvolution, is required to convert the 3D raw data matrix (RT vs *m/z* vs signal intensity) into a 2D matrix of metabolite features (i.e., unique *m/z* & retention time (RT) pairs) and signal intensity per sample [79, 243]. Several publicly available workflows including but not limited to MS-Dial [244, 245], Mzmine [246, 247] are available in addition to XCMS (extensible computational mass spectrometry) which is available both as an online platform and R package available via Bioconductor.org [242, 243]. Commercially available software from MS platform vendors are also available such as Compound Discoverer (Thermo Fisher Scientific),

MassHunter (Agilent) or Progenesis QI (Nonlinear dynamics, Waters Corporation, MA, USA). Due to its public availability and implementation within the R environment, which offers numerous freely available data handling/analysis tools, XCMS was the pre-processing software of choice employed in later chapters. Details below on the steps involved in pre-processing are therefore provided within the context of the XCMS software.

Prior to pre-processing via XCMS vendor specific file formats for MS files require conversion to an open format such as ".mzML"; this conversion can be achieved with freely available software such as Proteowizard [248, 249]. The pre-processing pipeline of XCMS can be subdivided as follows: peak detection (peak picking), peak alignment (RT correction), correspondence (peak grouping). Each step is described in detail by the following authors [242, 250-252] and an overview provided below. Peak detection (Figure 1.13), in XCMS first identifies regions of interest (ROI) within the chromatographic data. This is performed by specifying a narrow range in the *m/z* domain (e.g., 300.00-300.02) and identifying areas of high signal intensity along the retention time axis which are subdivided into scans (i.e., bins of a set time length). If a high signal intensity is detected for the *m/z* range across consecutive scans it is classified as a ROI. Experimentally measured *m/z* values will differ between consecutive scans but are expected to be distributed around the true value (forming a narrow Gaussian distribution). If the *m/z* values differ from the mean *m/z* of the ROI by less than the user defined ppm value then peaks are kept within the ROI, or else they are removed. Whether the ROI is considered a true peak and not noise is subject to passing a wavelet filter which is a second derivative Gaussian function, colloquially known as a 'Mexican hat' wavelet due to its shape. An extracted ion chromatogram (EIC, sum of intensities against time within a specified *m/z* range) for the region of interest has the wavelet filter applied, only the section of the ROI that falls within the wavelet is integrated (i.e., considered a true peak). The shape

of the Gaussian function as applied to the ROI results in peak tails not being integrated, these tails contribute a small proportion of the overall peak signal and may otherwise overlap with adjacent peaks. The width of the wavelet is user defined and, alongside ppm, is considered the most important parameter in accurate peak picking and integration. The integrated peak will have two definable parameters: peak height (maximum intensity of all datapoints that form the peak) and peak area (sum of all intensities for all datapoints). Peaks below a user defined signal to noise threshold and prefilter (minimum signal intensity to be integrated) are removed from the data, thus filtering out 'noise'. Anything that remains is a 'detected' peak which would have an associated *m/z* (median, minimum, and maximum) and retention time (median, minimum and maximum) and signal intensity. Dependent on the data and applied settings peaks may be artificially split or overlapping peaks may be incorporated as a single peak; however, XCMS affords the user numerous controls to manually optimise the workflow with software available (Isotopologue Parameter Optimization – IPO) to automate optimal parameter designations [253, 254]. These parameters also apply to proceeding steps of the XCMS pipeline such as peak alignment and retention time correction.

Peak alignment occurs on the integrated peaks as previously selected. Dependent on the user defined workflow, RT correction may be applied prior to or post peak alignment. A given compound subject to UHPLC-MS analysis may exhibit different RTs across analysed samples. RT correction is the process of normalising RT across samples, typically using QC samples as a point of reference, thus correcting any observed RT drift. This drift may not necessarily be a linear change so requires complex algorithms to correct such as Locally Estimated Scatter-plot Smoother (LOESS) [255-257] or ordered bijective interpolated warping (Obiwarp) [258]. Correspondence, or grouping, is the process of matching detected chromatographic peaks between samples across the retention time axis, within a user defined range ("bw" or

bandwidth), for a user defined *m/z* range (mzwid). Peaks that are within this *m/z* range and cluster at a given RT (+/- bw) are grouped and referred to as well-behaved peak groups. These grouped peaks constitute a metabolite feature, i.e., a unique RT and *m/z* pair observed across samples. Peaks are only integrated into a feature if the peaks are present in a user defined fraction of samples, e.g., minfrac = 0.4 only reports grouped peaks present across 40% or more of all samples as a feature. Finally, XCMS can export a matrix containing these features (normally thousands to low tens of thousands of features) and their measured intensities across samples as a simple 2D matrix ready for downstream processing and analysis e.g., normalisation, univariate analysis or multivariate analysis.

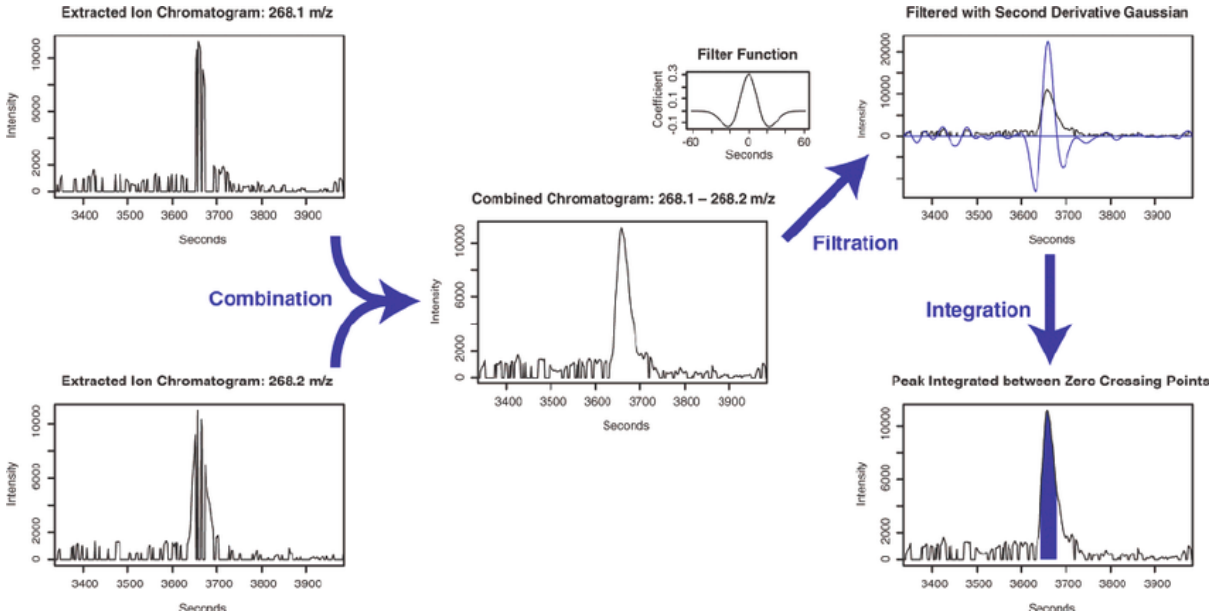


Figure 1.13: An illustration of peak grouping and integration within XCMS reproduced from [242]. Here, two *m/z* slices from a region of interest are overlaid to create a single combined peak prior to a second derivative gaussian function (wavelet filter) being applied to the peak. The area of the peak above the 0 (intensity) line and within the boundaries of the wavelet filter is integrated.

1.5. Metabolite annotation and identification

1.5.1. Levels of metabolite identification

In targeted metabolomics studies the metabolite's chemical structure is already known and data collection is optimised to allow confirmation of the chemical structure. In untargeted metabolomics studies data is collected to annotate or identify the metabolite rather than confirm its chemical structure and this is a more difficult process in untargeted metabolomics studies. For these reasons the terms annotation and identification are commonly used to communicate the confidence in a correct chemical structure being reported. Identification is defined by matching two or more datatypes for the experimental sample data to corresponding data from a chemical standard using the same analytical technique/method. For example, matching the MS/MS mass spectrum and retention time between a biological sample and chemical standard, applying the same experimental methods, provides a high level of confidence of the sample feature identity. Annotation defines matching one or more datatypes to a chemical standard data or theoretical data, though with these data are not necessarily being collected experimentally using the same analytical technique/method; for example, matching the experimental *m/z* to the monoisotopic mass of a metabolite in a metabolomics database. The Metabolomics Standards Initiative (MSI) has created a framework to more clearly communicate the confidence in a given metabolite annotation reported within the field [259, 260]. In broad terms, this system consists of four levels as follows:

- Level 1 – identified compounds with at least two matched independent orthogonal properties (e.g., accurate mass, RT, MS spectra) to a chemical reference

- Level 2 – putatively annotated compounds with at least two matched independent orthogonal properties to a reference database/spectral library
- Level 3 – putatively characterised compound classes compounds with at least one matched orthogonal property to a reference database/spectral library
- Level 4 – unknown compounds – reproducible *m/z* peaks that are otherwise unidentified

Different types of data or combinations of data can be used to annotate or identify a metabolite. For UHPLC-MS there are three types of data which can be collected and applied in the metabolite annotation workflow. The chromatographic retention time is collected for every metabolite feature for every sample and is dependent on the metabolite's physicochemical properties and the chromatographic properties (stationary phase, mobile phase etc). The *m/z* values of different metabolite features of the same metabolite can be applied to define the exact monoisotopic mass of the non-charged metabolite from which one or multiple molecular formulae can be calculated within a given mass error. These data are collected for every metabolite feature in every sample. The *m/z* values are dependent on the chemical structure and electrospray operating conditions. Thirdly, the MS/MS fragmentation mass spectrum for any metabolite feature can also be collected and this is dependent on the metabolite's chemical structure as well as the collision cell operating conditions. MS/MS data are not normally collected for all metabolite features. A fourth type of data that is sometimes used is ion mobility which is a rapid gas-phase separation technique with drift time or CCS being reported but is not discussed further as it is not relevant to subsequent chapters.

1.5.2. Application of m/z, RT, and MS/MS in metabolite annotation

The challenges of accurate metabolite annotation (identification) have been briefly touched upon in prior sections, to reiterate, the two major bottlenecks in accurate identification are the limited availability of metabolite standards relative to the total number of metabolites estimated to exist and incomplete metabolite databases (HMDB, LIPIDMAPS, etc.) which act as a vital reference. The challenges of metabolite annotation and available tools have been reviewed and discussed in the literature at length [118, 261-263] and are summarised below. Untargeted metabolomics workflows applying UHPLC-MS are reliant on metabolite databases to source metabolite annotations. Such databases allow for mass or *m/z* queries in which the measured accurate mass or *m/z* and potential adducts of a metabolite feature (a unique *m/z*, RT pair measured by MS) can be submitted and metabolites with a matching monoisotopic mass (within a user defined ppm range) and chemical formula are returned. A single metabolite, as measured in a typical UHPLC-MS assay, will correspond to numerous detected metabolite features of similar retention time (i.e., co-eluted) but differing *m/z* due to the presence of isotopes, adduct formation, multiply charged ions, neutral losses, dimerization to name but a few causes of this 'duplication' in detected signals. When observing a single metabolite feature in isolation it may not be known which adduct/losses/isotopes, etc. have occurred, introducing a degree of uncertainty even when measured at a high precision and accuracy. Thus, the number of annotations and formulae returned ('hits') from a database search may be greater if: allowing for a greater ppm tolerance between measured and database values, allowing for multiple adduct formations ([M+H], [M+NA], [M+K], [M+Mg] etc.), various neutral losses (e.g., H_2O, NH_3), and taking into account isotope abundance (e.g., probability of C^{12} being present vs C^{13}). Scrutinising *m/z* differences between co-eluting pairs of metabolite features can prove informative about the presence of specific

adduct/losses/isotopes, for example, two co-eluting metabolite features may differ in mass by 0.0034 which would indicate the presence of a single C^{13} in the heavier metabolite feature. In this fashion metabolite feature adduct/losses/isotopes can be deduced and these co-eluting metabolite features grouped together as they likely related to the same metabolite. As a collective group an accurate estimation of monoisotopic mass for the metabolite related to the metabolite features can be ascertained. Even if adducts/losses/isotopes are correctly deduced and a monoisotopic mass for the metabolite is obtained there can still be multiple putative annotations associated with the inferred mass. This is true even before considering structural isomers e.g., leucine and isoleucine, or chirality both of which MS1 data (accurate mass) is incapable of differentiating between whilst MS2 data (MS/MS fragmentation) has limited success in differentiating the former which can be supplemented with chromatography or ion mobility separation techniques [264]. Knowledge of the system being investigated may therefore be applied in an ad hoc approach to minimise the search space e.g., ignoring annotations relating to drugs and their derivatives if they are known to not be in the collected samples. Alternatively, there may be no database values that match the monoisotopic mass inferred from *m/z* values resulting in lack of annotation if no further action is taken. Retention time may be used to supplement the annotation process, however, this relies on the prior development of 'in-house' retention time libraries. Such libraries would be constructed by applying standardised UHPLC-MS assays to known chemical standards to ascertain the expected RT range of these metabolites on a particular UHPLC-MS system. Further, the assay(s) used to construct the RT library must be highly similar to those applied to the dataset requiring annotation as RT can vary greatly dependent on temperature, pH, mobile/stationary phase as discussed above in the chromatography section (see Section 1.3.1.). As a result, RT libraries are considered non-transferable which, coupled to the expense

of developing such libraries, make their wider application in annotation uncommon. Though efforts are being made by many in the field to apply machine learning algorithms to create retention time prediction models [263, 265-267] or focus on alternative metrics that can aid annotation such as 'retention order' i.e., the order in which different metabolite classes elute [268].

The more common approach to supplement annotation beyond database searches related to *m/z* is to apply MS/MS. As discussed above (see Section 1.3.3.6.) MS/MS spectra collected on the same platform under similar experiment platforms are reproducible between different labs, thus, public libraries have been constructed to store and disseminate these spectra. If a precursor ion has multiple possible annotations associated with its accurate mass then its MS/MS spectrum can be compared to that of metabolite standards of equal mass in a spectral library (e.g., METLIN) and assigned similarity scores to determine which annotation(s) is most likely (Figure 1.14). Spectral matching provides an additional orthogonal property of a metabolite feature to be gauged and provides greater confidence in its assigned annotation. However, the construction of spectral libraries is limited as reference spectra are derived from performing MS/MS on known chemical standards which, as mentioned previously, do not exist for the vast majority of metabolites detectable within humans alone. To supplement the databases *in silico* derived fragmentation spectra have been modelled based on predicting how a given molecular structure would likely fragment under specified experimental conditions. Even with this supplementation spectral libraries remain incomplete and so metabolite features may remain unannotated even when applying *m/z*, RT, and MS/MS parameters to derive a putative annotation list. Alternatively, should a metabolite feature have hits for both accurate mass and fragmentation spectra corresponding to a single annotation this does not necessarily correspond to a single metabolite nor is the annotation

definitive (see MSI classifications above). For instance, the diglyceride DG(40:2) has 5 known stereoisomers on the LIPIDMAPS database with differing fatty acids chain lengths and locations of C=C bonds on said chains. However, the commonly applied metrics of metabolite identification in untargeted studies: RT, accurate mass, MS/MS spectra, are incapable of discriminating which of these stereoisomers is present, thus, the annotation is provided at the species level (total carbon count of fatty acid chains : number of double bonds). The excessive prominence of unknowns in a dataset, or minimal confidence in their annotations, may limit the biological interpretations yielded from the dataset. Therefore, the success of an untargeted metabolomics analysis is reliant on successful annotation which requires the careful and diligent application of assistive software and community guided best practices.

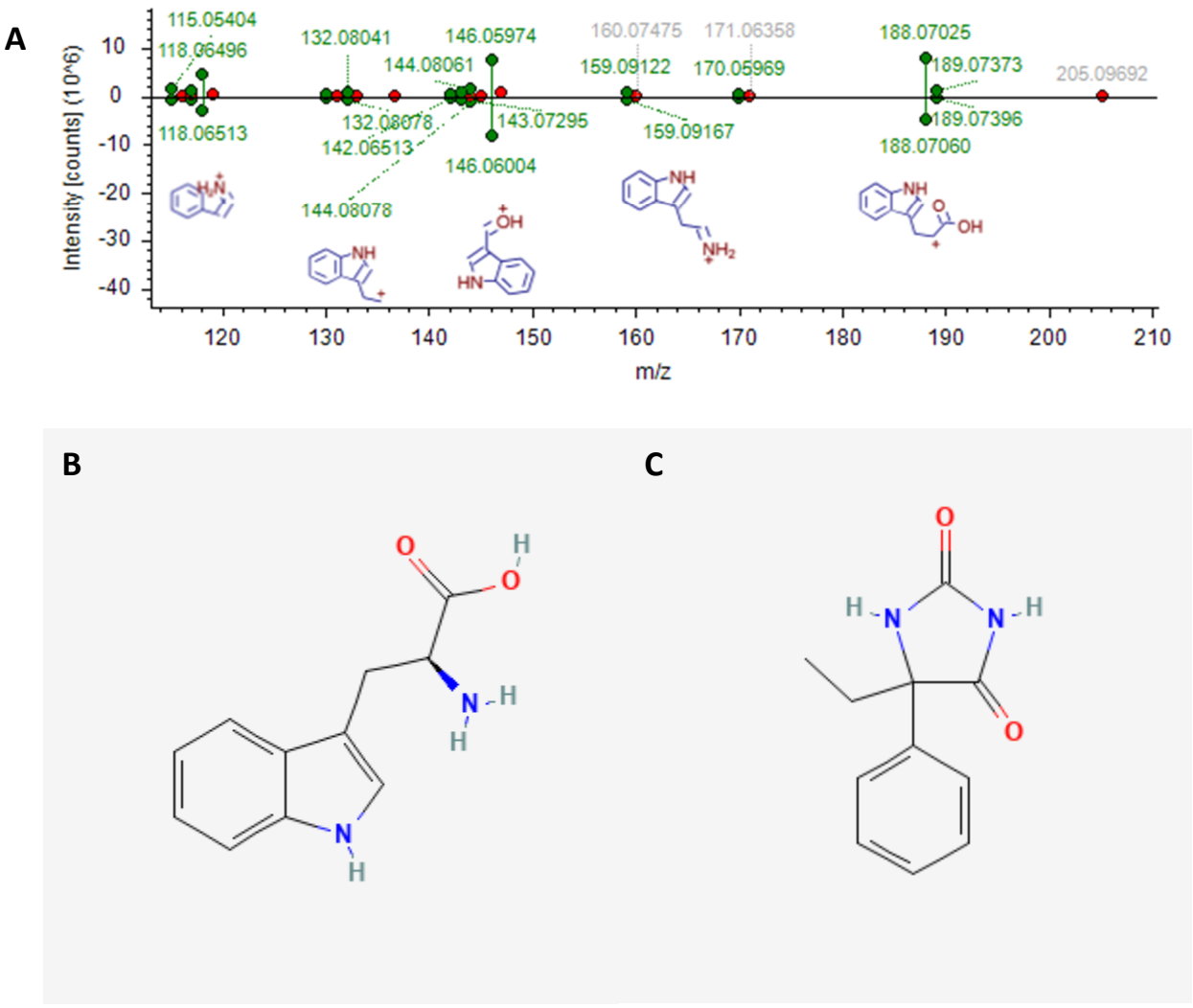


Figure 1.14: (A) An example MS/MS spectra (top) compared to a reference spectrum (bottom) for a precursor ion of *m/z* 205.09692 with multiple structural isomers (B-C, reproduced from PubChem). An ion of this mass could correspond to multiple structural isomers such as tryptophan (B) or nirvanol (C), thus cannot be putatively identified by accurate mass alone. Fragmentation spectra of the precursor ion closely matches that of reference spectra for tryptophan, evidenced by fragments matching the: conjoined ring structure of tryptophan, and the presence of COOH and NH2 functional groups. The closely matched spectra suggest the observed metabolite feature is most likely tryptophan and not nirvanol.

1.6. Conclusion

Investigating circadian regulation of the metabolome is an emerging field resulting in limited available data to understand how these two systems interact and limiting the ability to pose meaningful hypothesis; more studies are available with regards to diurnal regulation. What information is available in the literature for circadian studies is collected from male only/majority cohorts and in most circumstances provide only limited coverage of the metabolome due to the employment of targeted assays (see Chapter 2).

There is a need for a discovery-based study with broader coverage of the metabolome to extend our understanding of circadian regulation of metabolism, in healthy humans. Untargeted metabolomics datasets are large in size and difficult to handle and analyse, this complexity will only increase when investigating time series samples across multiple assays. A standardised workflow with custom built software tools will be required to aid efficient data handling and analysis whilst minimising manual handling to facilitate better reproducibility (see Chapter 4). Further, we must strive to study males and females in tandem to investigate the degree of sexual dimorphism, if any, in this regulation (see Chapters 5 & 6). Defining a baseline of 'normal' circadian regulation of the metabolome in humans may also be informative to alternate studies investigating the effects of shift work, circadian misalignment or sexual dimorphism under such conditions as a point of comparison.

1.6.1. Research objectives

The purpose of this thesis, in the first instance, is to expand the foundational knowledge of the emergent field of circadian regulated metabolism through reviewing current works in the field and the application of untargeted UHPLC-MS metabolomic assays. Such work required the development of a software pipeline that incorporates new and existing tools purpose built

to handle and analyse -omics datasets and detect circadian rhythms. These tools should be modular in nature and adaptable to other time series -omic style data so as to be easily applied to future datasets, thus overcoming much of the labour-intensive analysis associated with both time series and -omic datasets. The research performed during my PhD is subdivided as follows:

Chapter 2 contains the methods for subsequent chapters.

Chapter 3 is an abridged summary of circadian/diurnal metabolite rhythms in human plasma/serum. This summary is derived from a larger published review that sought to collate and provide commentary on all prior studies published in the field, i.e., applying metabolomics to observe time of day variation in humans. The review and its summary highlight current gaps in knowledge and potential future avenues of research in the field.

Chapter 4 encompasses the construction of the metabolomics-rhythm analysis software pipeline. The chapter introduces the power of the R environment and packages at its disposal before discussing key stages of the pipeline and associated challenges that needed to be overcome. This pipeline is applied in its entirety in Chapters 5 & 6 and in part for Chapter 7 to assist with data processing/handling, and rhythm analysis.

Chapters 5 & 6 detail discovery based metabolomic studies for a RPC or 'lipidomics' assay and a HILIC assay, respectively, both employed in positive and negative polarity mode for greater coverage of the metabolome. The assays were applied to human plasma samples collected under constant routine at a 2 h resolution across a 32-h time course from a population of 16 males and 16 females. This population comprises the largest human population subject to constant routine to date, for the purpose of metabolomics analysis, whilst benefitting from a balanced cohort with regards to sex and age/BMI matched participants at high sampling

resolution. This will also be the first instance of these four assays being employed within the same study. These chapters will employ the workflow created in Chapter 4 and were designed to generate a wealth of data, that is otherwise limited within the field, to better understand the relationship between the circadian system and metabolism. The samples were provided as part of a collaboration with the University of Surrey, with details of their collection provided in [64], my role was to perform the untargeted metabolomics assays on the samples and perform the subsequent data analysis.

Chapter 7 is focussed on studying a new animal model for Huntington's disease (HD). This newly developed model features transgenic sheep (*Ovis aries* - OVT73) which possess the human transgene for HD to investigate physiological differences between wild type controls and HD sheep at pre-symptomatic ages of 5 and 7 years of age. As HD is a late onset chronic disease sheep may offer better suitability for its study over rodent models due to the former possessing a longer lifespan. Patients with HD exhibit a range of symptoms inclusive of disrupted sleep and perturbed circadian rhythms in addition to perturbations to metabolism being reported. Investigating circadian regulation of metabolism within this model organism at two different ages in wild type controls and HD sheep will reveal longitudinal changes in circadian regulated metabolism in the latter. Identifying perturbed aspects of circadian metabolism may offer greater insight into the pathophysiology of HD at this pre-symptomatic stage. Data files were provided by the University of Surrey as part of a collaboration, my role was to perform rhythmic analysis on the provided (semi-targeted) metabolite dataset.

CHAPTER 2

MATERIALS & METHODS

2.1. Methods – Chapter 3

2.1.1. Literature search and inclusion criteria

The following methods are reproduced from [78], a review I authored and presented in part as Chapter 3 of this thesis.

Prior to starting the literature search, three database/search engines were selected based on differing breadth of literature curation and indexing features to facilitate specific searching. PubMed (NCBI) provides specific curation of biomedical literature alongside subject headings (indexing) in the form of Medical subject headings (MeSH) terms for improved specificity of returned literature from submitted searches. Web of Science was selected for offering wider curation of literature, which may potentially not be included on the MEDLINE database used by PubMed. However, Web of Science does not use subject headings which may reduce search specificity *vs.* PubMed but still allows for Boolean operators (e.g., and, or) to combine search terms. Lastly, Google Scholar was chosen due to a lack of specific curation and similar search tools as Web of Science. Search terms relating to the subject fields of interest were chosen and, where possible, matched to MeSH terms—it was assumed that MeSH terms such as metabolomics likely correspond to common key words associated with literature across databases. Various combinations of search terms were trialled to acquire a search result that was considered broad in terms of 'hits', i.e., tens to low hundreds, but not so broad as to be infeasible to read and evaluate. Two separate searches were performed, with the first search featuring 'circadian' as a key word (performed circa 21 April 2020) and the second search featuring 'diurnal' in lieu of circadian (performed circa 23 July 2020).

Yielded literature from the search was counted as relevant based on the article title, article keywords, or abstract broadly referring to metabolomics/metabolism and circadian

rhythms/chronobiology. Relevant literature collected from the database search were subject to inclusion criteria to assess if they were capable of addressing the outlined objectives of the review. Inclusion criteria consisted of three parameters, applied in the following order:

1. The literature details original research, i.e., no derivative work such as reviews
2. The research studied human participants over a time course
3. Employed any metabolomics platform to analyse samples collected across the time course.

After this assessment of relevant literature yielded from the database search, manual searching was performed. Manual searching consisted of reading literature that met inclusion criteria and checking reference lists which may refer to further relevant studies (based on title/keywords/abstract) which would then be subject to inclusion criteria.

2.2. Methods – Chapter 4

All code was written in RStudio version 1.4.1103 and executed through R version 4.0.5 (R 64 bit) on a Windows 10 (64 bit) operating system.

In addition to the attached base packages the following additional packages and their dependencies were employed for the listed purposes:

Pre-processing

xcms_3.12.0 [242, 251]

Metabolomics data analysis

MetaboAnalystR_3.2.0 [269] in conjunction with the Metaboanalyst web app (Version 5.0) [161]

Rhythm detection

DiscoRhythm_1.6.0 [270] – cosinor analysis [271]

MetaCycle_1.2.0 [272] – Jonckheere-Terpstra Kendall (JTK) [273] & autoregressive spectral estimation regression (ARS) analysis [274]

echo.find_4.0.1 – extended circadian harmonic oscillator (ECHO) analysis [275]

rain_1.24.0 – rhythmicity analysis incorporating nonparametric methods (RAIN) analysis [276]

BioCycle_0.9.3 [277] – BioCycle analysis [278]

DODR – Detection of differential rhythmicity [279]

compareRhythms_1.0.1 – differential rhythmicity analysis [280]

Data handling packages (strings/matrices/tables/arrays, etc.) sourced from the comprehensive R archive network (CRAN) [281].

dplyr_1.0.6

Matrix_1.3-2

readr_2.1.2

stringr_1.4.0

tibble_3.1.1

tidyr_1.2.0

SummarizedExperiment_1.20.0

Graphics packages sourced from CRAN [281].

ggplot2_3.3.5

ggrepel_0.9.1

png_0.1-7

RColorBrewer_1.1-2

VennDiagram_1.7.1

Outside of the R environment the following software applications were utilised:

Proteowizard [248] version 3.0.21034 MSConvert GUI (64 bit)

Parameters applied within XCMS are discussed below (see Section 2.3.2.7.). Parameters applied within the rhythm analysis are detailed in [282] and were the default assignments unless otherwise stated.

2.3. Methods – Chapters 5 & 6

2.3.1. Participant recruitment & sample collection

The following methods sections 2.2.1.1. – 2.2.1.5. relate to Chapters 5/6 and were performed by external persons to the PhD project but are included for context and clarity. Full details on cohort recruitment, screening, constant routine protocol, and sampling can be found here [64] and are summarised below.

2.3.1.1. Recruitment & pre-study

Participants met the following inclusion criteria: non-smokers, free of pre-existing medical conditions, were not taking prescription medication (excluding females on the oral contraceptive pill (OCP) in the active pill phase), no menstrual irregularities reported by

females prior to OCP use, no prior drug/alcohol abuse, no history of shift work or international travel across times zones in the month prior to the study. During the week prior to the in-laboratory session participants adhered to a self-selected sleep wake cycle within the following set parameters: bedtime between 22:00-01:00 hours, waketime between 06:00-09:00 hours, 8 hours sleep (opportunity). This was monitored via sleep diaries, time stamped voicemails made by participants to report going to bed/waking up, and actigraphy data. Participants were required to expose themselves to 15 minutes of outdoor light within 90 minutes of waking up. In the 72 hours preceding the in-laboratory session participants were requested to refrain from alcohol and caffeine consumption, exercise, and bright light in the evenings. Participants were required to meet pre-defined criteria for four issued questionnaires: Pittsburgh Sleep Quality Index (PSQI) [283] global score of ≤5, Beck's depression inventory [284] score of < 10, Epworth sleepiness scale [285] score of < 10, and Horne-Östberg [286] questionnaire not being classified as either morning/evening types.

2.3.1.2. Cohort

The cohort consisted of 32 participants (16 males, 16 females) aged and BMI matched. The average age of the cohort was 23.6 ± 3.6 years (mean ± SD) with an average BMI 23.9 ± 2.6 (mean ± SD). All females were on OCP, as previously described, with four different formulations as follows: 30 µg ethinylestradiol and 150 µg progestin (n = 10), 30 µg ethinylestradiol and 3000 µg progestin (n = 2), 35 µg ethinylestradiol and 250 µg progestin (n = 1) or 0 µg ethinylestradiol and 75 µg progestin (n = 1).

2.3.1.3. Constant routine

The in-laboratory protocol took place over a 4-day period (68 h), during which participants had no direct knowledge of clock time. The timeline of the in-laboratory protocol is illustrated in Figure 2.1 and was as follows:

Day 1-2: Participants arrived at ~16:00, between 19:00-23:00 participants were under dim light conditions (<5 lux), lay semi-recumbent in bed (awake) and were provided with two standardised meals. Between 23:00 – 07:00 participant had an 8 h Sleep opportunity under 0 lux light conditions.

Day 2-3: From 07:00 on D2 until 23:00 on D3 participants were under dim light conditions (<5 lux), lay semi-recumbent in bed (awake), and were provided with hourly isocaloric snacks with 100 ml of water. Hourly blood sampling started from 15:00 on D2 until 23:00 on D3 with 33 samples collected in total, (16 samples with a 2 h resolution used for metabolomics). Samples were collected via intravenous cannulae.

Day 3-4: From 23:00 on D3 until 12:00 on D4 participants were provided with a (recovery) sleep opportunity under 0 lux light conditions.

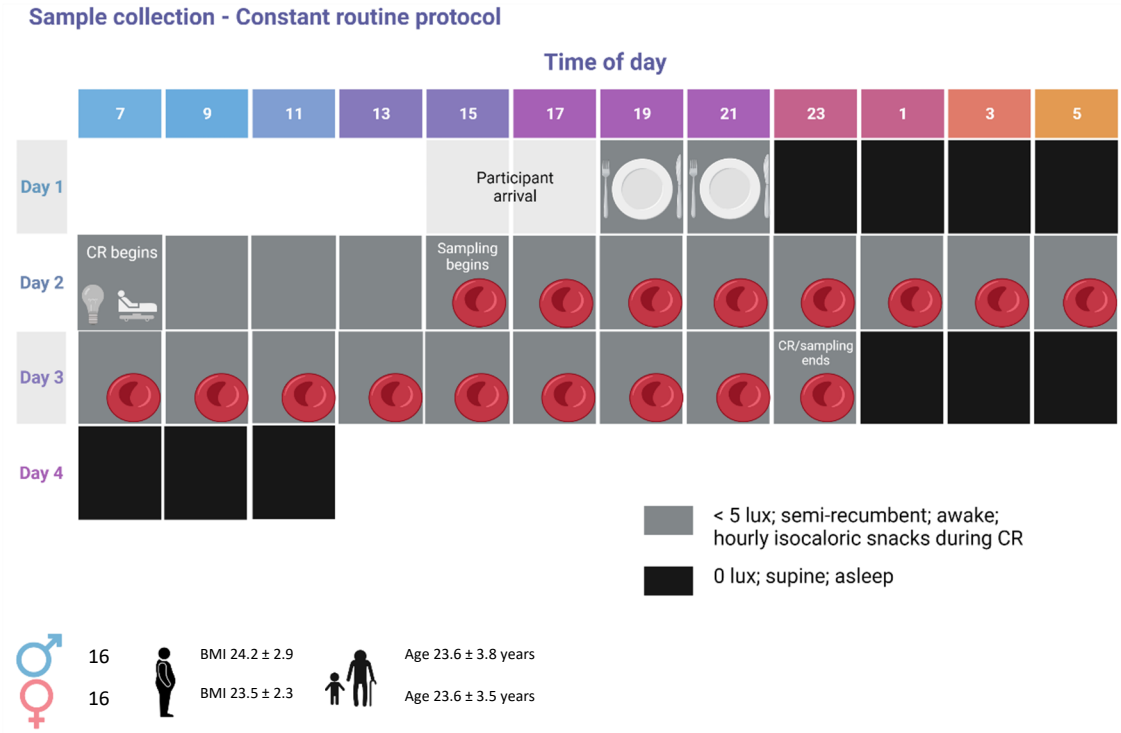


Figure 2.1: Infographic providing details on sample collection and the constant routine protocol – further details in [64].

2.3.1.4. Sample storage

Whole blood samples were collected into lithium heparin tubes, separated by centrifugation and the plasma fraction stored at -20°C, prior to melatonin/cortisol profiling or -80°C prior to UHPLC-MS analysis. Samples were later transferred, on dry ice, to the University of Birmingham and re-stored at -80°C prior to lipidomic/HILIC analysis.

2.3.1.5. Plasma melatonin and cortisol profiles

Plasma melatonin and cortisol concentrations were measured across the 33-h time course at an hourly resolution via radioimmunoassay. Plasma samples were incubated with an iodinated radio label (125I) and antiserum raised in sheep against a cortisol-3-(O-

carboxymethyl)oxime/bovine serum albumin conjugate, in a low pH buffer with no prior extraction performed [287]. All samples from any given participant were run in the same assay and samples were measured in duplicate. The limit of detection for the plasma melatonin assay was 3.9 pg/ml and inter-assay coefficient of variations (CVs) were 22.2% at 8.2 pg/ml, 10.4% at 33.2 pg/ml, 9.3% at 88.8 pg/ml and 9.0% at 121.4 pg/ml. The limit of detection for the plasma cortisol assay was 1.7 nmol/L and inter-assay CVs were 11.2% at 67.7 nmol/L, 11.0% at 520.0 nmol/L and 8.7% at 884.0 nmol/L [64].

2.3.2. Metabolomics – UHPLC-MS assays

The following methods pertaining to the metabolomics analysis, both Lipidomics and HILIC assays are adapted from Southam *et al* [155] and have been reproduced below.

2.3.2.1. Chemicals

Solvents and additives were all LC-MS grade or higher. Formic acid, acetonitrile, water, isopropanol from VWR International (Lutterworth, U.K); ammonium formate (99% purity) from Sigma-Aldrich (Poole, UK).

2.3.2.2. Instrumentation

The UHPLC-MS analysis was performed using a Vanquish UHPLC coupled to an electrospray Q Exactive Plus mass spectrometer (Thermo Scientific, San Jose, CA, U.S.A.) with Thermo ExactiveTune (2.8 SP1, build 2806) software controlling the instrument. All data were acquired in profile mode.

2.3.2.3. Quality control samples

A single pooled QC sample was prepared on a different day to biological sample extraction and analysis. All biological samples were thawed on ice and then equal aliquots (~13 µl) from

all 500 biological (plasma) samples ((16 timepoints X 32 participants) – 12 missing samples) were combined in a single 15 mL Falcon tube and vortexed (120 s). Post vortexing the pool was divided into aliquots, prior to freezing at -80^{o}C, to minimise the freeze thaw cycles required when making up QC samples across multiple days of sample preparation/analysis. The aliquot volume varied ranged from 700-1600 µl to correspond to the planned number of QC requiring preparation on a given day of sample preparation/analysis, with a single QC sample requiring 50 µl of the pooled sample.

2.3.2.4. Sample preparation

The untargeted lipidomics analysis, MS1 and MS2, employed two assays, a reversed phase (lipidomics) assay in both positive and negative ion mode as previously described [155]. For the extraction, samples were thawed/prepared on wet ice and 50 µL of plasma was transferred into a fresh 1.5 mL Eppendorf tube and 150 µL of isopropanol (IPA) added. Samples were vortexed (120 s) prior to centrifugation at 21,000 x g at 4^{o}C for 20 minutes. A 100 µL aliquot of the supernatant was transferred into a glass autosampler vial (VI-04-12-02RVG, 300 µl plastic with fused glass insert, Chromatography Direct, UK) and stored at 4^{o}C prior to UHPLC-MS analysis. The procedure for preparing QC and extraction (process) blank samples was the same as above with the exception of 50 µl of the QC pooled sample or water, respectively, used in lieu of plasma. The order in which samples were prepared and analysed was assigned via block randomisation, such that, all samples from a single participant were prepared and analysed together but in a random order that did not follow the linear time series, and that the order of these participant 'blocks' were themselves randomised to avoid sequential analysis of samples based upon sex or date of collection.

2.3.2.5. Lipidomics UHPLC-MS assay

Samples were analysed separately in positive and negative ion modes using a Vanquish UHPLC coupled to an electrospray Q Exactive Plus mass spectrometer (Thermo Scientific, San Jose, CA, U.S.A.). The assay applied a reversed phase method which used a Thermo Hypersil GOLD C_{18} column (2.1 ×100 mm, 1.9 µm) with two mobile phases, mobile phase A (10 mM ammonium formate in 60% acetonitrile / 40% water (v/v) + 0.1% formic acid) and mobile phase B (10 mM ammonium formate in 85.5% isopropanol / 9.5% acetonitrile / 5% water (v/v/v) + 0.1% formic acid). A 14 min gradient elution was applied as follows: t = 0.0, 20% B; t = 0.5, 20% B, t = 8.5, 100% B; t = 9.5, 100% B; t = 11.5, 20% B; t = 14.0, 20% B. All changes were linear (curve = 5), and the flow rate was 0.40ml min^{-1}. The column temperature of 55 °C and injection volume of 2 µL. Data were acquired in positive and negative ionisation modes separately within the mass range of 150–2000 *m/z* at resolution 70,000 (FWHM at *m/z* 200). Ion source parameters were set as follows: sheath gas = 48 arbitrary units, aux gas = 15 arbitrary units, sweep gas = 0 arbitrary units, spray voltage = 3.2 kV (positive ion mode) / 2.7 (negative ion mode), capillary temp. = 350 °C, aux gas heater temperature = 400 °C. Data dependent MS/MS in 'Discovery mode' was used for the MS/MS spectra acquisition using the following settings: resolution = 17,500 (FWHM at *m/z* 200); isolation width = 3.0 *m/z*; stepped normalised collision energies (stepped NCE) = 25, 60, 100%. Spectra were acquired in five different mass ranges: 150-510 *m/z*, 500-710 *m/z*, 700-860 *m/z*, 850-1010 *m/z*, 1000-2000 *m/z*. Each assay comprised of two batches with QC samples arranged as followed, batch 1: QC samples were analysed as injections 1-22 with the exception of injection 6 which was an extraction blank sample, QC injections 15-19 were applied for MS/MS data collection, biological samples were analysed as injections 23- 358 with every 7^{th} injection being a QC sample, injections 359-361 were 2 QC samples follow by an extraction blank making up the

end of batch 1. Batch 2: QC samples were analysed as injections 1-11 with the exception of injection 6 which was an extraction blank sample, biological samples were analysed as injections 12- 272 with every 7th injection being a QC sample, injections 273-275 were 2 QC samples follow by an extraction blank making up the end of batch 2.

2.3.2.6. HILIC UHPLC-MS assay

Samples were analysed separately in positive and negative ion modes using a Vanquish UHPLC coupled to an electrospray Q Exactive Plus mass spectrometer (Thermo Scientific, San Jose, CA, U.S.A.). The assay applied a hydrophilic interaction chromatography (HILIC) method which used a Thermo Accucore 150 Amide column (2.1 ×100 mm, 2.6 µm) with two mobile phases, mobile phase A (10 mM ammonium formate in 95% acetonitrile / 5% water (v/v) + 0.1% formic acid) and mobile phase B (10 mM ammonium formate in 50% acetonitrile/ 50% water (v/v) + 0.1% formic acid) for positive ion mode. Negative ion mode mobile phase solutions used 10 mM ammonium acetate in lieu of ammonium formate and 0.1% acetic acid in lieu of formic acid. For positive ion mode, a 14 min gradient elution was applied as follows: t = 0.0, 1% B; t = 1.0, 1% B; t = 3.0, 15% B; t = 6.0, 50% B; t = 9.0, 95% B; t = 10.0, 95% B; t = 10.5, 1% B; t = 14.0, 1% B. For negative ion mode a 15 min gradient elution was applied as follows: t = 0.0, 0% B; t = 2.1, 1% B; t = 4.1, 15% B; t = 7.1, 50% B; t = 10.1, 95% B; t = 11.0, 95% B; t = 11.5, 1% B; t = 15.0, 1% B. For both ion modes: all changes were linear (curve = 5), flow rate was 0.50 mL min−1, column temperature was 35 °C , and injection volume was 2 µL. Data were acquired in positive and negative ionisation modes separately within the mass range of 70–1050 *m/z* at resolution 70,000 (FWHM at *m/z* 200). Ion source parameters were set as follows: Sheath gas = 55 arbitrary units, aux gas = 14 arbitrary units, sweep gas = 4 arbitrary units, spray voltage = 3.2 kV (positive ion mode) / 2.7 (negative ion mode), capillary temp. = 380 °C, aux

gas heater temp. = 440 °C. Data dependent MS/MS in 'Discovery mode' was used for the MS/MS spectra acquisition applying the same settings utilised in the lipidomics assays (see Section 2.3.2.5.). Each assay comprised of two batches with the same QC/biological sample injection order as the lipidomics assays (see Section 2.3.2.5.).

2.3.2.7. Data deconvolution & quality controls

The resulting two sets of raw data per assay, positive (batch 1+2) & negative (batch 1 + 2) ion mode assays, were processed independently but subject to the same protocol outlined below. Post MS analysis, vendor format raw data files (.RAW) were converted to .mzML format via Proteowizard (v 3.0.2 – 64 bit build, [288]) prior to deconvolution through the 'xcms' R package (v 3.12.0, [242, 289, 290] with the following parameters: ppm=14, SNR = 10, sigma = 3, prefilter = c(3,100), minfrac = 0.2, peakwidth_low = 6, peakwidth_high = 30, mzdiff = 0.002, retcor_meth = obiwarp, noise_centroid = 0, mzCenterFun = "wMean", integrate = 1 , fitgauss = FALSE, verboseColumns = TRUE, bw = 0.25, mzwid = 0.01, gapInit = 0.4, gapExtend = 2.4, profstep = 1. Obiwarp retention time correction was applied and all other undefined parameters were left as their default assignments.

A data matrix of peak areas for metabolite features (*m/z*-retention time pairs) *vs.* samples were constructed. Data quality was assessed semi-quantitatively, identifying individual samples/participants which were a prominent source of intra-batch, i.e., beyond 95% confidence intervals for their respective batch within PCA performed via Metaboanalyst's online platform (v 5.0, [161]) as detailed below (section 3.4.4), or were known to be negatively impacted by technical issues during analysis. Individual QC samples that exhibited a high degree of variation, relative to other QC samples within the same batch, were identified and removed in addition to all conditioning QC samples (first 20/10 QC of batches 1/2,

respectively). Batch correction was then applied per assay matrix to mitigate inter-batch variation. This corrective step was performed by calculating average % difference of intensity between batch one and batch two QC samples, feature wise. The batch with the lower % difference was adjusted (QC and biological samples) so that the difference between average intensities became 0, e.g., if batch one QC had an average intensity 10% less than batch two QC for a given metabolite feature, the entirety of batch ones' intensity values were increased by 10% to match for that metabolite feature and then repeated for all metabolite features separately. Post batch correction the matrices were filtered as follows: any metabolite feature whose average intensity in the biological samples was <10× its average intensity of the process blank samples was deleted, any metabolite feature present in <70% of the QC samples was deleted; metabolite features with relative standard deviation (RSD) ≥ 30% across the intra-study QC samples were deleted. The code applied for XCMS and QC filtering was discussed as part of Chapter 4 and can be found in the respective appendix (section B).

2.3.2.8. Metabolite annotation

Putative metabolite annotation applying MS1 data was performed by applying the Python package BEAMSpy version 1.1.0 (https://github.com/computational-metabolomics/beamspy); with the following parameters applied: Maximum RT difference = 2, Grouping method = Pearson correlation, coefficient threshold = 0.5, P-value threshold = 0.05. Reference tables for adducts, isotopes and neutral losses were provided and are shown below (Tables 2.1-2.3). The *m/z* values of all experimentally observed peaks were searched directly, one at a time, against HMDB – The Serum Metabolome Database (https://hmdb.ca), KEGG – Human (hsa) database (https://www.genome.jp/kegg/) and LIPIDMAPS – full database (https://www.lipidmaps.org), and all matches within a 5 ppm mass error tolerance

were recorded. Multiple annotations (e.g., isomeric compounds) could be observed for a single detected metabolite feature. Here, the term putatively annotated metabolite refers to either single or multiple chemical compounds with the same retention time and accurate *m/z*. To generate more robust compound annotations, QC sample UHPLC-MS/MS data were matched to MS/MS databases using either LipidSearch software (lipid annotation; version 4.2.18, Thermo Fisher Scientific) or Compound Discoverer 3.1 software (polar compound and some lipid annotations; version 3.1, Thermo Fisher Scientific), as previously described by [155], with the exception of an extended retention time tolerance window, and defined below. LipidSearch features within the UHPLC-MS/MS data were searched against the entire in-silico HCD(cell) MS/MS database (5 ppm mass error). Only annotations graded A-C were used for annotation purposes (Grade A – all fatty acyl chains and class were completely identified; Grade B – some fatty acyl chains and the class were identified; Grade C – either the lipid class or some fatty acyls were identified). MS/MS annotations using Compound Discoverer were graded by HighChem HighRes algorithm within the Compound Discoverer software (annotations with a score >60 were retained for identification purposes). LipidSearch and Compound Discoverer annotations were aligned to the XCMS outputs using a 5 ppm mass error and 30 s (+/- 15) retention time tolerance window. Annotations were manually filtered to remove false positive annotations such as drugs and their derivatives known not to be taken by the participants or annotations deemed unlikely to be detected within human plasma. Utilising the LIPIDMAPS database, and with manual input, lipid annotations were transformed so that they were presented at the class level using the appropriate IUPAC abbreviation. Synonyms or chemical names not in common use, such as citrate/citric acid, were updated to reflect more common use nomenclature with duplicate entries removed. Applied code for annotation filtering is discussed as part of Chapter 4 and

can be found in the respective appendix (section B). Parameters applied for BEAMSpy are either listed above or provided in the below tables (Tables 2.1, 2.2, 2.3).

Table 2.1: Adduct list submitted to BEAMSpy.

Label	exact_mass	charge	ion_mode
[M+H]+	1.007276	1	pos
[M+NH_4]+	18.03383	1	pos
[M+Na]+	22.98922	1	pos
[M+Mg]+	23.98449	1	pos
[M+K]+	38.96316	1	pos
[M+Fe]+	55.93439	1	pos
[M+Cu]+	62.92905	1	pos
[M+2H]+	2.015101	1	pos
[M+3H]+	3.022926	1	pos
[M-H]-	-1.00728	1	neg
[M+35Cl]-	34.9694	1	neg
[M+Formate]-	44.9982	1	neg
[M+Acetate]-	59.01385	1	neg

Table 2.2: Neutral losses list submitted to BEAMSpy.

LABEL	mass_difference
H_2O	18.01057
CO	27.99492
NH_3	17.02655
C_2H_2	26.01565
C_2H_4	28.0313
CO2	43.98983
C_2H_4O	44.02621

Table 2.3: Isotope list submitted to BEAMSpy.

LABEL_X	label_y	mass_difference	abundance_x	abundance_y	ion_mode	charge
Li	(7Li)	1.00088	7.59	92.41	both	1
C	(13C)	1.00335	98.93	1.07	both	1
C	(13C)	0.501675	98.93	1.07	both	2
C	(13C)	0.33445	98.93	1.07	both	3
C(2)	(2x13C)	2.0067	99.8	0.2	both	1
C(3)	(3x13C)	3.01005	99.97	0.03	both	1
Mg1	(25Mg)	1.00008	78.99	10	both	1
Mg2	(26Mg)	1.99755	78.99	11.01	both	1
Mg1	(25Mg)	0.50004	78.99	10	both	2
Mg2	(26Mg)	0.998775	78.99	11.01	both	2
S	(34S)	1.9958	94.99	4.25	both	1
Cl	(37Cl)	1.99705	75.76	24.24	both	1
K	(41K)	1.99812	93.26	6.73	both	1
Ca	(44Ca)	3.99289	96.94	2.09	both	1
Fe1	(54Fe)	0.997665	91.75	5.85	both	2
Fe2	(57Fe)	0.500225	91.75	2.12	both	2
Fe1	(54Fe)	0.66511	91.75	5.85	both	3
Fe2	(57Fe)	0.333483	91.75	2.12	both	3
Cu	(65Cu)	1.99819	69.15	30.85	both	1
Cu	(65Cu)	0.999095	69.15	30.85	both	2
Se1	(76Se)	3.99731	49.61	9.37	both	1
Se2	(77Se)	2.99661	49.61	7.63	both	1
Se3	(78Se)	1.99921	49.61	23.77	both	1
Br	(81Br)	1.99795	50.69	49.31	both	1

2.3.3. Rhythm analysis

2.3.3.1. Z-score

The QC filtered matrices were Z-scored, these scores were calculated feature wise per participant using the general formula:

Z-score = (Observed value – mean) / Standard Deviation

Where the observed value was the relative abundance at a given time point for a participant, the mean being the averaged relative abundance across all 16 time points for the participant, and the standard deviation calculated from those same 16 time points. The Z-score equation was applied as detailed in Chapter 4 with code provided in the respective appendix section (section B).

2.3.3.2. DLMO correction

Post Z-scoring the matrices were corrected for DLMO so that the time series was measured relative to DLMO as opposed to social time as described previously [64]. DLMO was defined as the sample time point in which melatonin concentration exceeded 25% of the peak (highest) measured melatonin concentration measured on the evening of day 2 and was measured for each participant, as described above. For example, if melatonin concentration peaked at 80 pg/ml then the time point at which melatonin exceeded 25% of this value (20 pg/ml) is defined as DLMO. The time point at which DLMO occurred is defined as t = 0 on the DLMO scale with preceding and proceeding time points allocated negative and positive signage respectively.

2.3.3.3. Applied rhythm analysis algorithms

Prior to rhythm detection all feature intensity matrices were further filtered to remove any metabolite features with >50% missing values in any one participant such that, if any metabolite feature had missing values for >9 of 16 timepoints for >=1 participants the feature was removed across all participants within the group (i.e., male or female group). The remaining missing values were not imputed. To detect circadian periods (~24 h) and determine acrophase (rhythm peak), six rhythm analysis algorithms available as R packages were employed: (R – Version 4.0.5, Cosinor – *'DiscoRhythm'* (Version 1.6.0, [270]), JTK & ARS – *'MetaCycle'* (Version 1.2.0, [272]), ECHO – '*echo.find*' (Version 4.0.1, [275]), RAIN – *'rain'* (Version 1.24.0, [276]), and BioCycle – (Version 0.9.3, [277]), with default parameters applied unless otherwise stated as detailed in [282]. Applied code for rhythm analysis is discussed as part of Chapter 4 and can be found in the respective appendix (section B). Results from all six algorithms were cross referenced such that individual metabolite features only classified as 'rhythmic' with a ~24 h period if a significant result $p < 0.05$ (post Benjamini-Hochberg FDR correction) was obtained in at least three of the applied algorithms. The analysis was performed both on an individual level (per participant) and at the group level (male and females as separate groups). However, due to an unexpected error using ECHO with a single sample per time point it was not used for individual level analysis, the threshold for classifying metabolites as rhythmic remained the same. For the group level analysis rhythmic metabolites, subset from the intensity matrix, were ordered by ascending acrophase, as determined by cosinor analysis, and then plotted as heatmaps to visualise and confirm the presence of ~24 h rhythms in both groups. Heatmaps were produced via MetaboAnalystR [269] with no additional filtering or data transformations, as data was previously Z-scored,

except for group averaging data per time point. K Nearest neighbour (KNN, sample wise) was performed to impute missing values for the heatmaps.

Differential rhythmicity analysis, via DODR (Detection of differential rhythmicity, [279]) and compareRhythms [280] classification function, was employed to directly compare the rhythmic metabolite feature profiles, between males and females, of metabolite features classified as rhythmic (24 h period) by the above algorithms. Metabolite features were classified as follows in the model: 'same' – indicating the metabolite feature is deemed rhythmic in both groups with no significant difference in rhythm between groups, 'gain' – indicating the metabolite feature is deemed rhythmic in one group (females) with a significant difference in rhythm between groups, 'loss' – indicating the metabolite feature is deemed rhythmic in one group (males) with a significant difference in rhythm between groups, 'change' – indicating the metabolite feature is deemed rhythmic in both groups with a significant difference in rhythm between groups, 'arrhy' - indicating the metabolite feature is deemed non-rhythmic for 24 h rhythms in both groups. Rhythm analysis algorithms were applied using the code detailed in Chapter 4 with code provided in the respective appendix section (section B).

2.3.3.4. Multivariate analysis & descriptive statistics

Unless otherwise stated all data matrices submitted to Metaboanalyst for PCA analysis applied the following parameters: missing value imputation – KNN (sample wise), filtering - no further filtering, normalisation – log transformed & Pareto scaled. PCA plots were constructed with the default settings applied, specifically with PC1 (x axis) and PC2 (y axis) and displaying 95% confidence regions. Unsupervised multivariate PCA was used to determine outlier samples and to observe inter-group variation between sexes via an

unsupervised method. All t-tests were performed in Excel (Excel 2016 64-bit, Version 2208 Build 16.0.15601.20446) using the "Data Analysis" function applying the "Two-Sample Assuming Unequal Variances" method with a hypothesised difference of 0. All Spearman's rank correlation analyses were performed in Excel. The Spearman's rank coefficient was derived in Excel using the "=Correl()" function. The T-statistic was calculated as follows, "=(ABS(coefficient)*SQRT(n-2))/(SQRT(1-ABS(coefficient)^2))" where coefficient refers to the calculated Spearman's rank coefficient and 'n' the population size. The significance value was derived in Excel using the formula "=TDIST(T-statistic,n-2,2)".

2.3.3.5. Cohort changes

Post data collection nine participants were removed due to technical problems, further detailed in section 5.2.1, leaving a population of 23 participants (13 males, 10 females). As the rhythm analysis algorithms are applied independently to each group, i.e., hypothesis testing is performed independently for males and females, the decision was made to further reduce the population to 20 (10 males, 10 females) for the group level rhythm analysis. This was to mitigate concerns that if males presented with more rhythmic features than females, it was not the result of a larger sample size increasing the 'sensitivity' of the rhythm detection. The males removed from the group rhythm analysis data were chosen randomly by allocating each participant a value from 1-13 then drawing 3 random numbers from Excel's built-in random number generator ("=RANDBETWEEN(1,13)") to determine who should be removed. Differences in age and BMI remained non-significant ($p<0.05$) between males (age 22.7 ± 1.8 SD, BMI 25.3 ± 2.7 SD) and females (age 24.5 ± 3.7 SD, BMI 23.6 ± 2.2 SD) in this abridged cohort.

2.4. Methods – Chapter 7

2.4.1. Animal husbandry & sample collection/analysis

The following methods for Chapter 7 (sections 2.4.1.1. – 2.4.2.1.) were performed by external persons to the PhD project but are included for context and clarity. Full details on animal husbandry, sampling, and the metabolomics analysis can be found here [282] with key details summarised below.

2.4.1.1. Animals

The animal model utilised was the OVT73 line of *Ovis aries* that carries a full length human transgene that would typically induce juvenile onset of Huntington's disease (73 repeats of the CAG sequence). Sheep were of five years of age (14 HD, 13 control) or 7 years of age (16 HD, 14 control), thus age matched between control and HD groups at the two timepoints, and all were female (ewes). The seven year-old group included eight of the control and ten of the HD sheep that had previously been sampled at five years of age. The sheep were obtained from a mixed genotype (HD and control) flock maintained outdoors at the South Australian Research and Development Institute (SARDI) in accordance with OGTR approval NLRD-1037/2003 and NLRD-7574/2015. None of the sheep showed overt behavioural abnormalities of any kind and were considered to be at a 'pre-symptomatic' stage.

2.4.1.2. Pre-study & sample collection

Four weeks prior to sample collection sheep were fed a regulated diet. In advance of and during blood sampling sheep were housed in individual indoor metabolic crates with an indoor light/dark cycle matching the timing of the outdoor light/dark cycle. On the day of cannulation the lights remained off from sunset and remained off until blood sampling

collection was completed 45 h later. Samples were collected under dim light conditions using red head torches. Blood collection started at 13:00 the day after cannulae were placed with sampling occurring hourly (+/- 5 minutes) for the 27 h time course.

2.4.1.3. Sample storage

Post collection, blood samples were transferred to lithium heparin tubes then immediately centrifuged at 1620 g for 10 minutes at 4^{o}C. Plasma was aliquoted and stored at –20^{o}C for assays measuring melatonin/cortisol/urea and –80^{o}C prior to being shipped on dry ice to the UK for metabolomics analysis.

2.4.1.4. Plasma melatonin and cortisol profiles

Melatonin and cortisol concentrations were measured by radioimmunoassay as previously described [291].

2.4.1.5. Plasma urea profiles

Urea nitrogen measurements across the 27 h timeseries were taken using the ADVIA® 1800 Chemistry System (Siemens Healthcare Diagnostics Inc., Camberley, UK), method based on the Roch-Ramel enzymatic reaction using urease and glutamate dehydrogenase.

2.4.2. Metabolomics – (semi)-targeted assay

The metabolomics analysis utilised the Biocrates (Biocrates Life Sciences, Innsbruck, Austria) AbsoluteIDQ p180 kit and was performed according to manufacturer's guidelines as previously described [282, 292, 293].

2.4.2.1. Quality control

Sample order was randomised and 3 levels of quality controls (QC) were run on each 96-well plate. The levels of metabolites present in each QC were compared to the expected values and the RSD calculated. Data were normalised between batches using the results of quality control level 2 (QC2) repeats across the plate (n = 4) and between plates (n = 5) using Biocrates METIDQ software (QC2 correction). Metabolites where >25% concentrations were below the limit of detection (≪LLOQ) or above limit of quantification (>LOQ) or blank out of range, or the QC2 RSD was >30%, were excluded [293].

2.4.3. Rhythm analysis

Z-scoring, DLMO correction, applied rhythm analysis algorithms, and multivariate/descriptive statistics were performed as described in section 2.3.3.

CHAPTER 3

TIME OF DAY VARIATION OF THE HUMAN METABOLOME.

The following passages, results, and conclusions have been reproduced in part from prior published work, originally written by myself with planning/editing contributions from co-authors (project supervisors) as part of a literature review completed for my PhD. The 36-page document served as a brief introduction to investigating circadian/diurnal rhythms in humans and reviewed 29 studies that employed metabolomic platforms to investigate time of day changes (circadian or diurnal) in the metabolome. These studies took place under a variety of conditions and across five sample types (blood, urine, saliva, breath, skeletal muscle) with full commentary and discussion provided in the review and reproduced in an abridged format below.

3.1. Introduction

A favoured application of metabolomics is that of metabolic phenotyping, typically for biomarker discovery and better understanding of disease pathology within the context of the functions of metabolites observed in human metabolomes [79, 294], with such studies applied to the most prevalent chronic diseases within the human population such as cancer, diabetes and cardiovascular disease [295]. However, the concentration of metabolites in human biofluids and tissues is not static and varies across timescales of seconds to decades, driven by biological functions. Observed variation between such studies is inevitable due to biological variation within and between subjects as a result of genotype, environment, lifestyle, and the influence of biological rhythms, some of this variation being mitigated by control and monitoring of participant behaviour/diet/environment before and/or during the study [296]. In addition to biological variation, there is also analytical variation, resulting from sample selection/preparation [296, 297], sample degradation as a result of storage conditions and freeze–thaw cycles [214, 298, 299], instrument variation [300] and varied methods of data pre-treatment, statistical analysis and modelling [297, 301, 302]. Standard practices in the field such as quality control processes, batch correction, data normalisation, and matched cohorts mitigate or report some of this variation [79, 302] but the extent and rigor in which such practices apply vary from study to study, all of which impact the analysis and resulting biological interpretation [303]. This contributes to the 'reproducibility crisis', i.e., a recurring inability for external groups to reproduce published results, culminating in numerous independent research teams achieving conflicting results or inadvertently promoting false positive findings stemming from replicating effects and bias [301, 304-306]. Considine [301] puts forward an informed and comprehensive picture of the challenges of reproducibility and the implications for metabolic profiling and biomarker discovery. However, a key variable

which may explain some variation in the data within and between studies is repeatedly overlooked; that of diurnal variation and circadian control which regulates 24 h biological rhythms including the metabolome. This results in significant changes to observed concentrations of a range of metabolites and hormones dependent on the time of day, with a well known example being cortisol [307], with proline and leucine being further examples of highly rhythmic metabolites. Studies discussing these and further rhythms in metabolite concentrations and the context of these observations are discussed below. A number of detailed reviews and tutorials for biomarker discovery have also overlooked circadian control and diurnal variation of the metabolome as a consideration or influencing factor in study design and sample collection [113, 297, 308-312], with others only acknowledging circadian control as influencing metabolism in passing [313, 314] and fewer still suggesting that it be considered in study design [157]. The omission of time of day variation as an important feature in metabolomics study design likely results from the fact that work characterising circadian and diurnal influence through such platforms has almost exclusively occurred over the past ~15 years. The growing body of circadian metabolomics research should be of interest to the metabolomics community due to identifying metabolites that exhibit significant 24 h rhythmic behaviour as well as detailing the context and conditions in which this variation is observed. Such findings may prove informative when constructing or interrogating a biomarker panel, dissecting sources of inter-individual variation within studies as well as opening an additional line of enquiry when conflicting results arise between studies. It is timely that this body of work is thoroughly reviewed to establish the current knowledge base and consider the implications with regard to reproducibility especially in the construction of biomarker panels.

Whilst an in-depth knowledge of chronobiology is not required to appreciate the findings of the studies discussed within this review, a general understanding of biological rhythms, specifically circadian and diurnal rhythmicity, as well as internal biological clocks, their relationship with metabolism, and how they can be monitored will offer some insight into study design and the importance of context with regard to how results were collected. Such details are provided in the introductory chapter to this thesis as well as the original publication of this review [78].

Considering the broad context in which time of day variation can be observed, and the implications that this may have for reproducibility between studies, I reviewed all original research to date, at the time of writing, on human participants that applied either an untargeted or targeted metabolomics platform to observe time of day variation.

The objectives of the review [78] were as follows:

1. Establish the tissues in which time of day variation of metabolites have been observed, or failed to be observed, and the extent to which the metabolome is influenced.

2. Establish the source(s) for this observed daily variation and, if applicable, which metabolite classes are most susceptible.

3. Consider the implications of circadian/diurnal variation and the timing of sample collection on biomarker discovery and how this may undermine their potential clinical application.

3.2. Literature search – results and commentary

A summary of the literature search outcomes are shown in Table 3.1.

Table 3.1: Summary of literature search parameters and outcomes, reproduced from [78].

Search Terms	Database/Search Engine	'Hits'	Relevant Papers (Based on Abstract)	Met Inclusion Criteria †	
				Circadian Studies	Diurnal Studies
"Human(s)" "Circadian Rhythm OR Circadian Clocks" "Metabolomics OR Metabolome" ††	PubMed (NCBI)	70	133	6	19
	Web of Science	52			
"Metabolomic" "Circadian" "Rhythm" "Human" "Chronobiology"	Google Scholar	212			
N/A	Further manual searches	13			
"Human(s)" "Diurnal Variation OR Diurnal", "Metabolome OR Metabolomics" †††	PubMed (NCBI)	19	123 (Majority duplicates of prior search)	3	16
	Web of Science	28			
"Metabolomic" "Metabolome" "Diurnal" "Rhythm" "Human" "Chronobiology"	Google Scholar	92			
N/A	Further manual searches	0			

†A total of 29 novel papers met inclusion criteria, human studies employing a metabolomics platform taking multiple samples over a time course. †† Search terms corresponding to MeSH terms; ††† No MeSH terms corresponding to "Diurnal" OR "Diurnal variation". Two separate searches were performed, with the first search featuring 'circadian' as a key word (performed circa 21 April 2020) and the second search featuring 'diurnal' in lieu of circadian (highlighted—light grey, performed circa 23 July 2020). On the first search, the search terms had to be reduced when using Google Scholar due to insufficient 'hits' (n = 1 'hits') resulting from excess search terms. Manual searches consisted of looking for prior referenced work in collected literature that met inclusion criteria. NCBI – National center for Biotechnology information

For the purpose of brevity, only results pertaining to blood are reproduced in the main text of this thesis as other sample types are not relevant to subsequent chapters but commentary was provided in the original publication for urine, saliva, breath, skeletal muscle. Details relating to study design of reviewed literature (Table 3.2) are not listed within the main body of text unless relevant to the provided commentary.

Instead, key aspects of study design alongside results pertinent to the outlined objectives are summarised in tables organised by sample type with: urine, saliva, breath, skeletal muscle tables located in the appendix section (section A, Tables A1-A4) and blood (Table 3.2) in the main text. Furthermore, supplemental material containing compiled lists of metabolites showing time of day variation, and the context in which this variation took place (e.g., circadian, sleep deprivation, shift work), have been produced (Tables S1–6 in [78]) in addition to an overview of observed time of day changes across the studies (Figures S1 and S2 in [78]). Blood is the most investigated biofluid assessed for time of day variation of the metabolome, based on the outcome of the performed literature search, with 18 studies discussed here. The collated literature implicated various factors contributing to time of day variation and other findings of interest which are presented below as common subthemes, e.g., relating to shiftwork or health.

Table 3.2: A brief summary of study design, cohort details and results with regard to observed time of day variation of metabolites for relevant studies analysing plasma/serum, reproduced from [78].

Author(s)	Assay/Platform	Time Course Details	Study Setting/Conditions	Cohort Details	Rhythmic/Gradient Metabolites/ Features Observed	Rhythmic/Gradient Classes Primarily Observed
Park *et al*. (2009) [315]	Untargeted ^{1}H NMR	Diurnal variation 24 h, 1 h intervals between samples	'Inpatient' Standardised meals. Consistent light/dark cycle	N = 10, 5 males Age 22–83 BMI 18.5–32.6	34	Amino acids Lipids (unidentified)
Ang *et al*. (2012) [316]	Untargeted UPLC/Q-TOF MS (Reversed Phase)	Diurnal variation (25 h, 3 h intervals between samples)	'Inpatient' 17:8 wake/sleep, light/dark cycle. Hourly isocaloric meals Semi-recumbent position	N = 8 All male Age 53.6 ± 6.0 BMI 23.2 ± 1.4	203 features (19%) 34 metabolites	Amino acids Acylcarnitines LysoPEs LysoPCs
Dallmann *et al*. (2012) [317]	Untargeted GC-MS LC-MS (Reversed Phase)	Circadian variation (constant routine 40 h, 4 h intervals between samples)	'Inpatient' Standard constant routine parameters (see [318])	N = 10 (split into 2 equal groups, within which samples were pooled for each 4 h interval) All male Age 57.8 ± 1.0 & 61.0 ± 0.6 BMI 26.6 ± 0.6 & 25.1 ± 0.5	41 (15%)	Amino acids Glycerophospholipids Acylcarnitines Steroid hormones
Kasukawa *et al*. (2012) [319]	Untargeted LC-TOF MS (Reversed Phase)	Circadian variation (forced desynchrony 28 h, bookended by constant routine protocols (38 h each, 2 h intervals between samples)	'Inpatient' Standard constant routine parameters (with the exception of meals every 2 h (see [318]) Controlled light/dark cycles, temperature during forced desynchrony	N = 6 All male Aged 20–23	312 features (7%)	Amino acids Steroid hormones

Chua *et al*. (2013) [320]	Targeted Lipidomics LC-MS/MS (Reversed Phase)	Circadian variation (constant routine 37 h, 4 h intervals between samples at 5 h onwards of constant routine)	'Inpatient' Standard constant routine parameters (see [318])	N = 20 All male Age 24.4 ± 1.8 3 'Overweight' 17 'Healthy	35 (13.3%)	Glycerolipids Glycerophospholipids
Davies *et al*. (2014) [321]	Untargeted UPLC/Q-TOF MS/MS and targeted FIA-MS UPLC-MS/MS (Reversed Phase)	Diurnal variation (24 h). 24 h wake/sleep cycle *vs.* 24 h prolonged wakefulness, 2 h intervals between samples 48 h	'Inpatient' Standardised meals and mealtimes. Controlled light/dark cycle and activity/posture	N = 12 All male Age 23 ± 5, BMI 24.5 ± 2.3	109 (63.7%) sleep/wake 88 (51.5%) sleep deprivation 78 (45%) during both conditions	Amino acids Acylcarnitines LysoPCs Phosphatidylcholines Sphingolipids Fatty acids
Kim *et al*. (2014) [322]	Untargeted LC—TOF MS (Reversed Phase)	Diurnal variation Sampling 1, 3, 7, 9, 11, 14 h post-wake, first sample fasted.	'Inpatient' Standardised meals and mealtimes	N = 26 14 males Age 33 ±10.9 BMI 24.3 ±3.3	11 (9%)	LysoPCs Phosphatidylinositol
Chua *et al*. (2015) [323]	Targeted Lipidomics LC-MS/MS (Reversed Phase)	Circadian variation (constant routine 37 h, 4 h intervals between samples at 5 h onwards of constant routine)	'Inpatient' Standard constant routine parameters (see [318])	N = 20 All male Age 23 ± 5 BMI 24.5 ± 2.3	4 (1.5%) decreased during sleep deprivation 21 (5.5%) increased during sleep deprivation	Sphingomyelins TGs Phosphatidylcholines Phosphatidylinositol
Skarke *et al*. (2017) [324]	Targeted LC-MS/MS (HILIC)	Diurnal variation am *vs.* pm (48 h, 5 samples 12 h apart)	'Outpatient'	N = 6 All male Age 32.3 ± 3.6 BMI 25.2 ± 3.4	9 (5.4%)	

Isherwood *et al*. (2017) [325]	Targeted FIA-MS UPLC-MS/MS (Reversed Phase)	Diurnal variation (24 h—2 h intervals between samples)	'Inpatient' Controlled sleep/wake, light/dark cycle, and posture Hourly isocaloric meals	N = 23 All male BMI/Age Lean group 23.2 ± 1.4/53.6 ± 6.0 OW/OB 29.8 ± 2.3/51.0 ± 7.7 T2DM group 31 ± 1.6/57.3 ± 4.8	50/130 (38.5%) total 35—lean 39—OW/OB 20—T2DM	Amino acids Phosphatidylcholines LysoPCs Acylcarnitines
Gehrman *et al*. (2018) [326]	Targeted ^{1}H NMR	Diurnal variation (48–h - 2 h intervals between samples)	'Inpatient' Habitual sleep/wake cycle Hourly isocaloric meals	N = 30 20 male and 10 females (split equally into 2 groups) BMI < 29 Healthy Age 35.0 ± 7.5 Insomnia Age 37 ± 7.9	24 (total) 11 common to both groups 6 unique to healthy 7 unique to insomnia	Amino acids
Sato *et al*. (2018) [327]	Untargeted UHPLC-MS/MS GC-MS	Diurnal variation am *vs.* pm	'Outpatient' Standardised meals and mealtimes	N = 8 All male Age 30–45 BMI 27–32.5	532, 130, 349 features (50%, 12%, 33%) time of day, diet, time of day diet interaction, respectively. After HFD 13% features lost daily variation, 17% gained new daily variation After HCD 7% features lost daily variation 14% gained new daily variation	Amino acids Fatty acyls Glycerolipids Glycerophospholipids Sphingolipids Carbohydrates Xenobiotics

Skene *et al*. (2018) [70]	Targeted FIA-MS UPLC-MS/MS (Reversed Phase)	Circadian variation (constant routine 24 h, 11 samples at 1–3 h intervals) Day shift *vs.* night shift (simulation)) circadian *vs.* behavioural control	'Inpatient' Standard constant routine parameters (see [318]) During baseline & shift work—controlled sleep/wake, light/dark cycle, temperature. Standardised meals and mealtimes	Night shift: N = 7 6 males Age 27.6 ± 3.2 BMI 25.6 ± 3.3 Day shift: N = 7, 4 males Age 24.0 ± 2.2 BMI 25.9 ±3.4	65 (49.2%) across both shift patterns, 27 (20.5%) common to both	Amino acids LysoPCs Phosphatdylcholines Acylcarnitines Glycerophospholipids Sphingolipids
Grant *et al*. (2019) [328]	Untargeted & Targeted LC-QTOF/MS (HILIC)	Circadian variation (24 h) Circadian- *vs.* wake-dependent changes	'Inpatient' Standard constant routine parameters (see [318])	N = 13 9 males Age 25.0 ± 4.3 BMI 22.0 ± 2.1	Targeted: Group level 28/99 (28.3%) (rhythmic, rhythmic & linear) 4/99 (4%) linear Untargeted: Group level 361 (22%) rhythmic features 8% linear features Individual level 14% rhythmic profiles 4% linear profile	Amino acids Organic acids
Gu *et al*. (2019) [329]	Untargeted UHPLC-MS (Reversed phase) & GC-MS/MS	Diurnal variation (26–48 h) (48 h time course for N = 2, 26 h for N = 1 participants),	'Inpatient' Standardised meals and mealtimes Habitual sleep time (10 h sleep)	N = 3 2 males Age 20–31 BMI 18 < 29.9	100/663 (15.1%) rhythmic in at least 1 individual 26/663 (3.9%) rhythmic in at least 2 individuals.	Amino acids DGs Lysolipids Phospholipids Steroid lipids

Kervezee *et al*. (2019) [330]	Targeted DI-MS LC-MS/MS (Reversed phase)	Diurnal variation (24 h—2 h intervals between samples) Baseline *vs.* forced misalignment post-simulated shift work	'Inpatient' Controlled sleep/wake, light/dark cycle and hourly isocaloric meals during sampling periods	N = 9 8 males Age 22.6 ± 3.4 BMI 21.3 (19.6–23)	51 (39.2%) baseline 53 (40.8%) night shift 32 (24.6%) both, 24 phase shifted, 27 (21%) significantly changed post-night shift	Amino acids Fatty acids Organic acids Lysophospholipids PCs
Honma *et al*. (2020) [65]	Targeted FIA-MS UPLC-MS/MS (Reversed Phase)	Diurnal variation (70 h, 2 h intervals between samples) 16:8 wake/sleep cycle > 40 h prolonged wakefulness > 8 h recovery sleep	'Inpatient' Standardised meals and mealtimes. Controlled light/dark cycle and activity/posture	N = 12 All female Age 25 ± 4 BMI 24.9 ± 3.6	Total 97/130, 58 (44.6%) common for all conditions. Baseline 78 (60%) 8 unique. Sleep deprivation 76 (58.5%) 5 unique Recovery sleep 80 (61.5%) 5 unique	Glycerophospholipids Sphingolipids Amino acids Biogenic amines Acylcarnitines
Lusczek *et al*. (2020) [331]	Untargeted UHPLC/MS (Reversed Phase)	Diurnal variation (24 h—4 h intervals between samples)	'Inpatient' Self-selected light/dark, feeding/fasting, sleep/wake cycle for healthy participants	Healthy cohort N = 5 2 males, Age 45–72 BMI 22.4–33.3 ICU cohort N = 5 2 males Age 43–66 BMI 31.0–57.3	10 (16.7%) in healthy 0 in ICU	Amino acids Acyl carnitines LysoPEs

Table 3.2 Footnotes: Age and BMI are quoted in standard units, years and kg/m^2 , respectively. Where available, mean age/BMI ± 1 SD given. Significant changes in metabolites identified in studies performing am *vs.* pm (two-time point) comparison(s), should be considered as gradient changes ergo 'gradient metabolite', Significant changes in metabolites identified in studies over a >24 h time course with $n \geq 5$ should be considered as rhythmic changes ergo 'rhythmic metabolite', rhythmicity being detected by cosinor analysis and/or MetaCycle. Rhythmic/gradient features are denoted as such, otherwise the table refers to rhythmic/gradient metabolites. Rhythmic/gradient classes primarily observed are not an exhaustive list of all metabolite classes observed within a study but a summary of the most rhythmic classes, if any, for that particular study, as denoted by the author or inferred from provided data. Abbreviations: DG—diglyceride; DI-MS—direct infusion mass spectrometry; FIA-MS—flow injection analysis mass spectrometry; GC-MS—gas chromatography mass spectrometry; HILIC—hydrophilic interaction chromatography; LC-MS—liquid chromatography mass spectrometry; LysoPC—lysophosphatidylcholine; LysoPE—lysophosphatidylethanolamine; MS/MS—tandem mass spectrometry; NMR—nuclear magnetic resonance; PC—phosphatidylcholine; Q-TOF MS—quadruple time of flight mass spectrometry; SESI-MS—secondary electrospray ionisation mass spectrometry; TOF MS—time of flight mass spectrometry; U(H)PLC—ultra high performance liquid chromatography.

3.2.1. Circadian variation

Dallmann *et al* [317] was first to publish a constant routine study observing circadian control of the metabolome in plasma and saliva (latter discussed in [78]), with 41 (15%) of identified metabolites within plasma displaying circadian rhythms, of which, >75% comprised lipids, with nearly all peaking at subjective lunch time (when participants would expect their second daily meal under a typical sleep/wake, feeding/fasting cycle). These findings provide evidence that lipid metabolism is under circadian control. The amplitude of these observed rhythms differs between metabolites, with lactate abundance varying by ~66% and glutamate by ~40% across the 24 h period as but two examples, with other observed metabolites displaying even greater variation. This is a pertinent finding considering active discussion in the literature for these two metabolites as potential biomarkers and therapeutic targets for various diseases [332-338] and a consideration that can be applied to the findings of all the papers discussed below. In a separate constant routine study by Kasukawa *et al* published in the same year [319] it was reported that phenylalanine, tryptophan, and leucine were observed to display circadian rhythms. In total, eight rhythmic metabolites were observed within the study, a comparison to the 41 observed by Dallmann previously [317] was drawn and the conclusion reached was that the authors likely observed less rhythmic metabolites due to curated selection of features with minimal variability between participants. Chua *et al* [320] observed a comparable level of circadian rhythmic metabolites (35/263 metabolites, 13.3%) in their lipidomics assay, with triacylglycerides (TGs) and diacylglycerides (DGs) peaking in the morning. However, of these metabolites, only 12 (5%) to 86 (33%) displayed rhythmicity, with a median of 20% across all participants. Importantly, there were also significant differences in the phase of the rhythmic metabolites between participants, some as great as 12 hours. These phase differences were observed despite similar cortisol and melatonin rhythms

between participants, metabolite rhythms typically measured relative to DLMO or when cortisol levels peak to account for individual differences in circadian timing (chronotype). As age/sex/ethnicity were all controlled for, the authors concluded that the three observed 'lipidome phenotypes' may be a result of predetermined genetic differences. This is a reasonable assertion echoed by a prior review [339], suggesting that genetic variation in human clock genes could contribute to phenotypic differences and is supported by twin and familial studies stating 27–50% of the variance in diurnal preference/chronotype is attributable to genetic, not environmental, influence [340-343].

These findings suggest that a portion of the plasma metabolome/lipidome is under circadian influence (Table S7 in [78]), the extent of which is potentially confounded by inter-individual differences in diurnal preference/circadian timing (chronotype).

3.2.2. Sleep deprivation and prolonged wakefulness

The aforementioned studies employ a constant routine protocol enforcing wakefulness over the observed time period. In an interesting development, Davies *et al* [321], proceeded by Chua *et al* [323], Grant *et al* [328], and Honma *et al* [65], characterise the impact of wakefulness (total sleep deprivation) on the plasma metabolome (Table S8 in [78]). A distinction between these four studies is that both Chua and Grant performed their sample collection during a constant routine protocol where participants were subject to total sleep deprivation for the duration whilst Davies and Honma employed an entrained protocol with two and three phases, respectively, across consecutive days with a 8 h period of sleep permitted during the first 24 h of sampling proceeded by 24 h of total sleep deprivation with Honma monitoring a third day permitting 'recovery' sleep. The experimental design of both

Davies and Honma studies allowed for paired comparisons between test conditions (e.g., sleep *vs.* no sleep) for individual participants.

Of the 109/171 (64%) rhythmic metabolites observed by Davies *et al* (amino acids, biogenic amines, lipid groups), 95 (87%) of which peaked between 06:00 and 18:00 h, 21 (28%) lost rhythmicity during prolonged wakefulness alongside reduced/increased amplitude of metabolites that peaked within the day/night cycle, respectively. Furthermore 27/171 (16%) metabolites (three sphingolipids, eight acylcarnitines, 13 glycerophospholipids, tryptophan, serotonin, taurine) exhibited a significant increase during sleep deprivation *vs.* baseline sleep with serotonin exhibiting the largest change (44% ± 20%), illustrating a perturbation to metabolite rhythms coinciding with total sleep deprivation. These results were echoed by Chua *et al* [323] and Grant *et al* [328] who observed linear changes in 25/11 (9.5%/11.1%) metabolites, respectively, predominantly phosphatidylcholines/TGs in the former study (lipidomics assay) and amino acids in the latter (HILIC assay), as a result of sleep deprivation. Further agreement between Davies and Chua was observed for diacylphosphatidylcholines, which were affected by sleep deprivation.

Having performed untargeted/targeted HILIC assays, as opposed to reversed phase assays which were performed previously [320, 321], Grant *et al* [328] observed the impact of prolonged wakefulness on polar metabolites, with 28/99 (28.3%) displaying rhythmic, rhythmic/linear metabolites peaking in the biological night, of which 13 were amino acids, a similar result with regard to amino acids rhythms having been observed to some degree previously [292, 316, 317, 319]. Furthermore, Grant identified nine novel rhythmic polar metabolites (organic acids, nucleotides, and an amino acid) not observed in prior non-

polar/lipid studies. Group level analysis of their untargeted dataset mostly reflected what was observed in the targeted dataset (Table 3.2).

Honma *et al* [65] performed a similar study to Davies *et al* [321] with the exception of an all-female cohort (Davies studied an all-male cohort) and therefore offers a novel insight into sex-dependent differences resulting from prolonged wakefulness (Table S4 in [78]). At face value, the female cohort were comparable to the male cohort of Davies, with 58/130 (44.6%) metabolites being rhythmic and common across the three study days (baseline sleep, prolonged wakefulness, recovery sleep), and 97/130 (75%) were rhythmic on at least one of the three days but not rhythmic across all three. During sleep deprivation 15/130 (12%) of metabolites in female participants were significantly altered, of which 14 decreased in concentration, in contrast to Davies, in which 37/141 (26%) were significantly increased during sleep deprivation. Furthermore, a subset of 32 common rhythmic metabolites between the cohorts of both studies was analysed with regard to their mean acrophase (peak time) and it was observed that the acrophase in the female group was ~1 h later compared to the male group (female = 15:48 ± 0:40 h, male = 14:53 ± 0:42 h). Whilst the datasets are from different studies, they were both performed by the Skene laboratory (University of Surrey) with highly similar methodology, the same commercial kit to perform the targeted assay, similar cohorts (excluding sex) and subject to the same data processing and analysis, making these observations more compelling. Complementary to the findings of Davies and of Skene *et al* [70] (discussed below—see Section 3.2.3), the findings of Honma also illustrate how acute disruption of the sleep–wake cycle can result in changes to metabolite rhythms that persist after the disrupted behavioural cycle is restored.

3.2.3. Shift work

In work that runs parallel to the studies investigating prolonged wakefulness and sleep deprivation, Skene *et al* [70] and Kervezee *et al* [330] independently investigated and characterised the impact of circadian misalignment on the plasma metabolome brought about by simulated shift work (Table S10 in [78]). Skene compared simulated day (DS) *vs.* night (NS) shift work and focused on which metabolite rhythms are primarily driven by the central SCN clock or external behavioural cycles.

In the constant routine period succeeding the simulated shift work, Skene *et al* [70] observed 65/132 (49.2%) rhythmic metabolites following one or both conditions, with a further 19 metabolites (seven amino acids, 12 lysophosphatidylcholines) losing rhythmicity following NS and a second independent group of 19 metabolites (mostly phosphatdylcholines, acylcarnitines) gaining rhythmicity only after NS. There were 27 common rhythmic metabolites between test groups, only three of which maintained the same peak time between DS and NS. This was also a common trait with the circadian markers melatonin DLMO and cortisol, indicative to the authors that it is likely that these three metabolites (taurine, serotonin, sarcosine (N-methylglycine)) are strongly influenced by timing of the central SCN clock. Conversely, the remaining 24 rhythmic metabolites (mostly glycerophospholipids and sphingolipids) exhibited a significant shift in peak time between test conditions, with the majority experiencing a 12 h delay, i.e., inversed rhythms, and thus suggested to be strongly influenced by the shifted behavioural cycles (e.g., sleep/wake; feeding/fasting). This delay was maintained in the constant routine following cessation of the shift conditions demonstrating that endogenous metabolite rhythms can be induced and driven by external

behavioural cues which likely reflect peripheral oscillators dissociating from that of the SCN, as concluded by the authors.

Similar to Skene *et al* [70], Kervezee *et al* [330] observed a change in rhythmic metabolites in NS, with 51/130 (39.2%), 53 (40.8%), and 32 (24.6%) metabolites having displayed rhythmicity at baseline, NS, and in both datasets, respectively. Of the 32 metabolites rhythmic in both conditions (baseline and NS), 24 (75%) exhibited a phase shift, on average of 8.8 h, matching the shift in sleep pattern. Therefore, these changes are deemed more strongly influenced by behavioural cycles as opposed to circadian control with amino acids being an enriched group within the 24 metabolites exhibiting a phase shift. A further seven (~22%) metabolites, from the subset of 32, remained aligned with the non-shifted melatonin phase (measure of SCN phase). Therefore, this subset of seven metabolites (Table S10 in [78]) were deemed circadian SCN clock regulated, or at least not regulated by the sleep/wake cycle. The 19 metabolites that lost rhythmicity post-shift work simulation were predominantly lipids, similar findings shown by Davies *et al* [321] and Chua *et al* [323] when investigating the impact of sleep deprivation and corroborating with the results of Skene [70].

The observation that behaviourally induced rhythms can be retained and persist for at least 24 h in free-running constant routine conditions complements the findings of Davies *et al* [321] and Honma *et al* [65] thus shift work and atypical sleep patterns are important considerations for metabolomics studies prior to recruitment and sampling.

3.2.4. Diurnal (24 h) rhythms

Parallel to studies employing constant routine methodologies to investigate circadian rhythms are those which investigate diurnal rhythms, 24 h rhythms under entrained conditions. Three further studies, not investigating shift work or the impact of sleep, were

obtained from the performed literature search characterising time of day variation over >24 h time courses. Park *et al* [315] identified 34 metabolites exhibiting time of day variation between three time classes ('morning', 'afternoon' and 'night') covering a time course of 24 h. These distinct time classes were produced by averaging hourly 1H-NMR (nuclear magnetic resonance) spectra across all 10 participants for the 25 time points and by performing PCA analysis. Identification of time of day variation was performed with a bespoke two sample t-test, with correction for multiple testing, performed to test for significant differences in spectral regions between the three 'time classes'. Ang *et al* [316] showed time of day variation across 24 h in a range of metabolite classes using untargeted LC-MS (liquid chromatography mass spectrometry). They considered their work as an external validation of the Dallmann *et al* study [317], having replicated some of their findings despite some key differences in methodology (Table 3.2). In total, 34 rhythmic metabolites were identified with variation in abundance ranging from 49–81% (average 65%). The magnitude of this variation was similar to that observed by Dallmann *et al*, with amino acids and phospholipids being two highlighted rhythmic metabolite classes that displayed similar amplitudes. Gu *et al* [329] observed diurnal variation in three participants within their pilot study to perform individual-level analysis to better explore inter-individual variability as opposed to group-level analysis; with only Chua *et al* [320, 323] having observed individual rhythms and differences prior to this publication and Grant *et al* [328] also performing individual-level analysis and publishing in a similar timeframe. Gu observed diurnal rhythms in all three participants (Table 3.2, Table S2 in [78]) in addition to observing inter-individual differences in the acrophase of metabolite rhythms relative to DLMO, such differences may be masked in group-level analysis but are pertinent when investigating on the individual level.

3.2.5. Health status

It is well established that different physiological and pathological states caused by disease can lead to distinct metabolic profiles. Isherwood *et al* [292] adds to this understanding by observing unique time of day variations accompanying specific phenotypes (Table S11 in [78]). Their study observed differences in metabolic profiles of individuals with type 2 diabetes mellitus (T2DM), overweight/obese (OW/OB) non-diabetic individuals, and age-matched lean 'healthy' controls; the novel aspect of the study being the comparison of metabolic profiles between non-diabetic/diabetic individuals (age and weight matched) across multiple time points about a 24 h time course. In total, 50/130 (~38.5%) unique metabolites displayed 24 h rhythmicity of which 35, 39, and 20 were observed in lean, OW/OB and T2DM groups, respectively, from group-level analysis (Table S2 in [78]). Of these 50 rhythmic metabolites, 5 (10%) were unique to OW/OB and T2DM groups and 11 (22%) were unique to non-T2DM groups suggesting a change in metabolite rhythms associated with the onset or continuation of T2DM (Table S2 in [78]).

In a similar line of enquiry to Isherwood *et al*, Gehrman *et al* [326] investigated how diurnal rhythms may differ between participants confirmed to suffer with insomnia and matched healthy individuals ('good sleepers'). Unique metabolite profiles were observed for both test groups (Table S11 in [78]) with 29 metabolites being elevated/decreased in insomnia patients in the morning and/or night. Of these 29 metabolites, 13 were rhythmic. In total, 11 metabolites exhibited diurnal rhythms in both groups with a further six and seven unique rhythmic metabolites in healthy and insomniac participants, respectively. Phase changes were also noted with a phase advance (peaking earlier in the day) of acetone, proline, and a phase delay (peaking later in the day) of lactate, valine, isoleucine and 3-methyl-2-oxovalerate

for the insomniac participants. It remains unclear how many of these changes are associated with the underlying causes of insomnia as opposed to the symptoms, i.e., poor sleep quality/duration.

Lusczek *et al* [331], investigated whether intensive care unit (ICU) patients have their 24 h rhythms disrupted relative to healthy controls by profiling “circadian” rhythms in vital signs and plasma metabolites by analysing 60 “circadian” metabolites based on the work by Dallmann *et al* [317] and Ang *et al* [316] (who observed diurnal not circadian rhythms) and concluded that ICU patients experience desynchrony, leading to a loss of metabolite rhythmicity when compared to healthy non-ICU controls. The rhythmic metabolites observed by Dallmann and Ang were observed in a different and specific context, and do not serve as adequate or reliable markers for circadian desynchrony in the context of the study by Lusczek. Further measures taken to assess patient “circadian” rhythms included temperature, heart rate and blood pressure, and plasma cortisol of which only cortisol was deemed significantly rhythmic for group-level analysis of the ICU group. Lusczek suggests that this is indicative of a lack of coherence in circadian phases and amplitudes amongst the ICU patients. However, an alternative suggestion is that coherence of cortisol rhythms, is indicative of a similar phase angle of entrainment between participants and the variable temperature and blood pressure rhythms result from patients’ unique circumstances with regard to their disease status (cardiac or respiratory), physiology/trauma, concomitant medication and demographics [344, 345], disease state and physiology potentially leading to unique rhythmic metabolite profiles not necessarily related to desynchrony as established above.

3.2.6. Diet composition

The previously described studies have all collected blood samples at 2–5 h intervals across a minimum 24 h time course to define and visualise detected metabolite rhythms over a 24 h period via cosinor analysis (applies the least squares method to fit a sine wave to time series data) or MetaCycle (runs three separate algorithms to detect biological rhythms). Some studies may opt to characterise time of day variation over a shorter time course (<24 h) or with fewer samples per 24 h cycle (n ≤ 4 samples) and thus investigate variation in a morning *vs.* evening fashion to discern 'gradient' metabolites ('gradient' referring to a significant relative increase/decrease in concentration/abundance of a metabolite between two time points) such as the work performed by Sato *et al* [327]. To clarify, low resolution sampling of n ≤ 4 samples per 24 h cycle is insufficient to characterise 24 h metabolite rhythms and so morning *vs.* evening studies are inherently limited in determining time of day variation with results typically limited to t-test/ANOVA analysis to determine significant changes between time points. These 'gradient' metabolites may or may not be rhythmic and a higher resolution (n > 4 samples per 24 h cycle) is required to discern rhythmicity.

Sato *et al* [327] investigated the impact of a high-carbohydrate diet (HCD) and high-fat diet (HFD) of equal calorific content to observe the impact of nutritional challenge on the serum and skeletal muscle metabolome, whilst noting diurnal changes. Many time of day changes were observed as a result of the diet in the 1063 detected metabolite features (Table S12 in [78], Table 3.2). For the HFD and HCD diet conditions 85/50 metabolites displayed a gradient change at baseline only, 126/95 post-completion of diet conditions only, and 138/142 both before and after diet conditions, respectively. Metabolites relating to lipid metabolism were enriched in the HFD group regardless of time of day—perhaps unsurprisingly. Conversely, half

of all decreased metabolites in the evening as a result of HFD were related to amino acid metabolism despite equal calorific intake from protein between groups. Relative to HFD, HCD led to a decrease in metabolites relating to lipid metabolism, regardless of time of day. The authors suggested that increased serum insulin concentrations lead to suppressed lipolysis, explaining the observation, though the literature quoted in support of this statement observed this effect during high-activity (exercise) conditions. Sato *et al* have demonstrated that diet composition, with regard to calorific intake from major food groups, displays an interaction with time of day variation of the metabolome altering specific metabolite rhythms; in the context of this study lipid metabolism was most strongly impacted.

3.2.7. Morning versus evening studies

Kim *et al* [322] observed diurnal variation over 14 h in plasma (11 metabolites, 9%). The proportion of variance attributable to time of day in plasma was minimal *vs.* that attributed to patient 'effects' (age, sex, race, polycystic kidney disease ~40%) and residual variance (>50%). Data on sleep patterns, chronotype, work and light/dark history were not considered and so a proportion of this residual variance may well have a biological source, e.g., associated with chronotype, when considering the confidence in consistency and rigour by Kim *et al* of their employed methodology.

Skarke *et al* [324] observed diurnal variation in plasma (9/166, 5.4%) between 12 h samples over 48 h. Notably, less rhythmicity was observed when compared to the studies discussed above and may potentially be the result of reduced sampling frequency (12 hourly), the small cohort, or the purposeful monitoring of participants in their habitual routine undergoing unique and varied daily cycles. Despite the introduced inter-individual variation due to this setting, diurnal variation is still evident, suggestive of a high potency in temporal regulation

of specific metabolites driven by external and internal rhythms. The reduced number of rhythmic metabolites from both of these studies compared to those discussed earlier aptly illustrates the difficulties in monitoring time of day variation and the required rigour in experimental design. To see beyond the noise and identify significant changes as a factor of time (clocks and/or circadian time), it is of importance to control environmental conditions, patient demographics and lifestyle.

3.2.8. Summary of the field

Reviewing the literature has revealed the following key findings.

- The number of studies investigating time of day variation of the human metabolome, to date, is small (n = 29) with blood the most investigated sample type (n = 18).
- The majority of studies are either male only (n = 11) or male majority (n = 11) cohorts, with female only cohorts (n = 1), female majority cohorts (n = 2) or balanced cohorts (n = 4) being the minority.
- Endogenous metabolite rhythms, regulated by the circadian timing system, have been observed via constant routine studies in blood and saliva and include lipids and polar metabolites.
- Metabolite rhythms (timing and amplitude) may be sex dependent although sex has not been regularly investigated with regard to differences in 24 h metabolite rhythms.
- Diurnal 24 h metabolite rhythms potentially evoked by external cues, either environmental (e.g., light/dark cycle) or behavioural (e.g., sleep/wake;

feeding/fasting), have been observed in blood, urine, saliva, breath, and skeletal muscle.

- Acute changes in external cues, e.g., sleep/wake, feeding/fasting, activity/rest cycles and shift work, result in acute alterations to metabolite rhythms (timing and amplitude) that can persist after cessation of the change.
- Specific physiological phenotypes and healthy *vs.* diseased state are shown to result in unique diurnal rhythms alongside the expected metabolite profiles of each phenotype.
- Lipids, in particular glycerophospholipids, and amino acids are the most frequently observed rhythmic metabolite classes. Lipid rhythms have shown the most variation between individuals with differences in phase (timing).
- Lipid rhythms may feature class-dependent temporal separation based upon carbon chain length and degree of saturation.
- A subset of metabolites are repeatedly reported as undergoing significant time of day variation across studies. A total of 35 putatively identified metabolites having been observed in at least five studies (Table S1 in [78]) out of a total of 400 putatively identified across all studies.

Whilst blood was the most investigated biofluid (n=18), unsurprising considering its popularity within metabolomics as discussed above, only 5 of these studies applied a constant routine protocol to investigate circadian regulation of metabolites [70, 317, 319, 320, 328], the gold standard in assessing circadian rhythmicity [318]. Dallmann [317] and Skene [70] employed at least two assays within their respective studies (GC-MS/RPC/UHPLC-MS, FIA-

MS/UHPLC-MS) yet no study has yet to employ both a RPC and HILIC assay in an untargeted fashion, a common approach in untargeted metabolomics discussed previously. Furthermore, three of the five studies: Chua [320], Skene, and Grant [328], applied targeted assays which offer limited coverage of the metabolome. Finally, all of these studies were male dominated with Dallmann, Kasukawa [319] and Chua investigating male only cohorts whilst Skene and Grant had majority male cohorts (Table 3.2). Exclusion of females is often a deliberate decision as it removes the menstrual cycle, another biological rhythm, as a confounding factor. Nevertheless, there are solutions to include females within circadian studies such as monitoring/collecting samples from naturally cycling females at the same phase of their menstrual cycle e.g., the follicular phase. Alternatively, one could recruit females on oral contraceptives [346], however, this introduces an additional variable to consider during data analysis. As it becomes increasingly apparent that sexual dimorphism pervades the mammalian clock [347, 348], metabolism [349], and metabolic disease pathophysiology [350, 351], we can no longer consider males as an adequate proxy for the human population within circadian studies.

3.3. Investigating metabolite rhythms – the next steps

A common outcome for many of the studies reviewed above was the identification of rhythmic/gradient metabolites; with a second common feature for many studies being the analytical platform employed. Upon observing the methods of the reviewed studies it becomes quickly apparent that particular assays have been collectively favoured with variations of LC-MS the most commonly used across 19 studies (two of which applied HILIC assay and 17 applied reversed and/or normal phase assays) followed by NMR (six studies), GC-MS (gas chromatography mass spectrometry) (six studies), FIA-MS (Flow injection analysis

mass spectrometry) (four studies), SESI-MS (two studies), DI-MS (direct infusion mass spectrometry) (one study). Each of these platforms possesses its own advantages and limitations with regard to what metabolites can be detected thus influencing the collected dataset and metabolome coverage, resulting analysis and interpretation [79]. With this in mind, the curated list of metabolites undergoing time of day variation is somewhat limited and indeed biased towards lipids and non-polar metabolites which see better retention within reversed phase assays. By contrast polar metabolites may be underrepresented since they exhibit poor retention in these assays and are better retained and detected within HILIC assays. Further utilisation of HILIC methodology may yield additional rhythmic metabolites similar to the outcomes of the study by Grant *et al* [328]. It is also of interest to note that whilst various studies used multiple analytical platforms and assays in tandem, no study has yet to incorporate both HILIC and reversed phase assays in parallel producing a more 'complete' dataset and coverage of the metabolome. Further untargeted studies employing a greater diversity of UHPLC assays would benefit this growing body of work moving forward. Despite this potential limitation over 400 metabolites have been observed to be either significantly rhythmic or undergo significant time of day changes in a morning *vs.* evening fashion (Table S1 in [78]) based upon putative or definitive metabolite identification. Of these metabolites, 35 have been observed in at least five studies (Table S7 in[78]). This compiled information provides insight into metabolic pathways likely influenced by time of day variation with amino acids and their derivatives being amongst the most frequently observed rhythmic/gradient metabolites. This information provides an adequate starting point for the development of bespoke targeted assays to investigate such pathways, quantify observed rhythms in metabolite concentrations similar to Davies *et al* [321] and consider the biological significance of the rhythm, or lack thereof, under particular conditions such as shift work or

disease state [70, 292, 330]. Expanding upon the number of biofluids/tissues investigated to observe diurnal/circadian variation presents a challenge due to the need for repeated and regular sampling over a 24 h time course, methods used for single sample collection, such as a biopsy, may thus prove too invasive or impractical for repeated and regular sampling. However, ambulatory microdialysis sampling techniques capable of high resolution sampling in humans hold promise [352]. In the near future it may be more practical then to expand upon the base of work on the five currently investigated sample types of which serum/plasma have been favoured leaving saliva, urine, skeletal muscle, and breath under investigated by relative comparison.

A general challenge for identifying and validating biological rhythms in -omic datasets across studies is the distinction between rhythmic and non-rhythmic time series. Many detection algorithms or combination of algorithms have been proposed. Two of these have already been mentioned above (i.e., cosinor and MetaCycle) but many other workflows and applications such as Multi-Omics Selection with Amplitude Independent Criteria (MOSAIC) [353], Rhythmicity Analysis Incorporating Non-parametric methods (RAIN) [276], Extended Circadian Harmonic Oscillator (ECHO) [275] and others [279, 354] are available. As any one method is open to critique, MetaCycle, for example, is already combining a number of different methods to determine different regulation of rhythmicity in different groups of a study or between studies. Venn diagram analysis (VDA) employs any one of the methods above, e.g., RAIN, to identify changes in rhythmic items (transcripts/metabolites) between test groups. Recent findings, however, highlight inter-group and even inter-study comparisons where VDA has overestimated differences in rhythmic items [280]. The authors highlight the shortcomings of VDA and propose a novel approach to circumvent these issues implemented in the R package compareRhythms that compares circadian parameters

(amplitude and phase) between the groups under comparison. This allows compareRhythms to discern between metabolites that have remained the ”same” (rhythmic across test conditions) and have undergone a “change” (still rhythmic but phase/amplitude change between test conditions) whilst VDA cannot.

3.4. Chapter summary

In summary, the primary objective of this review was to establish the sample types in which time of day variation of metabolite concentrations have been reported using a metabolomics platform, with a focus on identifying rhythmic metabolites. The extent of this time of day variation on the complete metabolome has also been reported to highlight the number of detected metabolites which have been shown to vary with time. The metabolome of blood, urine, saliva, breath, and skeletal muscle are influenced by diurnal and/or circadian rhythms. This most likely extends to other human biofluids and tissues in a similar fashion to how gene transcripts are rhythmic across a range of tissues in mammals [74, 355]. Changes to external time cues (Zeitgeber), such as the light/dark and feeding/fasting cycle, result in changes to these rhythms and should thus be considered potential control variables where applicable in metabolomic studies, e.g., enforcing a specified light/dark, sleep/wake protocol for all participants, or setting meal times, excluding shift workers. Furthermore, it may be pertinent to collect additional metadata from participants such as history of shift work, sleep/wake times and a person’s chronotype, complete time/calendar date and geographical location in which samples were taken, all of which may influence rhythmic metabolite profiles, the resulting analysis and biological interpretation.

CHAPTER 4

DEVELOPMENT OF INTEGRATED DATA HANDLING & ANALYSIS TOOLS.

4.1. Introduction

As indicated by its -omics moniker metabolomics deals with large datasets which in their raw form may comprise gigabytes/terabytes of data and when processed produce a data matrix which contains thousands if not tens of thousands of metabolite features (*m/z*, RT pairs) in an untargeted study [356]. This large volume of data cannot be analysed without software assistance. The number of available software tools for assisting with metabolomics raw data processing and data analysis is rapidly increasing and covers aspects such as: raw data file format conversion, raw data processing, pre-processing, quality assessment and filtering, metabolite annotation and identification, and metabolite databases [357]. How such tools assist in converting raw data to a 2D data matrix was discussed in Chapter 1. Metabolomics therefore relies on a range of skills from other disciplines to correctly collect, process, and analyse data [79, 356]. In between these software assisted stages of data analysis, datasets will be manually handled e.g., check data quality/inspect data, save/archive, share/distribute. A potential concern at this juncture is inadvertent data handling errors that may be introduced into the data due to the multiple stages of processing and analysis in addition to potential handling by multiple individuals that may comprise a metabolomics research team or collaboration. This concern is shared by other fields transitioning into the realm of big data and developing tools to assist with data handling [358]. Human errors, with regards to metabolomics data handling, may range from random accidental edits or deletions of datapoint(s)/sample(s), systematic errors by incorrect application of data transformations/filters to all or subsets of the data, inaccurate labelling/file names or poor record keeping of applied analysis parameters to name but a few. Human errors may not be evident until the user notices the 'result' of an action does not align with their expectation and may not realise they are the cause of the fault in the first instance [359]. This logic is

applied to simple tasks; however, due to the size and scope of the typical metabolomics dataset, errors may go unnoticed due to the lack of a preconceived expected outcomes. For mistakes which are eventually realised it may not occur at the point of the mistakes inception but rather later in the workflow; at best this mistake is fixed with little cost to time or at worst goes unnoticed and compromises the final conclusions. Thus, there is a concern that manual data handling, beyond what is strictly necessary, may negatively impact the rigour and reproducibility of a metabolomics study. Even when free of error, reproducibility is already a concern within the field of metabolomics due to analytical and unexplained biological variation [78]. Furthermore, differing methods of data analysis, such as pre-processing, can lead to divergent results and subsequent different biological interpretations [360]; thus, set protocols with adequate record keeping need to be in place to facilitate reproducibility in addition to mitigating user errors. Whilst reproducibility and experimental rigour are important in any metabolomics study an argument could be made that they are most important within an untargeted study in which we are attempting to establish a foundational knowledge in which future validated studies may derive from. Considine and colleagues makes a statement on such an issue in their review on biomarker discovery discussing how flawed discovery based studies inevitably lead to failure of subsequent validation studies and poor reproducibility [301].

An additional challenge to consider when applying metabolomics to chronobiology is the time series/repeated measures in which samples were collected from and preparing data to undergo rhythm analysis. To detect ~24 h rhythms in relative abundance, multiple established rhythm detection algorithms may be employed and their results cross referenced to classify individual metabolite features as rhythmic for a ~24 h period. This cross-reference approach is taken to circumvent disadvantages associated with any single algorithm, as reviewed here

[361], and has been applied numerous times within the field [65, 362-370]. Utilising even a single rhythm detection algorithm poses a logistical challenge with regards to data handling as current commonly used rhythm detection algorithms are limited to a single data frame being input at a time and outputting a single result file. Whilst not a significant problem to surmount if only 1-2 data frames (e.g., control group, test group) need to be submitted the situation can quickly devolve if performing individual level analysis where each participant would be subject to each rhythm analysis algorithm independently. Solutions to this problem were limited with few pre-existing rhythm detection software packages incorporating multiple rhythm detection methods into a single analysis prior to the start of work on this thesis chapter; the primary example being an R package known as Metacycle which incorporates three different methods [272]. Recent developments in the field have seen additional R packages and applications address this problem such as DiscoRhythm [270] and BioDare2 [371]. A further consideration is the unique formatting and input data for the various available algorithms which would involve sub-setting or reformatting the feature intensity matrices prior to processing via rhythm detection and then collating the results in such a way that they can be interpreted easily by a human user. This then produces a logistical challenge in accurate record keeping and storage of data/results in addition to incurring a time cost to manually format input files and collate result files. Issues associated with the manual handling of data, such as those described above, can be circumvented by automation. This automation would require custom built software scripts written within or obtainable through open access platforms such as R, Python or directly within an operating systems command line/console, as developed for in metabolomics already [357]. By utilising an open access platform, scripts written and published by other individuals can be leveraged and used to construct custom use tools. Such tools would emulate processes that would otherwise be

manually performed on metabolomics data, e.g., sub-setting/filtering of data or applying data transformations, but the tools could iterate such tasks at a far greater efficiency than a human user whilst minimising error. Software tools are not error free, as they are only as good as their developer and the user, but an important difference is errors due to poorly written/executed code are traceable, reproducible and thus fixable. Whilst an initial time investment would be required to create tools to automate a given manual process a well written tool is adaptable and applicable to multiple datasets, hence the active community developing within R and Python would benefit. As metabolomics data, regardless of the initial MS platform employed or untargeted/targeted approach, is typically processed into a 2D matrix (intensity/concentration per feature by sample), generalised tools can be constructed around this standardised format. Additionally, this 2D matrix can be used a source file to automate the required formatting steps to transform the standardised 2D matrix format into a compatible matrix for the rhythm analysis algorithm(s) of choice. This would allow multiple rhythm detection algorithms to be applied and the results compiled with minimal effort and time investment from the user. Automated data processing via a programming language such as R would also minimise manual data handling of data, ergo reducing the risk of human error. Furthermore, as processing and data handling parameters would be present and defined in code and archivable, as opposed to being performed *ad hoc* by a human user and relying on manually produced notes, reporting these accurately is easily achievable. Finally, the time investment required to develop data handling tools would be offset by the time saved once said tools are applied to future metabolomics studies and would partially alleviate one of the major disadvantages of untargeted studies and applying numerous rhythm analysis algorithms which is time consuming data analysis.

The purpose of this chapter is to (1) use pre-existing tools (e.g., XCMS) develop a workflow to assist in preprocessing of metabolomics data and assembly of a feature intensity matrix and (2) develop tools to assist in: data transformation (Z-scoring), data subsetting & rhythm analysis (group-individual level), collation and visualisation of results.

4.2. Results & Discussion

Figure 4.1 displays the major steps with examples for the automated workflow. Processes within this workflow that are now fully automated or tool assisted to alleviate the workload burden and facilitate reproducibility are discussed below. Each stage of the pipeline can be applied individually when required, if the pre-requisite data exists (e.g., produced manually) or applied as a single workflow where subsequent steps inherit outputted data, directories etc., from previous steps. Whilst not necessary to use the code, understanding the definition of basic coding statements illuminates their inner workings. The workflow below commonly applied the following statements:

- 'for loop' – Iterates a process until a user defined condition has been met, e.g., find the highest recorded intensity value per male participant
- 'nested for loop' – a for loop within a for loop, e.g., find the highest recorded intensity value per male participant per time point
- 'if statement' – a conditional action e.g., if X < Y delete X
- 'else statement' – provides an alternative conditional action should the first 'if statement' condition not be met, if X < Y delete X, else delete Y
- Multiple 'if else' statements can be 'nested' within one another allowing for a hierarchy of conditions to be tested that can act to filter data whilst multiple nested for loops can be used to iteratively extract or transform data for examples.

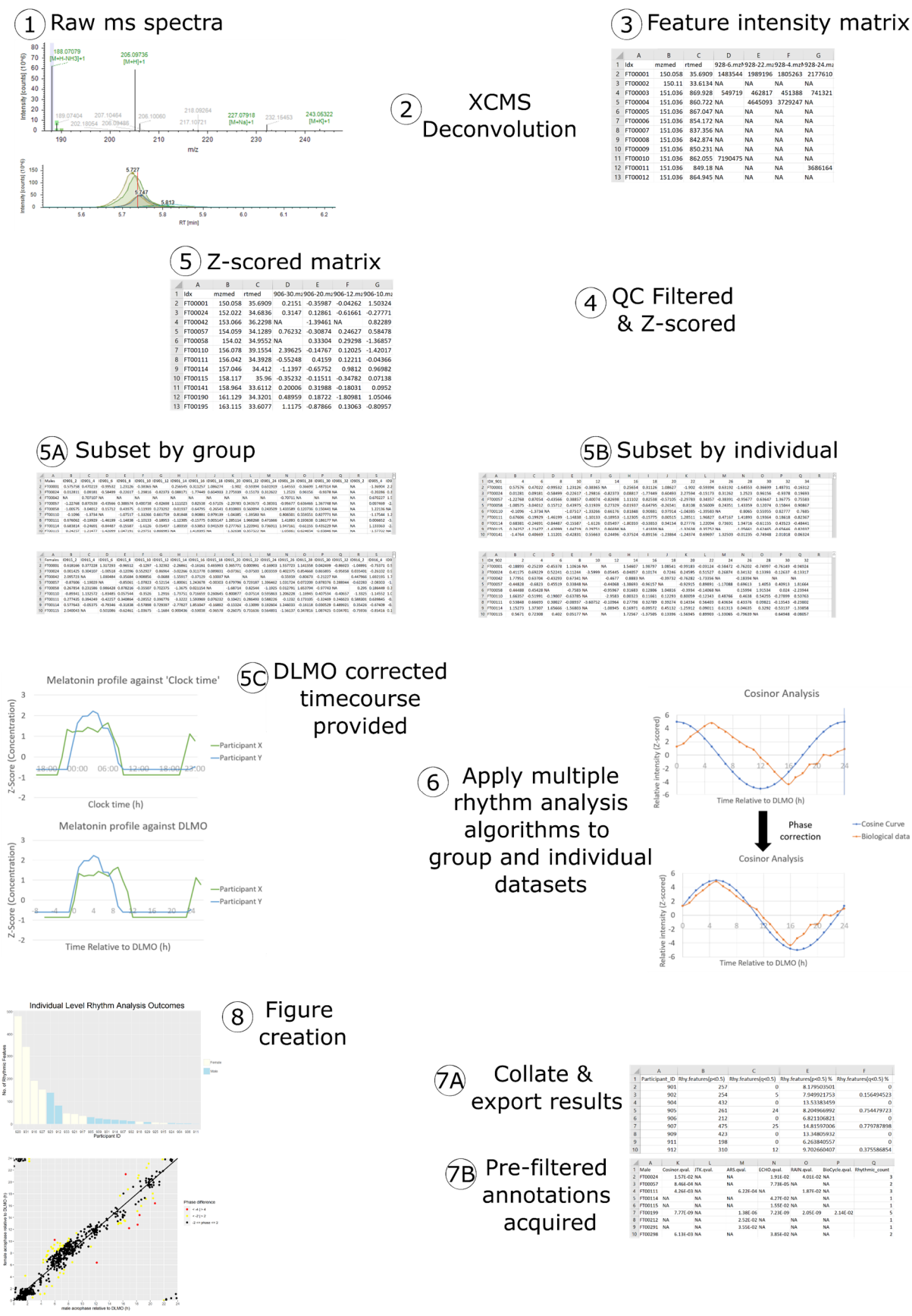


Figure 4.1: An illustrated workflow of the R coded pipeline. The pipeline starts with the deconvolution of collected data, transforms and arranges the data in preparation of rhythm analysis before finally collating results ready to be called upon to create illustrative figures. Additional tools to assist in filtering out or renaming compounds acquired from annotation software are also applied prior to collating results.

1-2: Pre-processing (deconvolution)

Pre-processing requires manual conversion of .raw vendor files to a compatible format (e.g., mzML) via Proteowizard and input of user defined parameters for XCMS processing. User defined inputs include the directory of mzML files, a directory to output results such as the feature intensity matrix, and labels required for data handling/organisation e.g., assay type, identifiers for QC and blank samples. Furthermore, the user can define parameters such as peak width and ppm tolerance, as described in Section 1.4.1., required by XCMS to process the data. The script will then run XCMS (data import, peak detection, peak alignment & RT correction, peak grouping(correspondence)), and export the feature intensity matrix, a record of the R session and the XCMSnExp set (an R object containing the outputs and history of the XCMS processing) which can be re-loaded into R to continue or edit the XCMS analysis. Dependent on the size of the dataset, running XCMS can require days to weeks of processing time on a standard desktop. The script here was written to provide feedback to the reader on progress and, if run on MacOS, can be run via parallel processing. A further quality of life implementation is the exporting of the XCMSnExp set object which can be re-loaded back into R in a matter of seconds to minutes allowing the original XCMS results to be more easily shared, re-visited (e.g, re-perform retention time correction, visualise spectra), or simply preserved as a record.

3-4: Quality control filtering

The quality control (QC) filtering script takes the feature intensity matrix, as outputted by XCMS, and filters the metabolite features based on three parameters commonly applied in metabolomics to improve data quality [163] (see Section 2.3.2.7.). The three parameters filter out metabolite features that exceed the allowable tolerances as set by the user allowing them

to tailor the tool to meet their needs. The parameters are defined as: <Missing_value_threshold_QC> assigns allowable percentage of missing values, per feature, across the intra-study QC samples, e.g., if 0.3, 30% missing values are tolerated or conversely 70% of values must be present across all QC's per feature; <RSD_threshold> assigns the RSD (relative standard deviation or coefficient of variation) cut off, applied per feature, across the intra-study QC samples e.g., if 0.3, features with a measured RSD > 30% are removed; <Blank_cont_threshold> filters any feature whose average intensity in the biological samples is below the average intensity of the process blank samples as a percentage assigned by the user, e.g., if 0.1 this is 10% or a factor of <10x). The filters are applied iteratively such that, if a metabolite feature fails the first parameter it is removed without need of checking the second. This also allows flexibility in the future to add additional filtering parameters without major restructuring of the code should the metabolomics community update standard practices in this regard. The QC filtered matrix is outputted as a .csv file compatible with Microsoft Excel. Prior to writing this code, QC filtering would have been a manual process performed on a .csv file in software such as Microsoft Excel. This would require the use of Excel's formulae and functions to calculate the above parameters per feature and manually delete rows failing these parameters. This process would need to be repeated independently for each assay on a per study basis and may even need to be re-performed if the original data was altered (e.g., re-processed via XCMS, outlier samples removed). The script also allows for specified QC samples to be excluded from the above calculations should they be identified as outliers, e.g., after checking data quality via PCA. The script and outputted file would then act as a record for what QC samples were used for filtering purposes and which, if any, were excluded.

4-5: Z-scoring – data transformation

Calculating a Z-score requires three pieces of information: n (the intensity value of a data point (Feature X for participant Y at time Z)), μ (the mean value of all datapoints (n's) for feature X for participant Y), σ (standard deviation across all datapoints (n's) for feature X for individual Y). The following formula was then applied (n - μ / σ) to calculate the Z score. The function was performed with the QC filtered dataset, as each sample contains a unique identifier corresponding to the participant the sample was taken from the QC filtered dataset was then subset to create a new data matrix per individual/group. A separate data frame can therefore be produced for participant/group A, B, C and so on, with this process being automated within the workflow.

The sub-setting of data prevents miscalculation of the mean/standard deviation and the subsequent Z-score by isolating the data to be transformed on a per participant basis. Once data was subset, average and standard deviations calculations are calculated, on a per feature basis, by initiating a 'for loop' across all features in the dataset and a nested 'for loop' used to perform the Z-score calculation per time point within a feature. The original index information used to subset the QC filtered matrix is applied in reverse such that the original non Z-scored intensity values for a participant are over written by the calculated Z-scored values. The Z-scored matrix is outputted as a .csv file compatible with Excel. Calculation of Z-score is possible without this code, for example within Excel, but requires careful use of formula to calculate the average, SD and Z-score per feature per participant. Incorrect co-ordinates in any of the Excel formula would result in an incorrect Z-score which at best may affect only a single feature for a participant or at worst a systematic miscalculation across the entirety of the dataset.

5A/B: Data sub-setting for rhythm analysis

To successfully perform individual level rhythm detection using algorithms such as cosinor, subset data matrices per participant are required. The same is true for group level analysis where separate matrices per group are required. This script subsets the Z-scored feature intensity matrix into designated groups, e.g., control vs test or male vs female, as well as per individual and creates separate labelled directories to store the data for each UHPLC-MS assay. Sub-setting requires R to be provided participant identifiers, as required for Z-scoring, and user defined groups provided as two vectors listing which participants belong to group1 or group2. As sub-setting may result in the creation of hundreds of subset matrices individual files are saved to unique directories based upon the MS assay and labelled by participant identifier or group. These directories can be re-formulated by subsequent steps as they are written to a standardised structure. Once again this code emulates a task that could be performed manually, however, manually subsetting data for different subjects and for multiple UHPLC-MS assays would result in the generation of hundreds of datasets. This would be a laborious time consuming task which may facilitate errors and mislabelling of data. Moreover, it is a task which may need to be repeated should the data preceding this step be re-processed in some way or when analysing a future dataset.

5C: Formatting for rhythm analysis

Different rhythm detection algorithms require sample (column) headers to be formatted in variable ways but ultimately contain a participant identifier and a (DLMO corrected) time point (e.g., MaleA_time0), for group level analysis, or just a time point for individual level analysis. The solution developed to format sample headers was to repurpose a file mandatory for rhythm analysis via JTK/ARS. Hereto referred to as the ‘study design’, which is a simple

table (produced manually) that contains a list of all sample identifiers, the participant/group labels, and time relative to DLMO the sample was collected (calculated manually). This design file can be used as a point of reference whenever R requires the aforementioned information e.g., a list of all male samples or all time points at DLMO 0 h and can also contain relevant metadata on each file e.g., sex, BMI. Thus, new headers with the required formatting and sample data can be created from the original participant identifiers by referring to the study design file making the process automatable within the workflow.

6: Rhythm analysis

An additional concern with rhythm detection algorithms is variable tolerance to missing values. A second filtering stage was implemented, prior to each rhythm detection algorithm being applied, that removes any metabolite features that possess more missing values than the set threshold per participant e.g., 0.5 (50%) (see Section 2.3.3.3.). It is possible that a given time point contains only missing values, e.g., due to a missing sample, which may cause issue with some of the applied rhythm detection algorithms (ARS, ECHO, RAIN, BioCycle). Therefore, additional checks, such as 'if statements', were employed to check for and correct issues in the data matrix prior to its submission for data analysis. Both the final data matrix submitted to the rhythm detection algorithm and the outputted results file are exported. This creates a record of what features were actually subject to the analysis and the outcomes. Different rhythm analysis algorithms require a variable amount of time to process data [361] which ranged from seconds to days for the data presented in Chapters 5-7. This script allows the user to click once and wait for the rhythm analysis to be completed automatically in its entirety as opposed to manually checking and submitting a file for analysis every few seconds/hours/days.

7A: Collating rhythm analysis results

The results file for a given rhythm analysis algorithm will contain a list of metabolite features, for that assay, and associated p-values denoting significance or lack thereof for rhythmicity (pre and post Benjamini-Hochberg FDR correction). Complicating matters is that, for individual analysis, each participant will have a separate results file specific to them per algorithm and metabolomics assay applied. Therefore, a system needs to be put in place that is capable of reading the multiple results files for an individual and counting/recording the significantly rhythmic metabolite features by extracting features that meet a user defined significance threshold (e.g., $p < 0.05$). For clear organisation of data a simple system of constructing a directory from a user designated root e.g., "C://StudyX_Results/" allowed for subsequent directories to be built up as required with information R was already provided such as the assay type e.g. "C://StudyX_Results /Lipids_pos" or the outcomes of a specific analysis for a given participant identifier e.g. "C://StudyX_Results /Lipids_pos/Rhythm_analysis/2D_cosinor/901". This file system is not the sole method in which data could be organised but strikes a balance of being easily navigable by a human user in an OS GUI, such as Windows file explorer, as file labels are informative and tiered. As R is responsible for assembling the directories it is capable of recreating directory addresses when required and can utilise tools in R such as 'list.files' than can locate existing files in a directory, in this instance, based on the study, the assay, or even the participant. Thus, R is capable of iteratively (re)creating the directories for all results files, importing and reading the results data with a simple 'for loop' function. As each data file is iteratively read into R the script extracts the required information and stores it in a purpose built data able that summaries the results into a more user friendly format, automatically. Moreover, organisation of the

results in this manner also ensures provenance is maintained for the analysis, allowing for results to be easily traceable to their source files should they ever need to be revisited.

7B: Annotation filtering

Whilst software assisted annotation of metabolomics data is a vastly more time efficient process than manual assignment it can suffer from a lack of accuracy and provide annotations irrelevant (false positives) to the study/sample being investigated. To elaborate, the annotation software BEAMSpy acquires annotations through accurate mass matching to public databases. A given metabolite feature (*m/z*, retention time pair) may have multiple 'hits' in such databases which could include drugs and their derivatives, specific dietary metabolites, or other such off-target annotations not anticipated within the sample. Additionally, there may be a degree of redundancy reported by the assistive software due to enantiomers, e.g., L-tryptophan and D-tryptophan, synonyms or the same metabolite but matched from different databases. Finally, the software may provide excessive precision, e.g., 9Z-octadecenoic acid, where the position of a double bond on a fatty acid cannot be determined by the accurate mass measurement used to procure the MS1 level annotation, thus the species level annotation (FA 18:1) should be provided. The removal of false positive metabolite annotations cannot be simply automated as what constitutes a false positive is context dependent and thus typically requires user experience and insight. However, if the user has prior knowledge of what they would like to exclude, e.g., drugs and derivatives, then an exclusion list could be provided to R to pre-filter metabolites on the exclusion list from the annotations saving the user time and effort. As annotations are returned by BEAMSpy as a 'string', with annotations separated by '|' the BEAMSpy annotations, per metabolite feature, can be converted to a list before being compared to a user provided exclusion list. Non-

overlapping metabolites, i.e., not on the exclusion list, can then be identified, reduced (duplicate entries removed), and saved to overwrite the original annotation values in the BEAMSpy results. A 'for loop' can iterate this process across all metabolite features in the dataset. Additionally, as the exclusion list is user provided it can be as broad or precise as the user requires or even adapted from publicly available databases such as DrugBank. This function also solves the redundancy problem of BEAMSpy reporting the same metabolite multiple times as it is present in multiple user selected databases to search. Public databases can also assist in other aspects of streamlining annotations. A public database for lipid annotations utilised by BEAMSpy, LIPIDMAPS, provides its database in a downloadable .csv format which contains all of its metabolite annotations and their species level annotation. This .csv database can be read into R and used as a reference table to convert the 'overly precise' annotations provided by BEAMSpy into the species level annotation. Applying the same principle described above a 'for loop' can be initiated to iteratively extract and list metabolite annotations, per feature and compare them to a reference list, though exact matches in spelling/spacing are required. In instances where there is a match between an annotation and the complete lipid name in the reference list R extracts the abbreviated lipid name from the reference list and overwrites the sourced full lipid name within the BEAMSpy results. Any duplications in annotations resulting from this process are removed as described above. Any lipids listed by their common name and not on the reference list will remain unchanged in the BEAMSpy results. As lipid names typically follow a standardised format it would be possible for R to identify a lipid and construct an abbreviated name from pre-defined rules that lack an entry on reference lists. However, this solution is imperfect as it would require a complex sex of rules to identify different lipid classes and assign the correctly

formatted abbreviation. Whilst software assisted, this step still requires significant user input and judgement.

8: Figure creation

Graphical figures are not produced automatically following step 7 as each study may have different requirements and needs with regards to how data is presented, hence it is a step that requires informed decisions by the user. Rather, step 8 involved writing 'templates' which could take data inputs either from the results of step 7 or manually produced data files to then produce high quality figures with the assistance of R packages MetaboanalystR or ggplot2. The templates provide a consistency in style and appearance, allow for quick edits or reproduction of the figures and, as figures are rendered in R, can be rendered at a dpi matching the users needs e.g., 300+ for publication. As the templates are a few lines of code that are a few kilobytes in size they can easily be shared to other R users.

A major challenge in writing the above workflow was adapting to the ever-evolving practices of rhythm detection. One such development has been the emphasis on favouring methods of differential rhythmicity analysis, e.g., DODR [279] and compareRhythms [280] , and avoiding once commonly employed practices such as Venn diagram analysis [280] as discussed above (see Section 3.3). Furthermore, over the short course of this thesis project additional R packages and webapps have become available which have attempted to offer multiple rhythm analysis functionality such as DiscoRhythm [270] which incorporates cosinor [271], JTK [273], ARS [274], and Lomb-Scargle [372] methods of rhythm analysis expanding on the initial three offered by Metacycle [272]. Further, GUI webapps such as BioDare2 have released which is a workflow that facilitates rhythm analysis and figure creation, similar to the workflow presented above, and is free to use on the condition that data submitted for analysis

becomes publicly available after a specified embargo period, thus the platform doubles as a public repository for all data submitted to the platform [371]. As raised in the introduction to this chapter performing rhythm analysis manually can be a tedious practice, so it is unsurprising that independent groups in the field are working to develop software to help address this problem. The issue then becomes multiple research groups working on their own solutions to address shared problems across the community, yet, with no centrally agreed best practice or methodology. A lack of community driven guidelines also extends to rhythm analysis software, of which there are numerous packages beyond those applied in the above workflow. As mentioned both in the introduction above and in Chapter 1, different rhythm analysis algorithms have differing advantages and disadvantages. As a result, it has become a common practice to apply multiple in conjunction with one another; however, no guidelines are present within the community to discuss best practice in this regard. This would be a less pressing issue if not for the broad differences in performance of the many available rhythm detection algorithms reported in the literature [361] and also observed during data analysis as part of this thesis (data not shown). It would benefit the chronobiology community to have an open consortium dedicated to discussion relating to standardising practices, creating community guidelines, and collaborating/pooling resources towards developing favoured platforms. Not too dissimilar to how in metabolomics there are centralised community led guidelines on best practice relating to quality control and assurance led by the mQACC consortium [373]. However, this would be no small feat within the field of chronobiology as it is a broad field that encompasses many disciplines from systems level biology to animal behaviour. The differing needs of each sub discipline and competing interests to self-promote in house tools over those developed by other research groups would likely hamper progress towards more standardised working practices in the field.

As the varying performance of rhythm analysis algorithms may be a concern in the community the workflow was designed, when collating results, to present p-values separately for each algorithm applied. This allows the user to assess the performance of each algorithm for their particular dataset, further, they could choose to only extract results from their preferred rhythm analysis algorithm(s) or use the cross-referenced results as performed within this thesis, i.e., rhythmic metabolites are defined as those with significant results ($p < 0.05$) post FDR Benjamini-Hochberg correction in at least three of the six applied rhythm detection algorithms.

4.3. Chapter summary

As metabolomics grows as a field there is a greater emphasis on its role in clinical practice, medicine or real time analysis which all require a high throughput. Advances in robotics have improved the throughput of large-scale sample preparation of blood based sample for UHPLC-MS analysis in proteomics [374] whilst the advent of UHPLC saw reduced analysis times per sample [375]. When processing metabolomics data, whether part of a chronobiology dataset or not, there are numerous data handling tasks ranging from simple re-labelling to more complex calculations which may need to be re-iterated *ad nauseum*. Many of these instances can be automated by simple code written or usable by a programming novice, alleviating the time burden of processing the data both in the present and for future datasets. Investment in people and skills may help improve metabolomics throughput in the data analysis stage of a metabolomics study just as improving technology has greatly improved throughput at the data acquisition stage. Currently the workflow is operated from within R via the command line but could see further development into an online GUI app via R package 'shiny' [376, 377] to facilitate its ease of use for those unfamiliar with R. Development of additional features,

e.g., incorporating additional methods of rhythm analysis and additional graphical templates, would help maintain relevancy of the workflow as other similar platforms are developed.

CHAPTER 5

INVESTIGATING METABOLITE RHYTHMS OF NON-POLAR METABOLITES VIA A UHPLC-MS LIPIDOMICS ASSAY APPLIED TO TIME SERIES PLASMA SAMPLES COLLECTED UNDER CONSTANT ROUTINE FROM MALES AND FEMALES – A DISCOVERY BASED STUDY.

5.1. Introduction

Within the emerging field of chronobiology applying metabolomics platforms, investigations into and evidence of the daily 24 h rhythms of lipids in blood have been published frequently [65, 70, 292, 315-317, 320-323, 327, 329-331, 337]. However, the majority of these studies investigate 24 h diurnal rhythms in free living individuals and not 24 h circadian rhythms defined through constant routine studies. Human circadian rhythms can be identified via constant routine studies which is a gold-standard protocol which controls and minimises the influence of Zeitgeber. Specifically, ambient conditions such as light, temperature, posture, activity, wakefulness, and diet of participants, which otherwise may mask circadian rhythms, are kept constant throughout this type of study. How a constant routine protocol is designed and typically employed is described in detail by Duffy & Dijk [318], with further details on the process of entrainment and the nature of un-entrained (free-running) rhythms provided in [19, 378]. Human studies, such as those by [70, 317, 319, 320, 328], typically employ small but highly controlled study groups to minimise intra- and inter-subject variation and restrict confounding factors which may otherwise mask the rhythms being observed.

The circadian rhythms observed in constant routine conditions will differ in their phase (timing) and amplitude between individuals, primarily as a result of their genetics, age, and sex [379-381]. Even if samples are taken at the same social time, e.g., 08:00 h, participants will likely express variation in the phase of their individual circadian rhythm; the relationship between the timing of an individual's circadian rhythm and the timing of a Zeitgeber being referred to as the phase angle of entrainment [379-381]. Much like the hands of an actual clock revealing the current time of day there are (bio)markers that reveal the circadian clock's phase, i.e., internal (biological) time. Widely used and validated central clock phase markers

are salivary/plasma melatonin, urinary 6-sulfatoxymelatonin (aMT6s) and plasma cortisol [382-384] and which are thought to reflect the timing of the SCN clock of an individual as discussed previously (see Section 1.1.2.). Biological rhythms with a period of 24 h that do not/are not known to meet the defining principles of circadian rhythms, as outlined in Chapter 1 (section 1.1.1.), may be referred to as diurnal rhythms, i.e., 24 h rhythms observable under entrained "real-life" conditions jointly driven by exogenous factors, e.g., light/dark, feeding/fasting, rest/activity cycles, and the endogenous circadian timing system.

Prior studies investigating circadian (i.e., under constant routine) lipid rhythms are sparse in number [70, 317, 319, 320] but have already observed circadian rhythms across a broad range of lipid classes such as: fatty acyls, glycerophospholipids, glycerolipids, sphingolipids and steroid hormones (see appendix Table C1 for a full list). These studies have predominantly investigated male cohorts. Only Skene *et al* [70] had female participants in their cohort (constant routine (CR) following night shift n = 7 (6 males, 1 females), CR following day shift n = 7 (4 males, 3 females)) but did not comment on sex as the effect of shiftwork was the purpose of the study and not sex differences. Additionally, Chua *et al* and Skene *et al* [70, 320] both performed targeted assays which purposefully limits the metabolite coverage of the analysis (see Section 1.2.2.2.) and in doing so may have limited the number of detected circadian regulated lipids. Whilst Dallmann *et al* [317] and Kasukawa *et al* [319] performed untargeted assays on males only, providing greater coverage of the sample (blood/saliva) metabolome, Dallmann only sampled at a 4 h resolution across a 40 h period whilst Kasukawa had a higher sampling resolution at 2 h across a 38 h time course but had a limited cohort size (n=6). Lower sampling resolutions or fewer replicates/smaller cohorts may negatively impact rhythm detection algorithms that identify circadian metabolites [361]. Other than the steroid hormone metabolism pathway, as reported by Kasukawa [319], no other lipid metabolite

pathways or processes were reported as significantly enriched within identified circadian metabolites. However, identified circadian metabolites detected in humans were consistent with those reported in rodents, as noted by Dallmann *et al* [317]. With reference to lipid metabolism in mouse models examples include: rate limiting enzymes for cholesterol and bile acid synthesis, adipose triacylglyceride lipase and hormone sensitive lipase (lipolysis), fatty acid oxidation, lipid synthesis and hydrolysis of triglycerides all being circadian regulated [385, 386]. Metabolically active tissues such as liver, kidney, brown adipose, white adipose and skeletal muscle possessed a high percentage of circadian regulated genes relative to six other tissues in a mouse model with 16%, 13%, 8%, 4%, 4% of detected genes being circadian regulated, respectively [74]. Rhythmically expressed genes in turn rhythmically regulate various aspects of metabolism about the 24 h day. Metabolic products can act as signalling molecules which feed back to the central circadian system with challenges to metabolism, e.g., caused by an ad libitum high fat diet, perturbing circadian regulation of metabolism versus controls in a mouse model [14]. Interestingly, the negative consequences associated with a high fat diet were not observed in mice when feeding times were time restricted in a previous study [387], moreover, time restricted feeding is being investigated as a means of treatment for metabolic disease [388]. It is apparent that there is a broad and complex relationship between the circadian system and the tissue-specific/global lipid metabolism in mammals, yet this is under investigated in humans specifically.

Females are often excluded from circadian studies as the menstrual cycle, another biological rhythm, is a confounding factor. Solutions to include females within circadian studies include sample collection from naturally cycling females at the same phase of their menstrual cycle e.g., the follicular phase or recruit females on oral contraceptives [346]. However, this introduces additional variables to consider during data analysis such as differing dose of

progesterone/oestrogen in available oral contraceptives or inconclusive links between contraceptive use and altered melatonin [389-394]. As it becomes increasingly apparent that sexual dimorphism pervades the mammalian clock [347, 348], metabolome [395], metabolic disease pathophysiology [350, 351], and potentially the circadian regulation of polar metabolites as observed by Grant *et al* [328] we can no longer consider males as an adequate proxy for the human population within circadian studies. There is a need for a discovery-based study with broad (untargeted) coverage of the metabolome at a high sampling resolution (<=2 h) to extend our understanding of circadian, not diurnal, regulation of metabolism, in both male and female healthy humans. Further, we must strive to study healthy males and females in tandem with equally sized cohorts to investigate the degree of sexual dimorphism, if any, in this circadian regulation as no prior study has done this. Defining a baseline of 'normal' circadian regulation of the metabolome in humans may also be informative to alternate studies investigating the effects of shift work, circadian misalignment, or sexual dimorphism under such conditions as a point of comparison.

Within this chapter, circadian metabolic rhythms are investigated in serum samples collected from both males and females every 2 hours over 32 hours under constant routine applying ultra high performance liquid chromatography-mass spectrometry metabolomics with a C_{18} reversed phase lipidomics assay. The goals of the study were to (1) identify and report on significantly 'rhythmic' non-polar (primarily lipid) metabolites with a ~24 h period observed under a constant routine and so not a result of exogenous cues and (2) identify and report sex dependent differences, both in terms of individual rhythmic metabolites and enriched metabolite classes.

5.2. Results

5.2.1. Data processing outcomes

Data was collected for 16 males and 16 females at a 2 h resolution across a 32 h time course, utilising two different UHPLC- MS assays to provide a broad detectable coverage of non-polar (lipid) metabolites. However, repeated malfunctions related to the gas supply of the UHPLC-MS system resulted in the automated shutdown of the system and premature termination of the autosampler sequence. As a consequence, a subset of samples were not analysed in a timely manner post preparation and, despite being cooled to 4^{o}C within the autosampler unit coupled to the UHPLC system, exhibited signs of sample evaporation. Thus, the affected samples were disregarded from further analysis for both the lipidomics and HILIC assays (see Chapter 6) leaving only the samples from 13 males and 10 females; all of which were used for the individual level analysis and a subset of 10 males and all 10 females used for the group level analysis. Differences in age and BMI remained non-significant between males (age 22.7 ± 1.8 SD, BMI 25.3 ± 2.7 SD) and females (age 24.5 ± 3.7 SD, BMI 23.6 ± 2.2 SD) in this abridged cohort.

5.2.2. Group level rhythm analysis

Results from the PCA displayed two distinct, though overlapped, clusters for each sex highlighting the inherent inter-sex variation between males and females at the metabolome level (Figure 5.1) as previously reported in the literature [395]. Dimorphism within this cohort has previously been reported with regards to melatonin and cortisol in which the peak concentration of plasma melatonin/cortisol is significantly higher in females compared to males, though acrophase (peak time) did not differ [64].

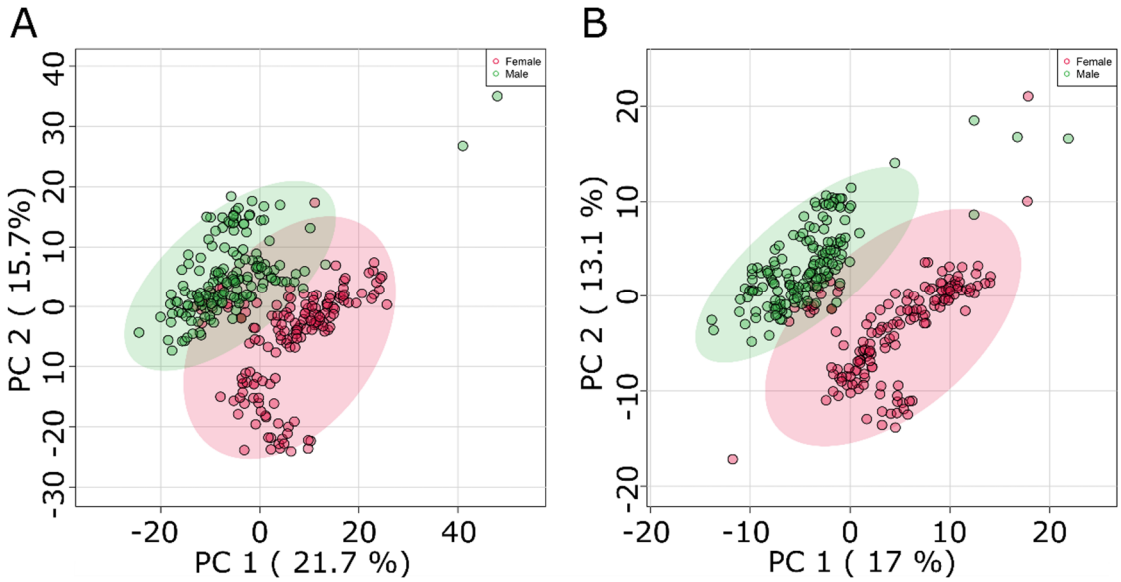


Figure 5.1: Principal components analysis in males (n = 10, green) and females (n = 10, red) for the lipidomics assay in (A) positive ion mode and (B) Negative ion mode, post quality filtering. PCA analysis reveals inter-group variation between males (n = 10) and females (n = 10) across both UHPLC-MS assays, primarily across PC1. This is indicative that inter-group variation, i.e., related to sex, is a prominent source of variation between groups. Variance appears greater within the female group compared to the male group, as indicated by the larger 95% confidence intervals, particularly along PC2.

Outcomes of filtering and details on the number of metabolite features subjected to rhythm analysis are summarised in Table 5.1. A full list of rhythmic metabolites annotated to a single metabolite class, across both assays for males and females are provided in the appendix (Table C2). A cross referencing approach was taken for the rhythm analysis such that metabolites were classified as rhythmic if they obtained a significant result ($p < 0.05$) post FDR correction in the majority of applied algorithms (>=3).

Table 5.1: Number of metabolite features (a) following XCMS raw data processing, (b) following quality filtering and (c) following group rhythm analysis.

Assay	No. of features yielded from raw data processing	No. of features post quality filtering	No. of features submitted to rhythm analysis male/female
Lipids – POS	19,747	3,334 (16.9%)	2,614 (13.2%) / 2331 (11.8%)
Lipids – NEG	10,370	1,851 (17.9%)	1,431 (13.8%) / 1301 (12.6%)
Total	30,117	5,185 (17.2%)	4,045 (13.4%) / 3632 (12.1%)

The inter-sex variation observed in the multivariate PCA scores plots was mirrored in the group level rhythm analysis. The female group presented with more rhythmic metabolites than males both in absolute terms (1073 vs 898, Figure 5.2A) as well as the percentage of rhythmic features out of those submitted for analysis (48.2% vs 37.2%, Figure 5.2A). Moreover, the female group presented with a larger number of rhythmic features classified as lipids across a range of different lipid classes (Figure 5.2C). The peak time (acrophase) of lipids differs between lipid class based on the below observations (Figure 5.2B), however, acrophase within each lipid class is similar for males and females with the greatest deviations between sexes observed within glycerolipids and sphingolipids. In some instances, these acrophase differences between rhythmic metabolite features exceeded the sampling resolution of 2 h.

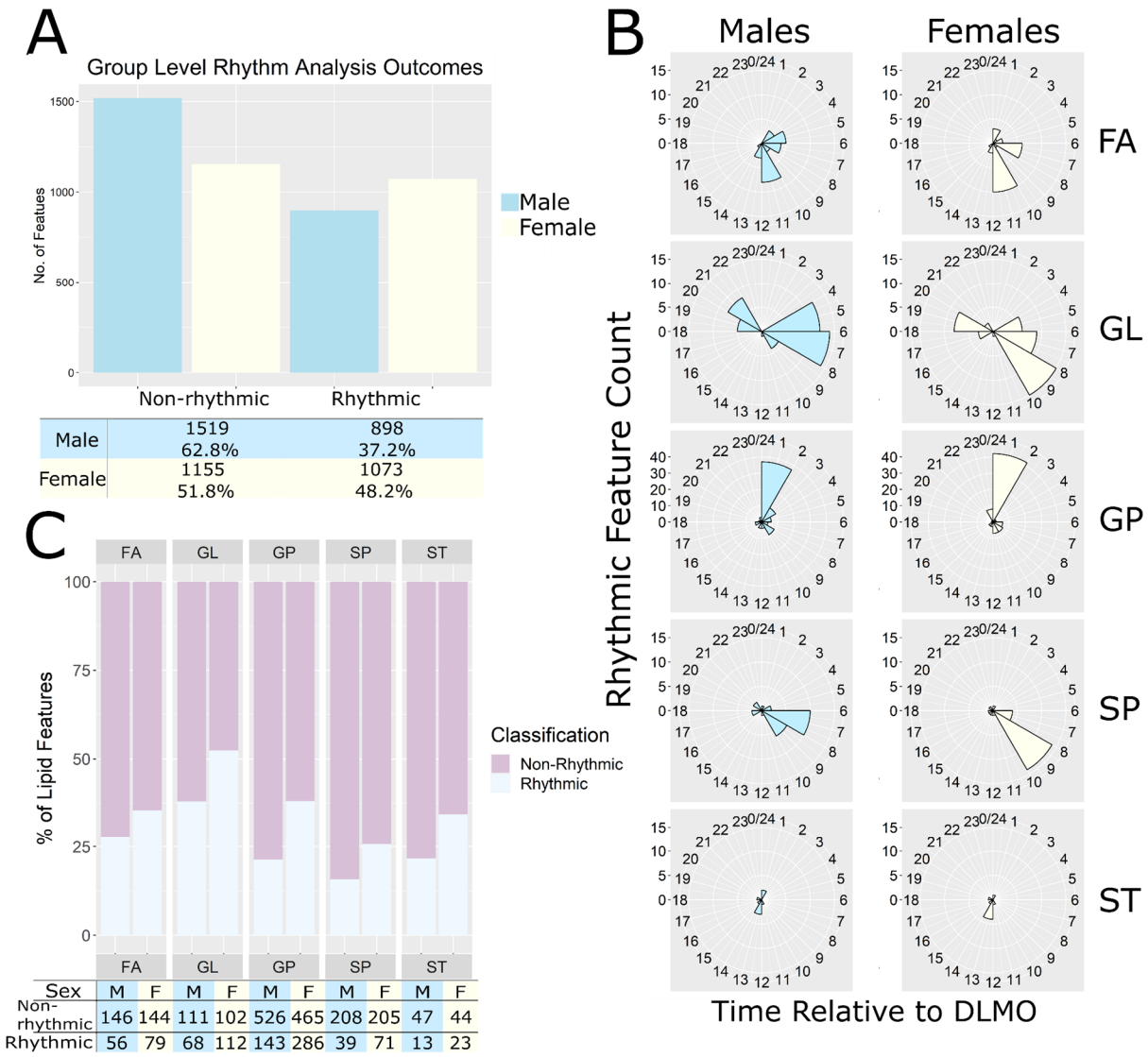


Figure 5.2: Group level rhythm analysis outcomes (A) Number of rhythmic metabolite features by sex (males: n = 10, females n =10) (B) acrophase (peak time) by sex per lipid class (C) stacked bar chart displaying % of rhythmicity in annotated (lipid) metabolites by sex per lipid class, with accompanying table displaying absolute number of metabolite features. (B & C) Only metabolite features corresponding to a single lipid class, obtained at species level (as defined here [87]), were included. Duplicate annotation entries were handled as follows: e.g., if Feature1 and Feature2 were both annotated as FA(4:0) and classified as rhythmic, this only counted as a single entry. However, if Feature1 was classified as rhythmic and Feature2 non-rhythmic it would be counted as an entry for both as it cannot be discerned if these example features correspond to the same lipid molecule or different isomers. M = Male, F = Female, FA = Fatty acyls, GL = Glycerolipids, GP = Glycerophospholipids, SP = Sphingolipids, ST = Sterol lipids.

A full list of putatively identified rhythmic metabolites observed in males and females where acrophase differed by >2 h between sex is provided in Table 5.2.

Table 5.2: A list of putatively identified metabolite features with an acrophase difference of >2 h between males and females.

Assay	Putative annotation(s)	Acrophase difference (h) (male relative to female)
Lipids - positive ion mode	PC(42:11) | PI(34:3)	-5
Lipids - negative ion mode	Cer(40:2;O2)	5
Lipids - positive ion mode	Hex2Cer(42:1;O2)	4
Lipids - negative ion mode	PS(40:9)	-3
Lipids - positive ion mode	TG(58:9)	3
Lipids - positive ion mode	LysoPS(O-18:1) | LysoPS(P-18:0)	3
Lipids - positive ion mode	Hex2Cer(44:1;O2) | Hex2Cer(42:0;O2)	-3
Lipids - positive ion mode	TG(54:5) | TG(52:4)	-3
Lipids - negative ion mode	Cer(42:0;O3)	-3
Lipids - positive ion mode	FA(18:3;O2)	3
Lipids - positive ion mode	FA(20:3;O2)	3
Lipids - negative ion mode	ST(19:2;O2;S)	3
Lipids - positive ion mode	TG(49:2)	-3
Lipids - positive ion mode	TG(51:4)	-3
Lipids - positive ion mode	TG(60:11)	2
Lipids - positive ion mode	HexCer(39:1;O2)	-2
Lipids - positive ion mode	TG(51:6)	2
Lipids - positive ion mode	TG(58:11)	2
Lipids - negative ion mode	LysoPI(18:2)	2
Lipids - positive ion mode	TG(51:2)	-2
Lipids - positive ion mode	LysoPG(22:6) | LysoPC(18:2)	-2

Metabolites are listed in order of descending magnitude, ignoring +/- . Acrophase is rounded to the nearest hour to match the resolution of the DLMO scale. Lipid annotations are given at the 'species level' as defined here [87]. Each entry, e.g., "PC(42:11) | PI(34:3)" corresponds to a single metabolite feature (*m/z*, RT pair). Cer – Ceramide, FA - Fatty acyl, PC – Glycerohosphatidylcholine, PG – Glycerophosphoglycerol, PI – Glycerophosphoinositol, PS – Glycerophosphoserine, ST - Sterol Lipid, TG - Triacylglyceride.

The majority of metabolites, for both males and females, appear to peak at or just after DLMO (DLMO = 0, +2) and once again at subjective morning (DLMO +6:+10) which corresponds to a decline in melatonin levels and increase in cortisol (Figure 5.3). These two acrophase clusters bookend the approximate start and end of subjective night-time when participants would otherwise be permitted to sleep if not for the enforced constant routine protocol.

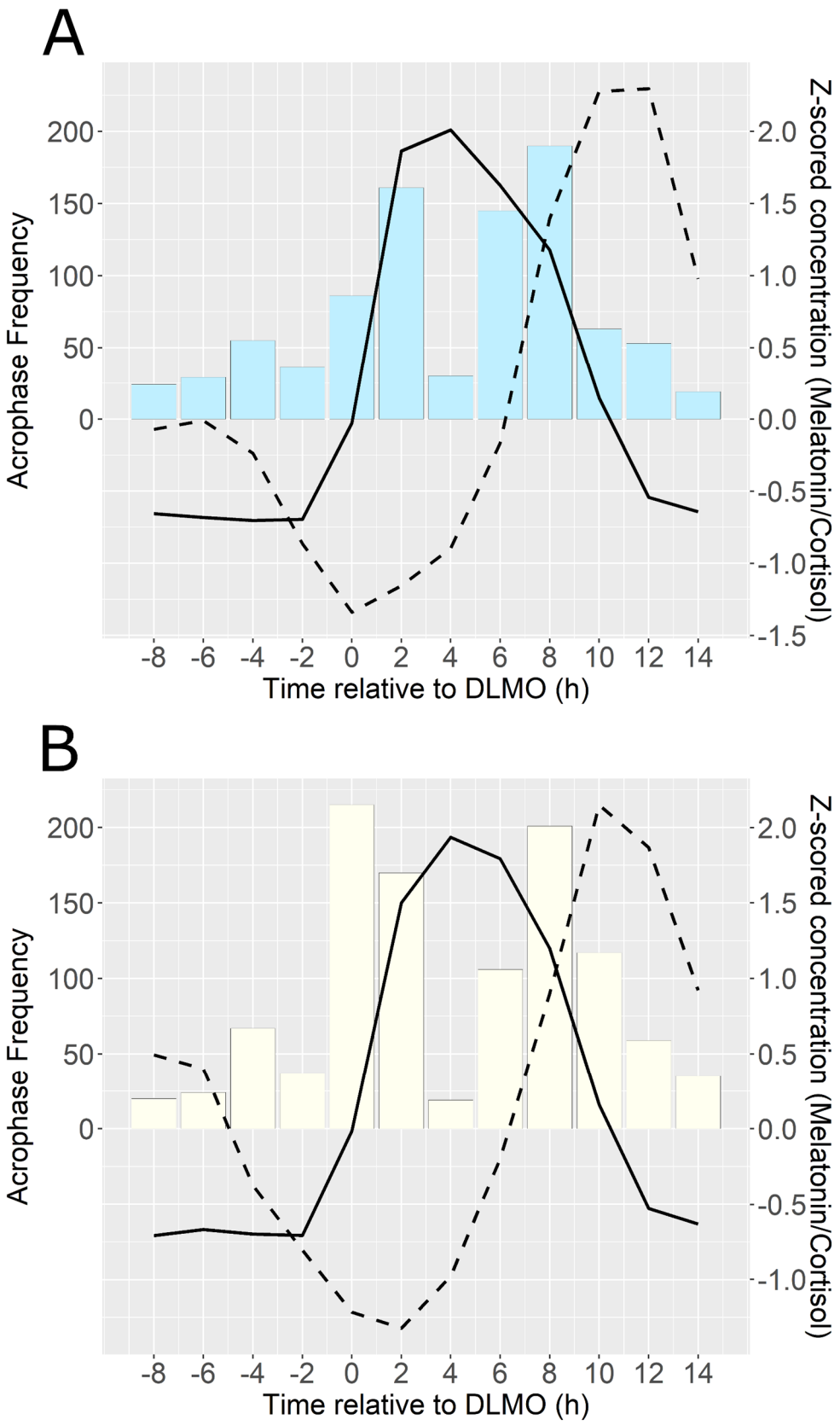


Figure 5.3: Distribution of annotated metabolite features acrophase relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) male lipidomics assay, (B) female lipidomics assay. Male cohort n = 10, female cohort n =10. For example, 204 metabolite features peaked 8 h after DLMO for the female group.

The acrophase of individual annotated rhythmic metabolites from the lipidomics assay, organised by metabolite class, are visualised below as heatmaps for males (Figure 5.4) and females (Figure 5.5).

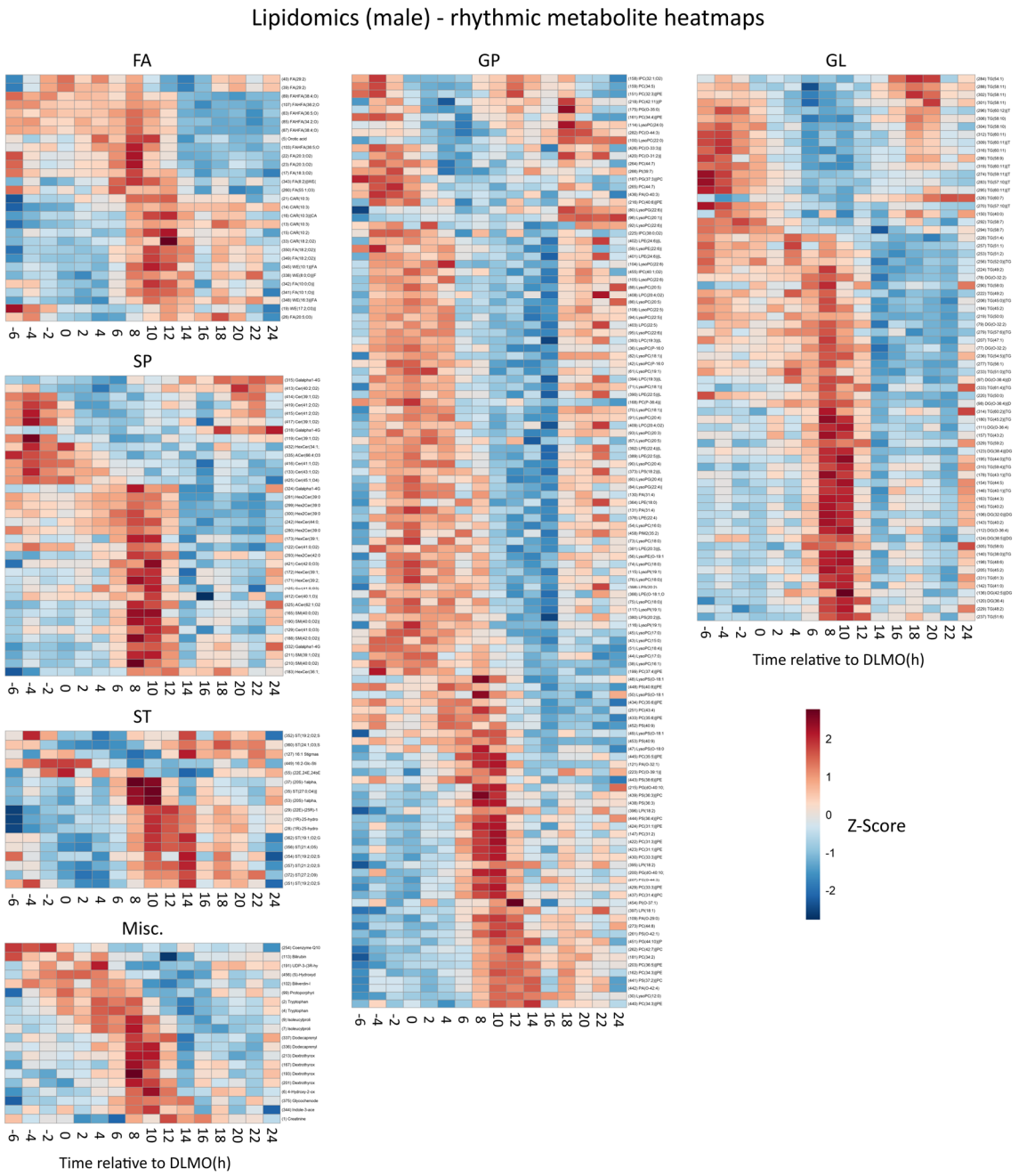


Figure 5.4: Male annotated rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged (male: n = 10) changes in relative abundance (red = high, blue = low) relative to dim light melatonin onset (DLMO, -6 : 24 h from left to right) and organised by metabolite class. FA = Fatty acyls (n = 29), GL = Glycerolipids (n = 69), GP = Glycerophospholipids (n = 121), SP = Sphingolipids (n = 36), ST = Sterol lipids (n = 17), Misc. = Miscellaneous metabolites not belonging to FA/GL/GP/SP/ST compound classes (n = 20). Note, each row corresponds to a unique metabolite feature, i.e., a unique m/z and RT; however, multiple features may have the same annotation e.g., Tryptophan. Only rhythmic metabolite features annotated to a single metabolite class are displayed. Index numbers can be cross referenced to Table C2 for additional details per metabolite feature.

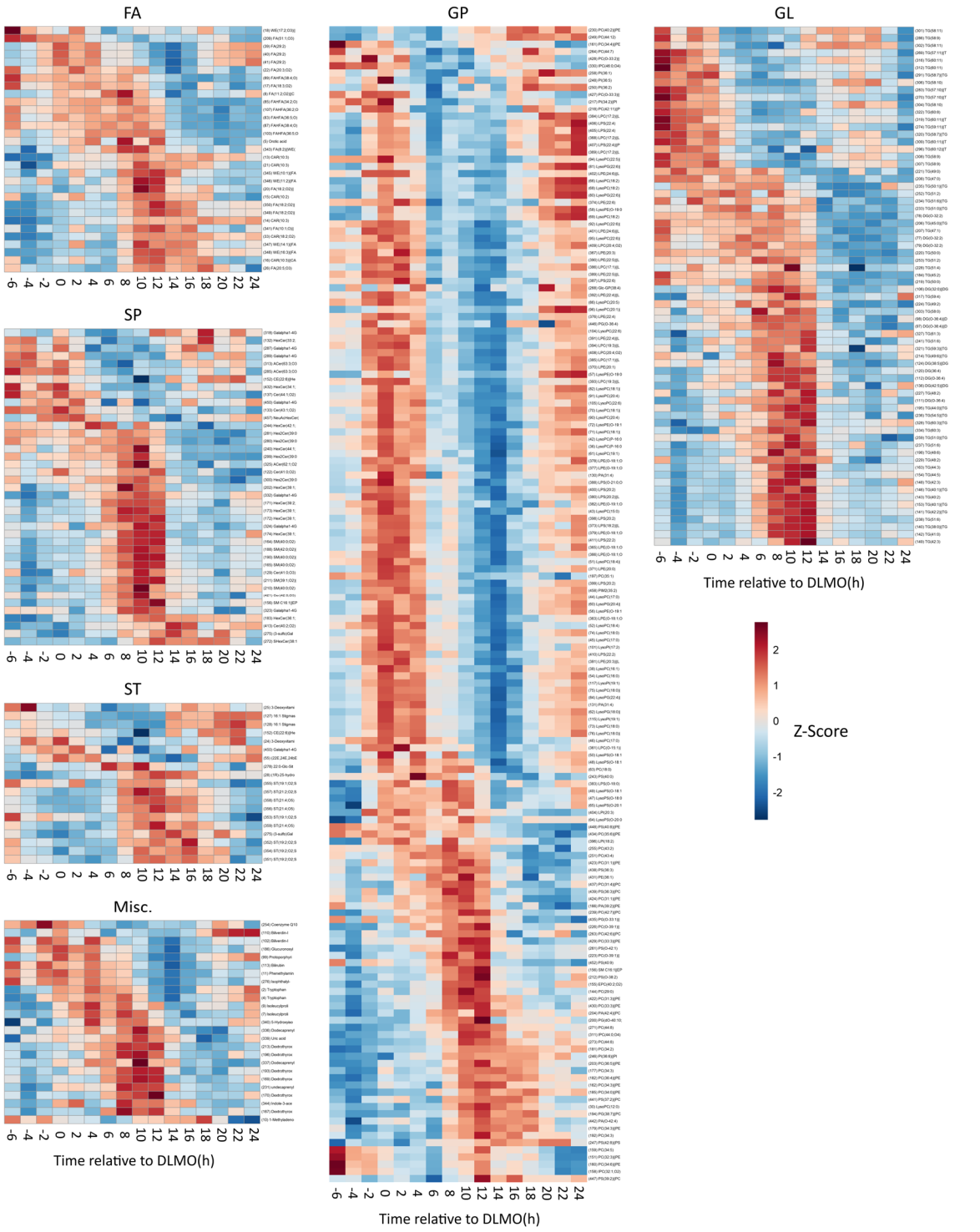


Figure 5.5: Female annotated rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged (female: n = 10) changes in relative abundance (red = high, blue = low) relative to dim light melatonin onset (DLMO, -6 : 24 h from left to right) and organised by metabolite class. FA = Fatty acyls (n = 31), GL = Glycerolipids (n = 70), GP = Glycerophospholipids (n = 161), SP = Sphingolipids (n = 41), ST = Sterol lipids (n = 19), Misc. = Miscellaneous metabolites not belonging to FA/GL/GP/SP/ST compound classes (n = 25). Note, each row corresponds to a unique metabolite feature, i.e., a unique m/z and RT; however, multiple features may have the same annotation e.g., Tryptophan. Only rhythmic metabolite features annotated to a single metabolite class are displayed.

Annotated rhythmic metabolites reported within this chapter which have previously been reported as rhythmic under constant routine are shown in Table 5.3.

Table 5.3: A list of putatively identified rhythmic (lipid) metabolites observed in both our group level analysis and the group level analysis of previous studies.

Metabolite annotation	Observed in
CAR(10:1)	[317]
CAR(12:0)	[317]
CAR(14:2)	[70]
CAR(8:0)	[317]
CE(19:0)	[317]
DG(36:1)	[320]
FA(10:0;O)	[317]
FA(16:1)	[317]
FA(17:1)	[317]
FA(18:2;O)	[317]
FA(20:2)	[317]
FA(20:3)	[317]
LysoPC(16:0)	[320]
LysoPC(16:1)	[70]
LysoPC(17:0)	[70]
LysoPC(18:0)	[70]
LysoPC(18:1)	[70]
LysoPC(18:2)	[70]
LysoPC(20:2)	[317]
LysoPC(20:3)	[70]
LysoPC(20:4)	[70]
PC(32:1)	[70]
PC(32:2)	[320]
PC(32:3)	[70]
PC(34:0)	[320]
PC(34:2)	[70, 320]
PC(34:3)	[70]
PC(34:4)	[70]
PC(36:0)	[70]
PC(36:4)	[70]
PC(36:5)	[70]
PC(36:6)	[70]
PC(38:4)	[70]
PC(40:2)	[70]
PC(42:2)	[70]
PC(O-34:3)	[70]
PE(34:1)	[317]
TG(48:2)	[320]

TG(50:2)	[320]
TG(51:2)	[320]
TG(52:0)	[320]
TG(52:1)	[320]
TG(52:4)	[320]
TG(52:5)	[320]
TG(54:1)	[320]
TG(54:3)	[320]
TG(54:4)	[320]
TG(54:5)	[320]

Metabolite IDs, as listed, may correspond to the same rhythmic metabolite features e.g., TG(52:4) | TG(54:5) (*m/z* ~ 882 RT ~585) or multiple rhythmic metabolite features e.g., TG(54:5) (*m/z* ~919 RT ~586). Lipid annotations are given at the 'species level' as defined here [87]. CAR – Acylcarnitine, CE – Cholesterol ester, DG – Diacylglyceride, FA – Fatty acid, PC – Glycerohosphatidylcholine, PE – Glycerophosphoethanolamines, TG – Triacylglyceride.

A total of 1384 non-polar metabolite features were identified as rhythmic in males and/or females by the independent group rhythm analysis tests. Of these 1384 metabolite features, differential rhythmicity analysis determined 51 metabolite features which were considered to have a significant difference in acrophase between sex ('change'), 44 were considered significantly rhythmic in females only ('gain') whilst 12 were considered significantly rhythmic in males only ('loss'). The remaining metabolite features were classified as not being significantly different between males and females ('same' - 1277 features). In brief, not all the observed differences in rhythmic metabolite features between males and females, as determined from the initial group level rhythm analysis, are distinct enough to be considered statistically significant between males and females ($p < 0.05$) but over 100 metabolites features are annotated to a single metabolite class (Table 5.4).

Table 5.4: A list of putatively identified rhythmic (lipid) metabolites also identified as differentially rhythmic between males and females.

Assay	Classification	Annotations (single class)
Lipids - negative ion mode	'change' – difference in acrophase/shape	ST(19:2;O2;S)
Lipids - negative ion mode	'change' – difference in acrophase/shape	ST(21:2;O2;S)
Lipids - negative ion mode	'change' – difference in acrophase/shape	Cer(40:2;O2)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	FA(20:5;O3)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	FA(29:2)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	FA(29:2)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	LysoPS(O-18:0) | PC(16:0) | LysoPS(O-18:1)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	(22E,24E,24bE)-1alpha,25-dihydroxy-26,27-dimethyl-22,23,24,24a,24b,24c-hexadehydro-24a,24b,24c-trihomovitamin D3
Lipids - Positive ion mode	'change' – difference in acrophase/shape	LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPC(18:0) | PC(O-18:0) | LysoPE(19:0) | LysoPE(21:0)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	LysoPC(18:1) | PC(O-18:1)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	LysoPC(18:0)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	LysoPI(19:1) | PA(27:2;O)
Lipids - Positive ion mode	'change' – difference in acrophase/shape	DG(36:4)
Lipids - negative ion mode	'gain' - Rhythmic in females	LysoPE(O-18:1;O) | LysoPC(15:0) | LysoPE(18:0) | LysoPS(O-19:0;O) | LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPE(19:0)
Lipids - negative ion mode	'gain' - Rhythmic in females	LysoPC(17:1) | LysoPE(20:1) | LysoPC(18:1) | PC(O-18:1) | LysoPS(21:0)
Lipids - negative ion mode	'gain' - Rhythmic in females	LysoPS(O-21:0;O) | LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0) | LysoPC(17:0) | LysoPE(20:0)
Lipids - negative ion mode	'gain' - Rhythmic in females	LysoPS(20:2)
Lipids - negative ion mode	'gain' - Rhythmic in females	NeuAcHexCer(34:1;O2)
Lipids - Positive ion mode	'gain' - Rhythmic in females	LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPC(18:0) | PC(O-18:0) | LysoPE(19:0) | LysoPE(21:0)

Lipids - Positive ion mode	'gain' - Rhythmic in females	SM(C16:1) | CerPE (38:1;O2) | CerPE (38:2;O2) | CerPE(40:2;O2) | SM(35:1;O2)
Lipids - Positive ion mode	'gain' - Rhythmic in females	PA(42:4) | PC(36:4) | PE(39:4)
Lipids - Positive ion mode	'gain' - Rhythmic in females	PS(O-38:2)
Lipids - Positive ion mode	'gain' - Rhythmic in females	undecaprenyl phosphate
Lipids - Positive ion mode	'gain' - Rhythmic in females	TG(51:6)
Lipids - Positive ion mode	'gain' - Rhythmic in females	PS(42:8) | PS(41:4) | PG(O-42:5)
Lipids - negative ion mode	'loss' - Rhythmic in males	Cer(39:1;O2) | Cer(41:2;O2)
Lipids - negative ion mode	'loss' - Rhythmic in males	Cer(39:1;O2) | Cer(41:2;O2)
Lipids - Positive ion mode	'loss' - Rhythmic in males	TG(58:7)
Lipids - Positive ion mode	'loss' - Rhythmic in males	TG(54:0) | TG(58:8) | TG(60:11) | TG(60:12)

Lipid annotations are given at the 'species level' as defined here [87]. Each entry, e.g., "PC(42:11) | PI(34:3)" corresponds to a single metabolite feature (*m/z*, RT pair). Cer – Ceramide, CerPE – Ceramide phosphoethanolamines, DG – Diacylglyceride, FA - Fatty acyl, NeuAcHexCer – N-Acetylneuraminic acid hexosylceramide, PA - Phosphatidic acid, PC – Glycerohosphatidylcholine, PE – Glycerophosphoethanolamines, PI – Glycerophosphoinositol, PS – Glycerophosphoserine, SM – Sphingomyelin, ST - Sterol Lipid, TG – Triacylglyceride.

5.2.3. Individual level analysis

Individual level rhythm analysis revealed a high degree of inter-individual variation within and between sexes with the number of rhythmic metabolite features ranging from 0-478 across the cohort (n=23) (Figure 5.6A). It is also apparent that, as a generalisation, most females presented with more rhythmic features than most males to the point of statistical significance ($p < 0.05$, one-tailed T-test) (Figure 5.6B). When comparing the number of rhythmic metabolites features, observed per participant, to various aspects of available metadata the number of rhythmic features did not correlate with age, BMI, melatonin peak/amplitude/area under the curve (AUC), order of sample collection, progesterone/oestrogen dose (relating to oral contraceptive, females only) but shared a moderate and statistically significant

correlation with cortisol peak/AUC ($p < 0.05$, Spearman's Rho 0.43 - 0.55). This trend is more pronounced in females compared to males though the correlation is not significant when either sex is viewed in isolation. Within the full cohort ($n = 23$) there was a statistically significant moderate correlation ($p < 0.05$, Spearman's Rho 0.47-0.51) between peak melatonin and cortisol levels, and melatonin and cortisol amplitude. When males ($n = 13$) are observed in isolation the latter trend is near significant ($p = 0.0518$) whilst the former trend is significant ($p < 0.05$) while neither trend is significant for females ($n = 10$) in isolation.

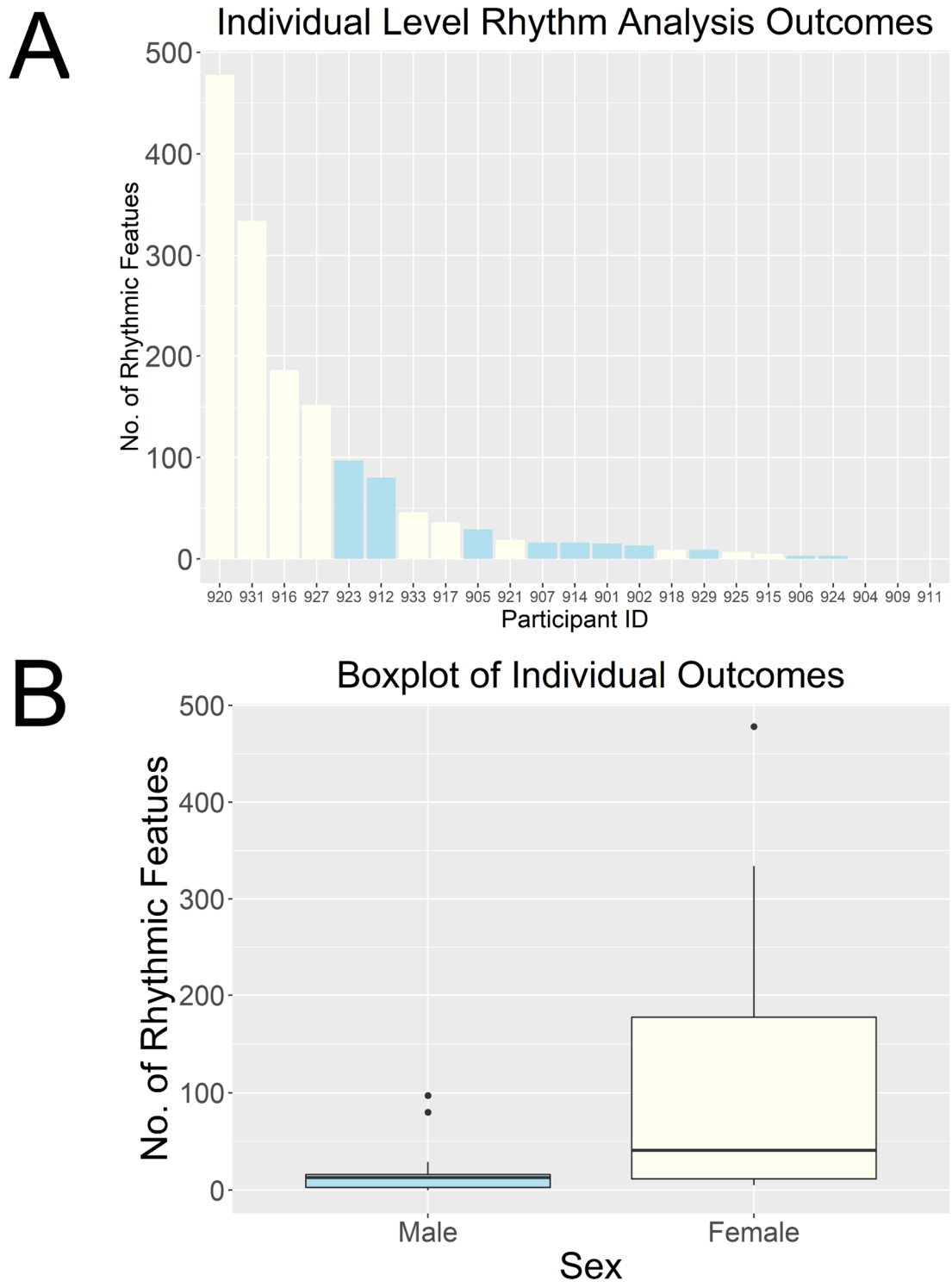


Figure 5.6: (A) – bar chart displaying number of rhythmic metabolites per individual determined by individual level rhythm analysis, (B) – Box and whisker plot of individual level analysis outcomes. Male cohort n = 13, female cohort n =10. A metabolite was classified as rhythmic for a ~24 h period if it achieved a significant result post FDR correction ($p < 0.05$) in at least 3 of the 5 applied rhythmic detection algorithms for individual level analysis: Cosinor, JTK, ARS, RAIN, BioCycle. Group means are significantly different ($p < 0.05$), one tailed T-test. Boxes are displayed with the default ggplot2 settings: The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles), The whiskers extend from the hinge to the largest/smallest value no further than 1.5 * interquartile range.

5.3. Discussion

The first objective of this study was to identify and report a list of circadian regulated non-polar (lipid) metabolites (~24h period) observed under constant routine study and therefore not influenced by external cues (i.e., diurnal rhythms). Across two lipidomics assays (C_{18} reversed phase UHPLC in positive and negative ion modes) a total of 5,185 metabolite features were reported after data quality filtering. For males and females respectively, 898 and 1073 rhythmic metabolite features (~24 h period) were reported and of these 292 and 343 rhythmic metabolites were annotated to a single metabolite class, respectively. All other annotated metabolites did not exhibit a ~24 h period. Whilst the number of rhythmic metabolites annotated may appear far less than the total number of metabolites detected across both UHPLC-MS assays (Appendix Table C2) this compiled list remains the largest and most comprehensive overview of the metabolome compared to any previously reported constant routine study in humans (as reviewed in Hancox *et al,* 2021 [78]) and contains metabolites not previously reported to be circadian. In addition, a further 606 and 730 unannotated rhythmic metabolite features (*m/z*- retention time pairs) with a ~24 h period in males and females respectively were identified. As methods of metabolite annotation improve and metabolite databases grow these features may be annotated in the future expanding on the already sizeable list of rhythmic metabolites. When compared to the five aforementioned circadian studies [70, 317, 319, 320, 323, 328] there are 48 potential matches between rhythmic metabolites observed in our data and the reported studies (Table 5.3). However, this is an approximation as matching annotations between datasets can be difficult e.g., FA(20:2), can represent multiple structural isomers and in the case of Skene *et al* [70] previous datasets include results from two constant routines that proceeded either simulated day shift or night shift work. The corroborating 48 rhythmic metabolites mostly comprise of

fatty acyls, triacylglycerides and (lyso)phosphatidylcholines and were predominantly observed by Chua *et al* [320] or Skene [70]. The former, Chua, performed a targeted lipidomics assay for glycerophospholipids (GP) as follows: phosphatidic acids (PA), phosphatidylcholines (PC), phosphatidylethanolamines (PE), phosphatidylglycerols (PG), phosphatidylinositols (PI), phosphatidylserines (PS) in addition to sphingolipids (SP) as follows: sphingomyelins (SM), ceramides (Cer), glucocyl-ceramides (GluCer), ganglioside mannosides 3 (GM3), and finally glycerolipids (GL) inclusive of diacylglycerides (DG) and triacylglycerides (TG). The latter, Skene, applied the commercially available AbsoluteIDQ® p180 kit which primarily targets (lyso)PCs, acyl carnitines, amino acids and biogenic amines. Grant *et al* applied [328] applied a HILIC assay and so little overlap was anticipated. Dallmann *et al* [317] and Kasukawa *et al* [319] both performed untargeted reversed phase assays with only the former identifying rhythmic lipids. Despite the broad coverage of the blood lipidome offered collectively from the above assays, inclusive of fatty acyls, GPs, SPs, and GLs, numerous rhythmic lipids were observed in the untargeted assay presented within this chapter not observed in these prior works. Rhythmic lipids observed in this chapters results but not prior studies include glycerophospholipids other than (lyso)phosphatidylcholines and the exception of PE(34:1) observed by Dallmann, DGs with the exception of DG(36:1) the only rhythmic DG observed by Chua, and sphingolipids. Examples of rhythmic sterol lipids and fatty acids were also observed in this chapters results but neither lipid class was not included in the targeted panels of prior studies whilst only the latter had examples identified in prior untargeted studies. Chua's assay assessed a broad range of lipid metabolites, with regards to lipid class, but circadian regulated metabolites were limited to PCs, TGs and a single DG as mentioned above. This is in contrast to the rhythm analysis presented within this chapter where rhythmic annotated metabolite

features were observed across a much larger range of metabolite classes (Figure 5.2) inclusive of glycerophospholipids other than (lyso)phosphatidylcholines, DGs, and sphingolipids.

The source of this discrepancy between the data presented in this chapter and that presented by Chua [320], who employed the most comprehensive lipidomic assay of prior work, may be the rhythm analysis methodology employed. To elaborate, Chua employed JTK, a rhythm detection algorithm known for its conservative nature in identifying rhythms which greatly limits the number of false positive findings but equally can result in an increase in false negatives [396]. Thus, there is high confidence the rhythmic metabolites Chua observed are truly rhythmic under constant routine. However, other rhythmic metabolites potentially belonging to other lipid classes were possibly not identified by JTK due to its conservative nature. The assay applied by Skene *et al* [70] assay had reduced coverage of lipid classes, relative to both Chua's targeted assay, but applied a less conservative method of rhythm detection than JTK in cosinor analysis. This resulted in Skene identifying 65 rhythmic metabolites over 2 constant routines (pre/post simulated shiftwork) but results relevant to lipids were limited to a select group of acylcarnitines, PCs, and SMs. This limits comparisons to the untargeted lipidomics assay presented within this thesis which provided broader coverage of metabolite classes. Whilst Dallmann *et* al [317] and Kasukawa *et al* [319] both performed untargeted assays, the design of these assays differed greatly to my own by employing different platforms, with Dallmann applying both LC and GC assays coupled to mass spectrometry and Kasukawa applying a single reversed phase LC assay on a ToF-MS. Further, Dallmann applied JTK to identify rhythmic metabolites whilst Kasukawa applied a bespoke analysis akin to cosinor analysis. Whilst the experimental designs of prior works and the untargeted assays presented are highly similar, the latter acts as a step forward by providing broader coverage of the blood lipidome than prior untargeted/targeted assays and

employed numerous rhythm detection methodologies to circumvent issues associated with any one method of analysis, e.g., the overly conservative nature of JTK.

The second objective was to identify sex dependent differences within the cohort at both an individual level and a group level, with regards to rhythmic metabolites. The most surprising result to come out of the analysis was the extent of inter-individual variation across the cohort ranging from hundreds of rhythmic metabolite features in some individuals to fewer than ten in others. Due to the limitation of a single sample per time point individual level rhythm analysis is more susceptible to extreme/missing values and both biological and analytical variation compared to group level analysis. Furthermore, the statistical power of rhythm detection algorithms is reduced when analysing a single individual compared to a larger population. Taking this into consideration it is likely that the individuals who had fewer rhythmic metabolites in the individual analysis presented above are not completely devoid of ~24 circadian rhythms but highlight challenges in detecting rhythms which are neither as robust nor easy to identify as the melatonin circadian rhythm, for example. Whilst it is possible the extent of inter-individual variation in the presented analysis is exaggerated, it is still present and this is not the first instance of such a finding. The aforementioned study by Chua *et al* [320] performed a constant routine protocol on 20 males, with comparable age and BMI to the cohort presented this chapter, with plasma samples taken at 4 h intervals across a 32 h time course. The study performed targeted metabolite analysis and observed inter-individual variation in the rhythmicity of GLs, GPs and SPs (range 13-86 rhythmic lipids across individuals) and acrophase differences up to 12 h. The inter-individual variation observed by Chua is not dissimilar in scale to that observed in the male subjects (1-138) presented in this chapter though more differences are observed in females (5-480).

In addition to this individual level variation was the group level variation in the number of rhythmic metabolite features (898 in males, 1073 in females) and elevated melatonin/cortisol amplitude in females compared to males previously reported [64]. Whilst the observed difference in melatonin/cortisol amplitude could be due to innate sexual dimorphism it is also possible that the source of this dimorphism is oral contraceptive usage, a variable exclusively applicable to females. All females in the cohort presented within Chapter 5/6 were using oral contraceptives. There is some evidence that suggests taking oral contraceptive leads to an increase in melatonin or cortisol [389-392]. However, oral contraceptive use is unlikely the only cause of elevated melatonin, as a significantly higher melatonin amplitude has been observed in females not on birth control relative to males [393] in addition to cohorts containing a mix of females both on and off hormonal contraceptives [394]. Therefore, it is expected this observation of increased rhythmic lipids in females is not solely caused by use of oral contraceptives and is an important observation discussed further below.

A detailed and comprehensive review by Bailey and Silver [347] discusses dimorphism of the circadian system, such as how key components of the central circadian system possess a dimorphic complement of oestrogen/androgen receptors which in turn have ligands which are also expressed in a dimorphic manner. It has been suggested that these dimorphisms shape the differential response observed in behavioural rhythms, such as activity levels, with females once again presenting higher amplitudes [348]. This line of enquiry fits into prior observations of higher amplitudes in melatonin and cortisol within our cohorts females compared to males [64], but there is insufficient evidence from our study to confirm whether these dimorphisms in the clock are the source of the increased amplitudes or rhythms observed in females. It is interesting to note however, that when specific behavioural rhythms are perturbed, such as during sleep deprivation, the metabolome is differentially impacted

between males and females with different metabolites gaining/losing rhythmicity [65]. By design the constant routine protocol employed places participants in a state of prolonged wakefulness and so male/female differences may in part relate to sleep, or lack thereof under constant routine. Furthermore, lipid metabolism appears structured around subjective night-time for our participants, when participants would normally sleep if not for the constant routine. Notably, GPs peaked at or just after DLMO and GLs/SPs 8 h later around subjective morning. Cell division is a process regulated by the circadian system with a key time gate being the transition between the G_2/M phases which is blocked by BMAL1/CLOCK activity, key components of the circadian clock that accumulate throughout the day but are repressed at night [397, 398]. GPs are major components for cell membranes and the peak observed at subjective night-time may, by design, coincide with 'peak demand' of these metabolites for cell growth and division. The subjective morning peak in GLs, the predominant form of storage for fatty acids, is proceeded by a peak in fatty acyls. As our participants did not consume a large meal at this time these peaks are not due to food intake but likely lipolysis with these GLs released from adipose tissue to yield fatty acids required to produce other lipids and/or energy production. Adipose triacylglycerol lipase (*ATGL*) and hormone sensitive lipase (*HSL*) genes produce enzymes that hydrolyse GLs to yield fatty acids and are activated directly by the BMAL1:CLOCK transcriptional complex, products of core clocks genes which accumulate during the day time [385]. Furthermore, melatonin plays a regulatory role of adipose tissue metabolism mediating processes such as lipolysis [399]. As a circadian process it could be assumed that lipolysis is at its lowest rate ~12 h after its peak at subjective morning which we might expect corresponds to the peak time of lipogenesis, the reformation of GLs which requires fatty acids. Furthermore, a second, albeit smaller cluster of GLs peaks at 19-20 h after DLMO, nearly 12 h after the morning peaks. What is most likely observed here is the

triglyceride/free fatty acid cycle (TG/FFA), a metabolic process that cycles between lipolysis and lipogenesis of TGs regulating the pool of free fatty acids in circulation, which in turn can act as signalling molecules for other metabolic processes [400]. Elevated levels of free fatty acids or perturbations to the TG/FFA cycle are associated with type 2 diabetes, obesity, metabolic syndrome, and contributes to cancer risk [401, 402], such conditions are also associated with circadian misalignment as discussed in this thesis (see Chapter 1/3). Time restricted feeding is considered an intervention that promotes more robust circadian rhythms, facilities weight loss and thus relives the symptoms of metabolic syndrome [388]. The TG/FFA cycle may be one such pathway assisted by time restricted feeding intervention or perturbed by circadian misalignment which may then contribute to the treatment/onset of diseases such as metabolic syndrome. If the TG/FFA cycle is relevant in the manner described above it is important to consider how males and females may differentially respond to time restricted feeding intervention or be differentially affected by circadian misalignment with regards to this cycle, particularly as melatonin plays a regulatory role which exhibits a dimorphic profile and thus may contribute to the differing circadian 'lipidomes' or phenotypes observed between healthy males and females. Our differential rhythmicity analysis confirmed at least 107 metabolite features exhibit a significantly ($p < 0.05$) different rhythmic profile between the sexes within the healthy young cohort investigated though it remains unclear how the sexually dimorphic melatonin/cortisol amplitudes may contribute to this observation.

Despite no definitive conclusion being provided with regards to the impact of the dimorphic profile or cortisol it should be considered highly relevant to lipid rhythms. Cushing's syndrome is a chronic condition defined by hypercortisolism, i.e., prolonged elevated levels of cortisol within the blood [403, 404]. Onset of Cushing's syndrome can mimic metabolic issues typically

associated with metabolic syndrome such as altered lipid metabolism, which can result in dyslipidaemia, as well as altered carbohydrate metabolism and insulin resistance [405, 406]. One such perturbation to lipid metabolism is increased adipocyte lipolysis, resulting in increased free fatty acids in circulation. The transcription of regulatory enzymes of lipolysis, such as the aforementioned *ATGL* and *HSL*, are increased by activation of adipocytes glucocorticoid receptors [407]. Whilst Cushing syndrome represents an extreme with regards to elevated glucocorticoids such as cortisol it is worth noting that females in our cohort, who had higher cortisol amplitude/peak (group level), also exhibited a greater number of circulating (rhythmic) lipid metabolite features inclusive of fatty acids and GLs suspected to be related to circadian regulation of lipolysis. As discussed previously the peak time of these fatty acid and GL rhythms predominantly corresponds to the (subjective) morning peak of cortisol. Unlike cases of Cushing's syndrome our cohort do not match the characteristic cortisol profile expected in the condition as the cohort's cortisol rhythm(s) still present a well-defined nadir in both sexes, loss of which can be diagnostic of Cushing's syndrome [404]. The pathology of Cushing's syndrome is strong evidence that cortisol, and potentially alterations to its rhythm, are strongly linked to changes in lipid metabolism, thus, dimorphism in the cortisol profiles of (healthy) males and females may therefore partially explain the apparent difference in lipid rhythms between the two groups of our cohort.

5.4. Chapter summary

Within, this chapter is the largest compiled list of circadian regulated non-polar (lipid) metabolites to date for a single study including both males and females, the majority of which have not been previously reported on and have underpinned the following findings. Chua *et al* [320] hypothesised in their closing statement that individuals may possess different

circadian metabolic phenotypes, due to the observed differences in circadian lipid regulation, these phenotypes having differing susceptibility to the effects of circadian disruption and associated onset of disease. Inter-individual differences across the cohort for circadian regulation of non-polar metabolites are observable in the data presented in this thesis and which has not previously been reported at the metabolic level at the time of writing. This concept of susceptible variability to circadian misalignment may partially explain why there is so much conflicting data on whether females have worse health outcomes due to shiftwork than their male counterparts. It validates the importance of studies, such as this one, that explore what is 'normal' circadian regulation in healthy individuals across different phenotypes and re-affirms the importance of individual level rhythm analysis, in tandem with group level rhythm analysis. We also report on novel observations that at the group level females in our cohort present with more rhythmic lipids metabolites than males building on prior observations that these females also have significantly higher melatonin and cortisol levels, both key circadian regulated hormones. Further, we conclude that lipid metabolism, particularly the TG/FFA is of great interest with regards to circadian regulation of metabolism due to the latter being directly regulated by melatonin and regulating other metabolic processes and cortisol's role in promoting lipolysis. Whilst it was already established that metabolism differs between males and females our data extends this observation to the circadian system and the timing of metabolic processes such as the differing peak times of GLs and fatty acids we believe associated with the TG/FFA cycle. A circadian regulated process that if perturbed has similar consequences to those associated with circadian misalignment.

CHAPTER 6

INVESTIGATING METABOLITE RHYTHMS OF WATER-SOLUBLE METABOLITES VIA A UHPLC-MS HILIC ASSAY APPLIED TO TIME SERIES PLASMA SAMPLES COLLECTED UNDER CONSTANT ROUTINE FROM MALES AND FEMALES – A DISCOVERY BASED STUDY.

6.1. Introduction

Much of the introduction from Chapter 5 (see Section 5.1) is applicable here with regards to identifying rhythmic metabolites, be they lipids (see Chapter 5) or water-soluble metabolites (in this chapter), and the lack of inclusion of females participants within the field. However, investigations into and evidence of the daily 24 h rhythms of water-soluble metabolites in blood have not featured as prominently as lipids with regards to human metabolite rhythms. Only two prior studies in this field have applied HILIC assays [324, 328], tailored to analyse water-soluble metabolite (see Section 1.3.1.2.), excluding the work presented (Chapter 5), with only the latter investigating circadian rhythms. Additional studies have observed more limited coverage of water-soluble metabolites, most frequently amino acids via NMR [315, 326], commercially available targeted assays (AbsoluteIDQ p180 kit by Biocrates) [65, 70, 321, 325] or reversed phase assays which typically favours retention of non-polar metabolites (see Section 1.3.1.3) [316, 317, 329-331]. Grant *et al* [328] is therefore the only published study to date which applies a metabolomics assay with the intent to observe circadian water-soluble metabolite rhythms and did so via both a targeted and untargeted approach with samples collected from a majority male cohort under constant routine. Grant reports on four metabolite features rhythmic in three of the four women in their cohort (n=13) concluding that sex based expression of metabolites are present in plasma but further research with larger sex balanced cohorts are required to further investigate circadian regulation of plasma metabolites. Due to this limitation pathway enrichment analysis was not performed separately for males and females to determine enriched metabolic pathways, with regards to circadian metabolites. However, phenylalanine and tyrosine metabolism was significantly enriched ($p < 0.05$, post FDR) for circadian metabolites for the group averaged analysis (full cohort, n =13) whilst the citric acid and urea cycles were significantly enriched ($p < 0.05$, post

FDR) for circadian metabolites as determined by the individual level analysis (circadian metabolites determined per individual). This lack of knowledge with regards to enriched metabolic pathways for water-soluble circadian metabolites is not just limited to humans but mammals as a whole. There is seemingly no literature available for a study which performs metabolomics specifically to measure circadian water-soluble metabolites and perform pathway enrichment analysis to monitor sex dependent differences. Whilst lipids perform a myriad of biological functions that make them an interesting point of research in terms of metabolomics, chronobiology and health, this should not negate the equally important role of water-soluble metabolites. Central carbon metabolism, glycolysis, citric acid (TCA cycle) are metabolic processes that form the foundations of metabolism within an organism and are comprised primarily of water-soluble metabolites [408]. If we wish to better understand the circadian regulation of metabolism in mammals/humans, investigating circadian polar metabolites is not just advised but rather it is necessary as failing to do so would be choosing to ignore core aspects of mammalian metabolism.

Within this chapter, circadian metabolic rhythms are investigated in serum samples collected from both males and females every 2 hours over 32 hours under constant routine applying ultra high performance liquid chromatography-mass spectrometry metabolomics with a HILIC assay. The goals of the study were to (1) identify and report on significantly 'rhythmic' water-soluble metabolites with a ~24 h period observed under a constant routine and so not a result of exogenous cues and (2) identify and report sex dependent differences, both in terms of individual rhythmic water-soluble metabolites and enriched metabolite classes.

6.2. Results

6.2.1. Data processing outcomes

As with the lipidomics dataset, HILIC data was collected for 16 males and 16 females at a 2 h resolution across a 32 h time course, utilising two different UHPLC- MS assays to provide a broad detectable coverage of water-soluble metabolites. However, as detailed in section 5.2.1, only 10 males and 10 females are reported on for the group level rhythm analysis and 13 males and 10 females for the individual level analysis. Differences in age and BMI remained non-significant between males (age 22.7 ± 1.8 SD, BMI 25.3 ± 2.7 SD) and females (age 24.5 ± 3.7 SD, BMI 23.6 ± 2.2 SD) in this abridged cohort.

6.2.2. Group level rhythm analysis

Similar to the observations made in Chapter 5, results from the PCA displayed two distinct albeit overlapped clusters for each sex highlighting inherent inter-sex variation between males and females at the metabolome level (Figure 6.1) as previously reported for water-soluble metabolites in the literature [395] and melatonin/cortisol within this cohort [64].

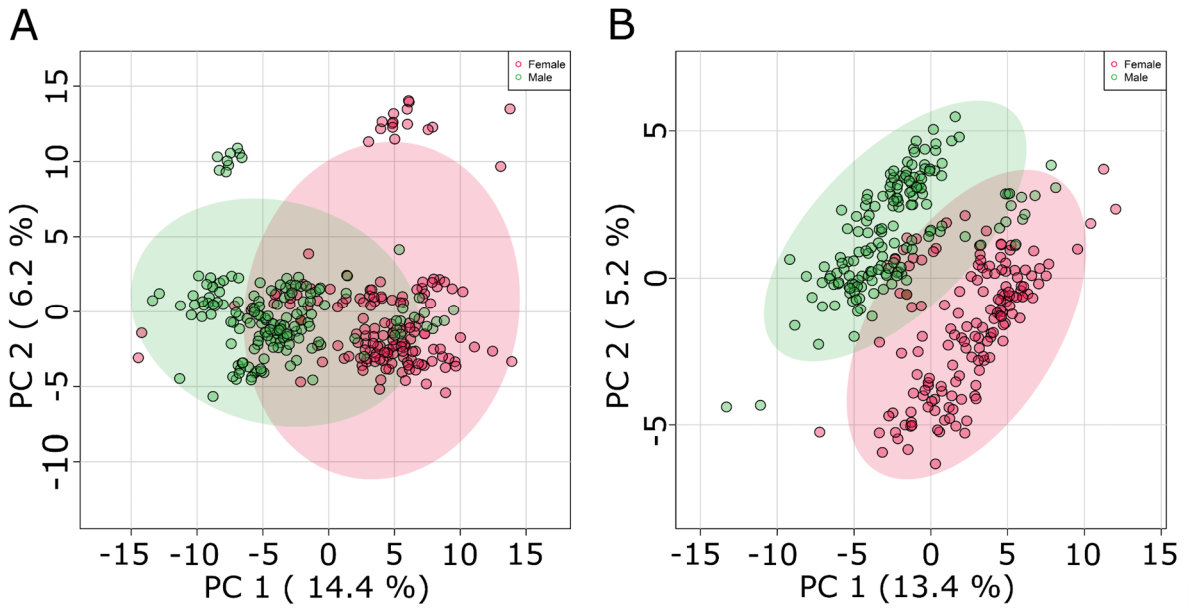


Figure 6.1: Principal components analysis of metabolites in males (n = 10, green) and females (n = 10, red) for the HILIC assay in (A) positive ion mode and (B) Negative ion mode, post QC filtering. PCA analysis reveals inter-group variation between males (n = 10) and females (n = 10) across both UHPLC-MS assays. Inter-group variation is evident, primarily along PC1 which displays separation of group clusters, more so within the negative ion mode assay.

Outcomes of filtering and details on the number of metabolite features subjected to rhythm analysis are summarised in Table 6.1. A full list of rhythmic metabolites annotated to a single metabolite class, across all assays for males and females are provided in Appendix Table D2. A cross referencing approach was taken for the rhythm analysis such that, metabolites were classified as rhythmic if they obtained a significant result ($p < 0.05$) post FDR correction in the majority of applied algorithms (>=3). This method of classification maintains parity with that employed in Chapter 5.

Table 6.1: Number of metabolite features (a) following XCMS raw data processing, (b) following quality filtering and (c) following group rhythm analysis.

Assay	No. of features yielded from raw data processing	No. of features post quality filtering	No. of features submitted to rhythm analysis male/female
HILIC – POS	17219	1605 (9.3%)	1173 (6.8%) / 1091 (6.3%)
HILIC - NEG	11317	902 (8.0%)	589 (5.2%) / 513 (4.5%)
Total	28,536	2,507 (8.8%)	1,762 (6.2%) / 1604 (5.6%)

The male group presented with more rhythmic metabolites than females both in absolute terms (179 vs 129, Figure 6.2) as well as the percentage of rhythmic features out of those submitted for analysis (~26.1% vs ~18.9%). For the majority of metabolite features rhythmic in both males and females the acrophase (peak) was equivalent, with less than 2 h difference between the sexes. Only a select few rhythmic metabolite features exhibited a difference in acrophase between males and females that exceeded the metabolomics sampling resolution of 2 h (n = 9).

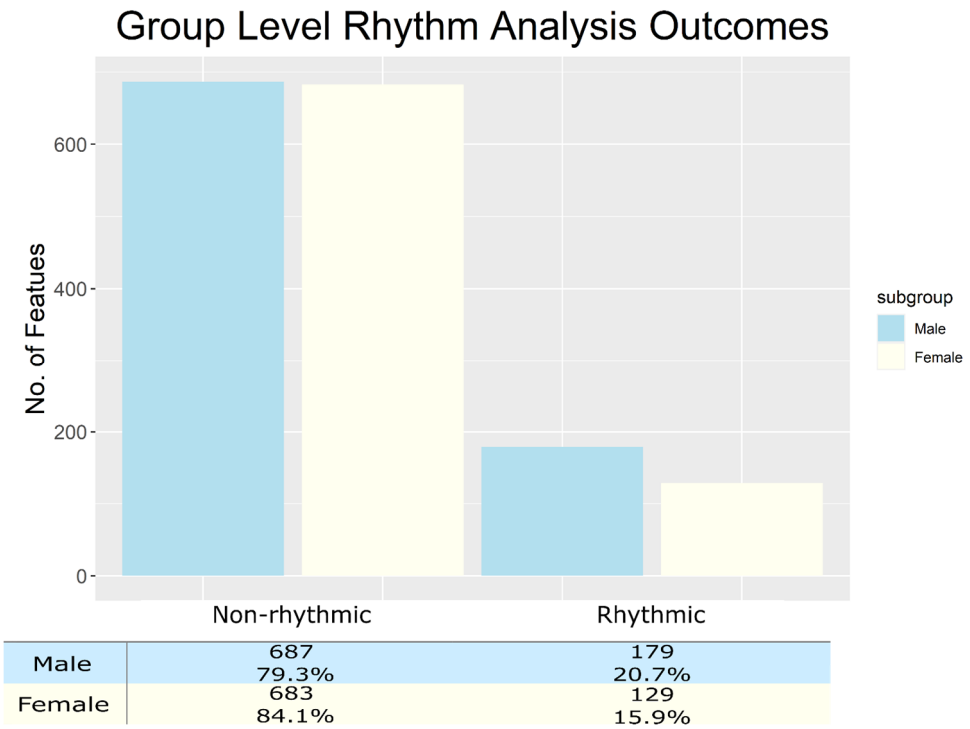


Figure 6.2: Group level rhythm analysis outcomes, number of rhythmic metabolite features by sex (males: n = 10, females n =10).

A full list of the putatively annotated rhythmic metabolites observed in males and females where acrophase differed by >2 h between sex is provided in Table 6.2.

Table 6.2: A list of putatively identified metabolite features with an acrophase difference of >2 h between males and females.

Assay	Putative annotation(s)	Acrophase difference (h) (male relative to female)
HILIC - positive ion mode	Serine	-2
HILIC - positive ion mode	Isoleucylproline | Leucylproline	2
HILIC - positive ion mode	1-Methylguanosine | (S)-malyl alpha-D-glucosaminide	-2

Metabolites are listed in order of descending magnitude, ignoring +/- . Acrophase is rounded to the nearest hour to match the resolution of the DLMO scale.

Despite the seemingly small scale of acrophase differences based on sex, when collectively observing the acrophase of water-soluble metabolites about the time course a striking distinction between males and females becomes apparent. In males most polar metabolites peak at or around 2-6 h after DLMO which corresponds to the middle of the subjective night,

the time at which melatonin levels are highest and when cortisol levels begin to increase approaching the subjective morning. By contrast, water-soluble metabolites in females aggregate around two time points between 0-4 hours after DLMO and 10-14 hours after DLMO, thus, display a very different acrophase distribution relative to males (Figure 6.3). Water-soluble metabolites in females exhibit a similar, acrophase distribution to that observed in lipid metabolites, albeit the water-soluble metabolites tend to peak a few hours later when compared to lipid metabolites.

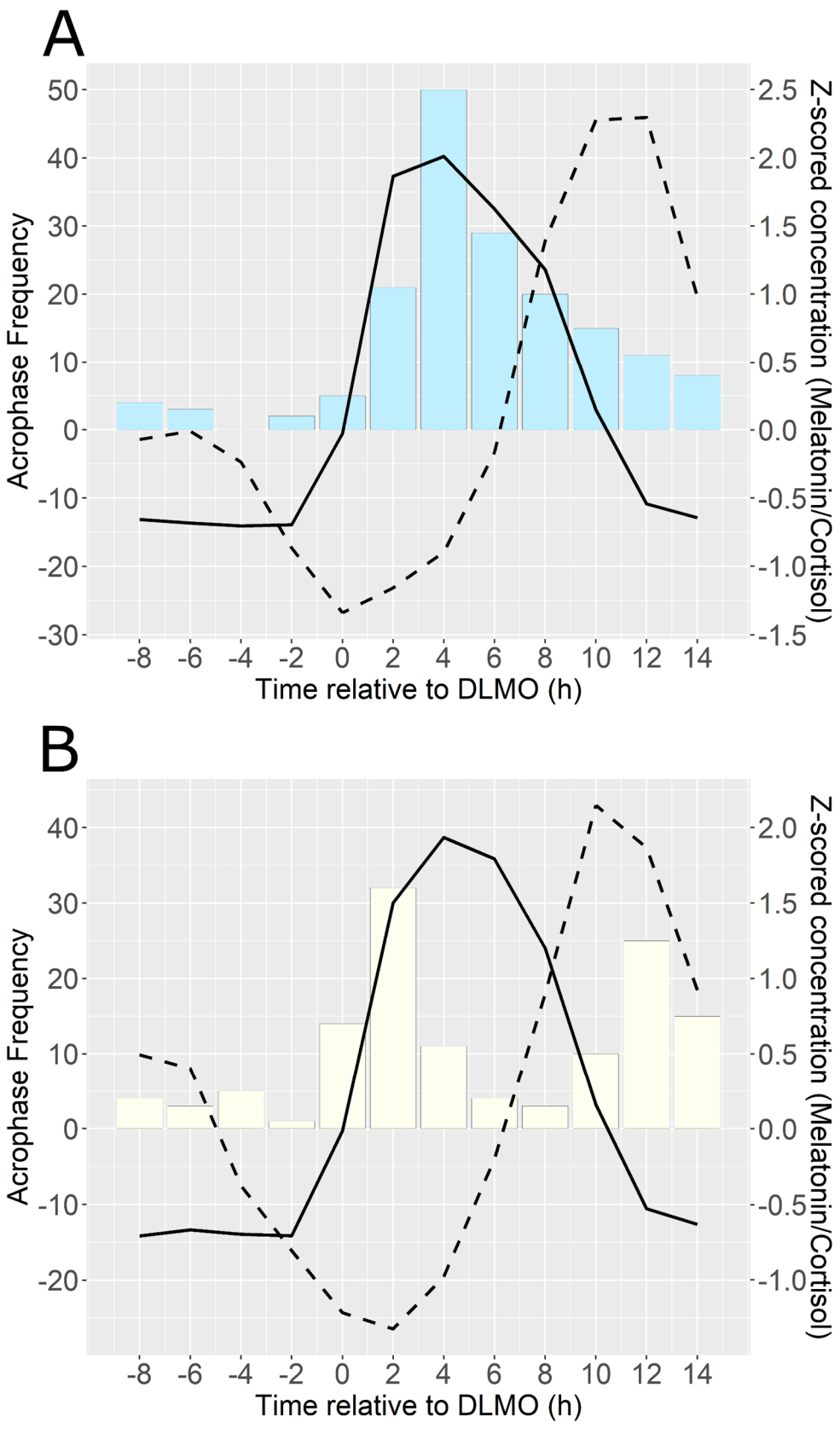


Figure 6.3: Acrophase distributions relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) male HILIC assay, (B) female HILIC assay. Male cohort n = 10, female cohort n =10. For example, 25 metabolite features peaked 12 h after DLMO for the female group.

The acrophase of annotated individual rhythmic metabolites from the HILIC assay are visualised below as heatmaps for males (Figure 6.4A) and females (Figure 6.4B).

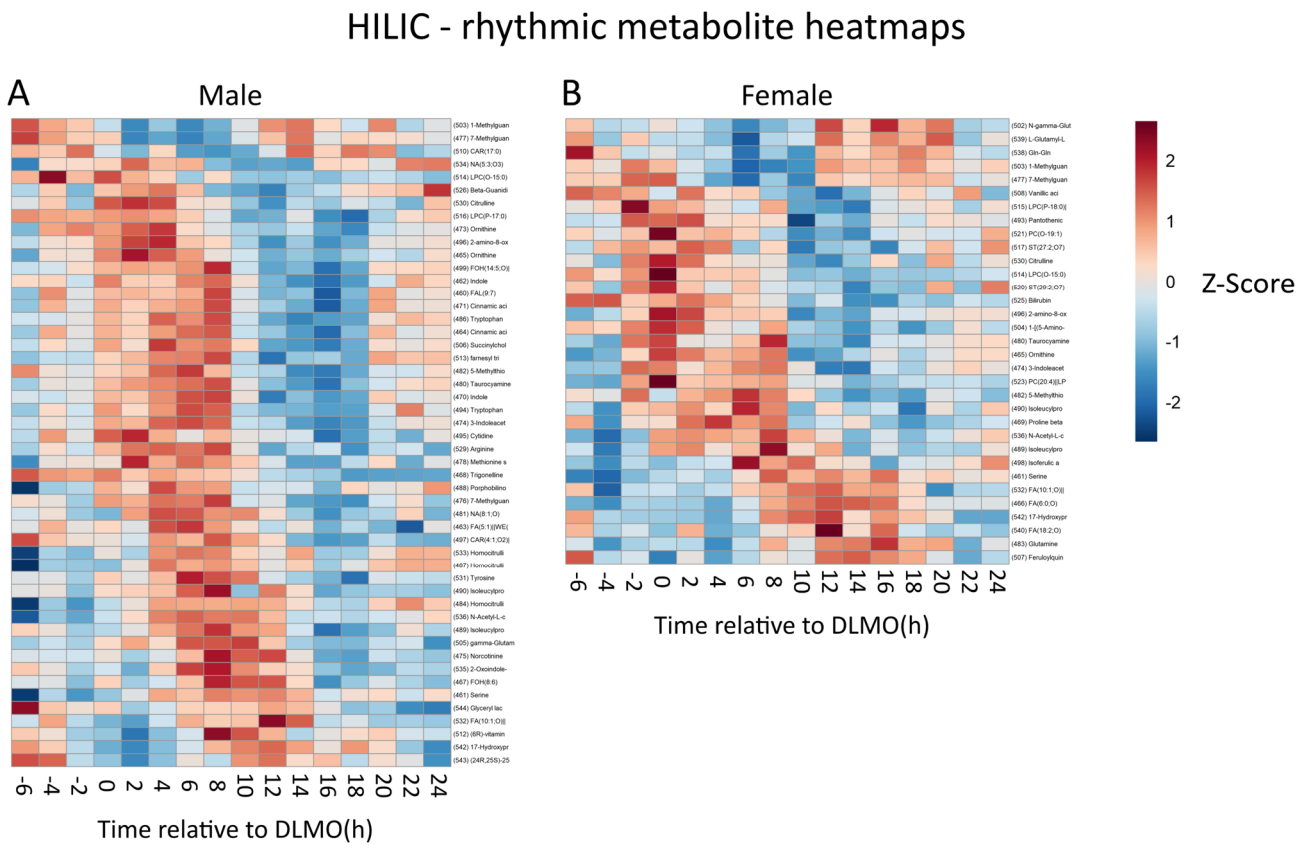


Figure 6.4: Male (A, left) and female (B, right) annotated rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged (male: n = 10, females: n = 10) changes in relative abundance (red = high, blue = low) relative to dim light melatonin onset (DLMO, -6 : 24 h from left to right). Male metabolites (n = 50), female metabolites (n = 33). Note, each row corresponds to a unique metabolite feature, i.e., a unique m/z and RT; however, multiple features may have the same annotation e.g., Tryptophan. Only rhythmic metabolite features annotated to a single metabolite class are displayed. Index numbers can be cross referenced to Table D2 for additional details per metabolite feature.

Annotated rhythmic metabolites reported within this chapter which have previously been reported as rhythmic under constant routine are shown in Table 6.3.

Table 6.3: A list of putatively identified rhythmic water-soluble metabolites observed in both our group level analysis and previous studies.

Metabolite annotation	Observed in
Alanine	[70]
Arginine	[70] [328]
Citrulline	[70, 328]
Glutamine	[70]
Histidine	[70]
Kynurenine	[70]
Lysine	[70, 328]
Ornithine	[70, 328]
Pantothenic acid	[328]
Tryptophan	[319]

Metabolite IDs, as listed, may correspond to multiple rhythmic metabolite features e.g., Glutamine (*m/z* ~185 & ~186 / RT 466).

A total of 257 water-soluble metabolite features were identified as rhythmic in males and/or females by the independent group rhythm analysis tests. Of these 257 metabolite features, differential rhythmicity analysis determined two metabolite features as having a statistically significant difference in acrophase between sex ('change'), Eleven metabolite features were considered significantly rhythmic in females only ('gain') whilst seven metabolite features were considered significantly rhythmic in males only ('loss'). The remaining metabolite features were classified as not being significantly different between males and females ('same' - 237 features). Of the 20 metabolite features identified as having significantly different rhythms between males and females, some have been successfully annotated to a single metabolite class (Table 6.4).

Table 6.4: A list of putatively identified rhythmic water-soluble metabolites also identified as differentially rhythmic between males and females.

Assay	Classification	Annotations (single class)
HILIC – Negative ion mode	'gain' - Rhythmic in females	Glutaminyl-Glutamine
HILIC – Positive ion mode	'gain' - Rhythmic in females	FA(6:0;O)
HILIC – Negative ion mode	'loss' - Rhythmic in males	Arginine

Lipid annotations are given at the 'species level' as defined here [87]. FA – Fatty acyl.

The group rhythm analysis results (Figure 6.2) would suggest more water-soluble metabolite features are rhythmic in males compared to females, whilst the opposite was true when observing lipids (see Chapter 5). However, the former observation is not definitive as differential rhythmicity analysis performed on water-soluble metabolites suggests the majority of rhythmic metabolite feature profiles, 237 of 257, are not significantly different males and females whilst 11 and 7 of these features are rhythmic only in females and males, respectively. Pathway enrichment analysis revealed the 'Arginine biosynthesis' pathway as being significantly enriched in both males and females (Table 6.5).

Table 6.5: Pathway enrichment analysis results for rhythmic putatively identified (non-lipid) metabolites. 'Hits' refers to the number of submitted entries matched to the corresponding pathway. Only significant results ($p < 0.05$) prior to FDR correction are shown.

	KEGG Pathway	Metabolites in pathway	'Hits'	FDR	Impact
Male	Arginine biosynthesis	14	3†	0.025784	0.36548
Female	Arginine biosynthesis	14	3††	0.010349	0.28934

† Enriched metabolites in males – Arginine, Citrulline, Ornithine.

†† Enriched metabolites in females – Citrulline, Glutamine , Ornithine.

6.2.3. Individual level rhythm analysis

Individual level rhythm analysis revealed a high degree of inter-individual variation within and between sexes, a similar observation was made with regards to lipids (see Chapter 5), with the number of rhythmic metabolites ranging from 0-41 across the cohort (n=23) (Figure 6.5A) which is smaller than the range observed in lipids (0-478 across the cohort (n=23), see Chapter 5). When observed as individuals males presented with more rhythmic water-soluble metabolites than females on average, though this difference in mean was non-significant ($p>0.05$). The number of rhythmic metabolites, observed per participant, across the full cohort did not correlate with age, BMI, melatonin or cortisol peak/amplitude/production, order of sample collection, or progesterone/oestrogen dose (relating to oral contraceptive, females only).

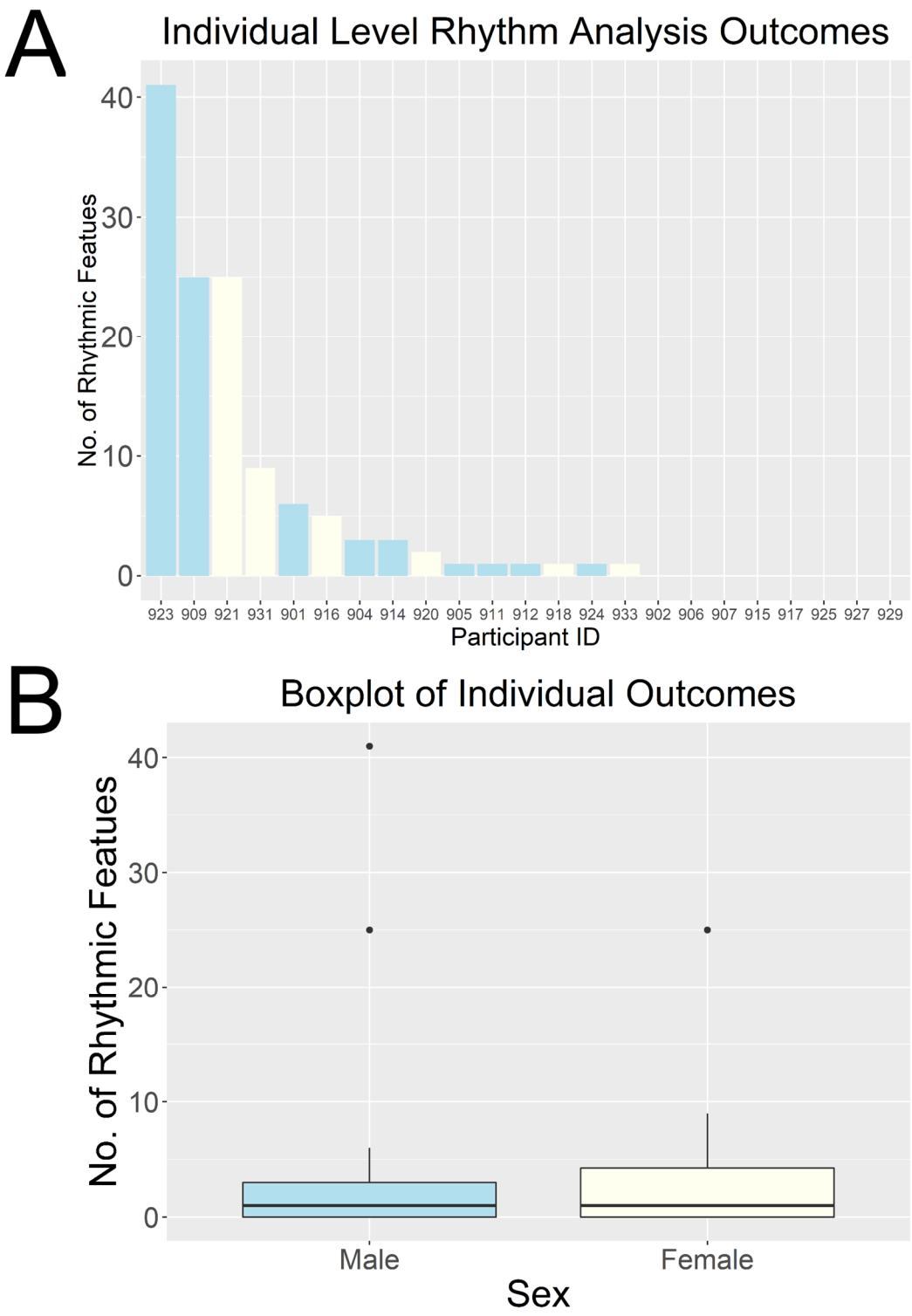


Figure 6.5: (A) – bar chart displaying number of rhythmic metabolites per individual determined by individual level rhythm analysis, (B) – Box and whisker plot of individual level analysis outcomes. Male cohort n = 13, female cohort n =10. A metabolite was classified as rhythmic for a ~24 h period if it achieved a significant result post FDR correction ($p < 0.05$) in at least 3 of the 5 applied rhythmic detection algorithms for individual level analysis: Cosinor, JTK, ARS, RAIN, BioCycle. Group means are significantly different ($p < 0.05$), one tailed T-test. Boxes are displayed with the default ggplot2 settings: The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles), The whiskers extend from the hinge to the largest/smallest value no further than 1.5 * interquartile range.

6.3. Discussion

The first objective of this study was to identify and report a list of circadian regulated water-soluble metabolites (~24h period) observed under a constant routine study and therefore not influenced by external cues (i.e., diurnal rhythms). Across two HILIC UHPLC assays (HILIC in positive and negative ion modes) a total of 2,507 metabolite features were reported after data quality filtering. For males and females respectively, 179 and 129 rhythmic metabolite features (~24 h period) were reported and of these 50 and 33 rhythmic metabolites were annotated to a single metabolite class, respectively. All other metabolite features did not exhibit a ~24 h period. In addition, a further 129/96 rhythmic metabolite features (*m/z*-retention time pairs) with a ~24 h period in males and females, respectively, were not structurally annotated. The lipidomics assays introduced in Chapter 5 yielded over double the metabolite features (5,185) of which a greater proportion were rhythmic in males (898, 37.2%) and females (1073, 48.2%). Prior lipidomic studies have established that the plasma metabolome is primarily made up of lipids, as measured by weight per unit volume [409]. Further, at least 588 distinct molecular lipid species have been characterised in blood [410], forming the foundation of curated human plasma lipid entries for public databases such as LIPIDMAPS. Many of these individual lipid species, e.g., TG(48:2), can have multiple conformations such as varying lengths of the three fatty acids further increasing the potential lipid diversity in human plasma. Detecting a greater number of lipid features compared to water-soluble metabolite features is therefore not wholly unexpected. What is of greater interest is why a greater proportion of the detected lipids are rhythmic versus that of the proportion of rhythmic water-soluble metabolites detected. Perhaps lipids regulation about a ~24 h cycle is vital to better fulfil their biological function vs water-soluble metabolites. Prior transcriptional studies have shown the liver, which regulates lipid homeostasis [411], has the

highest percentage of circadian regulated protein coding genes (16%, > 3000 genes) of the major organs supporting this notion [412]. As discussed in Chapter 5, metabolic cycles such as the triglyceride/free fatty acid cycle (TG/FFA) are evidenced as having circadian regulated enzymes. This is likely a means of temporally regulating energy storage/production to coincide with rest/activity, respectively. Both TGs, as well as diglycerides (DG) and fatty acids featured prominently within the rhythmic lipid dataset. Adipose tissue contains a diverse lipidome (>100 lipids) primarily comprised of TGs [413], thus, this one metabolic cycle could confer rhythmicity to a broad range of closely related metabolites such as TGs and downstream metabolites such as the resulting fatty acids yielded from lipolysis. These lipids would be transported to and from the liver/adipose tissue via the blood stream thus being detectable in plasma. To summarise this potential scenario, a single metabolic cycle (TG/FFA) confers rhythmicity to a broad range of closely related metabolites that are transported by blood and make up the majority of the blood metabolome. Therefore upon analysis, lipids appear vastly enriched with regards to the circadian regulated metabolite in comparison to water-soluble metabolites which constitute a smaller proportion of the sample metabolome and are sourced from a broader range of tissues/metabolic outputs.

Whilst previous circadian studies employing metabolomics did not apply HILIC assays to specifically observe water-soluble metabolites, with the exception of Grant [328] discussed below, some circadian water-soluble metabolites have previously been reported [78]. Not all of these prior reported metabolites corroborate with the data presented within this chapter, those that do are presented in Table 6.3. Isoleucine, leucine, phenylalanine, proline, sarcosine, serotonin, taurine, and threonine have all been reported as rhythmic in prior circadian studies but were not observed as rhythmic within this chapter. For the data presented within this chapter: isoleucine, leucine, sarcosine, serotonin, threonine were not

confidently annotated, phenylalanine and taurine (in females) were annotated but non-rhythmic (~24 h period), metabolite features annotated as proline, taurine (in males) were not submitted to rhythmic analysis due to excessive missing values. Notably, there is little overlap in the list of circadian metabolites between HILIC assays presented within this thesis (2507 metabolite features detected) with the only other circadian study which applied a HILIC assay (targeted assay – 99 metabolites, untargeted 1641 metabolite features detected) [328]. The HILIC assay presented found 'arginine biosynthesis' significantly enriched whilst Grant *et al.* found phenylalanine and tyrosine metabolism (group level analysis), and the citric acid and urea cycles (individual level analysis) post FDR correction in their targeted assay to be enriched. With only minor overlap between the results of this study and Grant's, the compiled list of rhythmic water-soluble metabolites presented here contains metabolites not previously reported to be circadian [78] yet is similar in size to the untargeted HILIC assay previously performed by Grant. Whilst Grant's study is similar to the study presented here it is important to note key differences in the experimental design which likely contribute to the lack of full corroboration between results, namely, the focus on sex dependent differences and different methods of rhythm analysis which can yield varying results as discussed in Chapter 4 and discussed in the literature [354, 361]. In their concluding remarks Grant suggested there were differences apparent between males and females, with regards to circadian regulated metabolites, but a future specifically designed study would be required to test this hypothesis. The data presented, corroborates with this suggestion, and sets a precedent, alongside the data presented in Chapter 5, that circadian regulation of metabolism differs between males and females which leads onto the second objective.

The second objective was to identify sex dependent differences within the cohort at both an individual level and a group level, with regards to rhythmic water-soluble metabolites. Whilst

inter-individual variation was apparent amongst the cohort there was no significant difference in the mean value of the number of rhythmic metabolites in males vs females. It is possible that the size of the cohort, which is large in the context of human circadian studies but small in terms of statistical power, reduces our ability to discern male vs female differences when inter-individual variation is so large. In other words, individual characteristics external to sex and other controlled variables (age, BMI) may be a source of individual level variation in the number of rhythmic metabolites observed that is equal to or exceeds the variation caused by sex, at least in the context of water-soluble metabolites. Such challenges associated with individual level rhythm analysis with the lipids data in Chapter 5 were also a point of discussion with regards to water-soluble metabolites (see Section 5.3). As constant routine studies require specialist facilities to perform and a demanding sampling protocol a study such as the one presented here is not easily upscaled and so the issue of larger cohort sizes is not a simple problem to address. As discussed in the introduction to Chapter 5 (see Section 5.1), smaller cohorts may actually be favoured to maintain more typically employed small but highly controlled study groups to restrict confounding factors which may otherwise mask the rhythms being observed, but this comes at the aforementioned cost of reduced statistical power due to the smaller cohort population. This is where non-circadian (e.g., single sampling) metabolic phenotyping studies with larger populations, such as a prior study by Dunn *et al* [395], may be beneficial to circadian research. Whilst such studies do not offer insight into circadian level differences they do act as a signpost to metabolites which differ in concentration based on a particular phenotype (e.g., sex) and how different phenotypes may interact (e.g., sex and blood pressure) which will be of interest in circadian studies. Case and point, Dunn identified creatinine, tyrosine, citric acid (citrate) and uric acid (urate) as expressing different serum concentrations between males

and females, among other water-soluble metabolites. The former two metabolites were identified as rhythmic within the male group whilst the latter was identified as rhythmic only in the female group within the data presented in this chapter. In reference to Grant's study [328] the citric and uric acid cycles in addition to tyrosine metabolism were enriched pathways. Further, citrulline, tryptophan, lysine, threonine, among other water-soluble metabolites, were identified by Dunn as metabolites which exhibit varying concentrations dependent on an interaction between sex and blood pressure. These metabolites have previously been identified as being rhythmic (see Appendix A5/D1) and the former two were rhythmic in both our male and female cohort at the group level yet were not rhythmic across all individuals in our individual level analysis. Not all metabolites identified by Dunn as exhibiting differences in concentration dependent on sex were differentially rhythmic between our male and female groups, however, some were illustrating the case that non-circadian population wide studies may benefit circadian studies by signposting metabolites of interest, e.g., with regards to sex, which may otherwise go unnoticed in the smaller populations employed in circadian studies. It is possible that complex interactions, such as sex and blood pressure, contribute to this inter-individual variation and even male vs female differences in unseen ways, particularly in this instance as no metadata is available on blood pressure for the cohort. As discussed in Chapter 1, metabolism and the circadian system are intertwined systems that regulate each other, therefore, phenotypic interactions which influence metabolism will influence the clock in turn and vice versa. Thus, population level studies may prove informative to circadian research as it may help explain why metabolites that exhibit daily rhythms in one seemingly healthy population but not in another, perhaps due to unforeseen variation caused by unexpected interactions between specific phenotypes.

Unfortunately, the pathway enrichment analysis results were limited within the study described here, with only arginine biosynthesis being enriched amongst metabolites which are circadian regulated. However, arginine metabolism and how arginine relates to other aspects of metabolism is reasonably understood [414]. Closely related metabolites to arginine have either previously been shown to be rhythmic (circadian/diurnal) in mammals (see appendix: section A Table A5 or Table S1 in [78]) or are rhythmic within the results presented within Chapters 5 & 6 (see appendix Table C2/D2). Such metabolites include ornithine, citrulline, urea, creatine, glutamate, glutamine and alpha-ketoglutarate. Thus, the circadian synthesis of arginine may confer rhythmicity to some of these metabolites, dependent upon its availability and dependent upon circadian rhythms of numerous enzyme gated reactions. Adjoining metabolic processes closely related to arginine such as: citrulline metabolism, urea (ornithine) cycle, and the citric acid (TCA) cycle may not have been enriched within the dataset presented in this chapter due to metabolites comprising the cycle not being detected or not classified as rhythmic within the cohort. However, these cycles were classified as rhythmic and enriched in Grant's study. Whilst definitive biological conclusions cannot be drawn, due to a lack of corroboration between the work of Grant and data presented here, the above listed pathways should be considered of interest with regards to future circadian studies investigating water-soluble metabolites to ascertain the rhythmic regulation more confidently, if any, of these pathways.

6.4. Chapter summary

Whilst the data presented within this chapter did not fully corroborate with Grant's findings, the only prior study investigating circadian regulation of metabolism in humans under constant routine, the data presented in this chapter was able to confirm their initial

observations of sex dependent differences in circadian regulation of polar metabolites. Moreover, taking the data presented in both Chapter 5 and 6 in hand, circadian regulation is more pervasive with regards to lipids and lipid metabolism and less so for water-soluble metabolites and associated pathways. As discussed across previous chapters there are a number of metabolic processes pertaining to lipid metabolism that are circadian regulated, the results presented fit this narrative yet should not undermine the need to further investigate water-soluble metabolites. Investigating circadian rhythms requires well designed sampling procedures and specialist facilities which allow for accurate measurement of metabolite abundance/concentration changes over time yet often can only employ small cohorts. Larger scale non-circadian population studies may serve as a useful guide for metabolites of interest, with regards to circadian studies, as they have the statistical power necessary to discern subtle differences between populations be it shiftwork vs non-shiftwork, male vs female, or interactions between biological or environmental factors. Once such metabolic differences between populations are identified then circadian studies can profile metabolic changes over the course of the day to investigate if there is a circadian component driving the metabolic differences between the population. Just as untargeted metabolic studies facilitate the production of future targeted studies, so too could single timepoint/paired sampling studies lay the groundwork for multiple sampling metabolic studies. Utilising pre-existing literature in this manner may serve to bolster the success of future circadian-metabolomic assays focussing on water-soluble metabolites and yield more informative results from pathway enrichment analysis which is currently hampered by restricted statistical power in identifying group level differences, identifying rhythmic metabolite features of interest, and successfully annotating said features.

CHAPTER 7

METABOLOMIC ANALYSIS OF PLASMA IN HUNTINGTON'S DISEASE TRANSGENIC SHEEP (OVIS ARIES) REVEALS PROGRESSIVE DIURNAL RHYTHM DYSREGULATION.

The following chapter contains text and figures reproduced from a publication in the Journal of Huntington's Disease in which I am a named author, see Spick, M., Hancox, T.P.M., Chowdhury, N.R., Middleton, B., Skene, D.J., Morton, A.J. (2023), "Metabolomic Analysis of Plasma in Huntington's Disease Transgenic Sheep (*Ovis aries*) Reveals Progressive Circadian Rhythm Dysregulation", Journal of Huntington's Disease, Vol 12(1), pp 31-42, doi: 10.3233/JHD-220552 [282]. Any analysis/commentary not originally authored by myself has not been included in this chapter. Where text/figures have been reproduced from the publication then this will be clearly indicated with the above reference. My role in the publication was to perform the rhythm analysis and create the subsequent figures and write discussion points related to this analysis. I was not involved in the experimental design, sample collection, metabolomics analysis, cortisol/melatonin/urea assays or analyses unrelated to the aforementioned rhythm detection.

7.1. Introduction

Neurodegenerative disorders are diseases of the nervous system characterised by the progressive loss of neuronal populations with a common mode of neurodegeneration being the aggregation of excess protein in nerve cells. Dependent on the neurodegenerative disease in question these protein aggregates may differ in composition and location and result in different cognitive or behavioural disorders which have clinical relevance in diagnosing the specific disorder, such as Alzheimer's disease, Parkinson's disease, or the focus of this study Huntington's disease (HD) [415, 416]. Unlike the former two diseases the onset of HD is not typically associated with the onset of old age and so the disease may present in children as well as young and old adults but primarily it manifests in middle age [416]. The age at which HD onsets, if at all, is dependent on the nature of the mutation which causes the disease. HD is caused by an autosomal dominantly inherited mutation of the *IT15* (*HTT*) gene which leads to multiple trinucleotide repeats (CAG) within the huntingtin protein (HTT) resulting in mutant huntingtin (mHTT) protein that contains an elongated polyglutamine repeat [416, 417]. Individual's with mHTT that contains > 39 CAG repeats are highly likely to develop HD with the median age of onset decreasing with an increasing number of CAG repeats where juvenile onset of the disease may be observed in individuals with >60 CAG repeats in mHTT [418, 419]. In HTT the number of CAG repeats is typically < 26, with intermediate alleles with 27-35 and 36-39 CAG repeats; the former is unlikely to result in HD but is considered genetically unstable and may cause HD in successive generations should the number of CAG repeats increase whilst the latter can result in HD that is not fully penetrant, i.e., different individuals may or may not develop HD symptoms [417, 420].

Both HTT and mHTT contain proteolytic sites for several different proteases, however, proteolytic cleavage of mHTT in cases of HD is enhanced leading to increased production and accumulation of small N-terminal fragments that contain the polyglutamine sequence (coded for by CAG repeats) [421]. Specifically, caspase cleavage on an aspartate residue at position 586 (D586) favours the production of insoluble cytotoxic aggregates that greatly contribute to neuronal cell death and disease progression [416, 422]. The loss of wild type function of HTT e.g., intracellular trafficking, maintaining macrophage, and potential scaffold protein in addition to toxic gain of function following proteolytic cleavage define the pathogenesis of HD [421-423]. The aforementioned N-terminal fragments can have numerous harmful effects on a cell including: translocation into the nucleus where they aggregate and dysregulate transcription; interactions with autophagy receptors blocking their binding to and subsequent degradation of damaged mitochondria; impairments to cell metabolism and energy balance (Figure 1) [416, 417, 421, 424]. The nature of this metabolic dysfunction in HD is not fully characterised, yet current evidence suggests a strong link between altered metabolism and disease progression. Notable findings include reduced glucose consumption/glucose accumulation within brain tissues in HD which is hypothesised to contribute to cell damage by increasing osmotic stress or chemical toxicity due to excess glucose or its metabolites synthesised from the polyol pathway (sorbitol, fructose) [425]. Further reported perturbations include: glycolysis, TCA (citric acid) cycle, urea (ornithine) cycle, pyruvate metabolism, phosphate & purine pathways, which may represent disrupted central carbon metabolism and a shift to alternate methods of energy production in HD, e.g., increased reliance on ketone bodies produced via glyceroneogenesis [416, 425, 426]. With regards to purine metabolism, uric acid is considered a neuroprotective agent but its concentration in serum and brain tissue decreases with HD progression [416]. Finally, perturbations to

tryptophan metabolism, specifically the kynurenine pathway may directly contribute to increased neurotoxicity. Quinolinic acid, a product of the kynurenine pathway, exhibits elevated levels in brain tissue of HD cases at the early stages of the disease, more so for areas of the brain most damaged by cell death. As a neuroactive metabolite these elevated levels are believed to contribute to excitotoxic neuronal damage. In tandem, levels of 3-hydroxykynurenine are increased to the point of inducing neurotoxicity specifically to striatal neurons. Compounding the impact of quinolinic acid are decreased levels of kynurenic acid in HD which acts as a neuroinhibitory metabolite thus affords some protection against excitotoxic agents such as quinolinic acid. Therefore, the lower levels of kynurenic acid observed in HD is believed to induce greater vulnerability to excitotoxic neuronal damage via quinolinic acid [427, 428]. It has also been hypothesised that the perturbations to the kynurenine pathway may promote susceptibility to glutamate induced excitotoxicity [429]. Simply summarised, the perturbed metabolism observed in HD appears to contribute to disease progression by promoting the accumulation of metabolites, e.g., those which are neuroactive, to the point of inducing toxicity whilst levels of protective agents such as uric acid or kynurenic acid are diminished. Damage to cell organelles, such as mitochondria, contribute to a lack of energy balance and may in part explain the lack of glucose utilisation and apparent shift in metabolism in an attempt to utilise alternative energy sources. The induced neuronal cell death manifests as numerous symptoms such as changes in behaviour or loss of fine motor control, moreover, HD may also cause neurodegeneration of the hypothalamus which houses the suprachiasmatic nucleus (SCN) the master circadian pacemaker (See Chapter 1) [417, 430]. Perturbations to the circadian system and sleep are therefore closely linked to HD pathology [430, 431].

HD is considered a rare disease, despite being the most common monogenic neurodegenerative disease in the west affecting ~10.6-13.7 per 100,000 in 'western' populations, with lower prevalence observed in non-white ethnic groups [417, 432]. HD can be pre-emptively tested for in patients where there is a known family history, else, patients will likely only receive a diagnosis after onset of symptoms [417, 432]. As a rare disease the pool of potential candidates for HD clinical trials is limited which restricts the frequency and scope of human clinical trials generating a need for an adequate animal model [433]. Rodents are the most commonly employed animal model to investigate HD, but questions have been raised on their efficacy for studying HD due to the differences in brain size and morphology vs larger brained mammals such as humans and their short lifespan is counterproductive to study a disease which develops progressively over an extended period of time. Sheep (*Ovis aries*) have therefore been proposed as a large animal model to circumvent these issues with a typical life span of ten years and possessing all the brain structures (nuclei) expected of a large mammal [433, 434]. Added benefits of a large animal model include ease of imaging (via PET or MRI) and drug/gene therapy agent delivery due to the larger brain size [434]. The transgenic HD sheep (*Ovis aries* – Merino breed) model OVT73 possesses a full length mHTT transgene with a 73 CAG repeat sequence. Carriers of this transgene do not exhibit brain abnormalities characteristic of HD pathology at age five but there is evidence of perturbed circadian rhythms (worsening with age) which manifests as altered activity cycles. Thus, sheep at this age are considered pre-symptomatic based on current guidelines of assessing HD progression but are not devoid of changes in comparison to their wild type counterparts. Case and point, metabolic differences between wild type and HD sheep have been reported at age five from post-mortem liver/brain tissue samples [434]. This would suggest that metabolic disruption and small alterations in behaviour, in this model at least, precedes atrophy of brain

tissue characteristic of HD progression. This model therefore affords the opportunity to investigate metabolic disturbances, as well as the progressive perturbations to the circadian system associated with the disease, in pre-symptomatic HD cases providing improved insight into HD pathology at the earliest stages.

The purpose of this study was to utilise the aforementioned HD sheep model (OVT73, *Ovis aries* – Merino breed) to (1) characterise differences in circadian regulated metabolism between control (wild type) and HD sheep (ewes) at ages of five and seven years and (2) ascertain metabolic pathways disrupted as a result of HD pathology at the pre-symptomatic stage. Time series blood samples were acquired under constant routine-esque conditions (darkness, fasted, restricted activity) with a sampling resolution of 1 h over 27 h. Samples were analysed via a commercially available targeted metabolomics assay (AbsoluteIDQ® p180 targeted metabolomics kit (Biocrates Life Sciences AG, Innsbruck, Austria)), to acquire metabolite concentration values across the time course for up to ~180 metabolites. As HD is progressive it was hypothesised that metabolic and circadian disruption would be greater in the older HD aged group versus wild type counterparts (control group) of the same age.

7.2. Results

7.2.1. Data processing outcomes

Data was collected for 13 control sheep and 14 HD sheep at five years of age (2015 cohort) and from 15 control sheep and 16 HD sheep at seven years of age (2017 cohort). Of the control sheep, eight were sampled at both five and seven years of age whilst ten HD sheep were sampled at both five and seven years of age, all other sheep were only sampled at a single age. Within the 2015 and 2017 cohorts a total of 133 and 128 metabolites, respectively, were detected and passed QC filtering checks. Of these metabolites, only one was excluded (Met-SO, Methionine sulfoxide) from rhythm analysis due to excessive missing values (see Methods 2.3.3.3.). Prior work established batch effects between samples analysed from the 2015 cohort and the 2017 cohort believed to be the result of changes made by the manufacturer in the applied Biocrates assays affecting the metabolomics analyses in unforeseen ways. Therefore, the decision was made to avoid inter-cohort comparisons (2015 *vs.* 2017) and rely on intra-cohort comparisons (control vs HD) at both ages of sheep separately [282].

7.2.2. Unsupervised multivariate analysis

A full list of rhythmic metabolites, alongside metabolite amplitude/midline estimating statistic of rhythm, are provided in Appendix Table E1. As in Chapters 5 and 6, a cross reference approach was taken for the rhythm analysis such that, metabolites were classified as rhythmic if they obtained a significant result ($p < 0.05$) post FDR correction in the majority of applied algorithms (>=3). Metabolites classified as rhythmic were subject to differential rhythmicity analysis and classified as ‘same’ (no significant difference in rhythm between control / HD), ‘change’ (significant difference ($p < 0.05$) in acrophase between control / HD), ‘gain’ (significant difference ($p < 0.05$) in rhythm between control / HD, only rhythmic in HD),

'loss' (significant difference ($p < 0.05$) in rhythm between control / HD, only rhythmic in control). Unsupervised analysis of the metabolome, outside of the time domain, for the 2015 and 2017 cohorts via PCA showed little inter-group difference (control vs HD) for either cohort with inter-group variation exceeded by intra-group variation.

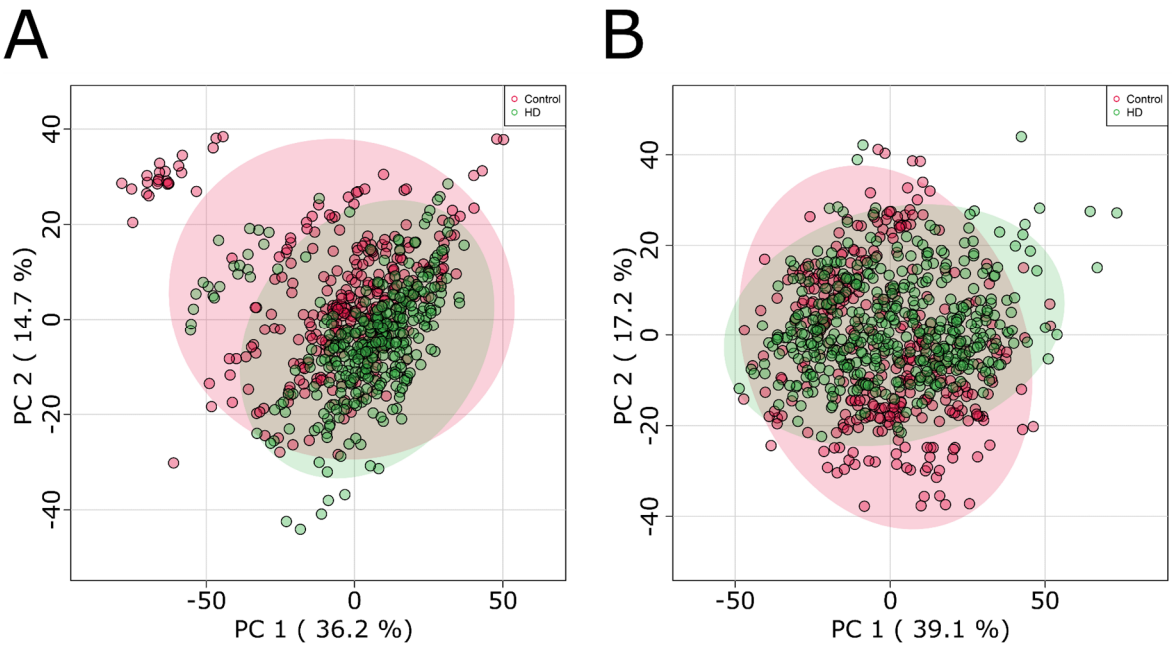
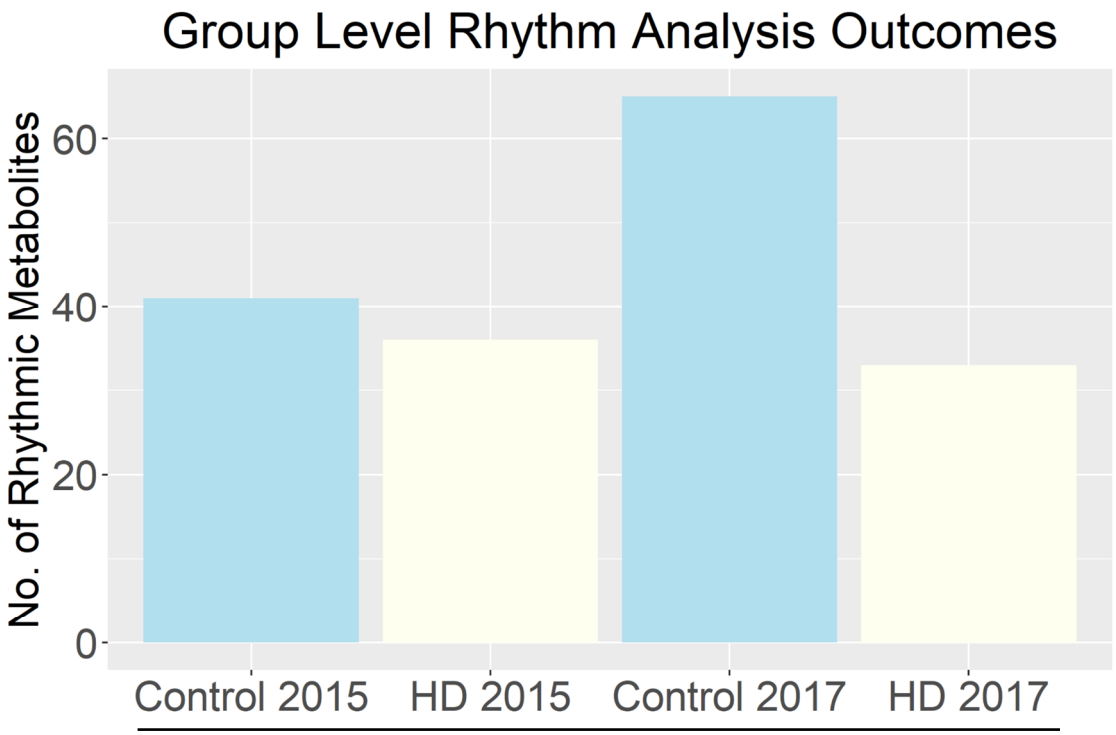


Figure 7.1: Principal components analysis of all metabolites across all time points for (A) 2015 control cohort (n = 13, red) & HD (n = 14, green) and (B) 2017 HD control cohort (n = 14, red) & HD cohort (n = 16, green). PCA analysis displays very little inter-group variation in this unsupervised analysis with the majority of samples from either group clustering together at both time points. Other sources of inter-individual variation appear to exceed the variation induced by the HD transgene at the pre-symptomatic stage when observing the serum metabolome as a whole in an unsupervised manner.

7.2.3. Group level rhythm analysis

Investigations into the time domain, such as the group rhythm analysis better reveal inter-group differences (Figure 7.2) where the control groups exhibit more rhythmic metabolites (five years of age = 41 metabolites, seven years of age = 65 metabolites) than the HD groups (five years of age = 36 metabolites, seven years of age = 33 metabolites) especially at seven years of age.



	2015	2017
Control	41	65
HD	36	33

Figure 7.2: Group level rhythm analysis outcomes, number of rhythmic metabolites by time point (2015, 2017) and group (control = white, HD = blue). Cohort size: 2015 control, n =13, 2015 HD, n = 14, 2017 control, n = 14, 2017 HD, n = 16

Furthermore, such inter-group differences present as differences in acrophase between control and HD groups with three and five metabolites exhibiting an acrophase difference of >2 h in the 2015 and 2017 cohort, respectively (Table 7.1).

Table 7.1: A list of metabolites with an acrophase difference of >2 h between control and HD sheep at 2015 (age five) or 2017 (age seven) time points.

Cohort	Metabolite IDs	Acrophase difference (h) (control relative to HD)
2015	LysoPC a C16:1	-4
2015	LysoPC a C18:1	-3
2015	LysoPC a C18:2	-3
2017	Asn (Asparagine)	7
2017	Phe (Phenylalanine)	5
2017	Urea	5
2017	Kynurenine	4
2017	C18 (Octadecanoylcarnitine)	2

Metabolites are listed by cohort then in order of descending magnitude, ignoring +/- . Acrophase is rounded to the nearest hour to match the resolution of the DLMO scale. Metabolite names maintain parity with naming conventions utilised by the applied Biocrates p180 targeted assay kit, details provided below. PC – Glycerohosphatidylcholine. ‘a’ or ‘aa’ – both moieties (or one in the case of lysoPCs) at the sn-1 and sn-2 positions are fatty acyl residues bound to the glycerol backbone by ester bonds. ‘ae’ - one of the moieties, either in the sn-1 or sn-2 position, is a fatty alcohol residue bound by an ether bond.

Visualisations of individual metabolite rhythms, via heatmaps (Figure 7.3) for all rhythmic metabolites display highly similar rhythmic patterns between control and HD sheep within the 2015 cohort yet there are perturbations to rhythm shape for numerous metabolites between control and HD sheep for the 2017 cohort, predominantly effecting glycerophospholipids (GP).

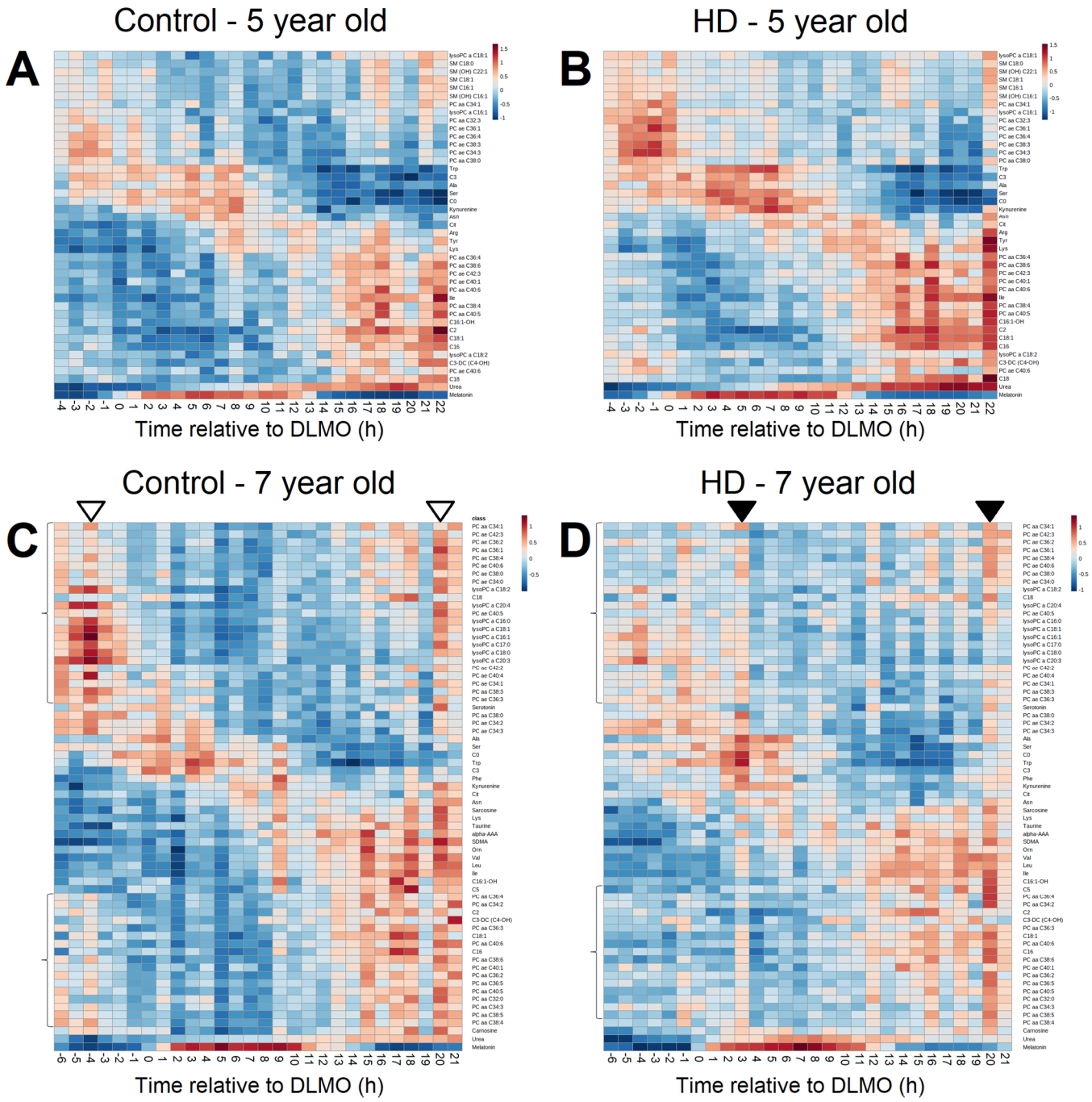


Figure 7.3: Reproduced in its entirety from [282]. Comparison of rhythmic metabolites in control sheep to the same metabolites in HD sheep. Rhythmic metabolites shown as heatmaps displaying Z-scored and group averaged changes in metabolite concentrations (red = high, blue = low) relative to dim light melatonin onset (DLMO) in 5-(control group, A; HD group, B) or 7-(control group, C; HD group, D) year old sheep. DLMO measurements were as follows: for 5-year-old control sheep (n = 13), DLMO= 17.6±4.5 h (A); 5-year-old HD sheep (n = 14), DLMO= 18.8±2.6 h (B); 7-year-old control sheep (n = 14), DLMO= 19.4±0.6 h (C); 7-year-old, HD Sheep (n = 16); DLMO 20.7±0.7 h (D). Brackets in C, D indicate metabolites with rhythm patterns that are distinctively different in control and HD sheep. Arrowheads in C, D indicate the approximate peak times of the rhythms.

Differential rhythmicity analysis did not identify any metabolite rhythms deemed significantly different between control and HD sheep in the 2015 cohort, meaning all 55 metabolites classified as rhythmic in both or either group were considered to have rhythms that did not significantly differ. For the same analysis performed on the 2017 cohort it was revealed that 46 metabolite rhythms did not significantly differ between control and HD cohort whilst 32 metabolite rhythms were deemed significantly different between control and HD sheep (22 – 'change', 2 – 'gain' , 8 – 'loss', Table 7.2). Metabolite rhythms classified as significantly 'changed' between the 2017 control and HD group were predominantly made up of GPs and lysoGPs, corroborating with the visual differences in GP rhythms observed within the produced heatmaps (Figure 7.3).

Table 7.2: A list of rhythmic metabolites also identified as differentially rhythmic between Control and HD sheep, 2017 (age seven) cohort.

Cohort	Classification	Metabolite
2017	'change' – difference in acrophase/shape	Ala (Alanine)
2017	'change' – difference in acrophase/shape	Asn (Asparagine)
2017	'change' – difference in acrophase/shape	C18:1 (Octadecenoylcarnitine)
2017	'change' – difference in acrophase/shape	Kynurenine
2017	'change' – difference in acrophase/shape	lysoPC a C16:0
2017	'change' – difference in acrophase/shape	lysoPC a C16:1
2017	'change' – difference in acrophase/shape	lysoPC a C18:0
2017	'change' – difference in acrophase/shape	lysoPC a C18:1
2017	'change' – difference in acrophase/shape	lysoPC a C18:2
2017	'change' – difference in acrophase/shape	lysoPC a C20:3
2017	'change' – difference in acrophase/shape	lysoPC a C20:4

2017	‘change' – difference in acrophase/shape	PC aa C32:0
2017	‘change' – difference in acrophase/shape	PC aa C34:4
2017	‘change' – difference in acrophase/shape	PC aa C36:1
2017	‘change' – difference in acrophase/shape	PC aa C36:2
2017	‘change' – difference in acrophase/shape	PC aa C38:3
2017	‘change' – difference in acrophase/shape	PC aa C38:4
2017	‘change' – difference in acrophase/shape	PC aa C38:5
2017	‘change' – difference in acrophase/shape	PC aa C38:6
2017	‘change' – difference in acrophase/shape	PC ae C36:1
2017	‘change' – difference in acrophase/shape	PC ae C36:4
2017	‘change' – difference in acrophase/shape	PC ae C42:3
2017	‘gain' - Rhythmic in HD	His (Histidine)
2017	‘gain' - Rhythmic in HD	Thr (Threonine)
2017	‘loss' - Rhythmic in control	alpha-AAA (alpha-Aminoadipic acid)
2017	‘loss' - Rhythmic in control	Carnosine
2017	‘loss' - Rhythmic in control	Cit (Citrulline)
2017	‘loss' - Rhythmic in control	Lys (Lysine)
2017	‘loss' - Rhythmic in control	PC aa C40:5
2017	‘loss' - Rhythmic in control	PC ae C40:4
2017	‘loss' - Rhythmic in control	PC ae C40:5
2017	‘loss' - Rhythmic in control	Sarcosine

Metabolite names maintain parity with naming conventions utilised by the applied Biocrates p180 targeted assay kit, details provided below. PC – Glycerohosphatidylcholine. ‘a’ or ‘aa’ – both moieties (or one in the case of lysoPCs) at the sn-1 and sn-2 positions are fatty acyl residues bound to the glycerol backbone by ester bonds. ‘ae’ - one of the moieties, either in the sn-1 or sn-2 position, is a fatty alcohol residue bound by an ether bond.

To ascertain when metabolites which were deemed as significantly different between the control and HD cohorts were peaking across the time course in relation to melatonin and cortisol, the following plots were produced (2015 cohort – Figure 7.4, 2017 – Figure 7.5). For both the 2015 and 2017 cohorts, control and HD sheep present with a robust melatonin rhythm with metabolites typically peaking around DLMO and approximately 12 h later (subjective morning) as melatonin levels begin to decline. Whilst melatonin was classified as significantly rhythmic across all groups, an unremarkable result due to its status as the gold standard circadian marker, cortisol was not classified as rhythmic for the control or HD sheep at either timepoint when it would be expected to follow a circadian rhythm.

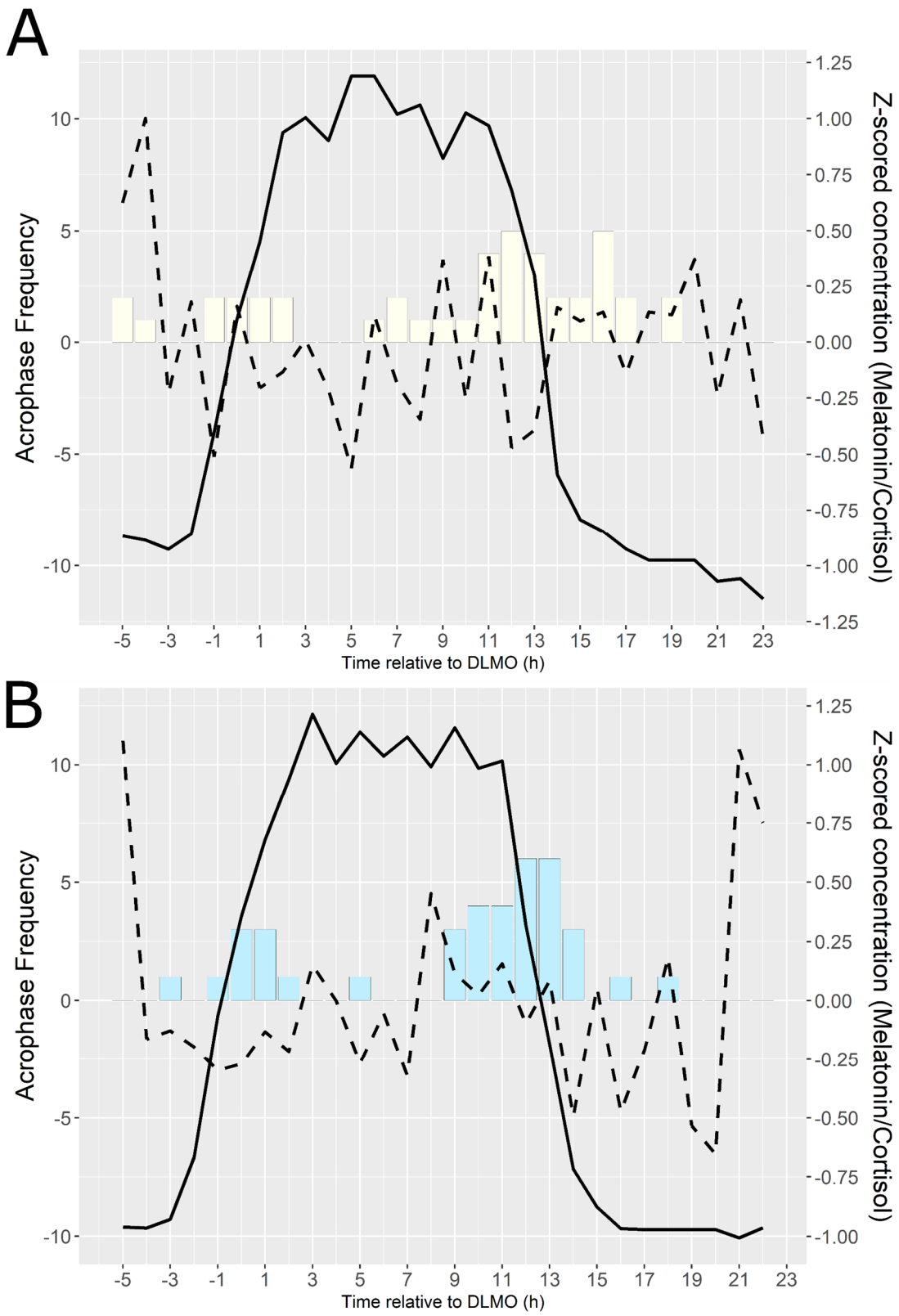


Figure 7.4: Acrophase distributions relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) Control sheep 2015, (B) HD sheep 2015. Control cohort n = 13, HD cohort n =14. For example, 5 metabolites peaked 12 h after DLMO for the control group.

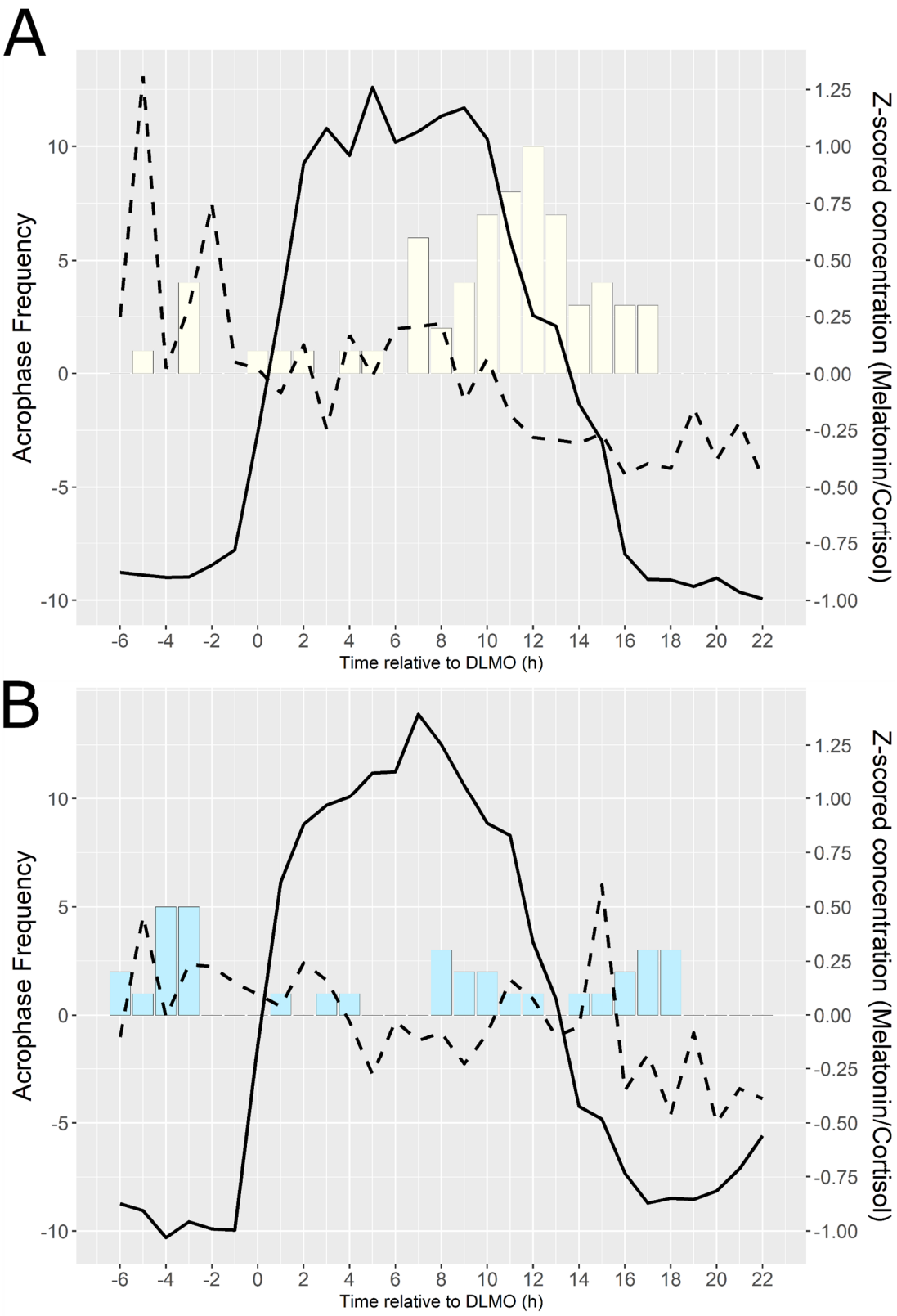


Figure 7.5: Acrophase distributions relative to DLMO overlaid with the melatonin (solid line) and cortisol (dashed line) rhythms for the (A) Control sheep 2017, (B) HD sheep 2017. Control cohort n = 14, HD cohort n =16. For example, 10 metabolites peaked 12 h after DLMO for the control group.

7.2.4. Pathway enrichment analysis

Pathway enrichment analysis for the 2015 cohort yielded two and three enriched pathways, post BH FDR correction, for the control and HD group respectively (Table 7.3). 'Aminoacyl-tRNA biosynthesis' and 'Arginine biosynthesis' were jointly enriched across both groups whilst 'Valine, leucine and isoleucine biosynthesis' was enriched only in the HD group.

Table 7.3: Pathway enrichment analysis results for rhythmic metabolites (non-lipid & lipid) in the 2015 (age five) cohort. 'Hits' refers to the number of submitted entries matched to the corresponding pathway. Only significant results (p < 0.05) post BH FDR correction are shown.

Control - 2015	Total	Hits	FDR	Impact
Aminoacyl-tRNA biosynthesis	48	7	4.28E-06	0
Arginine biosynthesis	14	3	0.00847	0.30457
HD - 2015	Total	Hits	FDR	Impact
Aminoacyl-tRNA biosynthesis	48	9	7.43E-08	0
Valine, leucine and isoleucine biosynthesis	8	4	4.61E-05	0
Arginine biosynthesis	14	3	0.014635	0.28934

The 2017 cohort yielded four and three enriched pathways, post BH FDR correction, for the control and HD group respectively (Table 7.4). Once again 'Aminoacyl-tRNA biosynthesis' was enriched in both groups whilst 'Valine, leucine and isoleucine biosynthesis', 'Arginine biosynthesis' and 'Tryptophan metabolism' were enriched only in the control groups whilst 'Phenylalanine, tyrosine and tryptophan biosynthesis' was rhythmic only in the HD group.

Table 7.4: Pathway enrichment analysis results for rhythmic metabolites (non-lipid & lipid) in the 2017 (age seven) cohort. 'Hits' refers to the number of submitted entries matched to the corresponding pathway. Only significant results (p < 0.05) post BH FDR correction are shown.

Control - 2017	Total	Hits	FDR	Impact
Aminoacyl-tRNA biosynthesis	48	8	5.87E-06	0
Valine, leucine and isoleucine biosynthesis	8	3	0.004833	0
Arginine biosynthesis	14	3	0.019873	0.28934
Tryptophan metabolism	41	4	0.038278	0.3695
HD - 2017	Total	Hits	FDR	Impact
Aminoacyl-tRNA biosynthesis	48	8	1.02E-07	0
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.018906	1

7.3. Discussion

The below discussion features excerpts, written by myself, contained in the original publication of which I am a co-author [282] while including edits and additions relating to new analysis and read literature performed by myself following the publication of the paper.

The differences between HD and control sheep across the two age groups were primarily evident in the circadian rhythm analysis of the two cohorts. While in the five year old sheep the number of rhythmic metabolites, acrophase, and rhythm shape, were highly similar the same did not hold true for the seven year old sheep. This result was most clearly observed by the outcome of the differential rhythmicity analysis which concluded that none of the metabolites tested were differentially rhythmic with the 2015 aged five cohort but that 32 metabolites were significantly differentially rhythmic in the 2017 aged seven cohort. Several metabolites, predominantly GPs in the HD group showed distinct alterations in normal

rhythms at DLMO+3 h, which corresponds to previous observations of disturbed evening behaviour in HD sheep when isolated by genotype [435]. Further, GPs have also been observed to follow circadian rhythms in humans in the untargeted lipidomics results presented (see Chapter 5). GPs almost exclusively peaked within 2 h of DLMO, indicating a high degree of temporal control of GPs in healthy humans. This observation in humans runs parallel to observations in the control sheep of either age, or the five year old HD sheep, where GPs tend to peak at or ~6 h prior to DLMO. Perturbations to GP rhythms observed in the 7 year old HD sheep are indicative of HD pathology in the ewe model, but are likely relevant to humans based on the parallel comparisons in healthy humans and the control sheep. Similar aberrations in day/night time activity have also been observed in mouse and pig models of HD and suspected in humans [435]. These aberrations in activity are believed to be linked to disruptions of the circadian system and present in other neurodegenerative diseases such as Parkinson's disease, Alzheimer's disease, and dementia [436-438]. Increased melatonin concentration in blood samples from HD sheep have been described previously, for this cohort yet acrophase did not differ relative to controls [439], despite prior evidence of delayed melatonin rhythms in HD discussed further below. Thus, the circadian changes described in the seven year old sheep here are observed despite there being no significant difference in melatonin 'phase' between control and HD sheep at either age though melatonin concentration was higher in HD sheep. This suggests that the observed differences in circadian phase described above are not directly regulated by the SCN, but rather may be regulated by aberrant behavioural cycles and/or timing of peripheral clocks. The liver is the most likely driver of this, due to its central role in lipid metabolism and lipids predominantly featuring a change in rhythm between control and HD sheep aged seven years. Prior work in HD mouse models have shown uncoupling of the liver peripheral clock with the central clock

(SCN) resulting in perturbed circadian rhythms of the liver and metabolic perturbations; thus, providing evidence that this is a possibility within the HD sheep model [440]. This concept is discussed in further detail below.

In the original publication relating to the ewe cohort presented [439], cortisol secretion was reported as significantly lower in HD ewes, but not rams, versus controls at both age groups (age five *vs.* age seven). However, within this thesis cortisol was not deemed rhythmic thus no comment on acrophase differences can be made in agreement with the original publication. As a circadian regulated hormone a cortisol rhythm would be expected in controls, though may be expected to be lost in HD, as shown in a HD mouse model [441], or exhibit increased concentration or delayed phase as observed in human HD cases [442-444]. Though not all studies report a significant change in cortisol profiles in HD participants [445] with Kalliolia *et al* suggesting the conflicting observations may relate to sampling protocols affecting results, i.e., more frequent sampling and participant arousal during sleep resulting in greater cortisol secretion. Further, Kalliolia notes that depression is a common co-morbidity of HD in humans and 30% of patients with depression have increased cortisol secretion whilst their cohort was reported as not having 'significant depression'. It is also possible that the conflicting results are in part caused by the multimodal regulation of cortisol, which exhibits both circadian and ultradian rhythms [445] and is regulated in response to stress via the hypothalamic–pituitary–adrenal (HPA) [446]. When phase advances are observed in cortisol is likely the result of perturbations to the circadian system (SCN), corroborated by observations of delayed circadian markers melatonin in humans with advanced HD [447]. As mentioned above, no such delay in melatonin was observed in the sheep model, however, a similar albeit non-significant delay in melatonin was observed [439] which may suggest the delay in melatonin/cortisol occurs at later stages in HD whilst uncoupling of peripheral clocks

such as the liver occur at earlier stages of HD. Whilst not observed in a single study it is interesting to question whether melatonin and cortisol remain phase locked in HD as they are in healthy individuals, as both have been observed to be delayed (phase advanced) in human HD cases [444, 447], or whether dysregulation of the circadian system extends to changes in the phase relationship between these two hormones. However, this assertion cannot be tested within the sheep model presented due to the aforementioned lack of a cortisol rhythm and no delay in melatonin phase; the former potentially masked by stress induced secretion of cortisol due to sampling or other factors as suggested by Kalliolia [445]. Observations with regards to elevated levels of cortisol in HD likely relate to independent changes inciting HPA hyperactivity resulting in the elevated cortisol baseline/peak observed in the human HD studies mentioned previously. However, data on the sheep (ewes) presented do not adhere to prior observations that baseline cortisol levels are elevated in HD. Being that the result is so atypical providing an explanation for the observation is difficult and may be why the original authors investigating melatonin and cortisol rhythms in the ewe cohort presented offered no further discussion on the matter themselves [439]. Hepatic dysfunction in HD, primarily observed in mouse models, have previously been reported on among other tissues. The most notable impacts appear to be progressive mitochondrial dysfunction, mirroring what occurs within neurons in HD, in addition to perturbations of the urea cycle and lactate metabolism with the latter perhaps related to the disruptions in central carbon metabolism and shift to alternate methods of energy production reported on in the Introduction (see Section 7.1) [448]. Evidence from our own results, coupled to the evidence compiled in the review by Chuang and Demontis [448] allude to dysfunction of circadian regulated metabolism in the liver. This could either be due to the breakdown of the SCNs innate ~24 h rhythm, which appears unlikely due to preservation of melatonin rhythms in all cohorts, or a

lack of communication between the central circadian system (SCN) and peripheral clocks such as the liver clock. As a peripheral clock the liver can be regulated by the SCN directly or via Zeitgeber such as feeding. A mouse model study observed the aforementioned aberrations in activity in HD mice vs wild type and were able to restore the typical daily rhythm and reverse metabolic abnormalities observed in the liver through scheduled feeding. Further, the study demonstrated, through in vitro cultures, that the liver peripheral clock still oscillates on a ~24 h cycle in HD mice. Thus, the circadian dysregulation observed in HD is likely due to uncoupling, i.e., a lack of communication between the liver clock and the SCN and not from the breakdown of their respective ~24 h rhythms. This hepatic mitochondrial dysfunction is not exclusively observed in the mouse model and has been observed in pre-symptomatic human HD cases [449]. The liver has been raised as a point of interest in HD across multiple animal models and now humans, thus, it appears highly likely that the liver as a peripheral clock and metabolically active organ is genuinely involved in HD pathology. A possible link between the liver dysfunction and HD is outlined below.

Glycerophospholipids, specifically phosphatidylcholines (PC) and lysophosphatidylcholines (lysoPC), are the most prominent metabolite classes displaying differences in rhythmicity within our data. The biological significance of these differences is not known. A potential metabolic process of interest, however, is the Land's cycle, a process which regulates interconversions between GPs and lysoGPs through deacetylation/acetylation and occurs jointly within the liver and the blood stream. Interestingly, temporal activity for specific enzymes that regulate the Land's cycle have been observed *in vitro* such as lysoPC acyltransferase (LPCAT) which regenerates PCs from lysoPCs in the presence of acyl-CoA [450]. Moreover, phospholipase A2 that is responsible for converting PCs to lysoPCs in the Land's cycle has been shown to have increased activity in leukocytes in HD patients and may

partly explain why serum concentrations of specific PCs decrease in conjunction with HD progression [451, 452]. Temporal regulation is a mechanism that can separate incompatible metabolic processes and is potentially required to some degree for optimal regulation of the Land's cycle and associated outcomes such as clearing excess lysoPCs from the blood stream and regenerating PCs [453]. It can be inferred from our data that there is some degree of temporal dysregulation in the HD sheep although its source must remain a matter for speculation but likely relates to uncoupling of peripheral clocks from the SCN. Further, it has already been shown that mHTT migrating from the brain to the liver via the bloodstream can affect gene expression of key enzymes for metabolic processes, e.g., by repressing transcription factors, [448] which likely contributes to the metabolic dysfunction alongside perturbed circadian regulation. Should phospholipase A2 activity be increased in HD patients, as previously observed, and/or temporal dysregulation of other components of the Land's cycle occurs, a potential consequence may be diminished circulating PCs and excessive circulating lysoPCs. Circulating lysoPCs have numerous roles and effects on different cell types, such as promoting inflammation and inducing demyelination both of which are associated with HD pathophysiology [452, 454]. In support of this hypothesis all seven lysoPC with significantly different rhythms ('change') between 7 year old control and HD sheep had a higher midline estimating statistic of rhythm (MESOR) in HD ranging from 8.6%-24.6%. The MESOR, calculated as part of the applied cosinor rhythm analysis, indicates an elevated base level of lysoPCs in the HD group vs controls. However, the same is also true for seven of the 10 PCs classified as 'change' with an increase in MESOR of 0.95-10.2% for HD compared to control; perhaps suggesting additional or more nuanced changes to the Lands cycle as outlined above. Of the aforementioned 7 lysoPCs, only three were rhythmic in the 5 year old cohort (lysoPC a C16:1, lysoPC a C18:1, lysoPC a C18:2), none were classified as significantly

different and all had a higher MESOR in the HD group vs controls (0.8-7.3%) whilst 6 of the 10 aforementioned PCs were rhythmic in the 5 year old cohort (PC aa C38:4, PC aa C38:6, PC aa C40:5, PC ae C36:1, PC ae C36:4, PC ae C42:3) none were classified as significantly different and five had a higher MESOR in the HD group vs controls (3.7-16.0%). Moving forward, the Land's cycle, and its potential temporal regulation of phospholipids, should be considered a metabolic pathway of interest in relation to HD progression.

The notable outcomes of this study were not limited to lipids, with some interesting data yielded from differential rhythmic analysis and pathway enrichment analysis of polar metabolites. Tryptophan metabolism was enriched in the 2017 control cohort which, as mentioned in the introduction, comprises the kynurenine pathway whose products have a profound effect on the progression of HD. Yet this pathway was not enriched among rhythmic metabolites in the 2017 HD cohort, nor the 2015 HD or control cohorts. However, of the 41 metabolites that comprise the KEGG 'Tryptophan metabolism' reference pathway only three were covered by targeted assay applied (tryptophan, serotonin, kynurenine) with melatonin measured separately. Thus, three of these four metabolites would have needed to be classified as rhythmic in any individual group, as was the case in 2017 control group, to meet the specified cutoff of being considered an enriched pathway. As the coverage of any specific metabolic pathway by the targeted assay applied is limited so too is the power of the pathway enrichment analysis. Not observing tryptophan metabolism enriched within the 2015 cohort may discourage the kynurenine pathway as a line of enquiry if not for the fact that kynurenine was classified as having a significantly altered rhythm between the 2017 control and HD cohorts. Whilst the evidence cannot be fully provided within this study, it is worth speculating that the reported disturbances in the kynurenine pathway within HD may be due to circadian disruption or loss of rhythm of metabolites that comprise the pathway. This may be due to

loss of rhythm of precursor metabolites, although tryptophan is such a precursor metabolite and was rhythmic in both HD and control groups, or loss of rhythm for enzymes bottle necking specific metabolite transitions potentially favouring some reactions such as the production of quinolinic acid whilst hampering others such as kynurenic acid. Excessive levels of the former and lowered levels of the latter were associated with HD progression as detailed in the introduction to this chapter (see Chapter 7.1). Other polar metabolites of note that lost rhythmicity in the 2017 HD cohort vs controls include lysine and alpha-Aminoadipic acid, both of which can be found in the lysine degradation pathway which, among multiple functions, produce precursors for the TCA (citric acid) cycle such as acetyl CoA. Within the introduction of this chapter (see Section 7.1) it was stated that central carbon metabolism (TCA cycle, glycolysis, pyruvate metabolism) are believed to be perturbed in HD. The loss of rhythm of lysine and alpha-Aminoadipic acid rhythms potentially contributes to these prior observations. It is worth noting however that the applied targeted panel did not contain the vast majority of metabolites that comprise these pathways, thus, the loss of lysine and alpha-Aminoadipic acid is not necessarily representative of system level alterations in metabolism. Further to this limitation, histidine and threonine, were rhythmic in HD but not control sheep for the 2017 cohort yet this gain in rhythm is not easily explained for two metabolites which are not closely related. Larger targeted panels or bespoke targeted assays for pathways of interest, such as those named above, should be considered in any follow up studies to (1) confirm the findings discussed and (2) expand on the limited pathway analysis afforded by the applied assay within this study. If possible, greater consideration should be given in advance with regards to mitigating batch affects to allow for longitudinal comparisons which were not possible in the current study. Further, the animal model in question has an expected lifespan of ~10 years, sampling at nine years of age, in addition to five and seven years, would

likely benefit future analysis by observing cases of HD in a more advanced stage. As surmised in the original publication of this chapter [282] it may also benefit follow up studies to include a mixed sex cohort of ewes and rams to identify sex dependent differences, if any, as recent reviews highlight conflicting reports on how sex may affect the severity of HD phenotype or onset of co-morbidities such as depression [455, 456]. As discussed in previous chapters, notably Chapter 1 and the discussion of Chapter 5, the mammalian circadian system and metabolism exhibit sexual dimorphism with sex also influencing brain development [456]. As circadian regulation of metabolism appears highly relevant to HD progression any factors which affect either system, such as sex, should be considered an avenue of further investigation.

7.4. Chapter summary

Taken as a whole the results suggest that the acrophase of the metabolite rhythms do not appear to shift by a large margin, as so few metabolites exhibited >2 h changes in acrophase in control vs HD comparisons. However, the rhythms do appear perturbed as visually illustrated in the heatmaps and inferred from the differential rhythmicity analysis which identified 22 metabolites as 'change' between 2017 control and HD sheep. The sheep at both age ranges are considered pre-symptomatic and is likely why these changes are so subtle and may in part explain why unsupervised analysis ignoring the time domain (e.g., PCA) offers limited insight into the apparent changes that have occurred between control and HD sheep at seven years of age. They do not exhibit large fold changes in metabolite concentrations nor complete ablation of circadian rhythms, at least not yet, but rather smaller alterations in the rhythm shape and in some instances acrophase. Such results would likely have been overlooked if only applying more standardised statistical analyses typically employed for

single time point sampling. This study therefore acts as an excellent example of interdisciplinary research that embodies the spirit of this entire thesis. By combining the power of metabolomics platforms with the bespoke sampling and time series analytics of circadian biologists new evidence has come to light with regards to HD progression.

Thesis Summary

The primary focus of the PhD research presented in this thesis has been on the detection and measurement of metabolite rhythms under the prevailing theme of circadian regulation of metabolism and how this regulation relates to various phenotypes such as health, disease, and sex. To fulfil this overarching theme original research was performed in the form of: a literature review on the emergent field of metabolomics applied in human circadian studies; application of untargeted UHPLC-MS metabolomics assays (lipidomics/HILIC) and a bespoke rhythm analysis workflow to characterise metabolite rhythms in a young healthy population in addition to identifying sex dependent differences; application of a targeted assay (Biocrates AbsoluteIDQ P180) to characterise metabolite rhythms in a sheep model of Huntington's disease at different ages.

The literature review, detailed in Chapter 3, was the first review to collate results and provide commentary for all studies applying metabolomics to investigate time of day variation (diurnal/circadian rhythms) of metabolites in humans. The review highlighted gaps in the literature detailed within this thesis, such as minimal application of HILIC assays and a lack of female representation in cohorts for this research area; in addition, it also highlighted metabolites that were consistently reported to be rhythmic across a broad range of experimental designs (Table A5). Identifying robustly rhythmic metabolites is of value to the field due to a growing need to accurately estimate internal biological time (circadian phase) within an individual for the purposes of chronotherapy. Chronotherapy is a broad term describing treatments designed to modulate the circadian system to treat or minimise the symptoms of circadian related health issues e.g., light therapy [466]. Additionally, it also refers to the timing of treatments such as drug/chemotherapy relative to an individual's circadian

phase that maximises treatment efficacy whilst minimising toxicity; such practices may also be described under the terms chronopharmacology and chronochemotherapy [466-468]. Thus, utilising chronotherapy as one facet of personalised medicine is dependent on our ability to accurately measure the circadian phase of an individual to facilitate optimal timing of treatment or modulation of circadian phase. As discussed within this thesis the current gold standard phase marker is melatonin to determine DLMO, which is reliant on multiple samples taken over a time course under dim light conditions. This methodology is considered suboptimal in a clinical setting [469]; yet, despite this suboptimal method and in the absence of a better phase marker DLMO is still recommended as a diagnostic tool for circadian rhythm sleep disorders [470]. Thus, single time point prediction of an individual's phase is considered a major stepping stone towards the application of chronotherapy in a real word setting. As detailed by Wittenbrink et al [469], the main benefits of single sampling are a reduction of (invasive) sample collection coupled to minimised time and cost of processing such samples resulting in a more efficient and therefore optimised method for clinical practice versus DLMO measurements. However, the challenge in single time point sampling is identifying and profiling robust phase markers (e.g., metabolites, transcripts) that are rhythmic (~24 h period) across a broad range of individuals. Further, phase markers would ideally differ in phase to one another but otherwise remain phase locked and exhibit similar phase, amplitude, and MESOR across individuals [319, 469, 472]. Rhythmic profiles of the phase markers acquired across a time series, from which acrophase, amplitude and MESOR could be measured, can be used as a point of reference to construct a molecular timetable [319, 471], or implemented into a machine learning algorithm as a training set [278, 469, 472, 473], with the intent of predicting biological time from minimal/single 'test' samples by either method. Robust phase markers would therefore need to be observable across a range of phenotypes (age, sex,

ethnicity, health/disease status) and under a range of environmental conditions (entrained – real world, constant routine). Therefore, the aforementioned results from the review and metabolites listed Table A5 are relevant in this regard. Furthermore, the results from Chapters 5/6 included group level rhythm analysis between male and female groups to determine sex dependent differences alongside individual level rhythm analysis. Subsequently, the outputs of this thesis are well placed to comment on rhythmic metabolites which are observed across multiple studies, similar between sexes, and observable across individuals, in turn highlighting candidate phase marker metabolites.

As shown and partially discussed in Chapters 5/6 metabolites from Table A5 such as ornithine, citrulline, tryptophan, LysoPC(16:0), LysoPC(17:0), PC(34:2), ST(21:2;O2;S) (potentially pregnenolone sulphate), ST(21:4;O5) (potentially cortisone) were annotated in the metabolomics data presented and were rhythmic in both male and female groups. There are more examples of rhythmic metabolite features in male and female groups which may correspond to the list in Table A5 (e.g., LysoPC(18:1), LysoPC(18:2), PC(32:1)) but such features have multiple annotations reducing the confidence in their identification and once again highlighting the challenges of metabolite annotation. The metabolites proline, leucine, and isoleucine are some of the most frequently reported rhythmic metabolites but were not successfully annotated and observed in Chapters 5/6, however the dipeptides isoleucylproline/leucylproline (annotated in tandem) were rhythmic in male and female groups. The aforementioned metabolites also exhibit a highly similar acrophase between male and female groups. Tryptophan, ornithine, citrulline, LysoPC(16:0), LysoPC(17:0), PC(34:2), isoleucylproline/leucylproline, ST(21:2;O2;S), ST(21:4;O5) all exhibited an acrophase difference between males and females ranging from 0.1 – 2 h, as determined by cosinor analysis. Therefore, the predicted acrophase difference for these metabolites was less

than the sampling resolution applied for the metabolomics analysis performed (2 h) indicative of the highly similar phase for these metabolites between male and female groups. However, the results do contain some discrepancies. One of two metabolite features annotated as ST(21:2;O2;S) was classified as having a 'change' in profile and possessed an acrophase difference between male and female groups; contrary to the cosinor acrophase estimation indicating the second metabolite feature annotated as ST(21:2;O2;S) exhibited no phase difference between groups. It is possible both features refer to different metabolites that are structural isomers, which could explain the difference in observations, but a more precise annotation cannot be obtained. Similar discrepancies are observed at the individual level where multiple metabolite features with the same annotation exhibit varying degrees of rhythmicity across the cohort. Within the individual level analysis ST(21:4;O5) was considered rhythmic in 17 of the 23 participants, ST(21:2;O2;S) rhythmic in 11 participants, isoleucylproline/leucylproline rhythmic in six participants, and tryptophan rhythmic in four participants. Prior to FDR correction tryptophan was rhythmic in 16 participants, followed by ornithine (pre-FDR) rhythmic in ten participants, PC(34:2) (pre-FDR) rhythmic in seven participants, and citrulline (pre-FDR) rhythmic in five participants. However, there were additional metabolite features annotated as ST(21:2;O2;S), isoleucylproline/leucylproline, tryptophan, ornithine, PC(34:2), and citrulline which presented as rhythmic in fewer individuals than the values detailed above. The most frequently observed rhythmic metabolites reported in Chapter 3 and observed in Chapters 5/6 exhibited some of the desirable characteristics of phase markers. The metabolites discussed in the previous paragraph, at the group level, possess similar acrophase between similarly aged healthy males and females. Furthermore, several of these metabolites are observed as rhythmic at the individual level, to varying degrees, despite the limitations and challenges of individual level

rhythm analysis discussed earlier in the thesis. Finally, these metabolites have been reported as rhythmic across multiple studies which include constant routine (circadian) and entrained (diurnal) studies, thus, we may expect their rhythms are robust across a range of environmental/behavioural factors (e.g., light/dark, sleep/wake). The metabolites listed in Table A5, such as the examples discussed above, are therefore ideal candidate phase markers pending further validation. Some of which, namely tryptophan, leucine, and cortisone, have already been used prior to assemble a molecular timetable [319].

Chapter 5 discussed some of the variability between males and females with regards to the magnitude and phase of rhythmic lipids, however, GPs almost exclusively peaked within 2 h after DLMO, coinciding with melatonin's peak, whilst SPs almost exclusively peaked in the subjective morning between 6-10 h after DLMO though acrophase for SPs tended to be later in the female group than the male group. Whilst not limited to a brief window GLs were partitioned into peaking in the subjective morning (4-10 h post DLMO) or evening (4-6 h prior to DLMO). These three lipid classes (GP, SP, GL) display strong temporal regulation, under constant routine at the group level, so should be considered metabolite classes of interest for candidate phase markers. Particularly if future studies can demonstrate that this strong temporal regulation of the different classes is observable under entrained (diurnal) conditions. A previous study partially corroborates the above observations and reported on GLs peaking both in the morning and the afternoon, though the same study could not discern a general time of day in which GPs peaked and observed SPs peaking predominately in the afternoon [320]. Despite this promise at the group level there is some difficulty observing such metabolites on the individual level; within this thesis no GP/SP/GL was rhythmic, post FDR, in more than five/two/five participants, respectively. Moreover, in many cases rhythmic metabolite features are only annotated confidently to the lipid class and not a single lipid

species. Whilst prior work has shown some lipids to be rhythmic and share a similar phase across individuals, such as TG(48:0), others present high variability or are even in antiphase between individuals such as SM(18:1/24:1) [320]. Despite these difficulties lipids may still prove a valuable source of phase markers beyond the frequently reported rhythms in non-lipid metabolites such as amino acids. Furthermore, a selection of lipids observed as rhythmic in this study have previously been observed in prior works as listed in Table 5.3. Thus, lipids remain a promising point of interest for the field due to their strong temporal regulation which illustrates their potential to act as metabolite phase markers and the role of circadian regulation with regards to lipid metabolism.

Future work should focus on validating the results of the untargeted assays presented, more specifically, via the application of (semi)-targeted assays focussing on one or more of the aforementioned lipid classes of interest. Such assays would provide quantitative information in the form of metabolite concentration over the time course and facilitate more accurate annotation. This in turn would produce more accurate rhythmic lipid profiles which may then be tested for use as phase markers, creating a preliminary catalogue of phase marker metabolites. Validation of the phase markers should include observing lipid rhythms under diurnal conditions to observe the degree in which the strong temporal regulation of lipid rhythms, initially observed under constant routine, persists under entrained conditions. Alternatively, focussing on lipid classes of interest such as GLs and FAs, that comprise the TG/FFA cycle, would yield further insight into metabolic pathways that appear strongly regulated by the circadian system. It would be of interest to see how such pathways, rhythmic in healthy participants under constant routine, behave under diurnal conditions or within shift workers. For example, should the TG/FFA cycle be perturbed within shift workers it would provide evidence for how lipid metabolism is perturbed due to shiftwork and develop our

understanding of the metabolic disruption shift work can cause. Chapter 7 is one such example of how a (semi)-targeted assay, with its quantitative data and known metabolite identities, greatly benefits the interpretation of rhythm analysis data and expands our understanding of metabolic changes in a health versus disease study design.

The combination of a large mammal model (OVT73), targeted metabolomics and the bespoke rhythm analysis workflow, developed in Chapter 4 and applied to data presented in Chapters 5-7, allowed for novel insight into the circadian regulation of the metabolome in pre-symptomatic HD vs controls. Most striking about these results was the lack of significant differences in metabolite rhythms between controls and HD sheep at age five, according to differential rhythmicity analysis, yet the stark difference in metabolite rhythms observed in the cohort at age seven. Chapter 7 therefore illustrates the subtle changes to the circadian metabolome in pre-symptomatic HD with the progression of age and acts as an excellent example of the advantages of time series sampling vs single sampling. Furthermore, it was hypothesised how the observed perturbations to lysoPCs and PCs may relate to the Land's cycle and how dysregulation of this cycle could result in increased regulation of lysoPCS. Differentially rhythmic lysoPCs all exhibited a higher MESOR in HD vs control groups at age seven, supporting this statement, moreover, elevated levels of circulating lysoPCs are associated with promoting inflammation and inducing demyelination both of which are associated with HD pathophysiology. Thus, Chapter 7 has established how circadian regulated metabolites differ between HD and healthy controls, in a large mammal model, and identified a metabolic pathway of interest that may be involved with HD pathology. Whilst the study benefited from quantitative data acquired from the applied targeted assay, its coverage was limited to a theoretical maximum of ~180 metabolites. Future metabolomics work on the OVT73 model may benefit from broader targeted assays to identify other metabolite classes

of interest with regards to HD pathology or bespoke targeted assays focussing on phospholipids, a lipid class of interest based on the data presented in Chapter 7.

To conclude, this thesis has collated prior knowledge in the research area of circadian regulated metabolism and acted on that foundational knowledge base having: (1) performed the broadest untargeted metabolomic assays on the largest balanced mix sex cohort sampled under constant routine to date; in turn compiling the largest dataset of annotated circadian regulated metabolites; (2) characterised sex dependent differences in circadian metabolite rhythms in healthy humans which had not previously been performed; (3) proposed metabolite phase markers and demonstrated that some frequently observed rhythmic metabolites have similar acrophase between males and females and are observable at the individual level; (4) expanded our understanding of how circadian regulated metabolism is affected by onset of HD in a sheep model; (5) developed an R workflow to facilitate more efficient data handling and analysis of time series metabolomics data.

REFERENCE LIST

[1] Takahashi, J. S. (2017). "Transcriptional architecture of the mammalian circadian clock", Nature Reviews Genetics, Vol. 18, Issue 3, pp. 164–179. DOI: 10.1038/nrg.2016.150.

[2] Lewy, A. J. and Sack, R. L. (1989). "The dim light melatonin onset as a marker for circadian phase position", Chronobiology International, Vol. 6, Issue 1, pp. 93–102. DOI: 10.3109/07420528909059144.

[3] Czeisler, C. A., Weitzman, E. D., Moore-Ede, M. C., Zimmerman, J. C. and Knauer, R. S. (1980). "Human sleep: Its duration and organization depend on its circadian phase", Science, Vol. 210, Issue 4475, pp. 1264–1267. DOI: 10.1126/SCIENCE.7434029.

[4] Strogatz, S. H., Kronauer, R. E., Czeisler, T. A. and Strogatz, S. H. (1986). "Circadian Regulation Dominates Homeostatic Control of Sleep Length and Prior Wake Length in Humans", Sleep, Vol. 9, Issue 2, pp. 353-364. DOI: 10.1093/sleep/9.2.353.

[5] Aschoff, J. (1965). "Circadian rhythms in man", Science, Vol. 148, Issue 3676, pp. 1427–1432. DOI: 10.1126/science.148.3676.1427.

[6] Qian, J., Dalla Man, C., Morris, C. J., Cobelli, C. and Scheer, F. A. J. L. (2018). "Differential effects of the circadian system and circadian misalignment on insulin sensitivity and insulin secretion in humans", Diabetes, obesity & metabolism, Vol. 20, Issue 10, pp. 2481–2485. DOI: 10.1111/DOM.13391.

[7] Vaze, K. M. and Sharma, V. K. (2013). "On the Adaptive Significance of Circadian Clocks for Their Owners", Chronobiology International, Vol. 30, Issue 4, pp. 413–433. DOI: 10.3109/07420528.2012.754457.

[8] Cox, K. H. and Takahashi, J. S. (2019). "Circadian clock genes and the transcriptional architecture of the clock mechanism", Journal of Molecular Endocrinology, Vol. 63, Issue 4, pp. 93–102. DOI: 10.1530/JME-19-0153.

[9] Pittendrigh, C. S. (1960). "Circadian rhythms and the circadian organization of living systems.", 1960. doi: 10.1101/SQB.1960.025.01.015.

[10] Czeisler, C. A., Duffy, J. F., Shanahan, T. L., Brown, E. N., Mitchell, J. F., et al. (1999). "Stability, precision, and near-24-hour period of the human circadian pacemaker", Science, Vol. 284, Issue 5423, pp. 2177–2181. DOI: 10.1126/SCIENCE.284.5423.2177.

[11] Mohawk, J. A., Green, C. B. and Takahashi, J. S. (2012). "Central and peripheral circadian clocks in mammals", Annual Review of Neuroscience, Vol. 35, pp. 445–462. DOI: 10.1146/annurev-neuro-060909-153128.

[12] Hankins, M. W., Peirson, S. N. and Foster, R. G. (2008). "Melanopsin: an exciting photopigment", Trends in Neurosciences, Vol. 31, Issue 1, pp. 27–36. DOI: 10.1016/j.tins.2007.11.002.

[13] Huang, W., Ramsey, K. M., Marcheva, B. and Bass, J. (2011). "Circadian rhythms, sleep, and metabolism", Journal of Clinical Investigation, Vol. 121, Issue 6, pp. 2133–2141. DOI: 10.1172/JCI46043.

[14] Dyar, K. A., Lutter, D., Artati, A., Ceglia, N. J., Liu, Y., et al. (2018). "Atlas of Circadian Metabolism Reveals System-wide Coordination and Communication between Clocks", Cell, Vol. 174, Issue 6, pp. 1571-1585. DOI: 10.1016/j.cell.2018.08.042.

[15] Hughes, S., Jagannath, A., Hankins, M. W., Foster, R. G. and Peirson, S. N. (2015). "Photic regulation of clock systems", Methods in Enzymology, Vol. 552, pp. 125–143. DOI: 10.1016/BS.MIE.2014.10.018.

[16] Chowdhury, D., Wang, C., Lu, A. P. and Zhu, H. L. (2019). "Understanding quantitative circadian regulations are crucial towards advancing chronotherapy", Cells, Vol. 8, Issue 8. DOI: 10.3390/cells8080883.

[17] Chung, S., Son, G. H. and Kim, K. (2011). "Circadian rhythm of adrenal glucocorticoid: its regulation and clinical implications", Biochimica et biophysica acta, Vol. 1812, Issue 5, pp. 581–591. DOI: 10.1016/J.BBADIS.2011.02.003.

[18] Bünning, E. (1974). "The physiological clock. Circadian rhythms and biological chronometry", Food and Cosmetics Toxicology, Vol. 12, Issue 5–6. DOI: 10.1016/0015-6264(74)90267-3.

[19] Aschoff, J. (1981). "Freerunning and Entrained Circadian Rhythms", in Biological Rhythms, Springer US, 1981, pp. 81–93. doi: 10.1007/978-1-4615-6552-9_6.

[20] Weaver, D. R. (1998). "The Suprachiasmatic Nucleus: A 25-Year Retrospective", Journal of Biological Rhythms, Vol. 13, Issue 2, pp. 100–112. DOI: 10.1177/074873098128999952.

[21] Moore, R. Y. (1973). "Retinohypothalamic projection in mammals: a comparative study", Brain Research, Vol. 49, Issue 2, pp. 403–409. DOI: 10.1016/0006-8993(73)90431-9.

[22] Moore, R. Y. and Lenn, N. J. (1972). "A retinohypothalamic projection in the rat", Journal of Comparative Neurology, Vol. 146, Issue 1, pp. 1–14. DOI: 10.1002/CNE.901460102.

[23] Hendrickson, A. E., Wagoner, N. and Cowan, W. M. (1972). "An autoradiographic and electron microscopic study of retino-hypothalamic connections", Zeitschrift für Zellforschung und mikroskopische Anatomie, Vol. 135, Issue 1, pp. 1–26. DOI: 10.1007/BF00307084.

[24] Stephan, F. K. and Zucker, I. (1972). "Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions", Proceedings of the National Academy of Sciences of the United States of America, Vol. 69, Issue 6, pp. 1583–1586. DOI: 10.1073/PNAS.69.6.1583.

[25] Moore, R. Y. and Eichler, V. B. (1972). "Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat", Brain Research, Vol. 42, Issue 1, pp. 201–206. DOI: 10.1016/0006-8993(72)90054-6.

[26] Rusak, B. and Zucker, I. (1979). "Neural regulation of circadian rhythms", Physiological Reviews, Vol. 59, Issue 3, pp. 449–526. DOI: 10.1152/PHYSREV.1979.59.3.449.

[27] Inouye, S. T. and Kawamura, H. (1979). "Persistence of circadian rhythmicity in a mammalian hypothalamic 'island' containing the suprachiasmatic nucleus", Proceedings of the National Academy of Sciences of the United States of America, Vol. 76, Issue 11, pp. 5962–5966. DOI: 10.1073/PNAS.76.11.5962.

[28] Green, D. J. and Gillette, R. (1982). "Circadian rhythm of firing rate recorded from single cells in the rat suprachiasmatic brain slice", Brain Research, Vol. 245, Issue 1, pp. 198–200. DOI: 10.1016/0006-8993(82)90361-4.

[29] Groos, G. and Hendriks, J. (1982). "Circadian rhythms in electrical discharge of rat suprachiasmatic neurones recorded in vitro", Neuroscience Letters, Vol. 34, Issue 3, pp. 283–288. DOI: 10.1016/0304-3940(82)90189-6.

[30] Shibata, S., Oomura, Y., Kita, H. and Hattori, K. (1982). "Circadian rhythmic changes of neuronal activity in the suprachiasmatic nucleus of the rat hypothalamic slice", Brain Research, Vol. 247, Issue 1, pp. 154–158. DOI: 10.1016/0006-8993(82)91041-1.

[31] Lehman, M. N., Silver, R., Gladstone, W. R., Kahn, R. M., Gibson, M., et al. (1987). "Circadian rhythmicity restored by neural transplant. Immunocytochemical characterization of the graft and its integration with the host brain", Journal of Neuroscience, Vol. 7, Issue 6, pp. 1626–1638. DOI: 10.1523/JNEUROSCI.07-06-01626.1987.

[32] Ralph, M. R., Foster, R. G., Davis, F. C. and Menaker, M. (1990). "Transplanted suprachiasmatic nucleus determines circadian period", Science, Vol. 247, Issue 4945, pp. 975–978. DOI: 10.1126/SCIENCE.2305266.

[33] Welsh, D. K., Logothetis, D. E., Meister, M. and Reppert, S. M. (1995). "Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms", Neuron, Vol. 14, Issue 4, pp. 697–706. DOI: 10.1016/0896-6273(95)90214-7.

[34] Lerner, A. B., Case, J. D., Takahashi, Y., Lee, T. H. and Mori, W. (1958). "Isolation of melatonin, the pineal gland factor that lightens melanocytes", Journal of the American Chemical Society, Vol. 80, Issue 10, p. 2587. DOI: 10.1021/JA01543A060.

[35] Wurtman, R. J., Axelrod, J. and Phillips, L. S. (1963). "Melatonin synthesis in the pineal gland: Control by light", Science, Vol. 142, Issue 3595, pp. 1071–1073. DOI: 10.1126/SCIENCE.142.3595.1071.

[36] Snyder, S. H., Zweig, M. and Axelrod, J. (1964). "Control of the circadian rhythm in serotonin content of the rat pineal gland", Life Sciences, Vol. 3, Issue 10, pp. 1175–1179. DOI: 10.1016/0024-3205(64)90137-7.

[37] Quay, W. B. (1964). "Circadian and Estrous Rhythms in Pineal Melatonin and 5-Hydroxy Indole-3-Acetic Acid", Proceedings of the Society for Experimental Biology and Medicine, Vol. 115, Issue 3, pp. 710–713. DOI: 10.3181/00379727-115-29014.

[38] Lewy, A. J., Wehr, T. A., Goodwin, F. K., Newsome, D. A. and Markey, S. P. (1980). "Light suppresses melatonin secretion in humans", Science, Vol. 210, Issue 4475, pp. 1267–1269. DOI: 10.1126/SCIENCE.7434030.

[39] Moore, R. Y. and Klein, D. C. (1974). "Visual pathways and the central neural control of a circadian rhythm in pineal serotonin N-acetyltransferase activity", Brain Research, Vol. 71, Issue 1, pp. 17–33. DOI: 10.1016/0006-8993(74)90188-7.

[40] Klein, D. C. and Moore, R. Y. (1979). "Pineal N-acetyltransferase and hydroxyindole-O-methyl-transferase: control by the retinohypothalamic tract and the suprachiasmatic nucleus", Brain Research, Vol. 174, Issue 2, pp. 245–262. DOI: 10.1016/0006-8993(79)90848-5.

[41] Cardinali, D. P., Vacas, M. I. and Boyer, E. E. (1979). "Specific Binding of Melatonin in Bovine Brain", Endocrinology, Vol. 105, Issue 2, pp. 437–441. DOI: 10.1210/ENDO-105-2-437.

[42] Glass, J. D. and Lynch, G. R. (1981). "Melatonin: Identification of Sites of Antigonadal Action in Mouse Brain", Science, Vol. 214, Issue 4522, pp. 821–823. DOI: 10.1126/SCIENCE.7292016.

[43] Vaněček, J., Pavlík, A. and Illnerová, H. (1987). "Hypothalamic melatonin receptor sites revealed by autoradiography", Brain Research, Vol. 435, Issue 1–2, pp. 359–362. DOI: 10.1016/0006-8993(87)91625-8.

[44] Laudon, M., Nir, I. and Zisapel, N. (1988). "Melatonin Receptors in Discrete Brain Areas of the Male Rat", Neuroendocrinology, Vol. 48, Issue 6, pp. 577–583. DOI: 10.1159/000125066.

[45] Redman, J., Armstrong, S. and Ng, K. T. (1983). "Free-running activity rhythms in the rat: Entrainment by melatonin", Science, Vol. 219, Issue 4588, pp. 1089–1091. DOI: 10.1126/SCIENCE.6823571.

[46] Middleton, B., Arendt, J. and Stone, B. M. (1997). "Complex Effects of Melatonin on Human Circadian Rhythms in Constant Dim Light", Journal of Biological Rhythms, Vol. 12, Issue 5, pp. 467–477. DOI: 10.1177/074873049701200508.

[47] Sack, R. L., Brandes, R. W., Kendall, A. R. and Lewy, A. J. (2000). "Entrainment of Free-Running Circadian Rhythms by Melatonin in Blind People", New England Journal of Medicine, Vol. 343, Issue 15, pp. 1070–1077. DOI: 10.1056/nejm200010123431503.

[48] Lockley, S. W., Skene, D. J., James, K., Thapan, K., Wright, J., et al. (2000). "Melatonin administration can entrain the free-running circadian system of blind subjects", The Journal of endocrinology, Vol. 164, Issue 1. DOI: 10.1677/JOE.0.164R001.

[49] Dubocovich, M. L. (2007). "Melatonin receptors: role on sleep and circadian rhythm regulation", Sleep medicine, Vol. 8 Suppl 3, pp. 34–42. DOI: 10.1016/J.SLEEP.2007.10.007.

[50] Mason, H. L., Hoehn, W. M., McKenzie, B. F. and Kendall, E. C. (1937). "Chemical studies of the suprarenal cortex III. The structures of compounds A, B, and H", J. Biol. Chem., Vol. 120, Issue 2, pp. 719–741. DOI: 10.1016/S0021-9258(18)45132-0

[51] Migeon, C. J., Tyler, F. H., Mahoney, J. P., Florentin, A. A., Castle, H., et al. (1956). "the diurnal variation of plasma levels and urinary excretion of 17-hydroxycorticosteroids in normal subjects, night workers and blind subjects*†", The Journal of Clinical Endocrinology & Metabolism, Vol. 16, Issue 5. DOI: 10.1210/jcem-16-5-622.

[52] Perkoff, G. T., Eik-nes, K., Nugent, C. A., Fred, H. L., Nimer, R. A., et al. (1959). "Studies of the diurnal variation of plasma 17-hydroxycorticosteroids in man.", The Journal of clinical endocrinology and metabolism, Vol. 19, Issue 4. DOI: 10.1210/jcem-19-4-432.

[53] Doe, R. P., Vennes, J. A., Flink, E. B., Engst, R., Johnson, P., et al. (1960). "Diurnal variation of 17-hydroxycorticosteroids, sodium, potassium, magnesium and creatinine in normal subjects and in cases of treated adrenal insufficiency and cushing's syndrome*†", The Journal of Clinical Endocrinology & Metabolism, Vol. 20, Issue 2. DOI: 10.1210/jcem-20-2-253.

[54] Orth, D. N. and Island, D. P. (1969). "Light Synchronization of the Circadian Rhythm in Plasma Cortisol (17-OHCS) Concentration in Man", The Journal of Clinical Endocrinology & Metabolism, Vol. 29, Issue 4, pp. 479–486. DOI: 10.1210/JCEM-29-4-479.

[55] Moberg, G. P., Scapagnini, U., de Groot, J. and Ganong, W. F. (1971). "Effect of Sectioning the Fornix on Diurnal Fluctuation in Plasma Corticosterone Levels in the Rat", Neuroendocrinology, Vol. 7, Issue 1, pp. 11–15. DOI: 10.1159/000121950.

[56] Roden, M., Koller, M., Pirich, K., Vierhapper, H. and Waldhauser, F. (1993). "The circadian melatonin and cortisol secretion pattern in permanent night shift workers", The American journal of physiology, Vol. 265, Issue 1 Pt 2. DOI: 10.1152/AJPREGU.1993.265.1.R261.

[57] Weibel, L. and Brandenberger, G. (2002). "The start of the quiescent period of cortisol remains phase locked to the melatonin onset despite circadian phase alterations in humans working the night schedule", Neuroscience Letters, Vol. 318, Issue 2, pp. 89–92. DOI: 10.1016/S0304-3940(01)02496-X.

[58] Kennaway, D. J. and Royles, P. (1986). "Circadian rhythms of 6-sulphatoxy melatonin, cortisol and electrolyte excretion at the summer and winter solstices in normal men and women", Acta Endocrinologica (Norway), Vol. 113, Issue 3, pp. 450–456. DOI: 10.1530/ACTA.0.1130450.

[59] Kennaway, D. J. and Van Dorp, C. F. (1991). "Free-running rhythms of melatonin, cortisol, electrolytes, and sleep in humans in Antarctica", https://doi.org/10.1152/ajpregu.1991.260.6.R1137, Vol. 260, Issue 6 29-6. DOI: 10.1152/AJPREGU.1991.260.6.R1137.

[60] Skene, D. J., Lockley, S. W., James, K. and Arendt, J. (1999). "Correlation between urinary cortisol and 6-sulphatoxymelatonin rhythms in field studies of blind subjects", Clinical Endocrinology, Vol. 50, Issue 6, pp. 715–719. DOI: 10.1046/J.1365-2265.1999.00714.X.

[61] Kripke, D. F., Youngstedt, S. D., Elliott, J. A., Tuunainen, A., Rex, K. M., et al. (2009). "Circadian Phase in Adults of Contrasting Ages", http://dx.doi.org/10.1080/07420520500180439, Vol. 22, Issue 4, pp. 695–709. DOI: 10.1080/07420520500180439.

[62] Buckley, T. M. and Schatzberg, A. F. (2010). "A pilot study of the phase angle between cortisol and melatonin in major depression – A potential biomarker?", Journal of Psychiatric Research, Vol. 44, Issue 2, pp. 69–74. DOI: 10.1016/J.JPSYCHIRES.2009.06.012.

[63] Van Puyvelde, M., Rietjens, G., Helmhout, P., Mairesse, O., Van Cutsem, J., et al. (2022). "The submariners' sleep study: a field investigation of sleep and circadian hormones during a 67-day submarine mission with a strict 6-h-on/6-h-off watch routine", Journal of Applied Physiology, Vol. 132, Issue 4. DOI: 10.1152/japplphysiol.00130.2021.

[64] Gunn, P. J., Middleton, B., Davies, S. K., Revell, V. L. and Skene, D. J. (2016). "Sex differences in the circadian profiles of melatonin and cortisol in plasma and urine matrices under constant routine conditions", Chronobiology International, Vol. 33, Issue 1, p. 39. DOI: 10.3109/07420528.2015.1112396.

[65] Honma, A., Revell, V. L., Gunn, P. J., Davies, S. K., Middleton, B., et al. (2020). "Effect of acute total sleep deprivation on plasma melatonin, cortisol and metabolite rhythms in females", European Journal of Neuroscience, Vol. 51, Issue 1, pp. 366–378. DOI: 10.1111/ejn.14411.

[66] Depner, C. M., Stothard, E. R. and Wright, K. P. (2014). "Metabolic consequences of sleep and circadian disorders", Current Diabetes Reports, Vol. 14, Issue 7. DOI: 10.1007/s11892-014-0507-z.

[67] Knoop, M., Stefani, O., Bueno, B., Matusiak, B., Hobday, R., et al. (2020). "Daylight: What makes the difference?", Lighting Research and Technology, Vol. 52, Issue 3, pp. 423–442. DOI: 10.1177/1477153519869758.

[68] Świątkiewicz, I., Woźniak, A. and Taub, P. R. (2021). "Time-restricted eating and metabolic syndrome: Current status and future perspectives", Nutrients, Vol. 13, Issue 1, pp. 1–24. DOI: 10.3390/nu13010221.

[69] Münch, M., Wirz-Justice, A., Brown, S. A., Kantermann, T., Martiny, K., et al. (2020). "The Role of Daylight for Humans: Gaps in Current Knowledge", Clocks & Sleep, Vol. 2, Issue 1, pp. 61–85. DOI: 10.3390/clockssleep2010008.

[70] Skene, D. J., Skornyakov, E., Chowdhury, N. R., Gajula, R. P., Middleton, B., et al. (2018). "Separation of circadian- and behavior-driven metabolite rhythms in humans provides a window on peripheral oscillators and metabolism.", Proceedings of the National Academy of Sciences of the United States of America, Vol. 115, Issue 30, pp. 7825–7830. DOI: 10.1073/pnas.1801183115.

[71] Parent-Thirion, A., Biletta, I., Cabrita, J., Vargas, O., Vermeylen, G., et al. (2017). "Eurofound (2017), 6th European Working Conditions Survey – Overview report (2017 update)" Publications Office of the European Union (Luxembourg). DOI: 10.2806/784968.

[72] US Bureau of Labor Statistics, (2019). "Job Flexibilities and Work Schedules — 2017-2018 Data From the American Time Use Survey", no. 202, Accessed: Sep. 26, 2022. [Online]. Available: www.bls.gov/tus

[73] Touitou, Y., Reinberg, A. and Touitou, D. (2017). "Association between light at night, melatonin secretion, sleep deprivation, and the internal clock: Health impacts and mechanisms of circadian disruption", Life Sciences, Vol. 173, pp. 94–106. DOI: 10.1016/J.LFS.2017.02.008.

[74] Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. and Hogenesch, J. B. (2014). "A circadian gene expression atlas in mammals: Implications for biology and medicine", Proceedings of the National Academy of Sciences of the United States of America, Vol. 111, Issue 45, pp. 16219–16224. DOI: 10.1073/pnas.1408886111.

[75] Cipolla-Neto, J., Amaral, F. G., Afeche, S. C., Tan, D. X. and Reiter, R. J. (2014). "Melatonin, energy metabolism, and obesity: a review", Journal of Pineal Research, Vol. 56, Issue 4, pp. 371-381. DOI: 10.1111/jpi.12137.

[76] Poggiogalle, E., Jamshed, H. and Peterson, C. M. (2018). "Circadian Regulation of Glucose, Lipid, and Energy Metabolism in Humans", Metabolism: clinical and experimental, Vol. 84, p. 11. DOI: 10.1016/J.METABOL.2017.11.017.

[77] Stenvers, D. J., Scheer, F. A. J. L., Schrauwen, P., la Fleur, S. E. and Kalsbeek, A. (2019). "Circadian clocks and insulin resistance", Nature Reviews Endocrinology, Vol. 15, Issue 2, pp. 75–89. DOI: 10.1038/s41574-018-0122-1.

[78] Hancox, T. P. M., Skene, D. J., Dallmann, R. and Dunn, W. B. (2021). "Tick-tock consider the clock: The influence of circadian and external cycles on time of day variation in the human metabolome—a review", Metabolites, Vol. 11, Issue 5. DOI: 10.3390/metabo11050328.

[79] Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R. and Griffin, J. L. (2011). "Systems level studies of mammalian metabolomes: The roles of mass spectrometry and nuclear

magnetic resonance spectroscopy", Chemical Society Reviews, Vol. 40, Issue 1, pp. 387–426. DOI: 10.1039/b906712b.

[80] Peironcely, J. E., Reijmers, T., Coulier, L., Bender, A. and Hankemeier, T. (2011). "Understanding and Classifying Metabolite Space and Metabolite-Likeness", PLOS ONE, Vol. 6, Issue 12. DOI: 10.1371/JOURNAL.PONE.0028966.

[81] Johnson, C. H., Ivanisevic, J. and Siuzdak, G. (2016). "Metabolomics: Beyond biomarkers and towards mechanisms", Nature Reviews Molecular Cell Biology, Vol. 17, Issue 7, pp. 451–459. DOI: 10.1038/nrm.2016.25.

[82] Vineis, P., Robinson, O., Chadeau-Hyam, M., Dehghan, A., Mudway, I., et al. (2020). "What is new in the exposome?", Environment international, Vol. 143. DOI: 10.1016/J.ENVINT.2020.105887.

[83] Heiles, S. (2021). "Advanced tandem mass spectrometry in metabolomics and lipidomics-methods and applications", Analytical and bioanalytical chemistry, Vol. 413, Issue 24, pp. 5927–5948. DOI: 10.1007/S00216-021-03425-1.

[84] "International Union of Pure and Applied Chemistry." Accessed: Mar. 02, 2023. [Online]. Available: https://iupac.org/

[85] Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., et al. (2016). "ClassyFire: automated chemical classification with a comprehensive, computable taxonomy", Journal of Cheminformatics, Vol. 8, p. 61. DOI: 10.1186/s13321-016-0174-y.

[86] Akhondi, S. A., Kors, J. A. and Muresan, S. (2012). "Consistency of systematic chemical identifiers within and between small-molecule databases", Journal of Cheminformatics, Vol. 4, Issue 12, p. 35. DOI: 10.1186/1758-2946-4-35.

[87] Liebisch, G., Fahy, E., Aoki, J., Dennis, E. A., Durand, T., et al. (2020). "Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures", Journal of Lipid Research, Vol. 61, Issue 12, pp. 1539–1555. DOI: 10.1194/jlr.S120001025.

[88] Quinn, R. A., Melnik, A. V., Vrbanac, A., Fu, T., Patras, K. A., et al. (2020). "Global chemical effects of the microbiome include new bile-acid conjugations", Nature, Vol. 579, Issue 7797, pp. 123–129. DOI: 10.1038/s41586-020-2047-9.

[89] Wishart, D. S., Guo, A. C., Oler, E., Wang, F., Anjum, A., et al. (2022). "HMDB 5.0: the Human Metabolome Database for 2022", Nucleic acids research, Vol. 50, Issue D1, pp. 622–631. DOI: 10.1093/NAR/GKAB1062.

[90] "LIPID MAPS." Accessed: Mar. 02, 2023. [Online]. Available: https://www.lipidmaps.org/

[91] Zhou, Z., Luo, M., Zhang, H., Yin, Y., Cai, Y., et al. (2022). "Metabolite annotation from knowns to unknowns through knowledge-guided multi-layer metabolic networking", Nature Communications, Vol. 13, Issue 1. DOI: 10.1038/S41467-022-34537-6.

[92] Kind, T. and Fiehn, O. (2007). "Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry", BMC Bioinformatics, Vol. 8, p. 105. DOI: 10.1186/1471-2105-8-105.

[93] Newman, J. C. and Verdin, E. (2014). "Ketone bodies as signaling metabolites", Trends in endocrinology and metabolism: TEM, Vol. 25, Issue 1, pp. 42–52. DOI: 10.1016/J.TEM.2013.09.002.

[94] Ricci, A., Di Betto, G., Bergamini, E., Buzzetti, E., Corradini, E., et al. (2022). "Iron Metabolism in the Disorders of Heme Biosynthesis", Metabolites, Vol. 12, Issue 9. DOI: 10.3390/METABO12090819.

[95] Nielsen, J. (2017). "Systems Biology of Metabolism", Annual review of biochemistry, Vol. 86, pp. 245–275. DOI: 10.1146/ANNUREV-BIOCHEM-061516-044757.

[96] Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M. (2023). "KEGG for taxonomy-based analysis of pathways and genomes", Nucleic Acids Research, Vol. 51, Issue D1, pp. 587–592. DOI: 10.1093/NAR/GKAC963.

[97] Kanehisa, M. and Goto, S. (2000). "KEGG: kyoto encyclopedia of genes and genomes", Nucleic acids research, Vol. 28, Issue 1, pp. 27–30. DOI: 10.1093/NAR/28.1.27.

[98] Medina, C. B., Mehrotra, P., Arandjelovic, S., Perry, J. S. A., Guo, Y., et al. (2020). "Metabolites released from apoptotic cells act as tissue messengers", Nature, Vol. 580. DOI: 10.1038/s41586-020-2121-3.

[99] Chen, L., Zhernakova, D. V., Kurilshikov, A., Andreu-Sánchez, S., Wang, D., et al. (2022). "Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome", Nature medicine, Vol. 28, Issue 11, pp. 2333–2343. DOI: 10.1038/S41591-022-02014-8.

[100] Wilson, A. S., Koller, K. R., Ramaboli, M. C., Nesengani, L. T., Ocvirk, S., et al. (2020). "Diet and the Human Gut Microbiome: An International Review", Digestive diseases and sciences, Vol. 65, Issue 3, pp. 723–740. DOI: 10.1007/S10620-020-06112-W.

[101] Sato, S., Dyar, K. A., Treebak, J. T., Jepsen, S. L., Ehrlich, A. M., et al. (2022). "Atlas of exercise metabolism reveals time-dependent signatures of metabolic homeostasis", Cell metabolism, Vol. 34, Issue 2, pp. 329-345. DOI: 10.1016/J.CMET.2021.12.016.

[102] Khoramipour, K., Sandbakk, Ø., Keshteli, A. H., Gaeini, A. A., Wishart, D. S., et al. (2022). "Metabolomics in Exercise and Sports: A Systematic Review", Sports medicine (Auckland, N.Z.), Vol. 52, Issue 3, pp. 547–583. DOI: 10.1007/S40279-021-01582-Y.

[103] Kelly, R. S., Kelly, M. P. and Kelly, P. (2020). "Metabolomics, physical activity, exercise and health: A review of the current evidence", Biochimica et biophysica acta. Molecular basis of disease, Vol. 1866, Issue 12. DOI: 10.1016/J.BBADIS.2020.165936.

[104] Rutledge, J., Oh, H. and Wyss-Coray, T. (2022). "Measuring biological age using omics data", Nature Reviews Genetics 2022 23:12, Vol. 23, Issue 12, pp. 715–727. DOI: 10.1038/s41576-022-00511-7.

[105] Costanzo, M., Caterino, M., Sotgiu, G., Ruoppolo, M., Franconi, F., et al. (2022). "Sex differences in the human metabolome", Biology of sex differences, Vol. 13, Issue 1. DOI: 10.1186/S13293-022-00440-4.

[106] Cook, D. J. and Nielsen, J. (2017). "Genome-scale metabolic models applied to human health and disease", Wiley Interdisciplinary Reviews: Systems Biology and Medicine, Vol. 9, Issue 6. DOI: 10.1002/WSBM.1393.

[107] Wang, Y., Eddy, J. A. and Price, N. D. (2012). "Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE", BMC Systems Biology, Vol. 6, p. 153. DOI: 10.1186/1752-0509-6-153.

[108] Ryu, J. Y., Kim, H. U. and Lee, S. Y. (2015). "Reconstruction of genome-scale human metabolic models using omics data", Integrative Biology, Vol. 7, Issue 8, pp. 859–868. DOI: 10.1039/C5IB00002E.

[109] Mardinoglu, A., Gatto, F. and Nielsen, J. (2013). "Genome-scale modeling of human metabolism - a systems biology approach", Biotechnology Journal, Vol. 8, Issue 9, pp. 985–996. DOI: 10.1002/BIOT.201200275.

[110] Gonçalves, E. and Frezza, C. (2021). "Genome and metabolome: chance and necessity", Genome Biology, Vol. 22, Issue 1. DOI: 10.1186/S13059-021-02501-0.

[111] Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., et al. (2013). "The Human Urine Metabolome", PLOS ONE, Vol. 8, Issue 9. DOI: 10.1371/JOURNAL.PONE.0073076.

[112] Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., et al. (2011). "The Human Serum Metabolome", PLOS ONE, Vol. 6, Issue 2. DOI: 10.1371/JOURNAL.PONE.0016957.

[113] Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., et al. (2010). "Targeted metabolomics for biomarker discovery", Angewandte Chemie - International Edition, Vol. 49, Issue 32, pp. 5426–5445. DOI: 10.1002/ANIE.200905579.

[114] Roberts, L. D., Souza, A. L., Gerszten, R. E. and Clish, C. B. (2012). "Targeted Metabolomics", Current Protocols in Molecular Biology, Vol. Ch-30, Issue Unit 30.2, pp. 1-24. DOI: 10.1002/0471142727.MB3002S98.

[115] Zhou, J. and Yin, Y. (2016). "Analyst CRITICAL REVIEW Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry", Cite this: Analyst, Vol. 141, p. 6362. DOI: 10.1039/c6an01753c.

[116] Cao, G., Song, Z., Hong, Y., Yang, Z., Song, Y., et al. (2020). "Large-scale targeted metabolomics method for metabolite profiling of human samples", Analytica Chimica Acta, Vol. 1125, pp. 144–151. DOI: 10.1016/J.ACA.2020.05.053.

[117] Zeki, Ö. C., Eylem, C. C., Reçber, T., Kır, S. and Nemutlu, E. (2020). "Integration of GC-MS and LC-MS for untargeted metabolomics profiling", Journal of pharmaceutical and biomedical analysis, Vol. 190. DOI: 10.1016/J.JPBA.2020.113509.

[118] Alseekh, S., Aharoni, A., Brotman, Y., Contrepois, K., D'Auria, J., et al. (2021). "Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices", Nature methods, Vol. 18, Issue 7, pp. 747–756. DOI: 10.1038/S41592-021-01197-1.

[119] Perez de Souza, L., Alseekh, S., Scossa, F. and Fernie, A. R. (2021). "Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research", Nature methods, Vol. 18, Issue 7, pp. 733–746. DOI: 10.1038/S41592-021-01116-4.

[120] Collins, S. L., Koo, I., Peters, J. M., Smith, P. B. and Patterson, A. D. (2021). "Current Challenges and Recent Developments in Mass Spectrometry–Based Metabolomics", Annual Reviews in

Analytical Chemistry, Vol. 14, Issue 1, pp. 467–487. DOI: 10.1146/ANNUREV-ANCHEM-091620-015205.

[121] Rey-Stolle, F., Dudzik, D., Gonzalez-Riano, C., Fernández-García, M., Alonso-Herranz, V., et al. (2022). "Low and high resolution gas chromatography-mass spectrometry for untargeted metabolomics: A tutorial", Analytica Chimica Acta, Vol. 1210, p. 339043. DOI: 10.1016/J.ACA.2021.339043.

[122] Keen, B., Cawley, A., Reedy, B. and Fu, S. (2022). "Metabolomics in clinical and forensic toxicology, sports anti-doping and veterinary residues", Drug testing and analysis, Vol. 14, Issue 5, pp. 794–807. DOI: 10.1002/DTA.3245.

[123] Locci, E., Bazzano, G., Chighine, A., Locco, F., Ferraro, E., et al. (2020). "Forensic NMR metabolomics: one more arrow in the quiver", Metabolomics : Official journal of the Metabolomic Society, Vol. 16, Issue 11. DOI: 10.1007/S11306-020-01743-6.

[124] Narduzzi, L., Dervilly, G., Audran, M., Le Bizec, B. and Buisson, C. (2020). "A role for metabolomics in the antidoping toolbox?", Drug testing and analysis, Vol. 12, Issue 6, pp. 677–690. DOI: 10.1002/DTA.2788.

[125] Szeremeta, M., Pietrowska, K., Niemcunowicz-Janica, A., Kretowski, A. and Ciborowski, M. (2021). "Applications of Metabolomics in Forensic Toxicology and Forensic Medicine", International journal of molecular sciences, Vol. 22, Issue 6, pp. 1–16. DOI: 10.3390/IJMS22063010.

[126] Alarcon-Barrera, J. C., Kostidis, S., Ondo-Mendez, A. and Giera, M. (2022). "Recent advances in metabolomics analysis for early drug development", Drug discovery today, Vol. 27, Issue 6, pp. 1763–1773. DOI: 10.1016/J.DRUDIS.2022.02.018.

[127] Aminov, R. (2022). "Metabolomics in antimicrobial drug discovery", Expert opinion on drug discovery, Vol. 17, Issue 9, pp. 1047–1059. DOI: 10.1080/17460441.2022.2113774.

[128] Wishart, D. S. (2016). "Emerging applications of metabolomics in drug discovery and precision medicine", Nature reviews. Drug discovery, Vol. 15, Issue 7, pp. 473–484. DOI: 10.1038/NRD.2016.32.

[129] Wolfender, J. L., Litaudon, M., Touboul, D. and Queiroz, E. F. (2019). "Innovative omics-based approaches for prioritisation and targeted isolation of natural products - new strategies for drug discovery", Natural product reports, Vol. 36, Issue 6, pp. 855–868. DOI: 10.1039/C9NP00004F.

[130] Kuhlisch, C. and Pohnert, G. (2015). "Metabolomics in chemical ecology", Natural product reports, Vol. 32, Issue 7, pp. 937–955. DOI: 10.1039/C5NP00003C.

[131] Liu, Y., Liu, Q., Krivoruchko, A., Khoomrung, S. and Nielsen, J. (2020). "Engineering yeast phospholipid metabolism for de novo oleoylethanolamide production", Nature chemical biology, Vol. 16, Issue 2, pp. 197–205. DOI: 10.1038/S41589-019-0431-2.

[132] Nadella, K. D., Marla, S. S. and Kumar, P. A. (2012). "Metabolomics in agriculture", Omics : a journal of integrative biology, Vol. 16, Issue 4, pp. 149–159. DOI: 10.1089/OMI.2011.0067.

[133] Iriti, M. and Vitalini, S. (2018). "Plant Metabolomics in the Global Scenario of Food Security: A Systems-Biology Approach for Sustainable Crop Production", International journal of molecular sciences, Vol. 19, Issue 7. DOI: 10.3390/IJMS19072094.

[134] Shi, T., Zhu, A., Jia, J., Hu, X., Chen, J., et al. (2020). "Metabolomics analysis and metabolite-agronomic trait associations using kernels of wheat (Triticum aestivum) recombinant inbred lines", The Plant journal : for cell and molecular biology, Vol. 103, Issue 1, pp. 279–292. DOI: 10.1111/TPJ.14727.

[135] Kumar, M., Patel, M. K., Kumar, N., Bajpai, A. B. and Siddique, K. H. M. (2021). "Metabolomics and Molecular Approaches Reveal Drought Stress Tolerance in Plants", International journal of molecular sciences, Vol. 22, Issue 17. DOI: 10.3390/IJMS22179108.

[136] Selamat, J., Rozani, N. A. A. and Murugesu, S. (2021). "Application of the Metabolomics Approach in Food Authentication", Molecules (Basel, Switzerland), Vol. 26, Issue 24. DOI: 10.3390/MOLECULES26247565.

[137] Crook, A. A. and Powers, R. (2020). "Quantitative NMR-Based Biomedical Metabolomics: Current Status and Applications", Molecules, Vol. 25, Issue 21. DOI: 10.3390/MOLECULES25215128.

[138] Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D. and McLean, J. A. (2016). "Untargeted metabolomics strategies – Challenges and Emerging Directions", Journal of the American Society for Mass Spectrometry, Vol. 27, Issue 12, p. 1897. DOI: 10.1007/S13361-016-1469-Y.

[139] Di Minno, A., Gelzo, M., Stornaiuolo, M., Ruoppolo, M. and Castaldo, G. (2021). "The evolving landscape of untargeted metabolomics", Nutrition, Metabolism and Cardiovascular Diseases, Vol. 31, Issue 6, pp. 1645–1652. DOI: 10.1016/J.NUMECD.2021.01.008.

[140] Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., et al. (2010). "Targeted Metabolomics for Biomarker Discovery", Angewandte Chemie International Edition, Vol. 49, Issue 32, pp. 5426–5445. DOI: 10.1002/anie.200905579.

[141] Khamis, M. M., Klemm, N., Adamko, D. J. and El-Aneed, A. (2018). "Comparison of accuracy and precision between multipoint calibration, single point calibration, and relative quantification for targeted metabolomic analysis", Analytical and Bioanalytical Chemistry, Vol. 410, Issue 23, pp. 5899–5913. DOI: 10.1007/s00216-018-1205-5.

[142] Pinu, F. R. (2018). "Grape and wine metabolomics to develop new insights using untargeted and targeted approaches", Fermentation, Vol. 4, Issue 4. DOI: 10.3390/fermentation4040092.

[143] Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-Mcintyre, S., et al. (2011). "Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry", Nature Protocols, Vol. 6, Issue 7, pp. 1060–1083. DOI: 10.1038/nprot.2011.335.

[144] Zhong, P., Wei, X., Li, X., Wei, X., Wu, S., et al. (2022). "Untargeted metabolomics by liquid chromatography-mass spectrometry for food authentication: A review", Comprehensive Reviews in Food Science and Food Safety, Vol. 21, Issue 3, pp. 2455–2488. DOI: 10.1111/1541-4337.12938.

[145] Forshed, J. (2017). "Experimental Design in Clinical 'Omics Biomarker Discovery", Journal of Proteome Research, Vol. 16, Issue 11, pp. 3954–3960. DOI: 10.1021/ACS.JPROTEOME.7B00418.

[146] Piras, C., Pibiri, M., Leoni, V. P., Balsamo, A., Tronci, L., et al. (2021). "Analysis of metabolomics profile in hypothyroid patients before and after thyroid hormone replacement", Journal of

Endocrinological Investigation, Vol. 44, Issue 6, pp. 1309–1319. DOI: 10.1007/S40618-020-01434-Y.

[147] Lee, T., Clavel, T., Smirnov, K., Schmidt, A., Lagkouvardos, I., et al. (2016). "Oral versus intravenous iron replacement therapy distinctly alters the gut microbiota and metabolome in patients with IBD", Gut, Vol. 66, Issue 5, pp. 863–871. DOI: 10.1136/GUTJNL-2015-309940.

[148] Alimajstorovic, Z., Mollan, S. P., Grech, O., Mitchell, J. L., Yiangou, A., et al. (2023). "Dysregulation of Amino Acid, Lipid, and Acylpyruvate Metabolism in Idiopathic Intracranial Hypertension: A Non-targeted Case Control and Longitudinal Metabolomic Study", Journal of Proteome Research, Vol. 22, Issue 4, pp. 1127–1137. DOI: 10.1021/ACS.JPROTEOME.2C00449.

[149] Varma, V. R., Oommen, A. M., Varma, S., Casanova, R., An, Y., et al. (2018). "Brain and blood metabolite signatures of pathology and progression in Alzheimer disease: A targeted metabolomics study", PLoS medicine, Vol. 15, Issue 1. DOI: 10.1371/JOURNAL.PMED.1002482.

[150] Villoslada, P., Alonso, C., Agirrezabal, I., Kotelnikova, E., Zubizarreta, I., et al. (2017). "Metabolomic signatures associated with disease severity in multiple sclerosis", Neurology(R) neuroimmunology & neuroinflammation, Vol. 4, Issue 2. DOI: 10.1212/NXI.0000000000000321.

[151] Sindelar, M., Stancliffe, E., Schwaiger-Haber, M., García-Sastre, A., Shriver, L. P., et al. (2021). "Longitudinal metabolomics of human plasma reveals prognostic markers of COVID-19 disease severity", Cell Reports Medicine, Vol. 2, p. 100369. DOI: 10.1016/j.xcrm.2021.100369.

[152] Yin, P., Lehmann, R. and Xu, G. (2015). "Effects of pre-analytical processes on blood samples used in metabolomics studies", Analytical and bioanalytical chemistry, Vol. 407, Issue 17, pp. 4879–4892. DOI: 10.1007/S00216-015-8565-X.

[153] Gil, A., Zhang, W., Wolters, J. C., Permentier, H., Boer, T., et al. (2018). "One- vs two-phase extraction: re-evaluation of sample preparation procedures for untargeted lipidomics in plasma samples", Analytical and Bioanalytical Chemistry, Vol. 410, Issue 23, p. 5859. DOI: 10.1007/S00216-018-1200-X.

[154] Fu, X., Calderón, C., Harm, T., Gawaz, M. and Lämmerhofer, M. (2022). "Advanced unified monophasic lipid extraction protocol with wide coverage on the polarity scale optimized for large-scale untargeted clinical lipidomics analysis of platelets", Analytica Chimica Acta, Vol. 1221, p. 340155. DOI: 10.1016/J.ACA.2022.340155.

[155] Southam, A. D., Haglington, L. D., Najdekr, L., Jankevics, A., Weber, R. J. M., et al. (2020). "Assessment of human plasma and urine sample preparation for reproducible and high-throughput UHPLC-MS clinical metabolic phenotyping", Analyst, Vol. 145, Issue 20, pp. 6511–6523. DOI: 10.1039/d0an01319f.

[156] Han, W. and Li, L. (2022). "Evaluating and minimizing batch effects in metabolomics", Mass Spectrometry Reviews, Vol. 41, Issue 3, pp. 421–442. DOI: 10.1002/MAS.21672.

[157] Dudzik, D., Barbas-Bernardos, C., García, A. and Barbas, C. (2018). "Quality assurance procedures for mass spectrometry untargeted metabolomics. a review", Journal of Pharmaceutical and Biomedical Analysis, Vol. 147, pp. 149–173. DOI: 10.1016/j.jpba.2017.07.044.

[158] Johnson, W. E., Li, C. and Rabinovic, A. (2007). "Adjusting batch effects in microarray expression data using empirical Bayes methods", Biostatistics (Oxford, England), Vol. 8, Issue 1, pp. 118–127. DOI: 10.1093/BIOSTATISTICS/KXJ037.

[159] Karpievitch, Y. V., Nikolic, S. B., Wilson, R., Sharman, J. E. and Edwards, L. M. (2014). "Metabolomics Data Normalization with EigenMS", PLOS ONE, Vol. 9, Issue 12, p. e116221. DOI: 10.1371/JOURNAL.PONE.0116221.

[160] Wehrens, R., Hageman, J. A., van Eeuwijk, F., Kooke, R., Flood, P. J., et al. (2016). "Improved batch correction in untargeted MS-based metabolomics", Metabolomics, Vol. 12, Issue 5. DOI: 10.1007/S11306-016-1015-8.

[161] Pang, Z., Chong, J., Zhou, G., Anderson De Lima Morais, D., Chang, L., et al. (2021). "MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights", Nucleic Acids Research, Vol. 49. DOI: 10.1093/nar/gkab382.

[162] Kirwan, J. A., Gika, H., Beger, R. D., Bearden, D., Dunn, W. B., et al. (2022). "Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management", Metabolomics, Vol. 18, Issue 9, p. 70. DOI: 10.1007/S11306-022-01926-3.

[163] Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., et al. (2018). "Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies", Metabolomics, Vol. 14, Issue 6, p. 72. DOI: 10.1007/S11306-018-1367-3.

[164] Chen, Y., Li, E. M. and Xu, L. Y. (2022). "Guide to Metabolomics Analysis: A Bioinformatics Workflow", Metabolites, Vol. 12, Issue 4. DOI: 10.3390/METABO12040357.

[165] Cambiaghi, A., Ferrario, M. and Masseroli, M. (2017). "Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration", Briefings in Bioinformatics, Vol. 18, Issue 3, pp. 498–510. DOI: 10.1093/BIB/BBW031.

[166] Anwardeen, N. R., Diboun, I., Mokrab, Y., Althani, A. A. and Elrayess, M. A. (2023). "Statistical methods and resources for biomarker discovery using metabolomics", BMC Bioinformatics 2023 24:1, Vol. 24, Issue 1, pp. 1–18. DOI: 10.1186/S12859-023-05383-0.

[167] Lever, J., Krzywinski, M. and Altman, N. (2017). "Points of Significance: Principal component analysis", Nature Methods, Vol. 14, Issue 7, pp. 641–642. DOI: 10.1038/NMETH.4346.

[168] Worley, B. and Powers, R. (2016). "PCA as a Practical Indicator of OPLS-DA Model Reliability", Current Metabolomics, Vol. 4, Issue 2, pp. 97–103. DOI: 10.2174/2213235x04666160613122429.

[169] Roca, M., Alcoriza, M. I., Garcia-Cañaveras, J. C. and Lahoz, A. (2021). "Reviewing the metabolome coverage provided by LC-MS: Focus on sample preparation and chromatography-A tutorial", Analytica Chimica Acta, Vol. 1147, pp. 38–55. DOI: 10.1016/J.ACA.2020.12.025.

[170] Emwas, A. H., Roy, R., McKay, R. T., Tenori, L., Saccenti, E., et al. (2019). "NMR Spectroscopy for Metabolomics Research", Metabolites 2019, Vol. 9, Issue 7, p. 123. DOI: 10.3390/METABO9070123.

[171] Plumb, R., Castro-Perez, J., Granger, J., Beattie, I., Joncour, K., et al. (2004). "Ultra-performance liquid chromatography coupled to quadrupole-orthogonal time-of-flight mass spectrometry", Rapid communications in mass spectrometry : RCM, Vol. 18, Issue 19, pp. 2331–2337. DOI: 10.1002/RCM.1627.

[172] Wishart, D. S., Cheng, L. L., Copié, V., Edison, A. S., Eghbalnia, H. R., et al. (2022). "NMR and Metabolomics—A Roadmap for the Future", Metabolites, Vol. 12, Issue 8. DOI: 10.3390/METABO12080678.

[173] Vinaixa, M., Schymanski, E. L., Neumann, S., Navarro, M., Salek, R. M., et al. (2016). "Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects", TrAC Trends in Analytical Chemistry, Vol. 78, pp. 23–35. DOI: 10.1016/J.TRAC.2015.09.005.

[174] Taylor, D. G., Inamdar, R. and Bushell, M. C. (1988). "NMR imaging in theory and in practice", Physics in Medicine and Biology, Vol. 33, Issue 6, pp. 635–669. DOI: 10.1088/0031-9155/33/6/001.

[175] Marion, D. (2013). "An Introduction to Biological NMR Spectroscopy", Molecular & Cellular Proteomics : MCP, Vol. 12, Issue 11, p. 3006. DOI: 10.1074/MCP.O113.030239.

[176] Horning, E. C. and VandenHeuvel, W. J. A. (1963). "Gas Chromatography", Annu. Rev. Biochem., Vol. 32, pp. 709–754. DOI: 10.1146/annurev.bi.32.070163.003425.

[177] Lísa, M., Cífková, E., Khalikova, M., Ovčačíková, M. and Holčapek, M. (2017). "Lipidomic analysis of biological samples: Comparison of liquid chromatography, supercritical fluid chromatography and direct infusion mass spectrometry methods", Journal of Chromatography A, Vol. 1525, pp. 96–108. DOI: 10.1016/J.CHROMA.2017.10.022.

[178] Wilson, J. N. (1940). "A Theory of Chromatography", Journal of the American Chemical Society, Vol. 62, Issue 6, pp. 1583–1591. DOI: 10.1021/JA01863A071.

[179] Joseph Weiss, B. (1943). "On the Theory of Chromatography", Journal of the Chemical Society, Vol. 1943, pp. 297-303. DOI: 10.1039/JR9430000297

[180] Lipsky, S.R. and Landowne, R.A. (1960). "Gas chromatography-biochemical applications", Annual Reviews of Biochemistry, Vol. 1960, Issue 29, pp. 649-668. DOI: 10.1146/annurev.bi.29.070160.003245.

[181] Walton, H. F. (1968). "Ion exchange chromatography", Analytical chemistry, Vol. 40, Issue 5, pp. 51–62. DOI: 10.1021/AC60261A020.

[182] Bruno, T. J. (1989). "A Supercritical Fluid Chromatograph for Physicochemical Studies", Journal of Research of the National Institute of Standards and Technology, Vol. 94, Issue 2, p. 105. DOI: 10.6028/JRES.094.013.

[183] READING, H. W. (1964). "Inter-Departmental Colloquium", Annals of The Royal College of Surgeons of England, Vol. 35, Issue 4, p. 234. DOI: 10.5005/jp/books/12054_37.

[184] Žuvela, P., Skoczylas, M., Jay Liu, J., Baczek, T., Kaliszan, R., et al. (2019). "Column Characterization and Selection Systems in Reversed-Phase High-Performance Liquid Chromatography", Chemical Reviews, Vol. 119, Issue 6, pp. 3674–3729. DOI: 10.1021/acs.chemrev.8b00246.

[185] Wagner, J. P., Schreiner, P. R., Schreiner, P. R. and Wagner, J. P. (2015). "London Dispersion London Dispersion in Molecular Chemistry-Reconsidering Steric Effects Angewandte Chemie Keywords: density functional theory · noncovalent interactions · van der Waals forces · weak interactions Angewandte Reviews", Angew.C hem. Int. Ed, Vol. 54, pp. 2274–12296. DOI: 10.1002/ange.201503476.

[186] Makowski, M. and Bogunia, M. (2020). "Influence of ionic strength on hydrophobic interactions in water: Dependence on solute size and shape", Journal of Physical Chemistry B, Vol. 124, Issue 46, pp. 10326–10336. DOI: 10.1021/ACS.JPCB.0C06399.

[187] Hao, Z., Xiao, B. and Weng, N. (2008). "Impact of column temperature and mobile phase components on selectivity of hydrophilic interaction chromatography (HILIC)", Journal of Separation Science, Vol. 31, Issue 9, pp. 1449–1464. DOI: 10.1002/JSSC.200700624.

[188] Layne, J. (2002). "Characterization and comparison of the chromatographic performance of conventional, polar-embedded, and polar-endcapped reversed-phase liquid chromatography stationary phases", Journal of Chromatography A, Vol. 957, Issue 2, pp. 149–164. DOI: 10.1016/S0021-9673(02)00193-0.

[189] Špánik, I. and Machyňáková, A. (2018). "Recent applications of gas chromatography with high-resolution mass spectrometry", Journal of Separation Science, Vol. 41, Issue 1, pp. 163–179. DOI: 10.1002/JSSC.201701016.

[190] Engel, K. M. and Schiller, J. (2021). "The value of coupling thin-layer chromatography to mass spectrometry in lipid research - a review", Journal of Chromatography B, Vol. 1185, p. 123001. DOI: 10.1016/J.JCHROMB.2021.123001.

[191] Morlock, G. E. (2021). "High-performance thin-layer chromatography combined with effect-directed assays and high-resolution mass spectrometry as an emerging hyphenated technology: A tutorial review", Analytica Chimica Acta, Vol. 1180, p. 338644. DOI: 10.1016/J.ACA.2021.338644.

[192] Ngere, J. B., Ebrahimi, K. H., Williams, R., Pires, E., Walsby-Tickle, J., et al. (2023). "Ion-Exchange Chromatography Coupled to Mass Spectrometry in Life Science, Environmental, and Medical Research", Analytical Chemistry, Vol. 95, Issue 1, pp. 152-166. DOI: 10.1021/acs.analchem.2c04298.

[193] Gordillo, R. (2021). "Supercritical fluid chromatography hyphenated to mass spectrometry for metabolomics applications", Journal of Separation Science, Vol. 44, Issue 1, pp. 448–463. DOI: 10.1002/JSSC.202000805.

[194] van de Velde, B., Guillarme, D. and Kohler, I. (2020). "Supercritical fluid chromatography – Mass spectrometry in metabolomics: Past, present, and future perspectives", Journal of Chromatography B, Vol. 1161, p. 122444. DOI: 10.1016/J.JCHROMB.2020.122444.

[195] Aslani, S. and Armstrong, D. W. (2022). "High information spectroscopic detection techniques for gas chromatography", Journal of Chromatography A, Vol. 1676, p. 463255. DOI: 10.1016/J.CHROMA.2022.463255.

[196] de Oliveira, J. B., de Oliveira, C. G. R. and Pereira, G. E. (2023). "Flavan-3-ol (Flavanol) Identification and Quantitation by High Performance Liquid Chromatography Coupled with Mass Spectrometry (HPLC-MSn)", M.B.M. de Castilhos (ed), Basic Protocols in Enology and Winemaking, Humana, New York, NY, pp. 87–100. DOI: 10.1007/978-1-0716-3088-4_8.

[197] Guo, Y. (2015). "Recent progress in the fundamental understanding of hydrophilic interaction chromatography (HILIC)", vol. 140, Issue 19, pp. 6452–6466. DOI: 10.1039/C5AN00670H.

[198] Buszewski, B. and Noga, S. (2012). "Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique", Analytical and bioanalytical chemistry, Vol. 402, Issue 1, pp. 231–247. DOI: 10.1007/S00216-011-5308-5.

[199] McCalley, D. V. (2015). "Study of retention and peak shape in hydrophilic interaction chromatography over a wide pH range", Journal of Chromatography A, Vol. 1411, pp. 41–49. DOI: 10.1016/J.CHROMA.2015.07.092.

[200] Bajad, S. U., Lu, W., Kimball, E. H., Yuan, J., Peterson, C., et al. (2006). "Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry", Journal of Chromatography A, Vol. 1125, Issue 1, pp. 76–88. DOI: 10.1016/j.chroma.2006.05.019.

[201] Karatapanis, A. E., Fiamegos, Y. C. and Stalikas, C. D. (2009). "HILIC separation and quantitation of water-soluble vitamins using diol column", Journal of separation science, Vol. 32, Issue 7, pp. 909–917. DOI: 10.1002/JSSC.200800525.

[202] Fairchild, J. N., Horvath, K., Gooding, J. R., Campagna, S. R. and Guiochon, G. (2010). "Two-dimensional liquid chromatography/mass spectrometry/mass spectrometry separation of water-soluble metabolites", Journal of chromatography. A, Vol. 1217, Issue 52, pp. 8161–8166. DOI: 10.1016/J.CHROMA.2010.10.068.

[203] Paglia, G. and Astarita, G. (2022). "A High-Throughput HILIC-MS-Based Metabolomic Assay for the Analysis of Polar Metabolites", Methods in Molecular Biology, Vol. 2396, pp. 137–159. DOI: 10.1007/978-1-0716-1822-6_11.

[204] Li, A., Hines, K. M. and Xu, L. (2020). "Lipidomics by HILIC-Ion Mobility-Mass Spectrometry", Methods in Molecular Biology, Vol. 2084, pp. 119–132. DOI: 10.1007/978-1-0716-0030-6_7.

[205] Tsui, H. W., Lin, S. Z., Hsu, Y. C. and Dai, F. J. (2022). "Retention modeling and adsorption mechanisms in reversed-phase liquid chromatography", Journal of Chromatography A, Vol. 1662, p. 462736. DOI: 10.1016/J.CHROMA.2021.462736.

[206] Rafferty, J. L., Zhang, L., Siepmann, J. I. and Schure, M. R. (2007). "Retention mechanism in reversed-phase liquid chromatography: A molecular perspective", Analytical Chemistry, Vol. 79, Issue 17, pp. 6551–6558. DOI: 10.1021/ac0705115.

[207] Nakamura, K., Saito, S. and Shibukawa, M. (2020). "Intrinsic difference between phenyl hexyl- and octadecyl-bonded silicas in the solute retention selectivity in reversed-phase liquid chromatography with aqueous mobile phase", Journal of Chromatography A, Vol. 1628, p. 461450. DOI: 10.1016/J.CHROMA.2020.461450.

[208] Jaochico, A., Sangaraju, D. and Shahidi-Latham, S. K. (2019). "A rapid derivatization based LC-MS/MS method for quantitation of short chain fatty acids in human plasma and urine", Bioanalysis, Vol. 11, Issue 8, pp. 741–753. DOI: 10.4155/BIO-2018-0241.

[209] Tang, D.-Q., Zou, L., Yin, X.-X. and Ong, C. N. (2014). "HILIC-MS for metabolomics: an attractive and complementary approach to RPLC-MS", Mass Spectrometry Reviews, Vol. 35, Issue 5, pp. 574-600. DOI: 10.1002/mas.21445.

[210] Lv, W., Guo, L., Zheng, F., Wang, Q., Wang, W., et al. (2020). "Alternate reversed-phase and hydrophilic interaction liquid chromatography coupled with mass spectrometry for broad coverage in metabolomics analysis", Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences, Vol. 1152. DOI: 10.1016/j.jchromb.2020.122266.

[211] Jandera, P. (2011). "Stationary and mobile phases in hydrophilic interaction chromatography: a review", Analytica Chimica Acta, Vol. 692, Issue 1–2, pp. 1–25. DOI: 10.1016/J.ACA.2011.02.047.

[212] Najdekr, L., Blanco, G. R. and Dunn, W. B. (2019). "Collection of untargeted metabolomic data for mammalian urine applying HILIC and reversed phase ultra performance liquid chromatography methods coupled to a Q exactive mass spectrometer", Methods in Molecular Biology, Vol. 1996, pp. 1–15. DOI: 10.1007/978-1-4939-9488-5_1.

[213] Martias, C., Baroukh, N., Mavel, S., Blasco, H., Lefèvre, A., et al. (2021). "Optimization of Sample Preparation for Metabolomics Exploration of Urine, Feces, Blood and Saliva in Humans Using Combined NMR and UHPLC-HRMS Platforms", Molecules (Basel, Switzerland), Vol. 26, Issue 14. DOI: 10.3390/MOLECULES26144111.

[214] Khamis, M. M., Adamko, D. J. and El-Aneed, A. (2017). "Mass spectrometric based approaches in urine metabolomics and biomarker discovery", Mass Spectrometry Reviews, Vol. 36, Issue 2, pp. 115–134. DOI: 10.1002/mas.21455.

[215] Nanita, S. C. and Kaldon, L. G. (2016). "Emerging flow injection mass spectrometry methods for high-throughput quantitative analysis", Analytical and Bioanalytical Chemistry, Vol. 408, Issue 1, pp. 23–33. DOI: 10.1007/S00216-015-9193-1.

[216] Dame, Z. T., Aziat, F., Mandal, R., Krishnamurthy, R., Bouatra, S., et al. (2015). "The human saliva metabolome", Metabolomics, Vol. 11, Issue 6, pp. 1864–1883. DOI: 10.1007/s11306-015-0840-5.

[217] Bauermeister, A., Mannochio-Russo, H., Costa-Lotufo, L. V., Jarmusch, A. K. and Dorrestein, P. C. (2022). "Mass spectrometry-based metabolomics in microbiome investigations", Nature Reviews Microbiology, Vol. 20, Issue 3, pp. 143–160. DOI: 10.1038/s41579-021-00621-9.

[218] Moskovets, E., Misharin, A., Laiko, V. and Doroshenko, V. (2016). "A comparative study on the analytical utility of atmospheric and low-pressure MALDI sources for the mass spectrometric characterization of peptides", Methods, Vol. 104, pp. 21–32. DOI: 10.1016/J.YMETH.2016.02.009.

[219] Lubin, A., Bajic, S., Cabooter, D., Augustijns, P. and Cuyckens, F. (2017). "Atmospheric Pressure Ionization Using a High Voltage Target Compared to Electrospray Ionization", Journal of the American Society for Mass Spectrometry, Vol. 28, Issue 2, pp. 286–293. DOI: 10.1007/s13361-016-1537-3.

[220] Kruve, A. and Kaupmees, K. (2017). "Adduct Formation in ESI/MS by Mobile Phase Additives", Journal of the American Society for Mass Spectrometry, Vol. 28, Issue 5, pp. 887–894. DOI: 10.1007/s13361-017-1626-y.

[221] Erngren, I., Haglöf, J., Engskog, M. K. R., Nestor, M., Hedeland, M., et al. (2019). "Adduct formation in electrospray ionisation-mass spectrometry with hydrophilic interaction liquid chromatography is strongly affected by the inorganic ion concentration of the samples", Journal of Chromatography A, Vol. 1600, pp. 174–182. DOI: 10.1016/J.CHROMA.2019.04.049.

[222] Thermo Fisher Scientific, (2010). "Part of Thermo Fisher Scientific Exactive TM Operating Manual", Accessed: Sep. 01, 2023. [Online]. Available: www.thermo.com/WEEERoHS.

[223] Thermo Fisher Scientific, (2017). "Exactive Series Exactive Plus, Exactive Plus EMR, Q Exactive, Q Exactive Focus, Q Exactive Plus, Q Exactive HF, and Q Exactive HF-X Operating Manual", Thermo Fisher Scientific, Issue April, Accessed: Apr. 13, 2023. [Online]. Available: https://assets.thermofisher.com/TFS-Assets/CMD/manuals/man-bre0012255-exactive-series-manbre0012255-en.pdf

[224] Boesl, U. (2017). "Time-of-flight mass spectrometry: Introduction to the basics", Mass Spectrometry Reviews, Vol. 36, Issue 1, pp. 86–109. DOI: 10.1002/MAS.21520.

[225] Nolting, D., Malek, R., Makarov, A., Alexander Makarov, C. and Fisher Scientific, T. (2019). "Ion traps in modern mass spectrometry", Mass Spectrometry Reviews, Vol. 38, Issue 2, pp. 150–168. DOI: 10.1002/MAS.21549.

[226] Eliuk, S. and Makarov, A. (2015). "Evolution of Orbitrap Mass Spectrometry Instrumentation", Annual Review of Analytical Chemistry, Vol. 8, pp. 61–80. DOI: 10.1146/annurev-anchem-071114-040325.

[227] Scigelova, M., Hornshaw, M., Giannakopulos, A. and Makarov, A. (2011). "Fourier transform mass spectrometry", Molecular and Cellular Proteomics, Vol. 10, Issue 7. DOI: 10.1074/MCP.M111.009431.

[228] Challen, B. and Cramer, R. (2022). "Advances in ionisation techniques for mass spectrometry-based omics research", PROTEOMICS, Vol. 22, Issue 15–16, p. 2100394. DOI: 10.1002/PMIC.202100394.

[229] Wilm, M. (2011). "Principles of Electrospray Ionization", Molecular & Cellular Proteomics : MCP, Vol. 10, Issue 7. DOI: 10.1074/MCP.M111.009407.

[230] Kebarle, P. and Verkcerk, U. H. (2009). "Electrospray: From ions in solution to ions in the gas phase, what we know now", Mass Spectrometry Reviews, Vol. 28, Issue 6, pp. 898–917. DOI: 10.1002/MAS.20247.

[231] Banerjee, S. and Mazumdar, S. (2012). "Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte", International Journal of Analytical Chemistry, Vol. 2012, pp. 1–40. DOI: 10.1155/2012/282574.

[232] Hopfgartner, G., Varesio, E., Tschäppät, V., Grivet, C., Bourgogne, E., et al. (2004). "Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules", Journal of Mass Spectrometry, Vol. 39, Issue 8, pp. 845–855. DOI: 10.1002/JMS.659.

[233] Hoffmann Edmond de, S. V. (2007). "Mass Spectrometry: principles and applications, 3rd Edition", John Wiley & Sons, Hoboken, NJ. ISBN: 9780470033104.

[234] Raffaelli, A. and Saba, A. (2021). "Ion scanning or ion trapping: Why not both?", Mass Spectrometry Reviews, Vol. 42, Issue 4, pp. 1152-1173. DOI: 10.1002/MAS.21746.

[235] Nikolaev, E. N., Kostyukevich, Y. I. and Vladimirov, G. N. (2016). "Fourier transform ion cyclotron resonance (FT ICR) mass spectrometry: Theory and simulations", Mass spectrometry reviews, Vol. 35, Issue 2, pp. 219–258. DOI: 10.1002/MAS.21422.

[236] Makarov, A., Denisov, E. and Lange, O. (2009). "Performance Evaluation of a High-field Orbitrap Mass Analyzer", JAM, Vol. 20, pp. 1391–1396. DOI: 10.1016/j.jasms.2009.01.005.

[237] Ghaste, M., Mistrik, R. and Shulaev, V. (2016). "Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap Based High Resolution Mass Spectrometry in Metabolomics and Lipidomics", International Journal of Molecular Sciences, Vol. 17, Issue 6. DOI: 10.3390/IJMS17060816.

[238] Murray, K. K. (2022). "Resolution and Resolving Power in Mass Spectrometry", Journal for the American Society of Mass Spectrometry, Vol. 33, Issue 12, pp 2342-2347. DOI: 10.1021/jasms.2c00216.

[239] Faktor J,Dvorakova M,Maryas J,Struharova I,and Bouchal P,. Identification and Characterisation of Pro-metastatic Targets, Pathways and Molecular Complexes Using a Toolbox of Proteomic Technologies Identifikace a charakterizace prometastatických cílů, drah a molekulárních komplexů s využitím proteomických technologií. Klin Onkol 25, 2–70 (2012). (2012). "Identification and Characterisation of Pro-metastatic Targets, Pathways and Molecular Complexes Using a Toolbox of Proteomic Technologies", Klinicka onkologie, Vol. 25, Issue 2, pp. 2S70–2S77.

[240] Mitchell Wells, J. and McLuckey, S. A. (2005). "Collision-Induced Dissociation (CID) of Peptides and Proteins", Methods in Enzymology, Vol. 402, pp. 148–185. DOI: 10.1016/S0076-6879(05)02005-7.

[241] Ichou, F., Schwarzenberg, A., Lesage, D., Alves, S., Junot, C., et al. (2014). "Comparison of the activation time effects and the internal energy distributions for the CID, PQD and HCD excitation modes", Journal of Mass Spectrometry, Vol. 49, Issue 6, pp. 498–508. DOI: 10.1002/jms.3365.

[242] Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. and Siuzdak, G. (2006). "XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification", Analytical Chemistry, Vol. 78, Issue 3, pp. 779–787. DOI: 10.1021/AC051437Y.

[243] Huan, T., Forsberg, E. M., Rinehart, D., Johnson, C. H., Ivanisevic, J., et al. (2017). "Systems biology guided by XCMS Online metabolomics", Nature methods, Vol. 14, Issue 5, p. 461. DOI: 10.1038/NMETH.4260.

[244] Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., et al. (2015). "MS-DIAL: Data Independent MS/MS Deconvolution for Comprehensive Metabolome Analysis", Nature methods, Vol. 12, Issue 6, p. 523. DOI: 10.1038/NMETH.3393.

[245] Tsugawa, H., Ikeda, K., Takahashi, M., Satoh, A., Mori, Y., et al. (2020). "A lipidome atlas in MS-DIAL 4", Nature Biotechnology 2020 38:10, Vol. 38, Issue 10, pp. 1159–1163. DOI: 10.1038/s41587-020-0531-2.

[246] Katajamaa, M. and Orešič, M. (2005). "Processing methods for differential analysis of LC/MS profile data", BMC Bioinformatics, Vol. 6, p. 179. DOI: 10.1186/1471-2105-6-179.

[247] Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., et al. (2023). "Integrative analysis of multimodal mass spectrometry data in MZmine 3", Nature Biotechnology 2023, pp. 1–3. DOI: 10.1038/s41587-023-01690-2.

[248] Kessner, D., Chambers, M., Burke, R., Agus, D. and Mallick, P. (2008). "ProteoWizard: open source software for rapid proteomics tools development", Bioinformatics (Oxford, England), Vol. 24, Issue 21, pp. 2534–2536. DOI: 10.1093/BIOINFORMATICS/BTN323.

[249] Adusumilli, R. and Mallick, P. (2017). "Data Conversion with ProteoWizard msConvert", Methods in molecular biology (Clifton, N.J.), Vol. 1550, pp. 339–368. DOI: 10.1007/978-1-4939-6747-6_23.

[250] Mahieu, N. G., Genenbacher, J. L. and Patti, G. J. (2016). "A Roadmap for the XCMS Family of Software Solutions in Metabolomics", Current opinion in chemical biology, Vol. 30, p. 87. DOI: 10.1016/J.CBPA.2015.11.009.

[251] Rainer, J. (2022). "LC/MS Preprocessing and Analysis with XCMS." Accessed: Mar. 23, 2023. [Online]. Available: http://sashimi.sourceforge.net/software_glossolalia.html

[252] Domingo-Almenara, X. and Siuzdak, G. (2020). "Metabolomics Data Processing Using XCMS", Methods in Molecular Biology, Vol. 2104, pp. 11–24. DOI: 10.1007/978-1-0716-0239-3_2.

[253] Libiseller, G., Dvorzak, M., Kleb, U., Gander, E., Eisenberg, T., et al. (2015). "IPO: a tool for automated optimization of XCMS parameters", BMC bioinformatics, Vol. 16, Issue 1. DOI: 10.1186/S12859-015-0562-8.

[254] Albóniga, O. E., González, O., Alonso, R. M., Xu, Y. and Goodacre, R. (2020). "Optimization of XCMS parameters for LC-MS metabolomics: an assessment of automated versus manual tuning and its effect on the final results", Metabolomics : Official journal of the Metabolomic Society, Vol. 16, Issue 1. DOI: 10.1007/S11306-020-1636-9.

[255] Cleveland, W. S., Grosse, E. and Shyu, W. M. (1992). "Local Regression Models", Statistical Models in S, Routledge, Oxfordshire, UK, pp. 309–376. DOI: 10.1201/9780203738535-8.

[256] Cleveland, W. S., Devlin, S. J. and Grosse, E. (1988). "Regression by local fitting: Methods, properties, and computational algorithms", Journal of Econometrics, Vol. 37, Issue 1, pp. 87–114. DOI: 10.1016/0304-4076(88)90077-2.

[257] Cleveland, W. S. and Devlin, S. J. (1988). "Locally weighted regression: An approach to regression analysis by local fitting", Journal of the American Statistical Association, Vol. 83, Issue 403, pp. 596–610. DOI: 10.1080/01621459.1988.10478639.

[258] Prince, J. T. and Marcotte, E. M. (2006). "Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping", Analytical Chemistry, Vol. 78, Issue 17, pp. 6140–6152. DOI: 10.1021/ac0605344.

[259] Fiehn, O., Robertson, D., Griffin, J., vab der Werf, M., Nikolau, B., et al. (2007). "The metabolomics standards initiative (MSI)", Metabolomics, Vol. 3, Issue 3, pp. 175–178. DOI: 10.1007/S11306-007-0070-6.

[260] Sumner, L. W., Alexander, A. E., Ae, A. N. L. A. A. W. N., Ae, D. B., Ae, M. H. B. M., et al. (2007). "Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)", Metabolomics, Vol. 3, pp. 211–221. DOI: 10.1007/s11306-007-0082-2.

[261] Chaleckis, R., Meister, I., Zhang, P. and Wheelock, C. E. (2019). "Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics", Current Opinion in Biotechnology, Vol. 55, pp. 44–50. DOI: 10.1016/J.COPBIO.2018.07.010.

[262] de Jonge, N. F., Mildau, K., Meijer, D., Louwen, J. J. R., Bueschl, C., et al. (2022). "Good practices and recommendations for using and benchmarking computational metabolomics metabolite annotation tools", Metabolomics, Vol. 18, Issue 12, p. 103. DOI: 10.1007/S11306-022-01963-Y.

[263] Domingo-Almenara, X., Montenegro-Burke, J. R., Benton, H. P. and Siuzdak, G. (2018). "Annotation: a computational solution for streamlining metabolomics analysis", Analytical chemistry, Vol. 90, Issue 1, p. 480. DOI: 10.1021/ACS.ANALCHEM.7B03929.

[264] Crotti, S., Menicatti, M., Pallecchi, M. and Bartolucci, G. (2021). "Tandem mass spectrometry approaches for recognition of isomeric compounds mixtures", Mass Spectrometry Reviews, Vol. 42, Issue 4, pp. 1244-1260. DOI: 10.1002/MAS.21757.

[265] Domingo-Almenara, X., Guijas, C., Billings, E., Montenegro-Burke, J. R., Uritboonthai, W., et al. (2019). "The METLIN small molecule dataset for machine learning-based retention time prediction", Nature Communications 2019 10:1, Vol. 10, Issue 1, pp. 1–9. DOI: 10.1038/s41467-019-13680-7.

[266] Bonini, P., Kind, T., Tsugawa, H., Barupal, D. K. and Fiehn, O. (2020). "Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics", Analytical chemistry, Vol. 92, Issue 11, pp. 7515–7522. DOI: 10.1021/ACS.ANALCHEM.9B05765.

[267] Stanstrup, J., Neumann, S. and Vrhovšek, U. (2015). "PredRet: prediction of retention time by direct mapping between multiple chromatographic systems", Analytical chemistry, Vol. 87, Issue 18, pp. 9421–9428. DOI: 10.1021/ACS.ANALCHEM.5B02287.

[268] Bach, E., Szedmak, S., Brouard, C., Böcker, S. and Rousu, J. (2018). "Liquid-chromatography retention order prediction for metabolite identification", Bioinformatics, Vol. 34, Issue 17, pp. 875–883. DOI: 10.1093/BIOINFORMATICS/BTY590.

[269] Chong, J. and Xia, J. (2018). "MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data", Bioinformatics, Vol. 34, Issue 24, pp. 4313–4314. DOI: 10.1093/BIOINFORMATICS/BTY528.

[270] Carlucci, M., Krisciunas, A., Li, H., Gibas, P., Koncevicius, K., et al. (2020). "DiscoRhythm: An easy-to-use web application and R package for discovering rhythmicity", Bioinformatics, Vol. 36, Issue 6, pp. 1952–1954. DOI: 10.1093/bioinformatics/btz834.

[271] Cornelissen, G. (2014). "Cosinor-based rhythmometry", Theoretical biology & medical modelling, Vol. 11, Issue 1. DOI: 10.1186/1742-4682-11-16.

[272] Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. and Hogenesch, J. B. (2016). "MetaCycle: An integrated R package to evaluate periodicity in large scale data", Bioinformatics, Vol. 32, Issue 21, pp. 3351–3353. DOI: 10.1093/bioinformatics/btw405.

[273] Hughes, M. E., Hogenesch, J. B. and Kornacker, K. (2010). "JTK-CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets", Journal of Biological Rhythms, Vol. 25, Issue 5, pp. 372–380. DOI: 10.1177/0748730410379711.

[274] Yang, R. and Su, Z. (2010). "Analyzing circadian expression data by harmonic regression based on autoregressive spectral estimation", Bioinformatics, Vol. 26, Issue 12, pp 168-174. DOI: 10.1093/BIOINFORMATICS/BTQ189.

[275] De Los Santos, H., Collins, E. J., Mann, C., Sagan, A. W., Jankowski, M. S., et al. (2020). "ECHO: An application for detection and analysis of oscillators identifies metabolic regulation on genome-wide circadian output", Bioinformatics, Vol. 36, Issue 3, pp. 773–781. DOI: 10.1093/bioinformatics/btz617.

[276] Thaben, P. F. and Westermark, P. O. (2014). "Detecting rhythms in time series with rain", Journal of Biological Rhythms, Vol. 29, Issue 6, pp. 391–400. DOI: 10.1177/0748730414553029.

[277] Ceglia, N., Liu, Y., Chen, S., Agostinelli, F., Eckel-Mahan, K., et al. (2018). "CircadiOmics: circadian omic web portal", Nucleic Acids Research, Vol. 46, pp. 157–162. DOI: 10.1093/nar/gky441.

[278] Agostinelli, F., Ceglia, N., Shahbaba, B., Sassone-Corsi, P. and Baldi, P. (2016). "What time is it? Deep learning approaches for circadian rhythms", Bioinformatics, Vol. 32, Issue 12, p. 8. DOI: 10.1093/BIOINFORMATICS/BTW243.

[279] Thaben, P. F. and Westermark, P. O. (2016). "Differential rhythmicity: detecting altered rhythmicity in biological data", Bioinformatics, Vol. 32, Issue 18, pp. 2800–2808. DOI: 10.1093/BIOINFORMATICS/BTW309.

[280] Pelikan, A., Herzel, H., Kramer, A. and Ananthasubramaniam, B. (2022). "Venn diagram analysis overestimates the extent of circadian rhythm reprogramming", The FEBS journal, Vol. 289, Issue 21, pp. 6605–6621. DOI: 10.1111/FEBS.16095.

[281] Hornik, K. (2012). "The Comprehensive R Archive Network", Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 4, Issue 4, pp. 394–398. DOI: 10.1002/WICS.1212.

[282] Spick, M., Hancox, T. P. M., Chowdhury, N. R., Middleton, B., Skene, D. J., et al. (2023). "Metabolomic Analysis of Plasma in Huntington's Disease Transgenic Sheep (Ovis aries) Reveals Progressive Circadian Rhythm Dysregulation", Journal of Huntington's Disease, Vol. 12, Issue 1, pp. 31–42. DOI: 10.3233/JHD-220552.

[283] Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R. and Kupfer, D. J. (1989). "The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research", Psychiatry Research, Vol. 28, Issue 2, pp. 193–213. DOI: 10.1016/0165-1781(89)90047-4.

[284] Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. and Erbaugh, J. (1961). "An Inventory for Measuring Depression", Archives of General Psychiatry, Vol. 4, Issue 6, pp. 561–571. DOI: 10.1001/ARCHPSYC.1961.01710120031004.

[285] Johns, M. W. (1991). "A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale", Sleep, Vol. 14, Issue 6, pp. 540–545. DOI: 10.1093/sleep/14.6.540.

[286] Horne, J. A. and Ostberg, O. (1976). "A self assessment questionnaire to determine Morningness Eveningness in human circadian rhythms", International Journal of Chronobiology, Vol. 4, Issue 2, pp. 97–110.

[287] Riad-Fahmy, D., Read, G. F., Gaskell, S. J., Dyas, J. and Hindawi, R. (1979). "A simple, direct radioimmunoassay for plasma cortisol, featuring a 125I radioligand and a solid-phase separation technique.", Clinical Chemistry, Vol. 25, Issue 5, pp. 665–668. DOI: 10.1093/clinchem/25.5.665.

[288] Chambers, M. C., MacLean, B., Burke, R., Amodei, D., Ruderman, D. L., et al. (2012). "A cross-platform toolkit for mass spectrometry and proteomics", Nature Biotechnology, Vol. 30, Issue 10, pp. 918–920. DOI: 10.1038/nbt.2377.

[289] Tautenhahn, R., Bottcher, C. and Neumann, S. (2008). "Highly sensitive feature detection for high resolution LC/MS", BMC Bioinformatics, Vol 9. Issue 504. DOI: 10.1186/1471-2105-9-504.

[290] Benton, H. P., Want, E. J. and Ebbels, T. M. D. (2010). "Data and text mining Correction of mass calibration gaps in liquid chromatography-mass spectrometry metabolomics data", BIOINFORMATICS APPLICATIONS NOTE, Vol. 26, pp. 2488–2489. DOI: 10.1093/bioinformatics/btq441.

[291] English, J., Arendt, J., Poulton, A. and Symons, A. M. (1987). "Short-term variations of circulating melatonin in the ewe", Journal of pineal research, Vol. 4, Issue 4, pp. 359–366. DOI: 10.1111/J.1600-079X.1987.TB00875.X.

[292] Isherwood, C. M., Van der Veen, D. R., Johnston, J. D. and Skene, D. J. (2017). "Twenty-four-hour rhythmicity of circulating metabolites: effect of body mass and type 2 diabetes.", FASEB journal : official publication of the Federation of American Societies for Experimental Biology, Vol. 31, Issue 12, pp. 5557–5567. DOI: 10.1096/fj.201700323R.

[293] Skene, D. J., Middleton, B., Fraser, C. K., Pennings, J. L. A., Kuchel, T. R., et al. (2017). "Metabolic profiling of presymptomatic Huntington's disease sheep reveals novel biomarkers", Scientific Reports 2017 7:1, Vol. 7, Issue 1, pp. 1–16. DOI: 10.1038/srep43030.

[294] Zhang, A., Sun, H., Yan, G., Wang, P. and Wang, X. (2015). "Review Article Metabolomics for Biomarker Discovery: Moving to the Clinic", Biomedical Research International, Vol. 2015. DOI: 10.1155/2015/354671.

[295] López-López, Á., López-Gonzálvez, Á., Barker-Tejeda, T. C. and Barbas, C. (2018). "A review of validated biomarkers obtained through metabolomics", Expert Review of Molecular Diagnostics, Vol. 18, Issue 6, pp. 557–575. DOI: 10.1080/14737159.2018.1481391.

[296] Álvarez-Sánchez, B., Priego-Capote, F. and Luque de Castro, M. D. (2010). "Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation", TrAC - Trends in Analytical Chemistry, Vol. 29, Issue 2, pp. 111–119. DOI: 10.1016/j.trac.2009.12.003.

[297] Zang, X., Monge, M. E. and Fernández, F. M. (2019). "Mass spectrometry-based non-targeted metabolic profiling for disease detection: Recent developments", TrAC - Trends in Analytical Chemistry, Vol. 118., pp. 158–169, 2019. DOI: 10.1016/j.trac.2019.05.030.

[298] León, Z., García-Cañaveras, J. C., Donato, M. T. and Lahoz, A. (2013). "Mammalian cell metabolomics: Experimental design and sample preparation", ELECTROPHORESIS, Vol. 34, Issue 19, pp. 2762–2775. DOI: 10.1002/ELPS.201200605.

[299] Hernandes, V. V., Barbas, C. and Dudzik, D. (2017). "A review of blood sample handling and pre-processing for metabolomics studies", ELECTROPHORESIS, Vol. 38, Issue 18, pp. 2232–2241. DOI: 10.1002/elps.201700086.

[300] Zelena, E., Dunn, W. B., Broadhurst, D., Francis-McIntyre, S., Carroll, K. M., et al. (2009). "Development of a robust and repeatable UPLC - MS method for the long-term metabolomic

study of human serum", Analytical Chemistry, Vol. 81, Issue 4, pp. 1357–1364. DOI: 10.1021/AC8019366.

[301] Considine, E. C. (2019). "The search for clinically useful biomarkers of complex disease: A data analysis perspective", Metabolites, Vol. 9, Issue 7, p. 126. DOI: 10.3390/metabo9070126.

[302] Di Guida, R., Engel, J., Allwood, J.W., Weber, R. J. M., Jones, M. R. et al., "Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling", Metabolomics, Vol. 12. DOI: 10.1007/s11306-016-1030-9.

[303] Vuckovic, D. (2012). "Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry", Analytical and Bioanalytical Chemistry, Vol. 403, Issue 6, pp. 1523–1548, 2012. DOI: 10.1007/s00216-012-6039-y.

[304] Ioannidis, J. P. A. (2008). "Why most discovered true associations are inflated", Epidemiology, Vol. 19, Issue 5, pp. 640–648. DOI: 10.1097/EDE.0b013e31818131e7.

[305] Ioannidis, J. P. A. (2008). "Effect of formal statistical significance on the credibility of observational Associations", American Journal of Epidemiology, Vol. 168, Issue 4, pp. 374–383. DOI: 10.1093/aje/kwn156.

[306] Rosenbaum, P. R. (2001). "Replicating effects and biases", American Statistician, Vol. 55, Issue 3, pp. 223–227. DOI: 10.1198/000313001317098220.

[307] Eshima, J., Davis, T. J., Bean, H. D., Fricks, J. and Smith, B. S. (2020). "A Metabolomic Approach for Predicting Diurnal Changes in Cortisol", Metabolites, Vol. 10, Issue 5. DOI: 10.3390/METABO10050194.

[308] Mastrangelo, A., Ferrarini, A., Rey-Stolle, F., García, A. and Barbas, C. (2015). "From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS", Analytica Chimica Acta, Vol. 900., pp. 21–35. DOI: 10.1016/j.aca.2015.10.001.

[309] Fiehn, O. (2016). "Metabolomics by Gas Chromatography–Mass Spectrometry: Combined Targeted and Untargeted Profiling", Current Protocols in Molecular Biology, Vol. 114, Issue 1, pp. 30.4.1-30.4.32. DOI: 10.1002/0471142727.mb3004s114.

[310] Gertsman, I. and Barshop, B. A. (2018). "Promises and Pitfalls of Untargeted Metabolomics", Journal of inherited metabolic disease, Vol. 41, Issue 3, p. 355. DOI: 10.1007/S10545-017-0130-7.

[311] Pezzatti, J., Boccard, J., Codesido, S., Gagnebin, Y., Joshi, A., et al. (2020). "Implementation of liquid chromatography–high resolution mass spectrometry methods for untargeted metabolomic analyses of biological samples: A tutorial", Analytica Chimica Acta, Vol. 1105., pp. 28–44. DOI: 10.1016/j.aca.2019.12.062.

[312] Manzi, M., Riquelme, G., Zabalegui, N. and Monge, M. E. (2020). "Improving diagnosis of genitourinary cancers: Biomarker discovery strategies through mass spectrometry-based metabolomics", Journal of Pharmaceutical and Biomedical Analysis, Vol. 178. DOI: 10.1016/j.jpba.2019.112905.

[313] Dettmer, K., Aronov, P. A. and Hammock, B. D. (2007). "Mass spectrometry-based metabolomics", Mass Spectrometry Reviews, Vol. 26, Issue 1, pp. 51–78. DOI: 10.1002/mas.20108.

[314] Monteiro, M. S., Carvalho, M., Bastos, M. L. and Guedes de Pinho, P. (2013). "Metabolomics Analysis for Biomarker Discovery: Advances and Challenges", Current Medicinal Chemistry, Vol. 20, Issue 2, pp. 257–271. DOI: 10.2174/092986713804806621.

[315] Park, Y., Kim, S. B., Wang, B., Blanco, R. A., Le, N.-A., et al. (2009). "Individual variation in macronutrient regulation measured by proton magnetic resonance spectroscopy of human plasma.", American journal of physiology. Regulatory, integrative and comparative physiology, Vol. 297, Issue 1, pp. 202-209. DOI: 10.1152/ajpregu.90757.2008.

[316] Ang, J. E., Revell, V., Mann, A., Mantele, S., Otway, D. T., et al. (2012). "Identification of human plasma metabolites exhibiting time-of-day variation using an untargeted liquid chromatography-mass spectrometry metabolomic approach.", Chronobiology international, Vol. 29, Issue 7, pp. 868–881. DOI: 10.3109/07420528.2012.699122.

[317] Dallmann, R., Viola, A. U., Tarokh, L., Cajochen, C. and Brown, S. A. (2012). "The human circadian metabolome", Proceedings of the National Academy of Sciences, Vol. 109, Issue 7, pp. 2625–2629. DOI: 10.1073/pnas.1114410109.

[318] Duffy, J. F. and Dijk, D. J. (2002). "Getting through to circadian oscillators: Why use constant routines?", Journal of Biological Rhythms, Vol. 17, Issue 1. pp. 4–13. DOI: 10.1177/074873002129002294.

[319] Kasukawa, T., Sugimoto, M., Hida, A., Minami, Y., Mori, M., et al. (2012). "Human blood metabolite timetable indicates internal body time.", Proceedings of the National Academy of Sciences of the United States of America, Vol. 109, Issue 37, pp. 15036–15041. DOI: 10.1073/pnas.1207768109.

[320] Chua, E. C. P., Shui, G., Lee, I. T. G., Lau, P., Tan, L. C., et al. (2013). "Extensive diversity in circadian regulation of plasma lipids and evidence for different circadian metabolic phenotypes in humans", Proceedings of the National Academy of Sciences of the United States of America, Vol. 110, Issue 35, pp. 14468–14473. DOI: 10.1073/pnas.1222647110.

[321] Davies, S. K., Ang, J. E., Revell, V. L., Holmes, B., Mann, A., et al. (2014). "Effect of sleep deprivation on the human metabolome", Proceedings of the National Academy of Sciences of the United States of America, Vol. 111, Issue 29, pp. 10761–10766. DOI: 10.1073/pnas.1402663111.

[322] Kim, K., Mall, C., Taylor, S. L., Hitchcock, S., Zhang, C., et al. (2014). "Mealtime, temporal, and daily variability of the human urinary and plasma metabolomes in a tightly controlled environment.", PloS one, Vol. 9, Issue 1. DOI: 10.1371/journal.pone.0086223.

[323] Chua, E. C.-P., Shui, G., Cazenave-Gassiot, A., Wenk, M. R. and Gooley, J. J. (2015). "Changes in Plasma Lipids during Exposure to Total Sleep Deprivation", Sleep, Vol. 38, Issue 11, pp. 1683–1691. DOI: 10.5665/sleep.5142.

[324] Skarke, C., Lahens, N. F., Rhoades, S. D., Campbell, A., Bittinger, K., et al. (2017). "A Pilot Characterization of the Human Chronobiome.", Scientific reports, Vol. 7, Issue 1, p. 17141. DOI: 10.1038/s41598-017-17362-6.

[325] Isherwood, C. M. (2016). "Metabolite and hormone rhythms : the effect of obesity type two diabetes and meal timing." ProQuest Dissertations Publishing, Ann Arbor, MI.

[326] Gehrman, P., Sengupta, A., Harders, E., Ubeydullah, E., Pack, A. I., et al. (2018). "Altered diurnal states in insomnia reflect peripheral hyperarousal and metabolic desynchrony: a preliminary study", SLEEP, Vol. 41, Issue 5. pp. 1–12. DOI: 10.1093/sleep/zsy043.

[327] Sato, S., Parr, E. B., Devlin, B. L., Hawley, J. A. and Sassone-Corsi, P. (2018). "Human metabolomics reveal daily variations under nutritional challenges specific to serum and skeletal muscle.", Molecular metabolism, Vol. 16, pp. 1–11. DOI: 10.1016/j.molmet.2018.06.008.

[328] Grant, L. K., Ftouni, S., Nijagal, B., De Souza, D. P., Tull, D., et al. (2019). "Circadian and wake-dependent changes in human plasma polar metabolites during prolonged wakefulness: A preliminary analysis", Scientific Reports, Vol. 9, Issue 1, pp. 1–14. DOI: 10.1038/s41598-019-40353-8.

[329] Gu, F., Klerman, E. B., Kim, S., Moore, S., Yu, K., et al. (2019). "Diurnal variation of metabolites in three individual participants", Chronobiology International, Vol. 36, Issue 3, pp. 332–342. DOI: 10.1080/07420528.2018.1541901.

[330] Kervezee, L., Cermakian, N. and Boivin, D. B. (2019). "Individual metabolomic signatures of circadian misalignment during simulated night shifts in humans.", PLoS biology, Vol. 17, Issue 6. DOI: 10.1371/journal.pbio.3000303.

[331] Lusczek, E. R., Parsons, L. S., Elder, J., Harvey, S. B., Skube, M., et al. (2020). "Metabolomics Pilot Study Identifies Desynchronization of 24-H Rhythms and Distinct Intra-patient Variability Patterns in Critical Illness: A Preliminary Report", Frontiers in Neurology, Vol. 11. DOI: 10.3389/fneur.2020.533915.

[332] Talasniemi, J. P., Pennanen, S., Savolainen, H., Niskanen, L. and Liesivuori, J. (2008). "Analytical investigation: Assay of d-lactate in diabetic plasma and urine", Clinical Biochemistry, Vol. 41, Issue 13, pp. 1099–1103. DOI: 10.1016/j.clinbiochem.2008.06.011.

[333] Acosta, S. and Nilsson, T. (2012). "Current status on plasma biomarkers for acute mesenteric ischemia", Journal of Thrombosis and Thrombolysis, Vol. 33, Issue 4, pp. 355–361. DOI: 10.1007/s11239-011-0660-z.

[334] Abbassi-Ghadi, N., Kumar, S., Huang, J., Goldin, R., Takats, Z., et al. (2013). "Metabolomic profiling of oesophago-gastric cancer: A systematic review", European Journal of Cancer, Vol. 49, Issue 17, pp. 3625–3637. DOI: 10.1016/j.ejca.2013.07.004.

[335] Ueland, P. M., Ulvik, A., Rios-Avila, L., Midttun, Ø. and Gregory, J. F. (2015). "Direct and Functional Biomarkers of Vitamin B6 Status", Annual Review of Nutrition, Vol. 35, Issue 1, pp. 33–70. DOI: 10.1146/annurev-nutr-071714-034330.

[336] Macrez, R., Stys, P. K., Vivien, D., Lipton, S. A. and Docagne, F. (2016). "Mechanisms of glutamate toxicity in multiple sclerosis: biomarker and therapeutic opportunities", The Lancet Neurology, Vol. 15, Issue 10, pp. 1089–1102. DOI: 10.1016/S1474-4422(16)30165-X.

[337] Boenzi, S. and Diodato, D. (2018). "Biomarkers for mitochondrial energy metabolism diseases", Essays in Biochemistry, Vol. 62, Issue 3, pp. 443–454. DOI: 10.1042/EBC20170111.

[338] Peoc'h, K. and Corcos, O. (2019). "Biomarkers for acute mesenteric ischemia diagnosis: State of the art and perspectives", Annales de Biologie Clinique, Vol. 77, Issue 4, pp. 415–421. DOI: 10.1684/abc.2019.1449.

[339] Takahashi, J. S., Hong, H. K., Ko, C. H. and McDearmon, E. L. (2008). "The genetics of mammalian circadian order and disorder: Implications for physiology and disease", Nature Reviews Genetics, Vol. 9, Issue 10, pp. 764–775. DOI: 10.1038/nrg2430.

[340] Watson, N. F., Buchwald, D. and Harden, K. P. (2013). "A twin study of genetic influences on diurnal preference and risk for alcohol use outcomes", Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, Vol. 9, Issue 12, pp. 1333–1339. DOI: 10.5664/JCSM.3282.

[341] Koskenvuo, M., Hublin, C., Partinen, M., Heikkilä, K. and Kaprio, J. (2007). "Heritability of diurnal type: a nationwide study of 8753 adult twin pairs", Journal of Sleep Research, Vol. 16, Issue 2, pp. 156–162. DOI: 10.1111/J.1365-2869.2007.00580.X.

[342] Barclay, N. L., Eley, T. C., Buysse, D. J., Archer, S. N. and Gregory, A. M. (2010). "Diurnal preference and sleep quality: same genes? A study of young adult twins", http://dx.doi.org/10.3109/07420521003663801, Vol. 27, Issue 2, pp. 278–296. DOI: 10.3109/07420521003663801.

[343] Von Schantz, M., Taporoski, T. P., Horimoto, A. R. V. R., Duarte, N. E., Vallada, H., et al. (2015). "Distribution and heritability of diurnal preference (chronotype) in a rural Brazilian family-based cohort, the Baependi study", Scientific Reports, Vol. 5. DOI: 10.1038/srep09214.

[344] Kelly, G. (2006). "Body Temperature Variability (Part 1): A Review of the History of Body Temperature and its Variability Due to Site Selection, Biological Rhythms, Fitness, and Aging", Alternative Medicine Review: A Journal of Clinical Therapeutic, Vol. 11, Issue 4, pp. 278-293.

[345] Kelly, G. S. (2007). "Body Temperature Variability (Part 2): Masking Influences of Body Temperature Variability and a review of body temperature variability in disease", Alternative Medicine Review: A Journal of Clinical Therapeutic, Vol. 12, Issue 1, pp. 49-62.

[346] Herljevic, M., Middleton, B., Thapan, K. and Skene, D. J. (2005). "Light-induced melatonin suppression: age-related reduction in response to short wavelength light", Experimental Gerontology, Vol. 40, Issue 3, pp. 237–242. DOI: 10.1016/J.EXGER.2004.12.001.

[347] Bailey, M. and Silver, R. (2014). "Sex differences in circadian timing systems: Implications for disease", Frontiers in Neuroendocrinology, Vol. 35, Issue 1, pp. 111–139. DOI: 10.1016/j.yfrne.2013.11.003.

[348] Anderson, S. T. and FitzGerald, G. A. (2020). "Sexual dimorphism in body clocks", Science, Vol. 369, Issue 6508, pp. 1164–1165. DOI: 10.1126/SCIENCE.ABD4964.

[349] Santosa, S. and Jensen, M. D. (2015). "The sexual dimorphism of lipid kinetics in humans", Frontiers in Endocrinology, Vol. 6. DOI: 10.3389/fendo.2015.00103.

[350] Palmer, B. F. and Clegg, D. J. (2015). "The sexual dimorphism of obesity", Molecular and Cellular Endocrinology, Vol. 402, pp. 113–119. DOI: 10.1016/j.mce.2014.11.029.

[351] Goossens, G. H., Jocken, J. W. E. and Blaak, E. E. (2021). "Sexual dimorphism in cardiometabolic health: the role of adipose tissue, muscle and liver", Nature Reviews Endocrinology, Vol. 17, Issue 1, pp. 47–66. DOI: 10.1038/s41574-020-00431-8.

[352] Bhake, R., Russell, G. M., Kershaw, Y., Stevens, K., Zaccardi, F., et al. (2020). "Continuous Free Cortisol Profiles in Healthy Men", The Journal of clinical endocrinology and metabolism, Vol. 105, Issue 4, pp. E1749–E1761. DOI: 10.1210/CLINEM/DGZ002.

[353] De Los Santos, H., Bennett, K. P. and Hurley, J. M. (2021). "MOSAIC: a joint modeling methodology for combined circadian and non-circadian analysis of multi-omics data", Bioinformatics (Oxford, England), Vol. 37, Issue 6, pp. 767–774. DOI: 10.1093/BIOINFORMATICS/BTAA877.

[354] Doherty, C. J. and Kay, S. A. (2010). "Circadian Control of Global Gene Expression Patterns", Annual Review of Genetics, Vol. 44, pp 419-444.. DOI: 10.1146/annurev-genet-102209-163432.

[355] Mure, L. S., Le, H. D., Benegiamo, G., Chang, M. W., Rios, L., et al. (2018). "Diurnal transcriptome atlas of a primate across major neural and peripheral tissues", Science, Vol. 359, Issue 6381. DOI: 10.1126/SCIENCE.AAO0318.

[356] Guo, J., Yu, H., Xing, S. and Huan, T. (2022). "Addressing big data challenges in mass spectrometry-based metabolomics", Chemical Communications, Vol. 58, Issue 72, pp. 9979–9990. DOI: 10.1039/D2CC03598G.

[357] O'Shea, K. and Misra, B. B. (2020). "Software tools, databases and resources in metabolomics: updates from 2018 to 2019", Metabolomics 2020 16:3, Vol. 16, Issue 3, pp. 1–23. DOI: 10.1007/S11306-020-01657-3.

[358] Vilela, B. and Villalobos, F. (2015). "letsR: a new R package for data handling and analysis in macroecology", Methods in Ecology and Evolution, Vol. 6, Issue 10, pp. 1229–1234. DOI: 10.1111/2041-210X.12401.

[359] Rizzo, A., Ferrante, D., Bagnara, S. (1995). "Handling Human Error", Hoc, P.C. Cacciabue, E. Hollnagel (Eds.) Expertise and Technology: Cognition & Human-Computer Cooperation, Lawrence Erlbaum Associates, Hillsdale, NJ, pp 195-212.

[360] Chen, Y., Xu, J., Zhang, R., Shen, G., Song, Y., et al. (2013). "Assessment of data pre-processing methods for LC-MS/MS-based metabolomics of uterine cervix cancer", The Analyst, Vol. 138, Issue 9, pp. 2669–2677. DOI: 10.1039/C3AN36818A.

[361] Mei, W., Jiang, Z., Chen, Y., Chen, L., Sancar, A., et al. (2021). "Genome-wide circadian rhythm detection methods: systematic evaluations and practical guidelines", Briefings in Bioinformatics, Vol. 22, Issue 3, pp. 1–13. DOI: 10.1093/BIB/BBAA135.

[362] Tong, X., Zhang, D., Arthurs, B., Li, P., Durudogan, L., et al. (2015). "Palmitate inhibits SIRT1-dependent BMAL1/CLOCK interaction and disrupts circadian gene oscillations in hepatocytes", PLoS ONE, Vol. 10, Issue 6. DOI: 10.1371/journal.pone.0130047.

[363] Loizides-Mangold, U., Perrin, L., Vandereycken, B., Betts, J. A., Walhin, J. P., et al. (2017). "Lipidomics reveals diurnal lipid oscillations in human skeletal muscle persisting in cellular myotubes cultured in vitro", Proceedings of the National Academy of Sciences of the United States of America, Vol. 114, Issue 41, pp. 8565–8574. DOI: 10.1073/pnas.1705821114.

[364] Abitbol, K., Debiesse, S., Molino, F., Mesirca, P., Bidaud, I., et al. (2017). "Clock-dependent and system-driven oscillators interact in the suprachiasmatic nuclei to pace mammalian circadian rhythms", PLoS ONE, Vol. 12, Issue 10. DOI: 10.1371/journal.pone.0187001.

[365] Wagner, P. M., Sosa Alderete, L. G., Gorné, L. D., Gaveglio, V., Salvador, G., et al. (2019). "Proliferative Glioblastoma Cancer Cells Exhibit Persisting Temporal Control of Metabolism

and Display Differential Temporal Drug Susceptibility in Chemotherapy", Molecular Neurobiology, Vol. 56, Issue 2, pp. 1276–1292. DOI: 10.1007/s12035-018-1152-3.

[366] Nobis, C. C., Laramée, G. D., Kervezee, L., de Sousa, D. M., Labrecque, N., et al. (2019). "The circadian clock of CD8 T cells modulates their early response to vaccination and the rhythmicity of related signaling pathways", Proceedings of the National Academy of Sciences of the United States of America, Vol. 116, Issue 40, pp. 20077–20086. DOI: 10.1073/PNAS.1905080116.

[367] Ruben, M. D., Francey, L. J., Guo, Y., Wu, G., Cooper, E. B., et al. (2019). "A large-scale study reveals 24-h operational rhythms in hospital treatment", Proceedings of the National Academy of Sciences of the United States of America, Vol. 116, Issue 42, pp. 20953–20958. DOI: 10.1073/PNAS.1909557116.

[368] Rabearivony, A., Li, H., Zhang, S., Chen, S., An, X., et al. (2020). "Housing temperature affects the circadian rhythm of hepatic metabolism and clock genes", Journal of Endocrinology, Vol. 247, Issue 2, pp. 183–195. DOI: 10.1530/JOE-20-0100.

[369] Egstrand, S., Mace, M. L., Morevati, M., Nordholm, A., Engelholm, L. H., et al. (2022). "Hypomorphic expression of parathyroid Bmal1 disrupts the internal parathyroid circadian clock and increases parathyroid cell proliferation in response to uremia", Kidney International, Vol. 101, Issue 6, pp. 1232–1250. DOI: 10.1016/J.KINT.2022.02.018.

[370] Berger, C. A. and Tarrant, A. M. (2023). "Sensory conflict disrupts circadian rhythms in the sea anemone Nematostella vectensis", eLife, Vol. 12. 2022.04.11.487933. DOI: 10.7554/eLife.81084.

[371] Zieliński, T., Hay, J. and Millar, A. J. (2022). "Period Estimation and Rhythm Detection in Timeseries Data Using BioDare2, the Free, Online, Community Resource", Methods in molecular biology (Clifton, N.J.), Vol. 2398, pp. 15–32. DOI: 10.1007/978-1-0716-1912-4_2.

[372] Glynn, E. F., Chen, J. and Mushegian, A. R. (2006). "Detecting periodic patterns in unevenly spaced gene expression time series using Lomb-Scargle periodograms", Bioinformatics (Oxford, England), Vol. 22, Issue 3, pp. 310–316. DOI: 10.1093/BIOINFORMATICS/BTI789.

[373] Beger, R. D., Dunn, W. B., Bandukwala, A., Bethan, B., Broadhurst, D., et al. (2019). "Towards quality assurance and quality control in untargeted metabolomics studies", Metabolomics, Vol. 15, Issue 1, pp. 1–5. DOI: 10.1007/S11306-018-1460-7.

[374] King, C. D., Kapp, K. L., Arul, A. B., Choi, M. J. and Robinson, R. A. S. (2022). "Advancements in automation for plasma proteomics sample preparation", Molecular Omics, Vol. 18, Issue 9, pp. 828–839. DOI: 10.1039/D2MO00122E.

[375] Vogeser, M. and Seger, C. (2008). "A decade of HPLC–MS/MS in the routine clinical laboratory — Goals for further developments", Clinical Biochemistry, Vol. 41, Issue 9, pp. 649–662. DOI: 10.1016/J.CLINBIOCHEM.2008.02.017.

[376] Jia, L., Yao, W., Jiang, Y., Li, Y., Wang, Z., et al. (2022). "Development of interactive biological web applications with R/Shiny", Briefings in Bioinformatics, Vol. 23, Issue 1. DOI: 10.1093/BIB/BBAB415.

[377] "Shiny" Accessed: Jan. 08, 2024. [Online]. Available: https://shiny.posit.co/

[378] Roenneberg, T., Daan, S. and Merrow, M. (2003). "The Art of Entrainment", Journal of Biological Rhythms. Vol. 18, Issue 3, pp. 183-194. DOI: 10.1177/0748730403253393.

[379] Merrow, M., Spoelstra, K. and Roenneberg, T. (2005). "The circadian cycle: daily rhythms from behaviour to genes", EMBO reports, Vol. 6, Issue 10, pp. 930–935. DOI: 10.1038/sj.embor.7400541.

[380] Duarte, L. L., Menna-Barreto, L., Miguel, M. A. L., Louzada, F., Araújo, J., et al. (2014). "Chronotype ontogeny related to gender", Brazilian Journal of Medical and Biological Research, Vol. 47, Issue 4, pp. 316–320. DOI: 10.1590/1414-431X20143001.

[381] Fischer, D., Lombardi, D. A., Marucci-Wellman, H. and Roenneberg, T. (2017). "Chronotypes in the US – Influence of age and sex", PLOS ONE, Vol. 12, Issue 6. DOI: 10.1371/journal.pone.0178782.

[382] Nowak, R., Mcmillen, I. C., Redman, J. and Short, R. V. (1987). "The correlation between serum and salivary melatonin concentrations and urinary 6-hydroxymelatonin sulphate excretion rates: two non-invasive techniques for monitoring human circadian rhythmicity", Clinical Endocrinology, Vol. 27, Issue 4, pp. 445–452. DOI: 10.1111/j.1365-2265.1987.tb01172.x.

[383] Klerman, E. B., Gershengorn, H. B., Duffy, J. F. and Kronauer, R. E. (2002). "Comparisons of the Variability of Three Markers of the Human Circadian Pacemaker", 2002. DOI: 10.1177/074873002129002474.

[384] Arendt, J. (2006). "Melatonin and human rhythms", Chronobiology International, Vol. 23, Issue 1-2, pp 21-37. DOI: 10.1080/07420520500464361.

[385] Gooley, J. J. (2016). "Circadian regulation of lipid metabolism.", The Proceedings of the Nutrition Society, Vol. 75, Issue 4, pp. 440–450. DOI: 10.1017/S0029665116000288.

[386] Panda, S. (2016). "Circadian physiology of metabolism", Science, Vol. 354, Issue 6315, pp. 1008–1015. DOI: 10.1126/science.aah4967.

[387] Hatori, M. and Panda, S. (2015). "Response of peripheral rhythms to the timing of food intake", in Methods in Enzymology, Vol. 552, pp. 145–161. DOI: 10.1016/bs.mie.2014.10.027.

[388] Wilkinson, M. J., Manoogian, E. N. C., Zadourian, A., Lo, H., Fakhouri, S., et al. (2020). "Ten-hour time-restricted eating reduces weight, blood pressure, and atherogenic lipids in patients with metabolic syndrome", Cell metabolism, Vol. 31, Issue 1, p. 92. DOI: 10.1016/J.CMET.2019.11.004.

[389] Cauter, E. Van,Golstein, J., Vanhaelst, L. and Leclercq, R. (1975). "Effects of Oral Contraceptive Therapy on the Orcadian Patterns of Cortisol and Thyrotropin (TSH)", European Journal of Clinical Investigation, Vol. 5, Issue 1, pp. 115–121. DOI: 10.1111/j.1365-2362.1975.tb00436.x.

[390] Webley, G. E. and Leidenberger, F. (1986). "The circadian pattern of melatonin and its positive relationship with progesterone in women", Journal of Clinical Endocrinology and Metabolism, Vol. 63, Issue 2. DOI: 10.1210/jcem-63-2-323.

[391] Hilli, J., Korhonen, T., Turpeinen, M., Hokkanen, J., Mattila, S., et al. (2008). "The effect of oral contraceptives on the pharmacokinetics of melatonin in healthy subjects with CYP1A2 g.-163C>A polymorphism", Journal of Clinical Pharmacology, Vol. 48, Issue 8. DOI: 10.1177/0091270008318669.

[392] Kostoglou-Athanassiou, I., Athanassiou, P., Treacher, D. F., Wheeler, M. J. and Forsling, M. L. (1998). "Neurohypophysial hormone and melatonin secretion over the natural and suppressed menstrual cycle in premenopausal women", Clinical Endocrinology, Vol. 49, Issue 2, pp. 209–216. DOI: 10.1046/J.1365-2265.1998.00504.X.

[393] Cain, S. W., Dennison, C. F., Zeitzer, J. M., Guzik, A. M., Khalsa, S. B. S., et al. (2010). "Sex differences in phase angle of entrainment and melatonin amplitude in humans", Journal of Biological Rhythms, Vol. 25, Issue 4. DOI: 10.1177/0748730410374943.

[394] Santhi, N., Lazar, A. S., McCabe, P. J., Lo, J. C., Groeger, J. A., et al. (2016). "Sex differences in the circadian regulation of sleep and waking cognition in humans", Proceedings of the National Academy of Sciences of the United States of America, Vol. 113, Issue 19, pp. E2730–E2739. DOI: 10.1073/PNAS.1521637113.

[395] Dunn, W. B., Lin, W., Broadhurst, D., Begley, P., Brown, M., et al. (2015). "Molecular phenotyping of a UK population: defining the human serum metabolome", Metabolomics, Vol. 11, Issue 1, p. 9. DOI: 10.1007/S11306-014-0707-1.

[396] Hutchison, A. L., Maienschein-Cline, M., Chiang, A. H., Tabei, S. M. A., Gudjonson, H., et al. (2015). "Improved Statistical Methods Enable Greater Sensitivity in Rhythm Detection for Genome-Wide Data", PLoS Computational Biology, Vol. 11, Issue 3, p. 1004094. DOI: 10.1371/JOURNAL.PCBI.1004094.

[397] Farshadi, E., van der Horst, G. T. J. and Chaves, I. (2020). "Molecular Links between the Circadian Clock and the Cell Cycle", Journal of Molecular Biology, Vol. 432, Issue 12, pp. 3515–3524. DOI: 10.1016/J.JMB.2020.04.003.

[398] Borgs, L., Beukelaers, P., Vandenbosch, R., Belachew, S., Nguyen, L., et al. (2009). "Cell circadian cycle: New role for mammalian core clock genes", Cell Cycle, Vol. 8, Issue 6, pp. 832–837. DOI: 10.4161/CC.8.6.7869.

[399] Guan, Q., Wang, Z., Cao, J., Dong, Y. and Chen, Y. (2022). "Mechanisms of Melatonin in Obesity: A Review", International Journal of Molecular Sciences, Vol. 23, Issue 1. DOI: 10.3390/IJMS23010218.

[400] Samovski, D., Jacome-Sosa, M. and Abumrad, N. A. (2023). "Fatty Acid Transport and Signaling: Mechanisms and Physiological Implications", https://doi.org/10.1146/annurev-physiol-032122-030352, Vol. 85, pp. 317–337. DOI: 10.1146/ANNUREV-PHYSIOL-032122-030352.

[401] Gong, Y., Dou, L. J. and Liang, J. (2014). "Link between obesity and cancer: role of triglyceride/free fatty acid cycling", European review for medical and pharmacological sciences, Vol. 18, Issue 19, pp. 2808–2820.

[402] Solis-Herrera, C., Triplitt, C., Cersosimo, E. and DeFronzo, R. A. (2021). "Pathogenesis of Type 2 Diabetes Mellitus", K.R. Feingold, B. Anawalt, M.R. Blackman, et al., (eds), Endotext [Internet], MDText.com, Inc. South Dartmouth, MA.

[403] Cai, Y., Ren, L., Tan, S., Liu, X., Li, C., et al. (2022). "Mechanism, diagnosis, and treatment of cyclic Cushing's syndrome: A review", Biomedicine & Pharmacotherapy, Vol. 153, p. 113301. DOI: 10.1016/J.BIOPHA.2022.113301.

[404] Raff, H. and Carroll, T. (2015). "Cushing's syndrome: from physiological principles to diagnosis and clinical care", The Journal of Physiology, Vol. 593, Issue Pt 3, p. 493. DOI: 10.1113/JPHYSIOL.2014.282871.

[405] Arnaldi, G., Scandali, V. M., Trementino, L., Cardinaletti, M., Appolloni, G., et al. (2010). "Pathophysiology of dyslipidemia in Cushing's syndrome", Neuroendocrinology, Vol. 92 Suppl 1, Issue SUPPL. 1, pp. 86–90. DOI: 10.1159/000314213.

[406] Vega-Beyhart, A., Iruarrizaga, M., Pané, A., García-Eguren, G., Giró, O., et al. (2021). "Endogenous cortisol excess confers a unique lipid signature and metabolic network", Journal of molecular medicine (Berlin, Germany), Vol. 99, Issue 8, pp. 1085–1099. DOI: 10.1007/S00109-021-02076-0.

[407] Salehidoost, R. and Korbonits, M. (2022). "Glucose and lipid metabolism abnormalities in Cushing's syndrome", Journal of Neuroendocrinology, Vol. 34, Issue 8, p. e13143. DOI: 10.1111/JNE.13143.

[408] Lioupi, A., Marinaki, M., Virgiliou, C., Begou, O., Gika, H., et al. (2023). "Probing the polar metabolome by UHPLC-MS", TrAC Trends in Analytical Chemistry, Vol. 161, p. 117014. DOI: 10.1016/J.TRAC.2023.117014.

[409] Quehenberger, O. and Dennis, E. A. (2011). "The Human Plasma Lipidome", The New England journal of medicine, Vol. 365, Issue 19, p. 1812. DOI: 10.1056/NEJMRA1104901.

[410] Quehenberger, O., Armando, A. M., Brown, A. H., Milne, S. B., Myers, D. S., et al. (2010). "Lipidomics reveals a remarkable diversity of lipids in human plasma1", Journal of Lipid Research, Vol. 51, Issue 11, pp. 3299–3305. DOI: 10.1194/jlr.M009449.

[411] Alves-Bezerra, M. and Cohen, D. E. (2017). "Triglyceride metabolism in the liver", Comprehensive Physiology, Vol. 8, Issue 1, p. 1. DOI: 10.1002/CPHY.C170012.

[412] Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. and Hogenesch, J. B. (2014). "A circadian gene expression atlas in mammals: Implications for biology and medicine", Proceedings of the National Academy of Sciences of the United States of America, Vol. 111, Issue 45, pp. 16219–16224. DOI: 10.1073/pnas.1408886111.

[413] Al-Sari, N., Suvitaival, T., Mattila, I., Ali, A., Ahonen, L., et al. (2020). "Lipidomics of human adipose tissue reveals diversity between body areas", PLoS ONE, Vol. 15, Issue 6. DOI: 10.1371/JOURNAL.PONE.0228521.

[414] Morris, S. M. (2016). "Arginine Metabolism Revisited", The Journal of Nutrition, Vol. 146, Issue 12, pp. 2579S-2586S. DOI: 10.3945/JN.115.226621.

[415] Dugger, B. N. and Dickson, D. W. (2017). "Pathology of Neurodegenerative Diseases", Cold Spring Harbor Perspectives in Biology, Vol. 9, Issue 7. DOI: 10.1101/CSHPERSPECT.A028035.

[416] Maszka, P., Kwasniak-Butowska, M., Cysewski, D., Slawek, J., Smolenski, R. T., et al. (2023). "Metabolomic Footprint of Disrupted Energetics and Amino Acid Metabolism in Neurodegenerative Diseases: Perspectives for Early Diagnosis and Monitoring of Therapy", Metabolites, Vol. 13, Issue 3, p. 369. DOI: 10.3390/METABO13030369.

[417] McColgan, P. and Tabrizi, S. J. (2018). "Huntington's disease: a clinical review", European Journal of Neurology, Vol. 25, Issue 1, pp. 24–34. DOI: 10.1111/ENE.13413.

[418] Seneca, S., Fagnart, D., Keymolen, K., Lissens, W., Hasaerts, D., et al. (2004). "Early onset Huntington disease: A neuronal degeneration syndrome", European Journal of Pediatrics, Vol. 163, Issue 12, pp. 717–721. DOI: 10.1007/s00431-004-1537-3.

[419] Brinkman, R. R., Mezei, M. M., Theilmann, J., Almqvist, E. and Hayden, M. R. (1997). "The likelihood of being affected with huntington disease by a particular age, for a specific CAG size", American Journal of Human Genetics, Vol. 60, Issue 5, pp. 1202–1210.

[420] Erkkinen, M. G., Kim, M. O. and Geschwind, M. D. (2018). "Clinical Neurology and Epidemiology of the Major Neurodegenerative Diseases", Cold Spring Harbor Perspectives in Biology, Vol. 10, Issue 4. DOI: 10.1101/CSHPERSPECT.A033118.

[421] Saudou, F. and Humbert, S. (2016). "The Biology of Huntingtin", Neuron, Vol. 89, Issue 5, pp. 910–926. DOI: 10.1016/J.NEURON.2016.02.003.

[422] Martin, D. D. O., Schmidt, M. E., Nguyen, Y. T., Lazic, N. and Hayden, M. R. (2019). "Identification of a novel caspase cleavage site in huntingtin that regulates mutant huntingtin clearance", FASEB Journal, Vol. 33, Issue 3, pp. 3190–3197. DOI: 10.1096/FJ.201701510RRR.

[423] O'Regan, G. C., Farag, S. H., Ostroff, G. R., Tabrizi, S. J. and Andre, R. (2020). "Wild-type huntingtin regulates human macrophage function", Scientific reports, Vol. 10, Issue 1. DOI: 10.1038/S41598-020-74042-8.

[424] Shirasaki, D. I., Greiner, E. R., Al-Ramahi, I., Gray, M., Boontheung, P., et al. (2012). "Network Organization of the Huntingtin Proteomic Interactome in Mammalian Brain", Neuron, Vol. 75, Issue 1, p. 41. DOI: 10.1016/J.NEURON.2012.05.024.

[425] Patassini, S., Begley, P., Xu, J., Church, S. J., Reid, S. J., et al. (2016). "Metabolite mapping reveals severe widespread perturbation of multiple metabolic processes in Huntington's disease human brain", Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, Vol. 1862, Issue 9, pp. 1650–1662. DOI: 10.1016/J.BBADIS.2016.06.002.

[426] Scholefield, M., Unwin, R. D. and Cooper, G. J. S. (2020). "Shared perturbations in the metallome and metabolome of Alzheimer's, Parkinson's, Huntington's, and dementia with Lewy bodies: A systematic review", Ageing Research Reviews, Vol. 63, p. 101152. DOI: 10.1016/J.ARR.2020.101152.

[427] Schwarcz, R., Bruno, J. P., Muchowski, P. J. and Wu, H. Q. (2012). "Kynurenines in the mammalian brain: when physiology meets pathology", Nature reviews. Neuroscience, Vol. 13, Issue 7, p. 465. DOI: 10.1038/NRN3257.

[428] Sapko, M. T., Guidetti, P., Yu, P., Tagle, D. A., Pellicciari, R., et al. (2006). "Endogenous kynurenate controls the vulnerability of striatal neurons to quinolinate: Implications for Huntington's disease", Experimental Neurology, Vol. 197, Issue 1, pp. 31–40. DOI: 10.1016/J.EXPNEUROL.2005.07.004.

[429] Castellano-Gonzalez, G., Jacobs, K. R., Don, E., Cole, N. J., Adams, S., et al. (2019). "Kynurenine 3-Monooxygenase Activity in Human Primary Neurons and Effect on Cellular Bioenergetics Identifies New Neurotoxic Mechanisms", Neurotoxicity research, Vol. 35, Issue 3, pp. 530–541. DOI: 10.1007/S12640-019-9997-4.

[430] Morton, A. J., Wood, N. I., Hastings, M. H., Hurelbrink, C., Barker, R. A., et al. (2005). "Disintegration of the Sleep-Wake Cycle and Circadian Timing in Huntington's Disease", The Journal of Neuroscience, Vol. 25, Issue 1, p. 157. DOI: 10.1523/JNEUROSCI.3842-04.2005.

[431] Kudo, T., Schroeder, A., Loh, D. H., Kuljis, D., Jordan, M. C., et al. (2011). "Dysfunctions in circadian behavior and physiology in mouse models of Huntington's disease", Experimental Neurology, Vol. 228, Issue 1, pp. 80–90. DOI: 10.1016/J.EXPNEUROL.2010.12.011.

[432] Stoker, T. B., Mason, S. L., Greenland, J. C., Holden, S. T., Santini, H., et al. (2022). "Huntington's disease: diagnosis and management", Practical Neurology, Vol. 22, Issue 1, pp. 32–41. DOI: 10.1136/PRACTNEUROL-2021-003074.

[433] Pouladi, M. A., Morton, A. J. and Hayden, M. R. (2013). "Choosing an animal model for the study of Huntington's disease", Nature Reviews Neuroscience 2013 14:10, Vol. 14, Issue 10, pp. 708–721. DOI: 10.1038/nrn3570.

[434] Morton, A. J. (2018). "Large-brained animal models of Huntington's disease: Sheep", Methods in Molecular Biology, Vol. 1780, pp. 221–239. DOI: 10.1007/978-1-4939-7825-0_12.

[435] Jennifer Morton, A., Rudiger, S. R., Wood, N. I., Sawiak, S. J., Brown, G. C., et al. (2014). "Early and progressive circadian abnormalities in Huntington's disease sheep are unmasked by social environment", Human Molecular Genetics, Vol. 23, Issue 13, pp. 3375–3383. DOI: 10.1093/HMG/DDU047.

[436] Nassan, M. and Videnovic, A. (2021). "Circadian rhythms in neurodegenerative disorders", Nature Reviews Neurology 2021 18:1, Vol. 18, Issue 1, pp. 7–24. DOI: 10.1038/s41582-021-00577-7.

[437] Khachiyants, N., Trinkle, D., Son, S. J. and Kim, K. Y. (2011). "Sundown Syndrome in Persons with Dementia: An Update", Psychiatry Investigation, Vol. 8, Issue 4, p. 275. DOI: 10.4306/PI.2011.8.4.275.

[438] Hoyt, K. R. and Obrietan, K. (2022). "Circadian clocks, cognition, and Alzheimer's disease: synaptic mechanisms, signaling effectors, and chronotherapeutics", Molecular Neurodegeneration, Vol. 17, Issue 1, p. 35. DOI: 10.1186/S13024-022-00537-9.

[439] Morton, A. J., Middleton, B., Rudiger, S., Bawden, C. S., Kuchel, T. R., et al. (2020). "Increased plasma melatonin in presymptomatic Huntington disease sheep (Ovis aries): Compensatory neuroprotection in a neurodegenerative disease?", Journal of pineal research, Vol. 68, Issue 2. DOI: 10.1111/JPI.12624.

[440] Morton, A. J. (2013). "Circadian and sleep disorder in Huntington's disease", Experimental Neurology, Vol. 243, pp. 34–44. DOI: 10.1016/J.EXPNEUROL.2012.10.014.

[441] Dufour, B. D. and McBride, J. L. (2016). "Corticosterone dysregulation exacerbates disease progression in the R6/2 transgenic mouse model of Huntington's disease", Experimental neurology, Vol. 283, Issue Pt A, p. 308. DOI: 10.1016/J.EXPNEUROL.2016.06.028.

[442] Saleh, N., Moutereau, S., Durr, A., Krystkowiak, P., Azulay, J. P., et al. (2009). "Neuroendocrine disturbances in Huntington's disease", PloS one, Vol. 4, Issue 3. DOI: 10.1371/JOURNAL.PONE.0004962.

[443] Adamczak-Ratajczak, A., Kupsz, J., Owecki, M., Zielonka, D., Sowinska, A., et al. (2017). "Circadian rhythms of melatonin and cortisol in manifest huntington's disease and in acute

cortical ischemic stroke", Journal of Physiology and Pharmacology, Vol. 68, Issue 4. DOI: 10.1186/s40101-016-0103-9.

[444] Aziz, N. A., Pijl, H., Frölich, M., Van Maurits Der Graaf, A. W., Roelfsema, F., et al. (2009). "Increased hypothalamic-pituitary-adrenal axis activity in Huntington's disease", The Journal of clinical endocrinology and metabolism, Vol. 94, Issue 4, pp. 1223–1228. DOI: 10.1210/JC.2008-2543.

[445] Kalliolia, E., Silajdžic, E., Nambron, R., Costelloe, S. J., Martin, N. G., et al. (2015). "A 24-Hour Study of the Hypothalamo-Pituitary Axes in Huntington's Disease", PLoS ONE, Vol. 10, Issue 10. DOI: 10.1371/JOURNAL.PONE.0138848.

[446] Shirbin, C. A., Chua, P., Churchyard, A., Hannan, A. J., Lowndes, G., et al. (2013). "The relationship between cortisol and verbal memory in the early stages of Huntington's disease", Journal of Neurology, Vol. 260, Issue 3, pp. 891–902. DOI: 10.1007/S00415-012-6732-Y.

[447] Aziz, N. A., Pijl, H., Frölich, M., Schröder-Van Der Elst, J. P., Van Der Bent, C., et al. (2009). "Delayed onset of the diurnal melatonin rise in patients with Huntington's disease", Journal of Neurology, Vol. 256, Issue 12, pp. 1961–1965. DOI: 10.1007/S00415-009-5196-1.

[448] Chuang, C. L. and Demontis, F. (2021). "Systemic manifestation and contribution of peripheral tissues to Huntington's disease pathogenesis", Ageing research reviews, Vol. 69, p. 101358. DOI: 10.1016/J.ARR.2021.101358.

[449] Stüwe, S. H., Goetze, O., Lukas, C., Klotz, P., Hoffmann, R., et al. (2013). "Hepatic mitochondrial dysfunction in manifest and premanifest Huntington disease", Neurology, Vol. 80, Issue 8, pp. 743–746. DOI: 10.1212/WNL.0B013E318282514E.

[450] Acosta-Rodríguez, V. A., Márquez, S., Salvador, G. A., Pasquaré, S. J., Gorné, L. D., et al. (2013). "Daily rhythms of glycerophospholipid synthesis in fibroblast cultures involve differential enzyme contributions", Journal of lipid research, Vol. 54, Issue 7, pp. 1798–1811. DOI: 10.1194/JLR.M034264.

[451] Mastrokolias, A., Pool, R., Mina, E., Hettne, K. M., van Duijn, E., et al. (2016). "Integration of targeted metabolomics and transcriptomics identifies deregulation of phosphatidylcholine metabolism in Huntington's disease peripheral blood samples", Metabolomics : Official journal of the Metabolomic Society, Vol. 12, Issue 8. DOI: 10.1007/S11306-016-1084-8.

[452] Cheng, M. L., Chang, K. H., Wu, Y. R. and Chen, C. M. (2016). "Metabolic disturbances in plasma as biomarkers for Huntington's disease", The Journal of nutritional biochemistry, Vol. 31, pp. 38–44. DOI: 10.1016/J.JNUTBIO.2015.12.001.

[453] Law, S. H., Chan, M. L., Marathe, G. K., Parveen, F., Chen, C. H., et al. (2019). "An Updated Review of Lysophosphatidylcholine Metabolism in Human Diseases", International journal of molecular sciences, Vol. 20, Issue 5. DOI: 10.3390/IJMS20051149.

[454] Casella, C., Lipp, I., Rosser, A., Jones, D. K. and Metzler-Baddeley, C. (2020). "A Critical Review of White Matter Changes in Huntington's Disease", Movement disorders : official journal of the Movement Disorder Society, Vol. 35, Issue 8, pp. 1302–1311. DOI: 10.1002/MDS.28109.

[455] Zalewska, T., Pawelec, P., Ziabska, K. and Ziemka-Nalecz, M. (2023). "Sexual Dimorphism in Neurodegenerative Diseases and in Brain Ischemia", Biomolecules, Vol. 13, Issue 1. DOI: 10.3390/BIOM13010026.

[456] Ullah, M. F., Ahmad, A., Bhat, S. H., Abu-Duhier, F. M., Barreto, G. E., et al. (2019). "Impact of sex differences and gender specificity on behavioral characteristics and pathophysiology of neurodegenerative disorders", Neuroscience & Biobehavioral Reviews, Vol. 102, pp. 95–105. DOI: 10.1016/J.NEUBIOREV.2019.04.003.

[457] Jerjes, W. K., Cleare, A. J., Peters, T. J. and Taylor, N. F. (2006). "Circadian rhythm of urinary steroid metabolites", Annals of Clinical Biochemistry, Vol. 43, Issue 4, pp. 287–294. DOI: 10.1258/000456306777695564.

[458] Walsh, M. C., Brennan, L., Malthouse, J. P. G., Roche, H. M. and Gibney, M. J. (2006). "Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans.", The American journal of clinical nutrition, Vol. 84, Issue 3, pp. 531–539. DOI: 10.1093/ajcn/84.3.531.

[459] Slupsky, C. M., Rankin, K. N., Wagner, J., Fu, H., Chang, D., et al. (2007). "Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles.", Analytical chemistry, Vol. 79, Issue 18, pp. 6995–7004. DOI: 10.1021/ac0708588.

[460] Giskeodegard, G. F., Davies, S. K., Revell, V. L., Keun, H. and Skene, D. J. (2015). "Diurnal rhythms in the human urine metabolome during sleep and total sleep deprivation.", Scientific reports, Vol. 5, p. 14843. DOI: 10.1038/srep14843.

[461] Papantoniou, K., Pozo, O. J., Espinosa, A., Marcos, J., Castaño-Vinyals, G., et al. (2015). "Increased and mistimed sex hormone production in night shift workers", Cancer Epidemiology Biomarkers and Prevention, Vol. 24, Issue 5, pp. 854–863. DOI: 10.1158/1055-9965.EPI-14-1271.

[462] Sinues, P. M.-L., Kohler, M. and Zenobi, R. (2013). "Monitoring diurnal changes in exhaled human breath.", Analytical chemistry, Vol. 85, Issue 1, pp. 369–373. DOI: 10.1021/ac3029097.

[463] Martinez-Lozano Sinues, P., Tarokh, L., Li, X., Kohler, M., Brown, S. A., et al. (2014). "Circadian variation of the human metabolome captured by real-time breath analysis.", PloS one, Vol. 9, Issue 12. DOI: 10.1371/journal.pone.0114422.

[464] Wilkinson, M., Maidstone, R., Loudon, A., Blaikley, J., White, I. R., et al. (2019). "Circadian rhythm of exhaled biomarkers in health and asthma", The European respiratory journal, Vol. 54, Issue 4. DOI: 10.1183/13993003.01068-2019.

[465] Held, N. M., Wefers, J., van Weeghel, M., Daemen, S., Hansen, J., et al. (2020). "Skeletal muscle in healthy humans exhibits a day-night rhythm in lipid metabolism.", Molecular metabolism, Vol. 37. DOI: 10.1016/j.molmet.2020.100989.

[466] Cardinali, D.P., Brown, G.M., Pandi-Perumal, S.R. (2021). "Chronotherapy", Handbook of Clinical Neurology, Vol. 179, pp. 357–370. DOI: 10.1016/B978-0-12-819975-6.00023-6

[467] Sancar, A., Van Gelder, R.N. (2021). "Clocks, cancer, and chronochemotherapy", Science, Vol. 371, Issue 6524. DOI: 10.1126/science.abb0738

[468] Lee, Y., Field, J.M., Sehgal, A. (2021). "Circadian rhythms, disease and chronotherapy", Journal of biological rhythms, Vol. 36, issue 6, pp. 503–531. DOI: 10.1177/07487304211044301

[469] Wittenbrink, N., Ananthasubramaniam, B., Münch, M., Koller, B., Maier, B., *et al*. (2018). "High-accuracy determination of internal circadian time from a single blood sample", The Journal of Clinical Investigation, Vol. 128, Issue 9. DOI: 10.1172/JCI120874

[470] Rahman, S.A., Kayumov, L., Tchmoutina, E.A., Shapiro, C.M. (2009). “Clinical efficacy of dim light melatonin onset testing in diagnosing delayed sleep phase syndrome”, Sleep Medicine, Vol. 10, Issue 5, pp. 549-555. DOI: 10.1016/j.sleep.2008.03.020

[471] Ueda, H.R., Chen, W., Minami, Y., Honma, S., Honma, K., *et al*. (2004). “Molecular-timetable methods for detection of body time and rhythm disorders from single-time-point genome-wide expression profiles”, PNAS, Vol. 101, Issue 31, pp. 11227–11232. DOI: 10.1073/pnas.0401882101

[472] Hughey, J.J., Hastie, T., Butte, A.J. (2016). “ZeitZeiger: supervised learning for high-dimensional data from an oscillatory system”, Nucleic acids research, Vol. 44, Issue 8. DOI: 10.1093/nar/gkw030

[473] Vlachou, D., Veretennikova, M., Usselmann, L., Vasilyev, V., Ott, S., *et al*. (2023). “TimeTeller: A tool to probe the circadian clock as a multigene dynamical system”, [online]. [Preprint] bioRxiv. [Viewed 11/01/24]. DOI: 10.1101/2023.03.14.532177

APPENDIX

Appendix Section A – Summary tables for metabolomic studies investigating time of day variation reproduced from [78].

Table A1: A brief summary of study design, cohort details and results with regard to observed time of day variation of metabolites for relevant studies analysing urine.

Author(s)	Assay/Platform	Time Course Details	Study Setting/Conditions	Cohort Details	Rhythmic/Gradient Metabolites/Features Observed	Rhythmic/Gradient Classes Primarily Observed
Jerjes *et al* (2006) [457]	Targeted GC-MS	Diurnal variation (24 h–3 h intervals between samples)		N = 20 10 males Age 32 ± 5.4 BMI 23.5 ± 2	9	Androgens Cortisol metabolites
Walsh *et al* (2006) [458]	Untargeted ^{1}H NMR	Diurnal variation am *vs.* pm	'Outpatient' Standardised meals	N = 60 30 males Age 19–69	1	
Slupsky *et al* (2007) [459]	Targeted ^{1}H NMR	Diurnal variation am *vs.* pm	'Outpatient'	N = 30 23 females Age 24.7 ± 2.7 BMI 22.7 ± 0.97	6	
Kim *et al* (2014) [322]	Untargeted LC—TOF MS (Reversed Phase)	Diurnal variation Sampling 1, 3, 7, 9, 11, 14 h post-wake, first sample fasted.	'Inpatient' Standardised meals and mealtimes	N = 26 14 males Age 33 ± 10.9 BMI 24.3 ± 3.3	135 (46%)	Glycerophospholipids LysoPCs Phosphatidylinositol
Giskeødegård *et al* (2015) [460]	Untargeted ^{1}H NMR	Diurnal variation (48 h) Samples at 2–4 h intervals when awake, 8 h overnight	'Inpatient' Standardised meals and mealtimes. Controlled light/dark cycle and activity/posture	N = 15 All male Age 23.7 ± 5.4	5 (15.6%)—sleep/wake cycle 7 (22%) during 24 h wakefulness During sleep deprivation 8 increased, 8 decreased	Amino acids Fatty acids
Papantoniou *et al* (2015) [461]	Targeted GC-MS	Diurnal variation (24 h)	'Outpatient' Day *vs.* night shift workers	N = 117 63 males Age 22–64 BMI 22.6–30.6	5 (31.3%) significantly different in premenopausal day *vs.* night workers	Progestagens Androgens

Footnotes: Footnotes: Age and BMI are quoted in standard units, years and kg/m^2 , respectively. Where available, mean age/BMI ± 1 SD given. Significant changes in metabolites identified in studies performing am *vs.* pm (two-time point) comparison(s), should be considered as gradient changes ergo 'gradient metabolite', Significant changes in metabolites identified in studies over a >24 h time course with $n \geq 5$ samples should be considered as rhythmic changes ergo 'rhythmic metabolite', rhythmicity being detected by cosinor analysis and/or MetaCycle. Rhythmic/gradient features are denoted as such, otherwise the table refers to rhythmic/gradient metabolites.

Rhythmic/gradient classes primarily observed are not an exhaustive list of all metabolite classes observed within a study but a summary of the most rhythmic classes, if any, for that particular study, as denoted by the author or inferred from provided data. Abbreviations: GC-MS—gas chromatography mass spectrometry; LC-MS—liquid chromatography mass spectrometry; LysoPC—lysophosphatidylcholine; NMR—nuclear magnetic resonance; TOF MS—time of flight mass spectrometry.

Table A2: A brief summary of study design, cohort details and results with regard to observed time of day variation of metabolites for relevant studies analysing saliva.

Authors	Assay/Platform	Time course details	Study setting/conditions	Cohort details	Rhythmic/gradient metabolites/features observed	Rhythmic/gradient classes primarily observed
Walsh *et al* (2006) [458]	Untargeted ^{1}H NMR	Diurnal variation am *vs.* pm	'Outpatient' Standardised meals	N = 60 30 males 30 females Age 19–69	1	No gradient metabolite classes identified
Dallmann *et al* (2012) [317]	Untargeted GC-MS LC-MS (Reversed Phase)	Circadian variation (constant routine 40 h, 4 h intervals between samples)	'Inpatient' Standard constant routine parameters (see Duffy and Dijk, 2002)	N = 10 (split into 2 equal groups within which samples were pooled for each 4 h interval) All male Age 57.8 ± 1.0 & 61.0 ± 0.6 BMI 26.6 ± 0.6 & 25.1 ± 0.5	29 (15%)	Amino acids
Dame *et al* (2015) [216]	Untargeted ^{1}H NMR	Diurnal variation sampling at prebreakfast *vs.* 2 h post-breakfast *vs.* 2 h post-lunch		N = 16 8 males & females Age (24–42) (only N = 2 took part in observation of diurnal variation)	8 (10.5%)	Amino acids
Skarke *et al* (2017) [324]	Targeted LC-MS/MS (HILIC)	Diurnal variation am *vs.* pm (48 h, 5 samples 12 h apart)	'Outpatient'	N = 6 All male Age 32.3 ± 3.6 BMI 25.2 ± 3.4	14 (5.6%)	Amino acids

Footnotes: See Table A1. Abbreviations: GC-MS—gas chromatography mass spectrometry; HILIC—hydrophilic interaction chromatography; LC-MS—liquid chromatography mass spectrometry; MS/MS—tandem mass spectrometry; NMR—nuclear magnetic resonance.

Table A3: A brief summary of study design, cohort details and results with regard to observed time of day variation of metabolites for relevant studies analysing breath.

Authors	Assay/Platform	Time Course Details	Study Setting/Conditions	Cohort Details	Rhythmic/Gradient Metabolites/Features Observed	Rhythmic/Gradient Classes Primarily Observed
Sinues *et al* (2012) [462]	Untargeted SESI-MS	Diurnal variation (4 time periods) 8:00−11:00, 11:00−13:00, 13:00−15:00, 15:00−18:00	'Outpatient'	N = 12 7 males 5 females	Diurnal changes observed but number of rhythmic features not reported	No metabolites structurally identified
Sinues *et al* (2014) [463]	Untargeted SESI-MS	Diurnal variation (24 h, 1 h intervals, 5–7 repeats per sample)	'Inpatient' Controlled laboratory conditions: hourly isocaloric meals, constant wakefulness, consistent light conditions	N = 3 2 males 1 female Age 33–38	40 (36%) of features (49% in N = 1)	No metabolites structurally identified
Wilkinson *et al* (2019) [464]	Untargeted GC-MS	Diurnal variation (24 h—4 time points: 16:00, 22:00, 04:00, 10:00)	Standardised meals and feeding schedule. Maintained habitual bedtime	Healthy N = 10 7 males 3 females Age 27.5–49.3 BMI 23.4–30.5 Asthma N = 9 7 males Age 26.0–49.5 BMI 22.3–27.2	Combined dataset 5/102 (4.9%) metabolites Asthma 3/102 (~2.9%) metabolites, 1 of which is unique to this group in addition to rhythmicity of exhaled nitric oxide fraction Healthy 2/102 (~2%) metabolites rhythmic and unique to this group	Volatile organic compounds

Footnotes: See Table A1. Abbreviations: GC-MS—gas chromatography mass spectrometry; SESI-MS – secondary electrospray ionisation.

Table A4: A brief summary of study design, cohort details and results with regard to observed time of day variation of metabolites for relevant studies analysing skeletal muscle.

Authors	Performed Assay	Time Course Details	Study Setting/Conditions	Cohort Details	Rhythmic/Gradient Metabolites/Features Observed	Rhythmic/Gradient Classes Primarily Observed
Loizides-Mangold *et al* (2017) [363]	Targeted (Lipidomics) LC-MS	Diurnal variation (24 h–4 h intervals between samples)	'Inpatient' Controlled sleep/wake, light/dark cycle, temperature. Isocaloric meals	N = 10, 9 males Age 29.9 ± 9.8 BMI 24.1 ± 2.7	106 of 1058 metabolites (10%)	TGs, PCs, PEs PIs, PSs, CLs Cers, GlcCers, SMs
Sato *et al* (2018) [327]	Untargeted UHPLC-MS/MS GC-MS	Diurnal variation am *vs.* pm	'Outpatient' Standardised meals and mealtimes	N = 8, All male Age 30–45 BMI 27–32.5	163 & 19 of 625 features (26% & 3%) as a result of time of day & diet, respectively	Amino acids Fatty acyls Glycerolipids Glycerophospholipids Sphingolipids Carbohydrates Xenobiotics
Held *et al* (2020) [465]	Semi-targeted Lipidomics UPLC/HRMS (reversed & normal phase)	Diurnal variation (24 h—5 h intervals between samples)	'Inpatient' Controlled sleep/wake, light/dark cycle. Standardised meals and mealtimes	N = 12, All male Age 22.2 ± 2.3 BMI 22.4 ± 2.0	126 of 971 (13%)	Glycerophospholipids TGs Sphingolipids DGs Sterol Lipids

Footnotes: See Table A1. Abbreviations: Cer—ceramide; CL—cardiolipin; DG—diglyceride; GC-MS—gas chromatography mass spectrometry; GlcCer—glucosylceramide; HRMS—high resolution mass spectrometry; LC-MS—liquid chromatography mass spectrometry; MS/MS—tandem mass spectrometry; PC—phosphatidylcholine; Pe—phosphatidylethanolamine; PI—phosphatidylinositol; PS—phosphatidylserine;; SM—sphingomyelin; TG—triglyceride; U(H)PLC—ultra high performance liquid chromatography.

Table A5: Putatively identified metabolites, observed in five or more human metabolomics time course studies, that underwent significant time of day variation (rhythmic/gradient metabolites) in ranked order.

Rank	Putative Identification of Rhythmic/Gradient Metabolites	InChIKey	Number of Studies Significant Changes were Observed in
1	Proline	ONIBWKKTOPOVIA-BYPYZUCNSA-N	11
2	Leucine	ROHFNLRQFUQHCH-YFKPBYRVSA-N	10
3	PC(32:0)	-	10
4	Phenylalanine	COLNVLDHVKWLRT-QMMMGPOBSA-N	9
5	Ornithine		9
6	Tyrosine	OUYCCCASQSFEME-QMMMGPOBSA-N	9
7	Glutamic acid	WHUUTDBJXJRKMK-VKHMYHEASA-N	8
8	Isoleucine	AGPKZVBTJJNPAG-WHFBIAKZSA-N	8
9	LysoPC(18:2) and/or LysoPE (18:2)	-	8
10	PC(34:3)	-	8
11	Citrulline	RHGKLRLOHDJJDR-BYPYZUCNSA-N	7
12	Taurine	XOAAWQZATWQOTB-UHFFFAOYSA-N	7
13	Tryptophan	QIVBCDIJIAJPQS-VIFPVBQESA-N	7
14	Valine	KZSNJWFQEVHDMF-BYPYZUCNSA-N	7
15	LysoPC(18:1)	-	6
16	LysoPC(16:0)	-	6
17	Aminoadipic acid	OYIFNHCXNCRBQI-BYPYZUCNSA-N	6
18	Citric acid	KRKNYBCHXYNGOX-UHFFFAOYSA-N	6
19	Cortisone	MFYSYFVPBJMHGN-ZPOLXVRWSA-N	6
20	Creatinine	DDRJAANPRJIHGJ-UHFFFAOYSA-N	6
21	Glycine	DHMQDGOQFOQNFH-UHFFFAOYSA-N	6
22	Kynurenine	YGPSJZOEDVAXAB-UHFFFAOYSA-N	6
23	PC(36:2)	-	6
24	Alanine	QNAYBMKLOCPYGJ-REOHCLBHSA-N	5
25	Cortisol	JYGXADMDTFJGBT-VWUMJDOOSA-N	5
26	Lysine	KDXKERNSBIXSRK-YFKPBYRVSA-N	5
27	LysoPC(17:0)	-	5
28	PC(34:1)	-	5
29	PC(34:2)	-	5
30	PC(32:1)	-	5
31	Pregnenolone sulfate	DIJBBUIOWGGQOP-OZIWPBGVSA-N	5
32	Sarcosine	FSYKKLYZXJSNPZ-UHFFFAOYSA-N	5
33	SM(20:2)	-	5
34	Threonine	AYFVYJQAPQTCCC-GBXIJSLDSA-N	5
35	Trimethylamine N-oxide (TMAO)	UYPYRKYUKCHHIB-UHFFFAOYSA-N	5

Footnotes: Common names have been assigned to act for synonyms reported in the literature; see Table S1 for further details. Putative identifications are ordered based on the frequency in which they are reported within the literature, with metabolites only listed here if observed in $n \geq 5$ studies; InChIKeys are provided where applicable. Abbreviations: PC—phosphatidylcholine; LysoPC—lysophosphatidylcholine; LysoPE—lysophosphatidylethanolamine; SM—sphingomyelin.

Appendix B: Below is a link to code written as part of the workflow described in Chapter 4, presented as R markdown files divided into specific sub functions, e.g., Z-scoring, and hosted on github.com. Access to the private repository is provided via gitfront.io. Documents in the .docx format will require downloading to be read.

https://gitfront.io/r/user-5413020/UbuiEekxtnBn/PhD-thesis/

Appendix C: Tables summarising annotated rhythmic lipid metabolites characterised in the current thesis and where these data corroborate with prior published studies.

Table C1: List of rhythmic metabolites, as reported, by prior circadian/metabolomics studies primarily investigating non-polar metabolites. Lists acquired from [78]

Dallmann *et al* (2012) [317]	Kasukawa *et al* (2012) [319]	Chua *et al* (2013) [446]	Skene *et al* (2018) - CR post day shift [70]	Skene *et al* (2018) - CR post night shift [70]
10-heptadecenoate (17:1n7)	Cortisone	TG 48:3	C18:1	C18:1
13-HODE + 9-HODE	Urocortisone-3-glucronide	TG 48:2	LysoPC a C18:1	LysoPC a C18:1
1-eicosadienoylglycerophosphocholine	Pregnanolone sulfate-like metabolite	TG 48:1	LysoPC a C18:2	LysoPC a C18:2
1-oleoylglycerol (1-monoolein)	(allo)Urocortisol-3-glucronide	TG 48:0	LysoPC a C26:0	LysoPC a C26:0
1-palmitoylglycerophosphoinositol		TG 50:4	PC aa C30:0	PC aa C30:0
21-hydroxypregnenolone disulfate		TG 50:3	PC aa C32:0	PC aa C32:0
2-linoleoylglycerophosphoethanolamine		TG 50:2	PC aa C32:1	PC aa C32:1
2-oleoylglycerophosphoethanolamine		TG 51:2	PC aa C32:3	PC aa C32:3
3-hydroxydecanoate		TG 52:6	PC aa C34:4	PC aa C34:4
5alpha-androstan-3alpha,17beta-diol disulfate		TG 52:5	PC aa C36:0	PC aa C36:0
arachidonate (20:4n6)		TG 52:4	PC aa C36:5	PC aa C36:5
cis-4-decenoyl carnitine		TG 52:3	PC aa C36:6	PC aa C36:6
cortisol		TG 52:2	PC aa C40:2	PC aa C40:2
cortisone		TG 52:1	PC aa C42:2	PC aa C42:2
dihomo-linoleate (20:2n6)		TG 52:0	PC ae C30:0	PC ae C30:0

dihomo-linolenate (20:3n3 or n6)		TG 53:3	PC ae C34:0	PC ae C34:0
docosadienoate (22:2n6)		TG 53:2	PC ae C34:3	PC ae C34:3
docosahexaenoate (DHA; 22:6n3)		TG 54:6	SM C18:0	SM C18:0
docosapentaenoate (n3 DPA; 22:5n3)		TG 54:5	SM C18:1	SM C18:1
docosapentaenoate (n6 DPA; 22:5n6)		TG 54:4	SM C20:2	SM C20:2
eicosapentaenoate (EPA; 20:5n3)		TG 54:3	LysoPC a C16:1	C14:2
laurylcarnitine		TG 54:2	LysoPC a C17:0	C2
linoleate (18:2n6)		TG 54:1	LysoPC a C18:0	C3-DC (C4-OH)
linolenate [alpha or gamma; (18:3n3 or 6)]		TG 56:3	LysoPC a C20:3	C5-DC (C6-OH)
nonadecanoate (19:0)		TG 60:0	LysoPC a C20:4	C9
octanoylcarnitine		DAG 34:1	LysoPC a C26:1	PC aa C34:1
palmitate (16:0)		DAG 36:3	LysoPC a C28:0	PC aa C34:2
palmitoleate (16:1n7)		DAG 36:2	PC aa C24:0	PC aa C34:3
pregn steroid monosulfate		DAG 36:1	PC ae C34:2	PC aa C36:2
pregnen-diol disulfate		PC 32:2	PC ae C36:0	PC aa C36:4
pregnenolone sulfate		PC O-34:3	PC ae C38:2	PC aa C38:4
alpha-hydroxyisocaproate		PC O-34:2	LysoPC 16:0	PC aa C40:5
glycerol 2-phosphate		PC O-34:1		PC ae C34:1
		PC O-36:3		PC ae C36:4
				PC ae C36:5
				PC ae C44:3

Footnotes: Metabolite names listed under Skene *et al* maintain parity with naming conventions utilised by the applied Biocrates p180 targeted assay kit, as follows. 'a' or 'aa' – both moieties (or one in the case of lysoPCs) at the sn-1 and sn-2 positions are fatty acyl residues bound to the glycerol backbone by ester bonds. 'ae' - one of the moieties, either in the sn-1 or sn-2 position, is a fatty alcohol residue bound by an ether bond. A lone capital C followed by integer(s) e.g., C2 or C18:1 corresponds to acyl carnitines of the

designated carbon length. A capital C proceeding another lipid, e.g., SM C20:2 corresponds to the carbon length of the designated lipid in this instance a sphingomyelin. Abbreviations: DG – diacylglycerol, PC - phosphatidylcholine, SM - sphingomyelin, TG – triacylglycerol.

Appendix Table C2: A full list of rhythmic metabolites, putatively annotated to a single metabolite class, across both lipidomics assays for males (n = 10) and females (n = 10).

Assay	Sex	*m/z*	RT	Putative annotation	Heatmap index	Acro-phase	MSI level
Lipidomics (+)	M	152.022223	35	Creatinine	(1)	23	3
Lipidomics (+)	M	188.0708345	35	Tryptophan	(2)	12	3
Lipidomics (+)	M	205.0966035	35	Tryptophan	(4)	12	3
Lipidomics (+)	M	218.9452946	593	Orotic acid	(5)	14	3
Lipidomics (+)	M	224.9453266	585	4-Hydroxy-2-oxoglutarate	(6)	17	3
Lipidomics (+)	M	229.1547825	34	Isoleucylproline | Leucylproline	(7)	14	3
Lipidomics (+)	M	251.1368002	35	Isoleucylproline | Leucylproline	(9)	14	3
Lipidomics (+)	M	310.2009858	39	CAR(10:3)	(13)	19	2
Lipidomics (+)	M	311.2047156	39	CAR(10:3)	(14)	18	3
Lipidomics (+)	M	312.21739	41	CAR(10:2)	(15)	19	2
Lipidomics (+)	M	313.219907	41	CAR(10:3) | CAR(10:2)	(16)	19	3
Lipidomics (+)	M	328.2488151	48	FA(18:3;O2)	(17)	15	3
Lipidomics (+)	M	333.2474263	40	WE(17:2;O3) | CAR(10:1;O) | CAR(8:1) | CAR(8:0)	(19)	24	3
Lipidomics (+)	M	336.217383	40	CAR(10:3)	(21)	17	3
Lipidomics (+)	M	356.28082	65	FA(20:3;O2)	(22)	15	3
Lipidomics (+)	M	357.283655	65	FA(20:3;O2)	(23)	15	3
Lipidomics (+)	M	413.1376837	39	FA(20:5;O3)	(26)	24	3
Lipidomics (+)	M	431.3528144	411	ST(28:2;O3) | ST(28:3;O3) | Various Vitamin D2/D3 derivatives	(28)	19	3
Lipidomics (+)	M	432.3561146	411	ST(28:2;O3) | ST(28:3;O3) | Various Vitamin D2/D3 derivatives	(29)	19	3
Lipidomics (+)	M	440.2783271	64	LysoPC(12:0)	(30)	21	2
Lipidomics (+)	M	453.3345445	411	ST(28:2;O3) | Various Vitamin D3 derivatives	(32)	19	3
Lipidomics (+)	M	456.3322197	53	CAR(18:2;O2)	(33)	20	3
Lipidomics (+)	M	465.3945736	561	ST(27:0;O4) | Various Vitamin D3 derivatives	(35)	17	3
Lipidomics (+)	M	480.3444051	187	LysoPC(P-16:0) | LysoPC(O-16:1)	(36)	9	2
Lipidomics (+)	M	491.4104867	561	Various Vitamin D3 derivatives	(37)	17	3
Lipidomics (+)	M	494.3223212	107	LysoPC(16:1)	(38)	10	2
Lipidomics (+)	M	497.3423858	141	FA(29:2)	(39)	10	3
Lipidomics (+)	M	498.3468163	141	FA(29:2)	(40)	10	3
Lipidomics (+)	M	502.3278539	187	LysoPC(P-16:0) | LysoPC(O-16:1)	(42)	9	3
Lipidomics (+)	M	504.3053358	121	LysoPC(15:0)	(43)	10	2
Lipidomics (+)	M	510.3554099	200	LysoPC(17:0)	(44)	10	2
Lipidomics (+)	M	510.3560246	189	LysoPC(17:0)	(45)	10	2
Lipidomics (+)	M	512.3340815	48	LysoPS(O-18:0) | PC(16:0) | LysoPS(O-18:1)	(47)	15	3
Lipidomics (+)	M	513.338597	107	LysoPS(O-18:1) | LysoPS(O-18:0)	(48)	13	3
Lipidomics (+)	M	514.3420036	48	LysoPS(O-18:1) | LysoPS(O-18:0)	(49)	14	3
Lipidomics (+)	M	514.3413237	107	LysoPS(O-18:1) | LysoPS(O-18:0)	(50)	14	3
Lipidomics (+)	M	516.3060071	107	LysoPC(18:4) | LysoPC(16:1) | LysoPE(19:1)	(51)	10	3
Lipidomics (+)	M	517.4257899	563	Various	(53)	17	3

				Vitamin D3 derivatives			
Lipidomics (+)	M	518.3202602	142	LysoPC(16:0)	(54)	10	2
Lipidomics (+)	M	519.3243239	157	(22E,24E,24bE)-1alpha,25-dihydroxy-26,27-dimethyl-22,23,24,24a,24b,24c-hexadehydro-24a,24b,24c-trihomovitamin D3	(55)	10	3
Lipidomics (+)	M	519.3254418	141	LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPC(18:0) | PC(O-18:0) | LysoPE(19:0) | LysoPE(21:0)	(56)	10	3
Lipidomics (+)	M	528.309279	126	LysoPE(22:6) | LysoPE(22:5)	(59)	8	3
Lipidomics (+)	M	534.2952679	157	LysoPG(20:4) | LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPE(19:0)	(60)	9	3
Lipidomics (+)	M	536.3726423	217	LysoPC(19:1)	(61)	9	2
Lipidomics (+)	M	542.3227164	88	LysoPC(20:5)	(67)	9	2
Lipidomics (+)	M	544.3362521	174	LysoPC(18:1) | PC(O-18:1)	(70)	9	3
Lipidomics (+)	M	545.3401958	174	LysoPC(18:1) | PC(O-18:1)	(71)	9	3
Lipidomics (+)	M	546.3517852	243	LysoPC(18:0)	(73)	10	2
Lipidomics (+)	M	546.3524255	227	LysoPC(18:0)	(74)	10	2
Lipidomics (+)	M	547.3557087	243	LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)	(75)	10	3
Lipidomics (+)	M	547.356496	227	LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)	(76)	10	3
Lipidomics (+)	M	551.5024448	600	DG(O-32:2)	(77)	14	3
Lipidomics (+)	M	551.5030262	591	DG(O-32:2)	(78)	13	3
Lipidomics (+)	M	552.5076118	600	DG(O-32:2)	(79)	13	3
Lipidomics (+)	M	558.2947071	121	LysoPG(22:6) | LysoPC(18:2)	(80)	6	3
Lipidomics (+)	M	560.3110168	174	LysoPC(18:1) | PC(O-18:1)	(82)	9	3
Lipidomics (+)	M	561.4881707	588	FAHFA(36:5;O) | FAHFA(34:4;O)	(83)	14	3
Lipidomics (+)	M	562.3274768	243	LysoPG(22:4) | LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)	(84)	10	3
Lipidomics (+)	M	563.5042472	596	FAHFA(34:2;O) | FAHFA(38:4;O)	(85)	14	3
Lipidomics (+)	M	564.3067741	88	LysoPC(20:5)	(86)	9	3
Lipidomics (+)	M	564.5076075	596	FAHFA(38:4;O)	(87)	14	3
Lipidomics (+)	M	565.3103184	88	LysoPC(20:5)	(88)	8	3
Lipidomics (+)	M	565.5195543	604	FAHFA(38:4;O) | FAHFA(36:2;O)	(89)	13	3
Lipidomics (+)	M	566.3215997	116	LysoPC(20:4)	(90)	9	3
Lipidomics (+)	M	567.3239769	116	LysoPC(20:4)	(91)	9	3
Lipidomics (+)	M	568.3376742	108	LysoPC(22:6) | LysoPC(20:3)	(92)	7	3
Lipidomics (+)	M	568.3374031	143	LysoPC(20:3)	(93)	9	3
Lipidomics (+)	M	570.3544513	125	LysoPC(22:5) | LysoPC(22:6) | LysoPC(20:2) | PC(O-20:2)	(94)	9	3
Lipidomics (+)	M	571.3593858	125	LysoPC(22:6) | LysoPC(22:5)	(95)	9	3
Lipidomics (+)	M	572.3690877	256	LysoPC(20:1) | PC(O-20:1)	(96)	6	3
Lipidomics (+)	M	577.5197662	600	DG(O-36:4) | DG(O-32:2)	(97)	14	3
Lipidomics (+)	M	578.5219613	600	DG(O-36:4) | DG(O-32:2)	(98)	15	3
Lipidomics (+)	M	580.2921562	141	Protoporphyrin | Protoporphyrin IX	(99)	10	3
Lipidomics (+)	M	580.4343927	372	LysoPC(22:0)	(100)	3	2
Lipidomics (+)	M	585.2724003	253	Biliverdin-IX-beta | Bilirubin	(102)	10	3
Lipidomics (+)	M	587.5035744	589	FAHFA(36:5;O) | FAHFA(34:1;O) | FAHFA(36:2;O)	(103)	14	3

Lipidomics (+)	M	590.3214035	108	LysoPC(22:6)	(104)	8	2
Lipidomics (+)	M	591.3257807	108	LysoPC(22:6)	(105)	8	3
Lipidomics (+)	M	591.4967621	510	DG(32:0) | DG(34:3) | DG(34:4)	(106)	16	3
Lipidomics (+)	M	591.5346124	604	FAHFA(36:2;O) | FAHFA(38:4;O)	(107)	13	3
Lipidomics (+)	M	592.3376722	125	LysoPC(22:5)	(108)	9	3
Lipidomics (+)	M	595.4720712	567	PA(O-29:0)	(109)	19	3
Lipidomics (+)	M	603.5336682	601	DG(O-36:4)	(111)	15	3
Lipidomics (+)	M	604.5374738	601	DG(O-36:4)	(112)	16	3
Lipidomics (+)	M	607.2525738	253	Bilirubin	(113)	8	3
Lipidomics (+)	M	608.4653866	414	LysoPC(24:0)	(114)	2	2
Lipidomics (+)	M	614.3406181	243	LysoPI(19:1) | PA(27:2;O)	(115)	10	3
Lipidomics (+)	M	614.3406266	227	LysoPI(19:1) | PA(27:2;O)	(116)	10	3
Lipidomics (+)	M	615.3435086	243	LysoPI(19:1) | PA(27:2;O)	(117)	10	3
Lipidomics (+)	M	631.5825631	526	Cer(39:1;O2)	(119)	5	3
Lipidomics (+)	M	634.5404813	482	DG(36:4)	(120)	17	2
Lipidomics (+)	M	634.4902434	511	PA(O-32:1)	(121)	15	3
Lipidomics (+)	M	639.6493499	546	Cer(41:0;O2)	(122)	15	3
Lipidomics (+)	M	645.5435963	528	DG(38:4) | DG(36:1) | DG(38:5)	(123)	15	3
Lipidomics (+)	M	646.5486125	527	DG(38:5) | DG(38:4) | DG(36:1)	(124)	16	3
Lipidomics (+)	M	654.6415718	533	Cer(41:0;O3) | Cer(41:1;O3)	(125)	15	3
Lipidomics (+)	M	671.5713134	591	16:1 Stigmasterol ester | 16:2 Sitosterol ester | CE(18:2) | zymosteryl oleate	(127)	3	3
Lipidomics (+)	M	676.6214521	533	Cer(41:0;O3)	(129)	16	3
Lipidomics (+)	M	682.3265786	243	PA(31:4)	(130)	10	3
Lipidomics (+)	M	683.3327021	243	PA(31:4)	(131)	10	3
Lipidomics (+)	M	687.6350407	553	Cer(43:1;O2)	(133)	9	3
Lipidomics (+)	M	699.5942009	546	DG(42:5) | DG(42:6)	(136)	17	3
Lipidomics (+)	M	710.6271986	547	TG(38:0) | TG(40:1) | TG(42:2)	(140)	16	3
Lipidomics (+)	M	711.6396081	547	TG(41:0)	(142)	16	3
Lipidomics (+)	M	713.5695738	535	TG(40:2)	(143)	16	3
Lipidomics (+)	M	714.5734454	535	TG(40:2)	(145)	16	3
Lipidomics (+)	M	715.5853898	547	TG(40:1) | TG(38:0) | TG(42:2)	(146)	16	3
Lipidomics (+)	M	716.521921	465	PC(31:2)	(147)	17	2
Lipidomics (+)	M	718.5935661	470	TG(40:0)	(150)	6	3
Lipidomics (+)	M	731.5393606	441	PC(32:3) | PE(35:3) | PC(32:2) | PE(35:2) | PA(37:3)	(151)	0	3
Lipidomics (+)	M	742.6098496	549	TG(44:5)	(154)	16	3
Lipidomics (+)	M	750.6622374	555	TG(43:2)	(157)	15	3
Lipidomics (+)	M	752.5101101	441	IPC(32:1;O2)	(158)	0	3
Lipidomics (+)	M	752.5233437	441	PC(34:5)	(159)	0	2
Lipidomics (+)	M	754.5351966	437	PC(34:4) | PE(37:4) | PC(30:1) | PC(32:1) | PE(33:1) | PE(35:1) | PE-NMe(34:1)	(161)	2	3
Lipidomics (+)	M	758.5659219	447	PC(34:3) | PE(37:3) | PC(34:2) | PE(37:2) | PE-NMe(36:2)	(162)	20	3
Lipidomics (+)	M	762.660623	550	TG(44:3)	(163)	16	2
Lipidomics (+)	M	765.677805	561	SM(40:0;O2)	(165)	15	3
Lipidomics (+)	M	768.7044158	581	Dextrothyroxine | Thyroxine	(167)	16	3

Lipidomics (+)	M	769.5922484	477	PC(P-36:4) | PC(O-36:5) | PA(O-40:1) | PC(O-34:1) | PE(O-37:1) | PC(P-36:3) | PC(O-36:4)	(168)	9	3
Lipidomics (+)	M	770.6509135	561	HexCer(39:2;O2) | HexCer(39:1;O2)	(171)	15	3
Lipidomics (+)	M	771.6574466	571	HexCer(39:1;O2)	(172)	15	3
Lipidomics (+)	M	772.6627237	571	HexCer(39:1;O2)	(173)	14	3
Lipidomics (+)	M	774.5602597	437	PG(O-35:0)	(175)	0	3
Lipidomics (+)	M	778.6915261	567	TG(43:1) | TG(45:2)	(178)	15	3
Lipidomics (+)	M	779.696004	567	TG(45:2) | TG(43:1)	(180)	15	3
Lipidomics (+)	M	780.5513923	428	PC(34:2)	(181)	20	2
Lipidomics (+)	M	783.5707374	448	HexCer(36:1;O4)	(183)	20	3
Lipidomics (+)	M	783.6468342	568	TG(45:2)	(184)	13	3
Lipidomics (+)	M	789.5574649	473	PG(37:3) | PC(P-36:4) | PC(O-36:5) | PA(O-42:5)	(187)	4	3
Lipidomics (+)	M	790.6888465	563	SM(42:0;O2) | SM(40:0;O2)	(188)	16	3
Lipidomics (+)	M	791.6920938	563	SM(40:0;O2) | SM(42:0;O2)	(190)	16	3
Lipidomics (+)	M	794.2834319	108	UDP-3-(3R-hydroxy-tetradecanoyl)-alphaD-glucosamine | UDP-3-O-(3-hydroxytetradecanoyl)-D-glucosamine	(191)	9	3
Lipidomics (+)	M	794.7173395	582	Dextrothyroxine | Thyroxine	(193)	16	3
Lipidomics (+)	M	795.7221523	582	TG(44:0) | TG(46:1)	(195)	15	3
Lipidomics (+)	M	796.6565844	563	TG(48:6)	(198)	16	3
Lipidomics (+)	M	797.5835229	471	PC(37:4) | PE(40:4) | PA(41:1) | PC(35:1) | PE(38:1) | PC(37:5) | PE(40:5) | PA(42:5)	(199)	11	3
Lipidomics (+)	M	798.6669997	572	PG(dO-40:10;1)	(200)	17	3
Lipidomics (+)	M	799.673023	581	Dextrothyroxine | Thyroxine	(201)	17	3
Lipidomics (+)	M	802.533416	429	PC(36:5) | PE(39:5)	(203)	20	3
Lipidomics (+)	M	805.7113679	568	TG(45:2)	(205)	16	3
Lipidomics (+)	M	808.7372872	586	TG(45:0) | TG(47:1)	(206)	13	3
Lipidomics (+)	M	809.7400418	586	TG(47:1)	(207)	14	3
Lipidomics (+)	M	817.7090409	564	SM(40:0;O2)	(210)	17	3
Lipidomics (+)	M	818.7126417	562	SM(39:1;O2) | SM(41:1;O2)	(211)	17	3
Lipidomics (+)	M	821.7322998	582	Dextrothyroxine | Thyroxine	(213)	15	3
Lipidomics (+)	M	826.6868003	572	PG(dO-40:10;1)	(215)	16	3
Lipidomics (+)	M	834.5984529	462	PC(40:6) | PE(43:6) | PC(38:3) | PE(41:3) | PA(44:3) | PC(40:7)	(216)	6	3
Lipidomics (+)	M	836.5356551	439	PC(42:11) | PI(34:3) | PI(34:2)	(218)	0	3
Lipidomics (+)	M	836.7830367	595	TG(50:0)	(219)	13	3
Lipidomics (+)	M	837.7810071	595	TG(50:0)	(220)	15	3
Lipidomics (+)	M	839.7066391	588	TG(49:2)	(222)	13	3
Lipidomics (+)	M	839.6584443	573	PC(O-39:1) | PE(O-42:1)	(223)	16	3
Lipidomics (+)	M	840.7112174	588	TG(49:2)	(224)	13	3
Lipidomics (+)	M	840.635669	498	IPC(38:0;O2)	(225)	8	3
Lipidomics (+)	M	842.7316638	596	TG(51:4)	(228)	12	3
Lipidomics (+)	M	842.669415	582	TG(48:2)	(229)	17	3
Lipidomics (+)	M	849.7894588	592	TG(51:0) | TG(51:1)	(233)	14	3
Lipidomics (+)	M	856.7473191	584	TG(54:5) | TG(52:4) | TG(52:5)	(236)	14	3

Lipidomics (+)	M	859.674865	563	TG(51:6)	(237)	19	3
Lipidomics (+)	M	865.7188722	578	HexCer(44:0;O2)	(242)	14	3
Lipidomics (+)	M	882.6971047	588	PC(43:4)	(251)	14	3
Lipidomics (+)	M	884.7138497	597	TG(51:2)	(253)	13	3
Lipidomics (+)	M	885.6772841	567	Coenzyme Q10	(254)	5	3
Lipidomics (+)	M	885.7880085	608	TG(52:0) | TG(54:3) | TG(54:4)	(256)	13	3
Lipidomics (+)	M	885.7283836	605	TG(51:1)	(257)	13	3
Lipidomics (+)	M	887.7860929	589	FA(55:1;O3)	(260)	17	3
Lipidomics (+)	M	888.7081828	571	PS(O-42:1)	(261)	19	3
Lipidomics (+)	M	888.6447637	557	PC(42:7) | PC(44:8)	(262)	20	3
Lipidomics (+)	M	889.6551062	513	PC(44:7)	(264)	4	3
Lipidomics (+)	M	890.6530102	513	PC(44:7)	(265)	5	3
Lipidomics (+)	M	896.5374777	452	PI(39:7)	(266)	4	3
Lipidomics (+)	M	899.7551982	575	PC(O-44:3)	(267)	17	3
Lipidomics (+)	M	913.7256439	579	TG(57:10) | TG(59:11) | TG(55:7) | TG(57:8)	(270)	5	3
Lipidomics (+)	M	914.6637518	557	PC(44:8)	(273)	19	3
Lipidomics (+)	M	914.7291254	579	TG(59:11) | TG(57:10) | TG(55:7) | TG(57:8)	(274)	4	3
Lipidomics (+)	M	920.8610144	609	TG(56:1)	(277)	14	3
Lipidomics (+)	M	923.8107287	612	TG(57:6) | TG(57:5)	(279)	14	3
Lipidomics (+)	M	935.7292709	597	Hex2Cer(39:0;O2)	(280)	14	3
Lipidomics (+)	M	936.7343937	597	Hex2Cer(39:0;O2)	(281)	14	3
Lipidomics (+)	M	937.6613304	555	PC(O-44:3)	(282)	3	3
Lipidomics (+)	M	939.7404342	580	TG(57:10) | TG(59:11) | TG(55:7) | TG(57:8)	(283)	4	3
Lipidomics (+)	M	944.7475222	568	TG(54:1)	(284)	2	3
Lipidomics (+)	M	946.7895749	577	TG(58:9)	(286)	4	3
Lipidomics (+)	M	948.7047651	559	TG(58:11)	(288)	3	3
Lipidomics (+)	M	948.9049793	626	TG(58:0)	(290)	13	3
Lipidomics (+)	M	951.8176653	590	TG(58:7)	(292)	6	3
Lipidomics (+)	M	952.7701474	609	Hex2Cer(42:0;O2)	(293)	15	3
Lipidomics (+)	M	952.8210718	589	TG(58:7)	(294)	8	3
Lipidomics (+)	M	953.7548565	587	TG(60:11) | TG(58:8) | TG(60:12) | TG(54:0)	(295)	5	3
Lipidomics (+)	M	954.7564496	587	TG(60:12) | TG(58:8) | TG(60:11) | TG(54:0)	(296)	3	3
Lipidomics (+)	M	961.7422926	597	Hex2Cer(39:0;O2)	(299)	14	3
Lipidomics (+)	M	962.7471466	597	Hex2Cer(39:0;O2)	(300)	14	3
Lipidomics (+)	M	963.6810375	559	TG(58:11)	(301)	3	3
Lipidomics (+)	M	964.6844691	559	TG(58:11)	(302)	3	3
Lipidomics (+)	M	965.6991552	569	TG(58:10)	(304)	4	3
Lipidomics (+)	M	965.9274274	630	TG(58:0)	(305)	16	3
Lipidomics (+)	M	966.701618	569	TG(58:10)	(306)	4	3
Lipidomics (+)	M	970.7861099	573	TG(60:11) | TG(59:7)	(309)	4	3
Lipidomics (+)	M	970.8797999	611	TG(59:4) | TG(58:0)	(310)	15	3
Lipidomics (+)	M	971.7901273	573	TG(60:11)	(312)	4	3
Lipidomics (+)	M	972.9049736	619	TG(60:2) | TG(62:3)	(314)	15	3
Lipidomics (+)	M	974.7486769	595	Galalpha1-4Galbeta-Cer(d42:2) | Hex2Cer(42:2;O2) | Manbeta1-4Glcbeta-Cer(d42:2)	(315)	1	3

Lipidomics (+)	M	975.7378346	573	TG(60:11)	(316)	4	3
Lipidomics (+)	M	975.7619131	602	Galalpha1-4Galbeta-Cer(d42:1) | Hex2Cer(42:1;O2) | Manbeta1-4Glcbeta-Cer(d42:1)	(318)	5	3
Lipidomics (+)	M	976.740635	573	TG(60:11) | TG(62:12)	(319)	4	3
Lipidomics (+)	M	978.7817485	609	Galalpha1-4Galbeta-Cer(d44:1) | Hex2Cer(44:1;O2) | Manbeta1-4Glcbeta-Cer(d44:1) | Hex2Cer(42:0;O2)	(324)	14	3
Lipidomics (+)	M	982.8973186	610	ACer(62:1;O2)	(325)	15	3
Lipidomics (+)	M	983.8009481	597	TG(60:7)	(326)	5	3
Lipidomics (+)	M	996.8437197	627	TG(59:2)	(329)	15	3
Lipidomics (+)	M	1000.926916	626	TG(61:3)	(331)	16	3
Lipidomics (+)	M	1004.798593	609	Galalpha1-4Galbeta-Cer(d44:1) | Hex2Cer(44:1;O2) | Manbeta1-4Glcbeta-Cer(d44:1) | Hex2Cer(42:0;O2)	(332)	17	3
Lipidomics (+)	M	1004.870628	620	TG(61:4) | TG(63:7)	(333)	15	3
Lipidomics (+)	M	1048.900544	541	ACer(66:4;O3)	(335)	6	3
Lipidomics (+)	M	1113.744897	608	Dodecaprenyl phosphate-galacturonic acid	(336)	15	3
Lipidomics (+)	M	1114.75282	608	Dodecaprenyl phosphate-galacturonic acid	(337)	15	3
Lipidomics (-)	M	159.1030079	44	WE(8:0;O) | FA(8:0;O)	(338)	20	3
Lipidomics (-)	M	185.1186816	50	FA(10:1;O) | WE(10:1;O)	(341)	21	3
Lipidomics (-)	M	187.1343016	57	FA(10:0;O) | WE(10:0;O)	(342)	20	3
Lipidomics (-)	M	199.0980259	38	FA(8:2) | WE(8:2) | FA(10:2;O2)	(343)	17	3
Lipidomics (-)	M	204.0669446	37	Indole-3-acetaldehyde | Indolelactic acid	(344)	18	3
Lipidomics (-)	M	229.1450713	43	WE(10:1) | FA(10:1) | FA(12:1;O2)	(345)	20	3
Lipidomics (-)	M	309.2081372	72	WE(16:3) | FA(16:3) | FA(17:3) | WE(17:3)	(348)	22	3
Lipidomics (-)	M	311.223695	88	FA(18:2;O2) | WE(16:2) | FA(16:2) | WE(17:2) | FA(17:2)	(349)	20	3
Lipidomics (-)	M	312.2271541	89	FA(18:2;O2) | WE(16:2) | FA(16:2) | WE(17:2) | FA(17:2)	(350)	20	3
Lipidomics (-)	M	367.1594453	37	ST(19:2;O2;S)	(351)	23	3
Lipidomics (-)	M	368.1627926	37	ST(19:2;O2;S)	(352)	0	3
Lipidomics (-)	M	369.1577337	37	ST(19:2;O2;S)	(354)	22	3
Lipidomics (-)	M	395.1638181	38	ST(21:4;O5)	(356)	21	3
Lipidomics (-)	M	395.1910812	40	ST(21:2;O2;S)	(357)	22	3
Lipidomics (-)	M	455.2489152	43	ST(24:1;O3;S)	(360)	0	3
Lipidomics (-)	M	465.2506008	40	ST(19:1;O2;GlcA)	(362)	21	3
Lipidomics (-)	M	480.3108607	142	LPE(18:0)	(364)	10	2
Lipidomics (-)	M	482.3169797	237	LPE(O-18:1;O) | LPC(15:0) | LPE(18:0)	(366)	10	3
Lipidomics (-)	M	511.2920012	39	ST(27:2;O9)	(372)	22	3
Lipidomics (-)	M	520.2665622	151	LPS(18:2) | LPE(18:3)	(373)	9	3
Lipidomics (-)	M	528.2645025	36	Glycochenodeoxycholate-3-sulfate	(375)	17	3
Lipidomics (-)	M	528.310183	114	LPE(22:4)	(376)	10	3
Lipidomics (-)	M	548.2975588	237	LPS(20:2) | LPE(20:3)	(380)	10	3
Lipidomics (-)	M	549.3017695	237	LPE(20:3) | LPS(20:2)	(381)	10	3
Lipidomics (-)	M	586.3160069	87	LPE(22:5) | LPC(20:5)	(389)	9	3

Lipidomics (-)	M	587.3200117	87	LPE(22:5) | LPC(20:5)	(390)	9	3
Lipidomics (-)	M	589.334992	114	LPE(22:4) | LPC(20:4)	(392)	9	3
Lipidomics (-)	M	590.3470248	142	LPC(19:3) | LPC(20:3)	(393)	9	3
Lipidomics (-)	M	591.3501493	142	LPC(19:3) | LPC(20:3)	(394)	9	3
Lipidomics (-)	M	595.29032	85	LPI(18:2)	(395)	17	3
Lipidomics (-)	M	596.2939081	86	LPI(18:2)	(396)	17	3
Lipidomics (-)	M	598.3096994	118	LPI(18:1)	(397)	18	3
Lipidomics (-)	M	608.3194147	142	LPS(20:2)	(398)	10	3
Lipidomics (-)	M	612.3315455	106	LPE(24:6) | LPC(22:6)	(401)	8	3
Lipidomics (-)	M	613.3348342	106	LPE(24:6) | LPC(22:6)	(402)	8	3
Lipidomics (-)	M	614.3475108	123	LPC(22:5)	(403)	9	3
Lipidomics (-)	M	634.3351938	173	LPC(20:4;O2)	(408)	9	3
Lipidomics (-)	M	635.3382854	173	LPC(20:4;O2)	(409)	9	3
Lipidomics (-)	M	650.6103454	529	Cer(40:1;O) | Cer(42:2;O) | Cer(41:1;O3) | Cer(43:2;O3)	(412)	15	3
Lipidomics (-)	M	654.561975	519	Cer(40:2;O2)	(413)	1	3
Lipidomics (-)	M	668.577572	526	Cer(39:1;O2) | Cer(41:2;O2)	(414)	2	3
Lipidomics (-)	M	669.5795262	526	Cer(41:2;O2) | Cer(39:1;O2)	(415)	3	3
Lipidomics (-)	M	670.5910592	540	Cer(41:1;O2)	(416)	8	3
Lipidomics (-)	M	678.60473	526	Cer(39:1;O2) | Cer(41:2;O2)	(417)	4	3
Lipidomics (-)	M	695.5952311	527	Cer(41:2;O2)	(419)	3	3
Lipidomics (-)	M	701.5346361	498	PC(O-31:2) | PE(O-34:2) | PnE(16:0/18:1(9Z))	(420)	3	3
Lipidomics (-)	M	702.6167983	539	Cer(42:0;O3)	(421)	15	3
Lipidomics (-)	M	712.4950317	448	PC(31:3) | PE(34:3)	(422)	17	3
Lipidomics (-)	M	716.5261433	486	PC(31:1) | PE(34:1)	(423)	17	3
Lipidomics (-)	M	717.5276759	486	PC(31:1) | PE(34:1)	(424)	17	3
Lipidomics (-)	M	723.6692992	559	Cer(45:1;O4) | Cer(44:1;O2) | Cer(43:1;O2)	(425)	9	3
Lipidomics (-)	M	726.5436846	501	PC(O-33:3) | PE(O-36:3)	(426)	3	3
Lipidomics (-)	M	740.5241183	469	PC(33:3) | PE(36:3)	(429)	17	3
Lipidomics (-)	M	741.5280562	469	PC(33:3) | PE(36:3)	(430)	17	3
Lipidomics (-)	M	760.5615136	456	HexCer(34:1;O3)	(432)	5	3
Lipidomics (-)	M	762.5089121	453	PC(35:6) | PE(38:6) | PE(O-38:7;O)	(433)	14	3
Lipidomics (-)	M	763.5114195	453	PC(35:6) | PE(38:6) | PE(O-38:7;O) | PG(O-33:2) | PG(O-32:2)	(434)	14	3
Lipidomics (-)	M	775.5444661	488	PA(O-40:3)	(436)	6	3
Lipidomics (-)	M	782.4962406	467	PC(31:4) | PC(33:5) | PC(35:6) | PE(34:4) | PE(36:5) | PE(38:6) | PE(O-38:7;O) | PC(34:6) | PE(37:6) | PC(30:4) | PC(32:5) | PE(33:4) | PE(35:5) | PC(36:6;O2) | PS(34:3) | PS(36:4) | PS(38:5)	(437)	18	3
Lipidomics (-)	M	784.5114879	486	PS(36:3)	(438)	16	2
Lipidomics (-)	M	785.5158205	486	PS(36:3) | PC(32:4) | PE(35:4) | PC(33:4) | PE(36:4) | PE(O-36:5;O)	(439)	16	3
Lipidomics (-)	M	800.5422587	456	PC(34:3) | PE(37:3) | PC(33:3) | PE(36:3) | PS(37:2)	(440)	21	3
Lipidomics (-)	M	801.5448248	456	PS(37:2) | PC(33:3) | PE(36:3) | PC(34:3) | PE(37:3)	(441)	20	3

Lipidomics (-)	M	801.5570608	456	PA(O-42:4)	(442)	20	3
Lipidomics (-)	M	807.5004667	462	PS(38:6) | PE(37:7) | PE(38:7)	(443)	16	3
Lipidomics (-)	M	808.5138956	469	PS(36:4) | PC(36:6;O2) | PS(38:5) | PC(32:5) | PE(35:5) | PC(34:6) | PE(37:6) | PC(33:5) | PC(35:6) | PE(36:5) | PE(38:6) | PE(O-38:7;O)	(444)	17	3
Lipidomics (-)	M	810.529297	489	PC(35:5) | PE(38:5) | PE(O-38:6;O) | PC(36:5;O2) | PS(38:4) | PC(34:5) | PE(37:5)	(445)	15	3
Lipidomics (-)	M	831.4994839	453	PS(40:8) | PE(40:9)	(448)	14	3
Lipidomics (-)	M	843.5870027	492	16:2-Glc-Stigmasterol | 16:3-Glc-Sitosterol | 18:3-Glc-cholesterol	(449)	8	3
Lipidomics (-)	M	869.5335457	456	PG(44:10) | PI(O-35:2)	(451)	19	3
Lipidomics (-)	M	874.4839871	462	PS(40:9)	(452)	14	3
Lipidomics (-)	M	875.4854308	462	PS(40:9)	(453)	14	3
Lipidomics (-)	M	899.5749121	499	PI(O-37:1)	(454)	18	3
Lipidomics (-)	M	908.6234287	454	IPC(40:1;O2)	(455)	8	3
Lipidomics (-)	M	996.2549515	114	(S)-Hydroxydecanoyl-CoA	(456)	9	3
Lipidomics (-)	M	1171.636267	156	PIM2(35:2)	(458)	10	3
Lipidomics (+)	F	188.0708345	35	Tryptophan	(2)	11	3
Lipidomics (+)	F	205.0966035	35	Tryptophan	(4)	12	3
Lipidomics (+)	F	218.9452946	593	Orotic acid	(5)	16	3
Lipidomics (+)	F	229.1547825	34	Isoleucylproline | Leucylproline	(7)	14	3
Lipidomics (+)	F	232.1545045	35	FA(11:2;O2) | CAR(4:1) | CAR(4:0)	(8)	13	3
Lipidomics (+)	F	251.1368002	35	Isoleucylproline | Leucylproline	(9)	14	3
Lipidomics (+)	F	282.119738	33	1-Methyladenosine | N6-Methyladenosine	(10)	23	3
Lipidomics (+)	F	299.136019	253	Phenethylamine glucuronide	(11)	10	3
Lipidomics (+)	F	310.2009858	39	CAR(10:3)	(13)	18	2
Lipidomics (+)	F	311.2047156	39	CAR(10:3)	(14)	20	3
Lipidomics (+)	F	312.21739	41	CAR(10:2)	(15)	19	2
Lipidomics (+)	F	313.219907	41	CAR(10:3) | CAR(10:2)	(16)	20	3
Lipidomics (+)	F	328.2488151	48	FA(18:3;O2)	(17)	13	3
Lipidomics (+)	F	332.2437513	40	WE(17:2;O3) | CAR(10:1;O) | CAR(8:1) | CAR(8:0)	(18)	2	3
Lipidomics (+)	F	335.2201673	58	FA(18:2;O2) | FA(20:5;O2) | FA(20:6;O2)	(20)	19	3
Lipidomics (+)	F	336.217383	40	CAR(10:3)	(21)	19	3
Lipidomics (+)	F	356.28082	65	FA(20:3;O2)	(22)	12	3
Lipidomics (+)	F	369.3526712	591	3-Deoxyvitamin D3	(24)	6	3
Lipidomics (+)	F	386.3786851	591	3-Deoxyvitamin D3	(25)	2	3
Lipidomics (+)	F	413.1376837	39	FA(20:5;O3)	(26)	22	3
Lipidomics (+)	F	431.3528144	411	ST(28:2;O3) | ST(28:3;O3) | Various Vitamin D2/D3 derivatives	(28)	20	3
Lipidomics (+)	F	440.2783271	64	LysoPC(12:0)	(30)	20	2
Lipidomics (+)	F	456.3322197	53	CAR(18:2;O2)	(33)	20	3
Lipidomics (+)	F	480.3444051	187	LysoPC(P-16:0) | LysoPC(O-16:1)	(36)	9	2
Lipidomics (+)	F	494.3223212	107	LysoPC(16:1)	(38)	10	2
Lipidomics (+)	F	497.3423858	141	FA(29:2)	(39)	9	3

Lipidomics (+)	F	498.3468163	141	FA(29:2)	(40)	9	3
Lipidomics (+)	F	499.349263	141	FA(29:2)	(41)	10	3
Lipidomics (+)	F	502.3278539	187	LysoPC(P-16:0) | LysoPC(O-16:1)	(42)	9	3
Lipidomics (+)	F	504.3053358	121	LysoPC(15:0)	(43)	9	2
Lipidomics (+)	F	510.3554099	200	LysoPC(17:0)	(44)	9	2
Lipidomics (+)	F	510.3560246	189	LysoPC(17:0)	(45)	10	2
Lipidomics (+)	F	510.3563909	172	LysoPC(17:0)	(46)	10	2
Lipidomics (+)	F	512.3340815	48	LysoPS(O-18:0) | PC(16:0) | LysoPS(O-18:1)	(47)	14	3
Lipidomics (+)	F	513.338597	107	LysoPS(O-18:1) | LysoPS(O-18:0)	(48)	11	3
Lipidomics (+)	F	514.3420036	48	LysoPS(O-18:1) | LysoPS(O-18:0)	(49)	13	3
Lipidomics (+)	F	514.3413237	107	LysoPS(O-18:1) | LysoPS(O-18:0)	(50)	10	3
Lipidomics (+)	F	516.3060071	107	LysoPC(18:4) | LysoPC(16:1) | LysoPE(19:1)	(51)	9	3
Lipidomics (+)	F	517.310513	107	LysoPC(18:4)	(52)	9	3
Lipidomics (+)	F	518.3202602	142	LysoPC(16:0)	(54)	10	2
Lipidomics (+)	F	519.3243239	157	(22E,24E,24bE)-1alpha,25-dihydroxy-26,27-dimethyl-22,23,24,24a,24b,24c-hexadehydro-24a,24b,24c-trihomovitamin D3	(55)	9	3
Lipidomics (+)	F	519.3254418	141	LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPC(18:0) | PC(O-18:0) | LysoPE(19:0) | LysoPE(21:0)	(56)	9	3
Lipidomics (+)	F	520.3361966	110	LysoPE(O-19:0;O)	(57)	8	3
Lipidomics (+)	F	521.3420014	110	LysoPE(O-19:0;O)	(58)	8	3
Lipidomics (+)	F	534.2952679	157	LysoPG(20:4) | LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPE(19:0)	(60)	9	3
Lipidomics (+)	F	536.3726423	217	LysoPC(19:1)	(61)	9	2
Lipidomics (+)	F	536.29661	157	LysoPG(18:0) | LysoPI(O-16:1)	(62)	10	3
Lipidomics (+)	F	538.351979	54	PC(18:0)	(63)	13	2
Lipidomics (+)	F	540.3662341	63	LysoPS(O-20:0) | PC(18:0) | LysoPS(O-20:1)	(64)	14	3
Lipidomics (+)	F	541.3709337	63	LysoPS(O-20:1) | LysoPS(O-20:0)	(65)	14	3
Lipidomics (+)	F	542.3199875	121	LysoPC(18:2)	(66)	8	3
Lipidomics (+)	F	542.3217051	110	LysoPC(18:2)	(68)	8	3
Lipidomics (+)	F	543.3239468	121	LysoPC(18:2)	(69)	8	3
Lipidomics (+)	F	544.3362521	174	LysoPC(18:1) | PC(O-18:1)	(70)	8	3
Lipidomics (+)	F	545.3401958	174	LysoPC(18:1) | PC(O-18:1)	(71)	9	3
Lipidomics (+)	F	546.350698	143	LysoPE(O-19:1;O) | LysoPC(16:0) | LysoPC(18:0) | PC(O-18:0) | LysoPE(19:0) | LysoPE(21:0)	(72)	9	3
Lipidomics (+)	F	546.3517852	243	LysoPC(18:0)	(73)	10	2
Lipidomics (+)	F	546.3524255	227	LysoPC(18:0)	(74)	10	2
Lipidomics (+)	F	547.3557087	243	LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)	(75)	10	3
Lipidomics (+)	F	547.356496	227	LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)	(76)	10	3
Lipidomics (+)	F	551.5024448	600	DG(O-32:2)	(77)	14	3
Lipidomics (+)	F	551.5030262	591	DG(O-32:2)	(78)	13	3
Lipidomics (+)	F	552.5076118	600	DG(O-32:2)	(79)	14	3
Lipidomics (+)	F	558.2947071	121	LysoPG(22:6) | LysoPC(18:2)	(80)	8	3
Lipidomics (+)	F	559.2991754	121	LysoPG(22:6) | LysoPC(18:2) | LysoPG(20:2)	(81)	8	3

Lipidomics (+)	F	560.3110168	174	LysoPC(18:1) | PC(O-18:1)	(82)	8	3
Lipidomics (+)	F	561.4881707	588	FAHFA(36:5;O) | FAHFA(34:4;O)	(83)	14	3
Lipidomics (+)	F	562.3274768	243	LysoPG(22:4) | LysoPC(18:0) | PC(O-18:0) | LysoPE(21:0)	(84)	10	3
Lipidomics (+)	F	563.5042472	596	FAHFA(34:2;O) | FAHFA(38:4;O)	(85)	14	3
Lipidomics (+)	F	564.3067741	88	LysoPC(20:5)	(86)	8	3
Lipidomics (+)	F	564.5076075	596	FAHFA(38:4;O)	(87)	15	3
Lipidomics (+)	F	565.5195543	604	FAHFA(38:4;O) | FAHFA(36:2;O)	(89)	12	3
Lipidomics (+)	F	566.3215997	116	LysoPC(20:4)	(90)	9	3
Lipidomics (+)	F	567.3239769	116	LysoPC(20:4)	(91)	8	3
Lipidomics (+)	F	568.3376742	108	LysoPC(22:6) | LysoPC(20:3)	(92)	8	3
Lipidomics (+)	F	570.3544513	125	LysoPC(22:5) | LysoPC(22:6) | LysoPC(20:2) | PC(O-20:2)	(94)	8	3
Lipidomics (+)	F	571.3593858	125	LysoPC(22:6) | LysoPC(22:5)	(95)	8	3
Lipidomics (+)	F	572.3690877	256	LysoPC(20:1) | PC(O-20:1)	(96)	8	3
Lipidomics (+)	F	577.5197662	600	DG(O-36:4) | DG(O-32:2)	(97)	16	3
Lipidomics (+)	F	578.5219613	600	DG(O-36:4) | DG(O-32:2)	(98)	15	3
Lipidomics (+)	F	580.2921562	141	Protoporphyrin | Protoporphyrin IX	(99)	10	3
Lipidomics (+)	F	584.2941201	107	LysoPI(17:2) | LysoPG(22:4)	(101)	10	3
Lipidomics (+)	F	585.2724003	253	Biliverdin-IX-beta | Bilirubin	(102)	9	3
Lipidomics (+)	F	587.5035744	589	FAHFA(36:5;O) | FAHFA(34:1;O) | FAHFA(36:2;O)	(103)	16	3
Lipidomics (+)	F	590.3214035	108	LysoPC(22:6)	(104)	8	2
Lipidomics (+)	F	591.3257807	108	LysoPC(22:6)	(105)	8	3
Lipidomics (+)	F	591.4967621	510	DG(32:0) | DG(34:3) | DG(34:4)	(106)	15	3
Lipidomics (+)	F	591.5346124	604	FAHFA(36:2;O) | FAHFA(38:4;O)	(107)	14	3
Lipidomics (+)	F	600.2801348	122	Biliverdin-IX-beta	(110)	7	3
Lipidomics (+)	F	603.5336682	601	DG(O-36:4)	(111)	17	3
Lipidomics (+)	F	604.5374738	601	DG(O-36:4)	(112)	16	3
Lipidomics (+)	F	607.2525738	253	Bilirubin	(113)	10	3
Lipidomics (+)	F	614.3406181	243	LysoPI(19:1) | PA(27:2;O)	(115)	10	3
Lipidomics (+)	F	615.3435086	243	LysoPI(19:1) | PA(27:2;O)	(117)	10	3
Lipidomics (+)	F	634.5404813	482	DG(36:4)	(120)	16	2
Lipidomics (+)	F	639.6493499	546	Cer(41:0;O2)	(122)	16	3
Lipidomics (+)	F	646.5486125	527	DG(38:5) | DG(38:4) | DG(36:1)	(124)	16	3
Lipidomics (+)	F	671.5713134	591	16:1 Stigmasterol ester | 16:2 Sitosterol ester | CE(18:2) | zymosteryl oleate	(127)	3	3
Lipidomics (+)	F	672.5750772	591	16:1 Stigmasterol ester | 16:2 Sitosterol ester | CE(18:2) | zymosteryl oleate	(128)	4	3
Lipidomics (+)	F	676.6214521	533	Cer(41:0;O3)	(129)	17	3
Lipidomics (+)	F	682.3265786	243	PA(31:4)	(130)	9	3
Lipidomics (+)	F	683.3327021	243	PA(31:4)	(131)	10	3
Lipidomics (+)	F	686.5547495	561	HexCer(33:2;O2)	(132)	3	3
Lipidomics (+)	F	687.6350407	553	Cer(43:1;O2)	(133)	9	3
Lipidomics (+)	F	699.5942009	546	DG(42:5) | DG(42:6)	(136)	16	3
Lipidomics (+)	F	700.6574514	560	Cer(44:1;O2)	(137)	6	3
Lipidomics (+)	F	710.6271986	547	TG(38:0) | TG(40:1) | TG(42:2)	(140)	18	3

Lipidomics (+)	F	711.630084	548	TG(42:2) | TG(40:1)	(141)	18	3
Lipidomics (+)	F	711.6396081	547	TG(41:0)	(142)	18	3
Lipidomics (+)	F	713.5695738	535	TG(40:2)	(143)	17	3
Lipidomics (+)	F	714.5064128	448	PC(29:0)	(144)	18	2
Lipidomics (+)	F	715.5853898	547	TG(40:1) | TG(38:0) | TG(42:2)	(146)	17	3
Lipidomics (+)	F	717.6005762	559	TG(42:3)	(148)	17	3
Lipidomics (+)	F	718.6093851	559	TG(42:3)	(149)	18	3
Lipidomics (+)	F	731.5393606	441	PC(32:3) | PE(35:3) | PC(32:2) | PE(35:2) | PA(37:3)	(151)	23	3
Lipidomics (+)	F	735.5469575	579	CE(22:6) | HexCer(35:2;O2)	(152)	5	3
Lipidomics (+)	F	741.5992456	549	TG(40:1) | TG(42:2)	(153)	18	3
Lipidomics (+)	F	742.6098496	549	TG(44:5)	(154)	17	3
Lipidomics (+)	F	744.617329	560	EPC(40:2;O2)	(155)	18	3
Lipidomics (+)	F	745.6252176	560	SM C16:1 | EPC(38:1;O2) | EPC(38:2;O2) | EPC(40:2;O2) | SM(35:1;O2)	(156)	18	3
Lipidomics (+)	F	752.5101101	441	IPC(32:1;O2)	(158)	23	3
Lipidomics (+)	F	752.5233437	441	PC(34:5)	(159)	22	2
Lipidomics (+)	F	753.5225776	441	PC(34:6) | PE(37:6) | PA(38:2) | PC(32:2) | PE(35:2) | PA(39:6) | PC(34:5) | PE(37:5)	(160)	23	3
Lipidomics (+)	F	754.5351966	437	PC(34:4) | PE(37:4) | PC(30:1) | PC(32:1) | PE(33:1) | PE(35:1) | PE-NMe(34:1)	(161)	1	3
Lipidomics (+)	F	758.5659219	447	PC(34:3) | PE(37:3) | PC(34:2) | PE(37:2) | PE-NMe(36:2)	(162)	20	3
Lipidomics (+)	F	762.660623	550	TG(44:3)	(163)	17	2
Lipidomics (+)	F	764.6735048	561	SM(40:0;O2)	(164)	17	3
Lipidomics (+)	F	765.677805	561	SM(40:0;O2)	(165)	17	3
Lipidomics (+)	F	766.5358658	463	PA(39:2) | PE(O-36:4) | PC(33:2) | PE(36:2) | LBPA(34:2) | PG(34:2) | PC(35:5) | PE(38:5) | PE(O-38:6;O) | PC(35:6) | PE(38:6) | PE(O-38:7;O) | PA(40:6) | PE(O-38:7)	(166)	16	3
Lipidomics (+)	F	768.7044158	581	Dextrothyroxine | Thyroxine	(167)	18	3
Lipidomics (+)	F	769.7081859	581	Dextrothyroxine | Thyroxine	(169)	17	3
Lipidomics (+)	F	770.7122461	581	Dextrothyroxine | Thyroxine	(170)	17	3
Lipidomics (+)	F	770.6509135	561	HexCer(39:2;O2) | HexCer(39:1;O2)	(171)	17	3
Lipidomics (+)	F	771.6574466	571	HexCer(39:1;O2)	(172)	17	3
Lipidomics (+)	F	772.6627237	571	HexCer(39:1;O2)	(173)	17	3
Lipidomics (+)	F	773.652937	571	HexCer(39:1;O2)	(174)	17	3
Lipidomics (+)	F	778.5339269	448	PC(34:3)	(177)	20	2
Lipidomics (+)	F	779.539408	447	PC(34:3) | PE(37:3) | PA(40:3)	(179)	21	3
Lipidomics (+)	F	780.5513923	428	PC(34:2)	(181)	19	2
Lipidomics (+)	F	782.5665405	448	PC(36:4) | PE(39:4)	(182)	20	3
Lipidomics (+)	F	783.5707374	448	HexCer(36:1;O4)	(183)	20	3
Lipidomics (+)	F	783.6468342	568	TG(45:2)	(184)	15	3
Lipidomics (+)	F	785.5766824	448	PC(34:0) | PE(37:0) | PE-NMe(36:0) | PC(35:4) | PE(38:4) | PE(O-38:5;O)	(185)	20	3
Lipidomics (+)	F	787.1916728	253	Glucuronosyletoposide	(186)	10	3
Lipidomics (+)	F	790.6888465	563	SM(42:0;O2) | SM(40:0;O2)	(188)	17	3

Lipidomics (+)	F	791.6920938	563	SM(40:0;O2) | SM(42:0;O2)	(190)	17	3
Lipidomics (+)	F	794.5085721	447	PC(34:3)	(192)	21	2
Lipidomics (+)	F	794.7173395	582	Dextrothyroxine | Thyroxine	(193)	16	3
Lipidomics (+)	F	795.5129335	447	PG(38:7) | PC(34:3) | PE(37:3) | PC(O-33:1) | PE(O-36:1)	(194)	21	3
Lipidomics (+)	F	795.7221523	582	TG(44:0) | TG(46:1)	(195)	17	3
Lipidomics (+)	F	796.7241915	582	Dextrothyroxine | Thyroxine	(196)	16	3
Lipidomics (+)	F	796.5845163	471	PC(35:1)	(197)	9	2
Lipidomics (+)	F	796.6565844	563	TG(48:6)	(198)	17	3
Lipidomics (+)	F	798.6669997	572	PG(dO-40:10;1)	(200)	19	3
Lipidomics (+)	F	799.6699669	572	HexCer(39:1;O2)	(202)	16	3
Lipidomics (+)	F	802.533416	429	PC(36:5) | PE(39:5)	(203)	20	3
Lipidomics (+)	F	805.5510184	448	PA(42:4) | PC(36:4) | PE(39:4)	(204)	18	3
Lipidomics (+)	F	808.7372872	586	TG(45:0) | TG(47:1)	(206)	14	3
Lipidomics (+)	F	809.7400418	586	TG(47:1)	(207)	14	3
Lipidomics (+)	F	810.7514567	595	TG(47:0)	(208)	11	3
Lipidomics (+)	F	815.7485355	612	FA(51:1;O3)	(209)	5	3
Lipidomics (+)	F	817.7090409	564	SM(40:0;O2)	(210)	18	3
Lipidomics (+)	F	818.7126417	562	SM(39:1;O2) | SM(41:1;O2)	(211)	17	3
Lipidomics (+)	F	819.6261538	553	PS(O-38:2)	(212)	18	3
Lipidomics (+)	F	821.7322998	582	Dextrothyroxine | Thyroxine	(213)	16	3
Lipidomics (+)	F	826.6924807	582	TG(49:6) | TG(48:2)	(214)	16	3
Lipidomics (+)	F	835.5313607	439	PI(34:2) | PI(34:3) | PA(41:1)	(217)	5	3
Lipidomics (+)	F	836.5356551	439	PC(42:11) | PI(34:3) | PI(34:2)	(218)	5	3
Lipidomics (+)	F	836.7830367	595	TG(50:0)	(219)	15	3
Lipidomics (+)	F	837.7810071	595	TG(50:0)	(220)	14	3
Lipidomics (+)	F	838.7847438	604	TG(49:0)	(221)	10	3
Lipidomics (+)	F	839.6584443	573	PC(O-39:1) | PE(O-42:1)	(223)	17	3
Lipidomics (+)	F	840.7112174	588	TG(49:2)	(224)	15	3
Lipidomics (+)	F	840.6577682	573	PC(O-39:1) | PE(O-42:1)	(226)	17	3
Lipidomics (+)	F	841.6659843	582	TG(48:2)	(227)	16	3
Lipidomics (+)	F	842.7316638	596	TG(51:4)	(228)	15	3
Lipidomics (+)	F	842.669415	582	TG(48:2)	(229)	17	3
Lipidomics (+)	F	843.6718187	527	PC(40:2) | PE(43:2)	(230)	0	3
Lipidomics (+)	F	847.6728759	565	undecaprenyl phosphate	(231)	17	3
Lipidomics (+)	F	849.7894588	592	TG(51:0) | TG(51:1)	(233)	13	3
Lipidomics (+)	F	854.7226908	591	TG(51:6) | TG(50:2)	(234)	13	3
Lipidomics (+)	F	856.7346607	599	TG(50:1) | TG(51:5)	(235)	12	3
Lipidomics (+)	F	856.7473191	584	TG(54:5) | TG(52:4) | TG(52:5)	(236)	17	3
Lipidomics (+)	F	859.674865	563	TG(51:6)	(237)	17	3
Lipidomics (+)	F	860.6805553	563	TG(51:6)	(238)	18	3
Lipidomics (+)	F	861.6256415	556	PC(42:7) | PC(44:8) | PGS	(239)	17	3
Lipidomics (+)	F	863.7045463	580	HexCer(44:1;O2) | Cer(51:1;O4)	(240)	16	3
Lipidomics (+)	F	863.6683756	565	TG(51:6)	(241)	16	3
Lipidomics (+)	F	865.6621672	574	PS(40:0)	(243)	13	3
Lipidomics (+)	F	866.6683847	574	HexCer(42:1;O4)	(244)	14	3

Lipidomics (+)	F	872.5297769	448	PI(36:6) | PI(35:2)	(246)	20	3
Lipidomics (+)	F	877.5693755	490	PS(42:8) | PS(41:4) | PG(O-42:5)	(247)	22	3
Lipidomics (+)	F	879.4952595	439	PI(36:5)	(248)	4	3
Lipidomics (+)	F	879.5755811	441	PC(44:12)	(249)	0	3
Lipidomics (+)	F	880.5906993	461	PI(36:2)	(250)	5	3
Lipidomics (+)	F	882.6971047	588	PC(43:4)	(251)	15	3
Lipidomics (+)	F	883.711353	597	TG(51:2)	(252)	13	3
Lipidomics (+)	F	884.7138497	597	TG(51:2)	(253)	15	3
Lipidomics (+)	F	885.6772841	567	Coenzyme Q10	(254)	6	3
Lipidomics (+)	F	885.71894	596	PC(43:2)	(255)	15	3
Lipidomics (+)	F	887.5599026	478	PI(36:1)	(258)	4	3
Lipidomics (+)	F	887.7483762	566	TG(51:0) | TG(53:1)	(259)	17	3
Lipidomics (+)	F	888.7081828	571	PS(O-42:1)	(261)	17	3
Lipidomics (+)	F	889.6584988	567	PC(42:6) | PC(44:7) | PGS	(263)	17	3
Lipidomics (+)	F	889.6551062	513	PC(44:7)	(264)	3	3
Lipidomics (+)	F	910.550393	457	Glc-GP(38:4) | PI(38:4)	(268)	8	3
Lipidomics (+)	F	912.7118187	570	TG(57:11) | TG(59:12)	(269)	2	3
Lipidomics (+)	F	913.7256439	579	TG(57:10) | TG(59:11) | TG(55:7) | TG(57:8)	(270)	3	3
Lipidomics (+)	F	913.6586548	558	PC(44:8)	(271)	19	3
Lipidomics (+)	F	914.5073434	447	SHexCer(38:1;O3)	(272)	22	3
Lipidomics (+)	F	914.6637518	557	PC(44:8)	(273)	19	3
Lipidomics (+)	F	914.7291254	579	TG(59:11) | TG(57:10) | TG(55:7) | TG(57:8)	(274)	4	3
Lipidomics (+)	F	915.5100362	446	(3-sulfo)Galbeta-Cer(d38:1/(2OH))	(275)	21	3
Lipidomics (+)	F	917.1489658	253	Isophthalyl-CoA	(276)	10	3
Lipidomics (+)	F	921.7495172	557	22:0-Glc-Sitosterol	(278)	16	3
Lipidomics (+)	F	935.7292709	597	Hex2Cer(39:0;O2)	(280)	15	3
Lipidomics (+)	F	936.7343937	597	Hex2Cer(39:0;O2)	(281)	15	3
Lipidomics (+)	F	939.7404342	580	TG(57:10) | TG(59:11) | TG(55:7) | TG(57:8)	(283)	3	3
Lipidomics (+)	F	945.9122271	627	ACer(63:3;O3)	(285)	5	3
Lipidomics (+)	F	946.7895749	577	TG(58:9)	(286)	1	3
Lipidomics (+)	F	947.7285642	594	Galalpha1-4Galbeta-Cer(d40:1) | Hex2Cer(40:1;O2) | Manbeta1-4Glcbeta-Cer(d40:1)	(287)	3	3
Lipidomics (+)	F	948.7337796	594	Galalpha1-4Galbeta-Cer(d40:1) | Hex2Cer(40:1;O2) | Manbeta1-4Glcbeta-Cer(d40:1)	(289)	4	3
Lipidomics (+)	F	950.8152394	589	TG(58:7) | TG(60:8)	(291)	3	3
Lipidomics (+)	F	954.7564496	587	TG(60:12) | TG(58:8) | TG(60:11) | TG(54:0)	(296)	5	3
Lipidomics (+)	F	961.7422926	597	Hex2Cer(39:0;O2)	(299)	16	3
Lipidomics (+)	F	962.7471466	597	Hex2Cer(39:0;O2)	(300)	16	3
Lipidomics (+)	F	963.6810375	559	TG(58:11)	(301)	1	3
Lipidomics (+)	F	964.6844691	559	TG(58:11)	(302)	1	3
Lipidomics (+)	F	964.9291197	630	TG(58:0)	(303)	15	3
Lipidomics (+)	F	965.6991552	569	TG(58:10)	(304)	4	3
Lipidomics (+)	F	966.701618	569	TG(58:10)	(306)	3	3
Lipidomics (+)	F	967.7153756	578	TG(58:9)	(307)	5	3
Lipidomics (+)	F	968.7167702	578	TG(58:9)	(308)	5	3

Lipidomics (+)	F	970.7861099	573	TG(60:11) | TG(59:7)	(309)	4	3
Lipidomics (+)	F	971.7308854	586	IPC(44:0;O4)	(311)	19	3
Lipidomics (+)	F	971.7901273	573	TG(60:11)	(312)	3	3
Lipidomics (+)	F	971.9242049	628	ACer(63:3;O3)	(313)	4	3
Lipidomics (+)	F	975.7378346	573	TG(60:11)	(316)	2	3
Lipidomics (+)	F	975.8307786	613	TG(59:4)	(317)	15	3
Lipidomics (+)	F	975.7619131	602	Galalpha1-4Galbeta-Cer(d42:1) | Hex2Cer(42:1;O2) | Manbeta1-4Glcbeta-Cer(d42:1)	(318)	0	3
Lipidomics (+)	F	976.740635	573	TG(60:11) | TG(62:12)	(319)	4	3
Lipidomics (+)	F	976.832231	590	TG(58:7) | TG(60:8)	(320)	4	3
Lipidomics (+)	F	977.8535303	620	TG(59:3) | TG(61:4) | TG(63:7)	(321)	16	3
Lipidomics (+)	F	977.8316427	590	TG(60:8)	(322)	4	3
Lipidomics (+)	F	977.775532	608	Galalpha1-4Galbeta-Cer(d44:1) | Hex2Cer(44:1;O2) | Manbeta1-4Glcbeta-Cer(d44:1) | Hex2Cer(42:0;O2)	(323)	18	3
Lipidomics (+)	F	978.7817485	609	Galalpha1-4Galbeta-Cer(d44:1) | Hex2Cer(44:1;O2) | Manbeta1-4Glcbeta-Cer(d44:1) | Hex2Cer(42:0;O2)	(324)	17	3
Lipidomics (+)	F	982.8973186	610	ACer(62:1;O2)	(325)	16	3
Lipidomics (+)	F	984.9108039	617	TG(61:3)	(327)	16	3
Lipidomics (+)	F	991.8656765	623	TG(60:3) | TG(62:4)	(328)	17	3
Lipidomics (+)	F	999.7632277	597	IPC(46:0;O4)	(330)	4	3
Lipidomics (+)	F	1004.798593	609	Galalpha1-4Galbeta-Cer(d44:1) | Hex2Cer(44:1;O2) | Manbeta1-4Glcbeta-Cer(d44:1) | Hex2Cer(42:0;O2)	(332)	17	3
Lipidomics (+)	F	1007.835996	623	TG(60:3)	(334)	17	3
Lipidomics (+)	F	1113.744897	608	Dodecaprenyl phosphate-galacturonic acid	(336)	16	3
Lipidomics (+)	F	1114.75282	608	Dodecaprenyl phosphate-galacturonic acid	(337)	16	3
Lipidomics (-)	F	167.0211565	33	Uric acid	(339)	16	3
Lipidomics (-)	F	183.0163458	38	5-Hydroxyisourate	(340)	14	3
Lipidomics (-)	F	185.1186816	50	FA(10:1;O) | WE(10:1;O)	(341)	20	3
Lipidomics (-)	F	199.0980259	38	FA(8:2) | WE(8:2) | FA(10:2;O2)	(343)	16	3
Lipidomics (-)	F	204.0669446	37	Indole-3-acetaldehyde | Indolelactic acid	(344)	17	3
Lipidomics (-)	F	229.1450713	43	WE(10:1) | FA(10:1) | FA(12:1;O2)	(345)	19	3
Lipidomics (-)	F	241.1452911	45	WE(11:2) | FA(11:2) | WE(12:2) | FA(12:2) | FA(13:2;O2)	(346)	19	3
Lipidomics (-)	F	285.2078635	73	WE(14:1) | FA(14:1) | WE(15:1) | FA(15:1) | FA(16:1;O2)	(347)	20	3
Lipidomics (-)	F	309.2081372	72	WE(16:3) | FA(16:3) | FA(17:3) | WE(17:3)	(348)	20	3
Lipidomics (-)	F	311.223695	88	FA(18:2;O2) | WE(16:2) | FA(16:2) | WE(17:2) | FA(17:2)	(349)	20	3
Lipidomics (-)	F	312.2271541	89	FA(18:2;O2) | WE(16:2) | FA(16:2) | WE(17:2) | FA(17:2)	(350)	19	3
Lipidomics (-)	F	367.1594453	37	ST(19:2;O2;S)	(351)	22	3
Lipidomics (-)	F	368.1627926	37	ST(19:2;O2;S)	(352)	22	3
Lipidomics (-)	F	369.1752467	39	ST(19:1;O2;S)	(353)	21	3

Lipidomics (-)	F	369.1577337	37	ST(19:2;O2;S)	(354)	22	3
Lipidomics (-)	F	370.1782943	39	ST(19:1;O2;S)	(355)	20	3
Lipidomics (-)	F	395.1638181	38	ST(21:4;O5)	(356)	21	3
Lipidomics (-)	F	395.1910812	40	ST(21:2;O2;S)	(357)	20	3
Lipidomics (-)	F	396.167425	38	ST(21:4;O5)	(358)	20	3
Lipidomics (-)	F	405.1926107	38	ST(21:4;O5)	(359)	21	3
Lipidomics (-)	F	464.3161463	267	LPC(O-15:1) | LPE(O-18:1)	(361)	10	3
Lipidomics (-)	F	480.3098476	237	LPE(O-18:1;O) | LPC(15:0) | LPE(18:0)	(363)	9	3
Lipidomics (-)	F	480.3110859	156	LPE(O-18:1;O) | LPC(15:0) | LPE(18:0)	(365)	9	3
Lipidomics (-)	F	482.3169797	237	LPE(O-18:1;O) | LPC(15:0) | LPE(18:0)	(366)	9	3
Lipidomics (-)	F	502.2956222	137	LPE(20:3)	(367)	8	3
Lipidomics (-)	F	504.3118477	120	LPC(17:2) | LPE(20:2)	(368)	7	3
Lipidomics (-)	F	505.3143411	120	LPC(17:2) | LPE(20:2)	(369)	8	3
Lipidomics (-)	F	506.3257461	173	LPE(20:1)	(370)	8	2
Lipidomics (-)	F	508.3415088	245	LPE(20:0)	(371)	9	2
Lipidomics (-)	F	520.2665622	151	LPS(18:2) | LPE(18:3)	(373)	9	3
Lipidomics (-)	F	524.2789742	102	LPE(22:6)	(374)	8	3
Lipidomics (-)	F	528.310183	114	LPE(22:4)	(376)	8	3
Lipidomics (-)	F	530.3043189	156	LPE(O-19:1;O) | LPC(16:0) | LPE(19:0)	(377)	9	3
Lipidomics (-)	F	540.3303627	142	LPE(O-18:1;O) | LPC(15:0) | LPE(18:0) | LPS(O-19:0;O) | LPE(O-19:1;O) | LPC(16:0) | LPE(19:0)	(378)	9	3
Lipidomics (-)	F	541.3351862	142	LPE(O-18:1;O) | LPC(15:0) | LPE(18:0) | LPE(O-19:1;O) | LPC(16:0) | LPE(19:0) | LPS(O-19:0;O)	(379)	9	3
Lipidomics (-)	F	548.2975588	237	LPS(20:2) | LPE(20:3)	(380)	9	3
Lipidomics (-)	F	549.3017695	237	LPE(20:3) | LPS(20:2)	(381)	10	3
Lipidomics (-)	F	554.3468956	202	LPE(O-19:1;O) | LPC(16:0) | LPE(19:0) | LPC(17:0) | LPE(20:0) | LPS(O-20:0;O)	(382)	9	3
Lipidomics (-)	F	556.3264039	48	LPS(O-18:0)	(383)	13	3
Lipidomics (-)	F	564.3295671	120	LPC(17:2) | LPE(20:2) | LPC(18:2)	(384)	7	3
Lipidomics (-)	F	566.3470727	156	LPC(17:1) | LPE(20:1) | LPC(18:1) | PC(O-18:1) | LPS(21:0)	(385)	8	3
Lipidomics (-)	F	566.3465904	173	LPC(17:1) | LPE(20:1) | LPC(18:1) | PC(O-18:1) | LPS(21:0)	(386)	8	3
Lipidomics (-)	F	568.2664405	110	LPS(22:6)	(387)	8	3
Lipidomics (-)	F	568.359419	245	LPS(O-21:0;O) | LPC(18:0) | PC(O-18:0) | LPE(21:0) | LPC(17:0) | LPE(20:0)	(388)	9	3
Lipidomics (-)	F	586.3160069	87	LPE(22:5) | LPC(20:5)	(389)	8	3
Lipidomics (-)	F	587.3200117	87	LPE(22:5) | LPC(20:5)	(390)	8	3
Lipidomics (-)	F	588.3303137	114	LPE(22:4) | LPC(20:4)	(391)	8	3
Lipidomics (-)	F	589.334992	114	LPE(22:4) | LPC(20:4)	(392)	8	3
Lipidomics (-)	F	590.3470248	142	LPC(19:3) | LPC(20:3)	(393)	8	3
Lipidomics (-)	F	591.3501493	142	LPC(19:3) | LPC(20:3)	(394)	8	3
Lipidomics (-)	F	596.2939081	86	LPI(18:2)	(396)	14	3
Lipidomics (-)	F	608.3194147	142	LPS(20:2)	(398)	9	3
Lipidomics (-)	F	608.318052	156	LPS(20:2)	(399)	9	3

Lipidomics (-)	F	609.3228819	156	LPS(20:2)	(400)	9	3
Lipidomics (-)	F	612.3315455	106	LPE(24:6) | LPC(22:6)	(401)	8	3
Lipidomics (-)	F	613.3348342	106	LPE(24:6) | LPC(22:6)	(402)	8	3
Lipidomics (-)	F	621.3058524	99	LPI(20:3)	(404)	14	3
Lipidomics (-)	F	632.3188398	120	LPS(22:4)	(405)	7	3
Lipidomics (-)	F	632.3190369	109	LPS(22:4)	(406)	7	3
Lipidomics (-)	F	633.3228812	120	LPS(22:4) | PE(25:4;O3)	(407)	7	3
Lipidomics (-)	F	634.3351938	173	LPC(20:4;O2)	(408)	8	3
Lipidomics (-)	F	635.3382854	173	LPC(20:4;O2)	(409)	8	3
Lipidomics (-)	F	636.3498522	245	LPS(22:2)	(410)	10	3
Lipidomics (-)	F	637.3533996	245	LPS(22:2)	(411)	9	3
Lipidomics (-)	F	654.561975	519	Cer(40:2;O2)	(413)	20	3
Lipidomics (-)	F	702.6167983	539	Cer(42:0;O3)	(421)	18	3
Lipidomics (-)	F	712.4950317	448	PC(31:3) | PE(34:3)	(422)	18	3
Lipidomics (-)	F	716.5261433	486	PC(31:1) | PE(34:1)	(423)	16	3
Lipidomics (-)	F	717.5276759	486	PC(31:1) | PE(34:1)	(424)	16	3
Lipidomics (-)	F	727.5468722	501	PC(O-33:3) | PE(O-36:3)	(427)	5	3
Lipidomics (-)	F	728.5633631	516	PC(O-33:2) | PE(O-36:2)	(428)	4	3
Lipidomics (-)	F	740.5241183	469	PC(33:3) | PE(36:3)	(429)	17	3
Lipidomics (-)	F	741.5280562	469	PC(33:3) | PE(36:3)	(430)	18	3
Lipidomics (-)	F	744.5557051	506	PE(36:1)	(431)	16	2
Lipidomics (-)	F	760.5615136	456	HexCer(34:1;O3)	(432)	6	3
Lipidomics (-)	F	763.5114195	453	PC(35:6) | PE(38:6) | PE(O-38:7;O) | PG(O-33:2) | PG(O-32:2)	(434)	14	3
Lipidomics (-)	F	765.5281565	463	PG(O-33:1) | PG(O-32:1) | EPC(38:2;O3)	(435)	17	3
Lipidomics (-)	F	782.4962406	467	PC(31:4) | PC(33:5) | PC(35:6) | PE(34:4) | PE(36:5) | PE(38:6) | PE(O-38:7;O) | PC(34:6) | PE(37:6) | PC(30:4) | PC(32:5) | PE(33:4) | PE(35:5) | PC(36:6;O2) | PS(34:3) | PS(36:4) | PS(38:5)	(437)	16	3
Lipidomics (-)	F	784.5114879	486	PS(36:3)	(438)	16	2
Lipidomics (-)	F	785.5158205	486	PS(36:3) | PC(32:4) | PE(35:4) | PC(33:4) | PE(36:4) | PE(O-36:5;O)	(439)	16	3
Lipidomics (-)	F	801.5448248	456	PS(37:2) | PC(33:3) | PE(36:3) | PC(34:3) | PE(37:3)	(441)	20	3
Lipidomics (-)	F	801.5570608	456	PA(O-42:4)	(442)	21	3
Lipidomics (-)	F	815.547397	467	PG(O-36:4)	(446)	8	3
Lipidomics (-)	F	829.5757855	480	PS(39:2) | PC(36:3) | PE(39:3) | PC(35:3) | PE(38:3)	(447)	23	3
Lipidomics (-)	F	831.4994839	453	PS(40:8) | PE(40:9)	(448)	14	3
Lipidomics (-)	F	860.6103602	454	Galalpha1-4Galbeta-Cer(d34:1) | Hex2Cer(34:1;O2) | Manbeta1-4Glcbeta-Cer(d34:1)	(450)	7	3
Lipidomics (-)	F	874.4839871	462	PS(40:9)	(452)	17	3
Lipidomics (-)	F	1035.658782	156	NeuAcHexCer(34:1;O2)	(457)	9	3
Lipidomics (-)	F	1171.636267	156	PIM2(35:2)	(458)	9	3

Footnotes: (+) / (-) indicate positive and negative ion modes respectively, M – Males, F – females, *m/z* – mass to charge ratio, RT – retention time. The 'Heatmap index' can be used to cross reference metabolites within this table to Figure 6.4, acrophase estimated via cosinor analysis as performed as part of the group level rhythm analysis. Lipid abbreviations: CAR – Acylcarnitine, Cer – Ceramide, CL – Cardiolipin, DG – Diacylglyceride, FA – Fatty acyl, FA(X:X) – Fatty acid, GL – Glycerolipid, GluCer – Glucocyl-ceramide, GM3 – Ganglioside mannoside 3, Hex2cer – Dihexosylceramide, NeuAcHexCer – N-Acetylneuraminic acid hexosylceramide, PA – Phosphatidic acid, PC – Phosphatidylcholine, PE – Phosphatidylethanolamine, PG – Phosphatidylglycerol, PI – Phosphatidylinositol, PS – Phosphatidylserine, SM – Sphingomyelin, SP – Sphingolipid, ST – Sterol lipid, TG – Triacylglyceride.

Appendix D Tables summarising annotated rhythmic water-soluble metabolites characterised in the current thesis and where these data corroborate with prior published studies.

Table D1: List of rhythmic metabolites, as reported, by prior circadian/metabolomics studies primarily investigating polar metabolites. Lists acquired from [78]
Grant *et al* (2018)
2-Methylcitric acid
Citrulline
Fumaric acid
L-Arginine
L-Leucine
L-Lysine
L-Methionine
L-Threonine
L-Valine
Sucrose-6-phosphate
4-Hydroxyphenylpyruvic acid
Citric acid
D-Glucuronic acid
L-Homoserine
L-Asparagine
L-Isoleucine
L-Tyrosine
Ornithine
dUMP
Glycyl-glycine
5-Hydroxylysine
Inosinic acid
Glucose 6-phosphate
Isocitric acid
L-Phenylalanine
L-Tryptophan
Pantothenic acid
Cysteic acid

Appendix Table D2: A full list of rhythmic metabolites, putatively annotated to a single metabolite class, across both HILIC assays for males (n = 10) and females (n = 10).

Assay	Sex	*m/z*	RT	Putative annotation	Heatmap index	Acro-phase	MSI level
HILIC (+)	M	103.0542	354	FAL(9:7)	(460)	11	3
HILIC (+)	M	106.0498	469	Serine	(461)	17	2
HILIC (+)	M	118.0648	344	Indole	(462)	11	3
HILIC (+)	M	123.0416	425	FA(5:1) | WE(5:1)	(463)	14	3
HILIC (+)	M	131.0491	354	Cinnamic acid	(464)	12	3
HILIC (+)	M	133.0971	479	Ornithine	(465)	11	2
HILIC (+)	M	136.0756	352	FOH(8:6)	(467)	16	3
HILIC (+)	M	138.0546	415	Trigonelline	(468)	13	2
HILIC (+)	M	144.0808	344	Indole	(470)	12	3
HILIC (+)	M	149.0597	354	Cinnamic acid	(471)	11	3
HILIC (+)	M	155.0795	479	Ornithine	(473)	11	2
HILIC (+)	M	157.076	344	3-Indoleacetonitrile	(474)	12	3
HILIC (+)	M	163.0865	352	Norcotinine	(475)	16	3
HILIC (+)	M	166.0723	122	7-Methylguanine | N2-Methylguanine	(476)	14	3
HILIC (+)	M	166.0723	309	7-Methylguanine | N2-Methylguanine	(477)	1	3
HILIC (+)	M	166.0531	481	Methionine sulfoxide	(478)	13	3
HILIC (+)	M	170.0599	344	Taurocyamine	(480)	12	3
HILIC (+)	M	175.1435	542	NA(8:1;O)	(481)	14	3
HILIC (+)	M	182.0561	179	5-Methylthio-D-ribose | 2,3-Dihydroxy-5-methylthio-4-pentenoic acid	(482)	12	3
HILIC (+)	M	190.1185	468	Homocitrulline	(484)	15	3
HILIC (+)	M	227.0789	344	Tryptophan	(486)	11	2
HILIC (+)	M	228.0742	468	Homocitrulline	(487)	14	2
HILIC (+)	M	229.1182	500	Porphobilinogen | 3-Methoxytyrosine	(488)	13	3
HILIC (+)	M	229.1546	468	Isoleucylproline | Leucylproline	(489)	15	2
HILIC (+)	M	230.157	468	Isoleucylproline | Leucylproline	(490)	15	3
HILIC (+)	M	244.0561	344	Tryptophan	(494)	12	2
HILIC (+)	M	246.1082	351	Cytidine	(495)	12	2
HILIC (+)	M	254.0801	479	2-amino-8-oxo-9,10-epoxy-decanoic acid | CAR(3:1)	(496)	11	3
HILIC (+)	M	262.1284	464	CAR(4:1;O2) | CAR(4:2;O2)	(497)	14	2
HILIC (+)	M	283.0744	344	FOH(14:5;O) | FA(14:4)	(499)	11	3
HILIC (+)	M	298.1146	309	1-Methylguanosine | (S)-malyl alpha-D-glucosaminide	(503)	1	3
HILIC (+)	M	311.1238	460	gamma-Glutamyltyrosine	(505)	15	3
HILIC (+)	M	346.1547	344	Succinylcholine	(506)	12	3
HILIC (+)	M	432.3869	26	CAR(17:0)	(510)	2	3
HILIC (+)	M	449.309	315	(6R)-vitamin D3 6,19-sulfur dioxide adduct | (6S)-vitamin D3 6,19-sulfur dioxide adduct | 1alpha,25-dihydroxy-24a-homo-22-thia-20-epivitamin D3 | 1alpha,25-dihydroxy-24a-homo-22-thiavitamin D3	(512)	20	3
HILIC (+)	M	463.1058	344	farnesyl triphosphate	(513)	12	3

HILIC (+)	M	468.3444	308	LPC(O-15:0)	(514)	8	2
HILIC (+)	M	494.3601	293	LPC(P-17:0)	(516)	10	2
HILIC (-)	M	130.0622	447	Beta-Guanidinopropionic acid | Creatine	(526)	9	3
HILIC (-)	M	173.1043	572	Arginine	(529)	12	2
HILIC (-)	M	174.0882	491	Citrulline	(530)	10	2
HILIC (-)	M	180.0668	407	Tyrosine	(531)	14	2
HILIC (-)	M	185.1183	53	FA(10:1;O) | WE(10:1;O) | FOH(10:2;O2) | FA(10:1;O) | FOH(9:2) | FOH(8:2) | 2-Octenal	(532)	20	3
HILIC (-)	M	188.1041	479	Homocitrulline	(533)	14	3
HILIC (-)	M	188.0209	447	NA(5:3;O3)	(534)	7	3
HILIC (-)	M	190.0511	75	2-Oxoindole-3-acetate | 5-Hydroxyindoleacetic acid	(535)	16	3
HILIC (-)	M	262.1047	513	N-Acetyl-L-citrulline	(536)	15	3
HILIC (-)	M	411.1848	26	17-Hydroxypregnenolone sulfate	(542)	22	2
HILIC (-)	M	451.2608	262	(24R,25S)-25,26-epoxy-1alpha,24-dihydroxy-27-norvitamin D3 | (24S,25R)-25,26-epoxy-1alpha,24-dihydroxy-27-norvitamin D3 | 1alpha,25-Dihydroxy-16-ene-19-nor-24-oxovitamin D3	(543)	23	3
HILIC (-)	M	487.3277	268	Glyceryl lactooleate	(544)	19	3
HILIC (+)	F	106.0498	469	Serine	(461)	19	2
HILIC (+)	F	133.0971	479	Ornithine	(465)	11	2
HILIC (+)	F	135.1016	539	FA(6:0;O)	(466)	20	3
HILIC (+)	F	144.1018	464	Proline betaine	(469)	13	2
HILIC (+)	F	157.076	344	3-Indoleacetonitrile	(474)	11	3
HILIC (+)	F	166.0723	309	7-Methylguanine | N2-Methylguanine	(477)	3	3
HILIC (+)	F	170.0599	344	Taurocyamine	(480)	11	3
HILIC (+)	F	182.0561	179	5-Methylthio-D-ribose | 2,3-Dihydroxy-5-methylthio-4-pentenoic acid	(482)	13	3
HILIC (+)	F	186.0356	466	Glutamine	(483)	22	2
HILIC (+)	F	229.1546	468	Isoleucylproline | Leucylproline	(489)	14	2
HILIC (+)	F	230.157	468	Isoleucylproline | Leucylproline	(490)	13	3
HILIC (+)	F	243.0949	426	Pantothenic acid	(493)	9	3
HILIC (+)	F	254.0801	479	2-amino-8-oxo-9,10-epoxy-decanoic acid | CAR(3:1)	(496)	10	3
HILIC (+)	F	277.0381	471	Isoferulic acid 3-sulfate | Dihydroferulic acid 4-O-sulfate	(498)	15	3
HILIC (+)	F	298.1008	529	N-gamma-Glutamylglutamine | N2-gamma-Glutamylglutamine	(502)	0	2
HILIC (+)	F	298.1146	309	1-Methylguanosine | (S)-malyl alpha-D-glucosaminide	(503)	3	3
HILIC (+)	F	309.1655	570	1-[(5-Amino-5-carboxypentyl)amino]-1-deoxyfructose	(504)	10	3
HILIC (+)	F	373.1155	537	Feruloylquinic acid	(507)	22	3
HILIC (+)	F	407.0024	478	Vanillic acid 4-O-glucuronide	(508)	7	3
HILIC (+)	F	468.3444	308	LPC(O-15:0)	(514)	9	2
HILIC (+)	F	484.3667	307	LPC(P-18:0) | LPC(O-18:1) | LPC(O-18:0)	(515)	8	3

HILIC (+)	F	498.3442	303	ST(27:2;O7)	(517)	9	3
HILIC (+)	F	527.3785	301	ST(29:2;O7)	(520)	9	3
HILIC (+)	F	536.3698	299	PC(O-19:1)	(521)	9	2
HILIC (+)	F	560.3354	308	PC(20:4) | LPC(20:4;O)	(523)	12	3
HILIC (+)	F	607.2508	96	Bilirubin	(525)	9	2
HILIC (-)	F	174.0882	491	Citrulline	(530)	9	2
HILIC (-)	F	185.1183	53	FA(10:1;O) | WE(10:1;O) | FOH(10:2;O2) | FA(10:1;O) | FOH(9:2) | FOH(8:2) | 2-Octenal	(532)	19	3
HILIC (-)	F	262.1047	513	N-Acetyl-L-citrulline	(536)	14	3
HILIC (-)	F	273.1205	519	Gln-Gln	(538)	1	2
HILIC (-)	F	274.1046	538	L-Glutamyl-L-glutamine	(539)	0	2
HILIC (-)	F	295.228	29	FA(18:2;O)	(540)	21	2
HILIC (-)	F	411.1848	26	17-Hydroxypregnenolone sulfate	(542)	21	2

Footnotes: (+) / (-) indicate + and - ion modes respectively, M – Ms, F – females, *m/z* – mass to charge ratio, RT – retention time. The 'Heatmap index' can be used to cross reference metabolites within this table to Figure 5.4 and Figure 5.5, acrophase estimated via cosinor analysis as performed as part of the group level rhythm analysis. Lipid abbreviations: CAR – Acylcarnitine, Cer – Ceramide, CL – Cardiolipin, DG – Diacylglyceride, FA – Fatty acyl, FA(X:X) – Fatty acid, GL – Glycerolipid, GluCer – Glucocyl-ceramide, GM3 – Ganglioside mannoside 3, Hex2cer – Dihexosylceramide, NeuAcHexCer – N-Acetylneuraminic acid hexosylceramide, PA – Phosphatidic acid, PC – Phosphatidylcholine, PE – Phosphatidylethanolamine, PG – Phosphatidylglycerol, PI – Phosphatidylinositol, PS – Phosphatidylserine, SM – Sphingomyelin, SP – Sphingolipid, ST – Sterol lipid, TG – Triacylglyceride.

Appendix E

Table E1: List of all rhythmic metabolites by age and group alongside amplitude and MESOR of fitted rhythms as determined by cosinor analysis.

COHORT	Group	Metabolite	Amplitude (μM)†	MESOR (μM)†
2015	Control	Alanine	13.45	124.41
2015	Control	Arginine	3.94	102.51
2015	Control	Asparagine	1.46	25.23
2015	Control	Citrulline	8.81	149.57
2015	Control	Isoleucine	7.44	69.34
2015	Control	Kynurenine	0.18	1.58
2015	Control	Lysine	9.89	130.46
2015	Control	Serine	21.48	153.78
2015	Control	Tryptophan	5.45	37.31
2015	Control	Tyrosine	1.82	39.18
2015	Control	Carnitine	7.69	51.28
2015	Control	Hexadecanoylcarnitine	0.01	0.09
2015	Control	Hydroxyhexadecanoylcarnitine	0.00	0.03
2015	Control	Octadecanoylcarnitine	0.02	0.15
2015	Control	Octadecenoylcarnitine	0.01	0.08
2015	Control	Acetylcarnitine	5.55	27.54
2015	Control	Propionylcarnitine	0.04	0.32
2015	Control	Hydroxybutyrylcarnitine	0.01	0.05
2015	Control	lysoPC a C16:1	0.05	1.03
2015	Control	lysoPC a C18:1	1.23	20.03
2015	Control	lysoPC a C18:2	0.58	10.00
2015	Control	Melatonin	††28.24	††30.22
2015	Control	PC aa C32:3	0.01	0.21
2015	Control	PC aa C34:1	2.20	66.68
2015	Control	PC aa C36:4	0.64	19.49
2015	Control	PC aa C38:0	0.05	0.93
2015	Control	PC aa C38:4	1.43	39.09
2015	Control	PC aa C38:6	0.88	19.60
2015	Control	PC aa C40:5	1.44	28.93
2015	Control	PC aa C40:6	1.50	41.22
2015	Control	PC ae C34:3	0.13	1.71
2015	Control	PC ae C36:1	0.63	13.31
2015	Control	PC ae C36:4	0.08	1.31
2015	Control	PC ae C38:3	0.09	1.57
2015	Control	PC ae C40:1	0.03	0.36
2015	Control	PC ae C40:6	0.05	2.61
2015	Control	PC ae C42:3	0.01	0.18
2015	Control	SM (OH) C16:1	0.32	10.96
2015	Control	SM (OH) C22:1	0.84	25.43
2015	Control	SM C16:1	0.26	7.53
2015	Control	SM C18:0	0.38	10.78
2015	Control	SM C18:1	0.19	5.87

2015	Control	Urea	286.35	5907.18
2015	HD	Alanine	12.16	139.76
2015	HD	Asparagine	1.03	25.70
2015	HD	Citrulline	4.22	134.67
2015	HD	Isoleucine	9.48	74.35
2015	HD	Leucine	18.89	121.78
2015	HD	Lysine	10.04	140.85
2015	HD	Ornithine	4.28	52.69
2015	HD	Serine	20.93	152.51
2015	HD	Threonine	6.48	200.26
2015	HD	Tryptophan	5.70	42.26
2015	HD	Tyrosine	2.08	44.33
2015	HD	Valine	13.36	161.41
2015	HD	alpha-Aminoadipic acid	0.68	6.53
2015	HD	Carnitine	8.71	54.73
2015	HD	Hexadecanoylcarnitine	0.01	0.10
2015	HD	Hydroxyhexadecanoylcarnitine	0.00	0.03
2015	HD	Octadecanoylcarnitine	0.01	0.15
2015	HD	Octadecenoylcarnitine	0.01	0.08
2015	HD	Acetylcarnitine	5.23	26.89
2015	HD	Propionylcarnitine	0.03	0.31
2015	HD	Hydroxybutyrylcarnitine	0.01	0.05
2015	HD	Valerylcarnitine	0.02	0.24
2015	HD	Carnosine	1.20	15.02
2015	HD	Kynurenine	0.20	1.64
2015	HD	lysoPC a C16:1	0.06	1.05
2015	HD	lysoPC a C18:1	1.06	21.60
2015	HD	lysoPC a C18:2	0.30	10.08
2015	HD	Melatonin	††40.37	††40.69
2015	HD	PC aa C34:2	2.49	65.83
2015	HD	PC aa C36:3	1.05	24.45
2015	HD	PC aa C36:4	1.21	21.84
2015	HD	PC aa C36:5	0.36	6.32
2015	HD	PC aa C38:4	2.46	40.76
2015	HD	PC aa C38:5	2.14	38.36
2015	HD	PC aa C38:6	1.86	19.60
2015	HD	PC aa C40:6	2.84	38.48
2015	HD	PC ae C38:0	0.03	0.75
2015	HD	Urea	430.49	5725.84
2017	Control	Alanine	13.06	144.73
2017	Control	Asparagine	0.69	25.60
2017	Control	Citrulline	1.70	93.93
2017	Control	Isoleucine	8.26	83.08
2017	Control	Leucine	14.21	165.95

2017	Control	Lysine	9.61	129.95
2017	Control	Ornithine	5.03	44.83
2017	Control	Phenylalanine	1.72	59.06
2017	Control	Serine	6.61	92.63
2017	Control	Tryptophan	3.95	38.36
2017	Control	Valine	14.62	189.06
2017	Control	alpha-Aminoadipic acid	0.80	6.88
2017	Control	Carnitine	3.80	42.34
2017	Control	Hexadecanoylcarnitine	0.01	0.15
2017	Control	Hydroxyhexadecanoylcarnitine	0.00	0.04
2017	Control	Octadecanoylcarnitine	0.01	0.21
2017	Control	Octadecenoylcarnitine	0.01	0.14
2017	Control	Acetylcarnitine	2.04	34.47
2017	Control	Propionylcarnitine	0.02	0.32
2017	Control	Hydroxybutyrylcarnitine	0.01	0.07
2017	Control	Valerylcarnitine	0.02	0.33
2017	Control	Carnosine	1.15	17.83
2017	Control	Kynurenine	0.05	1.80
2017	Control	lysoPC a C16:0	1.41	19.01
2017	Control	lysoPC a C16:1	0.06	0.79
2017	Control	lysoPC a C17:0	0.11	1.43
2017	Control	lysoPC a C18:0	1.31	15.10
2017	Control	lysoPC a C18:1	1.98	17.19
2017	Control	lysoPC a C18:2	0.94	8.15
2017	Control	lysoPC a C20:3	0.06	0.43
2017	Control	lysoPC a C20:4	0.33	2.95
2017	Control	Melatonin	††55.62	††60.78
2017	Control	PC aa C32:0	0.16	4.85
2017	Control	PC aa C34:1	3.08	88.05
2017	Control	PC aa C34:2	7.68	122.08
2017	Control	PC aa C34:3	0.30	6.82
2017	Control	PC aa C36:1	2.44	48.08
2017	Control	PC aa C36:2	5.96	111.45
2017	Control	PC aa C36:3	1.47	23.08
2017	Control	PC aa C36:4	1.52	24.67
2017	Control	PC aa C36:5	0.20	4.88
2017	Control	PC aa C38:0	0.02	0.70
2017	Control	PC aa C38:3	0.35	6.61
2017	Control	PC aa C38:4	2.43	47.21
2017	Control	PC aa C38:5	2.04	30.57
2017	Control	PC aa C38:6	1.33	16.72
2017	Control	PC aa C40:5	1.67	26.65
2017	Control	PC aa C40:6	2.52	36.06
2017	Control	PC ae C34:0	0.06	1.43
2017	Control	PC ae C34:1	0.20	5.17
2017	Control	PC ae C34:2	0.23	5.12

2017	Control	PC ae C34:3	0.09	1.71
2017	Control	PC ae C36:2	0.29	6.51
2017	Control	PC ae C36:3	0.10	2.65
2017	Control	PC ae C38:0	0.02	0.57
2017	Control	PC ae C38:4	0.10	2.55
2017	Control	PC ae C40:1	0.02	0.30
2017	Control	PC ae C40:4	0.03	0.67
2017	Control	PC ae C40:5	0.08	1.87
2017	Control	PC ae C40:6	0.09	2.03
2017	Control	PC ae C42:2	0.00	0.09
2017	Control	PC ae C42:3	0.01	0.12
2017	Control	Sarcosine	0.05	4.10
2017	Control	Symmetric dimethylarginine	0.10	1.89
2017	Control	Serotonin	0.63	6.97
2017	Control	Taurine	1.23	27.52
2017	Control	Urea	230.49	6151.01
2017	HD	Alanine	12.81	152.94
2017	HD	Asparagine	0.58	26.28
2017	HD	Histidine	1.12	72.64
2017	HD	Isoleucine	6.54	82.18
2017	HD	Phenylalanine	1.22	56.96
2017	HD	Serine	8.84	100.83
2017	HD	Threonine	2.65	170.33
2017	HD	Tryptophan	4.35	40.41
2017	HD	Tyrosine	1.11	46.92
2017	HD	Carnitine	5.55	44.11
2017	HD	Hexadecanoylcarnitine	0.01	0.12
2017	HD	Hydroxyhexadecanoylcarnitine	0.00	0.03
2017	HD	Octadecanoylcarnitine	0.01	0.17
2017	HD	Acetylcarnitine	2.06	30.51
2017	HD	Propionylcarnitine	0.02	0.31
2017	HD	Hydroxybutyrylcarnitine	0.01	0.08
2017	HD	Valerylcarnitine	0.01	0.31
2017	HD	Hydroxyvalerylcarnitine	0.00	0.06
2017	HD	Kynurenine	0.13	1.96
2017	HD	lysoPC a C18:1	0.86	21.65
2017	HD	lysoPC a C18:2	0.35	9.99
2017	HD	Melatonin	††64.98	††88.95
2017	HD	PC aa C28:1	0.01	0.37
2017	HD	PC aa C34:4	0.01	0.26
2017	HD	PC aa C36:3	0.77	22.43
2017	HD	PC ae C30:2	0.00	0.06
2017	HD	PC ae C34:2	0.29	5.14
2017	HD	PC ae C34:3	0.12	1.79
2017	HD	PC ae C36:0	0.03	1.09

2017	HD	PC ae C36:1	0.24	7.55
2017	HD	PC ae C36:3	0.15	2.67
2017	HD	PC ae C36:4	0.07	1.61
2017	HD	PC ae C38:3	0.04	0.94
2017	HD	PC ae C42:3	0.01	0.13
2017	HD	Urea	451.58	7254.42

Metabolite names maintain parity with naming conventions utilised by the applied Biocrates p180 targeted assay kit, details provided below. PC – glycerohosphatidylcholine. 'a' or 'aa' – both moieties (or one in the case of lysoPCs) at the sn-1 and sn-2 positions are fatty acyl residues bound to the glycerol backbone by ester bonds. 'ae' - one of the moieties, either in the sn-1 or sn-2 position, is a fatty alcohol residue bound by an ether bond.

† Calculated via cosinor analysis on non Z-scored (concentration - µM) data

†† Melatonin is measured on a picomolar scale